Dell Configuration Guide for the S4820T System

9.14.1.5

May 2019
Rev. A00

Notes, cautions, and warnings

NOTE: A NOTE indicates important information that helps you make better use of your product.

CAUTION: A CAUTION indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.

WARNING: A WARNING indicates a potential for property damage, personal injury, or death.

© 2018 - 2019 Dell Inc. or its subsidiaries. All rights reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

Contents

Chapter 1: About this Guide 31
Audience 31
Conventions 31
Related Documents 31

Chapter 2: Configuration Fundamentals 32
Accessing the Command Line 32
CLI Modes 32
Navigating CLI Modes 34
The do Command 37
Undoing Commands 37
Obtaining Help 38
Entering and Editing Commands 38
Command History 39
Filtering show Command Outputs 39
Multiple Users in Configuration Mode 40

Chapter 3: Getting Started 41
Console Access 41
Serial Console 41
Accessing the CLI Interface and Running Scripts Using SSH 42
Default Configuration 43
Configuring a Host Name 43
Accessing the System Remotely 43
Configure the Management Port IP Address 44
Configure a Management Route 44
Configuring a Username and Password 44
Configuring the Enable Password 45
Configuration File Management 46
Copy Files to and from the System 46
Mounting an NFS File System 46
Save the Running-Configuration 48
Configure the Overload Bit for a Startup Scenario 48
Viewing Files 48
Managing the File System 49
Enabling Software Features on Devices Using a Command Option 50
View Command History 50
Upgrading Dell EMC Networking OS 52
Using HTTP for File Transfers 52
Verify Software Images Before Installation 53

Chapter 4: Management 54
Configuring Privilege Levels 54
Removing a Command from EXEC Mode 55

Moving a Command from EXEC Privilege Mode to EXEC Mode....55
Allowing Access to CONFIGURATION Mode Commands....55
Allowing Access to Different Modes....55
Applying a Privilege Level to a Username....57
Applying a Privilege Level to a Terminal Line....57
Configuring Logging....57
Audit and Security Logs....57
Configuring Logging Format 59
Display the Logging Buffer and the Logging Configuration....59
Setting Up a Secure Connection to a Syslog Server....60
Log Messages in the Internal Buffer....61
Disabling System Logging....61
Sending System Messages to a Syslog Server....62
Configuring a UNIX System as a Syslog Server....62
Track Login Activity....62
Restrictions for Tracking Login Activity....62
Configuring Login Activity Tracking....62
Display Login Statistics....63
Limit Concurrent Login Sessions....64
Restrictions for Limiting the Number of Concurrent Sessions....64
Configuring Concurrent Session Limit....65
Enabling the System to Clear Existing Sessions....65
Enabling Secured CLI Mode....65
Changing System Logging Settings....66
Display the Logging Buffer and the Logging Configuration....67
Configuring a UNIX Logging Facility Level....67
Synchronizing Log Messages....68
Enabling Timestamp on Syslog Messages....69
File Transfer Services....71
Enabling the FTP Server....72
Configuring FTP Server Parameters....72
Configuring FTP Client Parameters....72
Terminal Lines....73
Denying and Permitting Access to a Terminal Line....73
Configuring Login Authentication for Terminal Lines....74
Setting Timeout for EXEC Privilege Mode....75
Using Telnet to get to Another Network Device....75
Lock CONFIGURATION Mode....76
Recovering from a Forgotten Password....76
Recovering from a Forgotten Enable Password....77
Recovering from a Failed Start....78
Restoring the Factory Default Settings....78
Restoring Factory Default Environment Variables....79
Dell EMC Networking OS Security Hardening....80
Dell EMC Networking OS Image Verification....80
Startup Configuration Verification....81
Configuring the root User Password....82
Locking Access to GRUB Interface....82
Enabling User Lockout for Failed Login Attempts....83

Chapter 5: 802.1X....84
Port-Authentication Process....85
EAP over RADIUS....86
Configuring 802.1X....87
Important Points to Remember....87
Enabling 802.1X....88
Configuring MAC addresses for a do1x Profile....89
Configuring Request Identity Re-Transmissions....90
Configuring a Quiet Period after a Failed Authentication....90
Forcibly Authorizing or Unauthorizing a Port....91
Re-Authenticating a Port....91
Configuring Timeouts....92
Configuring Dynamic VLAN Assignment with Port Authentication....93
Guest and Authentication-Fail VLANs....94
Configuring a Guest VLAN....95
Configuring an Authentication-Fail VLAN....95
Configuring dot1x Profile....96
Configuring the Static MAB and MAB Profile....96
Configuring Critical VLAN....97

Chapter 6: Access Control List (ACL) VLAN Groups and Content Addressable Memory (CAM)....99
Optimizing CAM Utilization During the Attachment of ACLs to VLANs....99
Guidelines for Configuring ACL VLAN Groups....100
Configuring ACL VLAN Groups and Configuring FP Blocks for VLAN Parameters....100
Configuring ACL VLAN Groups....100
Configuring FP Blocks for VLAN Parameters....101
Viewing CAM Usage....102
Allocating FP Blocks for VLAN Processes....103

Chapter 7: Access Control Lists (ACLs)....105
IP Access Control Lists (ACLs)....106
CAM Usage....106
Implementing ACLs on Dell EMC Networking OS....107
IP Fragment Handling....108
IP Fragments ACL Examples....109
Layer 4 ACL Rules Examples....109
Configure a Standard IP ACL....110
Configuring a Standard IP ACL Filter....111
Configure an Extended IP ACL....111
Configuring Filters with a Sequence Number....111
Configuring Filters Without a Sequence Number....112
Configure Layer 2 and Layer 3 ACLs....113
Assign an IP ACL to an Interface....114
Applying an IP ACL....114
Counting ACL Hits....114
Configure Ingress ACLs....115
Configure Egress ACLs....115
Applying Egress Layer 3 ACLs (Control-Plane)....116

IP Prefix Lists....116
Configuration Task List for Prefix Lists....117
ACL Remarks....120
Configuring a Remark....120
Deleting a Remark....121
ACL Resequencing....121
Resequencing an ACL or Prefix List....122
Route Maps....123
Important Points to Remember....123
Configuration Task List for Route Maps....123
Configuring Match Routes....125
Configuring Set Conditions....126
Configure a Route Map for Route Redistribution....127
Configure a Route Map for Route Tagging....128
Continue Clause....128
Logging of ACL Processes....128
Guidelines for Configuring ACL Logging....129
Configuring ACL Logging....130
Flow-Based Monitoring....130
Enabling Flow-Based Monitoring....132

Chapter 8: Bidirectional Forwarding Detection (BFD)....133
How BFD Works....133
BFD Packet Format....134
BFD Sessions....135
BFD Three-Way Handshake....136
Session State Changes....137
Important Points to Remember....137
Configure BFD....137
Configure BFD for Physical Ports....138
Configure BFD for Static Routes....139
Configure BFD for OSPF....142
Configure BFD for OSPFv3....145
Configure BFD for IS-IS....148
Configure BFD for BGP....150
Configure BFD for VRRP....156
Configuring Protocol Liveness....159
Troubleshooting BFD....159

Chapter 9: Border Gateway Protocol (BGP)....161
BGP IP version 4 (BGPv4) Overview....161
BGP Autonomous Systems....161
AS4 Number Representation....163
Four-Byte AS Numbers....165
BGP router ID....165
Sessions and Peers....165
Establish a Session....165
BGP Attributes for selecting Best Path....166
Multiprotocol BGP....171

BGP global and address family configuration....172
Implement BGP with Dell EMC Networking OS....173
Configuration Information....176
Basic BGP configuration tasks....177
Prerequisite for configuring a BGP network....177
Restrictions....177
Enabling BGP....177
Enabling four-byte autonomous system numbers....179
Changing a BGP router ID....180
Configuring AS4 Number Representations....180
Configuring a BGP peer....181
Example-Configuring BGP routing between peers....182
BGP peer group....183
Advanced BGP configuration tasks....190
Route-refresh and Soft-reconfiguration....190
Aggregating Routes....192
Filtering BGP....194
Configuring BGP Fast Fall-Over....199
Configuring Passive Peering....201
Enabling Graceful Restart....201
Redistributing Routes....202
Enabling Additional Paths....203
Configuring IP Community Lists....204
Configuring an IP Extended Community List....205
Configure BGP attributes....206
Enabling Multipath....209
Route Reflectors....210
Enabling Route Flap Dampening....211
Changing BGP keepalive and hold timers....213
Setting the extended timer....214
Enabling or disabling BGP neighbors....214
Route Map Continue....215
Configuring BGP Confederations....216
Configuring a BGP VRF address family....216
Maintaining Existing AS Numbers During an AS Migration....218
Allowing an AS Number to Appear in its Own AS Path....218
Enabling MBGP Configurations....219
MBGP support for IPv6....220
Configuring IPv6 MBGP between peers....220
Example-Configuring IPv4 and IPv6 neighbors....220
Configure IPv6 NH Automatically for IPv6 Prefix Advertised over IPv4 Neighbor....223
BGP Regular Expression Optimization....224
Debugging BGP....224

Chapter 10: Content Addressable Memory (CAM)....227
CAM Allocation....227
Test CAM Usage....229
View CAM Profiles....229
View CAM-ACL Settings....230
View CAM Usage....232

Configuring CAM Threshold and Silence Period ... 232
CAM Optimization ... 233
Troubleshoot CAM Profiling ... 233
QoS CAM Region Limitation ... 233

Chapter 11: Control Plane Policing (CoPP) ... 235
Configure Control Plane Policing ... 236
Configuring CoPP for Protocols ... 237
Configuring CoPP for CPU Queues ... 238
CoPP for OSPFv3 Packets ... 239
Configuring CoPP for OSPFv3 ... 242
Displaying CoPP Configuration ... 243

Chapter 12: Data Center Bridging (DCB) ... 245
Ethernet Enhancements in Data Center Bridging ... 245
Priority-Based Flow Control ... 246
Enhanced Transmission Selection ... 247
Data Center Bridging Exchange Protocol (DCBx) ... 248
Data Center Bridging in a Traffic Flow ... 249
Enabling Data Center Bridging ... 249
DCB Maps and its Attributes ... 250
Data Center Bridging: Default Configuration ... 250
Configuring Priority-Based Flow Control ... 251
Configuring Lossless Queues ... 251
Configuring PFC in a DCB Map ... 252
Applying a DCB Map on a Port ... 254
Configuring PFC without a DCB Map ... 254
Priority-Based Flow Control Using Dynamic Buffer Method ... 256
Behavior of Tagged Packets ... 257
Configuration Example for DSCP and PFC Priorities ... 257
Using PFC to Manage Converged Ethernet Traffic ... 258
Configure Enhanced Transmission Selection ... 259
ETS Prerequisites and Restrictions ... 259
Creating an ETS Priority Group ... 259
ETS Operation with DCBx ... 260
Configuring Bandwidth Allocation for DCBx CIN ... 260
Configuring ETS in a DCB Map ... 261
Hierarchical Scheduling in ETS Output Policies ... 262
Using ETS to Manage Converged Ethernet Traffic ... 263
Applying DCB Policies in a Switch Stack ... 263
Configure a DCBx Operation ... 263
DCBx Operation ... 263
DCBx Port Roles ... 263
DCB Configuration Exchange ... 265
Configuration Source Election ... 265
Propagation of DCB Information ... 265
Auto-Detection and Manual Configuration of the DCBx Version ... 266
DCBx Example ... 266
DCBx Prerequisites and Restrictions ... 267

Configuring DCBx....267
Verifying the DCB Configuration....270
Sample DCB Configuration....279
QoS dot1p Traffic Classification and Queue Assignment....281
Configuring the Dynamic Buffer Method....282

Chapter 13: Dynamic Host Configuration Protocol (DHCP)....283
DHCP Packet Format and Options....283
Assign an IP Address using DHCP....284
Implementation Information....285
Configure the System to be a DHCP Server....286
Configuring the Server for Automatic Address Allocation....286
Specifying a Default Gateway....287
Configure a Method of Hostname Resolution....288
Using DNS for Address Resolution....288
Using NetBIOS WINS for Address Resolution....288
Creating Manual Binding Entries....288
Debugging the DHCP Server....289
Using DHCP Clear Commands....289
Configure the System to be a Relay Agent....289
Configure the System to be a DHCP Client....291
DHCP Client Operation with Other Features....291
DHCP Client on a Management Interface....292
Configure the System for User Port Stacking (Option 230)....292
Configuring DHCP relay source interface....293
Global DHCP relay source IPv4 or IPv6 configuration....293
Interface level DHCP relay source IPv4 or IPv6 configuration....293
Configure Secure DHCP....294
Option 82 (DHCPv4 relay options)....295
DHCPv6 relay agent options....295
DHCP Snooping....296
Configuring the DHCP secondary-subnet....300
Drop DHCP Packets on Snooped VLANs Only....300
Dynamic ARP Inspection....301
Configuring Dynamic ARP Inspection....301
Source Address Validation....302
Enabling IP Source Address Validation....303
DHCP MAC Source Address Validation....303
Enabling IP+MAC Source Address Validation....303
Viewing the Number of SAV Dropped Packets....304
Clearing the Number of SAV Dropped Packets....304

Chapter 14: Equal Cost Multi-Path (ECMP)....305
ECMP for Flow-Based Affinity....305
Configuring the Hash Algorithm....305
Enabling Deterministic ECMP Next Hop....305
Configuring the Hash Algorithm Seed....306
Link Bundle Monitoring....306
Managing ECMP Group Paths....307

Creating an ECMP Group Bundle....307
Modifying the ECMP Group Threshold....307
RTAG7....308
Flow-based Hashing for ECMP....309

Chapter 15: FIP Snooping....312
Fibre Channel over Ethernet....312
Ensure Robustness in a Converged Ethernet Network....312
FIP Snooping on Ethernet Bridges....313
FIP Snooping in a Switch Stack....315
Using FIP Snooping....315
FIP Snooping Prerequisites....315
Important Points to Remember....315
Enabling the FCoE Transit Feature....316
Enable FIP Snooping on VLANs....316
Configure the FC-MAP Value....317
Configure a Port for a Bridge-to-Bridge Link....317
Configure a Port for a Bridge-to-FCF Link....317
Impact on Other Software Features....317
FIP Snooping Restrictions....318
Configuring FIP Snooping....318
Displaying FIP Snooping Information....319
FCoE Transit Configuration Example....324

Chapter 16: FIPS Cryptography....326
Configuration Tasks....326
Preparing the System....326
Enabling FIPS Mode....326
Generating Host-Keys....327
Monitoring FIPS Mode Status....327
Disabling FIPS Mode....327

Chapter 17: Force10 Resilient Ring Protocol (FRRP)....329
Protocol Overview....329
Ring Status....330
Multiple FRRP Rings....330
Important FRRP Points....331
Important FRRP Concepts....332
Implementing FRRP....332
FRRP Configuration....333
Creating the FRRP Group....333
Configuring the Control VLAN....333
Configuring and Adding the Member VLANs....334
Setting the FRRP Timers....335
Clearing the FRRP Counters....335
Viewing the FRRP Configuration....335
Viewing the FRRP Information....335
Troubleshooting FRRP....336
Sample Configuration and Topology....336

FRRP Support on VLT ... 337

Chapter 18: GARP VLAN Registration Protocol (GVRP) ... 340
Configure GVRP ... 341
Enabling GVRP Globally ... 341
Enabling GVRP on a Layer 2 Interface ... 342
Configure GVRP Registration ... 342
Configure a GARP Timer ... 343
RPM Redundancy ... 343

Chapter 19: High Availability (HA) ... 344
Component Redundancy ... 344
RPM Redundancy ... 344
Automatic and Manual Stack Unit Failover ... 346
Support for RPM Redundancy by Dell EMC Networking OS Version ... 347
Synchronization between Management and Standby Units ... 347
Configuring RPM Redundancy ... 347
Online Insertion and Removal ... 348
RPM Online Insertion and Removal ... 348
Linecard Online Insertion and Removal ... 348
Hitless Behavior ... 350
Graceful Restart ... 350
Software Resiliency ... 350
Software Component Health Monitoring ... 350
System Health Monitoring ... 350
Failure and Event Logging ... 351
Hot-Lock Behavior ... 351
Process Restartability ... 351
Enabling Process Restartability ... 352

Chapter 20: Internet Group Management Protocol (IGMP) ... 353
IGMP Protocol Overview ... 353
IGMP Version 2 ... 353
IGMP Version 3 ... 355
Configure IGMP ... 357
Viewing IGMP Enabled Interfaces ... 358
Selecting an IGMP Version ... 358
Viewing IGMP Groups ... 359
Adjusting Timers ... 359
Adjusting Query and Response Timers ... 359
Enabling IGMP Immediate-Leave ... 360
IGMP Snooping ... 360
Configuring IGMP Snooping ... 360
Removing a Group-Port Association ... 361
Disabling Multicast Flooding ... 361
Specifying a Port as Connected to a Multicast Router ... 361
Configuring the Switch as Querier ... 362
Fast Convergence after MSTP Topology Changes ... 362
Egress Interface Selection (EIS) for HTTP and IGMP Applications ... 362

Designating a Multicast Router Interface....369

Chapter 21: Interfaces....370
Interface Types....371
View Basic Interface Information....371
Resetting an Interface to its Factory Default State....373
Enabling Energy Efficient Ethernet....373
View EEE Information....374
Clear EEE Counters....378
Enabling a Physical Interface....378
Physical Interfaces....379
Configuration Task List for Physical Interfaces....379
Overview of Layer Modes....379
Configuring Layer 2 (Data Link) Mode....379
Configuring Layer 2 (Interface) Mode....380
Configuring Layer 3 (Network) Mode....380
Configuring Layer 3 (Interface) Mode....381
Egress Interface Selection (EIS)....381
Configuring EIS....381
Management Interfaces....382
Configuring Management Interfaces....382
Configuring a Management Interface on an Ethernet Port....383
VLAN Interfaces....384
Loopback Interfaces....385
Null Interfaces....385
Port Channel Interfaces....385
Port Channel Definition and Standards....385
Port Channel Benefits....386
Port Channel Implementation....386
Interfaces in Port Channels....386
Configuration Tasks for Port Channel Interfaces....387
Creating a Port Channel....387
Adding a Physical Interface to a Port Channel....387
Reassigning an Interface to a New Port Channel....389
Configuring the Minimum Oper Up Links in a Port Channel....389
Adding or Removing a Port Channel from a VLAN....390
Assigning an IP Address to a Port Channel....391
Deleting or Disabling a Port Channel....391
Load Balancing Through Port Channels....391
Changing the Hash Algorithm....391
Bulk Configuration....392
Interface Range....392
Bulk Configuration Examples....393
Defining Interface Range Macros....394
Define the Interface Range....394
Choosing an Interface-Range Macro....394
Monitoring and Maintaining Interfaces....395
Maintenance Using TDR....396
Configuring wavelength for 10–Gigabit SFP+ optics....396
Link Dampening....396

Enabling Link Dampening ... 399
Link Bundle Monitoring ... 400
Using Ethernet Pause Frames for Flow Control ... 400
Enabling Pause Frames ... 401
Configure the MTU Size on an Interface ... 401
Port-Pipes ... 402
Auto-Negotiation on Ethernet Interfaces ... 402
Setting the Speed of Ethernet Interfaces ... 402
Set Auto-Negotiation Options ... 404
View Advanced Interface Information ... 404
Configuring the Interface Sampling Size ... 405
Configuring the Traffic Sampling Size Globally ... 406
Dynamic Counters ... 407
Clearing Interface Counters ... 407
Compressing Configuration Files ... 408

Chapter 22: Internet Protocol Security (IPSec) ... 411
Configuring IPSec ... 411

Chapter 23: IPv4 Routing ... 413
IP Addresses ... 414
Configuration Tasks for IP Addresses ... 414
Assigning IP Addresses to an Interface ... 414
Configuring Static Routes ... 415
Adding description for IPv4 and IPv6 static routes ... 416
Configure Static Routes for the Management Interface ... 417
IPv4 Path MTU Discovery Overview ... 417
Using the Configured Source IP Address in ICMP Messages ... 418
Configuring the Duration to Establish a TCP Connection ... 418
Enabling Directed Broadcast ... 418
Resolution of Host Names ... 419
Enabling Dynamic Resolution of Host Names ... 419
Specifying the Local System Domain and a List of Domains ... 420
Configuring DNS with Traceroute ... 420
ARP ... 421
Configuration Tasks for ARP ... 421
Configuring Static ARP Entries ... 421
Enabling Proxy ARP ... 422
Clearing ARP Cache ... 422
ARP Learning via Gratuitous ARP ... 422
Enabling ARP Learning via Gratuitous ARP ... 422
ARP Learning via ARP Request ... 423
Configuring ARP Retries ... 423
ICMP ... 424
Configuration Tasks for ICMP ... 424
Enabling ICMP Unreachable Messages ... 424
ICMP Redirects ... 424
UDP Helper ... 425
Enabling UDP Helper ... 426

Configuring a Broadcast Address ... 426
Configurations Using UDP Helper ... 427
UDP Helper with Broadcast-All Addresses ... 427
UDP Helper with Subnet Broadcast Addresses ... 427
UDP Helper with Configured Broadcast Addresses ... 428
UDP Helper with No Configured Broadcast Addresses ... 428
Troubleshooting UDP Helper ... 429

Chapter 24: IPv6 Routing ... 430
Protocol Overview ... 430
Extended Address Space ... 430
Stateless Autoconfiguration ... 430
IPv6 Headers ... 431
IPv6 Header Fields ... 431
Extension Header Fields ... 433
Addressing ... 434
Implementing IPv6 with Dell EMC Networking OS ... 435
ICMPv6 ... 436
Path MTU discovery ... 437
IPv6 Neighbor Discovery ... 437
IPv6 Neighbor Discovery of MTU Packets ... 438
Configuration Task List for IPv6 RDNSS ... 438
Configuring the IPv6 Recursive DNS Server ... 438
Debugging IPv6 RDNSS Information Sent to the Host ... 439
Displaying IPv6 RDNSS Information ... 439
Secure Shell (SSH) Over an IPv6 Transport ... 440
Configuration Tasks for IPv6 ... 440
Adjusting Your CAM-Profile ... 440
Assigning an IPv6 Address to an Interface ... 441
Assigning a Static IPv6 Route ... 441
Configuring Telnet with IPv6 ... 442
SNMP over IPv6 ... 442
Displaying IPv6 Information ... 443
Displaying an IPv6 Interface Information ... 443
Showing IPv6 Routes ... 444
Showing the Running-Configuration for an Interface ... 445
Clearing IPv6 Routes ... 445
Disabling ND Entry Timeout ... 445
Configuring IPv6 RA Guard ... 446
Configuring IPv6 RA Guard on an Interface ... 447
Monitoring IPv6 RA Guard ... 448

Chapter 25: iSCSI Optimization ... 449
iSCSI Optimization Overview ... 449
Monitoring iSCSI Traffic Flows ... 450
Application of Quality of Service to iSCSI Traffic Flows ... 451
Information Monitored in iSCSI Traffic Flows ... 451
Detection and Auto-Configuration for Dell EqualLogic Arrays ... 452
Configuring Detection and Ports for Dell Compellent Arrays ... 452

Synchronizing iSCSI Sessions Learned on VLT-Lags with VLT-Peer ... 452
Enable and Disable iSCSI Optimization ... 453
Default iSCSI Optimization Values ... 453
iSCSI Optimization Prerequisites ... 454
Configuring iSCSI Optimization ... 454
Displaying iSCSI Optimization Information ... 456

Chapter 26: Intermediate System to Intermediate System ... 458
IS-IS Protocol Overview ... 458
IS-IS Addressing ... 458
Multi-Topology IS-IS ... 459
Transition Mode ... 459
Interface Support ... 459
Adjacencies ... 459
Graceful Restart ... 460
Timers ... 460
Implementation Information ... 460
Configuration Information ... 461
Configuration Tasks for IS-IS ... 461
Configuring the Distance of a Route ... 468
Changing the IS-Type ... 468
Redistributing IPv4 Routes ... 470
Redistributing IPv6 Routes ... 471
Configuring Authentication Passwords ... 472
Setting the Overload Bit ... 472
Debugging IS-IS ... 473
IS-IS Metric Styles ... 473
Configure Metric Values ... 474
Maximum Values in the Routing Table ... 474
Change the IS-IS Metric Style in One Level Only ... 474
Leaks from One Level to Another ... 475
Sample Configurations ... 476

Chapter 27: Link Aggregation Control Protocol (LACP) ... 479
Introduction to Dynamic LAGs and LACP ... 479
Important Points to Remember ... 479
LACP Modes ... 480
Configuring LACP Commands ... 480
LACP Configuration Tasks ... 480
Creating a LAG ... 481
Configuring the LAG Interfaces as Dynamic ... 481
Setting the LACP Long Timeout ... 481
Monitoring and Debugging LACP ... 482
Shared LAG State Tracking ... 482
Configuring Shared LAG State Tracking ... 483
Important Points about Shared LAG State Tracking ... 484
LACP Basic Configuration Example ... 485
Configure a LAG on ALPHA ... 485

Chapter 28: Layer 2 493
Manage the MAC Address Table 493
Clearing the MAC Address Table 493
Setting the Aging Time for Dynamic Entries 493
Configuring a Static MAC Address 494
Displaying the MAC Address Table 494
MAC Learning Limit 494
Setting the MAC Learning Limit 495
mac learning-limit Dynamic 495
mac learning-limit mac-address-sticky 495
mac learning-limit station-move 495
mac learning-limit no-station-move 496
Learning Limit Violation Actions 496
Setting Station Move Violation Actions 496
Recovering from Learning Limit and Station Move Violations 497
Disabling MAC Address Learning on the System 497
Enabling port security 497
NIC Teaming 498
Configure Redundant Pairs 499
Far-End Failure Detection 502
FEFD State Changes 502
Configuring FEFD 503
Enabling FEFD on an Interface 504
Debugging FEFD 505

Chapter 29: Link Layer Discovery Protocol (LLDP) 506
802.1AB (LLDP) Overview 506
Protocol Data Units 506
Optional TLVs 507
Management TLVs 507
TIA-1057 (LLDP-MED) Overview 509
TIA Organizationally Specific TLVs 509
Configure LLDP 513
CONFIGURATION versus INTERFACE Configurations 513
Enabling LLDP 514
Disabling and Undoing LLDP 514
Advertising TLVs 514
Viewing the LLDP Configuration 515
Viewing Information Advertised by Adjacent LLDP Neighbors 516
Examples of Viewing Information Advertised by Neighbors 516
Configuring LLDPDU Intervals 518
Configuring Transmit and Receive Mode 518
Configuring the Time to Live Value 519
Debugging LLDP 520
Relevant Management Objects 521

Chapter 30: Microsoft Network Load Balancing 525
Configuring a Switch for NLB 526

Enabling a Switch for Multicast NLB ... 527

Chapter 31: Multicast Source Discovery Protocol (MSDP) ... 528
Anycast RP ... 529
Implementation Information ... 530
Configure Multicast Source Discovery Protocol ... 530
Related Configuration Tasks ... 530
Enable MSDP ... 534
Manage the Source-Active Cache ... 535
Viewing the Source-Active Cache ... 535
Limiting the Source-Active Cache ... 535
Clearing the Source-Active Cache ... 535
Enabling the Rejected Source-Active Cache ... 536
Accept Source-Active Messages that Fail the RFP Check ... 536
Specifying Source-Active Messages ... 539
Limiting the Source-Active Messages from a Peer ... 540
Preventing MSDP from Caching a Local Source ... 540
Preventing MSDP from Caching a Remote Source ... 541
Preventing MSDP from Advertising a Local Source ... 541
Logging Changes in Peership States ... 542
Terminating a Peership ... 542
Clearing Peer Statistics ... 542
Debugging MSDP ... 543
MSDP with Anycast RP ... 543
Configuring Anycast RP ... 544
Reducing Source-Active Message Flooding ... 545
Specifying the RP Address Used in SA Messages ... 545
MSDP Sample Configurations ... 547

Chapter 32: Multiple Spanning Tree Protocol (MSTP) ... 550
Spanning Tree Variations ... 551
Implementation Information ... 551
Configure Multiple Spanning Tree Protocol ... 551
Related Configuration Tasks ... 551
Enable Multiple Spanning Tree Globally ... 552
Adding and Removing Interfaces ... 552
Creating Multiple Spanning Tree Instances ... 552
Influencing MSTP Root Selection ... 553
Interoperate with Non-Dell Bridges ... 554
Changing the Region Name or Revision ... 554
Modifying Global Parameters ... 554
Modifying the Interface Parameters ... 555
Setting STP path cost as constant ... 556
Configuring an EdgePort ... 557
Flush MAC Addresses after a Topology Change ... 557
MSTP Sample Configurations ... 557
Debugging and Verifying MSTP Configurations ... 561

Chapter 33: Multicast Features ... 563

Enabling IP Multicast ... 563
Implementation Information ... 563
Multicast Policies ... 564
IPv4 Multicast Policies ... 564
Understanding Multicast Traceroute (mtrace) ... 570
Printing Multicast Traceroute (mtrace) Paths ... 570
Supported Error Codes ... 572
mtrace Scenarios ... 573

Chapter 34: Object Tracking ... 579
Object Tracking Overview ... 579
Track Layer 2 Interfaces ... 580
Track Layer 3 Interfaces ... 580
Track IPv4 and IPv6 Routes ... 580
Set Tracking Delays ... 581
VRRP Object Tracking ... 582
Object Tracking Configuration ... 582
Tracking a Layer 2 Interface ... 582
Tracking a Layer 3 Interface ... 583
Track an IPv4/IPv6 Route ... 584
Displaying Tracked Objects ... 587

Chapter 35: Open Shortest Path First (OSPFv2 and OSPFv3) ... 589
Protocol Overview ... 589
Autonomous System (AS) Areas ... 589
Area Types ... 590
Networks and Neighbors ... 591
Router Types ... 591
Designated and Backup Designated Routers ... 592
Link-State Advertisements (LSAs) ... 592
Router Priority and Cost ... 593
OSPF with Dell EMC Networking OS ... 594
Graceful Restart ... 594
Fast Convergence (OSPFv2, IPv4 Only) ... 595
Multi-Process OSPFv2 with VRF ... 596
RFC-2328 Compliant OSPF Flooding ... 596
OSPF ACK Packing ... 597
Setting OSPF Adjacency with Cisco Routers ... 597
Configuration Information ... 598
Configuration Task List for OSPFv2 (OSPF for IPv4) ... 598
Sample Configurations for OSPFv2 ... 609
OSPFv3 NSSA ... 611
Configuration Task List for OSPFv3 (OSPF for IPv6) ... 611
Enabling IPv6 Unicast Routing ... 612
Applying cost for OSPFv3 ... 612
Assigning IPv6 Addresses on an Interface ... 612
Assigning Area ID on an Interface ... 613
Assigning OSPFv3 Process ID and Router ID Globally ... 613
Assigning OSPFv3 Process ID and Router ID to a VRF ... 613

Configuring Stub Areas....614
Configuring Passive-Interface....614
Redistributing Routes....614
Configuring a Default Route....615
Enabling OSPFv3 Graceful Restart....615
OSPFv3 Authentication Using IPsec....617
Troubleshooting OSPFv3....623

Chapter 36: Policy-based Routing (PBR)....624
Overview....624
Implementing PBR....625
Configuration Task List for Policy-based Routing....625
Create a Redirect List....626
Create a Rule for a Redirect-list....626
Apply a Redirect-list to an Interface using a Redirect-group....628
Show Redirect List Configuration....629
Sample Configuration....630

Chapter 37: PIM Sparse-Mode (PIM-SM)....635
Implementation Information....635
Protocol Overview....635
Requesting Multicast Traffic....635
Refuse Multicast Traffic....636
Send Multicast Traffic....636
Configuring PIM-SM....636
Related Configuration Tasks....637
Enable PIM-SM....637
Configuring S,G Expiry Timers....638
Configuring a Static Rendezvous Point....639
Overriding Bootstrap Router Updates....639
Configuring a Designated Router....639
Creating Multicast Boundaries and Domains....641

Chapter 38: PIM Source-Specific Mode (PIM-SSM)....642
Implementation Information....642
Configure PIM-SSM....642
Enabling PIM-SSM....643
Use PIM-SSM with IGMP Version 2 Hosts....643
Electing an RP using the BSR Mechanism....645
Enabling RP to Server Specific Multicast Groups....646

Chapter 39: Port Monitoring....648
Important Points to Remember....648
Port Monitoring....648
Configuring Port Monitoring....651
Enabling Flow-Based Monitoring....652
Remote Port Mirroring....653
Remote Port Mirroring Example....653
Configuring Remote Port Mirroring....654

Displaying Remote-Port Mirroring Configurations....656
Configuration procedure for Remote Port Mirroring....656
Encapsulated Remote Port Monitoring....660
ERPM Behavior on a typical Dell EMC Networking OS....662
Port Monitoring on VLT....663

Chapter 40: Private VLANs (PVLAN)....666
Private VLAN Concepts....666
Using the Private VLAN Commands....667
Configuration Task List....668
Creating PVLAN ports....668
Creating a Primary VLAN....668
Creating a Community VLAN....669
Creating an Isolated VLAN....670
Private VLAN Configuration Example....671
Inspecting the Private VLAN Configuration....672

Chapter 41: Per-VLAN Spanning Tree Plus (PVST+)....674
Protocol Overview....674
Implementation Information....675
Configure Per-VLAN Spanning Tree Plus....676
Enabling PVST+....676
Disabling PVST+....676
Influencing PVST+ Root Selection....677
Modifying Global PVST+ Parameters....678
Modifying Interface PVST+ Parameters....678
Configuring an EdgePort....679
PVST+ in Multi-Vendor Networks....680
Enabling PVST+ Extend System ID....680
PVST+ Sample Configurations....681

Chapter 42: Quality of Service (QoS)....683
Implementation Information....685
Port-Based QoS Configurations....685
Setting dot1p Priorities for Incoming Traffic....685
Honoring dot1p Priorities on Ingress Traffic....686
Configuring Port-Based Rate Policing....686
Configuring Port-Based Rate Shaping....687
Policy-Based QoS Configurations....688
Classify Traffic....688
Create a QoS Policy....691
Create Policy Maps....693
DSCP Color Maps....696
Enabling QoS Rate Adjustment....698
Enabling Strict-Priority Queueing....699
Weighted Random Early Detection....699
Creating WRED Profiles....700
Applying a WRED Profile to Traffic....700
Displaying Default and Configured WRED Profiles....700

Displaying WRED Drop Statistics....701
Displaying egress–queue Statistics....701
Pre-Calculating Available QoS CAM Space....702
Configuring Weights and ECN for WRED....702
Configuring WRED and ECN Attributes....704
Guidelines for Configuring ECN for Classifying and Color-Marking Packets....705
Sample configuration to mark non-ecn packets as “yellow” with Multiple traffic class....705
Classifying Incoming Packets Using ECN and Color-Marking....706
Sample configuration to mark non-ecn packets as “yellow” with single traffic class....707
Applying Layer 2 Match Criteria on a Layer 3 Interface....708
Applying DSCP and VLAN Match Criteria on a Service Queue....709
Enabling Buffer Statistics Tracking....709

Chapter 43: Routing Information Protocol (RIP)....712
Protocol Overview....712
RIPv1....712
RIPv2....712
Implementation Information....712
Configuration Information....713
Configuration Task List....713
RIP Configuration Example....719

Chapter 44: Remote Monitoring (RMON)....724
Implementation Information....724
Fault Recovery....724
Setting the RMON Alarm....724
Configuring an RMON Event....725
Configuring RMON Collection Statistics....726
Configuring the RMON Collection History....726

Chapter 45: Rapid Spanning Tree Protocol (RSTP)....727
Protocol Overview....727
Configuring Rapid Spanning Tree....727
Important Points to Remember....728
RSTP and VLT....728
Configuring Interfaces for Layer 2 Mode....728
Enabling Rapid Spanning Tree Protocol Globally....729
Adding and Removing Interfaces....731
Modifying Global Parameters....731
Enabling SNMP Traps for Root Elections and Topology Changes....732
Modifying Interface Parameters....732
Enabling SNMP Traps for Root Elections and Topology Changes....732
Influencing RSTP Root Selection....733
Configuring an EdgePort....733
Configuring Fast Hellos for Link State Detection....734

Chapter 46: Software-Defined Networking (SDN)....735

Chapter 47: Security....736

AAA Accounting....736
Configuration Task List for AAA Accounting....736
RADIUS Accounting....738
AAA Authentication....743
Configuration Task List for AAA Authentication....743
Obscuring Passwords and Keys....746
AAA Authorization....746
Privilege Levels Overview....746
Configuration Task List for Privilege Levels....747
RADIUS....751
RADIUS Authentication....751
Configuration Task List for RADIUS....752
TACACS+....755
Configuration Task List for TACACS+....755
TACACS+ Remote Authentication....757
Command Authorization....758
Protection from TCP Tiny and Overlapping Fragment Attacks....758
Enabling SCP and SSH....758
Using SCP with SSH to Copy a Software Image....759
Removing the RSA Host Keys and Zeroizing Storage....760
Configuring When to Re-generate an SSH Key....760
Configuring the SSH Server Key Exchange Algorithm....760
Configuring the HMAC Algorithm for the SSH Server....761
Configuring the HMAC Algorithm for the SSH Client....761
Configuring the SSH Server Cipher List....762
Configuring the SSH Client Cipher List....762
Secure Shell Authentication....763
Troubleshooting SSH....765
Telnet....765
VTY Line and Access-Class Configuration....766
VTY Line Local Authentication and Authorization....766
VTY Line Remote Authentication and Authorization....767
VTY MAC-SA Filter Support....767
Support for Change of Authorization and Disconnect Messages packets....767
Change of Authorization (CoA) packets....768
Disconnect Messages....768
Attributes....768
Error-cause Values....771
CoA Packet Processing....771
CoA or DM Discard....772
Disconnect Message Processing....772
Configuring DAC....773
Configuring the port number....773
Configuring shared key....773
Disconnecting administrative users logged in through RADIUS....773
Configuring CoA to bounce 802.1x enabled ports....774
Configuring CoA to re-authenticate 802.1x sessions....774
Terminating the 802.1x user session....775
Disabling 802.1x enabled port....776
Important points to remember....776

Configuring replay protection....777
Rate-limiting RADIUS packets....777
Configuring time-out value....777
Role-Based Access Control....777
Overview of RBAC....778
User Roles....780
AAA Authentication and Authorization for Roles....783
Role Accounting....785
Display Information About User Roles....786
Two Factor Authentication (2FA)....787
Handling Access-Challenge Message....788
Configuring Challenge Response Authentication for SSHv2....788
SMS-OTP Mechanism....788
Configuring the System to Drop Certain ICMP Reply Messages....789

Chapter 48: Service Provider Bridging....791
VLAN Stacking....791
Configure VLAN Stacking....792
Creating Access and Trunk Ports....793
Enable VLAN-Stacking for a VLAN....793
Configuring the Protocol Type Value for the Outer VLAN Tag....794
Configuring Dell EMC Networking OS Options for Trunk Ports....794
Debugging VLAN Stacking....795
VLAN Stacking in Multi-Vendor Networks....795
VLAN Stacking Packet Drop Precedence....799
Enabling Drop Eligibility....799
Honoring the Incoming DEI Value....799
Marking Egress Packets with a DEI Value....800
Dynamic Mode CoS for VLAN Stacking....800
Mapping C-Tag to S-Tag dot1p Values....802
Layer 2 Protocol Tunneling....802
Enabling Layer 2 Protocol Tunneling....804
Specifying a Destination MAC Address for BPDUs....805
Setting Rate-Limit BPDUs....805
Debugging Layer 2 Protocol Tunneling....805
Provider Backbone Bridging....805

Chapter 49: sFlow....807
Overview....807
Implementation Information....807
Enabling Extended sFlow....808
Enabling and Disabling sFlow on an Interface....809
Enabling sFlow Max-Header Size Extended....809
sFlow Show Commands....810
Displaying Show sFlow Global....810
Displaying Show sFlow on an Interface....811
Displaying Show sFlow on a Stack-unit....811
Configuring Specify Collectors....811
Changing the Polling Intervals....812

Back-Off Mechanism....812
sFlow on LAG ports....812
Enabling Extended sFlow....812
Important Points to Remember....813

Chapter 50: Simple Network Management Protocol (SNMP)....815
Protocol Overview....815
Implementation Information....816
SNMPv3 Compliance With FIPS....816
Configuration Task List for SNMP....817
Important Points to Remember....817
Set up SNMP....817
Creating a Community....818
Setting Up User-Based Security (SNMPv3)....818
Enable SNMPv3 traps....819
Reading Managed Object Values....819
Writing Managed Object Values....820
Configuring Contact and Location Information using SNMP....820
Subscribing to Managed Object Value Updates using SNMP....821
Enabling a Subset of SNMP Traps....822
Enabling an SNMP Agent to Notify Syslog Server Failure....823
Copy Configuration Files Using SNMP....824
Copying a Configuration File....826
Copying Configuration Files via SNMP....826
Copying the Startup-Config Files to the Running-Config....827
Copying the Startup-Config Files to the Server via FTP....827
Copying the Startup-Config Files to the Server via TFTP....828
Copy a Binary File to the Startup-Configuration....828
Additional MIB Objects to View Copy Statistics....828
Obtaining a Value for MIB Objects....829
MIB Support for Power Monitoring....829
MIB Support to Display the Available Memory Size on Flash....830
Viewing the Available Flash Memory Size....830
MIB Support to Display the Software Core Files Generated by the System....831
Viewing the Software Core Files Generated by the System....831
MIB Support to Display the Available Partitions on Flash....832
Viewing the Available Partitions on Flash....832
MIB Support to Display Egress Queue Statistics....833
MIB Support to ECMP Group Count....833
Viewing the ECMP Group Count Information....833
MIB Support for entAliasMappingTable....836
Viewing the entAliasMappingTable MIB....837
MIB Support for LAG....837
Viewing the LAG MIB....838
MIB support for Port Security....839
Global MIB objects for port security....839
MIB support for interface level port security....839
MIB objects for configuring MAC addresses....840
MIB objects for configuring MAC addresses....841
Manage VLANs using SNMP....842

Creating a VLAN....842
Assigning a VLAN Alias....842
Displaying the Ports in a VLAN....842
Add Tagged and Untagged Ports to a VLAN....843
Managing Overload on Startup....844
Enabling and Disabling a Port using SNMP....845
Fetch Dynamic MAC Entries using SNMP....845
Example of Deriving the Interface Index Number....846
MIB Objects for Viewing the System Image on Flash Partitions....847
Monitoring BGP sessions via SNMP....847
Monitor Port-Channels....849
Troubleshooting SNMP Operation....850
Transceiver Monitoring....850

Chapter 51: Stacking....852
Stacking Overview....852
Stack Management Roles....852
Stack Master Election....853
Virtual IP....853
Failover Roles....854
MAC Addressing on Stacks....854
Stacking LAG....855
Supported Stacking Topologies....855
High Availability on Stacks....856
Management Access on Stacks....856
Important Points to Remember....857
Stacking Installation Tasks....857
Create a Stack....857
Add Units to an Existing Stack....860
Split a Stack....863
Stacking Configuration Tasks....863
Assigning Unit Numbers to Units in an Stack....863
Creating a Virtual Stack Unit on a Stack....863
Displaying Information about a Stack....863
Influencing Management Unit Selection on a Stack....865
Managing Redundancy on a Stack....866
Resetting a Unit on a Stack....866
Verify a Stack Configuration....866
Displaying the Status of Stacking Ports....866
Remove Units or Front End Ports from a Stack....868
Removing a Unit from a Stack....868
Removing Front End Port Stacking....869
Troubleshoot a Stack....869
Recover from Stack Link Flaps....869
Recover from a Card Problem State on a Stack....870
Recover from a Card Mismatch State on a Stack....870

Chapter 52: Storm Control....872
Configure Storm Control....872

Configuring Storm Control from INTERFACE Mode....872
Configuring Storm Control from CONFIGURATION Mode....873

Chapter 53: Spanning Tree Protocol (STP)....874
Protocol Overview....874
Configure Spanning Tree....874
Important Points to Remember....875
Configuring Interfaces for Layer 2 Mode....876
Enabling Spanning Tree Protocol Globally....877
Adding an Interface to the Spanning Tree Group....878
Modifying Global Parameters....879
Modifying Interface STP Parameters....880
Enabling PortFast....880
Prevent Network Disruptions with BPDU Guard....881
Selecting STP Root....883
STP Root Guard....883
Root Guard Scenario....883
Configuring Root Guard....884
Enabling SNMP Traps for Root Elections and Topology Changes....885
Configuring Spanning Trees as Hitless....885
STP Loop Guard....885
Configuring Loop Guard....886
Displaying STP Guard Configuration....887

Chapter 54: SupportAssist....888
Configuring SupportAssist Using a Configuration Wizard....888
Configuring SupportAssist Manually....889
Configuring SupportAssist Activity....890
Configuring SupportAssist Company....892
Configuring SupportAssist Person....892
Configuring SupportAssist Server....893
Viewing SupportAssist Configuration....894

Chapter 55: System Time and Date....896
Network Time Protocol....896
Protocol Overview....897
Configure the Network Time Protocol....897
Enabling NTP....897
Configuring NTP Broadcasts....898
Disabling NTP on an Interface....898
Configuring a Source IP Address for NTP Packets....898
Configuring NTP Authentication....899
Configuring a Custom-defined Period for NTP time Synchronization....901
Dell EMC Networking OS Time and Date....901
Configuration Task List....901
Setting the Time and Date for the Switch Software Clock....901
Setting the Timezone....902
Set Daylight Saving Time....902
Setting Daylight Saving Time Once....902

Setting Recurring Daylight Saving Time....903

Chapter 56: Tunneling....905
Configuring a Tunnel....905
Configuring Tunnel Keepalive Settings....906
Configuring a Tunnel Interface....906
Configuring Tunnel Allow-Remote Decapsulation....907
Configuring Tunnel source anylocal Decapsulation....907

Chapter 57: Uplink Failure Detection (UFD)....909
Feature Description....909
How Uplink Failure Detection Works....910
UFD and NIC Teaming....911
Important Points to Remember....911
Configuring Uplink Failure Detection....912
Clearing a UFD-Disabled Interface....913
Displaying Uplink Failure Detection....914
Sample Configuration: Uplink Failure Detection....916

Chapter 58: Upgrade Procedures....917

Chapter 59: Virtual LANs (VLANs)....918
Default VLAN....918
Port-Based VLANs....919
VLANs and Port Tagging....919
Configuration Task List....920
Creating a Port-Based VLAN....920
Assigning Interfaces to a VLAN....921
Moving Untagged Interfaces....922
Assigning an IP Address to a VLAN....923
Configuring Native VLANs....923
Enabling Null VLAN as the Default VLAN....924

Chapter 60: Virtual Link Trunking (VLT)....925
Overview....925
VLT Terminology....927
Layer-2 Traffic in VLT Domains....928
Interspersed VLANs....929
VLT on Core Switches....929
Enhanced VLT....930
Configure Virtual Link Trunking....931
Important Points to Remember....931
Configuration Notes....932
Primary and Secondary VLT Peers....935
RSTP and VLT....935
VLT Bandwidth Monitoring....936
VLT and Stacking....936
VLT and IGMP Snooping....936
VLT IPv6....936

VLT Port Delayed Restoration....937
PIM-Sparse Mode Support on VLT....937
VLT Routing....939
Non-VLT ARP Sync....942
RSTP Configuration....942
Preventing Forwarding Loops in a VLT Domain....942
Sample RSTP configuration....943
Configuring VLT....943
PVST+ Configuration....952
Peer Routing Configuration Example....953
Dell-1 Switch Configuration....954
Dell-2 Switch Configuration....957
R1 Configuration....960
Access Switch A1 Configurations and Verification....961
eVLT Configuration Example....962
PIM-Sparse Mode Configuration Example....964
Verifying a VLT Configuration....965
Additional VLT Sample Configurations....968
Troubleshooting VLT....969
Reconfiguring Stacked Switches as VLT....970
Specifying VLT Nodes in a PVLAN....971
Configuring a VLT VLAN or LAG in a PVLAN....974
Creating a VLT LAG or a VLT VLAN....974
Associating the VLT LAG or VLT VLAN in a PVLAN....975
Proxy ARP Capability on VLT Peer Nodes....975
VLT Nodes as Rendezvous Points for Multicast Resiliency....977
Configuring VLAN-Stack over VLT....977
IPv6 Peer Routing in VLT Domains Overview....980
Configure BFD in VLT Domain....984
Sample BFD configuration in VLT domain....984

Chapter 61: VLT Proxy Gateway....989
Proxy Gateway in VLT Domains....989
LLDP VLT Proxy Gateway in a Square VLT Topology....992
Configuring a Static VLT Proxy Gateway....993
Configuring an LLDP VLT Proxy Gateway....993
VLT Proxy Gateway Sample Topology....994
VLT Domain Configuration....994
Dell-1 VLT Configuration....994
Dell-2 VLT Configuration....995
Dell-3 VLT Configuration....996
Dell-4 VLT Configuration....997

Chapter 62: Virtual Routing and Forwarding (VRF)....998
VRF Overview....998
VRF Configuration Notes....999
DHCP....1001
VRF Configuration....1001
Loading VRF CAM....1001

Creating a Non-Default VRF Instance ... 1001
Assigning an Interface to a VRF ... 1001
Assigning a Front-end Port to a Management VRF ... 1001
View VRF Instance Information ... 1002
Assigning an OSPF Process to a VRF Instance ... 1002
Configuring VRRP on a VRF Instance ... 1002
Configuring Management VRF ... 1003
Configuring a Static Route ... 1004
Sample VRF Configuration ... 1004
Route Leaking VRFs ... 1009
Dynamic Route Leaking ... 1010
Configuring Route Leaking without Filtering Criteria ... 1010
Configuring Route Leaking with Filtering ... 1012

Chapter 63: Virtual Router Redundancy Protocol (VRRP) ... 1015
VRRP Overview ... 1015
VRRP Benefits ... 1016
VRRP Implementation ... 1016
VRRP Configuration ... 1017
Configuration Task List ... 1017
Setting VRRP Initialization Delay ... 1025
VRRP for IPv6 Configuration ... 1025
Sample Configurations ... 1029
VRRP in a VRF Configuration ... 1032

Chapter 64: Debugging and Diagnostics ... 1038
Offline Diagnostics ... 1038
Important Points to Remember ... 1038
Running Offline Diagnostics ... 1038
Trace Logs ... 1041
Auto Save on Crash or Rollover ... 1041
Last Restart Reason ... 1042
Hardware Watchdog Timer ... 1042
Using the Show Hardware Commands ... 1042
Enabling Environmental Monitoring ... 1043
Recognize an Overtemperature Condition ... 1044
Troubleshoot an Over-temperature Condition ... 1045
Recognize an Under-Voltage Condition ... 1045
Troubleshoot an Under-Voltage Condition ... 1045
Buffer Tuning ... 1046
Troubleshooting Packet Loss ... 1046
Displaying Drop Counters ... 1047
Dataplane Statistics ... 1048
Display Stack Port Statistics ... 1049
Display Stack Member Counters ... 1050
Enabling Application Core Dumps ... 1052
Mini Core Dumps ... 1052
Enabling TCP Dumps ... 1053

Chapter 65: Standards Compliance....1055
IEEE Compliance....1055
RFC and I-D Compliance....1056
General Internet Protocols....1056
General IPv4 Protocols....1057
General IPv6 Protocols....1058
Border Gateway Protocol (BGP)....1059
Open Shortest Path First (OSPF)....1060
Intermediate System to Intermediate System (IS-IS)....1061
Routing Information Protocol (RIP)....1062
Multicast....1062
Network Management....1062
MIB Location....1068

Chapter 66: X.509v3....1069
Introduction to X.509v3 certificates....1069
X.509v3 support in 1070
Information about installing CA certificates....1071
Installing CA certificate....1072
Information about Creating Certificate Signing Requests (CSR)....1072
Creating Certificate Signing Requests (CSR)....1072
Information about installing trusted certificates....1073
Installing trusted certificates....1073
Transport layer security (TLS)....1073
Syslog over TLS....1074
Online Certificate Status Protocol (OSCP)....1074
Configuring OCSP setting on CA....1074
Configuring OCSP behavior....1075
Configuring Revocation Behavior....1075
Configuring OSCP responder preference....1075
Verifying certificates....1075
Verifying Server certificates....1075
Verifying Client Certificates....1076
Event logging....1076

1

About this Guide

This guide describes the protocols and features the Dell EMC Networking Operating System (OS) supports and provides configuration instructions and examples for implementing them. For complete information about all the CLI commands, see the *Dell EMC Command Line Reference Guide* for your system.

The S4820T platform is available with Dell EMC Networking OS version 8.3.19.0 and beyond. The S4820T platform is available with Dell EMC Networking OS version 8.3.19.0 and beyond. S4820T stacking is supported with Dell EMC Networking OS version 8.3.19.0 and beyond.

Though this guide contains information about protocols, it is not intended to be a complete reference. This guide is a reference for configuring protocols on Dell EMC Networking systems. For complete information about protocols, see the related documentation, including Internet Engineering Task Force (IETF) requests for comments (RFCs). The instructions in this guide cite relevant RFCs. The Standards Compliance chapter contains a complete list of the supported RFCs and management information base files (MIBs).

Topics:

- Audience
- Conventions
- Related Documents

Audience

This document is intended for system administrators who are responsible for configuring and maintaining networks and assumes knowledge in Layer 2 (L2) and Layer 3 (L3) networking technologies.

Conventions

This guide uses the following conventions to describe command syntax.

Keyword	Keywords are in Courier (a monospaced font) and must be entered in the CLI as listed.
parameter	Parameters are in italics and require a number or word to be entered in the CLI.
{X}	Keywords and parameters within braces must be entered in the CLI.
[X]	Keywords and parameters within brackets are optional.
x|y	Keywords and parameters separated by a bar require you to choose one option.
x||y	Keywords and parameters separated by a double bar allows you to choose any or all of the options.

Related Documents

For more information about the Dell EMC Networking switches, see the following documents:

- *Dell EMC Networking OS Command Line Reference Guide*
- *Dell EMC Networking OS Installation Guide*
- *Dell EMC Networking OS Quick Start Guide*
- *Dell EMC Networking OS Release Notes*

2

Configuration Fundamentals

The Dell EMC Networking Operating System (OS) command line interface (CLI) is a text-based interface you can use to configure interfaces and protocols.

The CLI is largely the same for each platform except for some commands and command outputs. The CLI is structured in modes for security and management purposes. Different sets of commands are available in each mode, and you can limit user access to modes using privilege levels.

In the Dell EMC Networking OS, after you enter a command, the command is added to the running configuration file. You can view the current configuration for the whole system or for a particular CLI mode. To save the current configuration, copy the running configuration to another location.

NOTE: Due to differences in hardware architecture and continued system development, features may occasionally differ between the platforms. Differences are noted in each CLI description and related documentation.

Topics:

- Accessing the Command Line
- CLI Modes
- The do Command
- Undoing Commands
- Obtaining Help
- Entering and Editing Commands
- Command History
- Filtering show Command Outputs
- Multiple Users in Configuration Mode

Accessing the Command Line

Access the CLI through a serial console port or a Telnet session.

When the system successfully boots, enter the command line in EXEC mode.

NOTE: You must have a password configured on a virtual terminal line before you can Telnet into the system. Therefore, you must use a console connection when connecting to the system for the first time.

```
telnet 172.31.1.53
Trying 172.31.1.53...
Connected to 172.31.1.53.
Escape character is '^]'.
Login: username
Password:
DellEMC>
```

CLI Modes

Different sets of commands are available in each mode.

A command found in one mode cannot be executed from another mode (except for EXEC mode commands with a preceding `do` command (refer to the `do` Command section).

You can set user access rights to commands and command modes using privilege levels.

For more information about privilege levels and security options, refer to the *Privilege Levels Overview* section in the Security chapter.

The Dell EMC Networking OS CLI is divided into three major mode levels:

- EXEC mode is the default mode and has a privilege level of 1, which is the most restricted level. Only a limited selection of commands is available, notably the `show` commands, which allow you to view system information.
- EXEC Privilege mode has commands to view configurations, clear counters, manage configuration files, run diagnostics, and enable or disable debug operations. The privilege level is 15, which is unrestricted. You can configure a password for this mode; refer to the *Configure the Enable Password* section in the Getting Started chapter.
- CONFIGURATION mode allows you to configure security features, time settings, set logging and SNMP functions, configure static ARP and MAC addresses, and set line cards on the system.

Beneath CONFIGURATION mode are submodes that apply to interfaces, protocols, and features. The following example shows the submode command structure. Two sub-CONFIGURATION modes are important when configuring the chassis for the first time:

- INTERFACE submode is the mode in which you configure Layer 2 and Layer 3 protocols and IP services specific to an interface. An interface can be physical (Management interface, 1 Gigabit Ethernet, 10 Gigabit Ethernet, 25 Gigabit Ethernet, 40 Gigabit Ethernet, 50 Gigabit Ethernet, or 100 Gigabit Ethernet) or logical (Loopback, Null, port channel, or virtual local area network [VLAN]).
- LINE submode is the mode in which you to configure the console and virtual terminal lines.

NOTE: At any time, entering a question mark (`?`) displays the available command options. For example, when you are in CONFIGURATION mode, entering the question mark first lists all available commands, including the possible submodes.

The CLI modes are:

```
EXEC
    EXEC Privilege
        CONFIGURATION
            AS-PATH ACL
            CONTROL-PLANE
            CLASS-MAP
            DCB POLICY
            DHCP
              DHCP POOL
            ECMP-GROUP
            EXTENDED COMMUNITY
            FRRP
            INTERFACE
              GROUP
              GIGABIT ETHERNET
              10 GIGABIT ETHERNET
              40 GIGABIT ETHERNET
              INTERFACE RANGE
              LOOPBACK
              MANAGEMENT ETHERNET
              NULL
              PORT-CHANNEL
              TUNNEL
              VLAN
              VRRP
            IP
              IPv6
              IP COMMUNITY-LIST
              IP ACCESS-LIST
                STANDARD ACCESS-LIST
                EXTENDED ACCESS-LIST
                MAC ACCESS-LIST
            LINE
              AUXILLIARY
              CONSOLE
              VIRTUAL TERMINAL
            LLDP
              LLDP MANAGEMENT INTERFACE
            MONITOR SESSION
            MULTIPLE SPANNING TREE
            OPENFLOW INSTANCE
            PVST
            PORT-CHANNEL FAILOVER-GROUP
            PREFIX-LIST
            PRIORITY-GROUP
            PROTOCOL GVRP
            QOS POLICY
```

```
            RSTP
            ROUTE-MAP
            ROUTER BGP
              BGP ADDRESS-FAMILY
            ROUTER ISIS
              ISIS ADDRESS-FAMILY
            ROUTER OSPF
            ROUTER OSPFV3
            ROUTER RIP
            SPANNING TREE
SUPPORTASSIST
            TRACE-LIST
            VLT DOMAIN
            VRRP
            UPLINK STATE GROUP
uBoot
```

Navigating CLI Modes

The Dell EMC Networking OS prompt changes to indicate the CLI mode.

The following table lists the CLI mode, its prompt, and information about how to access and exit the CLI mode. Move linearly through the command modes, except for the `end` command which takes you directly to EXEC Privilege mode and the `exit` command which moves you up one command mode level.

NOTE: Sub-CONFIGURATION modes all have the letters `conf` in the prompt with more modifiers to identify the mode and slot/port information.

Table 1. Dell EMC Networking OS Command Modes

CLI Command Mode	Prompt	Access Command
EXEC	`DellEMC>`	Access the router through the console or terminal line.
EXEC Privilege	`DellEMC#`	• From EXEC mode, enter the `enable` command. • From any other mode, use the `end` command.
CONFIGURATION	`DellEMC(conf)#`	• From EXEC privilege mode, enter the `configure` command. • From every mode except EXEC and EXEC Privilege, enter the `exit` command.
NOTE: Access all of the following modes from CONFIGURATION mode.		
AS-PATH ACL	`DellEMC(config-as-path)#`	`ip as-path access-list`
10 Gigabit Ethernet Interface	`DellEMC(conf-if-te-0/0)#`	`interface` (INTERFACE modes)
40 Gigabit Ethernet Interface	`DellEMC(conf-if-fo-0/0)#`	`interface` (INTERFACE modes)
Interface Group	`DellEMC(conf-if-group)#`	`interface(INTERFACE modes)`
Interface Range	`DellEMC(conf-if-range)#`	`interface` (INTERFACE modes)
Loopback Interface	`DellEMC(conf-if-lo-0)#`	`interface` (INTERFACE modes)
Management Ethernet Interface	`DellEMC(conf-if-ma-0/0)#`	`interface` (INTERFACE modes)
Null Interface	`DellEMC(conf-if-nu-0)#`	`interface` (INTERFACE modes)
Port-channel Interface	`DellEMC(conf-if-po-1)#`	`interface` (INTERFACE modes)

Table 1. Dell EMC Networking OS Command Modes (continued)

CLI Command Mode	Prompt	Access Command
Tunnel Interface	`DellEMC(conf-if-tu-1)#`	`interface` (INTERFACE modes)
VLAN Interface	`DellEMC(conf-if-vl-1)#`	`interface` (INTERFACE modes)
STANDARD ACCESS-LIST	`DellEMC(config-std-nacl)#`	`ip access-list standard` (IP ACCESS-LIST Modes)
EXTENDED ACCESS-LIST	`DellEMC(config-ext-nacl)#`	`ip access-list extended` (IP ACCESS-LIST Modes)
IP COMMUNITY-LIST	`DellEMC(config-community-list)#`	`ip community-list`
AUXILIARY	`DellEMC(config-line-aux)#`	`line` (LINE Modes)
CONSOLE	`DellEMC(config-line-console)#`	`line` (LINE Modes)
VIRTUAL TERMINAL	`DellEMC(config-line-vty)#`	`line` (LINE Modes)
STANDARD ACCESS-LIST	`DellEMC(config-std-macl)#`	`mac access-list standard` (MAC ACCESS-LIST Modes)
EXTENDED ACCESS-LIST	`DellEMC(config-ext-macl)#`	`mac access-list extended` (MAC ACCESS-LIST Modes)
MULTIPLE SPANNING TREE	`DellEMC(config-mstp)#`	`protocol spanning-tree mstp`
Per-VLAN SPANNING TREE Plus	`DellEMC(config-pvst)#`	`protocol spanning-tree pvst`
PREFIX-LIST	`DellEMC(conf-nprefixl)#`	`ip prefix-list`
RAPID SPANNING TREE	`DellEMC(config-rstp)#`	`protocol spanning-tree rstp`
REDIRECT	`DellEMC(conf-redirect-list)#`	`ip redirect-list`
ROUTE-MAP	`DellEMC(config-route-map)#`	`route-map`
ROUTER BGP	`DellEMC(conf-router_bgp)#`	`router bgp`
BGP ADDRESS-FAMILY	`DellEMC(conf-router_bgp_af)#` (for IPv4) `DellEMC(conf-routerZ_bgpv6_af)#` (for IPv6)	`address-family {ipv4 multicast | ipv6 unicast}` (ROUTER BGP Mode)
ROUTER ISIS	`DellEMC(conf-router_isis)#`	`router isis`
ISIS ADDRESS-FAMILY	`DellEMC(conf-router_isis-af_ipv6)#`	`address-family ipv6 unicast` (ROUTER ISIS Mode)
ROUTER OSPF	`DellEMC(conf-router_ospf)#`	`router ospf`
ROUTER OSPFV3	`DellEMC(conf-ipv6router_ospf)#`	`ipv6 router ospf`
ROUTER RIP	`DellEMC(conf-router_rip)#`	`router rip`
SPANNING TREE	`DellEMC(config-span)#`	`protocol spanning-tree 0`
TRACE-LIST	`DellEMC(conf-trace-acl)#`	`ip trace-list`
CLASS-MAP	`DellEMC(config-class-map)#`	`class-map`
CONTROL-PLANE	`DellEMC(conf-control-cpuqos)#`	`control-plane-cpuqos`
DHCP	`DellEMC(config-dhcp)#`	`ip dhcp server`
DHCP POOL	`DellEMC(config-dhcp-`*`pool-name`*`)#`	`pool` (DHCP Mode)

Table 1. Dell EMC Networking OS Command Modes (continued)

CLI Command Mode	Prompt	Access Command
ECMP	`DellEMC(conf-ecmp-group-`*`ecmp-group-id`*`)#`	`ecmp-group`
EIS	`DellEMC(conf-mgmt-eis)#`	`management egress-interface-selection`
FRRP	`DellEMC(conf-frrp-`*`ring-id`*`)#`	`protocol frrp`
LLDP	`DellEMC(conf-lldp)#` or `DellEMC(conf-if-`*`interface`*`-lldp)#`	`protocol lldp` (CONFIGURATION or INTERFACE Modes)
LLDP MANAGEMENT INTERFACE	`DellEMC(conf-lldp-mgmtIf)#`	`management-interface` (LLDP Mode)
LINE	`DellEMC(config-line-console)` or `DellEMC(config-line-vty)`	`line console` or`line vty`
MONITOR SESSION	`DellEMC(conf-mon-sess-`*`sessionID`*`)#`	`monitor session`
OPENFLOW INSTANCE	`DellEMC(conf-of-instance-`*`of-id`*`)#`	`openflow of-instance`
PORT-CHANNEL FAILOVER-GROUP	`DellEMC(conf-po-failover-grp)#`	`port-channel failover-group`
PRIORITY GROUP	`DellEMC(conf-pg)#`	`priority-group`
PROTOCOL GVRP	`DellEMC(config-gvrp)#`	`protocol gvrp`
QOS POLICY	`DellEMC(conf-qos-policy-out-ets)#`	`qos-policy-output`
SUPPORTASSIST	`DellEMC(support-assist)#`	`support-assist`
VLT DOMAIN	`DellEMC(conf-vlt-domain)#`	`vlt domain`
VRRP	`DellEMC(conf-if-`*`interface-type-slot/port`*`-vrid-`*`vrrp-group-id`*`)#`	`vrrp-group`
u-Boot	`Dell=>`	Press any key when the following line appears on the console during a system boot: `Hit any key to stop autoboot:`
UPLINK STATE GROUP	`DellEMC(conf-uplink-state-group-`*`groupID`*`)#`	`uplink-state-group`

The following example shows how to change the command mode from CONFIGURATION mode to PROTOCOL SPANNING TREE.

Example of Changing Command Modes

```
DellEMC(conf)#protocol spanning-tree 0
DellEMC(config-span)#
```

The do Command

You can enter an EXEC mode command from any CONFIGURATION mode (CONFIGURATION, INTERFACE, SPANNING TREE, and so on.) without having to return to EXEC mode by preceding the EXEC mode command with the `do` command.

The following example shows the output of the `do` command.

```
DellEMC(conf)#do show system brief

Stack MAC : 00:01:e8:00:66:64
Reload-Type                :   normal-reload [Next boot : normal-reload]

--  Stack Info  --
Unit  UnitType     Status         ReqTyp     CurTyp   Version     Ports
-------------------------------------------------------------------------
  0   Management   online         S4810      S4810    9.4(0.0)    64
  1   Member       not present
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
  8   Member       not present
  9   Member       not present
 10   Member       not present
 11   Member       not present

--  Power Supplies  --
Unit   Bay   Status       Type    FanStatus
-----------------------------------------------------------
  0     0    absent               absent
  0     1    up           UNKNOWN up

--  Fan  Status  --
Unit Bay   TrayStatus  Fan0    Speed   Fan1    Speed
-----------------------------------------------------------------
 0    0     up          up      9120    up      9120
 0    1     up          up      9120    up      9120

Speed in RPM

DellEMC(conf)#
```

Undoing Commands

When you enter a command, the command line is added to the running configuration file (running-config).

To disable a command and remove it from the running-config, enter the `no` command, then the original command. For example, to delete an IP address configured on an interface, use the `no ip address` *`ip-address`* command.

NOTE: Use the help or `?` command as described in Obtaining Help.

Example of Viewing Disabled Commands

```
DellEMC(conf)#interface tengigabitethernet 2/17
DellEMC(conf-if-te-2/17)#ip address 192.168.10.1/24
DellEMC(conf-if-te-2/17)#show config
!
  interface TenGigabitEthernet 2/17
  ip address 192.168.10.1/24
no shutdown
DellEMC(conf-if-te-2/17)#no ip address
DellEMC(conf-if-te-2/17)#show config
!
interface TenGigabitEthernet 2/17
```

```
no ip address
no shutdown
```

Layer 2 protocols are disabled by default. To enable Layer 2 protocols, use the `no disable` command. For example, in PROTOCOL SPANNING TREE mode, enter `no disable` to enable Spanning Tree.

Obtaining Help

Obtain a list of keywords and a brief functional description of those keywords at any CLI mode using the `?` or help command:

- To list the keywords available in the current mode, enter `?` at the prompt or after a keyword.
- Enter `?` after a command prompt to list all of the available keywords. The output of this command is the same as the `help` command.

```
DellEMC#?
bmp                             BMP commands
cd                              Change current directory
clear                           Reset functions
clock                           Manage the system clock
```

- Enter `?` after a partial keyword lists all of the keywords that begin with the specified letters.

```
DellEMC(conf)#cl?
class-map
clock
DellEMC(conf)#cl
```

- Enter `[space]?` after a keyword lists all of the keywords that can follow the specified keyword.

```
DellEMC(conf)#clock ?
summer-time       Configure summer (daylight savings) time
timezone          Configure time zone
DellEMC(conf)#clock
```

Entering and Editing Commands

Notes for entering commands.

- The CLI is not case-sensitive.
- You can enter partial CLI keywords.
  - Enter the minimum number of letters to uniquely identify a command. For example, you cannot enter `cl` as a partial keyword because both the `clock` and `class-map` commands begin with the letters “cl.” You can enter `clo`, however, as a partial keyword because only one command begins with those three letters.
- The TAB key auto-completes keywords in commands. Enter the minimum number of letters to uniquely identify a command.
- The UP and DOWN arrow keys display previously entered commands (refer to Command History).
- The BACKSPACE and DELETE keys erase the previous letter.
- Key combinations are available to move quickly across the command line. The following table describes these short-cut key combinations.

Short-Cut Key Combination	Action
CNTL-A	Moves the cursor to the beginning of the command line.
CNTL-B	Moves the cursor back one character.
CNTL-D	Deletes character at cursor.
CNTL-E	Moves the cursor to the end of the line.
CNTL-F	Moves the cursor forward one character.
CNTL-I	Completes a keyword.
CNTL-K	Deletes all characters from the cursor to the end of the command line.

Short-Cut Key Combination	Action
CNTL-L	Re-enters the previous command.
CNTL-N	Return to more recent commands in the history buffer after recalling commands with CTRL-P or the UP arrow key.
CNTL-P	Recalls commands, beginning with the last command.
CNTL-R	Re-enters the previous command.
CNTL-U	Deletes the line.
CNTL-W	Deletes the previous word.
CNTL-X	Deletes the line.
CNTL-Z	Ends continuous scrolling of command outputs.
Esc B	Moves the cursor back one word.
Esc F	Moves the cursor forward one word.
Esc D	Deletes all characters from the cursor to the end of the word.

Command History

The Dell EMC Networking OS maintains a history of previously-entered commands for each mode. For example:

- When you are in EXEC mode, the UP and DOWN arrow keys display the previously-entered EXEC mode commands.
- When you are in CONFIGURATION mode, the UP or DOWN arrows keys recall the previously-entered CONFIGURATION mode commands.

Filtering show Command Outputs

Filter the output of a `show` command to display specific information by adding `| [except | find | grep | no-more | save]` *`specified_text`* after the command.

The variable *`specified_text`* is the text for which you are filtering and it IS case sensitive unless you use the `ignore-case` sub-option.

Starting with Dell EMC Networking OS version 7.8.1.0, the `grep` command accepts an `ignore-case` sub-option that forces the search to case-insensitive. For example, the commands:

- `show run | grep Ethernet` returns a search result with instances containing a capitalized “Ethernet,” such as `interface TenGigabitEthernet`.
- `show run | grep ethernet` does not return that search result because it only searches for instances containing a non-capitalized “ethernet.”
- `show run | grep Ethernet ignore-case` returns instances containing both “Ethernet” and “ethernet.”

The `grep` command displays only the lines containing specified text. The following example shows this command used in combination with the `show system brief` command.

Example of the grep Keyword

```
DellEMC(conf)#do show system brief | grep 0
0      not present
```

NOTE: Dell EMC Networking OS accepts a space or no space before and after the pipe. To filter a phrase with spaces, underscores, or ranges, enclose the phrase with double quotation marks.

The `except` keyword displays text that does not match the specified text. The following example shows this command used in combination with the `show system brief` command.

Example of the `except` Keyword

```
DellEMC#show system brief | except 1

Stack MAC                   : 4c:76:25:e5:49:40
Reload-Type                 : normal-reload [Next boot : normal-reload]
```

The `find` keyword displays the output of the `show` command beginning from the first occurrence of specified text. The following example shows this command used in combination with the `show system brief` command.

Example of the `find` Keyword

The `display` command displays additional configuration information.

The `no-more` command displays the output all at once rather than one screen at a time. This is similar to the `terminal length` command except that the `no-more` option affects the output of the specified command only.

The `save` command copies the output to a file for future reference.

NOTE: You can filter a single command output multiple times. The `save` option must be the last option entered. For example: `DellEMC# command | grep regular-expression | except regular-expression | grep other-regular-expression | find regular-expression | save`.

Multiple Users in Configuration Mode

Dell EMC Networking OS notifies all users when there are multiple users logged in to CONFIGURATION mode.

A warning message indicates the username, type of connection (console or VTY), and in the case of a VTY connection, the IP address of the terminal on which the connection was established. For example:

- On the system that telnets into the switch, this message appears:

```
% Warning: The following users are currently configuring the system:
User "<username>" on line console0
```

- On the system that is connected over the console, this message appears:

```
% Warning: User "<username>" on line vty0 "10.11.130.2" is in configuration mode
```

If either of these messages appears, Dell EMC Networking recommends coordinating with the users listed in the message so that you do not unintentionally overwrite each other's configuration changes.

3

Getting Started

This chapter describes how you start configuring your system.

When you power up the chassis, the system performs a power-on self test (POST) and system then loads the Dell EMC Networking Operating System. Boot messages scroll up the terminal window during this process. No user interaction is required if the boot process proceeds without interruption.

When the boot process completes, the system status LEDs remain online (green) and the console monitor displays the EXEC mode prompt.

For details about using the command line interface (CLI), refer to the Accessing the Command Line section in the Configuration Fundamentals chapter.

Topics:

- Console Access
- Accessing the CLI Interface and Running Scripts Using SSH
- Default Configuration
- Configuring a Host Name
- Accessing the System Remotely
- Configuring the Enable Password
- Configuration File Management
- Managing the File System
- Enabling Software Features on Devices Using a Command Option
- View Command History
- Upgrading Dell EMC Networking OS
- Using HTTP for File Transfers
- Verify Software Images Before Installation

Console Access

The device has two management ports available for system access: a serial RS-232 /RJ-45 console port and an out-of-band (OOB) Ethernet port to manage the switch with an IP address.

Serial Console

The RJ-45/RS-232 console port is labeled on the upper right-hand side, as you face the I/O side of the chassis.

Figure 1. RJ-45 Console Port

Accessing the Console Port

To access the console port, follow these steps:

For the console port pinout, refer to Accessing the RJ-45 Console Port with a DB-9 Adapter.

1. Install an RJ-45 copper cable into the console port. Use a rollover (crossover) cable to connect the console port to a terminal server.
2. Connect the other end of the cable to the DTE terminal server.
3. Terminal settings on the console port cannot be changed in the software and are set as follows:
   - 9600 baud rate
   - No parity
   - 8 data bits
   - 1 stop bit
   - No flow control

Pin Assignments

You can connect to the console using a RJ-45 to RJ-45 rollover cable and a RJ-45 to DB-9 female DTE adapter to a terminal server (for example, a PC).

The pin assignments between the console and a DTE terminal server are as follows:

Table 2. Pin Assignments Between the Console and a DTE Terminal Server

Console Port	RJ-45 to RJ-45 Rollover Cable	RJ-45 to RJ-45 Rollover Cable	RJ-45 to DB-9 Adapter	Terminal Server Device
Signal	**RJ-45 Pinout**	**RJ-45 Pinout**	**DB-9 Pin**	**Signal**
RTS	1	8	8	CTS
NC	2	7	6	DSR
TxD	3	6	2	RxD
GND	4	5	5	GND
GND	5	4	5	GND
RxD	6	3	3	TxD
NC	7	2	4	DTR
CTS	8	1	7	RTS

Accessing the CLI Interface and Running Scripts Using SSH

In addition to the capability to access a device using a console connection or a Telnet session, you can also use SSH for secure, protected communication with the device. You can open an SSH session and run commands or script files. This method of connectivity is supported with S4810, S4048–ON, S3048–ON, S4820T, and Z9000 switches and provides a reliable, safe communication mechanism.

Entering CLI commands Using an SSH Connection

You can run CLI commands by entering any one of the following syntax to connect to a switch using the preconfigured user credentials using SSH:

```
ssh username@hostname <CLI Command>
```

or

```
echo <CLI Command> | ssh admin@hostname
```

The SSH server transmits the terminal commands to the CLI shell and the results are displayed on the screen non-interactively.

Executing Local CLI Scripts Using an SSH Connection

You can execute CLI commands by entering a CLI script in one of the following ways:

```
ssh username@hostname <CLIscript.file>
```

or

```
cat < CLIscript.file > | ssh admin@hostname
```

The script is run and the actions contained in the script are performed.

Following are the points to remember, when you are trying to establish an SSH session to the device to run commands or script files:

- There is an upper limit of 10 concurrent sessions in SSH. Therefore, you might expect a failure in executing SSH-related scripts.
- To avoid denial of service (DoS) attacks, a rate-limit of 10 concurrent sessions per minute in SSH is devised. Therefore, you might experience a failure in executing SSH-related scripts when multiple short SSH commands are executed.
- If you issue an interactive command in the SSH session, the behavior may not really be interactive.
- In some cases, when you use an SSH session, when certain show commands such as `show tech-support` produce large volumes of output, sometimes few characters from the output display are truncated and not displayed. This may cause one of the commands to fail for syntax error. In such cases, if you add few newline characters before the failed command, the output displays completely.

Execution of commands on CLI over SSH does not notice the errors that have occurred while executing the command. As a result, you cannot identify, whether a command has failed to be processed. The console output though is redirected back over SSH.

Default Configuration

Although a version of Dell EMC Networking OS is pre-loaded onto the system, the system is not configured when you power up the system first time (except for the default hostname, which is `DellEMC`). You must configure the system using the CLI.

Configuring a Host Name

The host name appears in the prompt. The default host name is `Dell`.

- Host names must start with a letter and end with a letter or digit.
- Characters within the string can be letters, digits, and hyphens.

To create a host name, use the `hostname` *`name`* command in Configuration mode.

```
Dell(conf)#hostname R1
R1(conf)#
```

Accessing the System Remotely

You can configure the system to access it remotely by Telnet or secure shell (SSH).

- The platform has a dedicated management port and a management routing table that is separate from the IP routing table.
- You can manage all Dell EMC Networking products in-band via the front-end data ports through interfaces assigned an IP address as well.

Accessing the System Remotely

Configuring the system for remote access is a three-step process, as described in the following topics:

1. Configure an IP address for the management port. Configure the Management Port IP Address
2. Configure a management route with a default gateway. Configure a Management Route
3. Configure a username and password. Configure a Username and Password

Configure the Management Port IP Address

To access the system remotely, assign IP addresses to the management ports.

1. Enter INTERFACE mode for the Management port.

   CONFIGURATION mode

   `interface ManagementEthernet slot/port`
2. Assign an IP address to the interface.

   INTERFACE mode

   `ip address ip-address/mask`
   - *ip-address*: an address in dotted-decimal format (A.B.C.D).
   - *mask*: a subnet mask in /prefix-length format (/ xx).
3. Enable the interface.

   INTERFACE mode

   `no shutdown`

Configure a Management Route

Define a path from the system to the network from which you are accessing the system remotely. Management routes are separate from IP routes and are only used to manage the system through the management port.

To configure a management route, use the following command.

- Configure a management route to the network from which you are accessing the system.

  CONFIGURATION mode

  `management route ip-address/mask gateway`
  - *ip-address*: the network address in dotted-decimal format (A.B.C.D).
  - *mask*: a subnet mask in /prefix-length format (/ xx).
  - *gateway*: the next hop for network traffic originating from the management port.

Configuring a Username and Password

To access the system remotely, configure a system username and password.

To configure a system username and password, use the following command.

- Configure a username and password to access the system remotely.

  CONFIGURATION mode

  `username name [access-class access-list-name] [nopassword | {password | secret | sha256-password} [encryption-type] password [dynamic-salt]] [privilege level] [role role-name]`
  - *name*: Enter a text string upto 63 characters long.
  - `access-class` *access-list-name*: Enter the name of a configured IP ACL.
  - `nopassword`: Allows you to configure an user without the password.
  - `password`: Allows you to configure an user with a password.
  - `secret`: Specify a secret string for an user.

- `sha256-password`: Uses sha256–based encryption method for password.
- *`encryption-type`*: Enter the encryption type for securing an user password. There are four encryption types.
  - 0 — input the password in clear text.
  - 5 — input the password that is already encrypted using MD5 encryption method.
  - 7 — input the password that is already encrypted using DES encryption method.
  - 8 — input the password that is already encrypted using sha256–based encryption method.
- *`password`*: Enter the password string for the user.
- `dynamic-salt`: Generates an additional random input to password encryption process whenever the password is configured.
- `privilege` *`level`*: Assign a privilege levels to the user. The range is from 0 to 15.
- `role` *`role-name`*: Assign a role name for the user.

Dell EMC Networking OS encrypts type 5 secret and type 7 password based on `dynamic-salt` option such that the encrypted password is different when an user is configured with the same password.

NOTE:

`dynamic-salt` option is shown only with `secret` and `password` options.

In `dynamic-salt` configuration, the length of type 5 secret and type 7 password is 32 and 16 characters more compared to the secret and password length without dynamic-salt configuration. An error message appears if the `username` command reaches the maximum length, which is 256 characters.

The `dynamic-salt` support for the user configuration is added in REST API. For more information on REST support, see *Dell EMC Networking Open Automation guide*.

Configuring the Enable Password

Access EXEC Privilege mode using the `enable` command. EXEC Privilege mode is unrestricted by default. Configure a password as a basic security measure.

There are three types of enable passwords:

- `enable password` is stored in the running/startup configuration using a DES encryption method.
- `enable secret` is stored in the running/startup configuration using MD5 encryption method.
- `enable sha256-password` is stored in the running/startup configuration using sha256-based encryption method (PBKDF2).

Dell EMC Networking recommends using the `enable sha256-password` password.

To configure an enable password, use the following command.

- Create a password to access EXEC Privilege mode.

  CONFIGURATION mode

  `enable [password | secret | sha256-password] [level` *`level`*`] [`*`encryption-type`*`] `*`password`*

  - *`level`*: is the privilege level, is 15 by default, and is not required.
  - *`encryption-type`*: specifies how you input the password, is 0 by default, and is not required.
    - 0 is to input the password in clear text.
    - 5 is to input a password that is already encrypted using MD5 encryption method. Obtain the encrypted password from the configuration file of another device.
    - 7 is to input a password that is already encrypted using DES encryption method. Obtain the encrypted password from the configuration file of another device.
    - 8 is to input a password that is already encrypted using sha256-based encryption method. Obtain the encrypted password from the configuration file of another device.

Configuration File Management

Files can be stored on and accessed from various storage media. Rename, delete, and copy files on the system from EXEC Privilege mode.

Copy Files to and from the System

The command syntax for copying files is similar to UNIX. The copy command uses the format `copy source-file-url destination-file-url`.

NOTE: For a detailed description of the copy command, refer to the *Dell EMC Networking OS Command Reference*.

- To copy a local file to a remote system, combine the file-origin syntax for a local file location with the file-destination syntax for a remote file location.
- To copy a remote file to Dell EMC Networking system, combine the file-origin syntax for a remote file location with the file-destination syntax for a local file location.

Table 3. Forming a `copy` Command

Location	***source-file-url* Syntax**	***destination-file-url* Syntax**
For a remote file location: FTP server	`copy ftp://username:password@{hostip | hostname}/filepath/filename`	`ftp://username:password@{hostip | hostname}/ filepath/filename`
For a remote file location: TFTP server	`copy tftp://{hostip | hostname}/filepath/ filename`	`tftp://{hostip | hostname}/filepath/filename`
For a remote file location: SCP server	`copy scp://{hostip | hostname}/filepath/ filename`	`scp://{hostip | hostname}/filepath/filename`

Important Points to Remember

- You may not copy a file from one remote system to another.
- You may not copy a file from one location to the same location.
- When copying to a server, you can only use a hostname if a domain name server (DNS) server is configured.

Example of Copying a File to an FTP Server

```
DellEMC#copy flash://Dell-EF-8.2.1.0.bin ftp://myusername:mypassword@10.10.10.10/
/Dell/Dell-EF-8.2.1.0
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
27952672 bytes successfully copied
```

Example of Importing a File to the Local System

```
core1#$//copy ftp://myusername:mypassword@10.10.10.10//Dell/
Dell-EF-8.2.1.0.bin flash://
Destination file name [Dell-EF-8.2.1.0.bin.bin]:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
26292881 bytes successfully copied
```

Mounting an NFS File System

This feature enables you to quickly access data on an NFS mounted file system. You can perform file operations on an NFS mounted file system using supported file commands.

This feature allows an NFS mounted device to be recognized as a file system. This file system is visible on the device and you can execute all file commands that are available on conventional file systems such as a Flash file system.

Before executing any CLI command to perform file operations, you must first mount the NFS file system to a mount-point on the device. Since multiple mount-points exist on a device, it is mandatory to specify the mount-point to which you want to load the system. The `/f10/mnt/nfs` directory is the root of all mount-points.

To mount an NFS file system, perform the following steps:

Table 4. Mounting an NFS File System

File Operation	Syntax	
To mount an NFS file system:	`mount nfs` *`rhost:path mount-point`* `username password`	

The foreign file system remains mounted as long as the device is up and does not reboot. You can run the file system commands without having to mount or un-mount the file system each time you run a command. When you save the configuration using the `write` command, the `mount` command is saved to the startup configuration. As a result, each time the device re-boots, the NFS file system is mounted during start up.

Table 5. Forming a `copy` Command

Location	***source-file-url*** **Syntax**	***destination-file-url*** **Syntax**
For a remote file location: NFS File System	`copy nfsmount://{<mount-point>}/filepath/filename} username:password`	`tftp://`*`{hostip | hostname}/filepath/filename`*

Important Points to Remember

- You cannot copy a file from one remote system to another.
- You cannot copy a file from one location to the same location.
- When copying to a server, you can only use a hostname if a domain name server (DNS) server is configured.

Example of Copying a File to current File System

```
DellEMC#copy tftp://10.16.127.35/dv-maa-test nfsmount://
Destination file name [dv-maa-test]:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!.!
44250499 bytes successfully copied
DellEMC#
DellEMC#copy ftp://10.16.127.35 nfsmount:
Source file name []: test.c
User name to login remote host: username
```

Example of Logging in to Copy from NFS Mount

```
DellEMC#copy nfsmount:///test flash:
Destination file name [test]: test2
!
5592 bytes successfully copied
DellEMC#
DellEMC#copy nfsmount:///test.txt ftp://10.16.127.35
Destination file name [test.txt]:
User name to login remote host: username
Password to login remote host:
!
```

Example of Copying to NFS Mount

```
DellEMC#copy flash://test.txt nfsmount:///
Destination file name [test.txt]:
!
15 bytes successfully copied
DellEMC#copy flash://test/capture.txt.pcap nfsmount:///
Destination file name [test.txt]:
!
15 bytes successfully copied
DellEMC#copy flash://test/capture.txt.pcap nfsmount:///username/snoop.pcap
```

```
!
24 bytes successfully copied
DellEMC#
DellEMC#copy tftp://10.16.127.35/username/dv-maa-test ?
flash:                  Copy to local file system ([flash://]filepath)
nfsmount:               Copy to nfs mount file system (nfsmount:///filepath)
running-config          remote host:
Destination file name [test.c]:
!
225 bytes successfully copied
DellEMC#
```

Save the Running-Configuration

The running-configuration contains the current system configuration. Dell EMC Networking recommends coping your running-configuration to the startup-configuration.

The commands in this section follow the same format as those commands in the Copy Files to and from the System section but use the filenames startup-configuration and running-configuration. These commands assume that current directory is the internal flash, which is the system default.

- Save the running-configuration to the startup-configuration on the internal flash of the primary RPM.

  EXEC Privilege mode

  ```
  copy running-config startup-config
  ```
- Save the running-configuration to an FTP server.

  EXEC Privilege mode

  ```
  copy running-config ftp:// username:password@{hostip | hostname}/filepath/ filename
  ```
- Save the running-configuration to a TFTP server.

  EXEC Privilege mode

  ```
  copy running-config tftp://{hostip | hostname}/ filepath/filename
  ```
- Save the running-configuration to an SCP server.

  EXEC Privilege mode

  ```
  copy running-config scp://{hostip | hostname}/ filepath/filename
  ```

  NOTE: When copying to a server, a host name can only be used if a DNS server is configured.

NOTE: When you load the startup configuration or a configuration file from a network server such as TFTP to the running configuration, the configuration is added to the running configuration. This does not replace the existing running configuration. Commands in the configuration file has precedence over commands in the running configuration.

Configure the Overload Bit for a Startup Scenario

For information about setting the router overload bit for a specific period of time after a switch reload is implemented, see the *Intermediate System to Intermediate System (IS-IS)* section in the *Dell Command Line Reference Guide* for your system.

Viewing Files

You can only view file information and content on local file systems.

To view a list of files or the contents of a file, use the following commands.

- View a list of files on the internal flash.

  EXEC Privilege mode

  ```
  dir flash:
  ```
- View the running-configuration.

  EXEC Privilege mode

  ```
  show running-config
  ```
- View the startup-configuration.

EXEC Privilege mode

```
show startup-config
```

The output of the `dir` command also shows the read/write privileges, size (in bytes), and date of modification for each file.

```
DellEMC#dir
Directory of flash:

 1 drw-    32768 Jan 01 1980 00:00:00 .
 2 drwx      512 Jul 23 2007 00:38:44 ..
 3 drw-     8192 Mar 30 1919 10:31:04 TRACE_LOG_DIR
 4 drw-     8192 Mar 30 1919 10:31:04 CRASH_LOG_DIR
 5 drw-     8192 Mar 30 1919 10:31:04 NVTRACE_LOG_DIR
 6 drw-     8192 Mar 30 1919 10:31:04 CORE_DUMP_DIR
 7 d---     8192 Mar 30 1919 10:31:04 ADMIN_DIR
 8 -rw- 33059550 Jul 11 2007 17:49:46 FTOS-EF-7.4.2.0.bin
 9 -rw- 27674906 Jul 06 2007 00:20:24 FTOS-EF-4.7.4.302.bin
10 -rw- 27674906 Jul 06 2007 19:54:52 boot-image-FILE
11 drw-     8192 Jan 01 1980 00:18:28 diag
12 -rw-     7276 Jul 20 2007 01:52:40 startup-config.bak
13 -rw-     7341 Jul 20 2007 15:34:46 startup-config
14 -rw- 27674906 Jul 06 2007 19:52:22 boot-image
15 -rw- 27674906 Jul 06 2007 02:23:22 boot-flash
--More--
```

View Configuration Files

Configuration files have three commented lines at the beginning of the file, as shown in the following example, to help you track the last time any user made a change to the file, which user made the changes, and when the file was last saved to the startup-configuration.

In the running-configuration file, if there is a difference between the timestamp on the “Last configuration change” and “Startup-config last updated,” you have made changes that have not been saved and are preserved after a system reboot.

Example of the `show running-config` Command

```
DellEMC#show running-config
Current Configuration ...
! Version 9.4(0.0)
! Last configuration change at Tue Mar 11 21:33:56 2014 by admin
! Startup-config last updated at Tue Mar 11 12:11:00 2014 by default
!
<output truncated for brevity>
```

Managing the File System

The Dell EMC Networking system can use the internal Flash, external Flash, or remote devices to store files.

The system stores files on the internal Flash by default but can be configured to store files elsewhere.

To view file system information, use the following command.

- View information about each file system.

  EXEC Privilege mode

  ```
  show file-systems
  ```

The output of the `show file-systems` command in the following example shows the total capacity, amount of free memory, file structure, media type, read/write privileges for each storage device in use.

```
DellEMC#show file-systems
Size(b)    Free(b)    Feature Type Flags Prefixes
520962048 213778432 dosFs2.0 USERFLASH rw flash:
127772672  21936128 dosFs2.0 USERFLASH rw slot0:
        -         -        - network   rw ftp:
```

```
-          -          - network    rw tftp:
-          -          - network    rw scp:
```

You can change the default file system so that file management commands apply to a particular device or memory.

To change the default directory, use the following command.

- Change the default directory.

  EXEC Privilege mode

  `cd directory`

Enabling Software Features on Devices Using a Command Option

The capability to activate software applications or components on a device using a command is supported on this platform.

Starting with Release 9.4(0.0), you can enable or disable specific software features or applications that need to run on a device by using a command attribute in the CLI interface. This enables effective, streamlined management and administration of applications and utilities that run on a device. You can employ this capability to perform an on-demand activation, or turn-off a software component or protocol. A feature configuration file generated for each image contains feature names, and denotes if this enabling or disabling method is available. You can enable or disable the VRF application globally across the system by using this capability.

Activate the VRF application on a device by using the `feature vrf` command in CONFIGURATION mode.

NOTE: The `no feature vrf` command is not supported on any of the platforms.

To enable the VRF feature and cause all VRF-related commands to be available or viewable in the CLI interface, use the following command. You must enable the VRF feature before you can configure its related attributes.

```
DellEMC(conf)# feature vrf
```

Based on if the VRF feature is identified as supported in the Feature Configuration file, configuration command `feature vrf` becomes available for usage. This command is stored in the running-configuration and precedes all other VRF-related configurations.

To display the state of Dell EMC Networking OS features:

```
DellEMC# show feature
```

Example of show feature output

For a particular target where VRF is enabled, the show output is similar to the following:

```
Feature          State
-----------------------
VRF              Enabled
```

View Command History

The command-history trace feature captures all commands entered by all users of the system with a time stamp and writes these messages to a dedicated trace log buffer.

The system generates a trace message for each executed command. No password information is saved to the file.

NOTE:

The timestamps display format of the `show command history` output changes based on the `service timestamps log datetime` configuration. The time format can be in uptime, local time zone time or UTC time.

If timestamp is disabled (`no service timestamps log`) then command history time format is shown with timestamp defaults (`service timestamps log datetime localtime`).

To view the command-history trace, use the `show command-history` command.

Example of the `show command-history` Command

Example 1: Default configuration `service timestamps log datetime` or `service timestamps log datetime localtime`

```
DellEMC(conf)#service timestamps log datetime
```

```
DellEMC# show command-history
         - Repeated 1 time.
[May 17 15:38:55]: CMD-(CLI):[service timestamps log datetime]by default from console
[May 17 15:41:40]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[May 17 15:41:45]: CMD-(CLI):[interface tengigabitethernet 1/1]by default from console
[May 17 15:41:47]: CMD-(CLI):[shutdown]by default from console
[May 17 15:41:50]: CMD-(CLI):[no shutdown]by default from console
[May 17 15:42:42]: CMD-(CLI):[show clock]by default from console
[May 17 15:42:52]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[May 17 15:43:08]: CMD-(CLI):[end]by default from console
[May 17 15:43:16]: CMD-(CLI):[show logging]by default from console
[May 17 15:43:22]: CMD-(CLI):[show command-history]by default from console
DellEMC#
```

Example 2: `service timestamps log datetime utc`

```
DellEMC(conf)#service timestamps log datetime utc
```

```
DellEMC# show command-history
         - Repeated 1 time.
[May 17 15:46:44]: CMD-(CLI):[configure]by default from console
         - Repeated 1 time.
[May 17 10:16:53]: CMD-(CLI):[service timestamps log datetime utc]by default from console
[May 17 10:17:05]: CMD-(CLI):[show clock]by default from console
[May 17 10:17:20]: CMD-(CLI):[show running-config]by default from console
[May 17 10:17:30]: CMD-(CLI):[interface tengigabitethernet 1/2]by default from console
[May 17 10:17:32]: CMD-(CLI):[shutdown]by default from console
[May 17 10:17:34]: CMD-(CLI):[no shutdown]by default from console
[May 17 10:17:40]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[May 17 10:17:46]: CMD-(CLI):[end]by default from console
[May 17 10:17:50]: CMD-(CLI):[show logging]by default from console
[May 17 10:17:56]: CMD-(CLI):[show command-history]by default from console
```

Example 3: `service timestamps log uptime`

```
DellEMC(conf)#service timestamps log uptime
```

```
DellEMC# show command-history
         - Repeated 1 time.
[May 17 10:20:37]: CMD-(CLI):[configure]by default from console
         - Repeated 1 time.
[1d0h24m]: CMD-(CLI):[service timestamps log uptime]by default from console
[1d0h24m]: CMD-(CLI):[interface tengigabitethernet 1/1]by default from console
[1d0h24m]: CMD-(CLI):[shutdown]by default from console
[1d0h24m]: CMD-(CLI):[no shutdown]by default from console
[1d0h25m]: CMD-(CLI):[end]by default from console
[1d0h25m]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[1d0h25m]: CMD-(CLI):[show clock]by default from console
[1d0h25m]: CMD-(CLI):[show version]by default from console
[1d0h25m]: CMD-(CLI):[show logging]by default from console
[1d0h25m]: CMD-(CLI):[show command-history]by default from console
```

Example 4: `no service timestamps log`

```
DellEMC(conf)#no service timestamps log
```

```
DellEMC# show command-history
         - Repeated 1 time.
[1d0h26m]: CMD-(CLI):[configure]by default from console
         - Repeated 1 time.
[May 17 15:53:10]: CMD-(CLI):[no service timestamps log]by default from console
```

```
[May 17 15:53:16]: CMD-(CLI):[write memory]by default from console
         - Repeated 3 times.
[May 17 15:53:22]: CMD-(CLI):[show logging]by default from console
         - Repeated 1 time.
[May 17 15:53:36]: CMD-(CLI):[write memory]by default from console
         - Repeated 5 times.
[May 17 15:53:44]: CMD-(CLI):[show logging]by default from console
[May 17 15:53:53]: CMD-(CLI):[show command-history]by default from console
[May 17 15:54:54]: CMD-(CLI):[end]by default from console
[May 17 15:55:00]: CMD-(CLI):[show logging]by default from console
[May 17 15:55:12]: CMD-(CLI):[show clock]by default from console
[May 17 15:55:22]: CMD-(CLI):[show running-config]by default from console
[May 17 15:55:27]: CMD-(CLI):[show command-history]by default from console
```

Upgrading Dell EMC Networking OS

To upgrade Dell EMC Networking Operating System (OS), refer to the Release Notes for the version you want to load on the system.

You can download the release notes of your platform at https://www.force10networks.com. Use your login ID to log in to the website.

Using HTTP for File Transfers

Stating with Release 9.3(0.1), you can use HTTP to copy files or configuration details to a remote server. To transfer files to an external server, use the **copy source-file-url http://host[:port]/file-path** command.

Enter the following *source-file-url* keywords and information:

- To copy a file from the internal FLASH, enter flash:// followed by the filename.
- To copy the running configuration, enter the keyword running-config.
- To copy the startup configuration, enter the keyword startup-config.
- To copy a file on the USB device, enter usbflash:// followed by the filename.

In the Dell EMC Networking OS release 9.8(0.0), HTTP services support the VRF-aware functionality. If you want the HTTP server to use a VRF table that is attached to an interface, configure that HTTP server to use a specific routing table. You can use the `ip http vrf` command to inform the HTTP server to use a specific routing table. After you configure this setting, the VRF table is used to look up the destination address.

NOTE: To enable HTTP to be VRF-aware, as a prerequisite you must first define the VRF.

You can specify either the management VRF or a nondefault VRF to configure the VRF awareness setting.

When you specify the management VRF, the copy operation that is used to transfer files to and from an HTTP server utilizes the VRF table corresponding to the Management VRF to look up the destination. When you specify a nondefault VRF, the VRF table corresponding to that nondefault VRF is used to look up the HTTP server.

However, these changes are backward-compatible and do not affect existing behavior; meaning, you can still use the `ip http source- interface` command to communicate with a particular interface even if no VRF is configured on that interface

NOTE: If the HTTP service is not VRF-aware, then it uses the global routing table to perform the look-up.

To enable an HTTP client to look up the VRF table corresponding to either management VRF or any nondefault VRF, use the `ip http vrf` command in CONFIGURATION mode.

- Configure an HTTP client with a VRF that is used to connect to the HTTP server.

  CONFIGURATION MODE

  `DellEMC(conf)#ip http vrf {management | <vrf-name>}`

Verify Software Images Before Installation

To validate the software image on the flash drive, you can use the MD5 message-digest algorithm or SHA256 Secure Hash Algorithm, after the image is transferred to the system but before the image is installed. The validation calculates a hash value of the downloaded image file on system's flash drive, and**,** optionally**,** compares it to a Dell EMC Networking published hash for that file.

The MD5 or SHA256 hash provides a method of validating that you have downloaded the original software. Calculating the hash on the local image file and comparing the result to the hash published for that file on iSupport provides a high level of confidence that the local copy is exactly the same as the published software image. This validation procedure, and the `verify {md5 | sha256}` command to support it, prevents the installation of corrupted or modified images.

The `verify {md5 | sha256}` command calculates and displays the hash of any file on the specified local flash drive. You can compare the displayed hash against the appropriate hash published on iSupport. Optionally, you can include the published hash in the `verify {md5 | sha256}` command, which displays whether it matches the calculated hash of the indicated file.

To validate a software image:

1. Download Dell EMC Networking OS software image file from the iSupport page to the local (FTP or TFTP) server. The published hash for that file displays next to the software image file on the iSupport page.
2. Go on to the Dell EMC Networking system and copy the software image to the flash drive, using the `copy` command.
3. Run the `verify {md5 | sha256} [ flash://]img-file [hash-value]` command. For example, `verify sha256 flash://FTOS-SE-9.5.0.0.bin`
4. Compare the generated hash value to the expected hash value published on the iSupport page.

To validate the software image on the flash drive after the image is transferred to the system, but before you install the image, use the `verify {md5 | sha256} [ flash://]img-file [hash-value]` command in EXEC mode.

- `md5`: MD5 message-digest algorithm
- `sha256`: SHA256 Secure Hash Algorithm
- `flash:` (Optional) Specifies the flash drive. The default uses the flash drive. You can enter the image file name.
- `hash-value`: (Optional). Specify the relevant hash published on iSupport.
- `img-file`: Enter the name of the Dell EMC Networking software image file to validate

Examples: Without Entering the Hash Value for Verification

MD5

```
DellEMC# verify md5 flash:file-name
```

SHA256

```
DellEMC# verify sha256 flash://file-name
```

Examples: Entering the Hash Value for Verification

MD5

```
DellEMC# verify md5 flash://file-name 275ceb73a4f3118e1d6bcf7d75753459
```

SHA256

```
DellEMC# verify sha256 flash://file-name
e6328c06faf814e6899ceead219afbf9360e986d692988023b749e6b2093e933
```

4

Management

This chapter describes the different protocols or services used to manage the Dell EMC Networking system.

Topics:

- Configuring Privilege Levels
- Configuring Logging
- Log Messages in the Internal Buffer
- Disabling System Logging
- Sending System Messages to a Syslog Server
- Track Login Activity
- Limit Concurrent Login Sessions
- Enabling Secured CLI Mode
- Changing System Logging Settings
- Display the Logging Buffer and the Logging Configuration
- Configuring a UNIX Logging Facility Level
- Synchronizing Log Messages
- Enabling Timestamp on Syslog Messages
- File Transfer Services
- Terminal Lines
- Setting Timeout for EXEC Privilege Mode
- Using Telnet to get to Another Network Device
- Lock CONFIGURATION Mode
- Recovering from a Forgotten Password
- Recovering from a Failed Start
- Restoring the Factory Default Settings
- Dell EMC Networking OS Security Hardening

Configuring Privilege Levels

Privilege levels restrict access to commands based on user or terminal line.

There are 16 privilege levels, of which three are pre-defined. The default privilege level is **1**.

Level	Description
Level 0	Access to the system begins at EXEC mode, and EXEC mode commands are limited to `enable`, `disable`, and `exit`.
Level 1	Access to the system begins at EXEC mode, and all commands are available.
Level 15	Access to the system begins at EXEC Privilege mode, and all commands are available.

For information about how access and authorization is controlled based on a user's role, see Role-Based Access Control.

Creating a Custom Privilege Level

Custom privilege levels start with the default EXEC mode command set. You can then customize privilege levels 2-14 by:

- restricting access to an EXEC mode command
- moving commands from EXEC Privilege to EXEC mode
- restricting access

A user can access all commands at his privilege level and below.

Removing a Command from EXEC Mode

To remove a command from the list of available commands in EXEC mode for a specific privilege level, use the `privilege exec` command from CONFIGURATION mode.

In the command, specify a level *greater* than the level given to a user or terminal line, then the first keyword of each command you wish to restrict.

Moving a Command from EXEC Privilege Mode to EXEC Mode

To move a command from EXEC Privilege to EXEC mode for a privilege level, use the `privilege exec` command from CONFIGURATION mode.

In the command, specify the privilege level of the user or terminal line and specify *all* keywords in the command to which you want to allow access.

Allowing Access to CONFIGURATION Mode Commands

To allow access to CONFIGURATION mode, use the `privilege exec level` *`level`* `configure` command from CONFIGURATION mode.

A user that enters CONFIGURATION mode remains at his privilege level and has access to only two commands, `end` and `exit`. You must individually specify each CONFIGURATION mode command you want to allow access to using the `privilege configure level` *`level`* command. In the command, specify the privilege level of the user or terminal line and specify *all* the keywords in the command to which you want to allow access.

Allowing Access to Different Modes

This section describes how to allow access to the INTERFACE, LINE, ROUTE-MAP, and ROUTER modes.

Similar to allowing access to CONFIGURATION mode, to allow access to INTERFACE, LINE, ROUTE-MAP, and ROUTER modes, you must first allow access to the command that enters you into the mode. For example, to allow a user to enter INTERFACE mode, use the `privilege configure level` *`level`* `interface tengigabitethernet` command.

Next, individually identify the INTERFACE, LINE, ROUTE-MAP or ROUTER commands to which you want to allow access using the `privilege {interface | line | route-map | router} level` *`level`* command. In the command, specify the privilege level of the user or terminal line and specify *all* the keywords in the command to which you want to allow access.

To remove, move or allow access, use the following commands.

The configuration in the following example creates privilege level 3. This level:

- removes the `resequence` command from EXEC mode by requiring a minimum of privilege level 4
- moves the `capture bgp-pdu max-buffer-size` command from EXEC Privilege to EXEC mode by requiring a minimum privilege level 3, which is the configured level for VTY 0
- allows access to CONFIGURATION mode with the `banner` command
- allows access to INTERFACE tengigabitethernet and LINE modes are allowed with no commands
- Remove a command from the list of available commands in EXEC mode.

  CONFIGURATION mode

  `privilege exec level` *`level`* `{`*`command`* `|...|` *`command`*`}`
- Move a command from EXEC Privilege to EXEC mode.

  CONFIGURATION mode

  `privilege exec level` *`level`* `{`*`command`* `|...|` *`command`*`}`
- Allow access to CONFIGURATION mode.

  CONFIGURATION mode

  `privilege exec level` *`level`* `configure`
- Allow access to INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode. Specify *all* the keywords in the command.

  CONFIGURATION mode

`privilege configure level level {interface | line | route-map | router} {command-keyword ||...|| command-keyword}`

- Allow access to a CONFIGURATION, INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode command.

  CONFIGURATION mode

  `privilege {configure |interface | line | route-map | router} level level {command ||...|| command}`

```
DellEMC(conf)#do show run priv
!
privilege exec level 3 capture
privilege exec level 3 configure
privilege exec level 4 resequence
privilege exec level 3 capture bgp-pdu
privilege exec level 3 capture bgp-pdu max-buffer-size
privilege configure level 3 line
privilege configure level 3 interface
DellEMC(conf)#do telnet 10.11.80.201
[telnet output omitted]
DellEMC#show priv
Current privilege level is 3.
DellEMC#?
capture         Capture packet
configure       Configuring from terminal
disable         Turn off privileged commands
enable          Turn on privileged commands
exit            Exit from the EXEC
ip              Global IP subcommands
monitor         Monitoring feature
mtrace          Trace reverse multicast path from destination to source
ping            Send echo messages
quit            Exit from the EXEC
show            Show running system information
[output omitted]
DellEMC#config
[output omitted]
DellEMC(conf)#do show priv
Current privilege level is 3.
DellEMC(conf)#?
end                 Exit from configuration mode
exit                Exit from configuration mode
interface           Select an interface to configure
line                Configure a terminal line
linecard            Set line card type
DellEMC(conf)#interface ?
fastethernet        Fast Ethernet interface
gigabitethernet     Gigabit Ethernet interface
loopback            Loopback interface
managementethernet  Management Ethernet interface
null                Null interface
port-channel        Port-channel interface
range               Configure interface range
sonet               SONET interface
tengigabitethernet  TenGigabit Ethernet interface
vlan                VLAN interface
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#?
end                 Exit from configuration mode
exit                Exit from interface configuration mode
DellEMC(conf-if-te-1/1)#exit
DellEMC(conf)#line ?
aux                 Auxiliary line
console             Primary terminal line
vty                 Virtual terminal
DellEMC(conf)#line vty 0
DellEMC(config-line-vty)#?
exit                Exit from line configuration mode
DellEMC(config-line-vty)#
DellEMC(conf)#interface group ?
gigabitethernet         GigabitEthernet interface IEEE 802.3z
tengigabitethernet     TenGigabit Ethernet interface
vlan                         VLAN keyword
```

```
DellEMC(conf)# interface group vlan 1 - 2 , tengigabitethernet 1/1
DellEMC(conf-if-group-vl-1-2,te-1/1)# no shutdown
DellEMC(conf-if-group-vl-1-2,te-1/1)# end
```

Applying a Privilege Level to a Username

To set the user privilege level, use the following command.

- Configure a privilege level for a user.
  CONFIGURATION mode
  `username` *`username`* `privilege` *`level`*

Applying a Privilege Level to a Terminal Line

To set a privilege level for a terminal line, use the following command.

- Configure a privilege level for a user.
  CONFIGURATION mode
  `username username privilege` *`level`*

NOTE: When you assign a privilege level between 2 and 15, access to the system begins at EXEC mode, but the prompt is `hostname#`, rather than `hostname>`.

Configuring Logging

The Dell EMC Networking OS tracks changes in the system using event and error messages.

By default, Dell EMC Networking OS logs these messages on:

- the internal buffer
- console and terminal lines
- any configured syslog servers

To disable logging, use the following commands.

- Disable all logging except on the console.
  CONFIGURATION mode
  `no logging on`
- Disable logging to the logging buffer.
  CONFIGURATION mode
  `no logging buffer`
- Disable logging to terminal lines.
  CONFIGURATION mode
  `no logging monitor`
- Disable console logging.
  CONFIGURATION mode
  `no logging console`

Audit and Security Logs

This section describes how to configure, display, and clear audit and security logs.

The following is the configuration task list for audit and security logs:

- Enabling Audit and Security Logs
- Displaying Audit and Security Logs

- Clearing Audit Logs

Enabling Audit and Security Logs

You enable audit and security logs to monitor configuration changes or determine if these changes affect the operation of the system in the network. You log audit and security events to a system log server, using the **logging extended** command in CONFIGURATION mode.

This command is available with or without RBAC enabled. For information about RBAC, see Role-Based Access Control.

Audit Logs

The audit log contains configuration events and information. The types of information in this log consist of the following:

- User logins to the switch.
- System events for network issues or system issues.
- Users making configuration changes. The switch logs who made the configuration changes and the date and time of the change. However, each specific change on the configuration is not logged. Only that the configuration was modified is logged with the user ID, date, and time of the change.
- Uncontrolled shutdown.

Security Logs

The security log contains security events and information. RBAC restricts access to audit and security logs based on the CLI sessions' user roles. The types of information in this log consist of the following:

- Establishment of secure traffic flows, such as SSH.
- Violations on secure flows or certificate issues.
- Adding and deleting of users.
- User access and configuration changes to the security and crypto parameters (not the key information but the crypto configuration)

Important Points to Remember

When you enabled RBAC and extended logging:

- Only the system administrator user role can execute this command.
- The system administrator and system security administrator user roles can view security events and system events.
- The system administrator user roles can view audit, security, and system events.
- Only the system administrator and security administrator user roles can view security logs.
- The network administrator and network operator user roles can view system events.

NOTE: If extended logging is disabled, you can only view system events, regardless of RBAC user role.

Example of Enabling Audit and Security Logs

```
DellEMC(conf)#logging extended
```

Displaying Audit and Security Logs

To display audit logs, use the `show logging auditlog` command in Exec mode. To view these logs, you must first enable the logging extended command. Only the RBAC system administrator user role can view the audit logs. Only the RBAC security administrator and system administrator user role can view the security logs. If extended logging is disabled, you can only view system events, regardless of RBAC user role. To view security logs, use the `show logging` command.

For information about the logging extended command, see Enabling Audit and Security Logs

Example of the `show logging auditlog` Command

```
DellEMC#show logging auditlog
May 12 12:20:25: DellEMC#: %CLI-6-logging extended by admin from vty0 (10.14.1.98)
May 12 12:20:42: DellEMC#: %CLI-6-configure terminal by admin from vty0 (10.14.1.98)
May 12 12:20:42: DellEMC#: %CLI-6-service timestamps log datetime by admin from vty0
(10.14.1.98)
```

For information about the logging extended command, see Enabling Audit and Security Logs

Example of the `show logging` Command for Security

```
DellEMC#show logging
Jun 10 04:23:40: %STKUNIT0-M:CP %SEC-5-LOGIN_SUCCESS: Login successful for user admin
on line vty0 ( 10.14.1.91 )
```

Clearing Audit Logs

To clear audit logs, use the `clear logging auditlog` command in Exec mode. When RBAC is enabled, only the system administrator user role can issue this command.

Example of the clear logging auditlog Command

```
DellEMC# clear logging auditlog
```

Configuring Logging Format

To display syslog messages in a RFC 3164 or RFC 5424 format, use the `logging version {0 | 1}` command in CONFIGURATION mode. By default, the system log version is set to `0`.

The following describes the two log messages formats:

- **0** – Displays syslog messages format as described in RFC 3164, The BSD syslog Protocol
- **1** – Displays syslog message format as described in RFC 5424, The SYSLOG Protocol

Example of Configuring the Logging Message Format

```
DellEMC(conf)#logging version ?
<0-1> Select syslog version (default = 0)
DellEMC(conf)#logging version 1
```

Display the Logging Buffer and the Logging Configuration

To display the current contents of the logging buffer and the logging settings for the system, use the `show logging` command in EXEC privilege mode. When RBAC is enabled, the security logs are filtered based on the user roles. Only the security administrator and system administrator can view the security logs.

Example of the `show logging` Command

```
DellEMC#show logging
syslog logging: enabled
   Console logging: level Debugging
   Monitor logging: level Debugging
   Buffer logging: level Debugging, 40 Messages Logged, Size (40960 bytes)
   Trap logging: level Informational
%IRC-6-IRC_COMMUP: Link to peer RPM is up
%RAM-6-RAM_TASK: RPM1 is transitioning to Primary RPM.
%RPM-2-MSG:CP1 %POLLMGR-2-MMC_STATE: External flash disk missing in 'slot0:'
%CHMGR-5-CARDDETECTED: Line card 0 present
%CHMGR-5-CARDDETECTED: Line card 2 present
%CHMGR-5-CARDDETECTED: Line card 4 present
%CHMGR-5-CARDDETECTED: Line card 5 present
```

```
%CHMGR-5-CARDDETECTED: Line card 8 present
%CHMGR-5-CARDDETECTED: Line card 10 present
%CHMGR-5-CARDDETECTED: Line card 12 present
%TSM-6-SFM_DISCOVERY: Found SFM 0
%TSM-6-SFM_DISCOVERY: Found SFM 1
%TSM-6-SFM_DISCOVERY: Found SFM 2
%TSM-6-SFM_DISCOVERY: Found SFM 3
%TSM-6-SFM_DISCOVERY: Found SFM 4
%TSM-6-SFM_DISCOVERY: Found SFM 5
%TSM-6-SFM_DISCOVERY: Found SFM 6
%TSM-6-SFM_DISCOVERY: Found SFM 7
%TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
%TSM-6-SFM_DISCOVERY: Found SFM 8
%TSM-6-SFM_DISCOVERY: Found 9 SFMs
%CHMGR-5-CHECKIN: Checkin from line card 5 (type EX1YB, 1 ports)
%TSM-6-PORT_CONFIG: Port link status for LC 5 => portpipe 0: OK portpipe 1: N/A
%CHMGR-5-LINECARDUP: Line card 5 is up
%CHMGR-5-CHECKIN: Checkin from line card 12 (type S12YC12, 12 ports)
%TSM-6-PORT_CONFIG: Port link status for LC 12 => portpipe 0: OK portpipe 1: N/A
%CHMGR-5-LINECARDUP: Line card 12 is up
%IFMGR-5-CSTATE_UP: changed interface Physical state to up: So 12/8
%IFMGR-5-CSTATE_DN: changed interface Physical state to down: So 12/8
```

To view any changes made, use the `show running-config logging` command in EXEC privilege mode.

Setting Up a Secure Connection to a Syslog Server

You can use reverse tunneling with the port forwarding to securely connect to a syslog server.

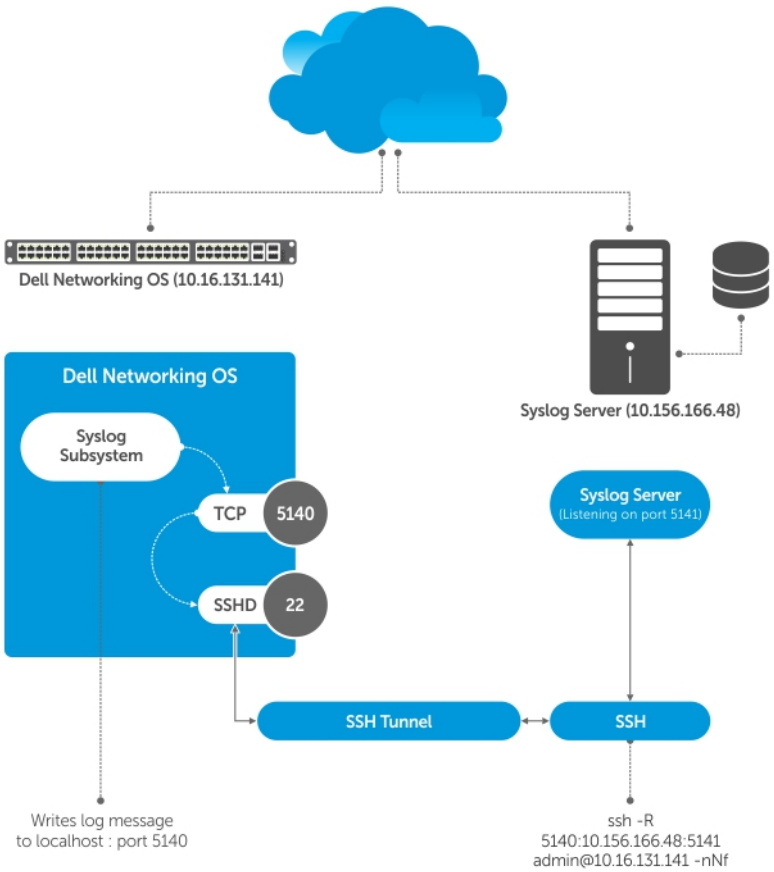


Figure 2. Setting Up a Secure Connection to a Syslog Server

Pre-requisites

To configure a secure connection from the switch to the syslog server:

1. On the switch, enable the SSH server

```
DellEMC(conf)#ip ssh server enable
```

2. On the syslog server, create a reverse SSH tunnel from the syslog server to the Dell OS switch, using following syntax:

```
ssh -R <remote port>:<syslog server>:<syslog server listen port> user@remote_host -nNf
```

In the following example the syslog server IP address is `10.156.166.48` and the listening port is `5141`. The switch IP address is `10.16.131.141` and the listening port is `5140`

```
ssh -R 5140:10.156.166.48:5141 admin@10.16.131.141 -nNf
```

3. Configure logging to a local host. *locahost* is "127.0.0.1" or "::1".

If you do not, the system displays an error when you attempt to enable role-based only AAA authorization.

```
DellEMC(conf)# logging localhost tcp port
DellEMC(conf)#logging 127.0.0.1 tcp 5140
```

Log Messages in the Internal Buffer

All error messages, except those beginning with `%BOOTUP (Message)`, are log in the internal buffer.

For example, `%BOOTUP:RPM0:CP %PORTPIPE-INIT-SUCCESS: Portpipe 0 enabled`

Configuration Task List for System Log Management

There are two configuration tasks for system log management:

- Disable System Logging
- Send System Messages to a Syslog Server

Disabling System Logging

By default, logging is enabled and log messages are sent to the logging buffer, all terminal lines, the console, and the syslog servers.

To disable system logging, use the following commands.

- Disable all logging except on the console.

  CONFIGURATION mode

  `no logging on`
- Disable logging to the logging buffer.

  CONFIGURATION mode

  `no logging buffer`
- Disable logging to terminal lines.

  CONFIGURATION mode

  `no logging monitor`
- Disable console logging.

  CONFIGURATION mode

  `no logging console`

Sending System Messages to a Syslog Server

To send system messages to a specified syslog server, use the following command. The following syslog standards are supported: RFC 5424 The SYSLOG Protocol, R.Gerhards and Adiscon GmbH, March 2009, obsoletes RFC 3164 and RFC 5426 Transmission of Syslog Messages over UDP.

- Specify the server to which you want to send system messages. You can configure up to eight syslog servers.

  CONFIGURATION mode

  `logging {ip-address | ipv6-address | hostname} {{udp {port}} | {tcp {port}}}`

  You can export system logs to an external server that is connected through a different VRF.

Configuring a UNIX System as a Syslog Server

To configure a UNIX System as a syslog server, use the following command.

- Configure a UNIX system as a syslog server by adding the following lines to */etc/syslog.conf* on the UNIX system and assigning write permissions to the file.
  - Add line on a 4.1 BSD UNIX system. `local7.debugging /var/log/ftos.log`
  - Add line on a 5.7 SunOS UNIX system. `local7.debugging /var/adm/ftos.log`

In the previous lines, `local7` is the logging facility level and debugging is the severity level.

Track Login Activity

Dell EMC Networking OS enables you to track the login activity of users and view the successful and unsuccessful login events.

When you log in using the console or VTY line, the system displays the last successful login details of the current user and the number of unsuccessful login attempts since your last successful login to the system, and whether the current user's permissions have changed since the last login. The system stores the number of unsuccessful login attempts that have occurred in the last 30 days by default. You can change the default value to any number of days from 1 to 30. By default, login activity tracking is disabled. You can enable it using the `login statistics enable` command from the configuration mode.

Restrictions for Tracking Login Activity

These restrictions apply for tracking login activity:

- Only the system and security administrators can configure login activity tracking and view the login activity details of other users.
- Login statistics is not applicable for login sessions that do not use user names for authentication. For example, the system does not report login activity for a telnet session that prompts only a password.

Configuring Login Activity Tracking

To enable and configure login activity tracking, follow these steps:

1. Enable login activity tracking.

   CONFIGURATION mode

   `login statistics enable`

   After enabling login statistics, the system stores the login activity details for the last 30 days.
2. (Optional) Configure the number of days for which the system stores the user login statistics. The range is from 1 to 30.

   CONFIGURATION mode

   `login statistics time-period days`

The following example enables login activity tracking. The system stores the login activity details for the last 30 days.

```
DellEMC(config)#login statistics enable
```

The following example enables login activity tracking and configures the system to store the login activity details for 12 days.

```
DellEMC(config)#login statistics enable
DellEMC(config)#login statistics time-period 12
```

Display Login Statistics

To view the login statistics, use the `show login statistics` command.

Example of the `show login statistics` Command

The `show login statistics` command displays the successful and failed login details of the current user in the last 30 days or the custom defined time period.

```
DellEMC#show login statistics

----------------------------------------------------------------------
User: admin
Last login time: 12:52:01 UTC Tue Mar 22 2016
Last login location: Line vty0 ( 10.16.127.143 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 30 day(s): 0
Successful login attempt(s) in last 30 day(s): 1
----------------------------------------------------------------------
```

Example of the `show login statistics all` command

The `show login statistics all` command displays the successful and failed login details of all users in the last 30 days or the custom defined time period.

```
DellEMC#show login statistics all

----------------------------------------------------------------------
User: admin
Last login time: 08:54:28 UTC Wed Mar 23 2016
Last login location: Line vty0 ( 10.16.127.145 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 30 day(s): 3
Successful login attempt(s) in last 30 day(s): 4
----------------------------------------------------------------------

----------------------------------------------------------------------
User: admin1
Last login time: 12:49:19 UTC Tue Mar 22 2016
Last login location: Line vty0 ( 10.16.127.145 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 30 day(s): 3
Successful login attempt(s) in last 30 day(s): 2
----------------------------------------------------------------------

----------------------------------------------------------------------
User: admin2
Last login time: 12:49:27 UTC Tue Mar 22 2016
Last login location: Line vty0 ( 10.16.127.145 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 30 day(s): 3
Successful login attempt(s) in last 30 day(s): 2
----------------------------------------------------------------------

----------------------------------------------------------------------
User: admin3
Last login time: 13:18:42 UTC Tue Mar 22 2016
Last login location: Line vty0 ( 10.16.127.145 )
Unsuccessful login attempt(s) since the last successful login: 0
```

```
Unsuccessful login attempt(s) in last 30 day(s): 3
Successful login attempt(s) in last 30 day(s): 2
```

Example of the `show login statistics user` *`user-id`* command

The `show login statistics user` *`user-id`* command displays the successful and failed login details of a specific user in the last 30 days or the custom defined time period.

```
DellEMC# show login statistics user admin

------------------------------------------------------------------
User: admin
Last login time: 12:52:01 UTC Tue Mar 22 2016
Last login location: Line vty0 ( 10.16.127.143 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 30 day(s): 0
Successful login attempt(s) in last 30 day(s): 1
------------------------------------------------------------------
```

The following is sample output of the `show login statistics unsuccessful-attempts` command.

```
DellEMC# show login statistics unsuccessful-attempts
There were 3 unsuccessful login attempt(s) for user admin in last 30 day(s).
```

The following is sample output of the `show login statistics unsuccessful-attempts time-period` *`days`* command.

```
DellEMC# show login statistics unsuccessful-attempts time-period 15
There were 0 unsuccessful login attempt(s) for user admin in last 15 day(s).
```

The following is sample output of the `show login statistics unsuccessful-attempts user` *`login-id`* command.

```
DellEMC# show login statistics unsuccessful-attempts user admin
There were 3 unsuccessful login attempt(s) for user admin in last 12 day(s).
```

The following is sample output of the `show login statistics successful-attempts` command.

```
DellEMC#show login statistics successful-attempts
There were 4 successful login attempt(s) for user admin in last 30 day(s).
```

Limit Concurrent Login Sessions

Dell EMC Networking OS enables you to limit the number of concurrent login sessions of users on VTY, auxiliary, and console lines. You can also clear any of your existing sessions when you reach the maximum permitted number of concurrent sessions.

By default, you can use all 10 VTY lines, one console line, and one auxiliary line. You can limit the number of available sessions using the `login concurrent-session limit` command and so restrict users to that specific number of sessions. You can optionally configure the system to provide an option to the users to clear any of their existing sessions.

Restrictions for Limiting the Number of Concurrent Sessions

These restrictions apply for limiting the number of concurrent sessions:

- Only the system and security administrators can limit the number of concurrent sessions and enable the clear-line option.
- Users can clear their existing sessions only if the system is configured with the `login concurrent-session clear-line enable` command.

Configuring Concurrent Session Limit

To configure concurrent session limit, follow this procedure:

- Limit the number of concurrent sessions for each user.

  CONFIGURATION mode

  `login concurrent-session limit` *`number-of-sessions`*

The following example limits the permitted number of concurrent login sessions to 4.

```
DellEMC(config)#login concurrent-session limit 4
```

Enabling the System to Clear Existing Sessions

To enable the system to clear existing login sessions, follow this procedure:

- Use the following command.

  CONFIGURATION mode

  `login concurrent-session clear-line enable`

The following example enables you to clear your existing login sessions.

```
DellEMC(config)#login concurrent-session clear-line enable
```

Example of Clearing Existing Sessions

When you try to log in, the following message appears with all your existing concurrent sessions, providing an option to close any one of the existing sessions:

```
$ telnet 10.11.178.14
Trying 10.11.178.14...
Connected to 10.11.178.14.
Escape character is '^]'.
Login: admin
Password:
Current sessions for user admin:
Line              Location
2  vty 0          10.14.1.97
3  vty 1          10.14.1.97
Clear existing session? [line number/Enter to cancel]:
```

When you try to create more than the permitted number of sessions, the following message appears, prompting you to close one of the existing sessions. If you close any of the existing sessions, you are allowed to login.

```
$ telnet 10.11.178.17
Trying 10.11.178.17...
Connected to 10.11.178.17.
Escape character is '^]'.
Login: admin
Password:

Maximum concurrent sessions for the user reached.
Current sessions for user admin:
Line              Location
2  vty 0          10.14.1.97
3  vty 1          10.14.1.97
4  vty 2          10.14.1.97
5  vty 3          10.14.1.97
Kill existing session? [line number/Enter to cancel]:
```

Enabling Secured CLI Mode

The secured CLI mode prevents the users from enhancing the permissions or promoting the privilege levels.

- Enter the following command to enable the secured CLI mode:

CONFIGURATION Mode

```
secure-cli enable
```

After entering the command, save the running-configuration. Once you save the running-configuration, the secured CLI mode is enabled.

If you do not want to enter the secured mode, do not save the running-configuration. Once saved, to disable the secured CLI mode, you need to manually edit the startup-configuration file and reboot the system.

Changing System Logging Settings

You can change the default settings of the system logging by changing the severity level and the storage location.

The default is to log all messages up to debug level, that is, all system messages. By changing the severity level in the logging commands, you control the number of system messages logged.

To specify the system logging settings, use the following commands.

- Specify the minimum severity level for logging to the logging buffer.

  CONFIGURATION mode

  ```
  logging buffered level
  ```
- Specify the minimum severity level for logging to the console.

  CONFIGURATION mode

  ```
  logging console level
  ```
- Specify the minimum severity level for logging to terminal lines.

  CONFIGURATION mode

  ```
  logging monitor level
  ```
- Specify the minimum severity level for logging to a syslog server.

  CONFIGURATION mode

  ```
  logging trap level
  ```
- Specify the minimum severity level for logging to the syslog history table.

  CONFIGURATION mode

  ```
  logging history level
  ```
- Specify the size of the logging buffer.

  CONFIGURATION mode

  ```
  logging buffered size
  ```

  NOTE: When you decrease the buffer size, Dell EMC Networking OS deletes all messages stored in the buffer. Increasing the buffer size does not affect messages in the buffer.
- Specify the number of messages that Dell EMC Networking OS saves to its logging history table.

  CONFIGURATION mode

  ```
  logging history size size
  ```

To view the logging buffer and configuration, use the `show logging` command in EXEC privilege mode, as shown in the example for Display the Logging Buffer and the Logging Configuration.

To view the logging configuration, use the `show running-config logging` command in privilege mode, as shown in the example for Configure a UNIX Logging Facility Level.

Display the Logging Buffer and the Logging Configuration

To display the current contents of the logging buffer and the logging settings for the system, use the `show logging` command in EXEC privilege mode. When RBAC is enabled, the security logs are filtered based on the user roles. Only the security administrator and system administrator can view the security logs.

Example of the `show logging` Command

```
DellEMC#show logging
syslog logging: enabled
   Console logging: level Debugging
   Monitor logging: level Debugging
   Buffer logging: level Debugging, 40 Messages Logged, Size (40960 bytes)
   Trap logging: level Informational
%IRC-6-IRC_COMMUP: Link to peer RPM is up
%RAM-6-RAM_TASK: RPM1 is transitioning to Primary RPM.
%RPM-2-MSG:CP1 %POLLMGR-2-MMC_STATE: External flash disk missing in 'slot0:'
%CHMGR-5-CARDDETECTED: Line card 0 present
%CHMGR-5-CARDDETECTED: Line card 2 present
%CHMGR-5-CARDDETECTED: Line card 4 present
%CHMGR-5-CARDDETECTED: Line card 5 present
%CHMGR-5-CARDDETECTED: Line card 8 present
%CHMGR-5-CARDDETECTED: Line card 10 present
%CHMGR-5-CARDDETECTED: Line card 12 present
%TSM-6-SFM_DISCOVERY: Found SFM 0
%TSM-6-SFM_DISCOVERY: Found SFM 1
%TSM-6-SFM_DISCOVERY: Found SFM 2
%TSM-6-SFM_DISCOVERY: Found SFM 3
%TSM-6-SFM_DISCOVERY: Found SFM 4
%TSM-6-SFM_DISCOVERY: Found SFM 5
%TSM-6-SFM_DISCOVERY: Found SFM 6
%TSM-6-SFM_DISCOVERY: Found SFM 7
%TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
%TSM-6-SFM_DISCOVERY: Found SFM 8
%TSM-6-SFM_DISCOVERY: Found 9 SFMs
%CHMGR-5-CHECKIN: Checkin from line card 5 (type EX1YB, 1 ports)
%TSM-6-PORT_CONFIG: Port link status for LC 5 => portpipe 0: OK portpipe 1: N/A
%CHMGR-5-LINECARDUP: Line card 5 is up
%CHMGR-5-CHECKIN: Checkin from line card 12 (type S12YC12, 12 ports)
%TSM-6-PORT_CONFIG: Port link status for LC 12 => portpipe 0: OK portpipe 1: N/A
%CHMGR-5-LINECARDUP: Line card 12 is up
%IFMGR-5-CSTATE_UP: changed interface Physical state to up: So 12/8
%IFMGR-5-CSTATE_DN: changed interface Physical state to down: So 12/8
```

To view any changes made, use the `show running-config logging` command in EXEC privilege mode.

Configuring a UNIX Logging Facility Level

You can save system log messages with a UNIX system logging facility.

To configure a UNIX logging facility level, use the following command.

- Specify one of the following parameters.
  CONFIGURATION mode
  `logging facility [facility-type]`
  - `auth` (for authorization messages)
  - `cron` (for system scheduler messages)
  - `daemon` (for system daemons)
  - `kern` (for kernel messages)
  - `local0` (for local use)
  - `local1` (for local use)
  - `local2` (for local use)

- `local3` (for local use)
- `local4` (for local use)
- `local5` (for local use)
- `local6` (for local use)
- `local7` (for local use)
- `lpr` (for line printer system messages)
- `mail` (for mail system messages)
- `news` (for USENET news messages)
- `sys9` (system use)
- `sys10` (system use)
- `sys11` (system use)
- `sys12` (system use)
- `sys13` (system use)
- `sys14` (system use)
- `syslog` (for syslog messages)
- `user` (for user programs)
- `uucp` (UNIX to UNIX copy protocol)

To view nondefault settings, use the `show running-config logging` command in EXEC mode.

```
DellEMC#show running-config logging
!
logging buffered 524288 debugging
service timestamps log datetime msec
service timestamps debug datetime msec
!
logging trap debugging
logging facility user
logging source-interface Loopback 0
logging 10.10.10.4
DellEMC#
```

Synchronizing Log Messages

You can configure Dell EMC Networking OS to filter and consolidate the system messages for a specific line by synchronizing the message output.

Only the messages with a severity at or below the set level appear. This feature works on the terminal and console connections available on the system.

1. Enter LINE mode.

   CONFIGURATION mode

   `line {console 0 | vty` *`number`* `[`*`end-number`*`] | aux 0}`

   Configure the following parameters for the virtual terminal lines:
   - *`number`*: the range is from zero (0) to 8.
   - *`end-number`*: the range is from 1 to 8.

   You can configure multiple virtual terminals at one time by entering a *number* and an *end-number*.

2. Configure a level and set the maximum number of messages to print.

   LINE mode

   `logging synchronous [level` *`severity-level`* `| all] [`*`limit`*`]`

   Configure the following optional parameters:
   - `level` *`severity-level`*: the range is from 0 to 7. The default is **2**. Use the `all` keyword to include all messages.
   - `limit`: the range is from 20 to 300. The default is **20**.

To view the logging synchronous configuration, use the `show config` command in LINE mode.

Enabling Timestamp on Syslog Messages

By default, syslog messages include a time/date stamp, taken from the datetime, stating when the error or message was created.

To enable timestamp, use the following command.

- Add timestamp to syslog messages.

  CONFIGURATION mode

  `service timestamps [log | debug] [datetime [localtime] [msec] [show-timezone] [utc] | uptime]`

  Specify the following optional parameters:

  - `datetime`: To view the timestamp in system local time that includes the local time zone.
  - `localtime`: You can add the keyword `localtime` to view timestamp in system local time that includes the local time zone.
  - `show-timezone`: Enter the keyword to include the time zone information in the timestamp.
  - `msec`: Enter the keyword `msec` to include milliseconds in the timestamp.
  - `uptime`: To view time since last boot.
  - `utc`: Enter the keyword `utc` to view timestamp in UTC time that excludes the local time zone.

  If you do not specify a parameter, Dell EMC Networking OS configures `datetime` as `localtime` by default.

To view the configuration, use the `show running-config logging` command in EXEC privilege mode.

To disable time stamping on syslog messages, use the `no service timestamps [log | debug]` command.

Example 1: Default configuration `service timestamps log datetime` or `service timestamps log datetime localtime`

```
DellEMC(conf)#service timestamps log datetime
```

```
DellEMC#show clock
15:42:42.804 IST Fri May 17 2019
```

```
DellEMC# show command-history
[May 17 15:38:55]: CMD-(CLI):[service timestamps log datetime]by default from console
[May 17 15:41:40]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[May 17 15:41:45]: CMD-(CLI):[interface tengigabitethernet 1/1]by default from console
[May 17 15:41:47]: CMD-(CLI):[shutdown]by default from console
[May 17 15:41:50]: CMD-(CLI):[no shutdown]by default from console
[May 17 15:42:42]: CMD-(CLI):[show clock]by default from console
[May 17 15:42:52]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[May 17 15:43:08]: CMD-(CLI):[end]by default from console
[May 17 15:43:16]: CMD-(CLI):[show logging]by default from console
[May 17 15:43:22]: CMD-(CLI):[show command-history]by default from console
DellEMC#
```

```
DellEMC#show logging
Syslog logging: enabled
    Console logging: disabled
    Monitor logging: level debugging
    Buffer logging: level debugging, 7 Messages Logged, Size (40960 bytes)
    Trap logging: level informational
    Last logging buffer cleared: May 17 15:38:38
May 17 15:43:08 %STKUNIT1-M:CP %SYS-5-CONFIG_I: Configured from  console
May 17 15:42:52 %STKUNIT1-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-
config in flash by default
May 17 15:41:53 %STKUNIT1-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 1/1
May 17 15:41:50 %STKUNIT1-M:CP %IFMGR-5-ASTATE_UP: Changed interface Admin state to up:
Te 1/1
May 17 15:41:47 %STKUNIT1-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te
1/1
May 17 15:41:47 %STKUNIT1-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to
down: Te 1/1
```

```
May 17 15:41:40 %STKUNIT1-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-
config in flash by default
```

Example 2: `service timestamps log datetime utc`

```
DellEMC(conf)#service timestamps log datetime utc
```

```
DellEMC#show clock
15:47:05.661 IST Fri May 17 2019
```

```
DellEMC# show command-history
[May 17 10:16:53]: CMD-(CLI):[service timestamps log datetime utc]by default from console
[May 17 10:17:05]: CMD-(CLI):[show clock]by default from console
[May 17 10:17:20]: CMD-(CLI):[show running-config]by default from console
[May 17 10:17:30]: CMD-(CLI):[interface tengigabitethernet 1/2]by default from console
[May 17 10:17:32]: CMD-(CLI):[shutdown]by default from console
[May 17 10:17:34]: CMD-(CLI):[no shutdown]by default from console
[May 17 10:17:40]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[May 17 10:17:46]: CMD-(CLI):[end]by default from console
[May 17 10:17:50]: CMD-(CLI):[show logging]by default from console
[May 17 10:17:56]: CMD-(CLI):[show command-history]by default from console
```

```
DellEMC# show logging
Syslog logging: enabled
    Console logging: disabled
    Monitor logging: level debugging
    Buffer logging: level debugging, 6 Messages Logged, Size (40960 bytes)
    Trap logging: level informational
    Last logging buffer cleared: May 17 15:46:36
May 17 10:17:46 %STKUNIT1-M:CP %SYS-5-CONFIG_I: Configured from  console
May 17 10:17:40 %STKUNIT1-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-
config in flash by default
May 17 10:17:37 %STKUNIT1-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 1/2
May 17 10:17:34 %STKUNIT1-M:CP %IFMGR-5-ASTATE_UP: Changed interface Admin state to up:
Te 1/2
May 17 10:17:32 %STKUNIT1-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te
1/2
May 17 10:17:32 %STKUNIT1-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to
down: Te 1/2
```

Example 3: `service timestamps log uptime`

```
DellEMC(conf)#service timestamps log uptime
```

```
DellEMC#show clock
15:51:47.534 IST Fri May 17 2019
```

```
DellEMC# show version |grep uptime
Dell EMC Networking OS uptime is 1 day(s), 0 hour(s), 25 minute(s)
```

```
DellEMC# show command-history
[1d0h24m]: CMD-(CLI):[service timestamps log uptime]by default from console
[1d0h24m]: CMD-(CLI):[interface tengigabitethernet 1/1]by default from console
[1d0h24m]: CMD-(CLI):[shutdown]by default from console
[1d0h24m]: CMD-(CLI):[no shutdown]by default from console
[1d0h25m]: CMD-(CLI):[end]by default from console
[1d0h25m]: CMD-(CLI):[write memory]by default from console
         - Repeated 1 time.
[1d0h25m]: CMD-(CLI):[show clock]by default from console
[1d0h25m]: CMD-(CLI):[show version]by default from console
[1d0h25m]: CMD-(CLI):[show logging]by default from console
[1d0h25m]: CMD-(CLI):[show command-history]by default from console
```

```
DellEMC# show logging
Syslog logging: enabled
    Console logging: disabled
    Monitor logging: level debugging
    Buffer logging: level debugging, 6 Messages Logged, Size (40960 bytes)
```

```
    Trap logging: level informational
    Last logging buffer cleared: May 17 15:50:31
1d0h25m %STKUNIT1-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-config in
flash by default
1d0h25m %STKUNIT1-M:CP %SYS-5-CONFIG_I: Configured from  console
1d0h24m %STKUNIT1-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 1/1
1d0h24m %STKUNIT1-M:CP %IFMGR-5-ASTATE_UP: Changed interface Admin state to up: Te 1/1
1d0h24m %STKUNIT1-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 1/1
1d0h24m %STKUNIT1-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 1/1
```

Example 4: `no service timestamps log`

```
DellEMC(conf)#no service timestamps log
```

```
DellEMC#show clock
15:55:12.246 IST Fri May 17 2019
```

```
DellEMC# show command-history
[May 17 15:53:10]: CMD-(CLI):[no service timestamps log]by default from console
[May 17 15:53:16]: CMD-(CLI):[write memory]by default from console
         - Repeated 3 times.
[May 17 15:53:22]: CMD-(CLI):[show logging]by default from console
         - Repeated 1 time.
[May 17 15:53:36]: CMD-(CLI):[write memory]by default from console
         - Repeated 5 times.
[May 17 15:53:44]: CMD-(CLI):[show logging]by default from console
[May 17 15:53:53]: CMD-(CLI):[show command-history]by default from console
[May 17 15:54:54]: CMD-(CLI):[end]by default from console
[May 17 15:55:00]: CMD-(CLI):[show logging]by default from console
[May 17 15:55:12]: CMD-(CLI):[show clock]by default from console
[May 17 15:55:22]: CMD-(CLI):[show running-config]by default from console
[May 17 15:55:27]: CMD-(CLI):[show command-history]by default from console
```

```
DellEMC# show logging
Syslog logging: enabled
    Console logging: disabled
    Monitor logging: level debugging
    Buffer logging: level debugging, 3 Messages Logged, Size (40960 bytes)
    Trap logging: level informational
    Last logging buffer cleared: May 17 15:52:54
%STKUNIT1-M:CP %SYS-5-CONFIG_I: Configured from  console
%STKUNIT1-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-config in flash by
default - repeated 3 times
%STKUNIT1-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-config in flash by
default
```

File Transfer Services

With Dell EMC Networking OS, you can configure the system to transfer files over the network using the file transfer protocol (FTP).

One FTP application is copying the system image files over an interface on to the system; however, FTP is not supported on virtual local area network (VLAN) interfaces.

The FTP and TFTP services are enhanced to support the VRF-aware functionality. If you want the FTP or TFTP server to use a VRF table that is attached to an interface, you must configure the FTP or TFTP server to use a specific routing table. You can use the `ip ftp vrf` *`vrf-name`* or `ip tftp vrf` *`vrf-name`* command to inform the FTP or TFTP server to use a specific routing table. After you configure this setting, the VRF table is used to look up the destination address. However, these changes are backward-compatible and do not affect existing behavior; meaning, you can still use the `source-interface` command to communicate with a particular interface even if no VRF is configured on that interface.

For more information about FTP, refer to RFC 959, *File Transfer Protocol*.

NOTE: To transmit large files, Dell EMC Networking recommends configuring the switch as an FTP server.

Configuration Task List for File Transfer Services

The configuration tasks for file transfer services are:

- Enable FTP Server (mandatory)
- Configure FTP Server Parameters (optional)
- Configure FTP Client Parameters (optional)

Enabling the FTP Server

To enable the system as an FTP server, use the following command.

To view FTP configuration, use the `show running-config ftp` command in EXEC privilege mode.

- Enable FTP on the system.

  CONFIGURATION mode

  `ftp-server enable`

```
DellEMC#show running ftp
!
ftp-server enable
ftp-server username nairobi password 0 zanzibar
DellEMC#
```

Configuring FTP Server Parameters

After you enable the FTP server on the system, you can configure different parameters.

To specify the system logging settings, use the following commands.

- Specify the directory for users using FTP to reach the system.

  CONFIGURATION mode

  `ftp-server topdir` *dir*

  The default is the internal flash directory.
- Specify a user name for all FTP users and configure either a plain text or encrypted password.

  CONFIGURATION mode

  `ftp-server username` *username* `password [`*encryption-type*`]` *password*

  Configure the following optional and required parameters:
  - *username*: enter a text string.
  - *encryption-type*: enter `0` for plain text or `7` for encrypted text.
  - *password*: enter a text string.

NOTE: You cannot use the `change directory (cd)` command until you have configured `ftp-server topdir`.

To view the FTP configuration, use the `show running-config ftp` command in EXEC privilege mode.

Configuring FTP Client Parameters

To configure FTP client parameters, use the following commands.

- Enter the following keywords and the interface information:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

CONFIGURATION mode

`ip ftp source-interface interface`

- Configure a password.

  CONFIGURATION mode

  `ip ftp password password`

- Enter a username to use on the FTP client.

  CONFIGURATION mode

  `ip ftp username name`

To view the FTP configuration, use the `show running-config ftp` command in EXEC privilege mode, as shown in the example for Enable FTP Server.

Terminal Lines

You can access the system remotely and restrict access to the system by creating user profiles.

Terminal lines on the system provide different means of accessing the system. The console line (console) connects you through the console port in the route processor modules (RPMs). The virtual terminal lines (VTYs) connect you through Telnet to the system. The auxiliary line (aux) connects secondary devices such as modems.

Denying and Permitting Access to a Terminal Line

Dell EMC Networking recommends applying only standard access control lists (ACLs) to deny and permit access to VTY lines.

- Layer 3 ACLs deny all traffic that is not explicitly permitted, but in the case of VTY lines, an ACL with no rules does not deny traffic.
- You cannot use the `show ip accounting access-list` command to display the contents of an ACL that is applied only to a VTY line.
- When you use the `access-class access-list-name` command without specifying the `ipv4` or `ipv6` attribute, both IPv4 as well as IPv6 rules that are defined in that ACL are applied to the terminal. This method is a generic way of configuring access restrictions.
- To be able to filter access exclusively using either IPv4 or IPv6 rules, use either the `ipv4` or `ipv6` attribute along with the `access-class access-list-name` command. Depending on the attribute that you specify (`ipv4` or `ipv6`), the ACL processes either IPv4 or IPv6 rules, but not both. Using this configuration, you can set up two different types of access classes with each class processing either IPv4 or IPv6 rules separately.

To apply an IP ACL to a line, Use the following command.

- Apply an ACL to a VTY line.

  LINE mode

  `access-class access-list-name [ipv4 | ipv6]`

  NOTE: If you already have configured generic IP ACL on a terminal line, then you cannot further apply IPv4 or IPv6 specific filtering on top of this configuration. Similarly, if you have configured either IPv4 or IPv6 specific filtering on a terminal line, you cannot apply generic IP ACL on top of this configuration. Before applying any of these configurations, you must first undo the existing configuration using the `no access-class access-list-name [ipv4 | ipv6]` command.

To view the configuration, use the `show config` command in LINE mode.

```
DellEMC(config-std-nacl)#show config
!
ip access-list standard myvtyacl
  seq 5 permit host 10.11.0.1
DellEMC(config-std-nacl)#line vty 0
DellEMC(config-line-vty)#show config
line vty 0
  access-class myvtyacl
```

```
DellEMC(conf-ipv6-acl)#do show run acl
!
```

```
ip access-list extended testdeny
seq 10 deny ip 30.1.1.0/24 any
seq 15 permit ip any any
!
ip access-list extended testpermit
seq 15 permit ip any any
!
ipv6 access-list testv6deny
seq 10 deny ipv6 3001::/64 any
seq 15 permit ipv6 any any
!
DellEMC(conf)#
DellEMC(conf)#line vty 0 0
DellEMC(config-line-vty)#access-class testv6deny ipv6
DellEMC(config-line-vty)#access-class testvpermit ipv4
DellEMC(config-line-vty)#show c
line vty 0
exec-timeout 0 0
access-class testpermit ipv4
access-class testv6deny ipv6
!
```

Configuring Login Authentication for Terminal Lines

You can use any combination of up to six authentication methods to authenticate a user on a terminal line.

A combination of authentication methods is called a method list. If the user fails the first authentication method, Dell EMC Networking OS prompts the next method until all methods are exhausted, at which point the connection is terminated. The available authentication methods are:

enable	Prompt for the enable password.
line	Prompt for the password you assigned to the terminal line. Configure a password for the terminal line to which you assign a method list that contains the line authentication method. Configure a password using the `password` command from LINE mode.
local	Prompt for the system username and password.
none	Do not authenticate the user.
radius	Prompt for a username and password and use a RADIUS server to authenticate.
tacacs+	Prompt for a username and password and use a TACACS+ server to authenticate.

1. Configure an authentication method list. You may use a mnemonic name or use the keyword default. The default authentication method for terminal lines is **local** and the default method list is **empty**.

   CONFIGURATION mode

   `aaa authentication login {`*`method-list-name`*` | default} [`*`method-1`*`] [`*`method-2`*`] [`*`method-3`*`] [`*`method-4`*`] [`*`method-5`*`] [`*`method-6`*`]`

2. Apply the method list from Step 1 to a terminal line.

   CONFIGURATION mode

   `login authentication {`*`method-list-name`*` | default}`

3. If you used the line authentication method in the method list you applied to the terminal line, configure a password for the terminal line.

   LINE mode

   `password`

In the following example, VTY lines 0-2 use a single authentication method, line.

```
DellEMC(conf)#aaa authentication login myvtymethodlist line
DellEMC(conf)#line vty 0 2
DellEMC(config-line-vty)#login authentication myvtymethodlist
DellEMC(config-line-vty)#password myvtypassword
DellEMC(config-line-vty)#show config
line vty 0
  password myvtypassword
login authentication myvtymethodlist
```

```
line vty 1
  password myvtypassword
login authentication myvtymethodlist
line vty 2
  password myvtypassword
login authentication myvtymethodlist
DellEMC(config-line-vty)#
```

Setting Timeout for EXEC Privilege Mode

EXEC timeout is a basic security feature that returns Dell EMC Networking OS to EXEC mode after a period of inactivity on the terminal lines.

To set timeout, use the following commands.

- Set the number of minutes and seconds. The default is **10 minutes** on the console and `30 minutes` on VTY. Disable EXEC time out by setting the timeout period to `0`.

  LINE mode

  `exec-timeout` *`minutes`* `[`*`seconds`*`]`
- Return to the default timeout values.

  LINE mode

  `no exec-timeout`

The following example shows how to set the timeout period and how to view the configuration using the `show config` command from LINE mode.

```
DellEMC(conf)#line con 0
DellEMC(config-line-console)#exec-timeout 0
DellEMC(config-line-console)#show config
line console 0
  exec-timeout 0 0
DellEMC(config-line-console)#
```

Using Telnet to get to Another Network Device

To telnet to another device, use the following commands.

NOTE: The device allows 120 Telnet sessions per minute, allowing the login and logout of 10 Telnet sessions, 12 times in a minute. If the system reaches this non-practical limit, the Telnet service is stopped for 10 minutes. You can use console and SSH service to access the system during downtime.

- Telnet to a device with an IPv4 or IPv6 address.

  EXEC Privilege

  `telnet [`*`ip-address`*`]`

  If you do not enter an IP address, Dell EMC Networking OS enters a Telnet dialog that prompts you for one.

  Enter an IPv4 address in dotted decimal format (A.B.C.D).

  Enter an IPv6 address in the format 0000:0000:0000:0000:0000:0000:0000:0000. Elision of zeros is supported.

```
DellEMC# telnet 10.11.80.203
Trying 10.11.80.203...
Connected to 10.11.80.203.
Exit character is '^]'.
Login:
Login: admin
Password:
DellEMC>exit
DellEMC#telnet 2200:2200:2200:2200:2200::2201
Trying 2200:2200:2200:2200:2200::2201...
Connected to 2200:2200:2200:2200:2200::2201.
Exit character is '^]'.
FreeBSD/i386 (freebsd2.force10networks.com) (ttyp1)
```

```
login: admin
DellEMC#
```

Lock CONFIGURATION Mode

Dell EMC Networking OS allows multiple users to make configurations at the same time. You can lock CONFIGURATION mode so that only one user can be in CONFIGURATION mode at any time (Message 2).

You can set two types of lockst: auto and manual.

- Set auto-lock using the `configuration mode exclusive auto` command from CONFIGURATION mode. When you set auto-lock, every time a user is in CONFIGURATION mode, all other users are denied access. This means that you can exit to EXEC Privilege mode, and re-enter CONFIGURATION mode without having to set the lock again.
- Set manual lock using the `configure terminal lock` command from CONFIGURATION mode. When you configure a manual lock, which is the default, you must enter this command each time you want to enter CONFIGURATION mode and deny access to others.

Viewing the Configuration Lock Status

If you attempt to enter CONFIGURATION mode when another user has locked it, you may view which user has control of CONFIGURATION mode using the `show configuration lock` command from EXEC Privilege mode.

You can then send any user a message using the `send` command from EXEC Privilege mode. Alternatively, you can clear any line using the `clear` command from EXEC Privilege mode. If you clear a console session, the user is returned to EXEC mode.

Example of Locking CONFIGURATION Mode for Single-User Access

```
DellEMC(conf)#configuration mode exclusive auto
BATMAN(conf)#exit
3d23h35m: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by console

DellEMC#config
! Locks configuration mode exclusively.
DellEMC(conf)#
```

If another user attempts to enter CONFIGURATION mode while a lock is in place, the following appears on their terminal (message 1): `% Error: User "" on line console0 is in exclusive configuration mode.`

If *any* user is already in CONFIGURATION mode when while a lock is in place, the following appears on their terminal (message 2): `% Error: Can't lock configuration mode exclusively since the following users are currently configuring the system: User "admin" on line vty1 ( 10.1.1.1 ).`

NOTE: The CONFIGURATION mode lock corresponds to a VTY session, not a user. Therefore, if you configure a lock and then exit CONFIGURATION mode, and another user enters CONFIGURATION mode, when you attempt to re-enter CONFIGURATION mode, you are denied access even though you are the one that configured the lock.

NOTE: If your session times out and you return to EXEC mode, the CONFIGURATION mode lock is unconfigured.

Recovering from a Forgotten Password

If you configure authentication for the console and you exit out of EXEC mode or your console session times out, you are prompted for a password to re-enter.

Use the following commands if you forget your password.

1. Log onto the system using the console.
2. Power-cycle the chassis by switching off all of the power modules and then switching them back on.
3. Hit any key to abort the boot process. You enter uBoot immediately, as indicated by the => prompt.

   (during bootup)

   hit any key

NOTE: You must enter the CLI commands. The system rejects them if they are copied and pasted.

4. Set the system parameters to ignore the startup configuration file when the system reloads.
   uBoot mode
   ```
   setenv stconfigignore true
   ```
5. To save the changes, use the `saveenv` command.
   uBoot mode
   ```
   saveenv
   ```
6. Reload the system.
   uBoot mode
   ```
   reset
   ```
7. Copy *startup-config.bak* to the running config.
   EXEC Privilege mode
   ```
   copy flash://startup-config.bak running-config
   ```
8. Remove all authentication statements you might have for the console.
   LINE mode
   ```
   no authentication login no password
   ```
9. Save the running-config.
   EXEC Privilege mode
   ```
   copy running-config startup-config
   ```
10. Set the system parameters to use the startup configuration file when the system reloads.
    uBoot mode
    ```
    setenv stconfigignore false
    ```
11. Save the running-config.
    EXEC Privilege mode
    ```
    copy running-config startup-config
    ```

Recovering from a Forgotten Enable Password

Use the following commands if you forget the enable password.

1. Log onto the system using the console.
2. Power-cycle the chassis by switching off all of the power modules and then switching them back on.
3. Hit any key to abort the boot process. You enter uBoot immediately, as indicated by the => prompt.
   (during bootup)
   hit any key

   NOTE: You must enter the CLI commands. The system rejects them if they are copied and pasted.
4. Set the system parameters to ignore the enable password when the system reloads.
   uBoot mode
   ```
   setenv enablepwdignore true
   ```
5. To save the changes, use the `saveenv` command.
   uBoot mode
   ```
   saveenv
   ```
6. Reload the system.
   uBoot mode
   ```
   reset
   ```
7. Configure a new enable password.
   CONFIGURATION mode

```
enable {password | secret | sha256-password}
```

8. Save the running-config to the startup-config.

   EXEC Privilege mode

   ```
   copy running-config startup-config
   ```

Recovering from a Failed Start

A system that does not start correctly might be attempting to boot from a corrupted Dell Networking OS image or from a mis-specified location.

In this case, you can restart the system and interrupt the boot process to point the system to another boot location. Use the `setenv` command, as described in the following steps. For details about the `setenv` command, its supporting commands, and other commands that can help recover from a failed start, see the *u-Boot* chapter in the *Dell Networking OS Command Line Reference Guide*.

1. Power-cycle the chassis (pull the power cord and reinsert it).
2. Hit any key to abort the boot process. You enter uBoot immediately, the => prompt indicates success.

   (during bootup)

   press any key
3. Assign the new location to the Dell Networking OS image it uses when the system reloads.

   uBoot mode

   ```
   setenv [primary_image f10boot location | secondary_image f10boot location | default_image f10boot location]
   ```
4. Assign an IP address to the Management Ethernet interface.

   uBoot mode

   ```
   setenv ipaddre address
   ```
5. Assign an IP address as the default gateway for the system.

   uBoot mode

   ```
   setenv gatewayip address
   ```
6. Reload the system.

   uBoot mode

   ```
   reset
   ```

Restoring the Factory Default Settings

Restoring the factory-default settings deletes the existing NVRAM settings, startup configuration, and all configured settings such as, stacking or fanout.

To restore the factory default settings, use the `restore factory-defaults stack-unit {stack–unit–number | all} {clear-all | nvram | bootvar}` command in EXEC Privilege mode.

CAUTION: There is no undo for this command.

Important Points to Remember

- When you restore all the units in a stack, these units are placed in standalone mode.
- When you restore a single unit in a stack, only that unit is placed in standalone mode. No other units in the stack are affected.
- When you restore the units in standalone mode, the units remain in standalone mode after the restoration.
- After the restore is complete, the units power cycle immediately.

The following example illustrates the **restore factory-defaults command** to restore the factory default settings.

```
DellEMC#restore factory-defaults stack-unit 0 nvram

    ********************************************************************
    *  Warning - Restoring factory defaults will delete the existing   *
    *  persistent settings (stacking, fanout, etc.)                    *
    *  After restoration the unit(s) will be powercycled immediately.  *
    *  Proceed with caution !                                          *
    ********************************************************************

Proceed with factory settings? Confirm [yes/no]:yes

-- Restore status --
Unit   Nvram     Config
------------------------
  0    Success

Power-cycling the unit(s).
....
```

Restoring Factory Default Environment Variables

The Boot line determines the location of the image that is used to boot up the chassis after restoring factory default settings. Ideally, these locations contain valid images, using which the chassis boots up.

When you restore factory-default settings, you can either use a flash boot procedure or a network boot procedure to boot the switch.

Important Points to Remember

- The Chassis remains in boot prompt if none of the partitions contain valid images.
- To enable TFTP boot after restoring factory default settings, you must stop the boot process in BLI.

In case the system fails to reload the image from the partition, perform the following steps:

1. Power-cycle the chassis (pull the power cord and reinsert it).
2. Hit any key to abort the boot process. You enter uBoot immediately, the => prompt indicates success.

   (during bootup)

   press any key
3. Assign the new location to the Dell EMC Networking OS image it uses when the system reloads.

   `uBoot mode`

   `=> setenv primary_boot` *`f10boot`*

   Boot variable (*f10boot*) can take the following values:
   - *flash0* — to boot from flash partition A.
   - *flash1* — to boot from flash partition B.
   - *tftp://server-ip/image-file-name* — to boot from the network.
4. Assign an IP address to the Management Ethernet interface.

   `uBoot mode`

   `=> setenv ipaddr` *`ip_address`*

   For example, *10.16.150.105*.

   `=> setenv netmask` *`mask`*

   For example, *255.255.0.0*.
5. Assign an IP address as the default gateway for the system.

   `uBoot mode`

```
=> setenv gatewayip gateway_ip_address
```

For example, *10.16.150.254*.

6. Save the modified environmental variables.

```
uBoot mode
=> saveenv
```

7. Reload the system.

```
uBoot mode
reset
```

Dell EMC Networking OS Security Hardening

The security of a network consists of multiple factors. Apart from access to the device, best practices, and implementing various security features, security also lies with the integrity of the device. If the software itself is compromised, all of the aforementioned methods become ineffective.

The Dell EMC Networking OS is enhanced verify whether the OS image and the startup configuration file are altered before loading. This section explains how to configure OS image and startup configuration verification.

Dell EMC Networking OS Image Verification

Dell EMC Networking OS comes with the OS image verification and the startup configuration verification features. When enabled, these features check the integrity of The OS image and the startup configuration that the system uses while the system reboots and loads only if they are intact.

Important Points to Remember

- The OS image verification feature is disabled by default on the Dell EMC Networking OS.
- The OS image verification feature is supported for images stored in the local system only.
- The OS image verification feature is not supported when the fastboot or the warmboot features are enabled on the system.
- If OS image verification fails after a reload, the system does not load the startup configuration. The System displays an appropriate error message until the `no verified boot` command is used on the system.
- After you enable The OS image verification feature, the system prompts you to enter The OS image hash when you upgrade the Dell EMC Networking OS to a later version. The system checks if your hash matches with The OS image hash only after reloading.
- After enabling The OS image verification feature, use the `verified boot hash` command to verify and store the hash value. If you don't store the hash value, you cannot reboot the device until you verify The OS image hash.

Enabling and Configuring OS Image Hash Verification

To enable and configure Dell EMC Networking OS image hash verification, follow these steps:

1. Enable the OS image hash verification feature.

   CONFIGURATION mode

   ```
   verified boot
   ```

2. Verify the hash checksum of the current OS image file on the local file system.

   EXEC Privilege

   ```
   verified boot hash system-image {A: | B:} hash-value
   ```

   You can get the hash value for your hashing algorithm from the Dell EMC iSupport page. You can use the MD5, SHA1, or SHA256 hash and the Dell EMC Networking OS automatically detects the type of hash.

   NOTE: The `verified boot hash` command is only applicable for OS images in the local file system.

3. Save the configuration.

EXEC Privilege

copy running-configuration startup-configuration

After enabling and configuring OS image hash verification, the device verifies the hash checksum of the OS boot image during every reload.

```
DellEMC# verified boot hash system-image A: 619A8C1B7A2BC9692A221E2151B9DA9E
```

Image Verification for Subsequent OS Upgrades

After enabling OS image hash verification, for subsequent Dell EMC Networking OS upgrades, you must enter the hash checksum of the new OS image file. To enter the hash checksum during upgrade, follow these steps:

- Use the following command to upgrade the Dell EMC Networking OS and enter the hash value when prompted.

  EXEC Privilege

  `upgrade system`

```
DellEMC# upgrade system tftp://10.16.127.35/FTOS-SE-9.11.0.1 A:
Hash Value: e42e2548783c2d5db239ea2fa9de4232
!!!!!!!!!!!!!!...
```

Startup Configuration Verification

Dell EMC Networking OS comes with startup configuration verification feature. When enabled, it checks the integrity of the startup configuration that the system uses while the system reboots and loads only if it is intact.

Important Points to Remember

- The startup configuration verification feature is disabled by default on the Dell EMC Networking OS.
- The feature is supported for startup configuration files stored in the local system only.
- The feature is not supported when the fastboot or the warmboot features are enabled on the system.
- If the startup configuration verification fails after a reload, the system does not load your startup configuration.
- After enabling the startup configuration verification feature, use the `verified boot hash` command to verify and store the hash value. If you don't store the hash value, you cannot reboot the device until you verify the image hash.
- If OS image verification fails, the system does not load your startup configuration and displays an error message until you remove the `verified boot` command from the configuration.

Dell EMC Networking OS Behavior after System Power-Cycle

If the system reboots due reasons such as power-cycle, the current startup configuration may be different than the one you verified the hash using the `verified boot hash` command. When the system comes up, the system may use the last-verified startup configuration.

Dell EMC Networking recommends backing up the startup configuration to a safe location after you use the `verified boot hash` command. When the startup configuration verification fails, you can restore it from the backup.

The system continues to display a message stating that startup configuration verification failed. You can disable the startup configuration feature either by disabling startup configuration verification or save the running configuration to the startup configuration and update the hash for the startup configuration.

Enabling and Configuring Startup Configuration Hash Verification

To enable and configure startup configuration hash verification, follow these steps:

1. Enable the startup configuration hash verification feature.

   CONFIGURATION mode

   `verified startup-config`

2. Generate the hash checksum for your startup configuration file.

EXEC Privilege

```
generate hash {md5 | sha1 | sha256} {flash://filename | startup-config}
```

3. Verify the hash checksum of the current startup configuration on the local file system.

   EXEC Privilege

   ```
   verified boot hash startup—config hash-value
   ```

   NOTE: The `verified boot hash` command is only applicable for the startup configuration file in the local file system.

After enabling and configuring startup configuration verification, the device verifies the hash checksum of the startup configuration during every reload.

```
DellEMC# verified boot hash startup—config 619A8C1B7A2BC9692A221E2151B9DA9E
```

Configuring the root User Password

For added security, you can change the root user password.

If you configure the `secure-cli` command on the system, the Dell EMC Networking OS resets any previously-configured root access password without displaying any warning message. With the `secure-cli` command enabled on the system, the CONFIGURATION mode does not display the `root access password` option.

To change the default root user password, follow these steps:

- Change the default root user password.

  CONFIGURATION mode

  ```
  root-access password [encryption-type] root-password
  ```

  Enter an encryption type for the root password.

  - `0` directs the system to store the password as clear text.
  - `7` directs the system to store the password with a dynamic salt.
  - `9` directs the system to encrypt the clear text password and store the encrypted password in an inaccessible location.

  When you configure the root access password, ensure that your password meets the following criteria:

  - A minimum of eight characters in length
  - A minimum of one lower case letter (a to z)
  - A minimum of one upper case letter (A to Z)
  - A minimum of one numeric character (0 to 9)
  - A minimum of one special character including a space (" !"#$%&'()*+,-./:;<=>?@[\]^_`{|}~")

```
DellEMC)# show running-config  | g root
root-access password 7 f4dc0cb9787722dd1084d17f417f164cc7f730d4f03d4f0215294cbd899614e3
```

Locking Access to GRUB Interface

You can configure the Dell EMC Networking OS to lock the GRUB interface using a password. If you configure a GRUB password, the system prompts for the password when you try to access the GRUB interface.

CAUTION: After configuring the boot access password, save it to a secure location. If you forget it, you will not be able to access the options in the startup menu. If you forget both the boot access password and the enable password, the system may become inaccessible.

To lock the GRUB interface using a password, use the following procedure:

- Configure a GRUB password.

  CONFIGURATION mode

  ```
  boot-access password [encryption-type] boot-password
  ```

  Enter an encryption type for the boot password.

- `0` directs the system to store the password as clear text.
- `7` directs the system to store the password with a dynamic salt.

When you configure the root access password, ensure that your password meets the following criteria:

- A minimum of eight characters in length
- A minimum of one lower case letter (a to z)
- A minimum of one upper case letter (A to Z)
- A minimum of one numeric character (0 to 9)
- A minimum of one special character including a space (" !"#$%&'()*+,-./:;<=>?@[\]^_`{|}~")

If you enable the boot access password, the system prompts for a password when you access the GRUB interface.

```
DellEMC(conf)#boot-access password 7 Hg$7^5HMoiY%
    **********************************************************************
    *  Warning - boot-access password will enable password protection in  *
    *  GRUB. Keep it safe. Forgetting this password and the CLI password  *
    *  may result in switch becoming inaccessible.                        *
    **********************************************************************

Do you want to configure boot-access password? Proceed [yes/no]:yes
DellEMC(conf)#
```

Enabling User Lockout for Failed Login Attempts

You can configure the system to lock out local users for a specific period for unsuccessful login attempts.

This feature enhances the security of the switch by locking out the local user account if there are more number of unsuccessful login attempts than what is configured using the *max-retry* parameter. To enable the user lock out feature, use the following commands:

Enable the user lockout feature.

CONFIGURATION

`password-attributes user-lockout-period` *minutes*

Enter the duration in minutes.

5

802.1X

802.1X is a port-based Network Access Control (PNAC) that provides an authentication mechanism to devices wishing to attach to a LAN or WLAN. A device connected to a port that is enabled with 802.1X is disallowed from sending or receiving packets on the network until its identity is verified (through a username and password, for example).

802.1X employs Extensible Authentication Protocol (EAP) to transfer a device's credentials to an authentication server (typically RADIUS) using a mandatory intermediary network access device, in this case, a Dell EMC Networking switch. The network access device mediates all communication between the end-user device and the authentication server so that the network remains secure. The network access device uses EAP-over-Ethernet (EAPOL) to communicate with the end-user device and EAP-over-RADIUS to communicate with the server.

NOTE: The Dell EMC Networking Operating System (OS) supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.

The following figures show how the EAP frames are encapsulated in Ethernet and RADIUS frames.

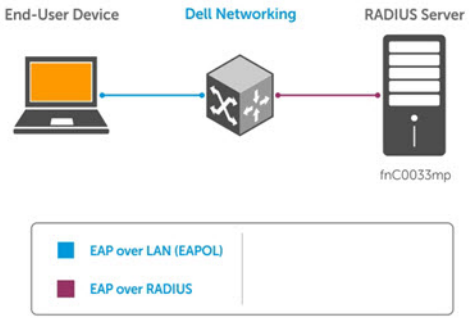


Figure 3. EAP Frames Encapsulated in Ethernet and RADUIS

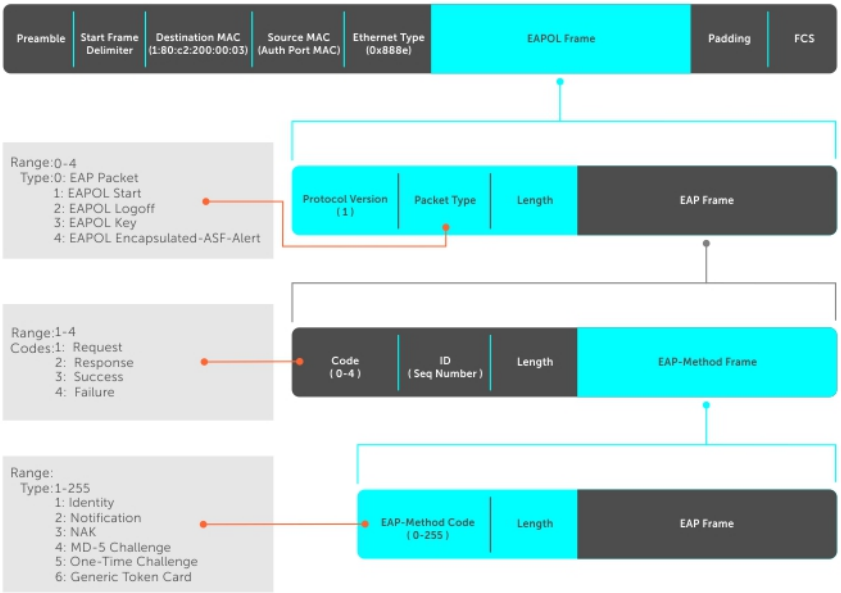


Figure 4. EAP Frames Encapsulated in Ethernet and RADUIS

The authentication process involves three devices:

- The device attempting to access the network is the **supplicant**. The supplicant is not allowed to communicate on the network until the authenticator authorizes the port. It can only communicate with the authenticator in response to 802.1X requests.
- The device with which the supplicant communicates is the **authenticator**. The authenticator is the gate keeper of the network. It translates and forwards requests and responses between the authentication server and the supplicant. The authenticator also changes the status of the port based on the results of the authentication process. The Dell EMC Networking switch is the authenticator.
- The authentication-server selects the authentication method, verifies the information the supplicant provides, and grants it network access privileges.

Ports can be in one of two states:

- Ports are in an **unauthorized** state by default. In this state, non-802.1X traffic cannot be forwarded in or out of the port.
- The authenticator changes the port state to authorized if the server can authenticate the supplicant. In this state, network traffic can be forwarded normally.

NOTE: The Dell EMC Networking switches place 802.1X-enabled ports in the unauthorized state by default.

Topics:

- Port-Authentication Process
- Configuring 802.1X
- Important Points to Remember
- Enabling 802.1X
- Configuring MAC addresses for a do1x Profile
- Configuring Request Identity Re-Transmissions
- Forcibly Authorizing or Unauthorizing a Port
- Re-Authenticating a Port
- Configuring Timeouts
- Configuring Dynamic VLAN Assignment with Port Authentication
- Guest and Authentication-Fail VLANs
- Configuring dot1x Profile
- Configuring the Static MAB and MAB Profile
- Configuring Critical VLAN

Port-Authentication Process

The authentication process begins when the authenticator senses that a link status has changed from down to up:

1. When the authenticator senses a link state change, it requests that the supplicant identify itself using an EAP Identity Request frame.
2. The supplicant responds with its identity in an EAP Response Identity frame.
3. The authenticator decapsulates the EAP response from the EAPOL frame, encapsulates it in a RADIUS Access-Request frame and forwards the frame to the authentication server.
4. The authentication server replies with an Access-Challenge frame. The Access-Challenge frame requests the supplicant to prove that it is who it claims to be, using a specified method (an EAP-Method). The challenge is translated and forwarded to the supplicant by the authenticator.
5. The supplicant can negotiate the authentication method, but if it is acceptable, the supplicant provides the Requested Challenge information in an EAP response, which is translated and forwarded to the authentication server as another Access-Request frame.
6. If the identity information provided by the supplicant is valid, the authentication server sends an Access-Accept frame in which network privileges are specified. The authenticator changes the port state to authorized and forwards an EAP Success frame. If the identity information is invalid, the server sends an Access-Reject frame. If the port state remains unauthorized, the authenticator forwards an EAP Failure frame.

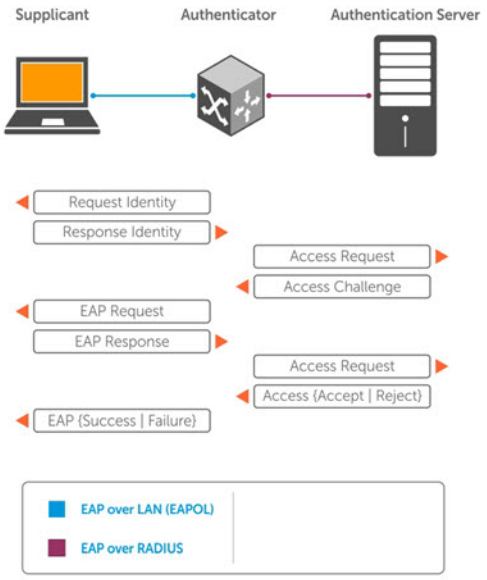


Figure 5. EAP Port-Authentication

EAP over RADIUS

802.1X uses RADIUS to shuttle EAP packets between the authenticator and the authentication server, as defined in RFC 3579.

EAP messages are encapsulated in RADIUS packets as a type of attribute in Type, Length, Value (TLV) format. The Type value for EAP messages is 79.

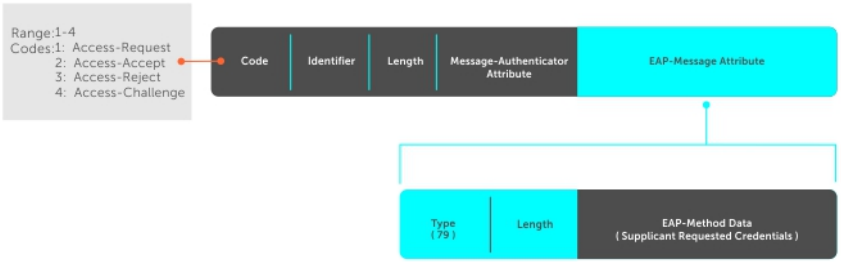


Figure 6. EAP Over RADIUS

RADIUS Attributes for 802.1X Support

Dell EMC Networking systems include the following RADIUS attributes in all 802.1X-triggered Access-Request messages:

Attribute 31	**Calling-station-id**: relays the supplicant MAC address to the authentication server.
Attribute 41	**NAS-Port-Type**: NAS-port physical port type. 15 indicates Ethernet.
Attribute 61	**NAS-Port**: the physical port number by which the authenticator is connected to the supplicant.
Attribute 81	**Tunnel-Private-Group-ID**: associate a tunneled session with a particular group of users.

Configuring 802.1X

Configuring 802.1X on a port is a one-step process.

For more information, refer to Enabling 802.1X.

Related Configuration Tasks

- Configuring Request Identity Re-Transmissions
- Forcibly Authorizing or Unauthorizing a Port
- Re-Authenticating a Port
- Configuring Timeouts
- Configuring a Guest VLAN
- Configuring an Authentication-Fail VLAN

Important Points to Remember

- Dell EMC Networking OS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.
- All platforms support only RADIUS as the authentication server.
- If the primary RADIUS server becomes unresponsive, the authenticator begins using a secondary RADIUS server, if configured.
- 802.1X is not supported on port-channels or port-channel members.
- The NAS-Port-Type attribute indicates the type of the physical port of the NAS which is authenticating the user. It is used in Access-Request packets. The value of this attribute is set as Ethernet (15) for both EAP and MAB supplicants.

Enabling 802.1X

Enable 802.1X globally.

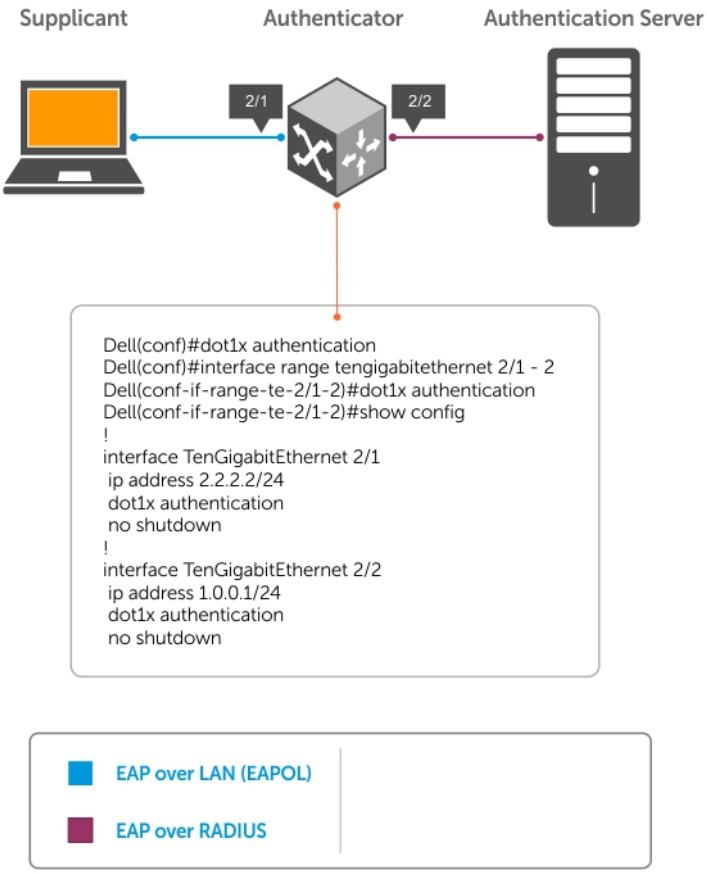


Figure 7. 802.1X Enabled

1. Enable 802.1X globally.
   CONFIGURATION mode
   `dot1x authentication`
2. Enter INTERFACE mode on an interface or a range of interfaces.
   INTERFACE mode
   `interface [range]`
3. Enable 802.1X on the supplicant interface only.
   INTERFACE mode
   `dot1x authentication`

Verify that 802.1X is enabled globally and at the interface level using the `show running-config | find dot1x` command from EXEC Privilege mode.

In the following example, the bold lines show that 802.1X is enabled.

```
DellEMC#show running-config | find dot1x
dot1x authentication
!
[output omitted]
!
interface TenGigabitEthernet 2/1
```

```
 no ip address
 dot1x authentication
 no shutdown
!
DellEMC#
```

To view 802.1X configuration information for an interface, use the `show dot1x interface` command.

In the following example, the bold lines show that 802.1X is enabled on all ports unauthorized by default.

```
DellEMC#show dot1x interface TenGigabitEthernet 2/1/

802.1x information on Te 2/1/:
------------------------------
Dot1x Status:       Enable
Port Control:               AUTO
Port Auth Status:     UNAUTHORIZED
Re-Authentication:          Disable
Untagged VLAN id:           None
Guest VLAN:                 Disable
Guest VLAN id:              NONE
Auth-Fail VLAN:             Disable
Auth-Fail VLAN id:          NONE
Auth-Fail Max-Attempts:     NONE
Mac-Auth-Bypass:            Disable
Mac-Auth-Bypass Only:       Disable
Tx Period:                  30 seconds
Quiet Period:               60 seconds
ReAuth Max:                 2
Supplicant Timeout:         30 seconds
Server Timeout:             30 seconds
Re-Auth Interval:           3600 seconds
Max-EAP-Req:                2
Host Mode:                  SINGLE_HOST
Auth PAE State:             Initialize
Backend State:              Initialize
```

Configuring MAC addresses for a do1x Profile

To configure a list of MAC addresses for a dot1x profile, use the `mac` command. You can configure 1 to 6 MAC addresses.

- Configure a list of MAC addresses for a dot1x profile.

  DOT1X PROFILE CONFIG (conf-dot1x-profile)

  `mac` *`mac-address`*

  `mac-address` — Enter the keyword `mac` and type up to the 48– bit MAC addresses using the nn:nn:nn:nn:nn:nn format. A maximum of 6 MAC addresses are allowed.

The following example configures 2 MAC addresses and then displays these addresses.

```
DellEMC(conf-dot1x-profile)#mac 00:50:56:AA:01:10 00:50:56:AA:01:11

DellEMC(conf-dot1x-profile)#show config
dot1x profile sample
 mac 00:50:56:aa:01:10
 mac 00:50:56:aa:01:11
DellEMC(conf-dot1x-profile)#
DellEMC(conf-dot1x-profile)#exit
DellEMC(conf)#
```

Configuring Request Identity Re-Transmissions

When the authenticator sends a Request Identity frame and the supplicant does not respond, the authenticator waits for 30 seconds and then re-transmits the frame.

The amount of time that the authenticator waits before re-transmitting and the maximum number of times that the authenticator re-transmits can be configured.

NOTE: There are several reasons why the supplicant might fail to respond; for example, the supplicant might have been booting when the request arrived or there might be a physical layer problem.

To configure re-transmissions, use the following commands.

- Configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame.

  INTERFACE mode

  `dot1x tx-period number`

  The range is from 1 to 65535 (1 year)

  The default is **30**.
- Configure the maximum number of times the authenticator re-transmits a Request Identity frame.

  INTERFACE mode

  `dot1x max-eap-req number`

  The range is from 1 to 10.

  The default is **2**.

The example in Configuring a Quiet Period after a Failed Authentication shows configuration information for a port for which the authenticator re-transmits an EAP Request Identity frame after 90 seconds and re-transmits for 10 times.

Configuring a Quiet Period after a Failed Authentication

If the supplicant fails the authentication process, the authenticator sends another Request Identity frame after 30 seconds by default. You can configure this period.

NOTE: The quiet period (`dot1x quiet-period`) is the transmit interval after a failed authentication; the Request Identity Re-transmit interval (`dot1x tx-period`) is for an unresponsive supplicant.

To configure a quiet period, use the following command.

- Configure the amount of time that the authenticator waits to re-transmit a Request Identity frame after a failed authentication.

  INTERFACE mode

  `dot1x quiet-period seconds`

  The range is from 1 to 65535.

  The default is **60 seconds**.

The following example shows configuration information for a port for which the authenticator re-transmits an EAP Request Identity frame:

- after 90 seconds and a maximum of 10 times for an unresponsive supplicant
- re-transmits an EAP Request Identity frame

The bold lines show the new re-transmit interval, new quiet period, and new maximum re-transmissions.

```
DellEMC(conf-if-range-Te-2/1)#dot1x tx-period 90
DellEMC(conf-if-range-Te-2/1)#dot1x max-eap-req 10
DellEMC(conf-if-range-Te-2/1)#dot1x quiet-period 120
DellEMC#show dot1x interface TenGigabitEthernet 2/1
802.1x information on Te 2/1:
-----------------------------
Dot1x Status:          Enable
Port Control:          AUTO
Port Auth Status:      UNAUTHORIZED
Re-Authentication:   Disable
Untagged VLAN id:      None
Tx Period:             90 seconds
```

```
Quiet Period:       120 seconds
ReAuth Max:         2
Supplicant Timeout: 30 seconds
Server Timeout:     30 seconds
Re-Auth Interval:   3600 seconds
Max-EAP-Req:        10
Auth Type:          SINGLE_HOST
Auth PAE State:     Initialize
Backend State:      Initialize
```

Forcibly Authorizing or Unauthorizing a Port

The 802.1X ports can be placed into any of the three states:

- **ForceAuthorized** — an authorized state. A device connected to this port in this state is never subjected to the authentication process, but is allowed to communicate on the network. Placing the port in this state is same as disabling 802.1X on the port.
- **ForceUnauthorized** — an unauthorized state. A device connected to a port in this state is never subjected to the authentication process and is not allowed to communicate on the network. Placing the port in this state is the same as shutting down the port. Any attempt by the supplicant to initiate authentication is ignored.
- **Auto** — an unauthorized state by default. A device connected to this port in this state is subjected to the authentication process. If the process is successful, the port is authorized and the connected device can communicate on the network. All ports are placed in the Auto state by default.

To set the port state, use the following command.

- Place a port in the ForceAuthorized, ForceUnauthorized, or Auto state.

  INTERFACE mode

  ```
  dot1x port-control {force-authorized | force-unauthorized | auto}
  ```

  The default state is **auto**.

The example shows configuration information for a port that has been force-authorized.

The bold line shows the new port-control state.

```
DellEMC(conf-if-Te-1/1)#dot1x port-control force-authorized
DellEMC(conf-if-Te-1/1)#show dot1x interface TenGigabitEthernet 1/1

802.1x information on Te 1/1:
-----------------------------
Dot1x Status:           Enable
Port Control:       FORCE_AUTHORIZED
Port Auth Status:       UNAUTHORIZED
Re-Authentication:      Disable
Untagged VLAN id:       None
Tx Period:              90 seconds
Quiet Period:           120 seconds
ReAuth Max:             2
Supplicant Timeout:     30 seconds
Server Timeout:         30 seconds
Re-Auth Interval:       3600 seconds
Max-EAP-Req:            10
Auth Type:              SINGLE_HOST
Auth PAE State:         Initialize
Backend State:          Initialize
Auth PAE State:         Initialize
Backend State:          Initialize
```

Re-Authenticating a Port

You can configure the authenticator for periodic re-authentication.

After the supplicant has been authenticated, and the port has been authorized, you can configure the authenticator to re-authenticate the supplicant periodically. If you enable re-authentication, the supplicant is required to re-authenticate every 3600 seconds by default, and you can configure this interval. You can configure the maximum number of re-authentications as well.

To configure re-authentication time settings, use the following commands:

- Configure the authenticator to periodically re-authenticate the supplicant.

  INTERFACE mode

  `dot1x reauthentication [interval] seconds`

  The range is from 1 to 31536000.

  The default is **3600**.
- Configure the maximum number of times the supplicant can be re-authenticated.

  INTERFACE mode

  `dot1x reauth-max number`

  The range is from 1 to 10.

  The default is **2**.

The bold lines show that re-authentication is enabled and the new maximum and re-authentication time period.

```
DellEMC(conf-if-Te-1/1)#dot1x reauthentication interval 7200
DellEMC(conf-if-Te-1/1)#dot1x reauth-max 10
DellEMC(conf-if-Te-1/1)#do show dot1x interface TenGigabitEthernet 1/1

802.1x information on Te 1/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:     UNAUTHORIZED
Re-Authentication:         Enable
Untagged VLAN id:          None
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:           10
Supplicant Timeout:        30 seconds
Server Timeout:            30 seconds
Re-Auth Interval:     7200 seconds
Max-EAP-Req:               10
Auth Type:                 SINGLE_HOST
Auth PAE State:            Initialize
Backend State:             Initialize
Auth PAE State:            Initialize
Backend State:             Initialize
```

Configuring Timeouts

If the supplicant or the authentication server is unresponsive, the authenticator terminates the authentication process after 30 seconds by default. You can configure the amount of time the authenticator waits for a response.

To terminate the authentication process, use the following commands:

- Terminate the authentication process due to an unresponsive supplicant.

  INTERFACE mode

  `dot1x supplicant-timeout seconds`

  The range is from 1 to 300.

  The default is **30**.
- Terminate the authentication process due to an unresponsive authentication server.

  INTERFACE mode

  `dot1x server-timeout seconds`

  The range is from 1 to 300.

  The default is **30**.

The example shows configuration information for a port for which the authenticator terminates the authentication process for an unresponsive supplicant or server after 15 seconds.

The bold lines show the new supplicant and server timeouts.

```
DellEMC(conf-if-Te-1/1)#dot1x port-control force-authorized
DellEMC(conf-if-Te-1/1)#do show dot1x interface TenGigabitEthernet 1/1

802.1x information on Te 1/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Disable
Guest VLAN id:             NONE
Auth-Fail VLAN:            Disable
Auth-Fail VLAN id:         NONE
Auth-Fail Max-Attempts:    NONE
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:     15 seconds
Server Timeout:      15 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10

Auth Type:                 SINGLE_HOST
Auth PAE State:            Initialize
Backend State:             Initialize
```

Enter the tasks the user should do after finishing this task (optional).

Configuring Dynamic VLAN Assignment with Port Authentication

Dell EMC Networking OS supports dynamic VLAN assignment when using 802.1X.

The basis for VLAN assignment is RADIUS attribute 81, Tunnel-Private-Group-ID. Dynamic VLAN assignment uses the standard dot1x procedure:

1. The host sends a dot1x packet to the Dell EMC Networking system
2. The system forwards a RADIUS REQEST packet containing the host MAC address and ingress port number
3. The RADIUS server authenticates the request and returns a RADIUS ACCEPT message with the VLAN assignment using Tunnel-Private-Group-ID

The illustration shows the configuration on the Dell EMC Networking system before connecting the end user device in black and blue text, and after connecting the device in red text. The blue text corresponds to the preceding numbered steps on dynamic VLAN assignment with 802.1X.

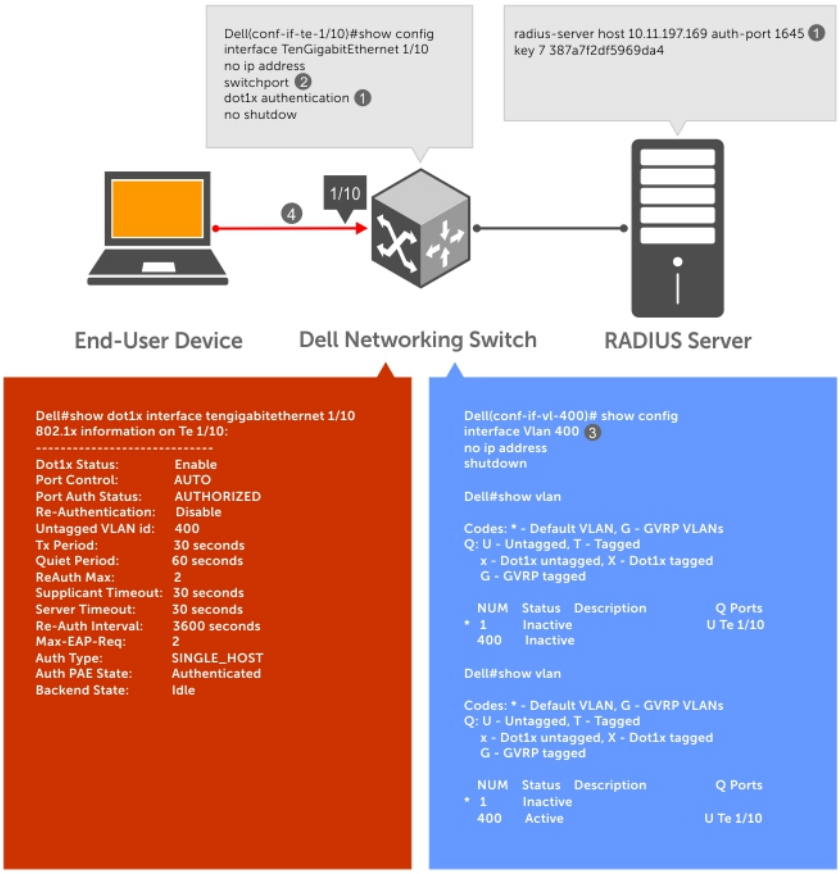


Figure 8. Dynamic VLAN Assignment

1. Configure 8021.x globally (refer to Enabling 802.1X) along with relevant RADIUS server configurations (refer to the illustration inDynamic VLAN Assignment with Port Authentication).
2. Make the interface a switchport so that it can be assigned to a VLAN.
3. Create the VLAN to which the interface will be assigned.
4. Connect the supplicant to the port configured for 802.1X.
5. Verify that the port has been authorized and placed in the desired VLAN (refer to the illustration in Dynamic VLAN Assignment with Port Authentication).

Guest and Authentication-Fail VLANs

Typically, the authenticator (the Dell system) denies the supplicant access to the network until the supplicant is authenticated. If the supplicant is authenticated, the authenticator enables the port and places it in either the VLAN for which the port is configured or the VLAN that the authentication server indicates in the authentication data.

NOTE: Ports cannot be dynamically assigned to the default VLAN.

If the supplicant fails authentication, the authenticator typically does not enable the port. In some cases this behavior is not appropriate. External users of an enterprise network, for example, might not be able to be authenticated, but still need access to the network. Also, some dumb-terminals, such as network printers, do not have 802.1X capability and therefore cannot authenticate themselves. To be able to connect such devices, they must be allowed access the network without compromising network security.

The Guest VLAN 802.1X extension addresses this limitation with regard to non-802.1X capable devices and the Authentication-fail VLAN 802.1X extension addresses this limitation with regard to external users.

- If the supplicant fails authentication a specified number of times, the authenticator places the port in the Authentication-fail VLAN.

- If a port is already forwarding on the Guest VLAN when 802.1X is enabled, the port is moved out of the Guest VLAN and the authentication process begins.

Configuring a Guest VLAN

If the supplicant does not respond within a determined amount of time ([reauth-max + 1] * tx-period, the system assumes that the host does not have 802.1X capability and the port is placed in the Guest VLAN.

NOTE: For more information about configuring timeouts, refer to Configuring Timeouts.

Configure a port to be placed in the Guest VLAN after failing to respond within the timeout period using the `dot1x guest-vlan` command from INTERFACE mode. View your configuration using the `show config` command from INTERFACE mode or using the `show dot1x interface` command from EXEC Privilege mode.

Example of Viewing Guest VLAN Configuration

```
DellEMC(conf-if-Te-2/1)#dot1x guest-vlan 200
DellEMC(conf-if-Te 2/1))#show config
!
interface TenGigabitEthernet 2/1
  switchport
  dot1x guest-vlan 200
  no shutdown
DellEMC(conf-if-Te 2/1))#
```

Configuring an Authentication-Fail VLAN

If the supplicant fails authentication, the authenticator re-attempts to authenticate after a specified amount of time.

NOTE: For more information about authenticator re-attempts, refer to Configuring a Quiet Period after a Failed Authentication.

You can configure the maximum number of times the authenticator re-attempts authentication after a failure (**3** by default), after which the port is placed in the Authentication-fail VLAN.

Configure a port to be placed in the VLAN after failing the authentication process as specified number of times using the `dot1x auth-fail-vlan` command from INTERFACE mode. Configure the maximum number of authentication attempts by the authenticator using the keyword `max-attempts` with this command.

Example of Configuring Maximum Authentication Attempts

```
DellEMC(conf-if-Te-2/1)#dot1x guest-vlan 200
DellEMC(conf-if-Te 2/1)#show config
!
interface TenGigabitEthernet 2/1
  switchport
  dot1x authentication
  dot1x guest-vlan 200
no shutdown
DellEMC(conf-if-Te-2/1)#

DellEMC(conf-if-Te-2/1)#dot1x auth-fail-vlan 100 max-attempts 5
DellEMC(conf-if-Te-2/1)#show config
!
interface TenGigabitEthernet 2/1
  switchport
  dot1x authentication
  dot1x guest-vlan 200
  dot1x auth-fail-vlan 100 max-attempts 5
no shutdown
DellEMC(conf-if-Te-2/1)#
```

View your configuration using the `show config` command from INTERFACE mode, as shown in the example in Configuring a Guest VLAN or using the `show dot1x interface` command from EXEC Privilege mode.

Example of Viewing Configured Authentication

```
802.1x information on Te 2/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:          Disabled
Guest VLAN id:        200
Auth-Fail VLAN:       Disabled
Auth-Fail VLAN id:     100
Auth-Fail Max-Attempts: 5
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:        15 seconds
Server Timeout:            15 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10
Auth Type:                 SINGLE_HOST

Auth PAE State:            Initialize
Backend State:             Initialize
```

Configuring dot1x Profile

You can configure a dot1x profile for defining a list of trusted supplicant MAC addresses. A maximum of 10 dot1x profiles can be configured. The profile name length is limited to 32 characters. The`dot1x profile {profile-name}` command sets the dot1x profile mode and you can enter profile-related commands, such as the `mac` command.

To configure a dot1x profile, use the following commands.

- Configure a dot1x profile.

  CONFIGURATION mode

  `dot1x profile {profile-name}`

  profile—name — Enter the dot1x profile name. The profile name length is limited to 32 characters.

```
DellEMC(conf)#dot1x profile test
DellEMC(conf-dot1x-profile)#

DellEMC#show dot1x profile

802.1x profile information
------------------------------
Dot1x Profile test
 Profile MACs
  00:00:00:00:01:11
```

Configuring the Static MAB and MAB Profile

Enable MAB (mac-auth-bypass) before using the `dot1x static-mab` command to enable static mab.

To enable static MAB and configure a static MAB profile, use the following commands.

- Configure static MAB and static MAB profile on dot1x interface.

  INTERFACE mode

  `dot1x static-mab profile profile-name`

  Eenter a name to configure the static MAB profile name. The profile name length is limited to a maximum of 32 characters.

```
DellEMC(conf-if-Te-2/1)#dot1x static-mab profile sample
DellEMC(conf-if-Te 2/1))#show config
!
```

```
interface TenGigabitEthernet 21
switchport
dot1x static-mab profile sample
no shutdown
DellEMC(conf-if-Te 2/1))#show dot1x interface TenGigabitEthernet 2/1


802.1x information on Te 2/1:
------------------------------
```

```
Dot1x Status:              Enable
Port Control:              Auto
Port Auth Status:          AUTHORIZED(STATIC-MAB)
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                 Enable
Guest VLAN id:             100
Auth-Fail VLAN:            Enable
Auth-Fail VLAN id:         200
Auth-Fail Max-Attempts:3
Critical VLAN:             Enable
Critical VLAN id:          300
Mac-Auth-Bypass Only:      Disable
Static-MAB:                Enable
Static-MAB Profile:        Sample
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:        30 seconds
Server Timeout:            30 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10
Auth Type:                 SINGLE_HOST
Auth PAE State:            Authenticated
Backend State:             Idle
```

Configuring Critical VLAN

By default, critical-VLAN is not configured. If authentication fails because of a server which is not reachable, user session is authenticated under critical-VLAN.

To configure a critical-VLAN for users or devices when authenticating server is not reachable, use the following command.

- Enable critical VLAN for users or devices

  INTERFACE mode

  `dot1x critical-vlan [{vlan-id}]`

  Specify a VLAN interface identifier to be configured as a critical VLAN. The VLAN ID range is 1– 4094.

```
DellEMC(conf-if-Te-2/1)#dot1x critical-vlan 300
DellEMC(conf-if-Te 2/1)#show config
!
interface TenGigabitEthernet 2/1
switchport
dot1x critical-vlan 300
no shutdown


DellEMC#show dot1x interface tengigabitethernet 2/1

802.1x information on Te 2/1:
-----------------------------------------------------
Dot1x Status:              Enable
Port Control:              AUTO
Port Auth Status:          AUTHORIZD(MAC-AUTH-BYPASS)
Critical VLAN       Enable
Critical VLAN id:     300
 Re-Authentication:        Disable
```

```
Untagged VLAN id:          400
Guest VLAN:                Enable
Guest VLAN id:             100
Auth-Fail VLAN:            Disable
Auth-Fail VLAN id:         NONE
Auth-Fail Max-Attempts:    NONE
Mac-Auth-Bypass:           Enable
Mac-Auth-Bypass Only:      Enable
Tx Period:                 3 seconds
Quiet Period:              60 seconds
ReAuth Max:                2
Supplicant Timeout:        30 seconds
Server Timeout:            30 seconds
Re-Auth Interval:          3600 seconds
Max-EAP-Req:               2
Host Mode:                 SINGLE_HOST
Auth PAE State:            Authenticated
Backend State:             Idle
```

6

Access Control List (ACL) VLAN Groups and Content Addressable Memory (CAM)

This section describes the access control list (ACL) virtual local area network (VLAN) group, and content addressable memory (CAM) enhancements.

Topics:

- Optimizing CAM Utilization During the Attachment of ACLs to VLANs
- Guidelines for Configuring ACL VLAN Groups
- Configuring ACL VLAN Groups and Configuring FP Blocks for VLAN Parameters
- Viewing CAM Usage
- Allocating FP Blocks for VLAN Processes

Optimizing CAM Utilization During the Attachment of ACLs to VLANs

To minimize the number of entries in CAM, enable and configure the ACL CAM feature. Use this feature when you apply ACLs to a VLAN (or a set of VLANs) and when you apply ACLs to a set of ports. The ACL CAM feature allows you to effectively use the Layer 3 CAM space with VLANs and Layer 2 and Layer 3 CAM space with ports.

To avoid using too much CAM space, configure ACL VLAN groups into a single group. A class identifier (Class ID) is assigned for each of the ACLs attached to the VLAN and this Class ID is used as an identifier or locator in the CAM space instead of the VLAN ID. This method of processing reduces the number of entries in the CAM area and saves memory space by using the Class ID for filtering in CAM instead of the VLAN ID.

When you apply an ACL separately on the VLAN interface, each ACL has a mapping with the VLAN and you use more CAM space. To maximize CAM space, create an ACL VLAN group and attach the ACL with the VLAN members.

The ACL manager application on the router processor (RP1) contains all the state information about all the ACL VLAN groups that are present. The ACL handler on the control processor (CP) and the ACL agent on the line cards do not contain any information about the group. After you enter the `acl-vlan-group` command, the ACL manager application performs the validation. If the command is valid, it is processed and sent to the agent, if required. If a configuration error is found or if the maximum limit has exceeded for the ACL VLAN groups present on the system, an error message displays. After you enter the `acl-vlan-group` command, the ACL manager application verifies the following parameters:

- Whether the CAM profile is set in virtual flow processing (VFP).
- Whether the maximum number of groups in the system is exceeded.
- Whether the maximum number of VLAN numbers permitted per ACL group is exceeded.
- When a VLAN member that is being added is already a part of another ACL group.

After these verification steps are performed, the ACL manager considers the command valid and sends the information to the ACL agent on the line card. The ACL manager notifies the ACL agent in the following cases:

- A VLAN member is added or removed from a group and previously associated VLANs exist in the group.
- The egress ACL is applied or removed from the group and the group contains VLAN members.
- VLAN members are added or deleted from a VLAN, which itself is a group member.
- A line card returns to the active state after going down and this line card contains a VLAN that is a member of an ACL group.
- The ACL VLAN group is deleted and it contains VLAN members.

The ACL manager does not notify the ACL agent in the following cases:

- The ACL VLAN group is created.
- The ACL VLAN group is deleted and it does not contain VLAN members.
- The ACL is applied or removed from a group and the ACL group does not contain a VLAN member.
- The description of the ACL group is added or removed.

Guidelines for Configuring ACL VLAN Groups

Keep the following points in mind when you configure ACL VLAN groups:

- The interfaces where you apply the ACL VLAN group function as restricted interfaces. The ACL VLAN group name identifies the group of VLANs that performs hierarchical filtering.
- You can add only one ACL to an interface at a time.
- When you attach an ACL VLAN group to the same interface, validation performs to determine whether the ACL is applied directly to an interface. If you previously applied an ACL separately to the interface, an error occurs when you attempt to attach an ACL VLAN group to the same interface.
- The maximum number of members in an ACL VLAN group is determined by the type of switch and its hardware capabilities. This scaling limit depends on the number of slices that are allocated for ACL CAM optimization. If one slice is allocated, the maximum number of VLAN members is 256 for all ACL VLAN groups. If two slices are allocated, the maximum number of VLAN members is 512 for all ACL VLAN groups.
- The maximum number of VLAN groups that you can configure also depends on the hardware specifications of the switch. Each VLAN group is mapped to a unique ID in the hardware. The maximum number of ACL VLAN groups supported is 31. Only a maximum of two components (iSCSI counters, Open Flow, ACL optimization, and so on) can be allocated virtual flow processing slices at a time.
- Port ACL optimization is applicable only for ACLs that are applied without the VLAN range.
- If you enable the ACL VLAN group capability, you cannot view the statistical details of ACL rules per VLAN and per interface. You can only view the counters per ACL only using the `show ip accounting access list` command.
- Within a port, you can apply Layer 2 ACLs on a VLAN or a set of VLANs. In this case, CAM optimization is not applied.
- To enable optimization of CAM space for Layer 2 or Layer 3 ACLs that are applied to ports, the port number is removed as a qualifier for ACL application on ports, and port bits are used. When you apply the same ACL to a set of ports, the port bitmap is set when the ACL flow processor (FP) entry is added. When you remove the ACL from a port, the port bitmap is removed.
- If you do not attach an ACL to any of the ports, the FP entries are deleted. Similarly, when the same ACL is applied on a set of ports, only one set of entries is installed in the FP, thereby saving CAM space. Enable optimization using the optimized option in the `ip access-group` command. This option is not valid for VLAN and link aggregation group (LAG) interfaces.

Configuring ACL VLAN Groups and Configuring FP Blocks for VLAN Parameters

This section describes how to optimize CAM blocks by configuring ACL VLAN groups that you can attach to VLAN interfaces. It also describes how to configure FP blocks for different VLAN operations.

Configuring ACL VLAN Groups

You can create an ACL VLAN group and attach the ACL with the VLAN members. The optimization is applicable only when you create an ACL VLAN group.

1. Create an ACL VLAN group.

   CONFIGURATION mode

```
acl-vlan-group {group name}
```

2. Add a description to the ACL VLAN group.
   CONFIGURATION (conf-acl-vl-grp) mode
   ```
   description description
   ```
3. Apply an egress IP ACL to the ACL VLAN group.
   CONFIGURATION (conf-acl-vl-grp) mode
   ```
   ip access-group {group name} out implicit-permit
   ```
4. Add VLAN member(s) to an ACL VLAN group.
   CONFIGURATION (conf-acl-vl-grp) mode
   ```
   member vlan {VLAN-range}
   ```
5. Display all the ACL VLAN groups or display a specific ACL VLAN group, identified by name.
   CONFIGURATION (conf-acl-vl-grp) mode
   ```
   show acl-vlan-group {group name | detail}
   ```

```
DellEMC#show acl-vlan-group detail

Group Name :
  TestGroupSeventeenTwenty
Egress IP Acl :
  SpecialAccessOnlyExpertsAllowed
Vlan Members :
  100,200,300

Group Name :
  CustomerNumberIdentificationEleven
Egress IP Acl :
  AnyEmployeeCustomerElevenGrantedAccess
Vlan Members :
  2-10,99

Group Name :
  HostGroup
Egress IP Acl :
  Group5
Vlan Members :
  1,1000
DellEMC#
```

Configuring FP Blocks for VLAN Parameters

To allocate the number of FP blocks for the various VLAN processes on the system, use the `cam-acl-vlan` command. To reset the number of FP blocks to the default, use the `no` version of this command. By default, 0 groups are allocated for the ACL in VLAN contentaware processor (VCAP). ACL VLAN groups or CAM optimization is not enabled by default. You also must allocate the slices for CAM optimization.

1. Allocate the number of FP blocks for VLAN operations.
   CONFIGURATION mode
   ```
   cam-acl-vlan vlanopenflow <0-2>
   ```
2. Allocate the number of FP blocks for VLAN iSCSI counters.
   CONFIGURATION mode
   ```
   cam-acl-vlan vlaniscsi <0-2>
   ```
3. Allocate the number of FP blocks for ACL VLAN optimization.
   CONFIGURATION mode
   ```
   cam-acl-vlan vlanaclopt <0-2>
   ```
4. View the number of FP blocks that is allocated for the different VLAN services.

EXEC Privilege mode

```
DellEMC#show cam-usage switch
Stackunit|Portpipe| CAM Partition  | Total CAM  |  Used CAM  |Available CAM
========|========|=================|============|============|=============
    1   |   0    | IN-L2 ACL       |      1536  |          0 |       1536
        |        | OUT-L2 ACL      |       206  |          9 |        197
Codes: * - cam usage is above 90%.
```

Viewing CAM Usage

View the amount of CAM space available, used, and remaining in each partition (including IPv4Flow and Layer 2 ACL sub-partitions) using the `show cam-usage` command in EXEC Privilege mode.

Display Layer 2, Layer 3, ACL, or all CAM usage statistics.

EXCE Privilege mode

```
show cam usage [acl | router | switch]
```

The following output shows CAM blocks usage for Layer 2 and Layer 3 ACLs and other processes that use CAM space:

```
DellEMC#show cam-usage
Stackunit|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available CAM
========|========|=================|=============|=============|==============
    1   |   0    | IN-L2 ACL       |      1536   |          0  |       1536
        |        | IN-L3 ACL       |      1024   |          1  |       1023
        |        | IN-L3 ECMP GRP  |      1024   |          0  |       1024
        |        | IN-L3 FIB       |     49152   |          3  |      49149
        |        | IN-V6 ACL       |         0   |          0  |          0
        |        | IN-NLB ACL      |         0   |          0  |          0
        |        | IPMAC ACL       |         0   |          0  |          0
        |        | OUT-L2 ACL      |       206   |          9  |        197
        |        | OUT-L3 ACL      |       178   |          9  |        169
        |        | OUT-V6 ACL      |       178   |          4  |        174
    2   |   0    | IN-L2 ACL       |      1536   |          0  |       1536
        |        | IN-L3 ACL       |      1024   |          1  |       1023
        |        | IN-L3 FIB       |     49152   |          3  |      49149
        |        | IN-V6 ACL       |         0   |          0  |          0
        |        | IN-NLB ACL      |         0   |          0  |          0
        |        | IPMAC ACL       |         0   |          0  |          0
        |        | OUT-L2 ACL      |       206   |          9  |        197
        |        | OUT-L3 ACL      |       178   |          9  |        169
        |        | OUT-V6 ACL      |       178   |          4  |        174
    3   |   0    | IN-L2 ACL       |      1536   |          0  |       1536
        |        | IN-L3 ACL       |      1024   |          1  |       1023
        |        | IN-L3 FIB       |     49152   |          3  |      49149
        |        | IN-V6 ACL       |         0   |          0  |          0
        |        | IN-NLB ACL      |         0   |          0  |          0
        |        | IPMAC ACL       |         0   |          0  |          0
        |        | OUT-L2 ACL      |       206   |          9  |        197
        |        | OUT-L3 ACL      |       178   |          9  |        169
        |        | OUT-V6 ACL      |       178   |          4  |        174
Codes: * - cam usage is above 90%.
```

The following output displays CAM space usage when you configure Layer 2 and Layer 3 ACLs:

```
DellEMC#show cam-usage acl
Stackunit|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available CAM
========|========|=================|=============|=============|==============
    1   |   0    | IN-L2 ACL       |      1536   |          0  |       1536
        |        | IN-L3 ACL       |      1024   |          1  |       1023
        |        | IN-L3 ECMP GRP  |      1024   |          0  |       1024
        |        | IN-V6 ACL       |         0   |          0  |          0
        |        | OUT-L2 ACL      |       206   |          9  |        197
        |        | OUT-L3 ACL      |       178   |          9  |        169
        |        | OUT-V6 ACL      |       178   |          4  |        174
    2   |   0    | IN-L2 ACL       |      1536   |          0  |       1536
        |        | IN-L3 ACL       |      1024   |          1  |       1023
```

```
        |         | IN-V6 ACL       |          0    |          0    |          0
        |         | OUT-L2 ACL      |        206    |          9    |        197
        |         | OUT-L3 ACL      |        178    |          9    |        169
        |         | OUT-V6 ACL      |        178    |          4    |        174
    3   |    0    | IN-L2 ACL       |       1536    |          0    |       1536
        |         | IN-L3 ACL       |       1024    |          1    |       1023
        |         | IN-V6 ACL       |          0    |          0    |          0
        |         | OUT-L2 ACL      |        206    |          9    |        197
        |         | OUT-L3 ACL      |        178    |          9    |        169
        |         | OUT-V6 ACL      |        178    |          4    |        174
Codes: * - cam usage is above 90%.
```

The following output displays CAM space usage for Layer 2 ACLs:

```
DellEMC#show cam-usage switch
Stackunit|Portpipe| CAM Partition    | Total CAM    |  Used CAM    |Available CAM
========|========|=================|==============|==============|==============
    1   |    0    | IN-L2 ACL       |       1536    |          0    |       1536
        |         | OUT-L2 ACL      |        206    |          9    |        197
    2   |    0    | IN-L2 ACL       |       1536    |          0    |       1536
        |         | OUT-L2 ACL      |        206    |          9    |        197
    3   |    0    | IN-L2 ACL       |       1536    |          0    |       1536
        |         | OUT-L2 ACL      |        206    |          9    |        197
        |         | IN-L3 ECMP GRP  |       1024    |          0    |       1024
Codes: * - cam usage is above 90%.
```

The following output displays CAM space usage for Layer 3 ACLs:

```
DellEMC#show cam-usage router
Stackunit|Portpipe| CAM Partition    | Total CAM    |  Used CAM    |Available CAM
========|========|=================|==============|==============|==============
    1   |    0    | IN-L3 FIB       |      49152    |          3    |      49149
        |         | IN-L3 ACL       |       1024    |          1    |       1023
        |         | IN-L3 ECMP GRP  |       1024    |          0    |       1024
        |         | IN-V6 ACL       |          0    |          0    |          0
        |         | OUT-L3 ACL      |        178    |          9    |        169
        |         | OUT-V6 ACL      |        178    |          4    |        174
    2   |    0    | IN-L3 FIB       |      49152    |          3    |      49149
        |         | IN-L3 ACL       |       1024    |          1    |       1023
        |         | IN-V6 ACL       |          0    |          0    |          0
        |         | OUT-L3 ACL      |        178    |          9    |        169
        |         | OUT-V6 ACL      |        178    |          4    |        174
    3   |    0    | IN-L3 FIB       |      49152    |          3    |      49149
        |         | IN-L3 ACL       |       1024    |          1    |       1023
        |         | IN-V6 ACL       |          0    |          0    |          0
        |         | OUT-L3 ACL      |        178    |          9    |        169
        |         | OUT-V6 ACL      |        178    |          4    |        174
Codes: * - cam usage is above 90%.
```

Allocating FP Blocks for VLAN Processes

The VLAN contentaware processor (VCAP) application is a pre-ingress CAP that modifies the VLAN settings before packets are forwarded. To support ACL CAM optimization, the CAM carving feature is enhanced. A total of four VCAP groups are present: two fixed groups and two dynamic groups. Of the two dynamic groups, you can allocate zero, one, or two FP blocks to iSCSI Counters, Open Flow, and ACL Optimization.

You can configure only two of these features at a time.

- To allocate the number of FP blocks for VLAN open flow operations, use the `cam-acl-vlan vlanopenflow <0-2>` command.
- To allocate the number of FP blocks for VLAN iSCSI counters, use the `cam-acl-vlan vlaniscsi <0-2>` command.
- To allocate the number of FP blocks for ACL VLAN optimization, use the `cam-acl-vlan vlanaclopt <0-2>` command.

To reset the number of FP blocks to the default, use the `no` version of these commands. By default, zero groups are allocated for the ACL in VCAP. ACL VLAN groups or CAM optimization is not enabled by default. You must also allocate the slices for CAM optimization.

To display the number of FP blocks that is allocated for the different VLAN services, use the `show cam-acl-vlan` command. After you configure the ACL VLAN groups, reboot the system to store the settings in nonvolatile storage. During CAM initialization, the chassis manager reads the NVRAM and allocates the dynamic VCAP regions.

7

Access Control Lists (ACLs)

This chapter describes access control lists (ACLs), prefix lists, and route-maps.

At their simplest, access control lists (ACLs), prefix lists, and route-maps permit or deny traffic based on MAC and/or IP addresses. This chapter describes implementing IP ACLs, IP prefix lists and route-maps. For MAC ACLS, refer to Layer 2.

An ACL is essentially a filter containing some criteria to match (examine IP, transmission control protocol [TCP], or user datagram protocol [UDP] packets) and an action to take (permit or deny). ACLs are processed in sequence so that if a packet does not match the criterion in the first filter, the second filter (if configured) is applied. When a packet matches a filter, the switch drops or forwards the packet based on the filter's specified action. If the packet does not match any of the filters in the ACL, the packet is dropped (implicit deny).

The number of ACLs supported on a system depends on your content addressable memory (CAM) size. For more information, refer to User Configurable CAM Allocation and CAM Optimization. For complete CAM profiling information, refer to Content Addressable Memory (CAM).

You can configure ACLs on VRF instances. In addition to the existing qualifying parameters, Layer 3 ACLs also incorporate VRF ID as one of the parameters. Using this new capability, you can also configure VRF based ACLs on interfaces.

NOTE: You can apply Layer 3 VRF-aware ACLs only at the ingress level.

You can apply VRF-aware ACLs on:

- VRF Instances
- Interfaces

In order to configure VRF-aware ACLs on VRF instances, you must carve out a separate CAM region. You can use the `cam-acl` command for allocating CAM regions. As part of the enhancements to support VRF-aware ACLs, the `cam-acl` command now includes the following new parameter that enables you to allocate a CAM region: *vrfv4acl*.

The order of priority for configuring user-defined ACL CAM regions is as follows:

- V4 ACL CAM
- VRF V4 ACL CAM
- L2 ACL CAM

With the inclusion of VRF based ACLs, the order of precedence of Layer 3 ACL rules is as follows:

- Port/VLAN based PERMIT/DENY Rules
- Port/VLAN based IMPLICIT DENY Rules
- VRF based PERMIT/DENY Rules
- VRF based IMPLICIT DENY Rules

NOTE: In order for the VRF ACLs to take effect, ACLs configured in the Layer 3 CAM region must have an implicit-permit option.

You can use the `ip access-group` command to configure VRF-aware ACLs on interfaces. Using the `ip access-group` command, in addition to a range of VLANs, you can also specify a range of VRFs as input for configuring ACLs on interfaces. The VRF range is from 1 to 511. These ACLs use the existing V4 ACL CAM region to populate the entries in the hardware and do not require you to carve out a separate CAM region.

NOTE: You can configure VRF-aware ACLs on interfaces either using a range of VLANs or a range of VRFs but not both.

Topics:

- IP Access Control Lists (ACLs)
- IP Fragment Handling
- Configure a Standard IP ACL
- Configure an Extended IP ACL
- Configure Layer 2 and Layer 3 ACLs
- Assign an IP ACL to an Interface
- Applying an IP ACL

- Configure Ingress ACLs
- Configure Egress ACLs
- IP Prefix Lists
- ACL Remarks
- ACL Resequencing
- Route Maps
- Important Points to Remember
- Logging of ACL Processes
- Flow-Based Monitoring

IP Access Control Lists (ACLs)

In Dell EMC Networking switch/routers, you can create two different types of IP ACLs: standard or extended.

A standard ACL filters packets based on the source IP packet. An extended ACL filters traffic based on the following criteria:

- IP protocol number
- Source IP address
- Destination IP address
- Source TCP port number
- Destination TCP port number
- Source UDP port number
- Destination UDP port number

For more information about ACL options, refer to the *Dell EMC Networking OS Command Reference Guide*.

For extended ACL, TCP, and UDP filters, you can match criteria on specific or ranges of TCP or UDP ports. For extended ACL TCP filters, you can also match criteria on established TCP sessions.

When creating an access list, the sequence of the filters is important. You have a choice of assigning sequence numbers to the filters as you enter them, or the Dell EMC Networking Operating System (OS) assigns numbers in the order the filters are created. The sequence numbers are listed in the display output of the `show config` and `show ip accounting access-list` commands.

Ingress and egress Hot Lock ACLs allow you to append or delete new rules into an existing ACL (already written into CAM) without disrupting traffic flow. Existing entries in the CAM are shuffled to accommodate the new entries. Hot lock ACLs are enabled by default and support both standard and extended ACLs and on all platforms.

NOTE: Hot lock ACLs are supported for Ingress ACLs only.

CAM Usage

The following section describes CAM allocation and CAM optimization.

- User Configurable CAM Allocation
- CAM Optimization

User Configurable CAM Allocation

Allocate space for IPV6 ACLs by using the `cam-acl` command in CONFIGURATION mode.

The CAM space is allotted in filter processor (FP) blocks. The total space allocated must equal 13 FP blocks. (There are 16 FP blocks, but System Flow requires three blocks that cannot be reallocated.)

Enter the `ipv6acl` allocation as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd numbered ranges.

If you want to configure ACL's on VRF instances, you must allocate a CAM region using the *vrfv4acl* option in the `cam-acl` command.

Save the new CAM settings to the startup-config (use `write-mem` or `copy run start`) then reload the system for the new settings to take effect.

CAM Optimization

When you enable this command, if a policy map containing classification rules (ACL and/or dscp/ ip-precedence rules) is applied to more than one physical interface on the same port-pipe, only a single copy of the policy is written (only one FP entry is used). When you disable this command, the system behaves as described in this chapter.

Test CAM Usage

This command applies to both IPv4 and IPv6 CAM profiles, but is best used when verifying QoS optimization for IPv6 ACLs.

To determine whether sufficient ACL CAM space is available to enable a service-policy, use this command. To verify the actual CAM space required, create a class map with all the required ACL rules, then execute the `test cam-usage` command in Privilege mode. The following example shows the output when executing this command. The status column indicates whether you can enable the policy.

Example of the `test cam-usage` Command

```
DellEMC#test cam-usage service-policy input asd stack-unit 1 port-set 0


Stack-unit|Portpipe|CAM Partition|Available CAM|Estimated CAM per Port|Status
--------------------------------------------------------------------------
       1|       1|     IPv4Flow|          232|                     0|Allowed
DellEMC#
```

Implementing ACLs on Dell EMC Networking OS

You can assign one IP ACL per interface. If you do not assign an IP ACL to an interface, it is not used by the software.

The number of entries allowed per ACL is hardware-dependent.

If counters are enabled on ACL rules that are already configured, those counters are reset when a new rule which is inserted or prepended or appended requires a hardware shift in the flow table. Resetting the counters to 0 is transient as the proginal counter values are retained after a few seconds. If there is no need to shift the flow in the hardware, the counters are not affected. This is applicable to the following features:

- L2 Ingress Access list
- L2 Egress Access list

In the Dell EMC Networking OS versions prior to 9.13(0.0), the system does not install any of your ACL rules if the available CAM space is lesser than what is required for your set of ACL rules. Effective with the Dell EMC Networking OS version 9.13(0.0), the system installs your ACL rules until all the allocated CAM memory is used. If there is no implicit permit in your rule, the Dell EMC Networking OS ensures that an implicit deny is installed at the end of your rule. This behavior is applicable for IPv4 and IPv6 ingress and egress ACLs.

NOTE: System access lists (system-flow entries) are pre-programmed in the system for lifting the control-plane packets destined for the local device which the CPU needs to process. The system access lists always override the user configured access lists. Even if you configure ACL to block certain hosts, control plane protocols such as, ARP, BGP, LACP, VLT, VRRP and so on, associated with such hosts cannot be blocked.

Assigning ACLs to VLANs

When you apply an ACL to a VLAN using single port-pipe, a copy of the ACL entries gets installed in the ACL CAM on the port-pipe. The entry looks for the incoming VLAN in the packet. When you apply an ACL on individual ports of a VLAN, separate copies of the ACL entries are installed for each port belonging to a port-pipe.

You can use the `log` keyword to log the details about the packets that match. The control processor becomes busy based on the number of packets that match the log entry and the rate at which the details are logged in. However, the route processor (RP) is unaffected. You can use this option for debugging issues related to control traffic.

ACL Optimization

If an access list contains duplicate entries, Dell EMC Networking OS deletes one entry to conserve CAM space.

Standard and extended ACLs take up the same amount of CAM space. A single ACL rule uses two CAM entries to identify whether the access list is a standard or extended ACL.

Determine the Order in which ACLs are Used to Classify Traffic

When you link class-maps to queues using the `service-queue` command, Dell EMC Networking OS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities).

As shown in the following example, class-map *cmap2* is matched against ingress packets before *cmap1*.

ACLs *acl1* and *acl2* have overlapping rules because the address range 20.1.1.0/24 is within 20.0.0.0/8. Therefore (without the keyword order), packets within the range 20.1.1.0/24 match positive against *cmap1* and are buffered in queue 7, though you intended for these packets to match positive against *cmap2* and be buffered in queue 4.

In cases where class-maps with overlapping ACL rules are applied to different queues, use the `order` keyword to specify the order in which you want to apply ACL rules. The order can range from 0 to 254. Dell EMC Networking OS writes to the CAM ACL rules with lower-order numbers (order numbers closer to 0) before rules with higher-order numbers so that packets are matched as you intended. By default, all ACL rules have an order of **255**.

Example of the `order` Keyword to Determine ACL Sequence

```
DellEMC(conf)#ip access-list standard acl1
DellEMC(config-std-nacl)#permit 20.0.0.0/8
DellEMC(config-std-nacl)#exit
DellEMC(conf)#ip access-list standard acl2
DellEMC(config-std-nacl)#permit 20.1.1.0/24 order 0
DellEMC(config-std-nacl)#exit
DellEMC(conf)#class-map match-all cmap1
DellEMC(conf-class-map)#match ip access-group acl1
DellEMC(conf-class-map)#exit
DellEMC(conf)#class-map match-all cmap2
DellEMC(conf-class-map)#match ip access-group acl2
DellEMC(conf-class-map)#exit
DellEMC(conf)#policy-map-input pmap
DellEMC(conf-policy-map-in)#service-queue 7 class-map cmap1
DellEMC(conf-policy-map-in)#service-queue 4 class-map cmap2
DellEMC(conf-policy-map-in)#exit
DellEMC(conf)#interface te 10/1
DellEMC(conf-if-te-10/1)#service-policy input pmap
```

IP Fragment Handling

Dell EMC Networking OS supports a configurable option to explicitly deny IP fragmented packets, particularly second and subsequent packets.

It extends the existing ACL command syntax with the `fragments` keyword for all Layer 3 rules applicable to all Layer protocols (permit/deny ip/tcp/udp/icmp).

- Both standard and extended ACLs support IP fragments.
- Second and subsequent fragments are allowed because a Layer 4 rule cannot be applied to these fragments. If the packet is to be denied eventually, the first fragment would be denied and hence the packet as a whole cannot be reassembled.
- Implementing the required rules uses a significant number of CAM entries per TCP/UDP entry.
- For IP ACL, Dell EMC Networking OS always applies implicit deny. You do not have to configure it.
- For IP ACL, Dell EMC Networking OS applies implicit permit for second and subsequent fragment just prior to the implicit deny.
- If you configure an *explicit* deny, the second and subsequent fragments do not hit the implicit permit rule for fragments.
- Loopback interfaces do not support ACLs using the `IP fragment` option. If you configure an ACL with the `fragments` option and apply it to a Loopback interface, the command is accepted but the ACL entries are not actually installed the offending rule in CAM.

IP Fragments ACL Examples

The following examples show how you can use ACL commands with the `fragment` keyword to filter fragmented packets.

The following configuration permits all packets (both fragmented and non-fragmented) with destination IP 10.1.1.1. The second rule does not get hit at all.

Example of Permitting All Packets on an Interface

```
DellEMC(conf)#ip access-list extended ABC
DellEMC(conf-ext-nacl)#permit ip any 10.1.1.1/32
DellEMC(conf-ext-nacl)#deny ip any 10.1.1.1/32 fragments
DellEMC(conf-ext-nacl)
```

To deny the second/subsequent fragments, use the same rules in a different order. These ACLs deny all second and subsequent fragments with destination IP 10.1.1.1 but permit the first fragment and non-fragmented packets with destination IP 10.1.1.1.

Example of Denying Second and Subsequent Fragments

```
DellEMC(conf)#ip access-list extended ABC
DellEMC(conf-ext-nacl)#deny ip any 10.1.1.1/32 fragments
DellEMC(conf-ext-nacl)#permit ip any 10.1.1.1/32
DellEMC(conf-ext-nacl)
```

Layer 4 ACL Rules Examples

The following examples show the ACL commands for Layer 4 packet filtering.

Permit an ACL line with L3 information only, and the `fragments` keyword is present: If a packet's L3 information matches the L3 information in the ACL line, the packet's FO is checked.

- If a packet's FO > 0, the packet is permitted.
- If a packet's FO = 0, the next ACL entry is processed.

Deny ACL line with L3 information only, and the `fragments` keyword is present: If a packet's L3 information does match the L3 information in the ACL line, the packet's FO is checked.

- If a packet's FO > 0, the packet is denied.
- If a packet's FO = 0, the next ACL line is processed.

In this first example, TCP packets from host 10.1.1.1 with TCP destination port equal to 24 are permitted. All others are denied.

Example of Permitting All Packets from a Specified Host

```
DellEMC(conf)#ip access-list extended ABC
DellEMC(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
DellEMC(conf-ext-nacl)#deny ip any any fragment
DellEMC(conf-ext-nacl)
```

In the following example, the TCP packets that are first fragments or non-fragmented from host 10.1.1.1 with TCP destination port equal to 24 are permitted. Additionally, all TCP non-first fragments from host 10.1.1.1 are permitted. All other IP packets that are non-first fragments are denied.

Example of Permitting Only First Fragments and Non-Fragmented Packets from a Specified Host

```
DellEMC(conf)#ip access-list extended ABC
DellEMC(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
DellEMC(conf-ext-nacl)#permit tcp host 10.1.1.1 any fragment
DellEMC(conf-ext-nacl)#deny ip any any fragment
DellEMC(conf-ext-nacl)
```

To log all the packets denied and to override the implicit deny rule and the implicit permit rule for TCP/ UDP fragments, use a configuration similar to the following.

Example of Logging Denied Packets

```
DellEMC(conf)#ip access-list extended ABC
DellEMC(conf-ext-nacl)#permit tcp any any fragment
DellEMC(conf-ext-nacl)#permit udp any any fragment
```

```
DellEMC(conf-ext-nacl)#deny ip any any log
DellEMC(conf-ext-nacl)
```

When configuring ACLs with the fragments keyword, be aware of the following.

When an ACL filters packets, it looks at the fragment offset (FO) to determine whether it is a fragment.

- FO = 0 means it is either the first fragment or the packet is a non-fragment.
- FO > 0 means it is dealing with the fragments of the original packet.

Configure a Standard IP ACL

To configure an ACL, use commands in IP ACCESS LIST mode and INTERFACE mode.

For a complete list of all the commands related to IP ACLs, refer to the *Dell EMC Networking OS Command Line Interface Reference Guide*. To set up extended ACLs, refer to Configure an Extended IP ACL.

A standard IP ACL uses the source IP address as its match criterion.

1. Enter IP ACCESS LIST mode by naming a standard IP access list.

   CONFIGURATION mode

   `ip access-list standard access-listname`

2. Configure a drop or forward filter.

   CONFIG-STD-NACL mode

   `seq sequence-number {deny | permit} {source [mask] | any | host ip-address} [count [byte] [dscp] [order] [monitor [session-id]] [fragments]`

NOTE: When assigning sequence numbers to filters, keep in mind that you might need to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five.

To view the rules of a particular ACL configured on a particular interface, use the `show ip accounting access-list ACL-name interface interface` command in EXEC Privilege mode.

The following is an example of viewing the rules of a specific ACL on an interface.

```
DellEMC#show ip accounting access-list ToOspf interface gig 1/6
Standard IP access list ToOspf
  seq 5 deny any
  seq 10 deny 10.2.0.0 /16
  seq 15 deny 10.3.0.0 /16
  seq 20 deny 10.4.0.0 /16
  seq 25 deny 10.5.0.0 /16
  seq 30 deny 10.6.0.0 /16
  seq 35 deny 10.7.0.0 /16
  seq 40 deny 10.8.0.0 /16
  seq 45 deny 10.9.0.0 /16
  seq 50 deny 10.10.0.0 /16
DellEMC#
```

The following example shows how the `seq` command orders the filters according to the sequence number assigned. In the example, filter 25 was configured before filter 15, but the `show config` command displays the filters in the correct order.

```
DellEMC(config-std-nacl)#seq 25 deny ip host 10.5.0.0 any log
DellEMC(config-std-nacl)#seq 15 permit tcp 10.3.0.0 /16 any monitor 300
DellEMC(config-std-nacl)#show config
!
ip access-list standard dilling
  seq 15 permit tcp 10.3.0.0/16 any monitor 300
  seq 25 deny ip host 10.5.0.0 any log
DellEMC(config-std-nacl)#
```

To delete a filter, use the `no seq sequence-number` command in IP ACCESS LIST mode. If you are creating a standard ACL with only one or two filters, you can let Dell EMC Networking OS assign a sequence number based on the order in which the filters are configured. The software assigns filters in multiples of 5.

Configuring a Standard IP ACL Filter

If you are creating a standard ACL with only one or two filters, you can let Dell EMC Networking OS assign a sequence number based on the order in which the filters are configured. The software assigns filters in multiples of five.

1. Configure a standard IP ACL and assign it a unique name.

   CONFIGURATION mode

   ```
   ip access-list standard access-list-name
   ```

2. Configure a drop or forward IP ACL filter.

   CONFIG-STD-NACL mode

   ```
   {deny | permit} {source [mask] | any | host ip-address} [count [byte] [dscp] [order]
   [monitor [session-id]] [fragments]
   ```

When you use the `log` keyword, the CP logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details. The following example shows a standard IP ACL in which Dell EMC Networking OS assigns the sequence numbers. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The `show config` command in IP ACCESS LIST mode displays the two filters with the sequence numbers 5 and 10.

```
DellEMC(config-route-map)#ip access standard acl1
DellEMC(config-std-nacl)#permit 10.1.0.0/16 monitor 177
DellEMC(config-std-nacl)#show config
!
ip access-list standard acl1
  seq 5 permit 10.1.0.0/16 monitor 177
DellEMC(config-std-nacl)#
```

To view all configured IP ACLs, use the `show ip accounting access-list` command in EXEC Privilege mode.

The following examples shows how to view a standard ACL filter sequence for an interface.

```
DellEMC#show ip accounting access example interface gig 4/12
Extended IP access list example
  seq 15 deny udp any any eq 111
  seq 20 deny udp any any eq 2049
  seq 25 deny udp any any eq 31337
  seq 30 deny tcp any any range 12345 12346
  seq 35 permit udp host 10.21.126.225 10.4.5.0 /28 monitor 300
  seq 40 permit udp host 10.21.126.226 10.4.5.0 /28
  seq 45 permit udp 10.8.0.0 /16 10.50.188.118 /31 range 1812 1813
  seq 50 permit tcp 10.8.0.0 /16 10.50.188.118 /31 eq 49 monitor 349
  seq 55 permit udp 10.15.1.0 /24 10.50.188.118 /31 range 1812 1813
```

To delete a filter, enter the `show config` command in IP ACCESS LIST mode and locate the sequence number of the filter you want to delete. Then use the `no seq sequence-number` command in IP ACCESS LIST mode.

Configure an Extended IP ACL

Extended IP ACLs filter on source and destination IP addresses, IP host addresses, TCP addresses, TCP host addresses, UDP addresses, and UDP host addresses.

The traffic passes through the filter in the order of the filter's sequence and hence you can configure the extended IP ACL by first entering IP ACCESS LIST mode, and then assigning a sequence number to the filter.

Configuring Filters with a Sequence Number

To configure filters with a sequence number, use the following commands.

1. Enter IP ACCESS LIST mode by creating an extended IP ACL.

   CONFIGURATION mode

   ```
   ip access-list extended access-list-name
   ```

2. Configure a drop or forward filter.

CONFIG-EXT-NACL mode

```
seq sequence-number {deny | permit} {ip-protocol-number | icmp | ip | tcp | udp} {source mask | any | host ip-address} {destination mask | any | host ip-address} [operator [portnumber ] [count [byte]] [order] [monitor [session-id]] [fragments]
```

When you use the `log` keyword, the CP logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details.

Configure Filters, TCP Packets

To create a filter for TCP packets with a specified sequence number, use the following commands.

1. Create an extended IP ACL and assign it a unique name.

   CONFIGURATION mode

   ```
   ip access-list extended access-list-name
   ```

2. Configure an extended IP ACL filter for TCP packets.

   CONFIG-EXT-NACL mode

   ```
   seq sequence-number {deny | permit} tcp {source mask | any | host ip-address} [count [byte]] [order] [monitor [session-id]] [fragments]
   ```

Configure Filters, UDP Packets

To create a filter for UDP packets with a specified sequence number, use the following commands.

1. Create an extended IP ACL and assign it a unique name.

   CONFIGURATION mode

   ```
   ip access-list extended access-list-name
   ```

2. Configure an extended IP ACL filter for UDP packets.

   CONFIG-EXT-NACL mode

   ```
   seq sequence-number {deny | permit} tcp {source mask | any | host ip-address} [count [byte]] [order] [monitor [session-id]] [fragments]
   ```

When you create the filters with a specific sequence number, you can create the filters in any order and the filters are placed in the correct order.

NOTE: When assigning sequence numbers to filters, you may have to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five or another number.

The example below shows how the `seq` command orders the filters according to the sequence number assigned. In the example, filter 15 was configured before filter 5, but the `show config` command displays the filters in the correct order.

```
DellEMC(config-ext-nacl)#seq 15 deny ip host 112.45.0.0 any log monitor 501
DellEMC(config-ext-nacl)#seq 5 permit tcp 12.1.3.45 0.0.255.255 any
DellEMC(config-ext-nacl)#show config
!
ip access-list extended dilling
   seq 5 permit tcp 12.1.0.0 0.0.255.255 any
   seq 15 deny ip host 112.45.0.0 any log monitor 501
DellEMC(config-ext-nacl)#
```

Configuring Filters Without a Sequence Number

If you are creating an extended ACL with only one or two filters, you can let Dell EMC Networking OS assign a sequence number based on the order in which the filters are configured. Dell EMC Networking OS assigns filters in multiples of five.

To configure a filter for an extended IP ACL without a specified sequence number, use any or all of the following commands:

- Configure a deny or permit filter to examine IP packets.

  CONFIG-EXT-NACL mode

`{deny | permit} {source mask | any | host ip-address} [count [byte]] [order] [monitor [session-id]] [fragments]`

- Configure a deny or permit filter to examine TCP packets.

  CONFIG-EXT-NACL mode

  `{deny | permit} tcp {source mask] | any | host ip-address}} [count [byte]] [order] [monitor [session-id]] [fragments]`
- Configure a deny or permit filter to examine UDP packets.

  CONFIG-EXT-NACL mode

  `{deny | permit} udp {source mask | any | host ip-address}} [count [byte]] [order] [monitor [session-id]] [fragments]`

When you use the `log` keyword, the CP logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details. The following example shows an extended IP ACL in which the sequence numbers were assigned by the software. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The `show config` command in IP ACCESS LIST mode displays the two filters with the sequence numbers 5 and 10.

```
DellEMC(config-ext-nacl)#deny tcp host 123.55.34.0 any
DellEMC(config-ext-nacl)#permit udp 154.44.123.34 0.0.255.255 host 34.6.0.0 monitor 111
DellEMC(config-ext-nacl)#show config
!
ip access-list extended nimule
  seq 5 deny tcp host 123.55.34.0 any
  seq 10 permit udp 154.44.0.0 0.0.255.255 host 34.6.0.0 monitor 111
DellEMC(config-ext-nacl)#
```

To view all configured IP ACLs and the number of packets processed through the ACL, use the `show ip accounting access-list` command in EXEC Privilege mode, as shown in the first example in Configure a Standard IP ACL Filter.

Configure Layer 2 and Layer 3 ACLs

Both Layer 2 and Layer 3 ACLs may be configured on an interface in Layer 2 mode.

If both L2 and L3 ACLs are applied to an interface, the following rules apply:

- When Dell EMC Networking OS routes the packets, only the L3 ACL governs them because they are not filtered against an L2 ACL.
- When Dell EMC Networking OS switches the packets, first the L3 ACL filters them, then the L2 ACL filters them.
- When Dell EMC Networking OS switches the packets, the egress L3 ACL filters the packet.

For the following features, if you enable counters on rules that have already been configured and a new rule is either inserted or prepended, all the existing counters are reset:

- L2 ingress access list
- L3 egress access list
- L2 egress access list

If a rule is simply appended, existing counters are not affected.

Table 6. L2 and L3 Filtering on Switched Packets

L2 ACL Behavior	L3 ACL Behavior	Decision on Targeted Traffic
Deny	Deny	L3 ACL denies.
Deny	Permit	L3 ACL permits.
Permit	Deny	L3 ACL denies.
Permit	Permit	L3 ACL permits.

NOTE: If you configure an interface as a vlan-stack access port, only the L2 ACL filters the packets. The L3 ACL applied to such a port does not affect traffic. That is, existing rules for other features (such as trace-list, policy-based routing [PBR], and QoS) are applied to the permitted traffic.

For information about MAC ACLs, refer to Layer 2.

Assign an IP ACL to an Interface

To pass traffic through a configured IP ACL, assign that ACL to a physical interface, a port channel interface, or a VLAN.

The IP ACL is applied to all traffic entering a physical or port channel interface and the traffic is either forwarded or dropped depending on the criteria and actions specified in the ACL.

The same ACL may be applied to different interfaces and that changes its functionality. For example, you can take ACL "ABCD" and apply it using the `in` keyword and it becomes an ingress access list. If you apply the same ACL using the `out` keyword, it becomes an egress access list. If you apply the same ACL to the Loopback interface, it becomes a Loopback access list.

Applying an IP ACL

To apply an IP ACL (standard or extended) to a physical or port channel interface, use the following commands.

1. Enter the interface number.

   CONFIGURATION mode

   `interface interface slot/port`
2. Configure an IP address for the interface, placing it in Layer-3 mode.

   INTERFACE mode

   `ip address ip-address`
3. Apply an IP ACL to traffic entering or exiting an interface.

   INTERFACE mode

   `ip access-group access-list-name {in} [implicit-permit] [vlan vlan-range | vrf vrf-range] [layer3]`

   NOTE:
   - The number of entries allowed per ACL is hardware-dependent. For detailed specification about entries allowed per ACL, refer to your line card documentation.
   - One of the usage scenarios is to avoid ACL being applied on the L2 traffic which comes in via ICL. The layer 3 keyword can be used at the VLAN level.
4. Apply rules to the new ACL.

   INTERFACE mode

   `ip access-list [standard | extended] name`

To view which IP ACL is applied to an interface, use the `show config` command in INTERFACE mode, or use the `show running-config` command in EXEC mode.

```
DellEMC(conf-if)#show conf
!
interface TenGigabitEthernet 1/1
  ip address 10.2.1.100 255.255.255.0
  ip access-group nimule in
  no shutdown
DellEMC(conf-if)#
```

To filter traffic on Telnet sessions, use only standard ACLs in the `access-class` command.

Counting ACL Hits

You can view the number of packets matching the ACL by using the `count` option when creating ACL entries.

1. Create an ACL that uses rules with the count option. Refer to Configure a Standard IP ACL Filter.
2. Apply the ACL as an inbound or outbound ACL on an interface.
3. `show ip accounting access-list`

   EXEC Privilege mode

   View the number of packets matching the ACL.

Configure Ingress ACLs

Ingress ACLs are applied to interfaces and to traffic entering the system.

These system-wide ACLs eliminate the need to apply ACLs onto each interface and achieves the same results. By localizing target traffic, it is a simpler implementation.

To create an ingress ACL, use the `ip access-group` command in EXEC Privilege mode. The example shows applying the ACL, rules to the newly created access group, and viewing the access list.

Example of Applying ACL Rules to Ingress Traffic and Viewing ACL Configuration

To specify ingress, use the `in` keyword. Begin applying rules to the ACL with the `ip access-list extended abcd` command. To view the access-list, use the `show` command.

```
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te1/1)#ip access-group abcd in
DellEMC(conf-if-te1/1)#show config
!
tengigabitethernet 1/1
  no ip address
  ip access-group abcd in
  no shutdown
DellEMC(conf-if-te1/1)#end
DellEMC#configure terminal
DellEMC(conf)#ip access-list extended abcd
DellEMC(config-ext-nacl)#permit tcp any any
DellEMC(config-ext-nacl)#deny icmp any any
DellEMC(config-ext-nacl)#permit 1.1.1.2
DellEMC(config-ext-nacl)#end
DellEMC#show ip accounting access-list
!
Extended Ingress IP access list abcd on tengigabitethernet 1/1
  seq 5 permit tcp any any
  seq 10 deny icmp any any
seq 15 permit 1.1.1.2
```

Configure Egress ACLs

Egress ACLs are applied to line cards and affect the traffic leaving the system. Configuring egress ACLs onto physical interfaces protects the system infrastructure from attack — malicious and incidental — by explicitly allowing only authorized traffic. These system-wide ACLs eliminate the need to apply ACLs onto each interface and achieves the same results. By localizing target traffic, it is a simpler implementation.

To restrict egress traffic, use an egress ACL. For example, when a denial of service (DOS) attack traffic is isolated to a specific interface, you can apply an egress ACL to block the flow from the exiting the box, thus protecting downstream devices.

To create an egress ACL, use the `ip access-group` command in EXEC Privilege mode. The example shows viewing the configuration, applying rules to the newly created access group, and viewing the access list.

NOTE: VRF based ACL configurations are not supported on the egress traffic.

Example of Applying ACL Rules to Egress Traffic and Viewing ACL Configuration

To specify ingress, use the `out` keyword. Begin applying rules to the ACL with the `ip access-list extended abcd` command. To view the access-list, use the `show` command.

```
DellEMC(conf)#interface TenGigabitEthernet 1/1
DellEMC(conf-if-te-1/1)#ip access-group abcd out
DellEMC(conf-if-te-1/1)#show config
!
TenGigabitEthernet 1/1
  no ip address
  ip access-group abcd out
  no shutdown
DellEMC(conf-if-te-1/1)#end
DellEMC#configure terminal
DellEMC(conf)#ip access-list extended abcd
```

```
DellEMC(config-ext-nacl)#permit tcp any any
DellEMC(config-ext-nacl)#deny icmp any any
DellEMC(config-ext-nacl)#permit 1.1.1.2
DellEMC(config-ext-nacl)#end
DellEMC#show ip accounting access-list
!
Extended Ingress IP access list abcd on tengigabitethernet 0/0
  seq 5 permit tcp any any
  seq 10 deny icmp any any
seq 15 permit 1.1.1.2

DellEMC#configure terminal
DellEMC(conf)#interface te 1/2
DellEMC(conf-if-te-1/2)#ip vrf forwarding blue
DellEMC(conf-if-te-1/2)#show config
!
interface TenGigabitEthernet 1/2
ip vrf forwarding blue
no ip address
shutdown
DellEMC(conf-if-te-1/2)#
DellEMC(conf-if-te-1/2)#
DellEMC(conf-if-te-1/2)#end
DellEMC#
```

Applying Egress Layer 3 ACLs (Control-Plane)

By default, packets originated from the system are not filtered by egress ACLs.

For example, if you initiate a ping session from the system and apply an egress ACL to block this type of traffic on the interface, the ACL does not affect that ping traffic. The Control Plane Egress Layer 3 ACL feature enhances IP reachability debugging by implementing control-plane ACLs for CPU-generated and CPU-forwarded traffic. Using permit rules with the `count` option, you can track on a per-flow basis whether CPU-generated and CPU-forwarded packets were transmitted successfully.

NOTE: The `ip control-plane [egress filter]` and the `ipv6 control-plane [egress filter]` commands are not supported.

1. Apply Egress ACLs to IPv4 system traffic.

   CONFIGURATION mode

   `ip control-plane [egress filter]`

2. Apply Egress ACLs to IPv6 system traffic.

   CONFIGURATION mode

   `ipv6 control-plane [egress filter]`

3. Create a Layer 3 ACL using permit rules with the `count` option to describe the desired CPU traffic.

   CONFIG-NACL mode

   `permit ip {`*`source mask`*` | any | host `*`ip-address`*`} {`*`destination mask`*` | any | host `*`ip-address`*`} count [monitor [`*`session-id`*`]]`

Dell EMC Networking OS Behavior: Virtual router redundancy protocol (VRRP) hellos and internet group management protocol (IGMP) packets are not affected when you enable egress ACL filtering for CPU traffic. Packets sent by the CPU with the source address as the VRRP virtual IP address have the interface MAC address instead of VRRP virtual MAC address.

IP Prefix Lists

IP prefix lists control routing policy. An IP prefix list is a series of sequential filters that contain a matching criterion (examine IP route prefix) and an action (permit or deny) to process routes. The filters are processed in sequence so that if a route prefix does not match the criterion in the first filter, the second filter (if configured) is applied. When the route prefix matches a filter, Dell EMC Networking OS drops or forwards the packet based on the filter's designated action. If the route prefix does not match any of the filters in the prefix list, the route is dropped (that is, implicit deny).

A route prefix is an IP address pattern that matches on bits within the IP address. The format of a route prefix is A.B.C.D/X where A.B.C.D is a dotted-decimal address and /X is the number of bits that should be matched of the dotted decimal

address. For example, in 112.24.0.0/16, the first 16 bits of the address 112.24.0.0 match all addresses between 112.24.0.0 to 112.24.255.255.

The following examples show permit or deny filters for specific routes using the `le` and `ge` parameters, where x.x.x.x/x represents a route prefix:

- To deny only /8 prefixes, enter `deny x.x.x.x/x ge 8 le 8`.
- To permit routes with the mask greater than /8 but less than /12, enter `permit x.x.x.x/x ge 8`.
- To deny routes with a mask less than /24, enter `deny x.x.x.x/x le 24`.
- To permit routes with a mask greater than /20, enter `permit x.x.x.x/x ge 20`.

The following rules apply to prefix lists:

- A prefix list without any permit or deny filters allows all routes.
- An "implicit deny" is assumed (that is, the route is dropped) for all route prefixes that do not match a permit or deny filter in a configured prefix list.
- After a route matches a filter, the filter's action is applied. No additional filters are applied to the route.

Implementation Information

In Dell EMC Networking OS, prefix lists are used in processing routes for routing protocols (for example, router information protocol [RIP], open shortest path first [OSPF], and border gateway protocol [BGP]).

NOTE: It is important to know which protocol your system supports prior to implementing prefix-lists.

Configuration Task List for Prefix Lists

To configure a prefix list, use commands in PREFIX LIST, ROUTER RIP, ROUTER OSPF, and ROUTER BGP modes.

Create the prefix list in PREFIX LIST mode and assign that list to commands in ROUTER RIP, ROUTER OSPF and ROUTER BGP modes.

The following list includes the configuration tasks for prefix lists, as described in the following sections.

- Configuring a prefix list
- Use a prefix list for route redistribution

For a complete listing of all commands related to prefix lists, refer to the *Dell EMC Networking OS Command Line Interface Reference Guide*.

Creating a Prefix List

To create a prefix list, use the following commands.

1. Create a prefix list and assign it a unique name.

   You are in PREFIX LIST mode.

   CONFIGURATION mode

   `ip prefix-list prefix-name`

2. Create a prefix list with a sequence number and a deny or permit action.

   CONFIG-NPREFIXL mode

   `seq sequence-number {deny | permit} ip-prefix [ge min-prefix-length] [le max-prefix-length]`

   The optional parameters are:

   - `ge min-prefix-length`: the minimum prefix length to match (from 0 to 32).
   - `le max-prefix-length`: the maximum prefix length to match (from 0 to 32).

If you want to forward all routes that do not match the prefix list criteria, configure a prefix list filter to permit all routes (`permit 0.0.0.0/0 le 32`). The "permit all" filter must be the last filter in your prefix list. To permit the default route only, enter `permit 0.0.0.0/0`.

The following example shows how the `seq` command orders the filters according to the sequence number assigned. In the example, filter 20 was configured before filter 15 and 12, but the `show config` command displays the filters in the correct order.

```
DellEMC(conf-nprefixl)#seq 20 permit 0.0.0.0/0 le 32
DellEMC(conf-nprefixl)#seq 12 deny 134.23.0.0 /16
DellEMC(conf-nprefixl)#seq 15 deny 120.23.14.0 /8 le 16
DellEMC(conf-nprefixl)#show config
!
ip prefix-list juba
  seq 12 deny 134.23.0.0/16
  seq 15 deny 120.0.0.0/8 le 16
  seq 20 permit 0.0.0.0/0 le 32
DellEMC(conf-nprefixl)#
```

NOTE: The last line in the prefix list Juba contains a "permit all" statement. By including this line in a prefix list, you specify that all routes not matching any criteria in the prefix list are forwarded.

To delete a filter, use the `no seq` *`sequence-number`* command in PREFIX LIST mode.If you are creating a standard prefix list with only one or two filters, you can let Dell EMC Networking OS assign a sequence number based on the order in which the filters are configured. The Dell EMC Networking OS assigns filters in multiples of five.

Creating a Prefix List Without a Sequence Number

To create a filter without a specified sequence number, use the following commands.

1. Create a prefix list and assign it a unique name.

   CONFIGURATION mode

   `ip prefix-list` *`prefix-name`*

2. Create a prefix list filter with a deny or permit action.

   CONFIG-NPREFIXL mode

   `{deny | permit}` *`ip-prefix`* `[ge` *`min-prefix-length`*`] [le` *`max-prefix-length`*`]`

   The optional parameters are:

   - `ge` *`min-prefix-length`*: is the minimum prefix length to be matched (0 to 32).
   - `le` *`max-prefix-length`*: is the maximum prefix length to be matched (0 to 32).

The example shows a prefix list in which the sequence numbers were assigned by the software. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The `show config` command in PREFIX LIST mode displays two filters with the sequence numbers 5 and 10.

```
DellEMC(conf-nprefixl)#permit 123.23.0.0 /16
DellEMC(conf-nprefixl)#deny 133.24.56.0 /8
DellEMC(conf-nprefixl)#show conf
!
ip prefix-list awe
  seq 5 permit 123.23.0.0/16
  seq 10 deny 133.0.0.0/8
DellEMC(conf-nprefixl)#
```

To delete a filter, enter the `show config` command in PREFIX LIST mode and locate the sequence number of the filter you want to delete, then use the `no seq sequence-number` command in PREFIX LIST mode.

Viewing Prefix Lists

To view all configured prefix lists, use the following commands.

- Show detailed information about configured prefix lists.

  EXEC Privilege mode

  `show ip prefix-list detail [`*`prefix-name`*`]`

- Show a table of summarized information about configured Prefix lists.

  EXEC Privilege mode

```
show ip prefix-list summary [prefix-name]
```

The following example shows the `show ip prefix-list detail` command.

```
DellEMC>show ip prefix detail
Prefix-list with the last deletion/insertion: filter_ospf
ip prefix-list filter_in:
count: 3, range entries: 3, sequences: 5 - 10
   seq 5 deny 1.102.0.0/16 le 32 (hit count: 0)
   seq 6 deny 2.1.0.0/16 ge 23 (hit count: 0)
   seq 10 permit 0.0.0.0/0 le 32 (hit count: 0)
ip prefix-list filter_ospf:
count: 4, range entries: 1, sequences: 5 - 10
   seq 5 deny 100.100.1.0/24 (hit count: 0)
   seq 6 deny 200.200.1.0/24 (hit count: 0)
   seq 7 deny 200.200.2.0/24 (hit count: 0)
   seq 10 permit 0.0.0.0/0 le 32 (hit count: 0)
```

The following example shows the `show ip prefix-list summary` command.

```
DellEMC>
DellEMC>show ip prefix summary
Prefix-list with the last deletion/insertion: filter_ospf
ip prefix-list filter_in:
count: 3, range entries: 3, sequences: 5 - 10
ip prefix-list filter_ospf:
count: 4, range entries: 1, sequences: 5 - 10
DellEMC>
```

Applying a Prefix List for Route Redistribution

To pass traffic through a configured prefix list, use the prefix list in a `route redistribution` command.

Apply the prefix list to all traffic redistributed into the routing process. The traffic is either forwarded or dropped, depending on the criteria and actions specified in the prefix list.

To apply a filter to routes in RIP, use the following commands.

- Enter RIP mode.

  CONFIGURATION mode

  `router rip`
- Apply a configured prefix list to incoming routes. You can specify an interface.

  If you enter the name of a nonexistent prefix list, all routes are forwarded.

  CONFIG-ROUTER-RIP mode

  `distribute-list prefix-list-name in [interface]`
- Apply a configured prefix list to outgoing routes. You can specify an interface or type of route.

  If you enter the name of a non-existent prefix list, all routes are forwarded.

  CONFIG-ROUTER-RIP mode

  `distribute-list prefix-list-name out [interface | connected | static | ospf]`

To view the configuration, use the `show config` command in ROUTER RIP mode, or the `show running-config rip` command in EXEC mode.

```
DellEMC(conf-router_rip)#show config
!
router rip
  distribute-list prefix juba out
  network 10.0.0.0
DellEMC(conf-router_rip)#router ospf 34
```

Applying a Filter to a Prefix List (OSPF)

To apply a filter to routes in open shortest path first (OSPF), use the following commands.

- Enter OSPF mode.
  CONFIGURATION mode
  `router ospf`
- Apply a configured prefix list to incoming routes. You can specify an interface.
  If you enter the name of a non-existent prefix list, all routes are forwarded.
  CONFIG-ROUTER-OSPF mode
  `distribute-list prefix-list-name in [interface]`
- Apply a configured prefix list to incoming routes. You can specify which type of routes are affected.
  If you enter the name of a non-existent prefix list, all routes are forwarded.
  CONFIG-ROUTER-OSPF mode
  `distribute-list prefix-list-name out [connected | rip | static]`

To view the configuration, use the `show config` command in ROUTER OSPF mode, or the `show running-config ospf` command in EXEC mode.

```
DellEMC(conf-router_ospf)#show config
!
router ospf 34
  network 10.2.1.1 255.255.255.255 area 0.0.0.1
  distribute-list prefix awe in
DellEMC(conf-router_ospf)#
```

ACL Remarks

While defining ACL rules, you can optionally include a remark to make the ACLs more descriptive. You can include a remark with a maximum of 80 characters in length.

The `remark` command is available in each ACL mode. You can configure up to 4294967291 remarks for a given IP ACL and 65536 remarks for a given MAC ACL.

You can include a remark with or without a remark number. If you do not enter a remark number, the remark inherits the sequence number of the last ACL rule. If there is no ACL rule when you enter a remark, the remark takes sequence number 5. If you configure two remarks with the same sequence number and different strings, the second one replaces the first string. You cannot configure two or more remarks with the same string and different sequence numbers.

To remove a remark, use the `no remark` command with the remark string and with or without the sequence number. If there is a matching string, the system deletes the remark.

Configuring a Remark

To write a remark for an ACL, follow these steps:

1. Create either an extended IPv4 or IPv6 ACL.
   CONFIGURATION mode
   `ip access-list {extended | standard} access-list-name`
   `ipv6 access-list {extended | standard} access-list-name`
2. Define the ACL rule.
   CONFIG-EXT-NACL mode or CONFIG-STD-NACL
   `seq sequence-number {permit | deny} options`
3. Write a remark.
   CONFIG-EXT-NACL mode or CONFIG-STD-NACL
   `remark [remark-number] remark-text`
   The remark number is optional.

The following example shows how to write a remark for an ACL rule:

```
Dell(config-ext-nacl)#ip access-list extended test
Dell(config-ext-nacl)# remark permit any ip
Dell(config-ext-nacl)# seq 10 permit ip any any
Dell(config-ext-nacl)#sh config
!
ip access-list extended test
 remark 10 permit any ip
 seq 10 permit ip any any
```

Deleting a Remark

To delete a remark, follow this procedure:

A standard IP ACL uses the source IP address as its match criterion.

- Use the `no` form of the following command:

  CONFIG-EXT-NACL mode or CONFIG-STD-NACL

  `no remark [remark-number] [remark-text]`

  The remark number is optional.

The following is an example of removing a remark.

```
Dell(config)#ip access-list extended test
Dell(config-ext-nacl)# remark 10 permit any ip
Dell(config-ext-nacl)# seq 10 permit ip any any
Dell(config-ext-nacl)#sh config
!
ip access-list extended test
 remark 10 permit any ip
 seq 10 permit ip any any
Dell(config-ext-nacl)#no remark 10
Dell(config-ext-nacl)#show config
!
ip access-list extended test
 seq 10 permit ip any any
```

ACL Resequencing

ACL resequencing allows you to re-number the rules and remarks in an access or prefix list.

The placement of rules within the list is critical because packets are matched against rules in sequential order. To order new rules using the current numbering scheme, use resequencing whenever there is no opportunity.

For example, the following table contains some rules that are numbered in increments of 1. You cannot place new rules between these packets, so apply resequencing to create numbering space, as shown in the second table. In the same example, apply resequencing if more than two rules must be placed between rules 7 and 10.

You can resequence IPv4 and IPv6 ACLs, prefixes, and MAC ACLs. No CAM writes happen as a result of resequencing, so there is no packet loss; the behavior is similar Hot-lock ACLs.

NOTE: ACL resequencing does not affect the rules, remarks, or order in which they are applied. Resequencing merely renumbers the rules so that you can place new rules within the list as needed.

Table 7. ACL Resequencing

Rules	Resequencing
Rules Before Resequencing:	seq 5 permit any host 1.1.1.1
	seq 6 permit any host 1.1.1.2
	seq 7 permit any host 1.1.1.3
	seq 10 permit any host 1.1.1.4

Table 7. ACL Resequencing (continued)

Rules	Resequencing
Rules After Resequencing:	seq 5 permit any host 1.1.1.1
	seq 10 permit any host 1.1.1.2
	seq 15 permit any host 1.1.1.3
	seq 20 permit any host 1.1.1.4

Resequencing an ACL or Prefix List

Resequencing is available for IPv4 and IPv6 ACLs, prefix lists, and MAC ACLs.

To resequence an ACL or prefix list, use the following commands. You must specify the list name, starting number, and increment when using these commands.

- IPv4, IPv6, or MAC ACL

  EXEC mode

  `resequence access-list {ipv4 | ipv6 | mac} {`*`access-list-name StartingSeqNum Step-to-Increment`*`}`

- IPv4 or IPv6 prefix-list

  EXEC mode

  `resequence prefix-list {ipv4 | ipv6} {`*`prefix-list-name StartingSeqNum Step-to-Increment`*`}`

Remarks and rules that originally have the same sequence number have the same sequence number after you apply the `resequence` command.

The example shows the resequencing of an IPv4 access-list beginning with the number 2 and incrementing by 2.

```
DellEMC(config-ext-nacl)# show config
!
ip access-list extended test
remark 4 XYZ
remark 5 this remark corresponds to permit any host 1.1.1.1
seq 5 permit ip any host 1.1.1.1
remark 9 ABC
remark 10 this remark corresponds to permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.2
seq 15 permit ip any host 1.1.1.3
seq 20 permit ip any host 1.1.1.4
DellEMC# end
DellEMC# resequence access-list ipv4 test 2 2
DellEMC# show running-config acl
!
ip access-list extended test
remark 2 XYZ
remark 4 this remark corresponds to permit any host 1.1.1.1
seq 4 permit ip any host 1.1.1.1
remark 6 this remark has no corresponding rule
remark 8 this remark corresponds to permit ip any host 1.1.1.2
seq 8 permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.3
seq 12 permit ip any host 1.1.1.4
```

Remarks that do not have a corresponding rule are incremented as a rule. These two mechanisms allow remarks to retain their original position in the list. The following example shows remark 10 corresponding to rule 10 and as such, they have the same number before and after the command is entered. Remark 4 is incremented as a rule, and all rules have retained their original positions.

```
DellEMC(config-ext-nacl)# show config
!
ip access-list extended test
remark 4 XYZ
remark 5 this remark corresponds to permit any host 1.1.1.1
seq 5 permit ip any host 1.1.1.1
```

```
remark 9 ABC
remark 10 this remark corresponds to permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.2
seq 15 permit ip any host 1.1.1.3
seq 20 permit ip any host 1.1.1.4
DellEMC# end
DellEMC# resequence access-list ipv4 test 2 2
DellEMC# show running-config acl
!
ip access-list extended test
remark 2 XYZ
remark 4 this remark corresponds to permit any host 1.1.1.1
seq 4 permit ip any host 1.1.1.1
remark 6 this remark has no corresponding rule
remark 8 this remark corresponds to permit ip any host 1.1.1.2
seq 8 permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.3
seq 12 permit ip any host 1.1.1.4
```

Route Maps

Although route maps are similar to ACLs and prefix lists in that they consist of a series of commands that contain a matching criterion and an action, route maps can modify parameters in matching packets.

Implementation Information

ACLs and prefix lists can only drop or forward the packet or traffic. Route maps process routes for route redistribution. For example, a route map can be called to filter only specific routes and to add a metric.

Route maps also have an “implicit deny.” Unlike ACLs and prefix lists; however, where the packet or traffic is dropped, in route maps, if a route does not match any of the route map conditions, the route is not redistributed.

The implementation of route maps allows route maps with the no match or no set commands. When there is no match command, all traffic matches the route map and the set command applies.

Important Points to Remember

- For route-maps with more than one match clause:
  - Two or more match clauses within the same route-map sequence have the *same* match commands (though the values are different), matching a packet against these clauses is a logical OR operation.
  - Two or more match clauses within the same route-map sequence have *different* match commands, matching a packet against these clauses is a logical AND operation.
- If no match is found in a route-map sequence, the process moves to the next route-map sequence until a match is found, or there are no more sequences.
- When a match is found, the packet is forwarded and no more route-map sequences are processed.
  - If a continue clause is included in the route-map sequence, the next or a specified route-map sequence is processed after a match is found.

Configuration Task List for Route Maps

Configure route maps in ROUTE-MAP mode and apply the maps in various commands in ROUTER RIP and ROUTER OSPF modes.

The following list includes the configuration tasks for route maps, as described in the following sections.

- Create a route map (mandatory)
- Configure route map filters (optional)
- Configure a route map for route redistribution (optional)
- Configure a route map for route tagging (optional)

Creating a Route Map

Route maps, ACLs, and prefix lists are similar in composition because all three contain filters, but route map filters do not contain the permit and deny actions found in ACLs and prefix lists.

Route map filters match certain routes and set or specific values.

To create a route map, use the following command.

- Create a route map and assign it a unique name. The optional `permit` and `deny` keywords are the actions of the route map.

  CONFIGURATION mode

  `route-map` *`map-name`* `[permit | deny]` [*`sequence-number`*]

  The default is **permit**.

  The optional `seq` keyword allows you to assign a sequence number to the route map instance.

The default action is **permit** and the default sequence number starts at **10**. When you use the keyword `deny` in configuring a route map, routes that meet the match filters are not redistributed.

To view the configuration, use the `show config` command in ROUTE-MAP mode.

```
DellEMC(config-route-map)#show config
!
route-map dilling permit 10
DellEMC(config-route-map)#
```

You can create multiple instances of this route map by using the `sequence number` option to place the route maps in the correct order. Dell EMC Networking OS processes the route maps with the lowest sequence number first. When a configured route map is applied to a command, such as `redistribute`, traffic passes through all instances of that route map until a match is found. The following is an example with two instances of a route map.

The following example shows matching instances of a route-map.

```
DellEMC#show route-map
route-map zakho, permit, sequence 10
  Match clauses:
  Set clauses:
route-map zakho, permit, sequence 20
  Match clauses:
    interface TenGigabitEthernet 1/1
  Set clauses:
    tag 35
    level stub-area
DellEMC#
```

To delete all instances of that route map, use the `no route-map` *`map-name`* command. To delete just one instance, add the sequence number to the command syntax.

```
DellEMC(conf)#no route-map zakho 10
DellEMC(conf)#end
DellEMC#show route-map
route-map zakho, permit, sequence 20
  Match clauses:
    interface TenGigabitEthernet 1/1
  Set clauses:
    tag 35
    level stub-area
DellEMC#
```

The following example shows a route map with multiple instances. The `show config` command displays only the configuration of the current route map instance. To view all instances of a specific route map, use the `show route-map` command.

```
DellEMC#show route-map dilling
route-map dilling, permit, sequence 10
  Match clauses:
  Set clauses:
route-map dilling, permit, sequence 15
  Match clauses:
    interface Loopback 23
```

```
   Set clauses:
     tag 3444
DellEMC#
```

To delete a route map, use the `no route-map map-name` command in CONFIGURATION mode.

Configure Route Map Filters

Within ROUTE-MAP mode, there are `match` and `set` commands.

- `match` commands search for a certain criterion in the routes.
- `set` commands change the characteristics of routes, either adding something or specifying a level.

When there are multiple `match` commands with the same parameter under one instance of route-map, Dell EMC Networking OS does a match between all of those `match` commands. If there are multiple `match` commands with different parameters, Dell EMC Networking OS does a match ONLY if there is a match among ALL the `match` commands. In the following example, there is a match if a route has any of the tag values specified in the `match` commands.

Example of the `match` Command to Match Any of Several Values

The following example shows using the `match` command to match any of several values.

```
DellEMC(conf)#route-map force permit 10
DellEMC(config-route-map)#match tag 1000
DellEMC(config-route-map)#match tag 2000
DellEMC(config-route-map)#match tag 3000
```

In the next example, there is a match *only* if a route has *both* of the specified characteristics. In this example, there a match only if the route has a tag value of 1000 *and* a metric value of 2000.

Also, if there are different instances of the same route-map, then it's sufficient if a permit match happens in any instance of that route-map.

Example of the `match` Command to Match All Specified Values

```
DellEMC(conf)#route-map force permit 10
DellEMC(config-route-map)#match tag 1000
DellEMC(config-route-map)#match metric 2000
```

In the following example, instance 10 permits the route having a tag value of 1000 and instances 20 and 30 deny the route having a tag value of 1000. In this scenario, Dell EMC Networking OS scans all the instances of the route-map for any permit statement. If there is a match anywhere, the route is permitted. However, other instances of the route-map deny it.

Example of the `match` Command to Permit and Deny Routes

```
DellEMC(conf)#route-map force permit 10
DellEMC(config-route-map)#match tag 1000

DellEMC(conf)#route-map force deny 20
DellEMC(config-route-map)#match tag 1000

DellEMC(conf)#route-map force deny 30
DellEMC(config-route-map)#match tag 1000
```

Configuring Match Routes

To configure match criterion for a route map, use the following commands.

- Match routes with the same AS-PATH numbers.

  CONFIG-ROUTE-MAP mode

  `match as-path as-path-name`
- Match routes with COMMUNITY list attributes in their path.

  CONFIG-ROUTE-MAP mode

  `match community community-list-name [exact]`
- Match routes whose next hop is a specific interface.

CONFIG-ROUTE-MAP mode

`match interface interface`

The parameters are:

- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
- For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
- For a port channel interface, enter the keywords `port-channel` then a number.
- For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

- Match destination routes specified in a prefix list (IPv4).
  CONFIG-ROUTE-MAP mode
  `match ip address prefix-list-name`
- Match destination routes specified in a prefix list (IPv6).
  CONFIG-ROUTE-MAP mode
  `match ipv6 address prefix-list-name`
- Match next-hop routes specified in a prefix list (IPv4).
  CONFIG-ROUTE-MAP mode
  `match ip next-hop {access-list-name | prefix-list prefix-list-name}`
- Match next-hop routes specified in a prefix list (IPv6).
  CONFIG-ROUTE-MAP mode
  `match ipv6 next-hop {access-list-name | prefix-list prefix-list-name}`
- Match source routes specified in a prefix list (IPv4).
  CONFIG-ROUTE-MAP mode
  `match ip route-source {access-list-name | prefix-list prefix-list-name}`
- Match source routes specified in a prefix list (IPv6).
  CONFIG-ROUTE-MAP mode
  `match ipv6 route-source {access-list-name | prefix-list prefix-list-name}`
- Match routes with a specific value.
  CONFIG-ROUTE-MAP mode
  `match metric metric-value`
- Match BGP routes based on the ORIGIN attribute.
  CONFIG-ROUTE-MAP mode
  `match origin {egp | igp | incomplete}`
- Match routes specified as internal or external to OSPF, ISIS level-1, ISIS level-2, or locally generated.
  CONFIG-ROUTE-MAP mode
  `match route-type {external [type-1 | type-2] | internal | level-1 | level-2 | local }`
- Match routes with a specific tag.
  CONFIG-ROUTE-MAP mode
  `match tag tag-value`

To create route map instances, use these commands. There is no limit to the number of `match` commands per route map, but the convention is to keep the number of match filters in a route map low. `Set` commands do not require a corresponding `match` command.

Configuring Set Conditions

To configure a set condition, use the following commands.

- Add an AS-PATH number to the beginning of the AS-PATH.
  CONFIG-ROUTE-MAP mode
  `set as-path prepend as-number [... as-number]`
- Generate a tag to be added to redistributed routes.

CONFIG-ROUTE-MAP mode

`set automatic-tag`

- Specify an OSPF area or ISIS level for redistributed routes.

  CONFIG-ROUTE-MAP mode

  `set level {backbone | level-1 | level-1-2 | level-2 | stub-area}`
- Specify a value for the BGP route's LOCAL_PREF attribute.

  CONFIG-ROUTE-MAP mode

  `set local-preference value`
- Specify a value for redistributed routes.

  CONFIG-ROUTE-MAP mode

  `set metric {+ | - | metric-value}`
- Specify an OSPF or ISIS type for redistributed routes.

  CONFIG-ROUTE-MAP mode

  `set metric-type {external | internal | type-1 | type-2}`
- Assign an IP address as the route's next hop.

  CONFIG-ROUTE-MAP mode

  `set next-hop ip-address`
- Assign an IPv6 address as the route's next hop.

  CONFIG-ROUTE-MAP mode

  `set ipv6 next-hop ip-address`
- Assign an ORIGIN attribute.

  CONFIG-ROUTE-MAP mode

  `set origin {egp | igp | incomplete}`
- Specify a tag for the redistributed routes.

  CONFIG-ROUTE-MAP mode

  `set tag tag-value`
- Specify a value as the route's weight.

  CONFIG-ROUTE-MAP mode

  `set weight value`

To create route map instances, use these commands. There is no limit to the number of `set` commands per route map, but the convention is to keep the number of set filters in a route map low. `Set` commands do not require a corresponding `match` command.

Configure a Route Map for Route Redistribution

Route maps on their own cannot affect traffic and must be included in different commands to affect routing traffic.

Route redistribution occurs when Dell EMC Networking OS learns the advertising routes from static or directly connected routes or another routing protocol. Different protocols assign different values to redistributed routes to identify either the routes and their origins. The metric value is the most common attribute that is changed to properly redistribute other routes into a routing protocol. Other attributes that can be changed include the metric type (for example, external and internal route types in OSPF) and route tag. Use the `redistribute` command in OSPF, RIP, ISIS, and BGP to set some of these attributes for routes that are redistributed into those protocols.

Route maps add to that redistribution capability by allowing you to match specific routes and set or change more attributes when redistributing those routes.

In the following example, the `redistribute` command calls the route map `static ospf` to redistribute only certain static routes into OSPF. According to the route map `static ospf`, only routes that have a next hop of interface 1/1 and that have a metric of 255 are redistributed into the OSPF backbone area.

NOTE: When re-distributing routes using route-maps, you must create the route-map defined in the `redistribute` command under the routing protocol. If you do not create a route-map, NO routes are redistributed.

Example of Calling a Route Map to Redistribute Specified Routes

```
router ospf 34
  default-information originate metric-type 1
  redistribute static metric 20 metric-type 2 tag 0 route-map staticospf
!
route-map staticospf permit 10
  match interface TenGigabitEthernet 1/1
  match metric 255
  set level backbone
```

Configure a Route Map for Route Tagging

One method for identifying routes from different routing protocols is to assign a tag to routes from that protocol.

As the route enters a different routing domain, it is tagged. The tag is passed along with the route as it passes through different routing protocols. You can use this tag when the route leaves a routing domain to redistribute those routes again. In the following example, the `redistribute ospf` command with a route map is used in ROUTER RIP mode to apply a tag of 34 to all internal OSPF routes that are redistributed into RIP.

Example of the `redistribute` Command Using a Route Tag

```
!
router rip
  redistribute ospf 34 metric 1 route-map torip
!
route-map torip permit 10
  match route-type internal
  set tag 34
!
```

Continue Clause

Normally, when a match is found, set clauses are executed, and the packet is then forwarded; no more route-map modules are processed.

If you configure the `continue` command at the end of a module, the next module (or a specified module) is processed even after a match is found. The following example shows a continue clause at the end of a route-map module. In this example, if a match is found in the route-map "test" module 10, module 30 is processed.

NOTE: If you configure the continue clause without specifying a module, the next sequential module is processed.

Example of Using the `continue` Clause in a Route Map

```
!
route-map test permit 10
match commu comm-list1
set community 1:1 1:2 1:3
set as-path prepend 1 2 3 4 5
continue 30!
```

Logging of ACL Processes

This functionality is supported on the S4820T platform.

To assist in the administration and management of traffic that traverses the device after being validated by the configured ACLs, you can enable the generation of logs for access control list (ACL) processes. Although you can configure ACLs with the required permit or deny filters to provide access to the incoming packet or disallow access to a particular user, it is also necessary to monitor and examine the traffic that passes through the device. To evaluate network traffic that is subjected to ACLs, configure the logs to be triggered for ACL operations. This functionality is primarily needed for network supervision and maintenance activities of the handled subscriber traffic.

When ACL logging is configured, and a frame reaches an ACL-enabled interface and matches the ACL, a log is generated to indicate that the ACL entry matched the packet.

When you enable ACL log messages, at times, depending on the volume of traffic, it is possible that a large number of logs might be generated that can impact the system performance and efficiency. To avoid an overload of ACL logs from being recorded, you can configure the rate-limiting functionality. Specify the interval or frequency at which ACL logs must be triggered and also the threshold or limit for the maximum number of logs to be generated. If you do not specify the frequency at which ACL logs must be generated, a default interval of 5 minutes is used. Similarly, if you do not specify the threshold for ACL logs, a default threshold of 10 is used, where this value refers to the number of packets that are matched against an ACL .

A Layer 2 or Layer 3 ACL contains a set of defined rules that are saved as flow processor (FP) entries. When you enable ACL logging for a particular ACL rule, a set of specific ACL rules translate to a set of FP entries. You can enable logging separately for each of these FP entries, which relate to each of the ACL entries configured in an ACL. Dell EMC Networking OS saves a table that maps each ACL entry that matches the ACL name on the received packet, sequence number of the rule, and the interface index in the database. When the configured maximum threshold has exceeded, log generation stops. When the interval at which ACL logs are configured to be recorded expires, a fresh interval timer starts and the packet count for that new interval commences from zero. If ACL logging was stopped previously because the configured threshold has exceeded, it is reenabled for this new interval.

The ACL application sends the ACL logging configuration information and other details, such as the action, sequence number, and the ACL parameters that pertain to that ACL entry. The ACL service collects the ACL log and records the following attributes per log message.

- For non-IP packets, the ACL name, sequence number, ACL action (permit or deny), source and destination MAC addresses, EtherType, and ingress interface are the logged attributes.
- For IP Packets, the ACL name, sequence number, ACL action (permit or deny), source and destination MAC addresses, source and destination IP addresses, and the transport layer protocol used are the logged attributes.
- For IP packets that contain the transport layer protocol as Transmission Control Protocol (TCP) or User Datagram Protocol (UDP), the ACL name, sequence number, ACL action (permit or deny), source and destination MAC addresses, source and destination IP addresses, and the source and destination ports (Layer 4 parameters) are also recorded.

If the packet contains an unidentified EtherType or transport layer protocol, the values for these parameters are saved as Unknown in the log message. If you also enable the logging of the count of packets in the ACL entry, and if the logging is deactivated in a specific interval because the threshold has exceeded, the count of packets that exceeded the logging threshold value during that interval is recorded when the subsequent log record (in the next interval) is generated for that ACL entry.

Guidelines for Configuring ACL Logging

This functionality is supported on the S4820T platform.

Keep the following points in mind when you configure logging of ACL activities:

- During initialization, the ACL logging application tags the ACL rule indices for which a match condition exists as being in-use, which ensures that the same rule indices are not reused by ACL logging again.
- The ACL configuration information that the ACL logging application receives from the ACL manager causes the allocation and clearance of the match rule number. A unique match rule number is created for the combination of each ACL entry, sequence number, and interface parameters.
- A separate set of match indices is preserved by the ACL logging application for the permit and deny actions. Depending on the action of an ACL entry, the corresponding match index is allocated from the particular set that is maintained for permit and deny actions.
- A maximum of 125 ACL entries with permit action can be logged. A maximum of 126 ACL entries with deny action can be logged.
- For virtual ACL entries, the same match rule number is reused. Similarly, when an ACL entry is deleted that was previously enabled for ACL logging, the match rule number used by it is released back to the pool or available set of match indices so that it can be reused for subsequent allocations.
- If you enabled the count of packets for the ACL entry for which you configured logging, and if the logging is deactivated in a specific interval owing to the threshold having exceeded, the count of packets that exceeded the logging threshold value during that interval is logged when the subsequent log record (in the next interval) is generated for that ACL entry.
- When you delete an ACL entry, the logging settings associated with it are also removed.
- ACL logging is supported for standard and extended IPv4 ACLs, IPv6 ACLs, and standard and extended MAC ACLs.

- For ACL entries applied on port-channel interfaces, one match index for every member interface of the port-channel interface is assigned. Therefore, the total available match indices of 251 are split (125 match indices for permit action and 126 match indices for the deny action).
- You can configure ACL logging only on ACLs that are applied to ingress interfaces; you cannot enable logging for ACLs on egress interfaces.
- The total available match rule indices is 255 with four match indices used by other modules, leaving 251 indices available for ACL logging.

Configuring ACL Logging

This functionality is supported on the S4820T platform.

To configure the maximum number of ACL log messages to be generated and the frequency at which these messages must be generated, perform the following steps:

NOTE: This example describes the configuration of ACL logging for standard IP access lists. You can enable the logging capability for standard and extended IPv4 ACLs, IPv6 ACLs, and standard and extended MAC ACLs.

1. Specify the maximum number of ACL logs or the threshold that can be generated by using the `threshold-in-msgs` *`count`* option with the seq, permit, or deny commands. Upon exceeding the specified maximum limit, the generation of ACL logs is terminated. You can enter a threshold in the range of 1-100. By default, 10 ACL logs are generated if you do not specify the threshold explicitly.

   CONFIG-STD-NACL mode

   `seq` *`sequence-number`* `{deny | permit} {source [mask] | any | host` *`ip-address`*`} [log [threshold-in-msgs` *`count`*`] ]`

2. Specify the interval in minutes at which ACL logs must be generated. You can enter an interval in the range of 1-10 minutes. The default frequency at which ACL logs are generated is 5 minutes. If ACL logging is stopped because the configured threshold has exceeded, it is re-enabled after the logging interval period elapses. ACL logging is supported for standard and extended IPv4 ACLs, IPv6 ACLs, and standard and extended MAC ACLs. Configure ACL logging only on ACLs that are applied to ingress interfaces; you cannot enable logging for ACLs that are associated with egress interfaces.

   CONFIG-STD-NACL mode

   `seq` *`sequence-number`* `{deny | permit} {source [mask] | any | host` *`ip-address`*`} [log [interval` *`minutes`*`]]`

Flow-Based Monitoring

Flow-based monitoring conserves bandwidth by monitoring only the specified traffic instead of all traffic on the interface. It is available for Layer 3 ingress traffic. You can specify the traffic that needs to be monitored using standard or extended access-lists. The flow-based monitoring mechanism copies packets that matches the ACL rules applied on the port and forwards (mirrors) them to another port. The source port is the monitored port (MD) and the destination port is the monitoring port (MG).

When a packet arrives at a port that is being monitored, the packet is validated against the configured ACL rules. If the packet matches an ACL rule, the system examines the corresponding flow processor to perform the action specified for that port. If the mirroring action is set in the flow processor entry, the destination port details, to which the mirrored information must be sent, are sent to the destination port.

Behavior of Flow-Based Monitoring

You can activate flow-based monitoring for a monitoring session using the `flow-based enable` command in the Monitor Session mode. When you enable this flow-based monitoring, traffic with particular flows that are traversing through the interfaces are examined in accordance with the applied ACLs. By default, flow-based monitoring is not enabled.

There are two ways in which you can enable flow-based monitoring in Dell EMC Networking OS. You can create an ACL and apply that ACL either to an interface that needs to be monitored or apply it in the monitor session context. If you apply the monitor ACL to an interface, the Dell EMC Networking OS mirrors the ingress traffic with an implicit deny applied at the end of the ACL. If you apply the ACL to the monitor section context, the Dell EMC Networking OS mirrors the ingress traffic with an implicit permit applied at the end of the ACL. This enables the other traffic to flow without being blocked by the ACL.

When you apply an ACL within the monitor session, it is applied to all source interfaces configured in the monitor session.

The Dell EMC Networking OS honors any permit or deny actions of the ACL rules used for flow-based mirroring. Packets that match a mirror ACL rule is denied or forwarded depending on the rule but the packet is mirrored. However, the user ACL has precedence over the mirror ACL.

The same source interface can be part of multiple monitor sessions.

Flow-based monitoring is supported for SPAN, RSPAN, and ERSPAN sessions. If there are overlapping rules between ACLs applied on different monitor sessions, the session with the highest monitor session ID takes precedence.

NOTE: You can apply only IPv4 ACL rules under monitor session context.

You must specify the `monitor` option with the `permit, deny`, or `seq` command for ACLs that are assigned to the source or the monitored port (MD) to enable the evaluation and replication of traffic that is traversing to the destination port. Enter the keyword `monitor` with the `seq`, `permit`, or `deny` command for the ACL rules to allow or drop IPv4, IPv6, ARP, UDP, EtherType, ICMP, and TCP packets. The ACL rule describes the traffic that you want to monitor, and the ACL in which you are creating the rule is applied to the monitored interface. Flow monitoring is supported for standard and extended IPv4 ACLs, standard and extended IPv6 ACLs, and standard and extended MAC ACLs.

CONFIG-STD-NACL mode

`seq` *`sequence-number`* `{deny | permit} {source [mask] | any | host` *`ip-address`*`} [count [byte]] [order] [fragments] [log [threshold-in-msgs count]] [monitor]`

If you configure the `flow-based enable` command and do not apply an ACL on the source port or the monitored port, both flow-based monitoring and port mirroring do not function.

You cannot apply the same ACL to an interface or a monitoring session context simultaneously.

The port mirroring application maintains a database that contains all monitoring sessions (including port monitor sessions). It has information regarding the sessions that are enabled for flow-based monitoring and those sessions that are not enabled for flow-based monitoring. It downloads monitoring configuration to the ACL agent whenever the ACL agent is registered with the port mirroring application or when flow-based monitoring is enabled.

The `show monitor session` *`session-id`* command displays the Type field in the output, which indicates whether a particular session is enabled for flow-monitoring.

Example Output of the `show` Command

```
DellEMC# show monitor session 1
SessID  Source          Destination          Dir  Mode  Source IP      Dest IP
DSCP  TTL    Drop  Rate    Gre-Protocol FcMonitor
------  ------          -----------          ---  ----  ---------      --------
----  ---    ----  ----    -----------  ---------
   1    Te 1/2          remote-ip            rx   Port   0.0.0.0        0.0.0.0          0
   0    No    N/A         N/A         yes
DellEMC#
```

The `show config` command has been modified to display monitoring configuration in a particular session.

Example Output of the `show` Command

The `show ip accounting` commands have been enhanced to display whether monitoring is enabled for traffic that matches with the rules of the specific ACL.

Example Output of the `show` Command

```
DellEMC# show ip accounting access-list
!
Extended Ingress IP access list kar on TenGigabitEthernet 1/1
Total cam count 1
   seq 5 permit ip 192.168.20.0/24 173.168.20.0/24 monitor

DellEMC#show ipv6 accounting access-list
    !
    Ingress IPv6 access list kar on TenGigabitEthernet 1/1
    Total cam count 1
     seq 5 permit ipv6 22::/24 33::/24 monitor
```

Enabling Flow-Based Monitoring

Flow-based monitoring is supported on the S4820T platform.

Flow-based monitoring conserves bandwidth by monitoring only specified traffic instead of all traffic on the interface. This feature is particularly useful when looking for malicious traffic. It is available for Layer 2 and Layer 3 ingress traffic. You can specify traffic using standard or extended access-lists.

1. Enable flow-based monitoring for a monitoring session.

   MONITOR SESSION mode

   `flow-based enable`

2. Define access-list rules that include the keyword `monitor`. Dell Networking OS only considers port monitoring traffic that matches rules with the keyword `monitor`.

   CONFIGURATION mode

   `ip access-list`

   For more information, see Access Control Lists (ACLs).

3. Apply the ACL to the monitored port.

   INTERFACE mode

   `ip access-group access-list`

To view an access-list that you applied to an interface, use the `show ip accounting access-list` command from EXEC Privilege mode.

```
DellEMC(conf)#monitor session 0
DellEMC(conf-mon-sess-0)#flow-based enable
DellEMC(conf)#ip access-list ext testflow
DellEMC(config-ext-nacl)#seq 5 permit icmp any any count bytes monitor
DellEMC(config-ext-nacl)#seq 10 permit ip 102.1.1.0/24 any count bytes monitor
DellEMC(config-ext-nacl)#seq 15 deny udp any any count bytes
DellEMC(config-ext-nacl)#seq 20 deny tcp any any count bytes
DellEMC(config-ext-nacl)#exit
DellEMC(conf)#interface TenGigabitEthernet 1/1
DellEMC(conf-if-te-1/1)#ip access-group testflow in
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  ip address 10.11.1.254/24
  ip access-group testflow in
  shutdown
DellEMC(conf-if-te-1/1)#exit
DellEMC(conf)#do show ip accounting access-list testflow
!
Extended Ingress IP access list testflow on TenGigabitEthernet 1/1
Total cam count 4
  seq 5 permit icmp any any monitor count bytes (0 packets 0 bytes)
  seq 10 permit ip 102.1.1.0/24 any monitor count bytes (0 packets 0 bytes)
  seq 15 deny udp any any count bytes (0 packets 0 bytes)
  seq 20 deny tcp any any count bytes (0 packets 0 bytes)
DellEMC(conf)#do show monitor session 0

DellEMC(conf-mon-sess-0)#do show monitor session 0
SessID  Source          Destination     Dir  Mode  Source IP      Dest IP         DSCP
TTL    Drop  Rate    Gre-Protocol FcMonitor
------  ------          -----------     ---  ----  ---------      --------        ----
---    ----  ----    -----------  ---------
  0     Te 1/1          Te 1/2          rx   Flow   0.0.0.0       0.0.0.0          0
0     No    N/A        N/A        yes
DellEMC#
```

8

Bidirectional Forwarding Detection (BFD)

BFD is a protocol that is used to rapidly detect communication failures between two adjacent systems. It is a simple and lightweight replacement for existing routing protocol link state detection mechanisms. It also provides a failure detection solution for links on which no routing protocol is used.

BFD is a simple hello mechanism. Two neighboring systems running BFD establish a session using a three-way handshake. After the session has been established, the systems exchange periodic control packets at sub-second intervals. If a system does not receive a hello packet within a specified amount of time, routing protocols are notified that the forwarding path is down.

BFD provides forwarding path failure detection times on the order of milliseconds rather than seconds as with conventional routing protocol hellos. It is independent of routing protocols, and as such, provides a consistent method of failure detection when used across a network. Networks converge faster because BFD triggers link state changes in the routing protocol sooner and more consistently because BFD eliminates the use of multiple protocol-dependent timers and methods.

BFD also carries less overhead than routing protocol hello mechanisms. Control packets can be encapsulated in any form that is convenient, and, on Dell EMC Networking routers, BFD agents maintain sessions that reside on the line card, which frees resources on the route processor. Only session state changes are reported to the BFD Manager (on the route processor), which in turn notifies the routing protocols that are registered with it.

BFD is an independent and generic protocol, which all media, topologies, and routing protocols can support using any encapsulation. Dell EMC Networking has implemented BFD at Layer 3 and with user datagram protocol (UDP) encapsulation. BFD is supported on static routing protocols and dynamic routing protocols such as VRRP, OSPF, OSPFv3, IS-IS, and BGP.

Topics:

- How BFD Works
- Important Points to Remember
- Configure BFD

How BFD Works

Two neighboring systems running BFD establish a session using a three-way handshake.

After the session has been established, the systems exchange control packets at agreed upon intervals. In addition, systems send a control packet anytime there is a state change or change in a session parameter. These control packets are sent without regard to transmit and receive intervals.

NOTE: The Dell EMC Networking Operating System (OS) does not support multi-hop BFD sessions.

If a system does not receive a control packet within an agreed-upon amount of time, the BFD agent changes the session state to Down. It then notifies the BFD manager of the change and sends a control packet to the neighbor that indicates the state change (though it might not be received if the link or receiving interface is faulty). The BFD manager notifies the routing protocols that are registered with it (clients) that the forwarding path is down and a link state change is triggered in all protocols.

NOTE: A session state change from Up to Down is the only state change that triggers a link state change in the routing protocol client.

BFD Packet Format

Control packets are encapsulated in user datagram protocol (UDP) packets. The following illustration shows the complete encapsulation of a BFD control packet inside an IPv4 packet.

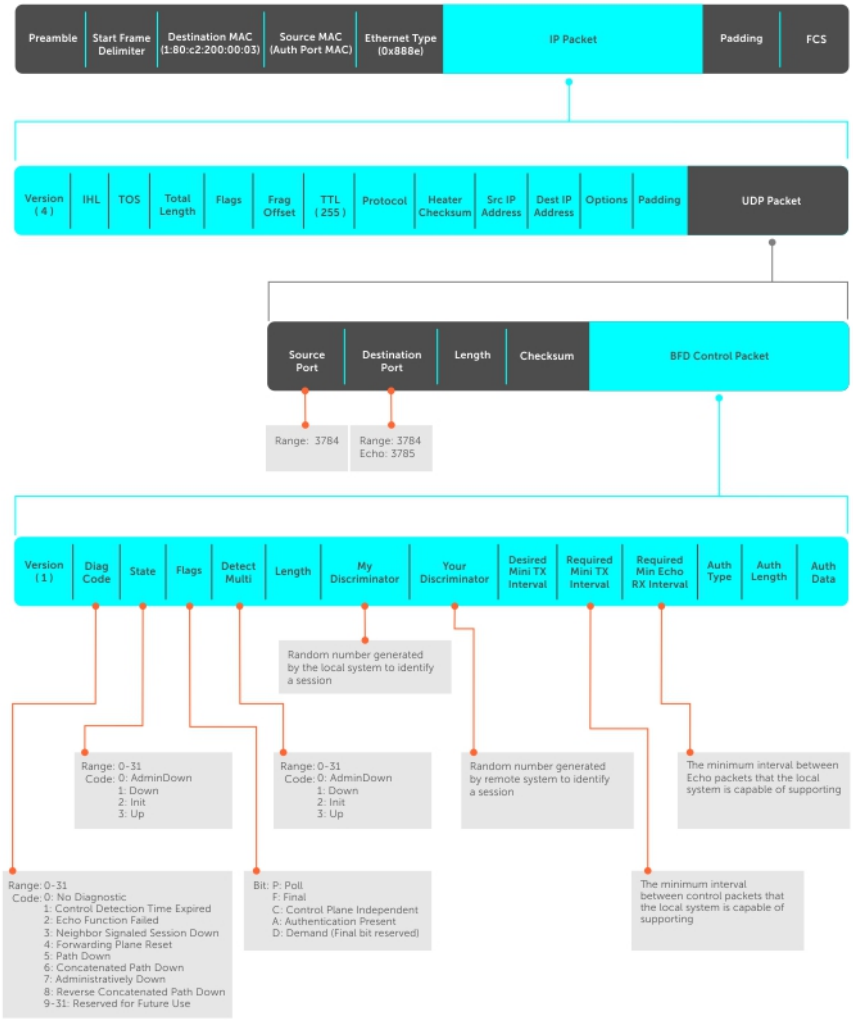


Figure 9. BFD in IPv4 Packet Format

Field	Description
Diagnostic Code	The reason that the last session failed.
State	The current local session state. Refer to BFD Sessions.
Flag	A bit that indicates packet function. If the poll bit is set, the receiving system must respond as soon as possible, without regard to its transmit interval. The responding system clears the poll bit and sets the final bit in its response. The poll and final bits are used during the handshake and in Demand mode (refer to BFD Sessions). **NOTE:** Dell EMC Networking OS does not currently support multi-point sessions, Demand mode, authentication, or control plane independence; these bits are always clear.
Detection Multiplier	The number of packets that must be missed in order to declare a session down.
Length	The entire length of the BFD packet.
My Discriminator	A random number generated by the local system to identify the session.

Field	Description
Your Discriminator	A random number generated by the remote system to identify the session. Discriminator values are necessary to identify the session to which a control packet belongs because there can be many sessions running on a single interface.
Desired Min TX Interval	The minimum rate at which the local system would like to send control packets to the remote system.
Required Min RX Interval	The minimum rate at which the local system would like to receive control packets from the remote system.
Required Min Echo RX	The minimum rate at which the local system would like to receive echo packets. **NOTE:** Dell EMC Networking OS does not currently support the echo function.
Authentication Type, Authentication Length, Authentication Data	An optional method for authenticating control packets. **NOTE:** Dell EMC Networking OS does not currently support the BFD authentication function.

Two important parameters are calculated using the values contained in the control packet.

Transmit Interval	Transmit interval is the agreed-upon rate at which a system sends control packets. Each system has its own transmit interval, which is the greater of the last received remote Desired TX Interval and the local Required Min RX Interval.
Detection time	Detection time is the amount of time that a system does not receive a control packet, after which the system determines that the session has failed. Each system has its own detection time. • In Asynchronous mode: Detection time is the remote Detection Multiplier multiplied by greater of the remote Desired TX Interval and the local Required Min RX Interval. • In Demand mode: Detection time is the local Detection Multiplier multiplied by the greater of the local Desired Min TX and the remote Required Min RX Interval.

BFD Sessions

BFD must be enabled on both sides of a link in order to establish a session.

The two participating systems can assume either of two roles:

Active	The active system initiates the BFD session. Both systems can be active for the same session.
Passive	The passive system does not initiate a session. It only responds to a request for session initialization from the active system.

A BFD session has two modes:

Asynchronous mode	In Asynchronous mode, both systems send periodic control messages at an agreed upon interval to indicate that their session status is Up.'
Demand mode	If one system requests Demand mode, the other system stops sending periodic control packets; it only sends a response to status inquiries from the Demand mode initiator. Either system (but not both) can request Demand mode at any time.

NOTE: Dell EMC Networking OS supports Asynchronous mode only.

A session can have four states: Administratively Down, Down, Init, and Up.

State	Description
Administratively Down	The local system does not participate in a particular session.
Down	The remote system is not sending control packets or at least not within the detection time for a particular session.

State	Description
Init	The local system is communicating.
Up	Both systems are exchanging control packets.

The session is declared down if:

- A control packet is not received within the detection time.
- Sufficient echo packets are lost.
- Demand mode is active and a control packet is not received in response to a poll packet.

BFD Three-Way Handshake

A three-way handshake must take place between the systems that participate in the BFD session.

The handshake shown in the following illustration assumes that there is one active and one passive system, and that this session is the first session established on this link. The default session state on both ports is Down.

1. The active system sends a steady stream of control packets that indicates that its session state is Down, until the passive system responds. These packets are sent at the desired transmit interval of the Active system. The Your Discriminator field is set to zero.
2. When the passive system receives any of these control packets, it changes its session state to Init and sends a response that indicates its state change. The response includes its session ID in the My Discriminator field and the session ID of the remote system in the Your Discriminator field.
3. The active system receives the response from the passive system and changes its session state to Up. It then sends a control packet indicating this state change. This is the third and final part of the handshake. Now the discriminator values have been exchanged and the transmit intervals have been negotiated.
4. The passive system receives the control packet and changes its state to Up. Both systems agree that a session has been established. However, because both members must send a control packet — that requires a response — anytime there is a state change or change in a session parameter, the passive system sends a final response indicating the state change. After this, periodic control packets are exchanged.

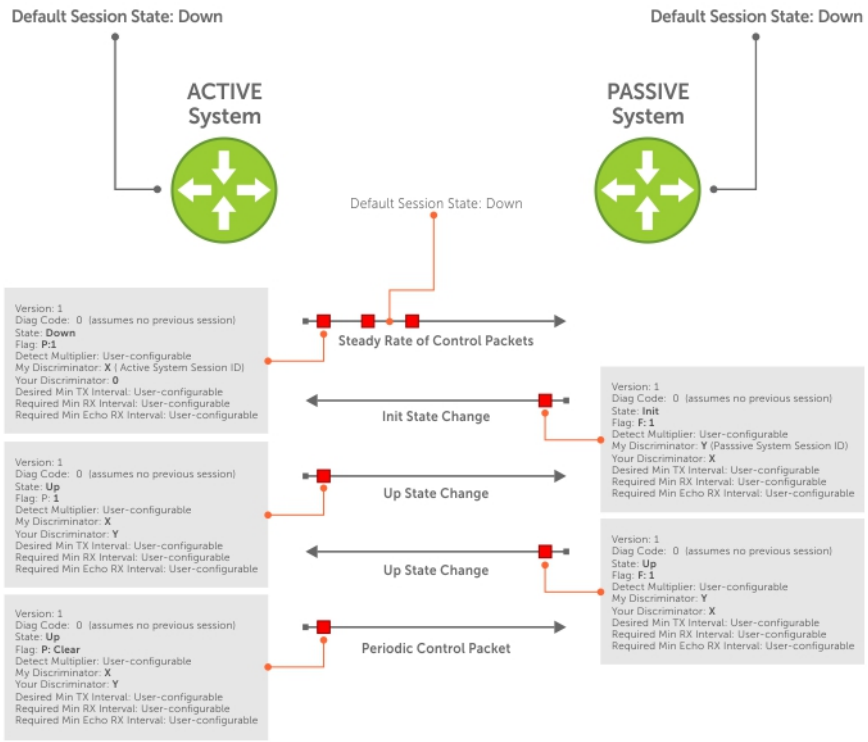


Figure 10. BFD Three-Way Handshake State Changes

Session State Changes

The following illustration shows how the session state on a system changes based on the status notification it receives from the remote system. For example, if a session on a system is down and it receives a Down status notification from the remote system, the session state on the local system changes to Init.

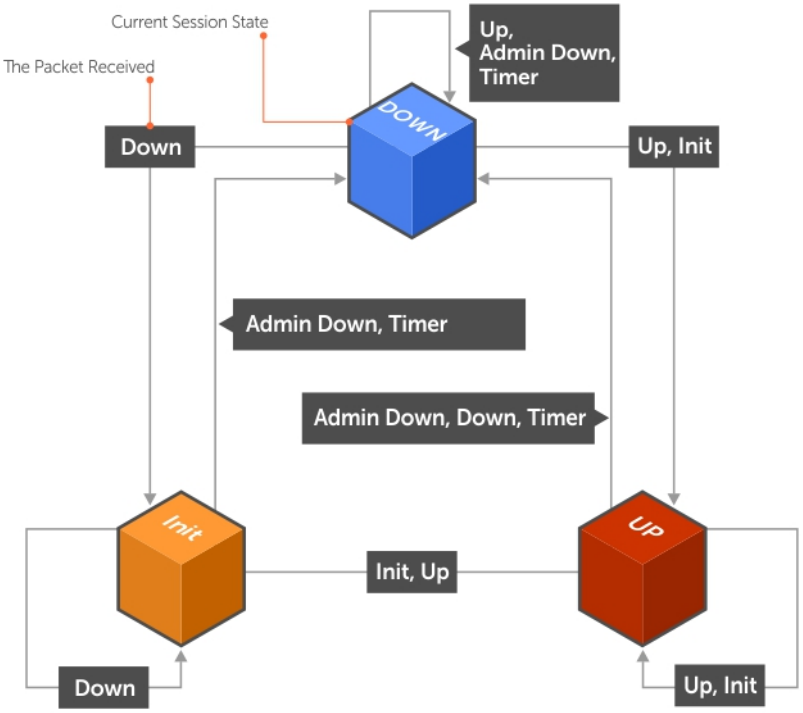


Figure 11. Session State Changes

Important Points to Remember

- Dell EMC Networking OS supports 128 sessions per stack unit at 300 minimum transmit and receive intervals with a multiplier of 3, 64 sessions at 150 minimum transmit and receive intervals with a multiplier of 3, and 32 sessions at 100 minimum transmit and receive intervals with a multiplier of 3.
- Enable BFD on both ends of a link.
- Demand mode, authentication, and the Echo function are not supported.
- BFD is not supported on multi-hop.
- Protocol Liveness is supported for routing protocols only.
- Dell EMC Networking OS supports static routes, OSPF, OSPFv3, IS-IS, BGP, IPv6 BGP, and VRRP protocols as BFD clients.
- Dell EMC Networking OS supports default and nondefault VRFs for IPv4 and IPv6 BGP.
- Dell EMC Networking OS does not recommend using 50 milliseconds as transmit and receive intervals in a production network.

Configure BFD

This section contains the following procedures.

- Configure BFD for Static Routes
- Configure BFD for OSPF

- Configure BFD for OSPFv3
- Configure BFD for IS-IS
- Configure BFD for BGP
- Configure BFD for VRRP
- Configuring Protocol Liveness

Configure BFD for Physical Ports

Configuring BFD for physical ports is supported on the C-Series and E-Series platforms only.

BFD on physical ports is useful when you do not enable the routing protocol. Without BFD, if the remote system fails, the local system does not remove the connected route until the first failed attempt to send a packet. When you enable BFD, the local system removes the route as soon as it stops receiving periodic control packets from the remote system.

Configuring BFD for a physical port is a two-step process:

1. Enable BFD globally.
2. Establish a session with a next-hop neighbor.

Related Configuration Tasks

- Viewing Physical Port Session Parameters.
- Disabling and Re-Enabling BFD.

Enabling BFD Globally

You must enable BFD globally on both routers.

To enable the BFD globally, use the following command.

- Enable BFD globally.

  CONFIGURATION mode

  `bfd enable`

To verify that BFD is enabled globally, use the `show running bfd` command.

The bold line shows that BFD is enabled.

```
R1(conf)#bfd ?
enable                Enable BFD protocol
protocol-liveness     Enable BFD protocol-liveness
R1(conf)#bfd enable

R1(conf)#do show running-config bfd
!
bfd enable
R1(conf)#
```

Viewing Physical Port Session Parameters

BFD sessions are configured with default intervals and a default role (active). Dell EMC Networking recommends maintaining the default values.

To view session parameters, use the `show bfd neighbors detail` command.

Example of Viewing Session Parameters

```
R1(conf-if-te-4/24)#bfd interval 100 min_rx 100 multiplier 4 role passive
R1(conf-if-te-4/24)#do show bfd neighbors detail

Session Discriminator: 1
Neighbor Discriminator: 1
Local Addr: 2.2.2.1
Local MAC Addr: 00:01:e8:09:c3:e5
```

```
Remote Addr: 2.2.2.2
Remote MAC Addr: 00:01:e8:06:95:a2
Int: TenGigabitEthernet 4/24
State: Up
Configured parameters:
  TX: 100ms, RX: 100ms, Multiplier: 4
Neighbor parameters:
  TX: 100ms, RX: 100ms, Multiplier: 3
Actual parameters:
  TX: 100ms, RX: 100ms, Multiplier: 4
Role: Passive
Delete session on Down: False
Client Registered: CLI
Uptime: 00:09:06
Statistics:
  Number of packets received from neighbor: 4092
  Number of packets sent to neighbor: 4093
  Number of state changes: 1
  Number of messages from IFA about port state change: 0
  Number of messages communicated b/w Manager and Agent: 7
```

Disabling and Re-Enabling BFD

BFD is enabled on all interfaces by default, though sessions are not created unless explicitly configured.

If you disable BFD, all of the sessions on that interface are placed in an Administratively Down state (the first message example), and the remote systems are notified of the session state change (the second message example).

To disable and re-enable BFD on an interface, use the following commands.

- Disable BFD on an interface.

  INTERFACE mode

  `no bfd enable`
- Enable BFD on an interface.

  INTERFACE mode

  `bfd enable`

If you disable BFD on a local interface, this message displays:

```
     R1(conf-if-te-4/24)#01:00:52: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed
session state to Ad
Dn for neighbor 2.2.2.2 on interface Te 4/24 (diag: 0)
```

If the remote system state changes due to the local state administration being down, this message displays:

```
     R2>01:32:53: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Down
for neighbor
2.2.2.1 on interface Te 2/1 (diag: 7)
```

Configure BFD for Static Routes

BFD offers systems a link state detection mechanism for static routes. With BFD, systems are notified to remove static routes from the routing table as soon as the link state change occurs, rather than waiting until packets fail to reach their next hop.

Configuring BFD for static routes is a three-step process:

1. Enable BFD globally.
2. Configure static routes on both routers on the system (either local or remote).
3. Configure an IP route to connect BFD on the static routes using the `ip route bfd` command.

Related Configuration Tasks

- Changing Static Route Session Parameters
- Disabling BFD for Static Routes

Establishing Sessions for Static Routes for Default VRF

Sessions are established for all neighbors that are the next hop of a static route on the default VRF.

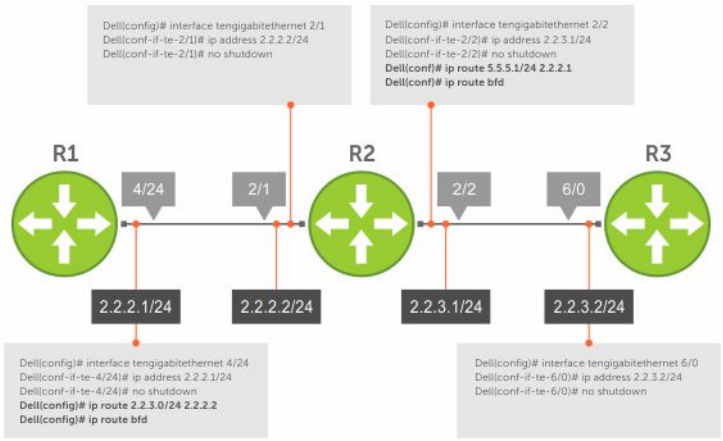


Figure 12. Establishing Sessions for Static Routes

To establish a BFD session, use the following command.

- Establish BFD sessions for all neighbors that are the next hop of a static route.

  CONFIGURATION mode

  `ip route bfd [prefix-list` *`prefix-list-name`*`] [interval` *`interval`* `min_rx` *`min_rx`* `multiplier` *`value`* `role {active | passive}]`

To verify that sessions have been created for static routes, use the `show bfd neighbors` command.

```
R1(conf)#ip route 2.2.3.0/24 2.2.2.2
R1(conf)#ip route bfd
R1(conf)#do show bfd neighbors

* - Active session role
Ad Dn - Admin Down
C - CLI
I - ISIS
O - OSPF
R - Static Route (RTM)
LocalAddr RemoteAddr Interface State Rx-int Tx-int Mult Clients
2.2.2.1   2.2.2.2    Te 4/24    Up    200    200    4    R
```

To view detailed session information, use the `show bfd neighbors detail` command.

Establishing Static Route Sessions on Specific Neighbors

You can selectively enable BFD sessions on specific neighbors based on a destination prefix-list.

When you establish a BFD session using the `ip route bfd` command, all the next-hop neighbors in the static route become part of the BFD session. Starting with Dell EMC Networking OS release 9.11.0.0, you can enable BFD sessions on specific next-hop neighbors. You can specify the next-hop neighbors to be part of a BFD session by including them in a prefix-list.

Prefix lists are used in route maps and route filtering operations. You can use prefix lists as an alternative to existing access lists (ACLs). A prefix is a portion of the IP address. Prefix lists constitute any number of bits in an IP address starting from the far left bit of the far left octet. By specifying the exactly number of bits in an IP address that belong to a prefix list, the prefix list can be used to aggregate addresses and perform some functions; for example, redistribution.

You can use the following options to enable or disable the BFD session:

- Permit – The permit option enables creation of a BFD session on the specified prefix list or prefix list range. The no permit option enables tear down of the BFD session if and only if the ACL has no permit entry that shares the same neighbor.
- Deny – The deny option prevents BFD sessions from getting created for the specified prefix list or prefix list range.

For more information on prefix lists, see IP Prefix Lists.

To enable BFD sessions on specific neighbors, perform the following steps:

Enter the following command to enable BFD session on specific next-hop neighbors:

CONFIGURATION

`ip route bfd prefix-list` *`prefix-list-name`*

The BFD session is established for the next-hop neighbors that are specified in the prefix-list.

- The absence of a prefix-list causes BFD sessions to be enabled on all the eligible next-hop neighbors.
- You can use only valid unicast address prefixes in the BFD prefix list. An erroneous prefix in a prefix-list causes the entire prefix-list to be rejected.
- A BFD session is enabled for the directly connected next-hop neighbor specified in the configured destination prefix list.
- If you attach an empty prefix-list, all the existing established BFD sessions are teared down. If a destination prefix or prefix range is not present in the prefix-list, then it is considered as an implicit deny.
- When a destination prefix is deleted from the prefix-list using the no permit option, the corresponding BFD session is torn down immediately. In this scenario, the BFD session tear down occurs only if the other destination prefixes in the prefix-list are not pointing to the same neighbor.
- The permit option enables creation of a BFD session for the specified static destination prefix or prefix range. The system prevents creation of BFD sessions for all other destination prefixes that are explicitly specified as Deny in the prefix list.
- If other destination prefixes in the prefix-list are pointing to the same neighbor, then the `no permit` or the `deny` option on a particular destination prefix neither creates a BFD session on a neighbor nor removes the static routes from the unicast database.
- BFD sessions created using any one IP prefix list are active at any given point in time. If a new prefix list is assigned, then BFD sessions corresponding to the older (existing) prefix list are replaced with the newer ones.
- Each time a prefix list is modified, only addition or deletion of new entries in that prefix list are processed for BFD session establishment or tear down.

Changing Static Route Session Parameters

BFD sessions are configured with default intervals and a default role.

The parameters you can configure are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured for all static routes. If you change a parameter, the change affects all sessions for static routes.

To change parameters for static route sessions, use the following command .

- Change parameters for all static route sessions.

  CONFIGURATION mode

  `ip route bfd [prefix-list` *`prefix-list-name`*`] interval` *`milliseconds`* `min_rx` *`milliseconds`* `multiplier` *`value`* `role [active | passive]`

To view session parameters, use the `show bfd neighbors detail` command.

Disabling BFD for Static Routes

If you disable BFD, all static route BFD sessions are torn down.

A final Admin Down packet is sent to all neighbors on the remote systems, and those neighbors change to the Down state.

To disable BFD for static routes, use the following command.

- Disable BFD for static routes.

  CONFIGURATION mode

  `no ip route bfd [prefix-list` *`prefix-list-name`*`] [interval` *`interval`* `min_rx` *`min_rx`* `multiplier` *`value`* `role {active | passive}]`

Configure BFD for OSPF

When you use BFD with OSPF, the OSPF protocol registers with the BFD manager. BFD sessions are established with all neighboring interfaces participating in OSPF. If a neighboring interface fails, the BFD agent notifies the BFD manager, which in turn notifies the OSPF protocol that a link state change has occurred.

Configuring BFD for OSPF is a two-step process:

1. Enable BFD globally.
2. Establish sessions with OSPF neighbors.

Related Configuration Tasks

- Changing OSPF Session Parameters
- Disabling BFD for OSPF

Establishing Sessions with OSPF Neighbors for the Default VRF

BFD sessions can be established with all OSPF neighbors at once or sessions can be established with all neighbors out of a specific interface. Sessions are only established when the OSPF adjacency is in the Full state.

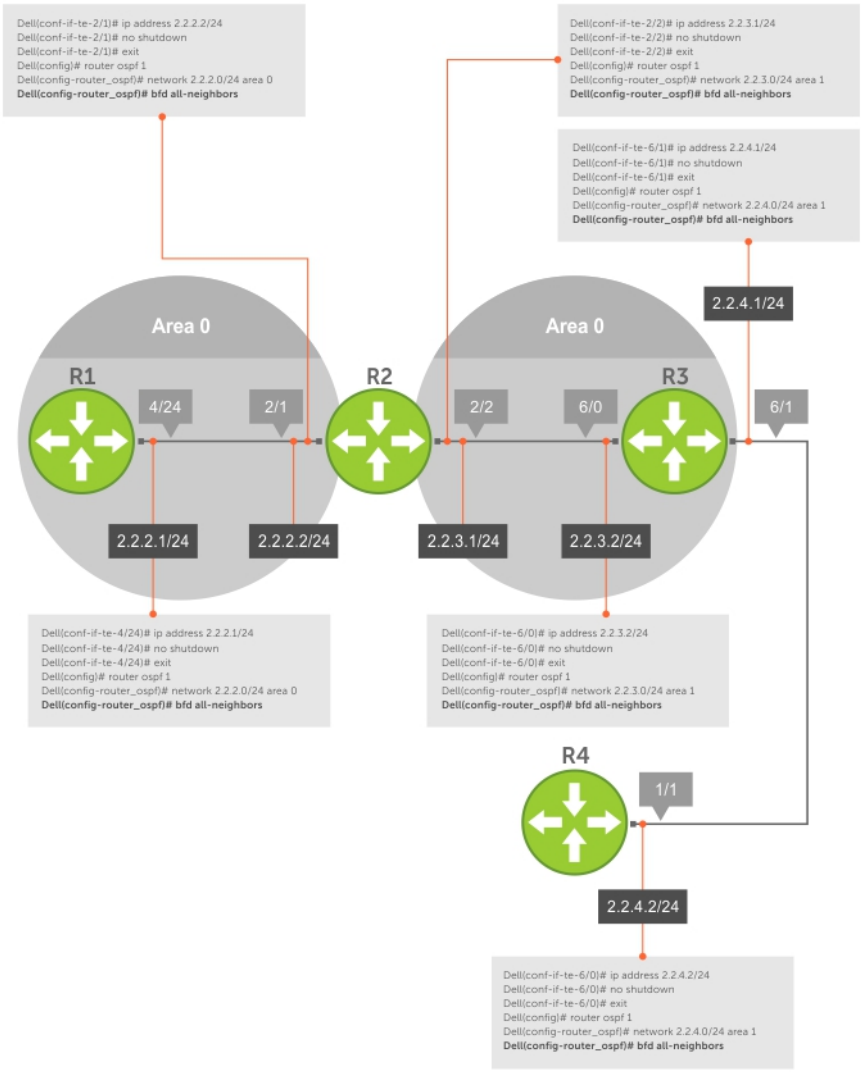


Figure 13. Establishing Sessions with OSPF Neighbors

To establish BFD with all OSPF neighbors or with OSPF neighbors on a single interface, use the following commands.

- Enable BFD globally.
  CONFIGURATION mode
  `bfd enable`
- Establish sessions with all OSPF neighbors.
  ROUTER-OSPF mode
  `bfd all-neighbors`
- Establish sessions with OSPF neighbors on a single interface.
  INTERFACE mode
  `ip ospf bfd all-neighbors`

To view the established sessions, use the `show bfd neighbors` command.

The bold line shows the OSPF BFD sessions.

```
R2(conf-router_ospf)#bfd all-neighbors
R2(conf-router_ospf)#do show bfd neighbors

*      - Active session role
Ad Dn - Admin Down
C      - CLI
I      - ISIS
O      - OSPF
R      - Static Route (RTM)

LocalAddr  RemoteAddr Interface State Rx-int Tx-int Mult Clients
* 2.2.2.2  2.2.2.1    Te 2/1    Up    100    100    3    O
* 2.2.3.1 2.2.3.2   Te 2/2   Up   100   100   3   O
```

Establishing Sessions with OSPF Neighbors for nondefault VRFs

To configure BFD in a nondefault VRF, follow this procedure:

- Enable BFD globally.
  CONFIGURATION mode
  `bfd enable`
- Establish sessions with all OSPF neighbors in a specific VRF.
  ROUTER-OSPF mode
  `bfd all-neighbors`
- Establish sessions with OSPF neighbors on a single interface in a specific VRF.
  INTERFACE mode
  `ip ospf bfd all-neighbors`
- to disable BFD on a specific OSPF enabled interface, use the `ip ospf bfd all-neighbors disable` command. You can also use the `no bfd enable` command to disable BFD on a specific interface.

The following example shows the configuration to establish sessions with all OSPF neighbors in a specific VRF:

```
router ospf 20 vrf VRF_blue
bfd all-neighbors
!
```

The following example shows the configuration to establish sessions with all OSPF neighbors on a single interface in a specific VRF:

```
int vlan 20
ip vrf forwarding vrf VRF_blue
ip ospf bfd all-neighbors
```

The following example shows the `show bfd neighbors` command output.

```
DellEMC# show bfd neighbors

*       - Active session role
Ad Dn   - Admin Down
B       - BGP
C       - CLI
I       - ISIS
O       - OSPF
O3      - OSPFv3
R       - Static Route (RTM)
M       - MPLS
V       - VRRP
VT      - Vxlan Tunnel

  LocalAddr   RemoteAddr           Interface State Rx-int Tx-int Mult Clients
* 1.0.1.1      1.0.1.2             Te 1/49/1 Up    200    200    3    B
* 3.3.3.3      192.168.122.135     Te 1/38   Up    1000   1000   3    VT
* 3.3.3.3      192.168.122.136     Te 1/42   Up    1000   1000   3    VT
* 3.3.3.3      192.168.122.137     Te 1/43   Up    1000   1000   3    VT
* 3.3.3.3      192.168.122.138     Te 1/38   Up    1000   1000   3    VT
* 3.3.3.3      192.168.122.139     Te 1/42   Up    1000   1000   3    VT
```

```
DellEMC# show bfd neighbors detail

Session Discriminator: 1
Neighbor Discriminator: 1
Local Addr: 10.1.3.2
Local MAC Addr: 00:01:e8:02:15:0e
Remote Addr: 10.1.3.1
Remote MAC Addr: 00:01:e8:27:2b:f1
Int: TenGigabitEthernet 1/3
State: Up
Configured parameters:
 TX: 100ms, RX: 100ms, Multiplier: 3
Neighbor parameters:
 TX: 250ms, RX: 300ms, Multiplier: 4
Actual parameters:
 TX: 300ms, RX: 250ms, Multiplier: 3
Role: Active
Delete session on Down: False
Client Registered: CLI
Uptime: 00:02:04
Statistics:
 Number of packets received from neighbor: 376
 Number of packets sent to neighbor: 314
 Number of state changes: 2
 Number of messages from IFA about port state change: 0
 Number of messages communicated b/w Manager and Agent: 6
DellEMC#
```

Changing OSPF Session Parameters

Configure BFD sessions with default intervals and a default role.

The parameters that you can configure are: `desired tx interval`, `required min rx interval`, `detection multiplier`, and `system role`. Configure these parameters for all OSPF sessions or all OSPF sessions on a particular interface. If you change a parameter globally, the change affects all OSPF neighbors sessions. If you change a parameter at the interface level, the change affects all OSPF sessions on that interface.

To change parameters for all OSPF sessions or for OSPF sessions on a single interface, use the following commands.

- Change parameters for OSPF sessions.

  ROUTER-OSPF mode

  `bfd all-neighbors interval` *`milliseconds`* `min_rx` *`milliseconds`* `multiplier` *`value`* `role [active | passive]`
- Change parameters for all OSPF sessions on an interface.

INTERFACE mode

```
ip ospf bfd all-neighbors interval milliseconds min_rx milliseconds multiplier value role
[active | passive]
```

To view session parameters, use the `show bfd neighbors detail` command.

Disabling BFD for OSPF

If you disable BFD globally, all sessions are torn down and sessions on the remote system are placed in a Down state.

If you disable BFD on an interface, sessions on the interface are torn down and sessions on the remote system are placed in a Down state. Disabling BFD does not trigger a change in BFD clients; a final Admin Down packet is sent before the session is terminated.

To disable BFD sessions, use the following commands.

- Disable BFD sessions with all OSPF neighbors.

  ROUTER-OSPF mode

  ```
  no bfd all-neighbors
  ```
- Disable BFD sessions with all OSPF neighbors on an interface.

  INTERFACE mode

  ```
  ip ospf bfd all-neighbors disable
  ```

Configure BFD for OSPFv3

BFD for OSPFv3 provides support for IPV6.

Configuring BFD for OSPFv3 is a two-step process:

1. Enable BFD globally.
2. Establish sessions with OSPFv3 neighbors.

Related Configuration Tasks

- Changing OSPFv3 Session Parameters
- Disabling BFD for OSPFv3

Establishing Sessions with OSPFv3 Neighbors

You can establish BFD sessions with all OSPFv3 neighbors at once or with all neighbors out of a specific interface. Sessions are only established when the OSPFv3 adjacency is in the Full state.

To establish BFD with all OSPFv3 neighbors or with OSPFv3 neighbors on a single interface, use the following commands.

- Establish sessions with all OSPFv3 neighbors.

  ROUTER-OSPFv3 mode

  ```
  bfd all-neighbors
  ```
- Establish sessions with OSPFv3 neighbors on a single interface.

  INTERFACE mode

  ```
  ipv6 ospf bfd all-neighbors
  ```

To view the established sessions, use the `show bfd neighbors` command.

The following example shows the `show bfd neighbors` command output for default VRF.

```
DellEMC#show bfd neighbors
*       - Active session role
Ad Dn   - Admin Down
B       - BGP
C       - CLI
I       - ISIS
O       - OSPF
```

```
O3       - OSPFv3
R        - Static Route (RTM)
M        - MPLS
V        - VRRP
VT       - Vxlan Tunnel


LocalAddr                   RemoteAddr                Interface  State Rx-int Tx-int Mult
Clients
* 1.1.1.1                   1.1.1.2                   Te 1/1     Up    200    200    3    O

* 2.1.1.1                   2.1.1.2                   Vl 2       Up    200    200    3    O

* fe80::2a0:c9ff:fe00:2     fe80::3617:98ff:fe34:12   Te 1/3     Up    200    200    3
O3

* fe80::2a0:c9ff:fe00:2     fe80::3617:98ff:fe34:12   Vl 2       Up    200    200    3
O3
DellEMC#
```

Establishing BFD Sessions with OSPFv3 Neighbors for nondefault VRFs

To configure BFD in a nondefault VRF, use the following procedure:

- Enable BFD globally.

  CONFIGURATION mode

  `bfd enable`
- Establish sessions with all OSPFv3 neighbors in a specific VRF.

  ROUTER-OSPFv3 mode

  `bfd all-neighbors`
- Establish sessions with the OSPFv3 neighbors on a single interface in a specific VRF.

  INTERFACE mode

  `ipv6 ospf bfd all-neighbors`
- To disable BFD on a specific OSPFv3 enabled interface, use the `ipv6 ospf bfd all-neighbors disable` command. You can also use the `no bfd enable` command to disable BFD on a specific interface.

NOTE: You can create upto a maximum of 128 BFD sessions (combination of OSPFv2 and OSPFv3 with a timer of 300*300*3) for both default and nondefault VRFs.

The following example shows the configuration to establish sessions with all OSPFv3 neighbors in a specific VRF:

```
ipv6 router ospf 20 vrf vrf1
bfd all-neighbors
!
```

The following example shows the configuration to establish sessions with all OSPFv3 neighbors on a single interface in a specific VRF:

```
interface vlan 102
ip vrf forwarding vrf vrf1
ipv6 ospf bfd all-neighbors
```

The following example shows the `show bfd vrf neighbors` command output for nondefault VRF:

```
DellEMC#show bfd vrf vrf1 neighbors
*        - Active session role
Ad Dn    - Admin Down
B        - BGP
C        - CLI
I        - ISIS
O        - OSPF
O3       - OSPFv3
R        - Static Route (RTM)
M        - MPLS
V        - VRRP
```

```
VT      - Vxlan Tunnel

  LocalAddr                RemoteAddr                Interface  State Rx-int Tx-int
Mult VRF Clients
* 10.1.1.1                 10.1.1.2                  Vl 100     Up    150    150    3
 511    O

* 11.1.1.1                 11.1.1.2                  Vl 101     Up    150    150    3
 511    O

* 12.1.1.1                 12.1.1.2                  Vl 102     Up    150    150    3
 511    O

* 13.1.1.1                 13.1.1.2                  Vl 103     Up    150    150    3
 511    O

* fe80::2a0:c9ff:fe00:2    fe80::3617:98ff:fe34:12   Vl 100     Up    150    150    3
 511    O3

* fe80::2a0:c9ff:fe00:2    fe80::3617:98ff:fe34:12   Vl 101     Up    150    150    3
 511    O3

* fe80::2a0:c9ff:fe00:2    fe80::3617:98ff:fe34:12   Vl 102     Up    150    150    3
 511    O3

* fe80::2a0:c9ff:fe00:2    fe80::3617:98ff:fe34:12   Vl 103     Up    150    150    3
 511    O3
DellEMC#
```

Changing OSPFv3 Session Parameters

Configure BFD sessions with default intervals and a default role.

The parameters that you can configure are: `desired tx interval`, `required min rx interval`, `detection multiplier`, and system `role`. Configure these parameters for all OSPFv3 sessions or all OSPFv3 sessions on a particular interface. If you change a parameter globally, the change affects all OSPFv3 neighbors sessions. If you change a parameter at the interface level, the change affects all OSPFv3 sessions on that interface.

To change parameters for all OSPFv3 sessions or for OSPFv3 sessions on a single interface, use the following commands.

To view session parameters, use the `show bfd neighbors detail` command.

- Change parameters for all OSPFv3 sessions.

  ROUTER-OSPFv3 mode

  `bfd all-neighbors interval` *`milliseconds`* `min_rx` *`milliseconds`* `multiplier` *`value`* `role [active | passive]`
- Change parameters for OSPFv3 sessions on a single interface.

  INTERFACE mode

  `ipv6 ospf bfd all-neighbors interval` *`milliseconds`* `min_rx` *`milliseconds`* `multiplier` *`value`* `role [active | passive]`

Disabling BFD for OSPFv3

If you disable BFD globally, all sessions are torn down and sessions on the remote system are placed in a Down state.

If you disable BFD on an interface, sessions on the interface are torn down and sessions on the remote system are placed in a Down state. Disabling BFD does not trigger a change in BFD clients; a final Admin Down packet is sent before the session is terminated.

To disable BFD sessions, use the following commands.

- Disable BFD sessions with all OSPFv3 neighbors.

  ROUTER-OSPFv3 mode

  `no bfd all-neighbors`
- Disable BFD sessions with OSPFv3 neighbors on a single interface.

INTERFACE mode

```
ipv6 ospf bfd all-neighbors disable
```

Configure BFD for IS-IS

When using BFD with IS-IS, the IS-IS protocol registers with the BFD manager on the RPM. BFD sessions are then established with all neighboring interfaces participating in IS-IS. If a neighboring interface fails, the BFD agent on the line card notifies the BFD manager, which in turn notifies the IS-IS protocol that a link state change occurred.

Configuring BFD for IS-IS is a two-step process:

1. Enable BFD globally.
2. Establish sessions for all or particular IS-IS neighbors.

Related Configuration Tasks

- Changing IS-IS Session Parameters
- Disabling BFD for IS-IS

Establishing Sessions with IS-IS Neighbors

BFD sessions can be established for all IS-IS neighbors at once or sessions can be established for all neighbors out of a specific interface.

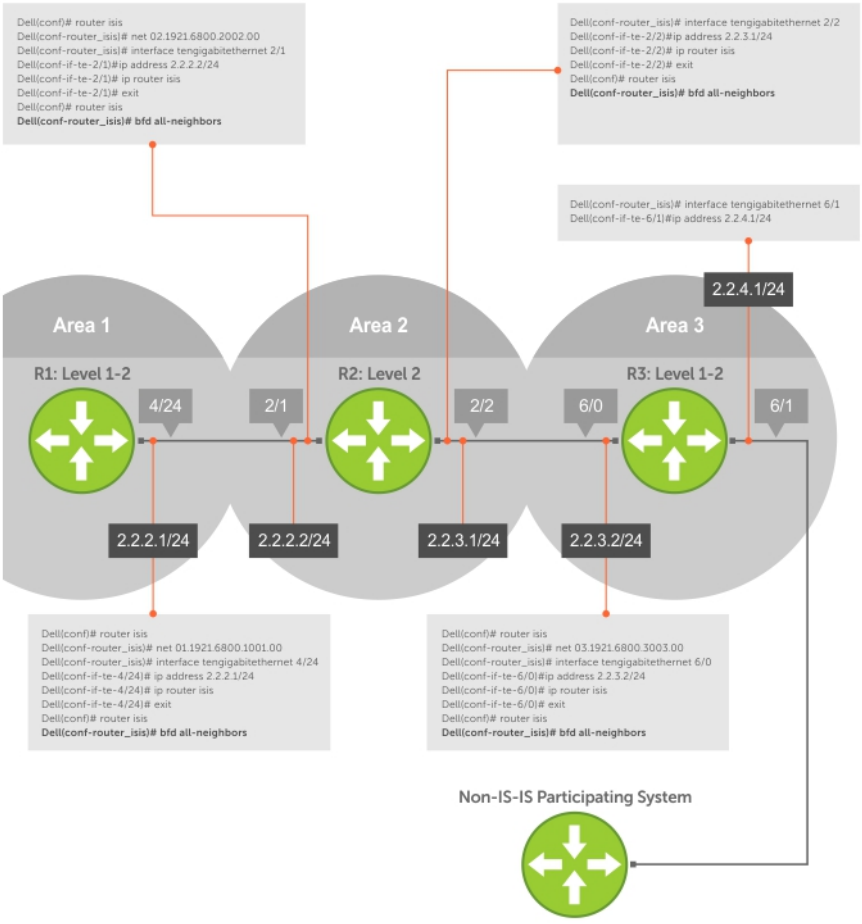


Figure 14. Establishing Sessions with IS-IS Neighbors

To establish BFD with all IS-IS neighbors or with IS-IS neighbors on a single interface, use the following commands.

- Establish sessions with all IS-IS neighbors.

ROUTER-ISIS mode

```
bfd all-neighbors
```

- Establish sessions with IS-IS neighbors on a single interface.

  INTERFACE mode

  ```
  isis bfd all-neighbors
  ```

To view the established sessions, use the `show bfd neighbors` command.

The bold line shows that IS-IS BFD sessions are enabled.

```
R2(conf-router_isis)#bfd all-neighbors
R2(conf-router_isis)#do show bfd neighbors

*      - Active session role
Ad Dn - Admin Down
C      - CLI
I    - ISIS
O      - OSPF
R      - Static Route (RTM)

LocalAddr  RemoteAddr Interface State Rx-int Tx-int Mult Clients
* 2.2.2.2  2.2.2.1    Te 2/1    Up    100    100    3    I
```

Changing IS-IS Session Parameters

BFD sessions are configured with default intervals and a default role.

The parameters that you can configure are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured for all IS-IS sessions or all IS-IS sessions out of an interface. If you change a parameter globally, the change affects all IS-IS neighbors sessions. If you change a parameter at the interface level, the change affects all IS-IS sessions on that interface.

To change parameters for all IS-IS sessions or for IS-IS sessions on a single interface, use the following commands.

To view session parameters, use the `show bfd neighbors detail` command, as shown in *Verifying BFD Sessions with BGP Neighbors Using the `show bfd neighbors` Command*.

- Change parameters for all IS-IS sessions.

  ROUTER-ISIS mode

  ```
  bfd all-neighbors interval milliseconds min_rx milliseconds multiplier value role [active | passive]
  ```

- Change parameters for IS-IS sessions on a single interface.

  INTERFACE mode

  ```
  isis bfd all-neighbors interval milliseconds min_rx milliseconds multiplier value role [active | passive]
  ```

Disabling BFD for IS-IS

If you disable BFD globally, all sessions are torn down and sessions on the remote system are placed in a Down state.

If you disable BFD on an interface, sessions on the interface are torn down and sessions on the remote system are placed in a Down state. Disabling BFD does not trigger a change in BFD clients; a final Admin Down packet is sent before the session is terminated.

To disable BFD sessions, use the following commands.

- Disable BFD sessions with all IS-IS neighbors.

  ROUTER-ISIS mode

  ```
  no bfd all-neighbors
  ```

- Disable BFD sessions with IS-IS neighbors on a single interface.

  INTERFACE mose

  ```
  isis bfd all-neighbors disable
  ```

Configure BFD for BGP

In a BGP core network, BFD provides rapid detection of communication failures in BGP fast-forwarding paths between internal BGP (iBGP) and external BGP (eBGP) peers for faster network reconvergence. BFD for BGP is supported on physical, port-channel, and VLAN interfaces. BFD for BGP does not support the BGP multihop feature.

Before configuring BFD for BGP, you must first configure BGP on the routers that you want to interconnect.

For example, the following illustration shows a sample BFD configuration on Router 1 and Router 2 that use eBGP in a transit network to interconnect AS1 and AS2. The eBGP routers exchange information with each other as well as with iBGP routers to maintain connectivity and accessibility within each autonomous system.

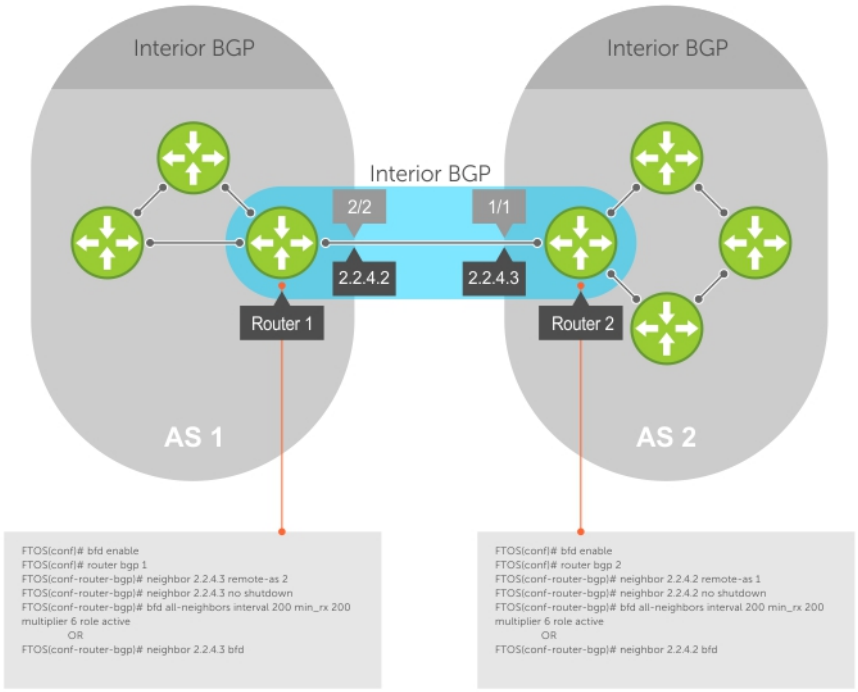


Figure 15. Establishing Sessions with BGP Neighbors

The sample configuration shows alternative ways to establish a BFD session with a BGP neighbor:

- By establishing BFD sessions with all neighbors discovered by BGP (the `bfd all-neighbors` command).
- By establishing a BFD session with a specified BGP neighbor (the `neighbor {ip-address | peer-group-name} bfd` command)

BFD packets originating from a router are assigned to the highest priority egress queue to minimize transmission delays. Incoming BFD control packets received from the BGP neighbor are assigned to the highest priority queue within the control plane policing (COPP) framework to avoid BFD packets drops due to queue congestion.

BFD notifies BGP of any failure conditions that it detects on the link. Recovery actions are initiated by BGP.

BFD for BGP is supported only on directly-connected BGP neighbors and in both BGP IPv4 and IPV6 networks. Up to 200 simultaneous BFD sessions are supported.

As long as each BFD for BGP neighbor receives a BFD control packet within the configured BFD interval for failure detection, the BFD session remains up and BGP maintains its adjacencies. If a BFD for BGP neighbor does not receive a control packet within the detection interval, the router informs any clients of the BFD session (other routing protocols) about the failure. It then depends on the individual routing protocols that uses the BFD link to determine the appropriate response to the failure condition. The typical response is to terminate the peering session for the routing protocol and reconverge by bypassing the failed neighboring router. A log message is generated whenever BFD detects a failure condition.

Prerequisites

Before configuring BFD for BGP, you must first configure the following settings:

- Configure BGP on the routers that you want to interconnect.

Establishing Sessions with BGP Neighbors for Default VRF

To establish sessions with either IPv6 or IPv4 BGP neighbors for the default VRF, follow these steps:

1. Enable BFD globally.
   CONFIGURATION mode
   `bfd enable`
2. Specify the AS number and enter ROUTER BGP configuration mode.
   CONFIGURATION mode
   `router bgp as-number`
3. Add a BGP neighbor or peer group in a remote AS.
   CONFIG-ROUTERBGP mode
   `neighbor {ip-address | peer-group name} remote-as as-number`
4. Enable the BGP neighbor.
   CONFIG-ROUTERBGP mode
   `neighbor {ip-address | peer-group-name} no shutdown`
5. Add a BGP neighbor or peer group in a remote AS.
   CONFIG-ROUTERBGP mode
   `neighbor {ipv6-address | peer-group name} remote-as as-number`
6. Enable the BGP neighbor.
   CONFIG-ROUTERBGP mode
   `neighbor { ipv6-address | peer-group-name} no shutdown`
7. To establish BFD sessions for IPv6 neighbors, specify the address family as IPv6.
   CONFIG-ROUTERBGP mode
   `address-family ipv6 unicast`
8. Activate the neighbor in IPv6 address family.
   CONFIG-ROUTERBGPv6_ADDRESSFAMILY mode
   `neighbor ipv6-address activate`
9. Configure parameters for a BFD session established with all neighbors discovered by BGP. Or establish a BFD session with a specified BGP neighbor or peer group using the default BFD session parameters.
   CONFIG-ROUTERBGP mode
   `bfd all-neighbors`

```
DellEMC(conf)#router bgp 1
DellEMC(conf-router_bgp)#neighbor 10.1.1.2 remote-as 2
DellEMC(conf-router_bgp)#neighbor 10.1.1.2 no shutdown
DellEMC(conf-router_bgp)#neighbor 20::2 remote-as 2
DellEMC(conf-router_bgp)#neighbor 20::2 no shutdown
DellEMC(conf-router_bgp)#address-family ipv6 unicast
DellEMC(conf-router_bgpv6_af)#neighbor 20::2 activate
DellEMC(conf-router_bgpv6_af)#exit
DellEMC(conf-router_bgp)#bfd all-neighbors
DellEMC(conf-router_bgp)#show config
!
router bgp 1
 neighbor 10.1.1.2 remote-as 2
 neighbor 10.1.1.2 no shutdown
 neighbor 20::2 remote-as 2
 neighbor 20::2 no shutdown
 bfd all-neighbors
 !
 address-family ipv6 unicast
```

```
   neighbor 20::2 activate
   exit-address-family
DellEMC(conf-router_bgp)#
```

Establishing Sessions with BGP Neighbors for Nondefault VRF

To establish sessions with either IPv6 or IPv4 BGP neighbors for nondefault VRFs, follow these steps:

1. Enable BFD globally.
   CONFIGURATION mode
   `bfd enable`
2. Specify the AS number and enter ROUTER BGP configuration mode.
   CONFIGURATION mode
   `router bgp as-number`
3. Specify the address family as IPv4.
   CONFIG-ROUTERBGP mode
   `address-family ipv4 vrf vrf-name`
4. Add an IPv4 BGP neighbor or peer group in a remote AS.
   CONFIG-ROUTERBGP_ADDRESSFAMILY mode
   `neighbor {ip-address | peer-group name} remote-as as-number`
5. Enable the BGP neighbor.
   CONFIG-ROUTERBGP_ADDRESSFAMILY mode
   `neighbor {ip-address | peer-group-name} no shutdown`
6. Add an IPv6 BGP neighbor or peer group in a remote AS.
   CONFIG-ROUTERBGP_ADDRESSFAMILY mode
   `neighbor {ipv6-address | peer-group name} remote-as as-number`
7. Enable the BGP neighbor.
   CONFIG-ROUTERBGP_ADDRESSFAMILY mode
   `neighbor { ipv6-address | peer-group-name} no shutdown`
8. Specify the address family as IPv6.
   CONFIG-ROUTERBGP_ADDRESSFAMILY mode
   `address-family ipv6 unicast vrf vrf-name`

   NOTE: Before performing this step, create the required VRF.

9. Activate the neighbor in IPv6 address family.
   CONFIG-ROUTERBGPv6_ADDRESSFAMILY mode
   `neighbor ipv6-address activate`
10. Configure parameters for a BFD session established with all neighbors discovered by BGP. Or establish a BFD session with a specified BGP neighbor or peer group using the default BFD session parameters.
    CONFIG-ROUTERBGP mode
    `bfd all-neighbors`

```
DellEMC(conf)#router bgp 1
DellEMC(conf-router_bgp)#address-family ipv4 vrf vrf1
DellEMC(conf-router_bgp_af)#neighbor 10.1.1.2 remote-as 2
DellEMC(conf-router_bgp_af)#neighbor 10.1.1.2 no shutdown
DellEMC(conf-router_bgp_af)#neighbor 20::2 remote-as 2
DellEMC(conf-router_bgp_af)#neighbor 20::2 no shutdown
DellEMC(conf-router_bgp_af)#address-family ipv6 unicast vrf vrf1
DellEMC(conf-router_bgpv6_af)#neighbor 20::2 activate
DellEMC(conf-router_bgpv6_af)#address-family ipv4 vrf vrf1
DellEMC(conf-router_bgp_af)#bfd all-neighbors
DellEMC(conf-router_bgp_af)#exit
DellEMC(conf-router_bgp)#show config
!
```

```
router bgp 1
 !
 address-family ipv4 vrf vrf1
  neighbor 10.1.1.2 remote-as 2
  neighbor 10.1.1.2 no shutdown
  neighbor 20::2 remote-as 2
  neighbor 20::2 no shutdown
  bfd all-neighbors
  exit-address-family
 !
 address-family ipv6 unicast vrf vrf1
  neighbor 20::2 activate
  exit-address-family
DellEMC(conf-router_bgp)#
```

Disabling BFD for BGP

You can disable BFD for BGP.

To disable a BFD for BGP session with a specified neighbor, use the first command. To remove the disabled state of a BFD for BGP session with a specified neighbor, use the second command.

The BGP link with the neighbor returns to normal operation and uses the BFD session parameters globally configured with the `bfd all-neighbors` command or configured for the peer group to which the neighbor belongs.

- Disable a BFD for BGP session with a specified neighbor.

  ROUTER BGP mode

  `neighbor {`*`ip-address | ipv6-address |peer-group-name`*`} bfd disable`
- Remove the disabled state of a BFD for BGP session with a specified neighbor.

  ROUTER BGP mode

  `no neighbor {`*`ip-address | ipv6-address | peer-group-name`*`} bfd disable`

Use BFD in a BGP Peer Group

You can establish a BFD session for the members of a peer group (the `neighbor peer-group-name bfd` command in ROUTER BGP configuration mode).

Members of the peer group may have BFD:

- Explicitly enabled (the `neighbor ip-address bfd` command)
- Explicitly disabled (the `neighbor ip-address bfd disable` command)
- Inherited (neither explicitly enabled or disabled) according to the current BFD configuration of the peer group. For information about BGP peer groups, refer to *Configure peer groups*.

If you explicitly enable (or disable) a BGP neighbor for BFD that belongs to a peer group:

- The neighbor does not inherit the BFD enable/disable values configured with the `bfd all-neighbors` command or configured for the peer group to which the neighbor belongs.
- The neighbor inherits only the global timer values that are configured with the `bfd all-neighbors` command (interval, min_rx, and multiplier).

If you explicitly enable (or disable) a peer group for BFD that has no BFD parameters configured (for example, advertisement interval) using the `neighbor` *`peer-group-name`* `bfd` command, the peer group inherits any BFD settings configured with the `bfd all-neighbors` command.

Displaying BFD for BGP Information

You can display related information for BFD for BGP.

To display information about BFD for BGP sessions on a router, use the following commands and refer to the following examples.

- Verify a BFD for BGP configuration.

  EXEC Privilege mode

  `show running-config bgp`

- Verify that a BFD for BGP session has been successfully established with a BGP neighbor. A line-by-line listing of established BFD adjacencies is displayed.

  EXEC Privilege mode

  `show bfd neighbors [interface] [detail]`
- Check to see if BFD is enabled for BGP connections.

  EXEC Privilege mode

  `show ip bgp summary`
- Displays routing information exchanged with BGP neighbors, including BFD for BGP sessions.

  EXEC Privilege mode

  `show ip bgp neighbors [ip-address]`

The following example shows verifying a BGP configuration.

```
R2# show running-config bgp
!
router bgp 2
   neighbor 1.1.1.2 remote-as 1
   neighbor 1.1.1.2 no shutdown
   neighbor 2.2.2.2 remote-as 1
   neighbor 2.2.2.2 no shutdown
   neighbor 3.3.3.2 remote-as 1
   neighbor 3.3.3.2 no shutdown
   bfd all-neighbors
```

The following example shows viewing all BFD neighbors.

```
R2# show bfd neighbors

*     - Active session role
Ad Dn - Admin Down
B     - BGP
C     - CLI
I     - ISIS
O     - OSPF
R     - Static Route (RTM)
M     - MPLS
V     - VRRP

LocalAddr  RemoteAddr Interface State Rx-int Tx-int Mult Clients
* 1.1.1.3  1.1.1.2    Te 6/1    Up    200    200    3    B
* 2.2.2.3  2.2.2.2    Te 6/2    Up    200    200    3    B
* 3.3.3.3  3.3.3.2    Te 6/3    Up    200    200    3    B
```

The following example shows viewing BFD neighbors with full detail.

The bold lines show the BFD session parameters: TX (packet transmission), RX (packet reception), and multiplier (maximum number of missed packets).

```
R2# show bfd neighbors detail

Session Discriminator: 9
Neighbor Discriminator: 10
Local Addr: 1.1.1.3
Local MAC Addr: 00:01:e8:66:da:33
Remote Addr: 1.1.1.2
Remote MAC Addr: 00:01:e8:8a:da:7b
Int: TenGigabitEthernet 6/1
State: Up
Configured parameters:
TX: 200ms, RX: 200ms, Multiplier: 3
Neighbor parameters:
TX: 200ms, RX: 200ms, Multiplier: 3
Actual parameters:
TX: 200ms, RX: 200ms, Multiplier: 3
Role: Active
Delete session on Down: True
Client Registered: BGP
Uptime: 00:07:55
```

```
Statistics:
Number of packets received from neighbor: 4762
Number of packets sent to neighbor: 4490
Number of state changes: 2
Number of messages from IFA about port state change: 0
Number of messages communicated b/w Manager and Agent: 5

Session Discriminator: 10
Neighbor Discriminator: 11
Local Addr: 2.2.2.3
Local MAC Addr: 00:01:e8:66:da:34
Remote Addr: 2.2.2.2
Remote MAC Addr: 00:01:e8:8a:da:7b
Int: TenGigabitEthernet 6/2
State: Up
Configured parameters:
TX: 200ms, RX: 200ms, Multiplier: 3
Neighbor parameters:
TX: 200ms, RX: 200ms, Multiplier: 3
Actual parameters:
TX: 200ms, RX: 200ms, Multiplier: 3
Role: Active
Delete session on Down: True
Client Registered: BGP
Uptime: 00:02:22
Statistics:
Number of packets received from neighbor: 1428
Number of packets sent to neighbor: 1428
Number of state changes: 1
Number of messages from IFA about port state change: 0
Number of messages communicated b/w Manager and Agent: 4
```

The following example shows viewing BFD summary information.

The bold line shows the message displayed when you enable BFD for BGP connections.

```
R2# show ip bgp summary
BGP router identifier 10.0.0.1, local AS number 2
BGP table version is 0, main routing table version 0
BFD is enabled, Interval 200 Min_rx 200 Multiplier 3 Role Active
3 neighbor(s) using 24168 bytes of memory

Neighbor AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down   State/Pfx

1.1.1.2  1  282     281     0      0   0    00:38:12  0
2.2.2.2  1  273     273     0      0   (0)  04:32:26  0
3.3.3.2  1  282     281     0      0   0    00:38:12  0
```

The following example shows viewing BFD information for a specified neighbor.

The bold lines show the message displayed when you enable a BFD session with different configurations:

- Message displays when you enable a BFD session with a BGP neighbor that inherits the global BFD session settings configured with the `global bfd all-neighbors` command.
- Message displays when you enable a BFD session with a BGP neighbor using the `neighbor` *`ip-address`* `bfd` command.
- Message displays when you enable a BGP neighbor in a peer group for which you enabled a BFD session using the `neighbor` *`peer-group-name`* `bfd` command

```
R2# show ip bgp neighbors 2.2.2.2

BGP neighbor is 2.2.2.2, remote AS 1, external link
  BGP version 4, remote router ID 12.0.0.4
  BGP state ESTABLISHED, in this state for 00:05:33
  Last read 00:00:30, last write 00:00:30
  Hold time is 180, keepalive interval is 60 seconds
  Received 8 messages, 0 in queue
    1 opens, 0 notifications, 0 updates
    7 keepalives, 0 route refresh requests
  Sent 9 messages, 0 in queue
    2 opens, 0 notifications, 0 updates
    7 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
```

```
  Minimum time before advertisements start is 0 seconds

Capabilities received from neighbor for IPv4 Unicast :
  MULTIPROTO_EXT(1)
  ROUTE_REFRESH(2)
  CISCO_ROUTE_REFRESH(128)

Capabilities advertised to neighbor for IPv4 Unicast :
  MULTIPROTO_EXT(1)
  ROUTE_REFRESH(2)
  CISCO_ROUTE_REFRESH(128)
```

Neighbor is using BGP global mode BFD configuration

```
For address family: IPv4 Unicast
BGP table version 0, neighbor version 0
Prefixes accepted 0 (consume 0 bytes), withdrawn 0 by peer, martian prefixes ignored 0
Prefixes advertised 0, denied 0, withdrawn 0 from peer

  Connections established 1; dropped 0
  Last reset never
Local host: 2.2.2.3, Local port: 63805
Foreign host: 2.2.2.2, Foreign port: 179
E1200i_R2#

R2# show ip bgp neighbors 2.2.2.3

BGP neighbor is 2.2.2.3, remote AS 1, external link
  Member of peer-group pg1 for session parameters
  BGP version 4, remote router ID 12.0.0.4
  BGP state ESTABLISHED, in this state for 00:05:33
  ...
  Neighbor is using BGP neighbor mode BFD configuration
  Peer active in peer-group outbound optimization
...

R2# show ip bgp neighbors 2.2.2.4

BGP neighbor is 2.2.2.4, remote AS 1, external link
  Member of peer-group pg1 for session parameters
  BGP version 4, remote router ID 12.0.0.4
  BGP state ESTABLISHED, in this state for 00:05:33
  ...
  Neighbor is using BGP peer-group mode BFD configuration
  Peer active in peer-group outbound optimization
  ...
```

Configure BFD for VRRP

When using BFD with VRRP, the VRRP protocol registers with the BFD manager on the route processor module (RPM). BFD sessions are established with all neighboring interfaces participating in VRRP. If a neighboring interface fails, the BFD agent on the line card notifies the BFD manager, which in turn notifies the VRRP protocol that a link state change occurred.

Configuring BFD for VRRP is a three-step process:

1. Enable BFD globally. Refer to Enabling BFD Globally.
2. Establish VRRP BFD sessions with all VRRP-participating neighbors.
3. On the master router, establish a VRRP BFD sessions with the backup routers. Refer to Establishing Sessions with All VRRP Neighbors.

Related Configuration Tasks

- Changing VRRP Session Parameters.
- Disabling BFD for VRRP.

Establishing Sessions with All VRRP Neighbors

BFD sessions can be established for all VRRP neighbors at once, or a session can be established with a particular neighbor.

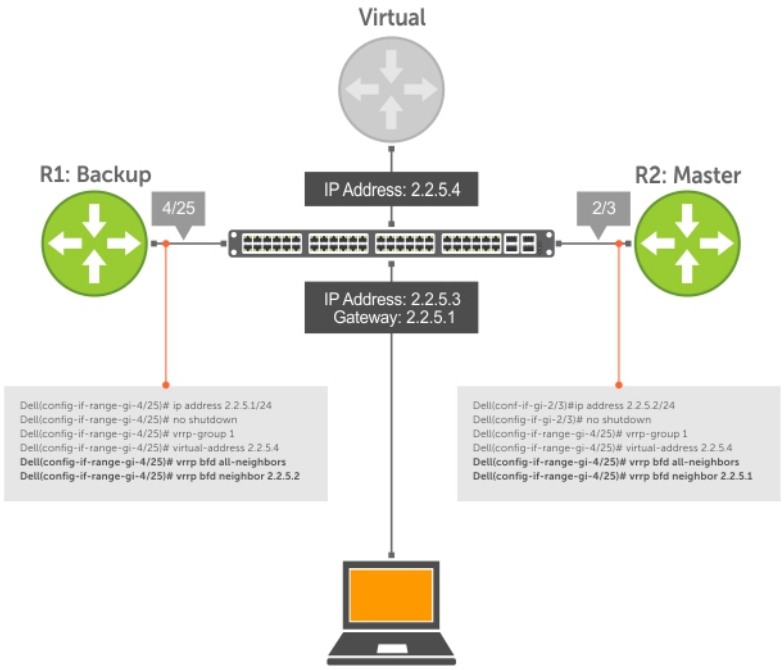


Figure 16. Establishing Sessions with All VRRP Neighbors

To establish sessions with all VRRP neighbors, use the following command.

- Establish sessions with all VRRP neighbors.

  INTERFACE mode

  `vrrp bfd all-neighbors`

Establishing VRRP Sessions on VRRP Neighbors

The master router does not care about the state of the backup router, so it does not participate in any VRRP BFD sessions.

VRRP BFD sessions on the backup router cannot change to the UP state. Configure the master router to establish an individual VRRP session the backup router.

To establish a session with a particular VRRP neighbor, use the following command.

- Establish a session with a particular VRRP neighbor.

  INTERFACE mode

  `vrrp bfd neighbor` *`ip-address`*

To view the established sessions, use the `show bfd neighbors` command.

The bold line shows that VRRP BFD sessions are enabled.

```
DellEMC(conf-if-te-4/25)#vrrp bfd all-neighbors
DellEMC(conf-if-te-4/25)#do show bfd neighbor

*      - Active session role
Ad Dn  - Admin Down
C      - CLI
I      - ISIS
O      - OSPF
R      - Static Route (RTM)
V      - VRRP
```

```
LocalAddr  RemoteAddr Interface State Rx-int Tx-int Mult Clients
* 2.2.5.1 2.2.5.2   Te 4/25  Down  1000   1000   3   V
```

To view session state information, use the `show vrrp` command.

The bold line shows the VRRP BFD session.

```
DellEMC(conf-if-te-4/25)#do show vrrp
------------------
TenGigabitEthernet 4/1, VRID: 1, Net: 2.2.5.1
VRF:0 default
State: Backup, Priority: 1, Master: 2.2.5.2
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 95, Bad pkts rcvd: 0, Adv sent: 933, Gratuitous ARP sent: 3
Virtual MAC address:
  00:00:5e:00:01:01
Virtual IP address:
  2.2.5.4
Authentication: (none)
BFD Neighbors:
RemoteAddr  State
2.2.5.2     Up
```

Changing VRRP Session Parameters

BFD sessions are configured with default intervals and a default role.

The parameters that you can configure are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. You can change parameters for all VRRP sessions or for a particular neighbor.

To change parameters for all VRRP sessions or for a particular VRRP session, use the following commands.

- Change parameters for all VRRP sessions.

  INTERFACE mode

  `vrrp bfd all-neighbors interval` *`milliseconds`* `min_rx` *`milliseconds`* `multiplier` *`value`* `role [active | passive]`
- Change parameters for a particular VRRP session.

  INTERFACE mode

  `vrrp bfd neighbor` *`ip-address`* `interval` *`milliseconds`* `min_rx` *`milliseconds`* `multiplier value role [active | passive]`

To view session parameters, use the `show bfd neighbors detail` command.

Disabling BFD for VRRP

If you disable any or all VRRP sessions, the sessions are torn down.

A final Admin Down control packet is sent to all neighbors and sessions on the remote system change to the Down state.

To disable all VRRP sessions on an interface, sessions for a particular VRRP group, or for a particular VRRP session on an interface, use the following commands.

- Disable all VRRP sessions on an interface.

  INTERFACE mode

  `no vrrp bfd all-neighbors`
- Disable all VRRP sessions in a VRRP group.

  VRRP mode

  `bfd disable`
- Disable a particular VRRP session on an interface.

  INTERFACE mode

  `no vrrp bfd neighbor` *`ip-address`*

Configuring Protocol Liveness

Protocol liveness is a feature that notifies the BFD manager when a client protocol is disabled.

When you disable a client, all BFD sessions for that protocol are torn down. Neighbors on the remote system receive an Admin Down control packet and are placed in the Down state.

To enable protocol liveness, use the following command.

- Enable Protocol Liveness.

  CONFIGURATION mode

  `bfd protocol-liveness`

Troubleshooting BFD

To troubleshoot BFD, use the following commands and examples.

To control packet field values or to examine the control packets in hexadecimal format, use the following command.

- Examine control packet field values.

  CONFIGURATION mode

  `debug bfd detail`

- Examine the control packets in hexadecimal format.

  CONFIGURATION

  `debug bfd packet`

The following example shows a three-way handshake using the `debug bfd detail` command.

```
     R1(conf-if-te-4/24)#00:54:38: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed
session state to
Down for neighbor 2.2.2.2 on interface Te 4/24 (diag: 0)
     00:54:38 : Sent packet for session with neighbor 2.2.2.2 on Te 4/24
     TX packet dump:
       Version:1, Diag code:0, State:Down, Poll bit:0, Final bit:0, Demand bit:0
       myDiscrim:4, yourDiscrim:0, minTx:1000000, minRx:1000000, multiplier:3, minEchoRx:0
     00:54:38 : Received packet for session with neighbor 2.2.2.2 on Te 4/24
     RX packet dump:
       Version:1, Diag code:0, State:Init, Poll bit:0, Final bit:0, Demand bit:0
       myDiscrim:6, yourDiscrim:4, minTx:1000000, minRx:1000000, multiplier:3, minEchoRx:0
     00:54:38: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Up for
neighbor 2.2.2.2
on interface Te 4/24 (diag: 0)
```

The following example shows hexadecimal output from the `debug bfd packet` command.

```
RX packet dump:
    20 c0 03 18 00 00 00 05 00 00 00 04 00 01 86 a0
    00 01 86 a0 00 00 00 00
    00:34:13 : Sent packet for session with neighbor 2.2.2.2 on Te 4/24
TX packet dump:
    20 c0 03 18 00 00 00 04 00 00 00 05 00 01 86 a0
    00 01 86 a0 00 00 00 00
00:34:14 : Received packet for session with neighbor 2.2.2.2 on Te 4/24
RX packet dump:
    20 c0 03 18 00 00 00 05 00 00 00 04 00 01 86 a0
    00 01 86 a0 00 00 00 00
00:34:14 : Sent packet for session with neighbor 2.2.2.2 on Te 4/24
TX packet dump:
    20 c0 03 18 00 00 00 04 00 00 00 05 00 01 86 a0
    00 01 86 a0 00 00 00 00
00:34:14 : Received packet for session with neighbor 2.2.2.2 on Te 4/24
RX packet dump:
    20 c0 03 18 00 00 00 05 00 00 00 04 00 01 86 a0
    00 01 86 a0 00 00 00 00
00:34:14 : Sent packet for session with neighbor 2.2.2.2 on Te 4/24
L
```

The output for the `debug bfd event` command is the same as the log messages that appear on the console by default.

9

Border Gateway Protocol (BGP)

Border Gateway Protocol (BGP) is an interdomain routing protocol that manages routing between edge routers.

BGP uses an algorithm to exchange routing information between switches enabled with BGP. BGP determines a path to reach a particular destination using certain attributes while avoiding routing loops. BGP selects a single path as the best path to a destination network or host. You can also influence BGP to select different path by altering some of the BGP attributes.

Topics:

- BGP IP version 4 (BGPv4) Overview
- Basic BGP configuration tasks
- Advanced BGP configuration tasks

BGP IP version 4 (BGPv4) Overview

This section provides a general description of BGPv4 as it is supported in the Dell EMC Networking Operating System (OS).

BGP protocol standards are listed in the Standards Compliance chapter.

BGP is an external gateway protocol (EGP) that transmits interdomain routing information within and between autonomous systems (AS). An AS is an network managed by a single entity or administration. BGPv4 includes support for 4-byte autonomous system numbers and multiprotocol extensions to allow BGP to exchange information. BGP also support multiple Layer 3 protocol address families including IPv4 and IPv6. The primary function of the BGP is to exchange network reachability information with other BGP systems. BGP generally operates with an internal gateway protocol (IGP) such as open shortest path first (OSPF) or router information protocol (RIP), allowing you to communicate to external ASs smoothly. BGP adds reliability to network connections by having multiple paths from one router to another.

BGP uses a path-vector routing algorithm to communicate network information with other BGP configured routers. The network information is exchanged as updates between BGP peers, which contains information about the network, attributes and AS that is used to reach destination network. BGP selects a single path as a best path to a destination using best path selection algorithm.

BGP Autonomous Systems

BGP autonomous systems (ASs) are a collection of nodes under common administration with common network routing policies. Devices in the AS use IGP to communicate with one another. For devices in different AS to communicate, they need to use EGP. BGP is the EGP which allows the devices to communicate.

Each AS has a number, which an internet authority already assigns. You do not assign the BGP number.

AS numbers (ASNs) are important because the ASN uniquely identifies each network on the internet. The Internet Assigned Numbers Authority (IANA) has reserved AS numbers 64512 through 65534 to be used for private purposes. IANA reserves ASNs 0 and 65535 and must not be used in a live environment.

You can group autonomous systems into three categories (multihomed, stub, and transit), defined by their connections and operation.

- **multihomed AS** — is one that maintains connections to more than one other AS. This group allows the AS to remain connected to the Internet in the event of a complete failure of one of their connections. However, this type of AS does not allow traffic from one AS to pass through on its way to another AS. A simple example of this group is seen in the following illustration.
- **stub AS** — is one that is connected to only one other AS.
- **transit AS** — is one that provides connections through itself to separate networks. For example, in the following illustration, Router 1 can use Router 2 (the transit AS) to connect to Router 4. Internet service providers (ISPs) are always transit ASs, because they provide connections from one network to another. The ISP is considered to be “selling transit service” to the customer network, so thus the term Transit AS.

The devices within an AS (AS1 or AS2, as seen in the following illustration) exchange routing information using Internal BGP (IBGP), whereas the devices in different AS communicate using External BGP (EBGP). IBGP provides routers inside the AS with

the knowledge to reach routers external to the AS. EBGP routers exchange information with other EBGP routers as well as IBGP routers to maintain connectivity and accessibility.

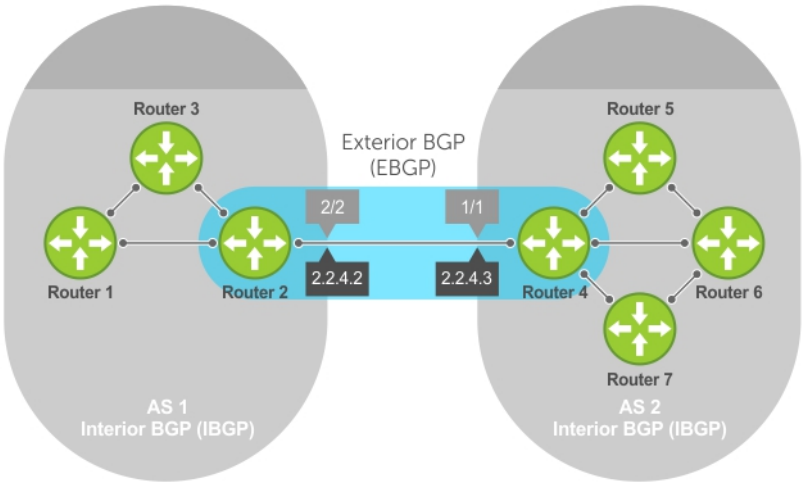


Figure 17. BGP Topology with autonomous systems (AS)

BGP version 4 (BGPv4) supports classless interdomain routing (CIDR) and aggregate routes and AS paths. BGP is a path vector protocol — a computer network in which BGP maintains the path that updated information takes as it diffuses through the network. Updates traveling through the network and returning to the same node are easily detected and discarded.

BGP does not use a traditional interior gateway protocol (IGP) matrix, but makes routing decisions based on path, network policies, and/or rulesets. Unlike most protocols, BGP uses TCP as its transport protocol.

Since each BGP router talking to another router is a session, an IBGP network needs to be in “full mesh.” This is a topology that has every router directly connected to every other router. Each BGP router within an AS must have IBGP sessions with all other BGP routers in the AS. For example, a BGP network within an AS needs to be in “full mesh.” As seen in the illustration below, four routers connected in a full mesh have three peers each, six routers have five peers each, and eight routers in full mesh have seven peers each.

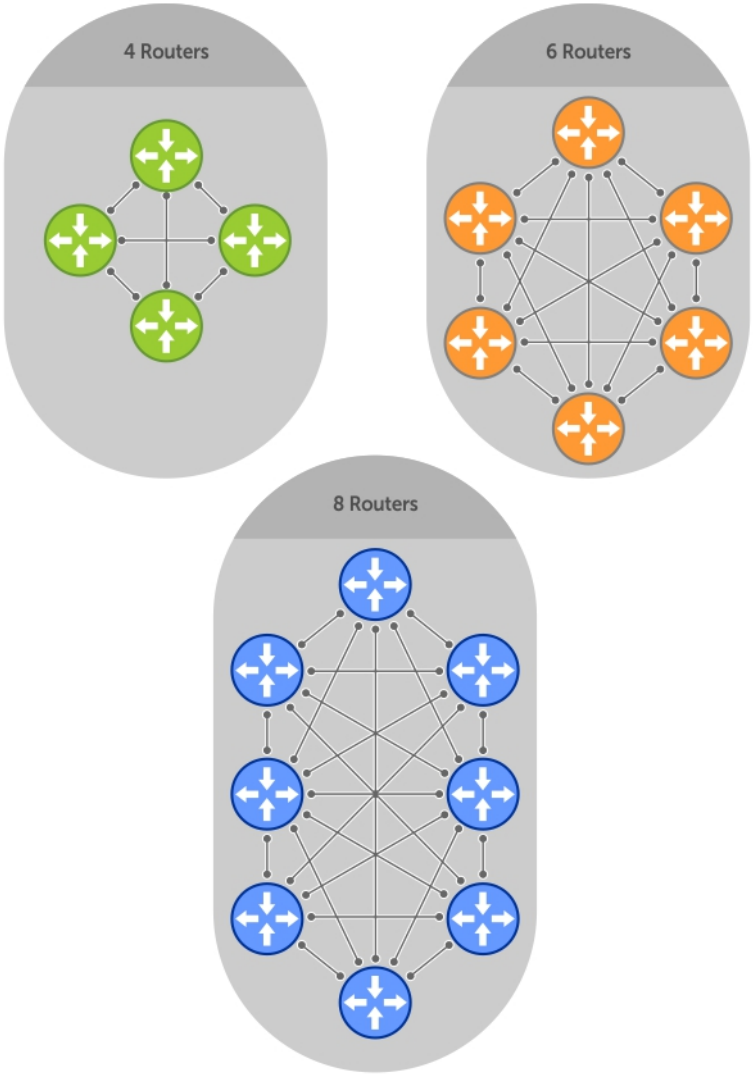


Figure 18. BGP Routers in Full Mesh

The number of BGP speakers each BGP peer must maintain increases exponentially. Network management quickly becomes impossible.

AS4 Number Representation

Dell EMC Networking OS supports multiple representations of 4-byte AS numbers: asplain, asdot+, and asdot.

NOTE: The ASDOT and ASDOT+ representations are supported only with the 4-Byte AS numbers feature. If 4-Byte AS numbers are not implemented, only ASPLAIN representation is supported.

ASPLAIN is the default method the system uses. With the ASPLAIN notation, a 32-bit binary AS number is translated into a decimal value.

- All AS numbers between 0 and 65535 are represented as a decimal number when entered in the CLI and when displayed in the `show` commands output.
- AS numbers larger than 65535 are represented using ASPLAIN notation. When entered in the CLI and when displayed in the `show` commands output, 65546 is represented as 65546.

ASDOT+ representation splits the full binary 4-byte AS number into two words of 16 bits separated by a decimal point (.): <high-order 16 bit value>.<low-order 16 bit value>. Some examples are shown in the following table.

- All AS numbers between 0 and 65535 are represented as a decimal number, when entered in the CLI and when displayed in the `show` commands outputs.
- AS Numbers larger than 65535 is represented using ASDOT notation as <higher 2 bytes in decimal>.<lower 2 bytes in decimal>. For example: AS 65546 is represented as 1.10.

ASDOT representation combines the ASPLAIN and ASDOT+ representations. AS numbers less than 65536 appear in integer format (asplain); AS numbers equal to or greater than 65536 appear in the decimal format (asdot+). For example, the AS number 65526 appears as 65526 and the AS number 65546 appears as 1.10.

Dynamic AS Number Notation Application

Dell EMC Networking OS applies the ASN notation type change dynamically to the running-config statements.

When you apply or change an notation, the type selected is reflected immediately in the running-configuration and the `show` commands (refer to the following two examples).

Example of Dynamic Changes in the Running Configuration When Using the `bgp asnotation` Command

```
ASDOT
DellEMC(conf-router_bgp)#bgp asnotation asdot
DellEMC(conf-router_bgp)#show conf
!
router bgp 100
bgp asnotation asdot
bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

DellEMC(conf-router_bgp)#do show ip bgp
BGP table version is 24901, local router ID is 172.30.1.57
<output truncated>

ASDOT+
DellEMC(conf-router_bgp)#bgp asnotation asdot+
DellEMC(conf-router_bgp)#show conf
!
router bgp 100
bgp asnotation asdot+
bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

DellEMC(conf-router_bgp)#do show ip bgp
BGP table version is 31571, local router ID is 172.30.1.57
<output truncated>

AS-PLAIN
DellEMC(conf-router_bgp)#bgp asnotation asplain
DellEMC(conf-router_bgp)#sho conf
!
router bgp 100
bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

DellEMC(conf-router_bgp)#do sho ip bgp
BGP table version is 34558, local router ID is 172.30.1.57
<output truncated>
```

Example of the Running Configuration When AS Notation is Disabled

```
AS NOTATION DISABLED
DellEMC(conf-router_bgp)#no bgp asnotation
DellEMC(conf-router_bgp)#sho conf
!
router bgp 100
bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>
```

```
DellEMC(conf-router_bgp)#do sho ip bgp
BGP table version is 28093, local router ID is 172.30.1.57
AS4 SUPPORT DISABLED
DellEMC(conf-router_bgp)#no bgp four-octet-as-support
DellEMC(conf-router_bgp)#sho conf
!
router bgp 100
neighbor 172.30.1.250 local-as 65057
DellEMC(conf-router_bgp)#do show ip bgp
BGP table version is 28093, local router ID is 172.30.1.57
```

Four-Byte AS Numbers

You can use the 4-Byte (32-bit) format when configuring autonomous system numbers (ASNs).

The 4-Byte support is advertised as a new BGP capability (4-BYTE-AS) in the OPEN message. If a 4-Byte BGP speaker has sent and received this capability from another speaker, all the messages will be 4-octet. The behavior of a 4-Byte BGP speaker is different with the peer depending on whether the peer is a 4-Byte or 2-Byte BGP speaker.

Where the 2-Byte format is 1-65535, the 4-Byte format is 1-4294967295. Enter AS numbers using the traditional format. If the ASN is greater than 65535, the dot format is shown when using the `show ip bgp` commands. For example, an ASN entered as 3183856184 appears in the `show` commands as 48581.51768; an ASN of 65123 is shown as 65123. To calculate the comparable dot format for an ASN from a traditional format, use ASN/65536. ASN%65536.

Traditional Format	DOT Format
65001	0.65501
65536	1.0
100000	1.34464
4294967295	65535.65535

When creating Confederations, all the routers in a Confederation must be either 4-Byte or 2-Byte identified routers. You cannot mix them.

Configure 4-byte AS numbers with the `bgp four-octet-as-support` command.

BGP router ID

A router ID is used to identify neighbors configured with BGP. The BGP router ID is a user given ID to a BGP router, which is unique 32-bit value IPv4 address. The router ID is the highest IP address of the Loopback interface and if no Loopback interfaces are configured, Dell EMC Networking OS selects the highest IP address of a physical interface on the router as the router ID.

Sessions and Peers

When two routers communicate using the BGP protocol, a BGP session is started. The two end-points of that session are called neighbors or peers. A local device does not identify a BGP peer automatically. You have to manually configure the connections between BGP running devices. BGP uses TCP connection to communicate with another BGP neighbor. After establishing a TCP connection between the peers, each peer exchange route information with other peer.

Establish a Session

Information exchange between peers is driven by events and timers. The focus in BGP is on the traffic routing policies.

In order to make decisions in its operations with other BGP peers, a BGP process uses a simple finite state machine that consists of six states: Idle, Connect, Active, OpenSent, OpenConfirm, and Established. For each peer-to-peer session, a BGP implementation tracks which of these six states the session is in. The BGP protocol defines the messages that each peer should exchange in order to change the session from one state to another.

State	Description
Idle	BGP initializes all resources, refuses all inbound BGP connection attempts, and initiates a TCP connection to the peer.
Connect	In this state the router waits for the TCP connection to complete, transitioning to the OpenSent state if successful. If that transition is not successful, BGP resets the ConnectRetry timer and transitions to the Active state when the timer expires.
Active	The router resets the ConnectRetry timer to zero and returns to the Connect state.
OpenSent	After successful OpenSent transition, the router sends an Open message and waits for one in return.
OpenConfirm	After the Open message parameters are agreed between peers, the neighbor relation is established and is in the OpenConfirm state. This is when the router receives and checks for agreement on the parameters of open messages to establish a session.
Established	Keepalive messages are exchanged next, and after successful receipt, the router is placed in the Established state. Keepalive messages continue to be sent at regular periods (established by the Keepalive timer) to verify connections.

After the connection is established, the router can now send/receive Keepalive, Update, and Notification messages to/from its peer.

Peer Groups

Peer groups are neighbors grouped according to common routing policies. They enable easier system configuration and management by allowing groups of routers to share and inherit policies.

Peer groups also aid in convergence speed. When a BGP process needs to send the same information to a large number of peers, the BGP process needs to set up a long output queue to get that information to all the proper peers. If the peers are members of a peer group however, the information can be sent to one place and then passed onto the peers within the group.

BGP Attributes for selecting Best Path

Routes learned using BGP have associated properties that are used to determine the best route to a destination when multiple paths exist to a particular destination.

These properties are referred to as BGP attributes, and an understanding of how BGP attributes influence route selection is required for the design of robust networks. This section describes the attributes that BGP uses in the route selection process:

- Weight
- Local Preference
- Multi-Exit Discriminators (MEDs)
- Origin
- AS Path
- Next Hop

NOTE: There are no hard coded limits on the number of attributes that are supported in the BGP. Taking into account other constraints such as the Packet Size, maximum number of attributes are supported in BGP.

`Communities`

BGP communities are sets of routes with one or more common attributes. Communities are a way to assign common attributes to multiple routes at the same time.

NOTE: Duplicate communities are not rejected.

Best Path Selection Criteria

Paths for active routes are grouped in ascending order according to their neighboring external AS number (BGP best path selection is deterministic by default, which means the `bgp non-deterministic-med` command is NOT applied).

The best path in each group is selected based on specific criteria. Only one "best path" is selected at a time. If any of the criteria results in more than one path, BGP moves on to the next option in the list. For example, two paths may have the same weights, but different local preferences. BGP sees that the Weight criteria results in two potential "best paths" and moves to local preference to reduce the options. If a number of best paths is determined, this selection criteria is applied to group's best to determine the ultimate best path.

In non-deterministic mode (the `bgp non-deterministic-med` command is applied), paths are compared in the order in which they arrive. This method can lead to Dell EMC Networking OS choosing different best paths from a set of paths, depending on the order in which they were received from the neighbors because MED may or may not get compared between the adjacent paths. In deterministic mode, Dell EMC Networking OS compares MED between the adjacent paths within an AS group because all paths in the AS group are from the same AS.

NOTE: The `bgp bestpath as-path multipath-relax` command is disabled by default, preventing BGP from load-balancing a learned route across two or more eBGP peers. To enable load-balancing across different eBGP peers, enable the `bgp bestpath as-path multipath-relax` command. A system error results if you configure the `bgp bestpath as-path ignore` command and the `bgp bestpath as-path multipath-relax` command at the same time. Only enable one command at a time.

The following illustration shows that the decisions BGP goes through to select the best path. The list following the illustration details the path selection criteria.

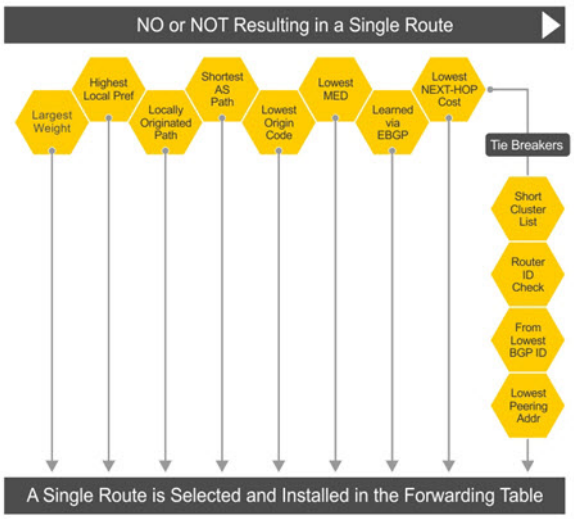


Figure 19. BGP Best Path Selection

Best Path Selection Details

1. Prefer the path with the largest WEIGHT attribute.
2. Prefer the path with the largest LOCAL_PREF attribute.
3. Prefer the path that was locally `Originated via a network` command, `redistribute` command or `aggregate-address` command.
   a. Routes originated with the `Originated via a network` or `redistribute` commands are preferred over routes originated with the `aggregate-address` command.
4. Prefer the path with the shortest AS_PATH (unless the `bgp bestpath as-path ignore` command is configured, then AS_PATH is not considered). The following criteria apply:
   a. An AS_SET has a path length of 1, no matter how many ASs are in the set.
   b. A path with no AS_PATH configured has a path length of 0.
   c. AS_CONFED_SET is not included in the AS_PATH length.
   d. AS_CONFED_SEQUENCE has a path length of 1, no matter how many ASs are in the AS_CONFED_SEQUENCE.
5. Prefer the path with the lowest ORIGIN type (IGP is lower than EGP, and EGP is lower than INCOMPLETE).
6. Prefer the path with the lowest multi-exit discriminator (MED) attribute. The following criteria apply:

a. This comparison is only done if the first (neighboring) AS is the same in the two paths; the MEDs are compared only if the first AS in the AS_SEQUENCE is the same for both paths.
b. If you entered the `bgp always-compare-med` command, MEDs are compared for all paths.
c. Paths with no MED are treated as "worst" and assigned a MED of 4294967295.

7. Prefer external (EBGP) to internal (IBGP) paths or confederation EBGP paths.
8. Prefer the path with the lowest IGP metric to the BGP if next-hop is selected when `synchronization` is disabled and only an internal path remains.
9. Dell EMC Networking OS deems the paths as equal and does not perform steps 9 through 11, if the following criteria is met:
   a. the IBGP multipath or EBGP multipath are configured (the `maximum-path` command).
   b. the paths being compared were received from the same AS with the same number of ASs in the AS Path but with different NextHops.
   c. the paths were received from IBGP or EBGP neighbor respectively.
10. If the `bgp bestpath router-id ignore` command is enabled and:
    a. if the Router-ID is the same for multiple paths (because the routes were received from the same route) skip this step.
    b. if the Router-ID is NOT the same for multiple paths, prefer the path that was first received as the Best Path. The path selection algorithm returns without performing any of the checks detailed here.
11. Prefer the external path originated from the BGP router with the lowest router ID. If both paths are external, prefer the oldest path (first received path). For paths containing a route reflector (RR) attribute, the originator ID is substituted for the router ID.
12. If two paths have the same router ID, prefer the path with the lowest cluster ID length. Paths without a cluster ID length are set to a 0 cluster ID length.
13. Prefer the path originated from the neighbor with the lowest address. (The neighbor address is used in the BGP neighbor configuration and corresponds to the remote peer used in the TCP connection with the local router.)

After a number of best paths is determined, this selection criteria is applied to group's best to determine the ultimate best path.

In non-deterministic mode (the `bgp non-deterministic-med` command is applied), paths are compared in the order in which they arrive. This method can lead to Dell EMC Networking OS choosing different best paths from a set of paths, depending on the order in which they were received from the neighbors because MED may or may not get compared between the adjacent paths. In deterministic mode, Dell EMC Networking OS compares MED between the adjacent paths within an AS group because all paths in the AS group are from the same AS.

Weight

The weight attribute is local to the router and is not advertised to neighboring routers.

If the router learns about more than one route to the same destination, the route with the highest weight is preferred. The route with the highest weight is installed in the IP routing table.

Local Preference

Local preference (LOCAL_PREF) represents the degree of preference within the entire AS. The higher the number, the greater the preference for the route.

Local preference (LOCAL_PREF) is one of the criteria used to determine the best path, so keep in mind that other criteria may impact selection, as shown in the illustration in Best Path Selection Criteria. For this example, assume that thelocal preference (LOCAL_PREF) is the only attribute applied. In the following illustration, AS100 has two possible paths to AS 200. Although the path through Router A is shorter (one hop instead of two), the LOCAL_PREF settings have the preferred path go through Router B and AS300. This is advertised to all routers within AS100, causing all BGP speakers to prefer the path through Router B.

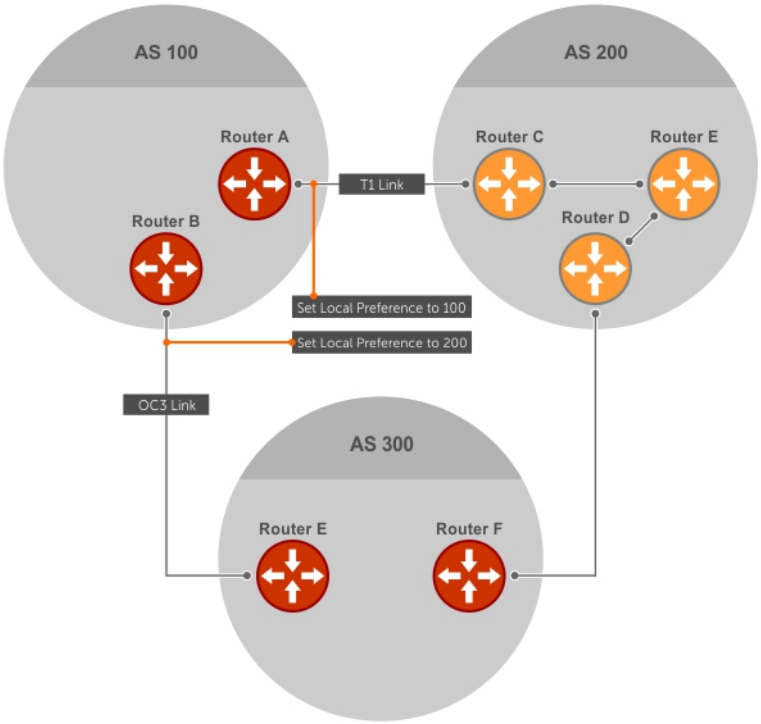


Figure 20. BGP Local Preference

Multi-Exit Discriminators (MEDs)

If two ASs connect in more than one place, a multi-exit discriminator (MED) can be used to assign a preference to a preferred path.

MED is one of the criteria used to determine the best path, so keep in mind that other criteria may impact selection, as shown in the illustration in Best Path Selection Criteria.

One AS assigns the MED a value and the other AS uses that value to decide the preferred path. For this example, assume the MED is the only attribute applied. In the following illustration, AS100 and AS200 connect in two places. Each connection is a BGP session. AS200 sets the MED for its T1 exit point to 100 and the MED for its OC3 exit point to 50. This sets up a path preference through the OC3 link. The MEDs are advertised to AS100 routers so they know which is the preferred path.

MEDs are non-transitive attributes. If AS100 sends an MED to AS200, AS200 does not pass it on to AS300 or AS400. The MED is a locally relevant attribute to the two participating ASs (AS100 and AS200).

NOTE: The MEDs are advertised across both links, so if a link goes down, AS 1 still has connectivity to AS300 and AS400.

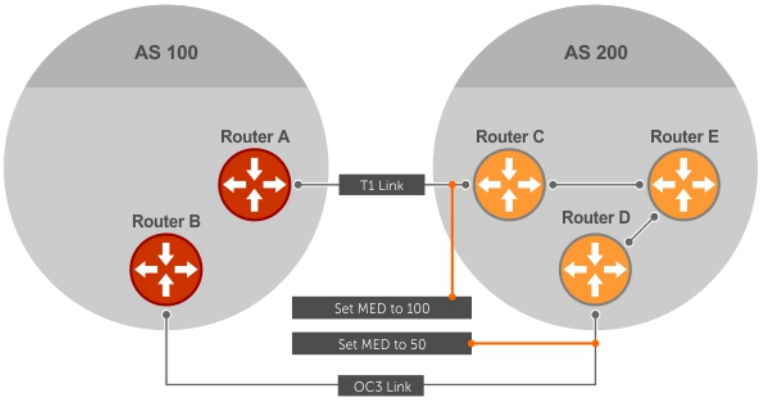


Figure 21. Multi-Exit Discriminators

NOTE: Configuring the `set metric-type internal` command in a route-map advertises the IGP cost as MED to outbound EBGP peers when redistributing routes. The configured `set metric` value overwrites the default IGP cost. If the outbound route-map uses MED, it overwrites IGP MED.

Origin

The origin indicates the origin of the prefix, or how the prefix came into BGP. There are three origin codes: IGP, EGP, INCOMPLETE.

Origin Type	Description
IGP	Indicates the prefix originated from information learned through an interior gateway protocol.
EGP	Indicates the prefix originated from information learned from an EGP protocol, which NGP replaced.
INCOMPLETE	Indicates that the prefix originated from an unknown source.

Generally, an IGP indicator means that the route was derived inside the originating AS. EGP generally means that a route was learned from an external gateway protocol. An INCOMPLETE origin code generally results from aggregation, redistribution, or other indirect ways of installing routes into BGP.

In Dell EMC Networking OS, these origin codes appear as shown in the following example. The question mark (?) indicates an origin code of INCOMPLETE (shown in bold). The lower case letter (i) indicates an origin code of IGP (shown in bold).

Example of Viewing Origin Codes

```
DellEMC#show ip bgp
BGP table version is 0, local router ID is 10.101.15.13
Status codes: s suppressed, d damped, h history, * valid, > best
Path source: I - internal, a - aggregate, c - confed-external, r - redistributed, n -
network
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network      Next Hop      Metric  LocPrf  Weight  Path
*> 7.0.0.0/29   10.114.8.33   0       0       18508   ?
*> 7.0.0.0/30   10.114.8.33   0       0       18508   ?
*> 9.2.0.0/16   10.114.8.33   10      0       18508   701 i
```

AS Path

The AS path is the list of all ASs that all the prefixes listed in the update have passed through.

The local AS number is added by the BGP speaker when advertising to a EBGP neighbor.

NOTE: Any update that contains the AS path number 0 is valid.

The AS path is shown in the following example. The origin attribute is shown following the AS path information (shown in bold).

Example of Viewing AS Paths

```
DellEMC#show ip bgp paths
Total 30655 Paths
Refcount Metric Path
 3        18508   701 3549 19421 i
 3        18508   701 7018 14990 i
 3        18508   209 4637 1221 9249 9249 i
 2        18508   701 17302 i
 26       18508   209 22291 i
 75       18508   209 3356 2529 i
 2        18508   209 1239 19265 i
 1        18508   701 2914 4713 17935 i
 162      18508   209 i
 2        18508   701 19878 ?
 31       18508   209 18756 i
 2        18508   209 7018 15227 i
 10       18508   209 3356 13845 i
 3        18508   209 701 6347 7781 i
 1        18508   701 3561 9116 21350 i
```

Next Hop

The next hop is the IP address used to reach the advertising router.

For EBGP neighbors, the next-hop address is the IP address of the connection between the neighbors. For IBGP, the EBGP next-hop address is carried into the local AS. A next hop attribute is set when a BGP speaker advertises itself to another BGP speaker outside its local AS and when advertising routes within an AS. The next hop attribute also serves as a way to direct traffic to another BGP speaker, rather than waiting for a speaker to advertise. When a next-hop BGP neighbor is unreachable, then the connection to that BGP neighbor goes down after hold down timer expiry. The connection flap can also be obtained immediately with Fallover enabled. BGP routes that contain the next-hop as the neighbor address are not sent to the neighbor. You can enable this feature using the `neighbor sender-side-loopdetect` command.

NOTE: For EBGP neighbors, the next-hop address corresponding to a BGP route is not resolved if the next-hop address is not the same as the neighbor IP address.

NOTE: The connection between a router and its next-hop BGP neighbor terminates immediately only if the router has received routes from the BGP neighbor in the past.

Multiprotocol BGP

MBGP is an extensison of BGP to carry routing information for multiple network-layer protocols such as IPv4 unicast and multicast, and IPv6 unicast. MBGP provides support for multiple network layer protocol address families that allows to define independent policy and peering configurations. MBGP carries different set of routes depending on the protocol. It carries routing information for IPv4 multicast and IPv6 unicast routes. Multiprotocol extensions for BGP (MBGP) is defined in IETF RFC 2858. MBGP allows different types of address families to be distributed in parallel.

MBGP allows information about the topology of the IP multicast-capable routers to be exchanged separately from the topology of normal IPv4 and IPv6 unicast routers. It allows a multicast routing topology different from the unicast routing topology.

MBGP uses either an IPv4 address configured on the interface (which is used to establish the IPv6 session) or a stable IPv4 address that is available in the box as the next-hop address. As a result, while advertising an IPv6 network, exchange of IPv4 routes does not lead to martian next-hop message logs.

NOTE: It is possible to configure BGP peers that exchange both unicast and multicast network layer reachability information (NLRI), but you cannot connect multiprotocol BGP with BGP. Therefore, you cannot redistribute multiprotocol BGP routes into BGP.

MBGP for IPv4 Multicast

PIM feature uses IPv4 multicast routing for data distribution. MBGP provides a link that is dedicated specific to multicast traffic. MBGP also allows a unicast routing apart from the multicast routing. A multicast routing protocol, such as PIM, uses both

the unicast and multicast BGP database to form a routing table for unicast and multicast. You can configure BGP peers that exchange both unicast and multicast Network Layer Reachability Information (NLRI) in which MBGP routes is redistributed into BGP. The default is IPv4 unicast.

IPv4 and IPv6 address family

The IPv4 address family configuration in Dell EMC Networking OS is used for identifying routing sessions for protocols that use IPv4 address. You can specify multicast within the IPv4 address family. The default of address family configuration is IPv4 unicast. You can configure the VRF instances for IPv4 address family configuration.

The IPv6 address family configuration is used for identifying routing sessions for protocols that use IPv6 address. You can specify unicast within the IPv6 address family. You can configure the VRF instances for IPv6 address family configuration.

BGP address family model

Dell EMC Networking OS supports BGP address family identifier (AFI) with MBGP to provide scalability for multiple AFI and subsequent address family identifier (SAFI) configurations. MBGP carries routing information for address families (IPv4 and IPv6) of multiple network-layer protocols. Each address family has a separated BGP database, so that you can configure BGP policy based on address family.

BGP global and address family configuration

Implementing BGP is divided into global and address family configuration.

BGP configuration command levels are grouped as high level and address family configuration. All independent commands are grouped at the beginning of the configuration and followed by separate sub commands specific to each address family.

Following is the order of BGP configuration:

- Global configuration — General configuration that is applied to BGP. For example, configuring BGP timer, Adding additional BGP paths, etc.
- Address family configuration — Configurations that are applied to a specific address family.

The BGP address family model consists of three address families such as IPv4 unicast, IPv4 multicast, and IPv6 unicast. The default address family is **IPv4 unicast**. The neighbors configured under CONFIGURATION-ROUTER-BGP mode is considered to be IPv4 unicast neighbor. The IPv4 routes are advertised or received irrespective of the address family configuration. To configure a neighbor for IPv4 multicast or IPv6 unicast, you have to create an appropriate address family and activate the neighbor in the respective address family configuration, as shown the following example configuration.

Example of BGP configuration command levels

Following is an example configuration, which explains the neighbor configuration for all the address families. Also, the configuration shows how to create address families (IPv4 and IPv6) and activate the neighbors in the address family.

```
DellEMC(conf)#router bgp 10
DellEMC(conf-router_bgp)#neighbor 20.20.20.1 remote-as 200
DellEMC(conf-router_bgp)#neighbor 20.20.20.1 no shutdown
DellEMC(conf-router_bgp)#neighbor 2001::1 remote-as 200
DellEMC(conf-router_bgp)#neighbor 2001::1 no shutdown
DellEMC(conf-router_bgp)#address-family ipv4 multicast
DellEMC(conf-router_bgp_af)#neighbor 20.20.20.1 activate
DellEMC(conf-router_bgp_af)#exit
DellEMC(conf-router_bgp)#address-family ipv6 unicast
DellEMC(conf-router_bgpv6_af)#neighbor 2001::1 activate
DellEMC(conf-router_bgpv6_af)#exit
```

BGP global configuration default values

By default, BGP is disabled. The following table displays the default values for BGP on Dell EMC Networking OS.

Table 8. BGP Default Values

Item	Default
BGP Neighbor Adjacency changes	All BGP neighbor changes are logged.
Fast External Fallover feature	Disabled
Graceful Restart feature	Disabled
Local preference	100
MED	0
Route Flap Damping Parameters	half-life = 15 minutes reuse = 750 suppress = 2000 max-suppress-time = 60 minutes
Distance	external distance = 20 internal distance = 200 local distance = 200
Timers	keepalive = 60 seconds holdtime = 180 seconds
Add-path	Disabled

Implement BGP with Dell EMC Networking OS

The following sections describe how to implement BGP on Dell EMC Networking OS.

Additional Path (Add-Path) Support

The add-path feature reduces convergence times by advertising multiple paths to its peers for the same address prefix without replacing existing paths with new ones. By default, a BGP speaker advertises only the best path to its peers for a given address prefix. If the best path becomes unavailable, the BGP speaker withdraws its path from its local RIB and recalculates a new best path. This situation requires both IGP and BGP convergence and can be a lengthy process. BGP add-path also helps switchover to the next new best path when the current best path is unavailable.

Advertise IGP Cost as MED for Redistributed Routes

When using multipath connectivity to an external AS, you can advertise the MED value selectively to each peer for redistributed routes. For some peers you can set the internal/IGP cost as the MED while setting others to a constant pre-defined metric as MED value.

Use the `set metric-type internal` command in a route-map to advertise the IGP cost as the MED to outbound EBGP peers when redistributing routes. The configured `set metric` value overwrites the default IGP cost.

By using the `redistribute` command with the `route-map` command, you can specify whether a peer advertises the standard MED or uses the IGP cost as the MED.

When configuring this functionality:

- If the `redistribute` command does not have `metric` configured and the BGP peer outbound route-map does have `metric-type internal` configured, BGP advertises the IGP cost as MED.

- If the `redistribute` command has `metric` configured (`route-map set metric` or `redistribute route-type metric`) and the BGP peer outbound route-map has `metric-type internal` configured, BGP advertises the metric configured in the `redistribute` command as MED.
- If BGP peer outbound route-map has `metric` configured, all other metrics are overwritten by this configuration.

NOTE: When redistributing static, connected, or OSPF routes, there is no `metric` option. Simply assign the appropriate route-map to the redistributed route.

The following table lists some examples of these rules.

Table 9. Redistributed Route Rules

Command Settings	BGP Local Routing Information Base	MED Advertised to Peer WITH route-map metric-type internal	MED Advertised to Peer WITHOUT route-map metric-type internal
redistribute isis (IGP cost = 20)	MED: IGP cost 20	MED = 20	MED = 0
redistribute isis route-map set metric 50	MED: IGP cost 50	MED: 50 MED: 50	MED: 50 MED: 50
redistribute isis metric 100	MED: IGP cost 100	MED: 100	MED: 100

Ignore Router-ID in Best-Path Calculation

You can avoid unnecessary BGP best-path transitions between external paths under certain conditions. The `bgp bestpath router-id ignore` command reduces network disruption caused by routing and forwarding plane changes and allows for faster convergence.

AS Number Migration

With this feature you can transparently change the AS number of an entire BGP network and ensure that the routes are propagated throughout the network while the migration is in progress.

When migrating one AS to another, perhaps combining ASs, an eBGP network may lose its routing to an iBGP if the ASN changes. Migration can be difficult as all the iBGP and eBGP peers of the migrating network must be updated to maintain network reachability. Essentially, Local-AS provides a capability to the BGP speaker to operate as if it belongs to "virtual" AS network besides its physical AS network.

The following illustration shows a scenario where Router A, Router B, and Router C belong to AS 100, 200, and 300, respectively. Router A acquired Router B; Router B has Router C as its customer. When Router B is migrating to Router A, it must maintain the connection with Router C without immediately updating Router C's configuration. Local-AS allows this behavior to happen by allowing Router B to appear as if it still belongs to Router B's old network (AS 200) as far as communicating with Router C is concerned.

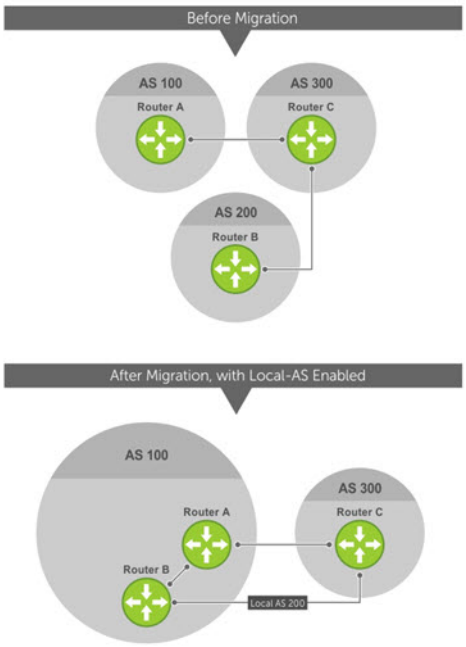


Figure 22. Before and After AS Number Migration with Local-AS Enabled

When you complete your migration, and you have reconfigured your network with the new information, disable this feature.

If you use the “no prepend” option, the Local-AS does not prepend to the updates received from the eBGP peer. If you do not select “no prepend” (the default), the Local-AS is added to the first AS segment in the AS-PATH. If an inbound route-map is used to prepend the as-path to the update from the peer, the Local-AS is added first. For example, consider the topology described in the previous illustration. If Router B has an inbound route-map applied on Router C to prepend "65001 65002" to the as-path, the following events take place on Router B:

1. Receive and validate the update.
2. Prepend local-as 200 to as-path.
3. Prepend "65001 65002" to as-path.

Local-AS is prepended before the route-map to give an impression that update passed through a router in AS 200 before it reached Router B.

BGP4 Management Information Base (MIB)

The FORCE10-BGP4-V2-MIB enhances support for BGP management information base (MIB) with many new simple network management protocol (SNMP) objects and notifications (traps) defined in *draft-ietf-idr-bgp4-mibv2-05*. To see these enhancements, download the MIB from the Dell website.

NOTE: For the *Force10-BGP4-V2-MIB* and other MIB documentation, refer to the Dell iSupport web page.

Important Points to Remember

- Because eBGP packets are not controlled by the ACL, packets from BGP neighbors cannot be blocked using the `deny ip` command.
- The *f10BgpM2AsPathTableEntry* table, *f10BgpM2AsPathSegmentIndex*, and *f10BgpM2AsPathElementIndex* are used to retrieve a particular ASN from the AS path. These indices are assigned to the AS segments and individual ASN in each segment starting from 0. For example, an AS path list of {200 300 400} 500 consists of two segments: {200 300 400} with segment index 0 and 500 with segment index 1. ASN 200, 300, and 400 are assigned 0, 1, and 2 element indices in that order.
- Unknown optional transitive attributes within a given path attribute (PA) are assigned indices in order. These indices correspond to the *f10BgpM2PathAttrUnknownIndex* field in the *f10BgpM2PathAttrUnknownEntry* table.
- Negotiation of multiple instances of the same capability is not supported. *F10BgpM2PeerCapAnnouncedIndex* and *f10BgpM2PeerCapReceivedIndex* are ignored in the peer capability lookup.

- Configure inbound BGP soft-reconfiguration on a peer for *f10BgpM2PrefixInPrefixesRejected* to display the number of prefixes filtered due to a policy. If you do enable `BGP soft-reconfig`, the denied prefixes are not accounted for.
- *F10BgpM2AdjRibsOutRoute* stores the pointer to the NLRI in the peer's Adj-Rib-Out.
- PA Index (*f10BgpM2PathAttrIndex* field in various tables) is used to retrieve specific attributes from the PA table. The Next-Hop, RR Cluster-list, and Originator ID attributes are not stored in the PA Table and cannot be retrieved using the `index passed in` command. These fields are not populated in *f10BgpM2PathAttrEntry*, *f10BgpM2PathAttrClusterEntry*, and *f10BgpM2PathAttrOriginatorIdEntry*.
- *F10BgpM2PathAttrUnknownEntry* contains the optional-transitive attribute details.
- Query for *f10BgpM2LinkLocalNextHopEntry* returns the default value for Link-local Next-hop.
- RFC 2545 and the *f10BgpM2Rfc2545Group* are not supported.
- An SNMP query displays up to 89 AS paths. A query for a larger AS path count displays as "..." at the end of the output.
- SNMP set for BGP is not supported. For all peer configuration tables (*f10BgpM2PeerConfigurationGroup*, *f10BgpM2PeerRouteReflectorCfgGroup*, and *f10BgpM2PeerAsConfederationCfgGroup*), an SNMP set operation returns an error. Only SNMP queries are supported. In addition, the *f10BgpM2CfgPeerError*, *f10BgpM2CfgPeerBgpPeerEntry*, and *f10BgpM2CfgPeerRowEntryStatus* fields are to hold the SNMP set status and are ignored in SNMP query.
- The AFI/SAFI is not used as an index to the *f10BgpM2PeerCountersEntry* table. The BGP peer's AFI/ SAFI (IPv4 Unicast or IPv6 Multicast) is used for various outbound counters. Counters corresponding to IPv4 Multicast cannot be queried.
- The *f10BgpM2[Cfg]PeerReflectorClient* field is populated based on the assumption that route-reflector clients are not in a full mesh if you enable `BGP client-2-client reflection` and that the BGP speaker acting as reflector advertises routes learned from one client to another client. If disabled, it is assumed that clients are in a full mesh and there is no need to advertise prefixes to the other clients.
- High CPU utilization may be observed during an SNMP walk of a large BGP Loc-RIB.
- To avoid SNMP timeouts with a large-scale configuration (large number of BGP neighbors and a large BGP Loc-RIB), Dell EMC Networking recommends setting the timeout and retry count values to a relatively higher number. For example, t = 60 or r = 5.
- To return all values on an snmpwalk for the *f10BgpM2Peer sub-OID*, use the `-C c` option, such as `snmpwalk -v 2c -C c -c public<IP_address><OID>`.
- An SNMP walk may terminate pre-maturely if the index does not increment lexicographically. Dell EMC Networking recommends using options to ignore such errors.
- Multiple BPG process instances are not supported. Thus, the *f10BgpM2PeerInstance* field in various tables is not used to locate a peer.
- Multiple instances of the same NLRI in the BGP RIB are not supported and are set to zero in the SNMP query response.
- The *f10BgpM2NlriIndex* and *f10BgpM2AdjRibsOutIndex* fields are not used.
- Carrying MPLS labels in BGP is not supported. The *f10BgpM2NlriOpaqueType* and *f10BgpM2NlriOpaquePointer* fields are set to zero.
- 4-byte ASN is supported. The *f10BgpM2AsPath4byteEntry* table contains 4-byte ASN-related parameters based on the configuration.
- If a received update route matches with a local prefix, then that route is discarded. This behavior results from an incorrect BGP configuration. To overcome this issue, you can trigger a route refresh after you properly configure BGP.
- If all the IP interfaces are in non-default VRFs, then you must have at least one interface in default VRF in order to configure a routing process that works with non-default VRFs.

Traps (notifications) specified in the BGP4 MIB draft `<draft-ietf-idr-bgp4-mibv2-05.txt>` are not supported. Such traps (*bgpM2Established* and *bgpM2BackwardTransition*) are supported as part of RFC 1657.

Configuration Information

The software supports BGPv4 as well as the following:

- deterministic multi-exit discriminator (MED) (default)
- a path with a missing MED is treated as worst path and assigned an MED value of (0xffffffff)
- the community format follows RFC 1998
- delayed configuration (the software at system boot reads the entire configuration file prior to sending messages to start BGP peer sessions)

The following are not yet supported:

- auto-summarization (the default is no auto-summary)
- synchronization (the default is no synchronization)

Basic BGP configuration tasks

The following sections describe how to configure a basic BGP network and the basic configuration tasks that are required for the BGP to be up and running.

Following are the basic configuration tasks required for BGP:

- Enabling BGP
- Configuring the router ID
- Configuring local AS number
- Configuring AS4 number representation
- Configuring a BGP peer
- Configuring a BGP peer group

Prerequisite for configuring a BGP network

You should be familiar with the overview of BGP before proceeding with configuring a basic BGP network. For more information about the BGP overview, see Border Gateway Protocol IPv4 (BGPv4) Overview.

Restrictions

Dell EMC Networking OS supports only one BGP routing configuration and autonomous system (AS), but supports multiple address family configuration.

Enabling BGP

By default, BGP is disabled on the system. Dell EMC Networking OS supports one autonomous system (AS) and assigns the AS number (ASN). To enable the BGP process and begin exchanging information, assign an AS number and use commands in ROUTER BGP mode to configure a BGP neighbor.

To establish BGP sessions and route traffic, configure at least one BGP neighbor or peer.

In BGP, routers with an established TCP connection are called neighbors or peers. After a connection is established, the neighbors exchange full BGP routing tables with incremental updates afterward. In addition, neighbors exchange KEEPALIVE messages to maintain the connection.

In BGP, neighbor routers or peers can be classified as internal or external. External BGP peers must be connected physically to one another (unless you enable the EBGP multihop feature), while internal BGP peers do not need to be directly connected. The IP address of an EBGP neighbor is usually the IP address of the interface directly connected to the router. First, the BGP process determines if all internal BGP peers are reachable, then it determines which peers outside the AS are reachable.

Following is the sample configuration steps to enable BGP, configure a BGP router-id and network for a router. The same configurations have to be repeated with appropriate changes in the IP addresses for a peer or router to achieve BGP session between two devices. In the below configuration example, no address family is configured. So, the routing information for the IPv4 unicast address family is advertised by default.

1. Assign an AS number and enter ROUTER BGP mode.

   CONFIGURATION mode

   ```
   router bgp as-number
   ```

   - *as-number*: from 0 to 65535 (2 Byte) or from 1 to 4294967295 (4 Byte) or 0.1 to 65535.65535 (Dotted format).

   Only one AS is supported per system.

   NOTE: If you enter a 4-Byte AS number, 4-Byte AS support is enabled automatically.

2. Add a BGP neighbor or peer and AS number.

   CONFIG-ROUTER-BGP mode

`neighbor {ip-address | ipv6-address| peer-group name} remote-as as-number`

- *ip-address*: IPv4 address of the neighbor
- *ipv6-address*: IPv6 address of the neighbor
- *peer-group name*: Name of the peer group. It can contain 16 characters.
- *as-number*: Autonomous number

NOTE: Neighbors that are defined using the `neighbor remote-as` command in the CONFIGURATION-ROUTER-BGP mode exchange IPv4 unicast address prefixes only.

3. Enable the BGP neighbor.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} no shutdown`

NOTE: When you change the configuration of a BGP neighbor, always reset it by entering the `clear ip bgp *` command in EXEC Privilege mode.

To view the BGP configuration, enter `show config` in CONFIGURATION ROUTER BGP mode. To view the BGP status, use the `show ip bgp summary` command in EXEC Privilege mode. The example shows that the summary with a 2-byte AS number using the `show ip bgp summary` command.

Following is the configuration steps:

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 65535
DellEMC(conf-router_bgp)# neighbor 20.20.20.1 remote-as 20
DellEMC(conf-router_bgp)# neighbor 20.20.20.1 no shutdown
DellEMC(conf-router_bgp)#exit
DellEMC(conf)#
```

The following example shows verifying the BGP configuration using the `show running-config bgp` command..

```
DellEMC#show running-config bgp
!
router bgp 65535
 neighbor 20.20.20.1 remote-as 20
 neighbor 20.20.20.1 no shutdown
DellEMC#
```

Examples of the `show ip bgp` Commands

The following example shows the `show ip bgp summary` command output.

```
R2#show ip bgp summary
BGP router identifier 1.1.1.1, local 65535
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 40960 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      20                 0        0          0    0     0 00:00:00 0
```

For the router's identifier, Dell EMC Networking OS uses the highest IP address of the Loopback interfaces configured. Because Loopback interfaces are virtual, they cannot go down, thus preventing changes in the router ID. If you do not configure Loopback interfaces, the highest IP address of any interface is used as the router ID.

To view the status of BGP neighbors, use the `show ip bgp neighbors` command in EXEC Privilege mode as shown in the first example. For BGP neighbor configuration information, use the `show running-config bgp` command in EXEC Privilege mode as shown in the second example.

NOTE: The `showconfig` command in CONFIGURATION ROUTER BGP mode gives the same information as the `show running-config bgp` command.

The following example displays two neighbors: one is an external internal BGP neighbor and the second one is an internal BGP neighbor. The first line of the output for each neighbor displays the AS number and states whether the link is an external or internal (shown in bold).

The third line of the `show ip bgp neighbors` output contains the BGP State. If anything other than ESTABLISHED is listed, the neighbor is not exchanging information and routes. For more information about using the `show ip bgp neighbors` command, refer to the *Dell EMC Networking OS Command Line Interface Reference Guide*.

The following example shows the `show ip bgp neighbors` command output.

```
DellEMC#show ip bgp neighbors
BGP neighbor is 20.20.20.1, remote AS 20, external link
  BGP remote router ID 1.1.1.1
  BGP state ESTABLISHED, in this state for 00:05:22
  Last read 00:00:00, Last write 00:00:07
  Hold time is 90, keepalive interval is 30 seconds
  Received 18 messages, 0 in queue
     7 opens, 6 notifications, 0 updates
     5 keepalives, 0 route refresh requests
  Sent 26 messages, 0 in queue
     7 opens, 0 notifications, 0 updates
     19 keepalives, 0 route refresh requests

  Route refresh request: received 0, sent messages 0
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

   For address family: IPv4 Unicast
  BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
  InQ : Added 0, Replaced 0, Withdrawn 0
  OutQ : Added 0, Withdrawn 0
  Allow local AS number 0 times in AS-PATH attribute
  Prefixes accepted 0, withdrawn 0 by peer, martian prefixes ignored 0
  Prefixes advertised 0, denied 0, withdrawn 0 from peer

  Connections established 0; dropped 0
  Last reset never
        No active TCP connection
```

Enabling four-byte autonomous system numbers

You can enable 4-byte support for configuring autonomous system numbers (ASN).

To enable 4-byte support for the BGP process, use the following command.

NOTE: When creating BGP confederations, all the routers in the Confederation must be a 4-byte or 2-byte identified routers. You cannot mix them.

- Enable 4-byte support for the BGP process.
  CONFIG-ROUTER-BGP mode
  `bgp four-octect-as-support`

  NOTE: If you enter a 4-byte AS number, 4-byte AS support is enabled automatically. This command is optional. Use it only if you support 4-Byte AS numbers or if you support AS4 number representation.

  Disable 4-Byte support and return to the default 2-Byte format by using the `no bgp four-octet-as-support` command. You cannot disable 4-Byte support if you currently have a 4-Byte ASN configured.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 80000
DellEMC(conf-router_bgp)# bgp four-octet-as-support
DellEMC(conf-router_bgp)# neighbor 20.20.20.1 remote-as 200
DellEMC(conf-router_bgp)# neighbor 20.20.20.1 no shutdown
DellEMC(conf-router_bgp)#exit
```

To view BGP configuration, use the `show running-config bgp` command in EXEC Privilege mode.

The following example shows the `show ip bgp summary` command output (4–byte AS number displays).

```
R2#show ip bgp summary
BGP router identifier 1.1.1.1, local 80000
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
```

```
1 neighbor(s) using 40960 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      200                 0        0          0    0     0 00:00:00 0
```

Changing a BGP router ID

BGP uses the configured router ID to identify the devices in the network.

By default, the router ID is the highest IP address of the Loopback interface. If no Loopback interfaces are configured, the highest IP address of a physical interface on the router is used as the BGP router ID. To change the default BGP router ID, use the following command.

- Change the BGP router ID of a BGP router

  ROUTER BGP mode

  `bgp router-id` *`ip-address`*

  `ip-address`- IP address in dotted decimal format.

  NOTE: Do not specify an IP address which is used by another device as a router ID.

  Peering sessions are reset when you change the router ID of a BGP router. Upon changing the router ID, the system automatically restarts the BGP instance for the configuration to take effect.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# bgp router-id 1.1.1.1
```

Following is the sample output of `show ip bgp ipv4 multicast summary` command.

```
DellEMC# show ip bgp summary
BGP router identifier 1.1.1.1, local AS number 400
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 8192 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      200                 0        0          0    0     0 00:00:00 0
DellEMC#
```

Configuring AS4 Number Representations

Enable one type of AS number representation: ASPLAIN, ASDOT+, or ASDOT.

Term	Description
ASPLAIN	Default method for AS number representation. With the ASPLAIN notation, a 32–bit binary AS number is translated into a decimal value.
ASDOT+	A representation splits the full binary 4-byte AS number into two words of 16 bits separated by a decimal point (.): <high-order 16 bit value>.<low-order 16 bit value>.
ASDOT	A representation combines the ASPLAIN and ASDOT+ representations. AS numbers less than 65536 appear in integer format (asplain); AS numbers equal to or greater than 65536 appear using the decimal method (asdot+). For example, the AS number 65526 appears as 65526 and the AS number 65546 appears as 1.10.

NOTE: The ASDOT and ASDOT+ representations are supported only with the 4-Byte AS numbers feature. If you do not implement 4-Byte AS numbers, only ASPLAIN representation is supported.

Only one form of AS number representation is supported at a time. You cannot combine the types of representations within an AS.

To configure AS4 number representations, use the following commands.

- Enable ASPLAIN AS Number representation.

  CONFIG-ROUTER-BGP mode

`bgp asnotation asplain`

NOTE: ASPLAIN is the default method Dell EMC Networking OS uses and does not appear in the configuration display.

- Enable ASDOT AS Number representation.
  CONFIG-ROUTER-BGP mode
  `bgp asnotation asdot`
- Enable ASDOT+ AS Number representation.
  CONFIG-ROUTER-BGP mode
  `bgp asnotation asdot+`

The following example shows the `bgp asnotation asplain` command output.

```
DellEMC(conf-router_bgp)#bgp asnotation asplain
DellEMC(conf-router_bgp)#sho conf
!
router bgp 100
  bgp four-octet-as-support
  neighbor 172.30.1.250 remote-as 18508
  neighbor 172.30.1.250 local-as 65057
  neighbor 172.30.1.250 route-map rmap1 in
  neighbor 172.30.1.250 password 7 5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
  neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

The following example shows the `bgp asnotation asdot` command output.

```
DellEMC(conf-router_bgp)#bgp asnotation asdot
DellEMC(conf-router_bgp)#sho conf
!
router bgp 100
  bgp asnotation asdot
bgp four-octet-as-support
  neighbor 172.30.1.250 remote-as 18508
  neighbor 172.30.1.250 local-as 65057
  neighbor 172.30.1.250 route-map rmap1 in
  neighbor 172.30.1.250 password 7 5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
  neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

The following example shows the `bgp asnotation asdot+` command output.

```
DellEMC(conf-router_bgp)#bgp asnotation asdot+
DellEMC(conf-router_bgp)#sho conf
!
router bgp 100
  bgp asnotation asdot+
bgp four-octet-as-support
  neighbor 172.30.1.250 remote-as 18508
  neighbor 172.30.1.250 local-as 65057
  neighbor 172.30.1.250 route-map rmap1 in
  neighbor 172.30.1.250 password 7 5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
  neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Configuring a BGP peer

To configure a BGP neighbor or peer, you must provide the IP address and the AS number of each neighbor since BGP does not discover the neighbor or peer. Neighbors that are present with the same AS communicate using IBGP while the neighbors that are present in different AS communicate using EBGP. To allow the neighbor to establish a BGP session, you have to enable the neighbor by providing `neighbor no shutdown` command. To configure BGP between two peers, use the following commands.

Following are the steps to configure a BGP peer. The following example configuration demonstrates how the configure a neighbor and IPv4 multicast address family. The default is IPv4 unicast address family.

- Enter the router configuration mode and the AS number.

CONFIG mode

`router bgp as-number`

- Add the IP address of the neighbor for the specified autonomous system.

  CONFIG-ROUTER-BGP mode

  `neighbor {ip-address | ipv6-address | peer-group-name} remote-as as-number`
- Enable the neighbor.

  CONFIG-ROUTERBGP mode

  `neighbor ip-address | ipv6-address | peer-group-name no shutdown`
- Specify the IPv4 address family configuration.

  CONFIG-ROUTER-BGP mode

  `address-family ipv4 [multicast | vrf vrf-name]`

  `multicast` — Specifies the IPv4 multicast address family.

  `vrf vrf-name` — Specifies the name of VRF instance associated with the IPv4 address-family configuration.
- Enable the neighbor to exchange prefixes for IPv4 unicast address family.

  CONFIG-ROUTER-BGP-AF (Address Family) mode

  `neighbor {ip-address | ipv6-address | peer-group-name} activate`

  NOTE: Neighbors have to be activated using `neighbor activate` command in the respective address family. To exchange other address prefix types (IPv4 multicast or IPv6 unicast), the neighbors must be activated under the respective address family configuration such as `address-family ipv4 multicast` (for IPv4 multicast) and`address-family ipv6 unicast`(for IPv6).

```
DellEMC(conf)# router bgp 10
DellEMC(conf-router_bgp)# neighbor 20.20.20.1 remote-as 200
DellEMC(conf-router_bgp)# neighbor 20.20.20.1 no shutdown
DellEMC(conf-router_bgp)# address-family ipv4 multicast
DellEMC(conf-router_bgp_af)# neighbor 20.20.20.1 activate
DellEMC(conf-router_bgp_af)# exit
```

Following is the sample output of `show ip bgp ipv4 multicast summary` command.

```
DellEMC# show ip bgp ipv4 multicast summary
BGP router identifier 1.1.1.1, local AS number 10
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 8192 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      200                 0        0          0    0     0 00:00:00 0
DellEMC#
```

Example-Configuring BGP routing between peers

The following example show how BGP is configured between two peers.

Figure 23. BGP topology with two AS

To support your own IP addresses, interfaces, names, and so on, you can copy and paste from these examples to your CLI. Be sure that you make the necessary changes.

Example-Configuring BGP routing between peers

Example of enabling BGP in Router A

Following is an example to enable BGP configuration in the router A.

```
RouterA# configure terminal
RouterA(conf)# router bgp 40000
RouterA(conf-router_bgp)# bgp router-id 10.1.1.99
RouterA(conf-router_bgp)# timers bgp 80 130
RouterA(conf-router_bgp)# neighbor 192.168.1.1 remote-as 45000
RouterA(conf-router_bgp)# exit
RouterA(conf)#
```

The neighbors configured under CONFIGURATION-ROUTER-BGP mode and are defined using the `neighbor remote-as` command exchange only IPv4 unicast address prefixes.

Example of enabling BGP in Router B

Following is an example to enable BGP configuration in the router B.

```
RouterB# configure terminal
RouterB(conf)# router bgp 45000
RouterB(conf-router_bgp)# bgp router-id 172.17.1.99
RouterB(conf-router_bgp)# timers bgp 70 120
RouterB(conf-router_bgp)# neighbor 192.168.1.2 remote-as 40000
RouterB(conf-router_bgp)# exit
RouterB(conf)#
```

The `show ip bgp summary` displays BGP configuration. Following is the sample output for `show ip bgp summary` command for router A.

```
RouterA#show ip bgp summary
BGP router identifier 10.1.1.99, local 40000
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 40960 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.1.1     45000              10       20          0    0     0 00:06:11 0
```

Following is the sample output for `show ip bgp summary` command for Router B.

```
RouterB#show ip bgp summary
BGP router identifier 172.17.1.99, local 45000
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 40960 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.1.2     40000             10       20          0    0     0 00:06:11 0
```

BGP peer group

To configure multiple BGP neighbors at one time, create and populate a BGP peer group.

If you want to configure multiple BGP neighbors with similar policies (route maps, distribute lists, and so on), you can group the respective neighbors under a peer group. A peer group contains a set of BGP neighbors which share common configurations. You can group neighbors that share configurations in a BGP peer group to simplify neighbor configuration. An advantage of peer groups is that members of a peer group inherit the configuration properties of the group and share same update policy.

Dell EMC Networking OS allows a maximum of 256 peer groups.

Tasks involved in configuring a BGP peer group:

- Create a peer group by assigning it a name
- Adding members (neighbors) to the peer group

Configuration rules in a peer group:

- You must create a peer group first before adding the neighbors in the peer group.
- If you remove any configuration parameters from a peer group, it will apply to all the neighbors configured under that peer group.
- If you have not configured a parameter for an individual neighbor in the peer group, the neighbor uses the value configured in the peer group.
- If you reset any parameter for an individual neighbor, it will override the value set in the peer group.
- If you add a configuration parameters to a peer group that already has neighbors, the parameter value is applied to all the neighbors which does not have the value set. If a neighbor has a value explicitly set, the value overrides the value set for the peer group.
- You cannot remove the remote AS configuration from the peer group, if the peer group has neighbors.
- If you remove any parameter value in a peer group, the value of the parameter is set to the default for all the neighbors in the peer group, unless there is an explicit parameter set for an individual neighbor.

After you create a peer group, you can configure route policies for it. For information about configuring route policies for a peer group, refer to Filtering BGP Routes. See Example-Configuring BGP peer groups for configuring multiple BGP neighbors and enabling peer groups.

Configuring Peer Groups

To configure a peer group, use the following commands.

1. Enter the router configuration mode and the AS number.
   CONFIG mode
   `router bgp as-number`
2. Create a peer group by assigning a name to it.
   CONFIG-ROUTER-BGP mode
   `neighbor peer-group-name peer-group`
3. Enable the peer group.
   CONFIG-ROUTER-BGP mode
   `neighbor peer-group-name no shutdown`
   By default, all peer groups are disabled.
4. Create a BGP neighbor.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address} remote-as as-number`
5. Enable the neighbor.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address} no shutdown`
   NOTE: You can use `neighbor shutdown` command to disable a BGP neighbor or a peer group.
6. Add an enabled neighbor to the peer group.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} peer-group peer-group-name`
7. Add a neighbor as a remote AS.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} remote-as as-number`
   - *peer-group Name*: 16 characters.
   - *as-number*: the range is from 0 to 65535 (2-Byte) or 1 to 4294967295 | 0.1 to 65535.65535 (4-Byte) or 0.1 to 65535.65535 (Dotted format)

   To add an external BGP (EBGP) neighbor, configure the `as-number` parameter with a number different from the BGP as-number configured in the `router bgp as-number` command.

To add an internal BGP (IBGP) neighbor, configure the `as-number` parameter with the same BGP as-number configured in the `router bgp as-number` command.

After you create a peer group, you can use any of the commands beginning with the keyword `neighbor` to configure that peer group.

When you add a peer to a peer group, it inherits all the peer group's configured parameters.

A neighbor cannot become part of a peer group if it has any of the following commands configured:

- `neighbor advertisement-interval`
- `neighbor distribute-list out`
- `neighbor filter-list out`
- `neighbor next-hop-self`
- `neighbor route-map out`
- `neighbor route-reflector-client`
- `neighbor send-community`

A neighbor may keep its configuration after it was added to a peer group if the neighbor's configuration is more specific than the peer group's and if the neighbor's configuration does not affect outgoing updates.

NOTE: When you configure a new set of BGP policies for a peer group, *always* reset the peer group by entering the `clear ip bgp peer-group` *`peer-group-name`* command in EXEC Privilege mode.

To view the configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode.

When you create a peer group, it is disabled (shutdown). The following example shows the creation of a peer group (zanzibar) (in bold). To enable a peer group, use the `neighbor` *`peer-group-name`* `no shutdown` command.

```
DellEMC(conf)# router bgp 45
DellEMC(conf-router_bgp)# neighbor zanzibar peer-group
DellEMC(conf-router_bgp)# neighbor zanzibar no shutdown
DellEMC(conf-router_bgp)# neighbor 10.1.1.1 remote-as 65535
DellEMC(conf-router_bgp)# neighbor 10.1.1.1 no shutdown
DellEMC(conf-router_bgp)# neighbor 10.14.8.60 remote-as 18505
DellEMC(conf-router_bgp)# neighbor 10.14.8.60 no shutdown
DellEMC(conf-router_bgp)# neighbor 2001::1 peer-group zanzibar
```

Following is the `show config` output for the above sample configuration.

```
DellEMC# show config
!
router bgp 45
  neighbor zanzibar peer-group
  neighbor zanzibar no shutdown
  neighbor 10.1.1.1 remote-as 65535
  neighbor 10.1.1.1 shutdown
  neighbor 10.14.8.60 remote-as 18505
  neighbor 10.14.8.60 no shutdown
  neighbor 2001::1 peer-group zanzibar
DellEMC#
```

To disable a peer group, use the `neighbor` *`peer-group-name`* `shutdown` command in CONFIGURATION-ROUTER-BGP mode. The configuration of the peer group is maintained, but it is not applied to the peer group members. When you disable a peer group, all the peers within the peer group that are in the ESTABLISHED state move to the IDLE state.

To view the status of peer groups, use the `show ip bgp peer-group` command in EXEC Privilege mode, as shown in the following example.

```
DellEMC>show ip bgp peer-group
Peer-group zanzibar
BGP version 4
Minimum time between advertisement runs is 5 seconds

For address family: IPv4 Unicast
BGP neighbor is zanzibar, peer-group internal,
Number of peers in this group 1
Peer-group members (* - outbound optimized):
  2001::1
```

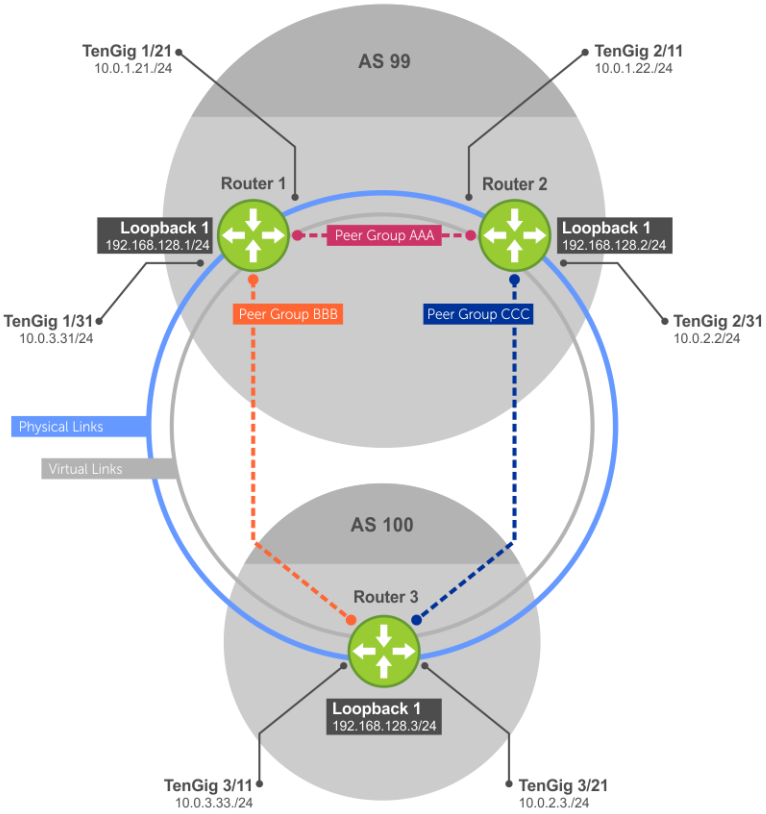
Example-Configuring BGP peer groups

The following example configurations show how to enable BGP and set up some peer groups. These examples are not comprehensive directions. They are intended to give you some guidance with typical configurations.

To support your own IP addresses, interfaces, names, and so on, you can copy and paste from these examples to your CLI. Be sure that you make the necessary changes.

The following illustration shows the configurations described on the following examples. These configurations show how to create BGP areas using physical and virtual links. They include setting up the interfaces and peers groups with each other.

Figure 24. BGP peer group example configurations



Example of Enabling BGP (Router 1)

```
R1# conf
R1(conf)#int loop 0
R1(conf-if-lo-0)#ip address 192.168.128.1/32
R1(conf-if-lo-0)#no shutdown
R1(conf-if-lo-0)#show config
!
interface Loopback 0
ip address 192.168.128.1/24
no shutdown
R1(conf-if-lo-0)#int gi 1/21
R1(conf-if-gi-1/21)#ip address 10.0.1.21/24
R1(conf-if-gi-1/21)#no shutdown
R1(conf-if-gi-1/21)#show config
!
interface GigabitEthernet 1/21
ip address 10.0.1.21/24
no shutdown
R1(conf-if-gi-1/21)#int gi 1/31
R1(conf-if-gi-1/31)#ip address 10.0.3.31/24
R1(conf-if-gi-1/31)#no shutdown
```

```
R1(conf-if-gi-1/31)#show config
!
interface GigabitEthernet 1/31
ip address 10.0.3.31/24
no shutdown
R1(conf-if-gi-1/31)#exit
R1(conf)#ip route 192.168.128.2/32 10.0.1.22
R1(conf)#router bgp 99
R1(conf-router_bgp)#neighbor 192.168.128.2 remote 99
R1(conf-router_bgp)#neighbor 192.168.128.2 no shut
R1(conf-router_bgp)#neighbor 192.168.128.2 update-source loop 0
R1(conf-router_bgp)#neighbor 10.0.3.33 remote 100
R1(conf-router_bgp)#neighbor 10.0.3.33 no shut
R1(conf-router_bgp)#show config
!
router bgp 99
network 192.168.128.0/24
neighbor 192.168.128.2 remote-as 99
neighbor 192.168.128.2 update-source Loopback 0
neighbor 10.0.3.33 no shutdown
neighbor 10.0.3.33 remote-as 100
```

Example of Enabling BGP (Router 2)

```
R2# conf
R2(conf)#int loop 0
R2(conf-if-lo-0)#ip address 192.168.128.2/32
R2(conf-if-lo-0)#no shutdown
R2(conf-if-lo-0)#show config
!
interface Loopback 0
ip address 192.168.128.2/24
no shutdown
R2(conf-if-lo-0)#int te 2/11
R2(conf-if-te-2/11)#ip address 10.0.1.22/24
R2(conf-if-te-2/11)#no shutdown
R2(conf-if-te-2/11)#show config
!
interface TengigabitEthernet 2/11
ip address 10.0.1.22/24
no shutdown
R2(conf-if-te-2/11)#int te 2/31
R2(conf-if-te-2/31)#ip address 10.0.2.2/24
R2(conf-if-te-2/31)#no shutdown
R2(conf-if-te-2/31)#show config
!
interface TengigabitEthernet 2/31
ip address 10.0.2.2/24
no shutdown
R2(conf-if-te-2/31)#exit
R2(conf)#ip route 192.168.1.1/32 10.0.1.21
R2(conf)#
R2(conf-if-te-2/31)#router bgp 99
R2(conf-router_bgp)#neighbor 192.168.128.1 remote 99
R2(conf-router_bgp)#neighbor 192.168.128.1 no shut
R2(conf-router_bgp)#neighbor 192.168.128.1 update-source loop 0
R2(conf-router_bgp)#neighbor 10.0.2.3 remote 100
R2(conf-router_bgp)#neighbor 10.0.2.3 no shut
R2(conf-router_bgp)#show config
!
router bgp 99
neighbor 192.168.128.1 remote 99
neighbor 192.168.128.1 no shut
neighbor 192.168.128.1 update-source loop 0
neighbor 10.0.2.3 remote 100
neighbor 10.0.2.3 no shut
```

Example of Enabling BGP (Router 3)

```
R3#conf
R3(conf)#int te 3/11
R3(conf-if-te-3/11)#ip address 10.0.3.33/24
```

```
R3(conf-if-te-3/11)#no shutdown
R3(conf-if-te-3/11)#show config
!
interface TengigabitEthernet 3/11
ip address 10.0.3.33/24
no shutdown
R3(conf-if-te-3/11)#int te 3/21
R3(conf-if-te-3/21)#ip address 10.0.2.3/24
R3(conf-if-te-3/21)#no shutdown
R3(conf-if-te-3/21)#show config
!
interface TengigabitEthernet 3/21
ip address 10.0.2.3/24
no shutdown
R3(conf-if-te-3/21)#
R3(conf-if-te-3/21)#router bgp 100
R3(conf-router_bgp)#show config
!
router bgp 100
R3(conf-router_bgp)#neighbor 10.0.3.31 remote 99
R3(conf-router_bgp)#neighbor 10.0.3.31 no shut
R3(conf-router_bgp)#neighbor 10.0.2.2 remote 99
R3(conf-router_bgp)#neighbor 10.0.2.2 no shut
R3(conf-router_bgp)#show config
!
router bgp 100
neighbor 10.0.3.31 remote 99
neighbor 10.0.3.31 no shut
neighbor 10.0.2.2 remote 99
neighbor 10.0.2.2 no shut
```

Example of Enabling Peer Groups (Router 1)

```
conf
R1(conf)#router bgp 99
R1(conf-router_bgp)# network 192.168.128.0/24
R1(conf-router_bgp)# neighbor AAA peer-group
R1(conf-router_bgp)# neighbor AAA no shutdown
R1(conf-router_bgp)# neighbor BBB peer-group
R1(conf-router_bgp)# neighbor BBB no shutdown
R1(conf-router_bgp)# neighbor 192.168.128.2 peer-group AAA
R1(conf-router_bgp)# neighbor 192.168.128.3 peer-group BBB
R1(conf-router_bgp)#
R1(conf-router_bgp)#show config
!
router bgp 99
network 192.168.128.0/24
neighbor AAA peer-group
neighbor AAA no shutdown
neighbor BBB peer-group
neighbor BBB no shutdown
neighbor 192.168.128.2 remote-as 99
neighbor 192.168.128.2 peer-group AAA
neighbor 192.168.128.2 update-source Loopback 0
neighbor 192.168.128.2 no shutdown
neighbor 192.168.128.3 remote-as 100
neighbor 192.168.128.3 peer-group BBB
neighbor 192.168.128.3 update-source Loopback 0
neighbor 192.168.128.3 no shutdown
R1#
R1#show ip bgp summary
BGP router identifier 192.168.128.1, local AS number 99
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 96 bytes of memory
2 BGP AS-PATH entrie(s) using 74 bytes of memory
2 neighbor(s) using 8672 bytes of memory
Neighbor        AS    MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.2   99    23       24          0       0     0  00:06:11 0
```

Example of Enabling Peer Groups (Router 2)

```
R2#conf
R2(conf)#router bgp 99
R2(conf-router_bgp)# neighbor CCC peer-group
R2(conf-router_bgp)# neighbor CC no shutdown
R2(conf-router_bgp)# neighbor BBB peer-group
R2(conf-router_bgp)# neighbor BBB no shutdown
R2(conf-router_bgp)# neighbor 192.168.128.1 peer AAA
R2(conf-router_bgp)# neighbor 192.168.128.1 no shut
R2(conf-router_bgp)# neighbor 192.168.128.3 peer BBB
R2(conf-router_bgp)# neighbor 192.168.128.3 no shut
R2(conf-router_bgp)#show conf
!
router bgp 99
network 192.168.128.0/24
neighbor AAA peer-group
neighbor AAA no shutdown
neighbor BBB peer-group
neighbor BBB no shutdown
neighbor 192.168.128.1 remote-as 99
neighbor 192.168.128.1 peer-group CCC
neighbor 192.168.128.1 update-source Loopback 0
neighbor 192.168.128.1 no shutdown
neighbor 192.168.128.3 remote-as 100
neighbor 192.168.128.3 peer-group BBB
neighbor 192.168.128.3 update-source Loopback 0
neighbor 192.168.128.3 no shutdown
R2(conf-router_bgp)#end
R2#
R2#show ip bgp summary
BGP router identifier 192.168.128.2, local AS number 99
BGP table version is 2, main routing table version 2
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory
Neighbor        AS    MsgRcvd  MsgSent  TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.1   99    140      136      2       0    (0)  00:11:24 1
192.168.128.3  100    138      140      2       0    (0)  00:18:31 1
```

Example of Enabling Peer Groups (Router 3)

```
R3#conf
R3(conf)#router bgp 100
R3(conf-router_bgp)# neighbor AAA peer-group
R3(conf-router_bgp)# neighbor AAA no shutdown
R3(conf-router_bgp)# neighbor CCC peer-group
R3(conf-router_bgp)# neighbor CCC no shutdown
R3(conf-router_bgp)# neighbor 192.168.128.2 peer-group BBB
R3(conf-router_bgp)# neighbor 192.168.128.2 no shutdown
R3(conf-router_bgp)# neighbor 192.168.128.1 peer-group BBB
R3(conf-router_bgp)# neighbor 192.168.128.1 no shutdown
R3(conf-router_bgp)#
R3(conf-router_bgp)#end
R3#show ip bgp summary
BGP router identifier 192.168.128.3, local AS number 100
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory
Neighbor        AS    MsgRcvd  MsgSent  TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.1   99    93       99       1       0    (0)  00:00:15 1
192.168.128.2   99    122      120      1       0    (0)  00:00:11 1
```

Advanced BGP configuration tasks

The following sections describe how to configure the advanced (optional) BGP configuration tasks.

Route-refresh and Soft-reconfiguration

BGP soft-reconfiguration allows for faster and easier route changing.

Changing routing policies typically requires a reset of BGP sessions (the TCP connection) for the policies to take effect. Such resets cause undue interruption to traffic due to hard reset of the BGP cache and the time it takes to re-establish the session. BGP soft-reconfiguration allows for policies to be applied to a session without clearing the BGP Session. Soft-reconfiguration can be done on a per-neighbor basis and can either be inbound or outbound.

BGP soft-reconfiguration clears the policies without resetting the TCP connection.

To reset a BGP connection using BGP soft-reconfiguration, use the `clear ip bgp` command in EXEC Privilege mode at the system prompt.

When you change the BGP inbound policy locally, you need to process the updates received from a peer. The route-refresh capability allows the local peer to reset inbound information dynamically by exchanging route-refresh requests to supporting peers. When an inbound policy configuration is changed, then the Dell EMC Networking OS sends a refresh request message asking the peer to re-advertise updates. Upon receiving the refresh request, the peer advertises the required information after applying the outbound filters. Route-refresh is negotiated during the BGP session establishment and will be used only if both the BGP peers support this capability.

To determine whether a BGP router supports this capability, use the `show ip bgp neighbors` command. If a router supports the route-refresh capability, the following message displays:

```
Received route refresh capability from peer.
```

If you specify a BGP peer group by using the *peer-group-name* argument, all members of the peer group inherit the characteristic configured with this command.

- Enable soft-reconfiguration for the BGP neighbor specified.

  CONFIG-ROUTER-BGP mode

  `neighbor {ip-address | ipv6-address | peer-group-name} soft-reconfiguration inbound`

  BGP stores all the updates received by the neighbor but does not reset the peer-session.

  Entering this command starts the storage of updates, which is required to do inbound soft-reconfiguration. Outbound BGP soft-reconfiguration does not require inbound soft-reconfiguration to be enabled.
- Reset BGP connection using soft-reconfiguration.

  EXEC Privilege mode

  `clear ip bgp [vrf vrf-name] [* | neighbor-address | as-number | peer-group-name] {soft {in | out}}`

  - `*`: Clears all peers.
  - *neighbor-address*: Clears the IPv4 or IPv6 neighbor of this IP address.
  - *as-number*: Clears the peers AS numbers.
  - *peer-group-name*: Clears all members of the specified peer group.

The example enables inbound soft-reconfiguration for the neighbor 10.108.1.1. All updates received from this neighbor are stored unmodified, regardless of the inbound policy. When inbound soft-reconfiguration is done later, the stored information is used to generate a new set of inbound updates

```
DellEMC(conf)# router bgp 100
DellEMC(conf-router_bgp)# neighbor 10.108.1.1 remote-as 20
DellEMC(conf-router_bgp)# neighbor 10.108.1.1 soft-reconfiguration inbound
DellEMC(conf-router_bgp)# exit
```

Route-refresh

This section explains how the soft-reconfiguration and route-refresh works. Soft-reconfiguration has to be configured explicitly for a neighbor unlike route refresh, which is automatically negotiated between BGP peers when establishing a peer session.

The route-refresh updates will be sent, only if the `neighbor soft-reconfiguration inbound` command is not configured in a BGP neighbor and when you do a soft reset using `clear ip bgp {`*`neighbor-address | peer-group-name`*`} soft in` command.

If the `neighbor soft-reconfiguration inbound` command is already configured in the BGP neighbor, the route refresh updates will not be sent. If you want to use route-refresh updates instead, remove the `neighbor soft-reconfiguration inbound` configuration and do a hard reset once, using `clear ip bgp {* | `*`as-number | ipv4| ipv6 | neighbor-address | peer-group-name`*`}` command. If you remove `neighbor soft-reconfiguration inbound` configuration for an individual neighbor (not part of peer-group), it is sufficient to do a hard reset only for the individual neighbor. If the neighbor is part of a peer-group and when `neighbor soft-reconfiguration inbound` is removed from the peer group, you need to do a hard reset for the peer-group.

If `neighbor soft-reconfiguration inbound` command is not configured ever in the router, then doing a soft reset is enough for the route-refresh updates to be sent.

Route-refresh updates for IPv4 and IPv6 prefixes

This section explains the route-refresh functionality in different combinations for IPv4 or IPv6 prefix configured with IPv4 or IPv6 neighbors.

By default, the IPv4 prefixes is sent for all the neighbors irrespective of IPv4 address family is enabled or disabled. Following are the different scenarios during which the route-refresh message is sent for the corresponding neighbors:

By default, the route-refresh message for IPv4 prefixes is sent for both the IPv4 and IPv6 neighbors. Following is an example configuration in which IPv4 prefixes are enabled by default and the corresponding route-refresh message:

```
DellEMC(conf-router_bgp)# show config
!
router bgp 100
 redistribute connected
 neighbor 20.1.1.2 remote-as 200
 neighbor 20.1.1.2 no shutdown
 neighbor 20::2 remote-as 200
 neighbor 20::2 no shutdown
 !
 address-family ipv6 unicast
  redistribute connected
  exit-address-family
 !
```

```
DellEMC(conf-router_bgp)#do clear ip bgp 20.1.1.2 soft in
May  8 15:28:11 : BGP: 20.1.1.2 sending ROUTE_REFRESH AFI/SAFI (1/1)
May  8 15:28:12 : BGP: 20.1.1.2 UPDATE rcvd packet len 56
May  8 15:28:12 : BGP: 20.1.1.2 rcvd UPDATE w/ attr: origin ?, path  200, nexthop
20.1.1.2, metric 0,
<output truncated>
DellEMC(conf-router_bgp)#do clear ip bgp 20::2 soft in
May  8 15:28:04 : BGP: 20::2 sending ROUTE_REFRESH AFI/SAFI (1/1)
May  8 15:28:04 : BGP: 20::2 UPDATE rcvd packet len 56
May  8 15:28:04 : BGP: 20::2 rcvd UPDATE w/ attr: origin ?, path  200, nexthop 20.1.1.2,
metric 0,
<output truncated>
```

In the below example, under the IPv6 address family configuration, only the IPv6 neighbor is enabled using `neighbor `*`ipv6-address`*` activate` command. If you execute, `clear ip bgp `*`neighbor-ipv4-address`*` soft in` command, only the IPv4 route-refresh update is sent. If you execute `clear ip bgp `*`neighbor-ipv6-address`*` soft in`command, both the IPv4 and IPv6 route-refresh updates are sent. Following is an example configuration in which IPv6 prefixes is enabled for a IPv6 neighbor and the corresponding route-refresh message:

```
DellEMC(conf-router_bgp)# show config
!
router bgp 100
```

```
 redistribute connected
 neighbor 20.1.1.2 remote-as 200
 neighbor 20.1.1.2 no shutdown
 neighbor 20::2 remote-as 200
 neighbor 20::2 no shutdown
 !
 address-family ipv6 unicast
  redistribute connected
  neighbor 20::2 activate
  exit-address-family
 !
```

```
DellEMC(conf-router_bgp)#do clear ip bgp 20.1.1.2 soft in
May  8 15:28:11 : BGP: 20.1.1.2 sending ROUTE_REFRESH AFI/SAFI (1/1)
May  8 15:28:12 : BGP: 20.1.1.2 UPDATE rcvd packet len 56
May  8 15:28:12 : BGP: 20.1.1.2 rcvd UPDATE w/ attr: origin ?, path  200, nexthop
20.1.1.2, metric 0,
<output truncated>
DellEMC(conf-router_bgp)#do clear ip bgp 20::2 soft in
May  8 15:40:08 : BGP: 20::2 sending ROUTE_REFRESH AFI/SAFI (1/1)
May  8 15:40:08 : BGP: 20::2 sending ROUTE_REFRESH AFI/SAFI (2/1)
May  8 15:40:08 : BGP: 20::2 UPDATE rcvd packet len 56
May  8 15:40:08 : BGP: 20::2 rcvd UPDATE w/ attr: origin ?, path  200, nexthop 20.1.1.2,
metric 0,
```

Controlling route-refresh updates

You can control route-refresh updates for IPv4 and IPv6 prefixes.

Following is the list of commands to control the route-refresh updates for IPv4 and IPv6 prefixes:

Table 10. Command for respective route-refresh updates

Command	Route-refresh Updates
`clear ip bgp` *`neighbor-ipv4-address`* `ipv4 unicast soft in`	Only IPv4 route-refresh updates using IPv4 neighbor
`clear ip bgp` *`neighbor-ipv6-address`* `ipv6 unicast soft in`	Only IPv6 route-refresh updates using IPv6 neighbor
`clear ip bgp` *`neighbor-ipv4-address`* `ipv6 unicast soft in`	Only IPv6 route-refresh updates using IPv4 neighbor
`clear ip bgp` *`neighbor-ipv6-address`* `ipv4 unicast soft in`	Only IPv4 route-refresh updates using IPv6 neighbor

Aggregating Routes

When BGP peers exchange information about the networks, a large BGP routing tables is created. Dell EMC Networking OS provides multiple ways to aggregate routes to reduced the size of the BGP routing table which helps to improve the network performance by reducing the convergence time. You can configure BGP to configure and advertise aggregated routes. At least one specific route of the aggregate must be in the routing table for the configured aggregate to become active.

To aggregate routes, use the following command.

AS_SET includes AS_PATH and community information from the routes included in the aggregated route.

- Assign the IP address and mask of the prefix to be aggregated.
  CONFIG-ROUTER-BGP mode

  `aggregate-address` *`ip-address`* `mask [advertise-map` *`map-name`*`] [as-set] [attribute-map` *`map-name`*`] [summary-only] [suppress-map` *`map-name`*`]`

  - `as-set`- Specify that the advertised path of this route is an AS_SET.
  - `summary-only`-Create aggregate route and suppress advertisements of specific routes to all neighbors.

- `suppress-map map-name`-Create aggregate route by suppressing the advertisements of specific routes.
- `advertise-map map-name`-Create aggregate route by advertising specific routes.

Configuring BGP aggregate routes

To create an aggregate route entry in the BGP routing table, use the following commands.

The aggregate route is advertised from the autonomous system.

- Enter the router configuration mode and the AS number for the specific BGP routing process.

  CONFIG mode

  `router bgp as-number`
- Create an aggregate entry in the BGP routing table.

  CONFIG-ROUTER-BGP mode

  `aggregate—address address-mask`

  Use the `aggregate-address` command without any keywords to create an aggregate entry if any specific BGP routes are available in the specified range.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 100
DellEMC(conf-router_bgp)# aggregate-address 10.1.1.0/24
DellEMC(conf-router_bgp)# exit
DellEMC(conf)#
```

Following is the sample output of `show ip bgp` command.

```
DellEMC# show ip bgp
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
BGP local router ID is 30.30.30.30
Status codes: s suppressed, S stale, d dampened, h history, * valid, > best
Path source: I - internal, a - aggregate, c - confed-external, r - redistributed
             n - network,  D - denied,  S - stale
Origin codes: i - IGP, e - EGP, ? - incomplete

    Network            Next Hop            Metric     LocPrf Weight Path
*>r 3.3.3.0/24         0.0.0.0                  0             32768 ?
s>r 10.1.1.0/30        0.0.0.0                  0             32768 ?
*>r 30.30.30.30/32     0.0.0.0                  0             32768 ?
*>a 10.1.1.0/24        0.0.0.0                                32768 ?
```

Suppressing BGP aggregate routes

The routes that are suppressed are not advertised to any of the BGP neighbors. You can suppress the aggregate routes using the `suppress-map` or `summary-only` options in the `aggregate-address` configuration. To suppress the advertisement of the aggregate routes using BGP, use the following commands.

- Create an aggregate entry and suppress the advertisement of specific routes to all neighbors.

  CONFIG-ROUTER-BGP mode

  `aggregate—address address-mask summary-only`
- Create an aggregate entry and suppress the advertisement of specific routes.

  CONFIG-ROUTER-BGP mode

  `aggregate—address address-mask suppress-map map-name`

  The `suppress-map` keyword creates the aggregate route but suppress the advertisement of specified routes. The routes that are suppressed are not advertised to the neighbors. You can use `match` clause of route maps to selectively suppress the specific route from the aggregate routes.

Following is the sample configuration to suppress the advertisement of specific aggregate routes.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 100
```

```
DellEMC(conf-router_bgp)# aggregate-address 10.1.1.0/24 suppress-map map1
DellEMC(conf-router_bgp)# exit
DellEMC(conf)#
```

The route-map named `map1` can have any action such as `permit` and sequence number configured, so that the advertisement of aggregate routes can be suppressed based on the set action in the route-map.

Following is the sample configuration to suppress the advertisement of specific aggregate routes to all neighbors.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 100
DellEMC(conf-router_bgp)# aggregate-address 10.1.1.0 255.255.255.0 summary-only
DellEMC(conf-router_bgp)# exit
DellEMC(conf)#
```

Filtering BGP

The following section describes the methods used to filter the updates received from BGP neighbors.

Following are the filtering methods of BGP updates:

- Filtering using IP prefix lists
- Filtering using route maps
- Filtering using AS-PATH information
- Filtering using community lists

Regular Expressions as Filters

Regular expressions are used to filter AS paths or community lists. A regular expression is a special character used to define a pattern that is then compared with an input string.

For an AS-path access list, as shown in the previous commands, if the AS path matches the regular expression in the access list, the route matches the access list.

The following lists the regular expressions accepted in Dell EMC Networking OS.

Regular Expression	Definition
^ (caret)	Matches the beginning of the input string. Alternatively, when used as the first character within brackets [^], this matches any number except the ones specified within the brackets.
$ (dollar)	Matches the end of the input string.
. (period)	Matches any single character, including white space.
* **(asterisk)**	Matches 0 or more sequences of the immediately previous character or pattern.
+ (plus)	Matches 1 or more sequences of the immediately previous character or pattern.
? (question)	Matches 0 or 1 sequence of the immediately previous character or pattern.
( ) (parenthesis)	Specifies patterns for multiple use when one of the multiplier metacharacters follows: asterisk *, plus sign +, or question mark ?
[ ] (brackets)	Matches any enclosed character and specifies a range of single characters.
- (hyphen)	Used within brackets to specify a range of AS or community numbers.
_ (underscore)	Matches a ^, a $, a comma, a space, or a {, or a }. Placed on either side of a string to specify a literal and disallow substring matching. You can precede or follow numerals enclosed by underscores by any of the characters listed.
| (pipe)	Matches characters on either side of the metacharacter; logical OR.

As seen in the following example, the expressions are displayed when using the `show` commands. To view the AS-PATH ACL configuration, use the `show config` command in CONFIGURATION AS-PATH ACL mode and the `show ip as-path-access-list` command in EXEC Privilege mode.

For more information about this command and route filtering, refer to Filtering BGP Routes.

The following example applies access list Eagle to routes inbound from BGP peer 10.5.5.2. Access list Eagle uses a regular expression to deny routes originating in AS 32. The first lines shown in bold create the access list and filter. The second lines shown in bold are the regular expression shown as part of the access list filter.

Example of Using Regular Expression to Filter AS Paths

```
DellEMC(config)#router bgp 99
DellEMC(conf-router_bgp)#neigh AAA peer-group
DellEMC(conf-router_bgp)#neigh AAA no shut
DellEMC(conf-router_bgp)#show conf
!
router bgp 99
  neighbor AAA peer-group
  neighbor AAA no shutdown
  neighbor 10.155.15.2 remote-as 32
  neighbor 10.155.15.2 shutdown
DellEMC(conf-router_bgp)#neigh 10.155.15.2 filter-list 1 in
DellEMC(conf-router_bgp)#ex

DellEMC(conf)#ip as-path access-list Eagle
DellEMC(config-as-path)#deny 32$
DellEMC(config-as-path)#ex
DellEMC(conf)#router bgp 99
DellEMC(conf-router_bgp)#neighbor AAA filter-list Eagle in
DellEMC(conf-router_bgp)#show conf
!
router bgp 99
  neighbor AAA peer-group
  neighbor AAA filter-list Eaglein
  neighbor AAA no shutdown
  neighbor 10.155.15.2 remote-as 32
  neighbor 10.155.15.2 filter-list 1 in
  neighbor 10.155.15.2 shutdown
DellEMC(conf-router_bgp)#ex
DellEMC(conf)#ex

DellEMC#show ip as-path-access-lists
ip as-path access-list Eagle
 deny 32$
DellEMC#
```

Filtering BGP Routes

Filtering routes allows you to implement BGP policies.

You can use either IP prefix lists, route maps, AS-PATH ACLs or IP community lists (using a route map) to control which routes the BGP neighbor or peer group accepts and advertises. Prefix lists filter routes based on route and prefix length, while AS-Path ACLs filter routes based on the ASN. Route maps can filter and set conditions, change attributes, and assign update policies.

NOTE: Dell EMC Networking OS supports up to 255 characters in a set community statement inside a route map.

NOTE: You can create inbound and outbound policies. Each of the commands used for filtering has `in` and `out` parameters that you must apply. In Dell EMC Networking OS, the order of preference varies depending on whether the attributes are applied for inbound updates or outbound updates.

For inbound and outbound updates the order of preference is:

- prefix lists (using the `neighbor distribute-list` command)
- AS-PATH ACLs (using the `neighbor filter-list` command)
- route maps (using the `neighbor route-map` command)

Prior to filtering BGP routes, create the prefix list, AS-PATH ACL, or route map.

For configuration information about prefix lists, AS-PATH ACLs, and route maps, refer to Access Control Lists (ACLs).

NOTE: When you configure a new set of BGP policies, to ensure the changes are made, always reset the neighbor or peer group by using the `clear ip bgp` command in EXEC Privilege mode.

Filtering BGP using IP prefix lists

An IP prefix lists contains a list of networks. When applying an IP prefix list to a BGP neighbor, you can able to send or receive the routes whose destination is in the IP prefix list.

Filtering BGP routes based on IP prefix lists involves the following steps:

- Create a prefix list. The steps 1 to 3 in the following configuration commands shows how to create a prefix list.
- Applying the created prefix list to a BGP neighbor. The steps 4 and 5 show how the prefix list is applied to a neighbor.

To configure an IP prefix list, use the following commands.

1. Create a prefix list and assign it a name.
   CONFIGURATION mode
   `ip prefix-list prefix-name`
2. Create multiple prefix list filters with a deny or permit action.
   CONFIG-PREFIX LIST mode
   `seq sequence-number {deny | permit} {any | ip-prefix [ge | le] }`
   - `ge`: minimum prefix length to be matched.
   - `le`: maximum prefix length to me matched.

   For information about configuring prefix lists, refer to Access Control Lists (ACLs).
3. Return to CONFIGURATION mode.
   CONFIG-PREFIX LIST mode
   `exit`
4. Enter ROUTER BGP mode.
   CONFIGURATION mode
   `router bgp as-number`
5. Filter routes based on the criteria in the configured prefix list.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} distribute-list prefix-list-name {in | out}`
   Configure the following parameters:
   - *ip-address* or *ipv6-address* or *peer-group-name*: enter the neighbor's IPv4 or IPv6 address or the peer group's name.
   - *prefix-list-name*: enter the name of a configured prefix list.
   - `in`: apply the prefix list to inbound routes.
   - `out`: apply the prefix list to outbound routes.

As a reminder, the following are rules concerning prefix lists:

- If the prefix list contains no filters, all routes are permitted.
- If none of the routes match any of the filters in the prefix list, the route is denied. This action is called an implicit deny. (If you want to forward all routes that do not match the prefix list criteria, you must configure a prefix list filter to permit all routes. For example, you could have the following filter as the last filter in your prefix list permit 0.0.0.0/0 le 32).
- After a route matches a filter, the filter's action is applied. No additional filters are applied to the route.

```
DellEMC# configure terminal
DellEMC(conf)# ip prefix-list route10
DellEMC(conf)# seq 10 permit 10.10.10.0/24 ge 32
DellEMC(conf)#exit
DellEMC(conf)# router bgp
DellEMC(conf-router_bgp)#neighbor 10.10.10.2 distribute-list route10 out
```

The commands show the configuration of an IP prefix list named `route10`, which permits routes to network 10.10.10.0/24. The `neighbor` command configures to use IP prefix list `route10` to determine, which routes to be distributed to the neighbor

10.10.10.2. So the routes from 10.10.10.1/24 network is distributed to neighbor 10.10.10.2 since the IP prefix list `route10` explicitly permits the routes to be distributed to the neighbor.

To view the BGP configuration, use the `show config` command in ROUTER BGP mode. To view a prefix list configuration, use the `show ip prefix-list detail` or `show ip prefix-list summary` commands in EXEC Privilege mode.

Filtering BGP Routes Using Route Maps

To filter routes using a route map, use these commands.

1. Create a route map and assign it a name.
   CONFIGURATION mode
   `route-map map-name [permit | deny] [sequence-number]`
2. Create multiple route map filters with a match or set action.
   CONFIG-ROUTE-MAP mode
   `{match | set}`
   For information about configuring route maps, see Access Control Lists (ACLs).
3. Return to CONFIGURATION mode.
   CONFIG-ROUTE-MAP mode
   `exit`
4. Enter ROUTER BGP mode.
   CONFIGURATION mode
   `router bgp as-number`
5. Filter routes based on the criteria in the configured route map.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} route-map map-name {in | out}`
   Configure the following parameters:
   - *ip-address* or *ipv6-address* or *peer-group-name*: enter the neighbor's IPv4 or IPv6 address or the peer group's name.
   - *map-name*: enter the name of a configured route map.
   - `in`: apply the route map to inbound routes.
   - `out`: apply the route map to outbound routes.

```
DellEMC# configure terminal
DellEMC(conf)# route-map route1 permit
DellEMC(conf-route-map)# set as-path prepend 400
DellEMC(conf-route-map)#exit
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 route-map route1 out
DellEMC(conf-router_bgp)#
```

In the above example, a route-map named route1 is created with a permit clause. Within the route-map, a set clause is configured to prepend an AS is configured. This configuration adds the configured AS number to the AS-Path for each route. The configured route-map is applied to the outbound routes of the neighbor 10.10.10.1.

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the `show route-map` command in EXEC Privilege mode.

Filtering on an AS-Path Attribute

You can use the BGP attribute, AS_PATH, to manipulate routing policies.

The AS_PATH attribute contains a sequence of AS numbers representing the route's path. As the route traverses an AS, the ASN is prepended to the route. You can manipulate routes based on their AS_PATH to affect interdomain routing. By identifying certain ASN in the AS_PATH, you can permit or deny routes based on the number in its AS_PATH.

AS-PATH ACLs use regular expressions to search AS_PATH values. AS-PATH ACLs have an "implicit deny." This means that routes that do not meet a deny or match filter are dropped.

To configure an AS-PATH ACL to filter a specific AS_PATH value, use these commands in the following sequence.

1. Assign a name to a AS-PATH ACL and enter AS-PATH ACL mode.
   CONFIGURATION mode
   `ip as-path access-list as-path-name`
2. Enter the parameter to match BGP AS-PATH for filtering.
   CONFIG-AS-PATH mode
   `{deny | permit} filter parameter`
   This is the filter that is used to match the AS-path. The entries can be any format, letters, numbers, or regular expressions.
   You can enter this command multiple times if multiple filters are desired.
   For accepted expressions, refer to Regular Expressions as Filters.
3. Return to CONFIGURATION mode.
   AS-PATH ACL mode
   `exit`
4. Enter ROUTER BGP mode.
   CONFIGURATION mode
   `router bgp as-number`
5. Use a configured AS-PATH ACL for route filtering and manipulation.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} filter-list as-path-name {in | out}`
   If you assign an non-existent or empty AS-PATH ACL, the software allows all routes.

To view all BGP path attributes in the BGP database, use the `show ip bgp paths` command in EXEC Privilege mode.

```
DellEMC#show ip bgp paths
Total 386 Paths
Refcount Metric Path
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
20       0            64801 i
--More--
```

Filtering Routes with Community Lists

To use an IP community list or IP extended community list to filter routes, you must apply a match community filter to a route map and then apply that route map to a BGP neighbor or peer group.

1. Enter the ROUTE-MAP mode and assign a name to a route map.
   CONFIGURATION mode
   `route-map map-name [permit | deny] [sequence-number]`
2. Configure a match filter for all routes meeting the criteria in the IP community or IP extended community list.

CONFIG-ROUTE-MAP mode

`match {community community-list-name [exact] | extcommunity extcommunity-list-name [exact]}`

3. Return to CONFIGURATION mode.

   CONFIG-ROUTE-MAP mode

   `exit`

4. Enter ROUTER BGP mode.

   CONFIGURATION mode

   `router bgp as-number`

   AS-number: 0 to 65535 (2-Byte) or 1 to 4294967295 (4-Byte) or 0.1 to 65535.65535 (Dotted format)

5. Apply the route map to the neighbor or peer group's incoming or outgoing routes.

   CONFIG-ROUTER-BGP mode

   `neighbor {ip-address | ipv6-address | peer-group-name} route-map map-name {in | out}`

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 remote-as 500
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 route-map route2 in
DellEMC(conf-router_bgp)# exit
DellEMC(conf)# route-map route2 permit 10
DellEMC(conf-route-map)# match community 1
DellEMC(conf-route-map)# set community 99
DellEMC(conf-route-map)# exit
DellEMC(conf)# ip community-list 1 permit 100
DellEMC(conf)# exit
DellEMC#
```

In the above example, add a BGP neighbor to the AS 400 and the route-map called `route2` applied to inbound routes from the BGP neighbor at 10.10.10.1. A route map `route2` is created with a permit clause and the route's community attribute is matched to communities in community list 1. A community list 1 that permits routes with a communities attribute of 100.

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the `show route-map` command in EXEC Privilege mode.

To view which BGP routes meet an IP community or IP extended community list's criteria, use the `show ip bgp {community-list | extcommunity-list}` command in EXEC Privilege mode.

Configuring BGP Fast Fall-Over

By default, a BGP session is governed by the hold time.

BGP routers typically carry large routing tables, so frequent session resets are not desirable. The BGP fast fall-over feature reduces the convergence time while maintaining stability. The connection to a BGP peer is immediately reset if a link to a directly connected external peer fails.

When you enable fall-over, BGP tracks IP reachability to the peer remote address and the peer local address. Whenever either address becomes unreachable (for example, no active route exists in the routing table for peer IPv6 destinations/local address), BGP brings down the session with the peer.

The BGP fast fall-over feature is configured on a per-neighbor or peer-group basis and is disabled by default.

To enable the BGP fast fall-over feature, use the following command.

- Enable BGP fast fall-over.

  CONFIG-ROUTER-BGP mode

  `neighbor {ip-address | ipv6-address | peer-group-name} fall-over`

  To disable fast fall-over, use the `[no] neighbor [ip-address | ipv6-address | peer-group] fall-over` command in CONFIGURATION ROUTER BGP mode.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 remote-as 500
```

```
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 fall-over
DellEMC(conf-router_bgp)# exit
DellEMC(conf-router_bgp)#
```

To verify that you enabled fast fall-over on a particular BGP neighbor, use the `show ip bgp neighbors` command. Because fast fall-over is disabled by default, it appears only if it has been enabled (shown in bold).

```
DellEMC#show ip bgp neighbors

BGP neighbor is 10.10.10.1, remote AS 500, internal link
  Member of peer-group test for session parameters
  BGP version 4, remote router ID 30.30.30.5
  BGP state ESTABLISHED, in this state for 00:19:15
  Last read 00:00:15, last write 00:00:06
  Hold time is 180, keepalive interval is 60 seconds
  Received 52 messages, 0 notifications, 0 in queue
  Sent 45 messages, 5 notifications, 0 in queue
  Received 6 updates, Sent 0 updates
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 5 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)
```

fall-over enabled

```

Update source set to Loopback 0

Peer active in peer-group outbound optimization

For address family: IPv4 Unicast
BGP table version 52, neighbor version 52
4 accepted prefixes consume 16 bytes
Prefix advertised 0, denied 0, withdrawn 0

Connections established 6; dropped 5
Last reset 00:19:37, due to Reset by peer

Notification History
  'Connection Reset' Sent : 5 Recv: 0

Local host: 20.20.20.2, Local port: 65519
Foreign host: 10.10.10.1, Foreign port: 179

DellEMC#
```

To verify that fast fall-over is enabled on a peer-group, use the `show ip bgp peer-group` command (shown in bold).

```
DellEMC#show ip bgp peer-group

Peer-group test
  fall-over enabled
  BGP version 4
  Minimum time between advertisement runs is 5 seconds

  For address family: IPv4 Unicast
  BGP neighbor is test
  Number of peers in this group 1
  Peer-group members (* - outbound optimized):
    10.10.10.1*

DellEMC#

router bgp 400
```

```
   neighbor test peer-group
   neighbor test fall-over
   neighbor test no shutdown
   neighbor 10.10.10.1 remote-as 500
   neighbor 10.10.10.1 fall-over
   neighbor 10.10.10.1 update-source Loopback 0
   neighbor 10.10.10.1 no shutdown
DellEMC#
```

Configuring Passive Peering

When you enable a peer-group, the software sends an OPEN message to initiate a TCP connection.

If you enable passive peering for the peer group, the software does not send an OPEN message, but it responds to an OPEN message.

When a BGP neighbor connection with authentication configured is rejected by a passive peer-group, Dell EMC Networking OS does not allow another passive peer-group on the same subnet to connect with the BGP neighbor. To work around this, change the BGP configuration or change the order of the peer group configuration.

You can constrain the number of passive sessions accepted by the neighbor. The `limit` keyword allows you to set the total number of sessions the neighbor will accept, between 2 and 265. The default is **256** sessions.

1. Configure a peer group that does not initiate TCP connections with other peers.
   CONFIG-ROUTER-BGP mode
   `neighbor peer-group-name peer-group passive limit`
   Enter the `limit` keyword to restrict the number of sessions accepted.
2. Assign a subnet to the peer group.
   CONFIG-ROUTER-BGP mode
   `neighbor peer-group-name subnet subnet-number mask`
   The peer group responds to OPEN messages sent on this subnet.
3. Enable the peer group.
   CONFIG-ROUTER-BGP mode
   `neighbor peer-group-name no shutdown`
4. Create and specify a remote peer for BGP neighbor.
   CONFIG-ROUTER-BGP mode
   `neighbor peer-group-name remote-as as-number`

Only after the peer group responds to an OPEN message sent on the subnet does its BGP state change to ESTABLISHED. After the peer group is ESTABLISHED, the peer group is the same as any other peer group.

For more information about peer groups, refer to Configure Peer Groups.

Enabling Graceful Restart

Use this feature to lessen the negative effects of a BGP restart.

Dell EMC Networking OS advertises support for this feature to BGP neighbors through a capability advertisement. You can enable graceful restart by router and/or by peer or peer group.

NOTE: By default, BGP graceful restart is disabled.

The default role for BGP is as a receiving or restarting peer. If you enable BGP, when a peer that supports graceful restart resumes operating, Dell EMC Networking OS performs the following tasks:

- Continues saving routes received from the peer if the peer advertised it had graceful restart capability. Continues forwarding traffic to the peer.
- Flags routes from the peer as Stale and sets a timer to delete them if the peer does not perform a graceful restart.
- Deletes all routes from the peer if forwarding state information is not saved.
- Speeds convergence by advertising a special update packet known as an end-of-RIB marker. This marker indicates the peer has been updated with all routes in the local RIB.

If you configure your system to do so, Dell EMC Networking OS can perform the following actions during a hot failover:

- Save all forwarding information base (FIB) and content addressable memory (CAM) entries on the line card and continue forwarding traffic while the secondary route processor module (RPM) is coming online.
- Advertise to all BGP neighbors and peer-groups that the forwarding state of all routes has been saved. This prompts all peers to continue saving the routes they receive and to continue forwarding traffic.
- Bring the secondary RPM online as the primary and re-open sessions with all peers operating in No Shutdown mode.
- Defer best path selection for a certain amount of time. This helps optimize path selection and results in fewer updates being sent out.

To enable graceful restart, use the following command.

`bgp graceful-restart [restart-time seconds] [stale-path-time seconds] [role receiver-only]`

To return to the default, use the `no bgp graceful-restart` command.

- Enable graceful restart for the BGP node.

  CONFIG-ROUTER-BGP mode

  `bgp graceful-restart`
- Set maximum restart time, in seconds, to restart and bring-up all the peers.

  CONFIG-ROUTER-BGP mode

  `bgp graceful-restart [restart-time time-in-seconds]`

  The default is **120 seconds**.
- Set maximum time, in seconds, to retain the restarting peer's stale paths.

  CONFIG-ROUTER-BGP mode

  `bgp graceful-restart [stale-path-time time-in-seconds]`

  The default is **360 seconds**.
- Enable the local router to support graceful restart as a receiver only.

  CONFIG-ROUTER-BGP mode

  `bgp graceful-restart [role receiver-only]`

The below example configuration shows how to enable the BGP graceful restart.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# bgp graceful-restart
DellEMC(conf-router_bgp)# exit
```

Redistributing Routes

In addition to filtering routes, you can add routes from other routing instances or protocols to the BGP process. You can configure the device to redistribute ISIS, OSPF, static, or directly connected routes into BGP process using the `redistribute` command.

To add routes from other routing instances or protocols, use any of the following commands in ROUTER BGP mode.

- Include directly connected or user-configured (static) routes into BGP.

  ROUTER BGP or CONF-ROUTER_BGPv6_ AF mode

  `redistribute {connected | static} [route-map map-name]`

  Configure the following parameters:
  - `connected`: Indicate that you are redistributing routes to directly connected routes into BGP.
  - `static`: Indicate that you are redistributing static routes into BGP.
  - `route-map map-name`: Specify the name of a configured route map to be consulted before adding the connected or static route.
- Include specific ISIS routes into BGP.

  ROUTER BGP or CONF-ROUTER_BGPv6_ AF mode

  `redistribute isis [level-1 | level-1-2 | level-2] [metric value] [route-map map-name]`

  Configure the following parameters:
  - `isis`: Indicate that you are redistributing ISIS routes into BGP.
  - `level-1`, `level-1-2`, or `level-2`: Assign all redistributed routes to a level. The default is **level-2**.

- `metric value`: The value is from 0 to 16777215. The default is **0**.
- `route-map map-name`: Specify the name of a configured route map to be consulted before adding the ISIS route.

- Include specific OSPF routes into BGP.

  ROUTER BGP or CONF-ROUTER_BGPv6_ AF mode

  `redistribute ospf process-id [match external {1 | 2} | match internal] [metric-type {external | internal}] [route-map map-name]`

  Configure the following parameters:
  - `ospf`: Indicates that you are redistributing OSPF routes in BGP.
  - `process-id`: the range is from 1 to 65535.
  - `match internal, external1, and external2`: Specify the type of OSPF routes to be redistributed into BGP.
  - `metric-type`: external or internal.
  - `route-map map-name`: Specify the name of a configured route map to be consulted before adding the OSPF route.

```
DellEMC# configure terminal
DellEMC(conf)# ip route 10.0.0.0 255.0.0.0 null 0
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp_af)# redistribute static
DellEMC(conf-router_bgp_af)# exit
```

The above configuration example show how to create a static route and redistribute the static routes into the BGP routing table.

Enabling Additional Paths

The additional path allows the advertisement of more paths in addition to the best path. Enabling additional path allows the advertisement of multiple paths for the same address prefix without the new paths replacing any previous paths. The additional path feature is disabled by default.

NOTE: Dell EMC Networking OS recommends not to use multipath and add path simultaneously in a route reflector.

To allow multiple paths sent to peers, use the following commands.

1. Allow the advertisement of multiple paths (send, receive or both).

   CONFIG-ROUTER-BGP or CONFIG-ROUTER-BGP-AF mode

   `bgp add-path [both | enable | receive | send] path-count`

   Configure the following parameters:
   - `both`: Indicate that the system sends and accepts multiple paths from peers.
   - `enable`: Indicate that the system enables add-path support for the node.
   - `send`: Indicate that the system sends multiple paths to peers.
   - `receive`: Indicate that the system accepts multiple paths from peers.
   - `path-count`: Indicate that the system sends multiple paths to peers. The range is from 2 to 64.

2. Allow the specified neighbor or peer group to allow multiple path advertisements.

   CONFIG-ROUTER-BGP or CONFIG-ROUTER-BGP-AF mode

   `neighbor [ip-address | ipv6-address | peer-group-name] add-path [send | receive | both] path-count`

   NOTE: The `path-count` parameter controls the number of paths that are advertised, not the number of paths that are received.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 remote-as 500
DellEMC(conf-router_bgp)# bgp add-path both 2
DellEMC(conf-router_bgp)# address-family ipv4 multicast
DellEMC(conf-router_bgp_af)# neighbor 10.10.10.1 activate
```

```
DellEMC(conf-router_bgp_af)# neighbor 10.10.10.1 add-path both 3
DellEMC(conf-router_bgp_af)# exit
```

The above configuration example shows how to enable BGP additional paths to be sent and received with a maximum of two additional paths to the peers. You can configure the neighbor to send and receive additional paths using the `neighbor add-path`command at the address family configuration level.

Configuring IP Community Lists

Within Dell EMC Networking OS, you have multiple methods of manipulating routing attributes.

One attribute you can manipulate is the COMMUNITY attribute. This attribute is an optional attribute that is defined for a group of destinations. In Dell EMC Networking OS, you can assign a COMMUNITY attribute to BGP routers by using an IP community list. After you create an IP community list, you can apply routing decisions to all routers meeting the criteria in the IP community list.

IETF RFC 1997 defines the COMMUNITY attribute and the predefined communities of INTERNET, NO_EXPORT_SUBCONFED, NO_ADVERTISE, and NO_EXPORT. All BGP routes belong to the INTERNET community. In the RFC, the other communities are defined as follows:

- All routes with the NO_EXPORT_SUBCONFED (0xFFFFFF03) community attribute are not sent to CONFED-EBGP or EBGP peers, but are sent to IBGP peers within CONFED-SUB-AS.
- All routes with the NO_ADVERTISE (0xFFFFFF02) community attribute must not be advertised.
- All routes with the NO_EXPORT (0xFFFFFF01) community attribute must not be advertised outside a BGP confederation boundary, but are sent to CONFED-EBGP and IBGP peers.

Dell EMC Networking OS also supports BGP Extended Communities as described in RFC 4360 — BGP Extended Communities Attribute.

To configure an IP community list, use these commands.

1. Create a community list and enter COMMUNITY-LIST mode.

   CONFIGURATION mode

   `ip community-list` *`community-list-name`*

2. Configure a community list by denying or permitting specific community numbers or types of community.

   CONFIG-COMMUNITYLIST mode

   `{deny | permit}` *`{community-number`* `| local-AS | no-advertise | no-export | quote-regexp` *`regular-expression-list`* `| regexp` *`regular-expression}`*

   - *`community-number`*: use AA:NN format where AA is the AS number (2 Bytes or 4 Bytes) and NN is a value specific to that autonomous system.
   - `local-AS`: routes with the COMMUNITY attribute of NO_EXPORT_SUBCONFED.
   - `no-advertise`: routes with the COMMUNITY attribute of NO_ADVERTISE.
   - `no-export`: routes with the COMMUNITY attribute of NO_EXPORT.
   - `quote-regexp`: then any number of regular expressions. The software applies all regular expressions in the list.
   - `regexp`: then a regular expression.

To view the configuration, use the `show config` command in CONFIGURATION COMMUNITY-LIST or CONFIGURATION EXTCOMMUNITY LIST mode or the `show ip {`*`community-lists | extcommunity-list`*`}` command in EXEC Privilege mode.

```
DellEMC#show ip community-lists
ip community-list standard 1
  deny 701:20
  deny 702:20
  deny 703:20
  deny 704:20
  deny 705:20
  deny 14551:20
  deny 701:112
  deny 702:112
  deny 703:112
  deny 704:112
  deny 705:112
  deny 14551:112
  deny 701:667
```

```
   deny 702:667
   deny 703:667
   deny 704:666
   deny 705:666
   deny 14551:666
DellEMC#
```

Configuring an IP Extended Community List

To configure an IP extended community list, use these commands.

1. Create a extended community list and enter the EXTCOMMUNITY-LIST mode.

   CONFIGURATION mode

   `ip extcommunity-list` *`extcommunity-list-name`*

2. Two types of extended communities are supported.

   CONFIG-COMMUNITY-LIST mode

   `{permit | deny} {{rt | soo}` *`{ASN:NN | IPADDR:N}`* `| regex` *`REGEX-LINE`*`}`

   Filter routes based on the type of extended communities they carry using one of the following keywords:

   - `rt`: route target.
   - `soo`: route origin or site-of-origin. Support for matching extended communities against regular expression is also supported. Match against a regular expression using the following keyword.
   - `regexp`: regular expression.

To set or modify an extended community attribute, use the `set extcommunity {rt | soo} {ASN:NN | IPADDR:NN}` command.

To view the configuration, use the `show config` command in CONFIGURATION COMMUNITY-LIST or CONFIGURATION EXTCOMMUNITY LIST mode or the `show ip {`*`community-lists | extcommunity-list`*`}` command in EXEC Privilege mode.

```
DellEMC#show ip community-lists
ip community-list standard 1
   deny 701:20
   deny 702:20
   deny 703:20
   deny 704:20
   deny 705:20
   deny 14551:20
   deny 701:112
   deny 702:112
   deny 703:112
   deny 704:112
   deny 705:112
   deny 14551:112
   deny 701:667
   deny 702:667
   deny 703:667
   deny 704:666
   deny 705:666
   deny 14551:666
DellEMC#
```

Configure BGP attributes

Following sections explain how to configure the BGP attributes such as MED, COMMUNITY, WEIGHT, and LOCAL_PREFERENCE.

Changing MED Attributes

By default, Dell EMC Networking OS uses the MULTI_EXIT_DISC or MED attribute when comparing EBGP paths received from different BGP neighbors or peers from the same AS for the same route. You can configure the device to compare the MED attributes from neighbors or peers in different AS using the `bgp always-compare-med` command.

To change the usage of MED attribute to choose the best path among the advertised, enter any or all of the following commands.

- Enable MED comparison in the paths from neighbors with different ASs.

  CONFIG-ROUTER-BGP mode

  ```
  bgp always-compare-med
  ```

  By default, this comparison is not performed.
- Change the bestpath MED selection.

  CONFIG-ROUTER-BGP mode

  ```
  bgp bestpath med {confed | missing-as-best}
  ```

  - `confed`: Chooses the bestpath MED comparison of paths learned from BGP confederations.
  - `missing-as-best`: Treats a path without a MED value as the most preferred one.

To view the non-default values, use the `show config` command in CONFIGURATION ROUTER BGP mode.

Manipulating the COMMUNITY Attribute

A COMMUNITY attribute indicates that all the routes with that attribute belong to the same community grouping. In addition to permitting or denying routes based on the values of the COMMUNITY attributes, you can manipulate the COMMUNITY attribute value and send the COMMUNITY attribute with the route information.

By default, Dell EMC Networking OS does not send the COMMUNITY attribute to a BGP neighbor or a peer group.

To send the COMMUNITY attribute to BGP neighbors, use the following command.

- Enable the software to send the router's COMMUNITY attribute to the BGP neighbor or peer group specified.

  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | ipv6-address | peer-group-name} send-community [extended | standard]
  ```

`extended`- Allows to send the extended community attribute to a BGP neighbor or peer group.

`standard`- Allows to send the standard community attribute to a BGP neighbor or peer group.

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode.

If you want to remove or add a specific COMMUNITY number from a BGP path, you must create a route map with one or both of the following statements in the route map. Then apply that route map to a BGP neighbor or peer group.

1. Enter ROUTE-MAP mode and assign a name to a route map.

   CONFIGURATION mode

   ```
   route-map map-name [permit | deny] [sequence-number]
   ```
2. Configure a set filter to delete all COMMUNITY numbers in the IP community list.

   CONFIG-ROUTE-MAP mode

   ```
   set comm-list community-list-name delete
   ```

   OR

   ```
   set community {community-number | local-as | no-advertise | no-export | none}
   ```

   Configure a community list by denying or permitting specific community numbers or types of community.

- `community-number`: use AA:NN format where AA is the AS number (2 or 4 Bytes) and NN is a value specific to that autonomous system.
- `local-AS`: routes with the COMMUNITY attribute of NO_EXPORT_SUBCONFED and are not sent to EBGP peers.
- `no-advertise`: routes with the COMMUNITY attribute of NO_ADVERTISE and are not advertised.
- `no-export`: routes with the COMMUNITY attribute of NO_EXPORT.
- `none`: remove the COMMUNITY attribute.
- `additive`: add the communities to already existing communities.

3. Return to CONFIGURATION mode.

   CONFIG-ROUTE-MAP mode

   ```
   exit
   ```

4. Enter the ROUTER BGP mode.

   CONFIGURATION mode

   ```
   router bgp as-number
   ```

5. Apply the route map to the neighbor or peer group's incoming or outgoing routes.

   CONFIG-ROUTER-BGP mode

   ```
   neighbor {ip-address | ipv6-address | peer-group-name} route-map map-name {in | out}
   ```

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the `show route-map` command in EXEC Privilege mode.

To view BGP routes matching a certain community number or a pre-defined BGP community, use the `show ip bgp community` command in EXEC Privilege mode.

```
DellEMC>show ip bgp community
BGP table version is 3762622, local router ID is 10.114.8.48
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

Network             Next Hop  Metric  LocPrf Weight Path
* i 3.0.0.0/8       195.171.0.16       100    0      209 701 80 i
*>i 4.2.49.12/30    195.171.0.16       100    0      209 i
* i 4.21.132.0/23   195.171.0.16       100    0      209 6461 16422 i
*>i 4.24.118.16/30  195.171.0.16       100    0      209 i
*>i 4.24.145.0/30   195.171.0.16       100    0      209 i
*>i 4.24.187.12/30  195.171.0.16       100    0      209 i
*>i 4.24.202.0/30   195.171.0.16       100    0      209 i
*>i 4.25.88.0/30    195.171.0.16       100    0      209 3561 3908 i
*>i 6.1.0.0/16      195.171.0.16       100    0      209 7170 1455 i
*>i 6.2.0.0/22      195.171.0.16       100    0      209 7170 1455 i
*>i 6.3.0.0/18      195.171.0.16       100    0      209 7170 1455 i
*>i 6.4.0.0/16      195.171.0.16       100    0      209 7170 1455 i
*>i 6.5.0.0/19      195.171.0.16       100    0      209 7170 1455 i
*>i 6.8.0.0/20      195.171.0.16       100    0      209 7170 1455 i
*>i 6.9.0.0/20      195.171.0.16       100    0      209 7170 1455 i
*>i 6.10.0.0/15     195.171.0.16       100    0      209 7170 1455 i
*>i 6.14.0.0/15     205.171.0.16       100    0      209 7170 1455 i
*>i 6.133.0.0/21    205.171.0.16       100    0      209 7170 1455 i
*>i 6.151.0.0/16    205.171.0.16       100    0      209 7170 1455 i
--More--
```

Changing the LOCAL_PREFERENCE Attribute

In Dell EMC Networking OS, you can change the value of the LOCAL_PREFERENCE attribute, so that the preferred path can be changed.

To change the default values of this attribute for all routes received by the router, use the following command.

- Change the LOCAL_PREF value.

  CONFIG-ROUTER-BGP mode

  ```
  bgp default local-preference value
  ```

value: the range is from 0 to 4294967295. The default is **100**.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf_router_bgp)# neighbor 10.10.10.1 remote-as 500
DellEMC(conf_router_bgp)# bgp default local-preference 150
DellEMC(conf_router_bgp)# exit
```

In the above example configuration, the default LOCAL_PREFERENCE value is changed to 150 for all the updates from AS 500 to AS 400. The default value is 100.

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode or the `show running-config bgp` command in EXEC Privilege mode.

Manipulating the LOCAL_PREFERENCE attribute

A more flexible method for manipulating the LOCAL_PREFERENCE attribute value is to use a route-map.

To change the usage of LOCAL_PREFERENCE attribute to choose the best path among the advertised, enter any or all of the following commands.

1. Enter the ROUTE-MAP mode and assign a name to a route-map.
   CONFIGURATION mode
   `route-map map-name [permit | deny] [sequence-number]`
2. Change LOCAL_PREF value for routes meeting the criteria of this route-map.
   CONFIG-ROUTE-MAP mode
   `set local-preference value`
3. Return to CONFIGURATION mode.
   CONFIG-ROUTE-MAP mode
   `exit`
4. Enter ROUTER BGP mode.
   CONFIGURATION mode
   `router bgp as-number`
5. Apply the route-map to the neighbor or peer group's incoming or outgoing routes.
   CONFIG-ROUTER-BGP mode
   `neighbor {ip-address | ipv6-address | peer-group-name} route-map map-name {in | out}`

```
DellEMC# configure terminal
DellEMC(conf)# route-map route1 permit 10
DellEMC(conf-route-map)# set local-preference 140
DellEMC(conf-route-map)# exit
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# neighbor 10.10.10.2 route-map route1 in
```

In the above example configuration, a route-map named `route1` is created with a clause to set the specified local preference value, which is 140.

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode. To view a route-map configuration, use the `show route-map` command in EXEC Privilege mode.

Configuring the local System or a Different System to be the Next Hop for BGP-Learned Routes

You can configure the local router or a different router as the next hop for BGP-learned routes.

To change how the NEXT_HOP attribute is used, enter the first command. To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode or the `show running-config bgp` command in EXEC Privilege mode.

You can also use route maps to change this and other BGP attributes. For example, you can include the second command in a route map to specify the next hop address.

- Disable next hop processing and configure the router (route reflector) as the next hop for a BGP neighbor.
  CONFIG-ROUTER-BGP mode
  `neighbor {ip-address | ipv6-address | peer-group-name} next-hop-self [all]`
  If you do not use the `all` keyword, the next hop of only eBGP-learned routes is updated by the route reflector. If you use the `all` keyword, the next hop of both eBGP- and iBGP-learned routes are updated by the route reflector.
- Sets the next hop address.
  CONFIG-ROUTE-MAP mode
  `set next-hop ip-address`
  If the `set next-hop` command is applied on the out-bound interface using a route map, it takes precedence over the `neighbor next-hop-self` command.

Changing the WEIGHT Attribute

To change how the WEIGHT attribute is used, enter the first command. You can also use route maps to change this and other BGP attributes. For example, you can include the second command in a route map to specify the next hop address.

- Assign a weight to the neighbor connection.
  CONFIG-ROUTER-BGP mode
  `neighbor {ip-address | ipv6-address | peer-group-name} weight weight`
  weight: the range is from 0 to 65535. The default is **0**.
- Sets weight for the route.
  CONFIG-ROUTE-MAP mode
  `set weight weight`
  weight: the range is from 0 to 65535.
  NOTE: The weight assigned using the `set weight` command under route map configuration override the weight assigned using the `neighbor weight` command.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 remote-as 500
DellEMC(conf-router_bgp)# neighbor 10.10.10.1 weight 150
DellEMC(conf-router_bgp)# exit
```

In the above configuration example, the weight attribute for routes received from the neighbor 10.10.10.1 is set to 150.

To view BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode or the `show running-config bgp` command in EXEC Privilege mode.

Enabling Multipath

By default, the software allows one path to a destination. You can enable multipath to allow up to 64 parallel paths to a destination.

You can configure the maximum number of parallel routes (multipath support) to a destination in BGP.

NOTE: Dell EMC Networking recommends *not* using multipath and add path simultaneously in a route reflector.

To allow more than one path, use the following command.

- Enable multiple parallel paths.
  CONFIG-ROUTER-BGP mode
  `maximum-paths {ebgp | ibgp} number`
  Configure the following parameters:
  - `ebgp`: Enable multipath support for external BGP routes.
  - `ibgp`: Enable multipath support for internal BGP routes.

- *number*: Maximum number of parallel paths. The range is from 2 to 64.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# maximum-paths ibgp 5
DellEMC(conf-router_bgp)# exit
```

In the above example configuration, the maximum number of parallel internal BGP routes is set to 5, so that only 5 routes can be installed in a routing table.

The `show ip bgp network` command includes multipath information for that network.

Route Reflectors

Route reflectors reorganize the iBGP core into a hierarchy and allow some route advertisement rules.

NOTE: Do not use route reflectors (RRs) in the forwarding path. In iBGP, hierarchal RRs maintaining forwarding plane RRs could create routing loops.

Route reflection divides iBGP peers into two groups: client peers and nonclient peers. A route reflector and its client peers form a route reflection cluster. Because BGP speakers announce only the best route for a given prefix, route reflector rules are applied after the router makes its best path decision.

- If a route was received from a nonclient peer, reflect the route to all client peers.
- If the route was received from a client peer, reflect the route to all nonclient and all client peers.

To illustrate how these rules affect routing, refer to the following illustration and the following steps. Routers B, C, D, E, and G are members of the same AS (AS100). These routers are also in the same Route Reflection Cluster, where Router D is the Route Reflector. Router E and H are client peers of Router D; Routers B and C and nonclient peers of Router D.

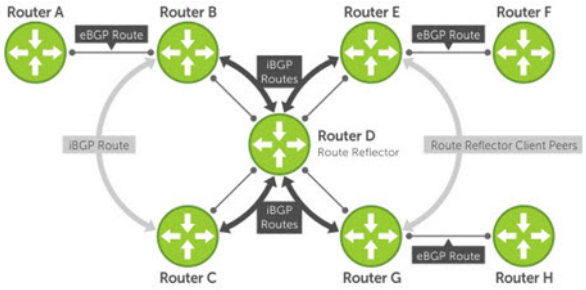


Figure 25. BGP Router Rules

1. Router B receives an advertisement from Router A through eBGP. Because the route is learned through eBGP, Router B advertises it to all its iBGP peers: Routers C and D.
2. Router C receives the advertisement but does not advertise it to any peer because its only other peer is Router D, an iBGP peer, and Router D has already learned it through iBGP from Router B.
3. Router D does not advertise the route to Router C because Router C is a nonclient peer and the route advertisement came from Router B who is also a nonclient peer.
4. Router D does reflect the advertisement to Routers E and G because they are client peers of Router D.
5. Routers E and G then advertise this iBGP learned route to their eBGP peers Routers F and H.

Configuring BGP Route Reflectors

BGP route reflectors are intended for ASs with a large mesh; they reduce the amount of BGP control traffic.

NOTE: Dell EMC Networking recommends *not* using multipath and add path simultaneously in a route reflector.

With route reflection configured properly, IBGP routers are not fully meshed within a cluster but all receive routing information.

Configure clusters of routers where one router is a concentration router and the others are clients who receive their updates from the concentration router.

To configure a route reflector, use the following commands.

- Assign a cluster ID or an IP address to a router reflector cluster.

  CONFIG-ROUTER-BGP mode

  `bgp cluster-id ip-address | number`

  - *ip-address*: IP address as the route reflector cluster ID.
  - *number*: A route reflector cluster ID as a number from 1 to 4294967295.

  You can have multiple clusters in an AS. When a BGP cluster contains only one route reflector, the cluster ID is the route reflector's router ID. For redundancy, a BGP cluster may contain two or more route reflectors. Without a cluster ID, the route reflector cannot recognize route updates from the other routes reflector within the cluster.
- Configure the local router as a route reflector and the specified neighbors or peer group as members of the cluster.

  CONFIG-ROUTER-BGP mode

  `neighbor {ip-address | ipv6-address | peer-group-name} route-reflector-client`

When you enter this command for the first time, the router configures as a route reflector and the specified BGP neighbors configure as clients in the route reflector cluster. When you remove all clients of a route reflector using the `no neighbor route-reflector-client` command, the router no longer functions as a route reflector. When you enable a route reflector, Dell EMC Networking OS automatically enables route reflection to all clients. To disable route reflection between all clients in this reflector, use the `no bgp client-to-client reflection` command in CONFIGURATION ROUTER BGP mode. All clients must be fully meshed before you disable route reflection.

To view a route reflector configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode or the `show running-config bgp` in EXEC Privilege mode.

Enabling Route Flap Dampening

When EBGP routes become unavailable, they "flap" and the router issues both WITHDRAWN and UPDATE notices.

A flap is when a route:
- is withdrawn
- is readvertised after being withdrawn
- has an attribute change

The constant router reaction to the WITHDRAWN and UPDATE notices causes instability in the BGP process. To minimize this instability, you may configure penalties (a numeric value) for routes that flap. When that penalty value reaches a configured limit, the route is not advertised, even if the route is up. In Dell EMC Networking OS, that penalty value is 1024. As time passes and the route does not flap, the penalty value decrements or is decayed. However, if the route flaps again, it is assigned another penalty.

The penalty value is cumulative and penalty is added under following cases:
- Withdraw
- Readvertise
- Attribute change

When dampening is applied to a route, its path is described by one of the following terms:
- history entry — an entry that stores information on a downed route
- dampened path — a path that is no longer advertised
- penalized path — a path that is assigned a penalty

To configure route flap dampening parameters, set dampening parameters using a route map, clear information on route dampening and return suppressed routes to active state, view statistics on route flapping, or change the path selection from the default mode (deterministic) to non-deterministic, use the following commands.

- Enable route dampening.

  CONFIG-ROUTER-BGP mode

  `bgp dampening [half-life | reuse | suppress max-suppress-time] [route-map map-name]`

  Enter the following optional parameters to configure route dampening parameters:
  - *half-life*: the range is from 1 to 45. Number of minutes after which the Penalty is decreased. After the router assigns a Penalty of 1024 to a route, the Penalty is decreased by half after the half-life period expires. The default is **15 minutes**.
  - *reuse*: the range is from 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is less than the reuse value, the flapping route is once again advertised (or no longer suppressed). Withdrawn routes are removed from history state. The default is **750**.

- *suppress*: the range is from 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is greater than the suppress value, the flapping route is no longer advertised (that is, it is suppressed). The default is **2000**.)
- *max-suppress-time*: the range is from 1 to 255. The maximum number of minutes a route can be suppressed. The default is four times the half-life value. The default is **60 minutes**.
- `route-map` *map-name*: name of a configured route map. Only match commands in the configured route map are supported. Use this parameter to apply route dampening to selective routes.

- Enter the following optional parameters to configure route dampening.

  CONFIG-ROUTE-MAP mode

  `set dampening` *half-life reuse suppress max-suppress-time*
  - `half-life`: the range is from 1 to 45. Number of minutes after which the Penalty is decreased. After the router assigns a Penalty of 1024 to a route, the Penalty is decreased by half after the half-life period expires. The default is **15 minutes**.
  - *reuse*: the range is from 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is less than the reuse value, the flapping route is once again advertised (or no longer suppressed). The default is **750**.
  - *suppress*: the range is from 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is greater than the suppress value, the flapping route is no longer advertised (that is, it is suppressed). The default is **2000**.
  - *max-suppress-time*: the range is from 1 to 255. The maximum number of minutes a route can be suppressed. The default is four times the half-life value. The default is **60 minutes**.
- Clear all information or only information on a specific route.

  EXEC Privilege

  `clear ip bgp [vrf` *vrf-name*`] dampening [`*ip-address mask*`]`
- View all flap statistics or for specific routes meeting the following criteria.

  EXEC or EXEC Privilege mode

  `show ip bgp [vrf` *vrf-name*`] flap-statistics [`*ip-address* `[`*mask*`]] [filter-list` *as-path-name*`] [regexp` *regular-expression*`]`
  - *ip-address* `[`*mask*`]`: enter the IP address and mask.
  - `filter-list` *as-path-name*: enter the name of an AS-PATH ACL.
  - `regexp` *regular-expression*: enter a regular express to match on.

  By default, the path selection in Dell EMC Networking OS is deterministic, that is, paths are compared irrespective of the order of their arrival. You can change the path selection method to non-deterministic, that is, paths are compared in the order in which they arrived (starting with the most recent). Furthermore, in non-deterministic mode, the software may not compare MED attributes though the paths are from the same AS.
- Change the best path selection method to non-deterministic.

  Change the best path selection method to non-deterministic.

  CONFIG-ROUTER-BGP mode

  `bgp non-deterministic-med`

  NOTE: When you change the best path selection method, path selection for existing paths remains unchanged until you reset it by entering the `clear ip bgp` command in EXEC Privilege mode.

To view the BGP configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode or the `show running-config bgp` command in EXEC Privilege mode.

The following example shows how to configure values to reuse or restart a route. In the following example, `default = 15` is the set time before the value decrements, `bgp dampening 2 ?` is the set re-advertise value, `bgp dampening 2 2000 ?` is the suppress value, and `bgp dampening 2 2000 3000 ?` is the time to suppress a route. Default values are also shown.

```
DellEMC(conf-router_bgp)#bgp dampening ?
<1-45> Half-life time for the penalty (default = 15)
route-map Route-map to specify criteria for dampening
<cr>
DellEMC(conf-router_bgp)#bgp dampening 2 ?
<1-20000>     Value to start reusing a route (default = 750)
DellEMC(conf-router_bgp)#bgp dampening 2 2000 ?
<1-20000>     Value to start suppressing a route (default = 2000)
DellEMC(conf-router_bgp)#bgp dampening 2 2000 3000 ?
<1-255>       Maximum duration to suppress a stable route (default = 60)
DellEMC(conf-router_bgp)#bgp dampening 2 2000 3000 10 ?
```

```
route-map     Route-map to specify criteria for dampening
<cr>
```

To view a count of dampened routes, history routes, and penalized routes when you enable route dampening, look at the seventh line of the `show ip bgp summary` command output, as shown in the following example (bold).

```
DellEMC>show ip bgp summary
BGP router identifier 10.114.8.131, local AS number 65515
BGP table version is 855562, main routing table version 780266
122836 network entrie(s) and 221664 paths using 29697640 bytes of memory
34298 BGP path attribute entrie(s) using 1920688 bytes of memory
29577 BGP AS-PATH entrie(s) using 1384403 bytes of memory
184 BGP community entrie(s) using 7616 bytes of memory
Dampening enabled. 0 history paths, 0 dampened paths, 0 penalized paths

Neighbor    AS    MsgRcvd MsgSent TblVer  InQ OutQ Up/Down State/PfxRcd
10.114.8.34 18508 82883   79977   780266  0   2 00:38:51   118904
10.114.8.33 18508 117265  25069   780266  0   20 00:38:50  102759
DellEMC>
```

To view which routes are dampened (non-active), use the `show ip bgp dampened-routes` command in EXEC Privilege mode.

Changing BGP keepalive and hold timers

BGP uses timers to control the activity of sending the keepalive messages to its neighbors or peers. Also, you can adjust the interval of how long the device has to wait for a keepalive messge from a neighbor before declaring the peer dead. To configure BGP timers, use either or both of the following commands.

To change the BGP timers for all neighbors, use `timers bgp` command. To change the BGP keepalive and holdtime timers for a specific neighbor or peer, use the `neighbor timers` command.

NOTE: Timer values configured with the `neighbor timers` command override the timer values configured with the `timers bgp` command.

When two neighbors, configured with different `keepalive` and `holdtime` values, negotiate for new values, the resulting values are as follows:

- the lower of the `holdtime` values is the new `holdtime` value, and
- whichever is the lower value; one-third of the new `holdtime` value, or the configured `keepalive` value is the new `keepalive` value.

- Configure timer values for a BGP neighbor or peer group.

  CONFIG-ROUTER-BGP mode

  `neighbors {ip-address | ipv6-address | peer-group-name} timers keepalive holdtime`

  - *keepalive*: Time interval, in seconds, between keepalive messages sent to the neighbor routers. The range is from 1 to 65535. The default is **60 seconds**.
  - *holdtime*: Time interval, in seconds, between the last keepalive message and declaring the BGP peer is dead. The range is from 3 to 65536. The default is **180 seconds**.
- Configure timer values for all neighbors.

  CONFIG-ROUTER-BGP mode

  `timers bgp keepalive holdtime`

  - *keepalive*: Time interval, in seconds, between keepalive messages sent to the neighbor routers. The range is from 1 to 65535. The default is **60 seconds**.
  - *holdtime*: Time interval, in seconds, between the last keepalive message and declaring the BGP peer is dead. The range is from 3 to 65536. The default is **180 seconds**.

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 400
DellEMC(conf-router_bgp)# network 10.1.1.0 mask 255.255.255.0
DellEMC(conf-router_bgp)# timers bgp 80 120
DellEMC(conf-router_bgp)# exit
```

In the above example configuration, the BGP timers are set with keepalive time as 80 seconds with which the system sends keepalive messages to the BGP peer and holdtime as 120 seconds with which the system waits for a message from the BGP peer before concluding that the peer is dead.

To view non-default values, use the `show config` command in CONFIGURATION ROUTER BGP mode or the `show running-config bgp` command in EXEC Privilege mode.

Setting the extended timer

To configure BGP idle hold time, use the following commands.

Timer values configured with the `neighbor timers extended` command override the timer values configured with the `timers bgp extended` command.

The peer remains in idle state based on the configured `idle-holdtime`. The less the `idle-holdtime`, lesser the peer remains in idle state.

NOTE: For the new `idle-holdtime` to take effect, you need to shutdown the respective peer manually using `neighbor shutdown` command and enable the peer again.

- Configure `idle-holdtime` values for a BGP neighbor or peer group.

  CONFIG-ROUTER-BGP mode

  `neighbors {`*`ip-address`* `|` *`ipv6-address`* `|` *`peer-group-name`*`} timers extended` *`idle-holdtime`*

  `idle-holdtime`: the range is from 1 to 32767. Time interval, in seconds, during which the peer remains in idle state. The default is **15 seconds**.
- Configure `idle-holdtime` values for all BGP neighbors.

  CONFIG-ROUTER-BGP mode

  `timers bgp extended` *`idle holdtime`*

  `idle-holdtime`: the range is from 1 to 32767. Time interval, in seconds, during which the peer remains in idle state. The default is **15 seconds**.

Enabling or disabling BGP neighbors

You can enable or disable all the configured BGP neighbors using the `shutdown all` command in ROUTER BGP mode.

To disable all the configured BGP neighbors:

1. Enter the `router bgp` mode using the following command:

   CONFIGURATION Mode

   ```
   router bgp as-number
   ```
2. In ROUTER BGP mode, enter the following command:

   ROUTER BGP Mode

   ```
   shutdown all
   ```

   You can use the `no shutdown all` command in the ROUTER BGP mode to re-enable all the BGP interface.

You can also enable or disable BGP neighbors corresponding to the IPv4 unicast or multicast address families and the IPv6 unicast address family.

To enable or disable BGP neighbors corresponding to the IPv4 unicast address families, use the following commands:

1. Enter the router bgp mode using the following command:

   CONFIGURATION Mode

   ```
   router bgp as-number
   ```

2. Shut down the BGP neighbors corresponding to the IPv4 unicast address family using the following command:

```
shutdown address-family-ipv4-unicast
```

To enable or disable BGP neighbors corresponding to IPv4 multicast address family:

1. Enter the router bgp mode using the following command:

   CONFIGURATION Mode

   ```
   router bgp as-number
   ```

2. Shut down the BGP neighbors corresponding to IPv4 multicast address family using the following command:

   ROUTER-BGP Mode

   ```
   shutdown address-family-ipv4-multicast
   ```

To enable or disable BGP neighbors corresponding to the IPv6 unicast address family:

1. Enter the `router bgp` mode using the following command:

   CONFIGURATION Mode

   ```
   router bgp as-number
   ```

2. Shut down the BGP neighbors corresponding to the IPv6 unicast address family using the following command:

   ROUTER-BGP Mode

   ```
   shutdown address-family-ipv6-unicast
   ```

When you configure BGP, you must explicitly enable the BGP neighbors using the following commands:

```
neighbor {ip-address | peer-group name} remote-as as-number
neighbor {ip-address | peer-group-name} no shutdown
```

For more information on enabling BGP, see Enabling BGP.

When you use the `shutdown all` command in global configuration mode, this command takes precedence over the `shutdown address-family-ipv4-unicast`, `shutdown address-family-ipv4-multicast`, and `shutdown address-family-ipv6-unicast` commands. Irrespective of whether the BGP neighbors are disabled earlier, the `shutdown all` command brings down all the configured BGP neighbors.

When you issue the `no shutdown all` command, all the BGP neighbor neighbors are enabled. However, when you re-enable all the BGP neighbors in global configuration mode, only the neighbors that were not in disabled state before the global shutdown come up.

Meaning, BGP neighbors corresponding to the IPv4 unicast or multicast address family and the IPv6 unicast address family that were explicitly disabled before the global shutdown remains in disabled state. Use the `no shutdown address-family-ipv4-unicast`, `no shutdown address-family-ipv4-multicast`, or `no shutdown address-family-ipv6-unicast` commands to enable these neighbors.

NOTE: This behavior applies to all BGP neighbors. Meaning, BGP neighbors that were explicitly disabled before global shutdown also remain in disabled state. Enable these neighbors individually using the `no shutdown` command.

Route Map Continue

The BGP route map continue feature, `continue [sequence-number]`, (in ROUTE-MAP mode) allows movement from one route-map entry to a specific route-map entry (the sequence number).

If you do not specify a sequence number, the continue feature moves to the next sequence number (also known as an “implied continue”). If a match clause exists, the continue feature executes only after a successful match occurs. If there are no successful matches, continue is ignored.

Match a Clause with a Continue Clause

The continue feature can exist without a match clause.

Without a match clause, the continue clause executes and jumps to the specified route-map entry. With a match clause and a continue clause, the match clause executes first and the continue clause next in a specified route map entry. The continue clause launches only after a successful match. The behavior is:

- A successful match with a continue clause—the route map executes the set clauses and then goes to the specified route map entry after execution of the continue clause.
- If the next route map entry contains a continue clause, the route map executes the continue clause if a successful match occurs.
- If the next route map entry does not contain a continue clause, the route map evaluates normally. If a match does not occur, the route map does not continue and falls-through to the next sequence number, if one exists

Set a Clause with a Continue Clause

If the route-map entry contains sets with the continue clause, the set actions operation is performed first followed by the continue clause jump to the specified route map entry.

- If a set actions operation occurs in the first route map entry and then the same set action occurs with a different value in a subsequent route map entry, the last set of actions overrides the previous set of actions with the same `set` command.
- If the `set community additive` and `set as-path prepend` commands are configured, the communities and AS numbers are prepended.

Configuring BGP Confederations

Another way to organize routers within an AS and reduce the mesh for IBGP peers is to configure BGP confederations.

As with route reflectors, BGP confederations are recommended only for IBGP peering involving many IBGP peering sessions per router. Basically, when you configure BGP confederations, you break the AS into smaller sub-AS, and to those outside your network, the confederations appear as one AS. Within the confederation sub-AS, the IBGP neighbors are fully meshed and the MED, NEXT_HOP, and LOCAL_PREF attributes are maintained between confederations.

To configure BGP confederations, use the following commands.

- Specifies the confederation ID.

  CONFIG-ROUTER-BGP mode

  `bgp confederation identifier` *`as-number`*

  - *`as-number`*: from 0 to 65535 (2 Byte) or from 1 to 4294967295 (4 Byte).

- Specifies which confederation sub-AS are peers.

  CONFIG-ROUTER-BGP mode

  `bgp confederation peers` *`as-number [... as-number]`*

  - *`as-number`*: from 0 to 65535 (2 Byte) or from 1 to 4294967295 (4 Byte).

  All Confederation routers must be either 4 Byte or 2 Byte. You cannot have a mix of router ASN support.

To view the configuration, use the `show config` command in CONFIGURATION ROUTER BGP mode.

Configuring a BGP VRF address family

To perform BGP configuration between two neighbors that must exchange IPv6 or IPv4 VRF information, use the following commands.

Following are the steps to configure BGP VRF address-family between two peers.

- Configure a VRF routing table.

  CONFIG mode

  `ip vrf` *`vrf-name`*

  For more information on VRF configuration, see Virtual Routing and Forwarding (VRF).

- Enter the router configuration mode and the AS number.

  CONFIG mode

  `router bgp as-number`
- Specify the IPv4 address family configuration.

  CONFIG-ROUTER-BGP mode

  `address-family {ipv4 [multicast] | ipv6 unicast} vrf vrf-name`

  `vrf vrf-name` — Specifies the name of VRF instance associated with the IPv4 or IPv6 address-family configuration.
- Add the IP address of the neighbor in the specified AS to the IPv4 MBGP neighbor table.

  CONFIG-ROUTER-BGP-AF mode

  `neighbor {ip-address | ipv6-address | peer-group-name} remote-as as-number`
- Enable the neighbor.

  CONFIG-ROUTER-BGP mode

  `neighbor ip-address no shutdown`
- Specify the number of prefixes that can be received from a neighbor.

  CONFIG-ROUTER-BGP-AF mode

  `neighbor {ip-address | ipv6-address | peer-group-name} maximum—prefix maximum [threshold] [warning-only]as-number`

The following are the sample steps performed to configure a VRF, and VRF address families for IPv4 (unicast and multicast) and IPv6.

```
DellEMC(conf)# ip vrf vrf1
DellEMC(conf-vrf)# exit
DellEMC(conf)# router bgp 100
DellEMC(conf-router_bgp)# address-family ipv4 vrf vrf1
DellEMC(conf-router_bgp_af)# neighbor 50.0.0.2 remote-as 200
DellEMC(conf-router_bgp_af)# neighbor 50.0.0.2 maximum-prefix 10000 warning-only
DellEMC(conf-router_bgp_af)# neighbor 50.0.0.2 no shutdown
DellEMC(conf-router_bgp_af)# exit
DellEMC(conf-router_bgp)# address-family ipv6 unicast vrf vrf1
DellEMC(conf-router_bgpv6_af)# neighbor 50.0.0.2 activate
DellEMC(conf-router_bgpv6_af)# exit
DellEMC(conf-router_bgp)# address-family ipv4 multicast vrf vrf1
DellEMC(conf-router_bgpv6_af)# neighbor 50.0.0.2 activate
DellEMC(conf-router_bgp)# exit
```

Following is the output of `show ip bgp vrf vrf1 summary` command for the above configuration.

```
DellEMC#show ip bgp vrf vrf1 summary
BGP router identifier 1.1.1.1, local AS number 100
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 16384 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
50.0.0.2        200                 0        0          0    0     0 00:00:00 0
```

Following is the output of `show running-config bgp` command for the above configuration.

```
DellEMC#show running-config bgp
!
router bgp 10
 bgp router-id 1.1.1.1
 network 10.10.21.0/24
 bgp four-octet-as-support
 neighbor 20.20.20.1 remote-as 65550
 neighbor 20.20.20.1 no shutdown
 !
 address-family ipv4 vrf vrf1
  neighbor 50.0.0.2 maximum-prefix 10000 warning-only
  neighbor 50.0.0.2 remote-as 200
  neighbor 50.0.0.2 no shutdown
  exit-address-family
 !
 address-family ipv4 multicast vrf vrf1
```

```
   neighbor 50.0.0.2 activate
   exit-address-family
  !
  address-family ipv6 unicast vrf vrf1
   neighbor 50.0.0.2 activate
   exit-address-family
DellEMC#
```

Maintaining Existing AS Numbers During an AS Migration

The local-as feature smooths out the BGP network migration operation and allows you to maintain existing ASNs during a BGP network migration.

When you complete your migration, be sure to reconfigure your routers with the new information and disable this feature.

- Allow external routes from this neighbor.

  CONFIG-ROUTERBGP mode

  `neighbor {ip-address | ipv6-address | peer-group-name} local-as as number [no prepend]`

  - *peer-group-name*: 16 characters.
  - *AS-number*: 0 to 65535 (2-Byte) or 1 to 4294967295 (4-Byte) or 0.1 to 65535.65535 (Dotted format).
  - `No Prepend`: specifies that local AS values are not prepended to announcements from the neighbor.

  Format: IPv4 Address: A.B.C.D and IPv6 address: X:X:X:X::X.

  You must Configure Peer Groups *before* assigning it to an AS. This feature is not supported on passive peer groups.

The first line in bold shows the actual AS number. The second two lines in bold show the local AS number (6500) maintained during migration.

To disable this feature, use the `no neighbor local-as` command in CONFIGURATION ROUTER BGP mode.

```
R2(conf-router_bgp)#show conf
!
router bgp 65123
  bgp router-id 192.168.10.2
  network 10.10.21.0/24
  network 10.10.32.0/24
  network 100.10.92.0/24
  network 192.168.10.0/24
  bgp four-octet-as-support
  neighbor 10.10.21.1 remote-as 65123
  neighbor 10.10.21.1 filter-list Laura in
  neighbor 10.10.21.1 no shutdown
  neighbor 10.10.32.3 remote-as 65123
  neighbor 10.10.32.3 no shutdown
  neighbor 100.10.92.9 remote-as 65192
 neighbor 100.10.92.9 local-as 6500
 neighbor 100.10.92.9 no shutdown
  neighbor 192.168.10.1 remote-as 65123
  neighbor 192.168.10.1 update-source Loopback 0
  neighbor 192.168.10.1 no shutdown
  neighbor 192.168.12.2 remote-as 65123
  neighbor 192.168.12.2 update-source Loopback 0
  neighbor 192.168.12.2 no shutdown
R2(conf-router_bgp)#
```

Allowing an AS Number to Appear in its Own AS Path

This command allows you to set the number of times a particular AS number can occur in the AS path.

The allow-as feature permits a BGP speaker to allow the ASN to be present for a specified number of times in the update received from the peer, even if that ASN matches its own. The AS-PATH loop is detected if the local ASN is present more than the specified number of times in the command.

- Allow this neighbor ID to use the AS path the specified number of times.

  CONFIG-ROUTER-BGP mode

  `neighbor {ip-address | ipv6-address | peer-group-name} allowas-in number`

- *peer-group-name*: 16 characters.
- *Number*: 1 through 10.

Format: IP Address: A.B.C.D and IPv6 adress: X:X:X:X::X.

You must Configure Peer Groups before assigning it to an AS.

The lines shown in bold are the number of times ASN 65123 can appear in the AS path (**allows–in 9**).

To disable this feature, use the `no neighbor allow-as in number` command in CONFIGURATION ROUTER BGP mode.

```
R2(conf-router_bgp)#show conf
!
router bgp 65123
  bgp router-id 192.168.10.2
  network 10.10.21.0/24
  network 10.10.32.0/24
  network 100.10.92.0/24
  network 192.168.10.0/24
  bgp four-octet-as-support
  neighbor 10.10.21.1 remote-as 65123
  neighbor 10.10.21.1 filter-list Laura in
  neighbor 10.10.21.1 no shutdown
  neighbor 10.10.32.3 remote-as 65123
  neighbor 10.10.32.3 no shutdown
  neighbor 100.10.92.9 remote-as 65192
  neighbor 100.10.92.9 local-as 6500
  neighbor 100.10.92.9 no shutdown
  neighbor 192.168.10.1 remote-as 65123
  neighbor 192.168.10.1 update-source Loopback 0
  neighbor 192.168.10.1 no shutdown
  neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 allowas-in 9
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2(conf-router_bgp)#R2(conf-router_bgp)#
```

Enabling MBGP Configurations

Multiprotocol BGP (MBGP) is an enhanced BGP that carries IP multicast routes. BGP carries two sets of routes: one set for unicast routing and one set for multicast routing. The routes associated with multicast routing are used by the protocol independent multicast (PIM) to build data distribution trees.

Dell EMC Networking OS MBGP is implemented per RFC 1858. You can enable the MBGP feature per router and/or per peer/peer-group.

The default is **IPv4 Unicast** routes.

When you configure a peer to support IPv4 multicast, Dell EMC Networking OS takes the following actions:

- Send a capability advertisement to the peer in the BGP Open message specifying IPv4 multicast as a supported AFI/SAFI (Subsequent Address Family Identifier).
- If the corresponding capability is received in the peer's Open message, BGP marks the peer as supporting the AFI/SAFI.
- When exchanging updates with the peer, BGP sends and receives IPv4 multicast routes if the peer is marked as supporting that AFI/SAFI.
- Exchange of IPv4 multicast route information occurs through the use of two new attributes called MP_REACH_NLRI and MP_UNREACH_NLRI, for feasible and withdrawn routes, respectively.
- If the peer has not been activated in any AFI/SAFI, the peer remains in Idle state.

Most Dell EMC Networking OS BGP IPv4 unicast commands are extended to support the IPv4 multicast RIB using extra options to the command. For a detailed description of the MBGP commands, refer to the *Dell EMC Networking OS Command Line Interface Reference Guide*.

MBGP support for IPv6

MBGP suports IPv6 with same features and functionality as IPv4 BGP.

MBGP for IPv6 supports IPv6 address-family and Network Layer Reachability Information (NLRI) and next hop that uses IPv6 address.

Configuring IPv6 MBGP between peers

To configure IPv6 MBGP, use the following commands.

Following are the steps to configure IPv6 MBGP between two peers. The neighbors that are configured using `neighbor remote-as` command exchange only the IPv4 unicast address prefixes. In order to exchange IPv6 address prefixes, you have to activate the neighbors using `neighbor activate` command inside the address-family configuration.

NOTE: When configuring a BGP peer under IPv6 address-family, both IPv4 and IPv6 prefixes are sent by default. If you do not want a peer to exchange IPv4 prefixes, manually deactivate the peer with the `no neighbor activate` command under the CONFIGURATION-ROUTER-BGP mode.

- Enter the router configuration mode and the AS number.
  CONFIG mode
  `router bgp as-number`
- Add the IP address of the neighbor for the specified autonomous system.
  CONFIG-ROUTER-BGP mode
  `neighbor {ip-address | ipv6-address | peer-group-name} remote-as as-number`
- Specify the IPv6 address family configuration.
  CONFIG-ROUTER-BGP mode
  `address-family ipv6 [unicast | vrf vrf-name]`
  `unicast` — Specifies the IPv6 unicast address family. The default address-family is IPv6 unicast.
  `vrf vrf-name` — Specifies the name of VRF instance associated with the IPv6 address-family configuration.
- Enable the neighbor to exchange prefixes for IPv6 unicast address family.
  CONFIG-ROUTER-BGP-AF (Address Family) mode
  `neighbor {ip-address | ipv6-address | peer-group-name} activate`

```
DellEMC(conf)# router bgp 10
DellEMC(conf-router_bgp)# neighbor 2001::1 remote-as 200
DellEMC(conf-router_bgp)# neighbor 2001::1 no shutdown
DellEMC(conf-router_bgp)# address-family ipv6 unicast
DellEMC(conf-router_bgpv6_af)#neighbor 2001::1 activate
DellEMC(conf-router_bgpv6_af)#exit
```

Following is the output of `show ip bgp ipv6 unicast summary` command for the above configuration example.

```
DellEMC#show ip bgp ipv6 unicast summary
BGP router identifier 1.1.1.1, local AS number 10
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 8192 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
2001::1         200                 0        0          0    0     0 00:00:00 0
DellEMC#
```

Example-Configuring IPv4 and IPv6 neighbors

The following example configurations show how to enable BGP and set up some peer under IPv4 and IPv6 address families.

To support your own IP addresses, interfaces, names, and so on, you can copy and paste from these examples to your CLI. Be sure that you make the necessary changes.

Example-Configuring IPv4 and IPv6 neighbors

Example of enabling BGP and address family configuration in router (R1)

Following is an example to enable BGP and address family configuration for the neighbor R2 (20.20.20.2) in the router R1.

```
R1(conf)# router bgp 10
R1(conf-router_bgp)# neighbor 20.20.20.2 remote-as 200
R1(conf-router_bgp)# neighbor 20.20.20.2 no shutdown
R1(conf-router_bgp)# neighbor 2001::2 remote-as 200
R1(conf-router_bgp)# neighbor 2001::2 no shutdown
R1(conf-router_bgp)# neighbor 30.30.30.1 remote-as 20
R1(conf-router_bgp)# neighbor 30.30.30.1 no shutdown
R1(conf-router_bgp)# neighbor 3000::1 remote-as 300
R1(conf-router_bgp)# neighbor 3000::1 no shutdown
R1(conf-router_bgp)# no neighbor 3000::1 activate
R1(conf-router_bgp)# address-family ipv4 multicast
R1(conf-router_bgp_af)# neighbor 20.20.20.2 activate
R1(conf-router_bgp_af)# exit
R1(conf-router_bgp)# address-family ipv6 unicast
R1(conf-router_bgpv6_af)# neighbor 20.20.20.2 activate
R1(conf-router_bgpv6_af)# neighbor 2001::2 activate
R1(conf-router_bgpv6_af)#exit
```

The neighbors configured under CONFIGURATION-ROUTER-BGP mode and are defined using the `neighbor remote-as` command exchange only IPv4 unicast address prefixes. If you want a neighbor to exchange other prefixes such as IPv4 multicast or IPv6 unicast, you have to explicitly activate the respective neighbor using `neighbor activate` command in the respective IPv4 multicast or IPv6 unicast address family configuration. In the above example configuration, activating the neighbor (20.20.20.2) under the IPv4 multicast address family enables the neighbor to exchange IPv4 muticast prefixes. Similarly, activating the neighbors (20.20.20.2 and 2001::2) under the IPv6 unicast address family enables the neighbor to exchange IPv6 unicast prefixes. The neighbor (30.30.30.1) is activated by default for exchanging IPv4 unicast address prefixes, but will not exchange IPv4 multicast and IPv6 unicast address prefixes since they are not activated under the respective address family. If you want the neighbor (30.30.30.1) to exchange IPv4 multicast and/or IPv6 unicast prefixes, you have to explicitly active the neighbor using `neighbor activate` command.

If you do not want a neighbor to exchange IPv4 unicast prefixes, you have to manually deactivate the peer with the `no neighbor activate` command under the CONFIGURATION-ROUTER-BGP mode. If any neighbor is already activated to exchange IPv4 multicast or IPv6 unicast prefixes, exchanging of prefixes can be deactivated using `no neighbor activate` command under the IPv4 multicast or IPv6 unicast address family. In the above example configuration, the peer (3000::1) is deactivated from exchanging IPv4 unicast prefixes.

The `show ip bgp summary` or `show ip bgp ipv4 unicast summary` displays IPv4 unicast address family configuration. In a dual stack scenario, to view the specific IPv4 multicast or IPv6 unicast address family configuration, use `show ip bgp ipv4 multicast summary` or `show ip bgp ipv6 unicast summary` commands.

Following is the sample output for `show ip bgp summary` command.

```
R1#show ip bgp summary
BGP router identifier 1.1.1.1, local 10
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
3 neighbor(s) using 40960 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.2      200                10       20          0    0     0 00:06:11 0
30.30.30.1       20                 0        0          0    0     0 00:00:00 0
2001::2         200                40       45          0    0     0 00:03:14 0
```

The same output will be displayed when using `show ip bgp ipv4 unicast summary` command.

Following is the sample output of `show ip bgp ipv4 multicast summary` command.

```
R1# show ip bgp ipv4 multicast summary
BGP router identifier 1.1.1.1, local AS number 10
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 24576 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.2      200                10       20          0    0     0 00:06:11 0
R1#
```

Following is the output of `show ip bgp ipv6 unicast summary` command for the above configuration example.

```
R1#show ip bgp ipv6 unicast summary
BGP router identifier 1.1.1.1, local AS number 10
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
2 neighbor(s) using 24576 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.2     200               10       20          0    0     0 00:06:11 0
2001::2        200               40       45          0    0     0 00:03:14 0
R1#
```

Example of enabling BGP and address family configuration in router (R2)

Following is an example to enable BGP and address family configuration for the neighbor R1 (20.20.20.1) in the router R2.

```
R2(conf)# router bgp 200
R2(conf-router_bgp)# neighbor 20.20.20.1 remote-as 10
R2(conf-router_bgp)# neighbor 20.20.20.1 no shutdown
R2(conf-router_bgp)# neighbor 2001::1 remote-as 10
R2(conf-router_bgp)# neighbor 2001::1 no shutdown
R2(conf-router_bgp)# address-family ipv4 multicast
R2(conf-router_bgp_af)# neighbor 20.20.20.1 activate
R2(conf-router_bgp_af)# exit
R2(conf-router_bgp)# address-family ipv6 unicast
R2(conf-router_bgpv6_af)# neighbor 20.20.20.1 activate
R2(conf-router_bgpv6_af)# neighbor 2001::1 activate
R2(conf-router_bgpv6_af)#exit
```

Following is the sample output for `show ip bgp summary` command.

```
R2#show ip bgp summary
BGP router identifier 2.2.2.2, local 200
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
2 neighbor(s) using 40960 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      10                10       20          0    0     0 00:06:11 0
2001::1         10                40       45          0    0     0 00:03:14 0
```

Following is the sample output of `show ip bgp ipv4 multicast summary` command.

```
R2# show ip bgp ipv4 multicast summary
BGP router identifier 2.2.2.2, local AS number 200
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
1 neighbor(s) using 24576 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      10                10       20          0    0     0 00:06:11 0
R2#
```

Following is the output of `show ip bgp ipv6 unicast summary` command for the above configuration example.

```
R2#show ip bgp ipv6 unicast summary
BGP router identifier 2.2.2.2, local AS number 200
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
2 neighbor(s) using 24576 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
20.20.20.1      10                10       20          0    0     0 00:06:11 0
2001::1         10                40       45          0    0     0 00:03:14 0
R2#
```

Configure IPv6 NH Automatically for IPv6 Prefix Advertised over IPv4 Neighbor

You can configure the system to pick the next hop IPv6 address dynamically for IPv6 prefix advertised over an IPv4 neighbor configured under IPv6 address family. If there is no IPv6 address configured on the local interface, the system uses the IPv4 mapped IPv6 address. If there are multiple IPv6 addresses configured on the interface, the system uses the lowest IPv6 address configured on that interface. If the configuration is already present and a new IPv6 address, which is lower than the lowest existing address, is assigned to one of the peer interfaces, that address is not used as the NH until you flap the interface manually.

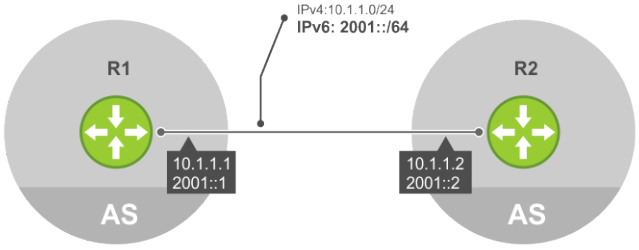


Figure 26. Configure IPv6 NH to advertise IPv6 prefix over IPv4 neighbor

To enable BGP to pick the next hop IPv6 address automatically for IPv6 prefix advertised over an IPv4 neighbor, use the following commands.

- Enable the system to pick the next hop IPv6 address dynamically for IPv6 prefix advertised over an IPv4 neighbor.

  ROUTER-BGP mode mode

  `neighbor {ip-address | peer-group-name} auto-local-address`

  Enter either the neighbor IP address or the name of the peer group.

  NOTE: If outbound policy is applied along with auto-local-address enabled, then policy takes precedence.

The following example configuration demonstrates how to configure BGP to automatically pick IPv6 address for IPv6 prefix advertised over an IPv4 neighbor.

Example configuration performed in R1

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 655
DellEMC(conf-router_bgp)# neighbor 10.1.1.2 remote-as 20
DellEMC(conf-router_bgp)# neighbor 10.1.1.2 auto-local-address
DellEMC(conf-router_bgp)# neighbor 10.1.1.2 no shutdown
DellEMC(conf-router_bgp)# bgp router-id 1.1.1.1
DellEMC(conf-router_bgp)# address-family ipv6 unicast
DellEMC(conf-router_bgpv6_af)# neighbor 10.1.1.2 activate
DellECM(conf-router_bgpv6_af)# exit
```

Configuring the auto-local-address for a neighbor will dynamically pick the local BGP interface IPv6 address (2001::1/64) as a the next hop for all the updates over IPv4 neighbor configured under IPv6 address family. If the auto-local-address is not configured, the IPv4 mapped IPv6 address (10.1.1.1) as a next hop.

Following is the `show running-config` command output for the above configuration.

```
DellEMC# show running-config bgp
!
router bgp 655
 bgp router-id 1.1.1.1
 neighbor 10.1.1.2 remote-as 20
 neighbor 10.1.1.2 auto-local-address
 neighbor 10.1.1.2 no shutdown
 !
 address-family ipv6 unicast
  neighbor 10.1.1.2 activate
```

```
  exit-address-family
 !
```

Example configuration performed in R2

```
DellEMC# configure terminal
DellEMC(conf)# router bgp 20
DellEMC(conf-router_bgp)# neighbor 10.1.1.1 remote-as 655
DellEMC(conf-router_bgp)# neighbor 10.1.1.1 no shutdown
 DellEMC(conf-router_bgp)# address-family ipv6 unicast
DellEMC(conf-router_bgpv6_af)# neighbor 10.1.1.1 activate
DellECM(conf-router_bgpv6_af)# exit
```

Following is the `show running-config` command output for the above configuration.

```
DellEMC# show running-config bgp
!
router bgp 20
neighbor 10.1.1.1 remote-as 655
neighbor 10.1.1.1 no shutdown
!
address-family ipv6 unicast
  neighbor 10.1.1.1 activate
  exit-address-family
!
```

Following is the `show ip bgp ipv6 unicast` command output for the above configuration.

```
DellEMC# show ip bgp ipv6 unicast
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
BGP local router ID is 1.1.1.1
Status codes: s suppressed, S stale, d dampened, h history, * valid, > best
Path source: I - internal, a - aggregate, c - confed-external, r - redistributed
             n - network,  D - denied,  S - stale
Origin codes: i - IGP, e - EGP, ? - incomplete

     Network            Next Hop            Metric     LocPrf Weight Path
*> 2001::/64      2001::1         0         0 655 ?
*>  3001::/64           3001::1                  0                  0 655 ?
*>  4001::/64           3001::1                  0                  0 655 ?
*>  5001::/64           3001::1                  0                  0 655 ?
```

BGP Regular Expression Optimization

Dell EMC Networking OS optimizes processing time when using regular expressions by caching and re-using regular expression evaluated results, at the expense of some memory in RP1 processor.

BGP policies that contain regular expressions to match against as-paths and communities might take a lot of CPU processing time, thus affect BGP routing convergence. Also, `show bgp` commands that get filtered through regular expressions can to take a lot of CPU cycles, especially when the database is large.

This feature is turned on by default. If necessary, use the `bgp regex-eval-optz-disable` command in CONFIGURATION ROUTER BGP mode to disable it.

Debugging BGP

To enable BGP debugging, use any of the following commands.

- View all information about BGP, including BGP events, keepalives, notifications, and updates.

  EXEC Privilege mode

  `debug ip bgp [`*`ip-address`*` | peer-group `*`peer-group-name`*`] [in | out]`
- View information about BGP route being dampened.

  EXEC Privilege mode

  `debug ip bgp dampening [in | out]`

- View information about local BGP state changes and other BGP events.

  EXEC Privilege mode

  `debug ip bgp [ip-address | peer-group peer-group-name] events [in | out]`
- View information about BGP KEEPALIVE messages.

  EXEC Privilege mode

  `debug ip bgp [ip-address | peer-group peer-group-name] keepalive [in | out]`
- View information about BGP notifications received from or sent to neighbors.

  EXEC Privilege mode

  `debug ip bgp [ip-address | peer-group peer-group-name] notifications [in | out]`
- View information about BGP updates and filter by prefix name.

  EXEC Privilege mode

  `debug ip bgp [ip-address | peer-group peer-group-name] updates [in | out] [prefix-list name]`
- Enable soft-reconfiguration debug.

  EXEC Privilege mode

  `debug ip bgp {ip-address | peer-group-name} soft-reconfiguration`

  To enhance debugging of soft reconfig, use the `bgp soft-reconfig-backup` command only when route-refresh is not negotiated to avoid the peer from resending messages.

  In-BGP is shown using the `show ip protocols` command.

Dell EMC Networking OS displays debug messages on the console. To view which debugging commands are enabled, use the `show debugging` command in EXEC Privilege mode.

To disable a specific `debug` command, use the keyword `no` then the `debug` command. For example, to disable debugging of BGP updates, use `no debug ip bgp updates` command.

To disable all BGP debugging, use the `no debug ip bgp` command.

To disable all debugging, use the `undebug all` command.

Storing Last and Bad PDUs

Dell EMC Networking OS stores the last notification sent/received and the last bad protocol data unit (PDU) received on a per peer basis. The last bad PDU is the one that causes a notification to be issued.

In the following example, the last seven lines shown in bold are the last PDUs.

Example of the `show ip bgp neighbor` Command to View Last and Bad PDUs

```
DellEMC(conf-router_bgp)#do show ip bgp neighbors 1.1.1.2

BGP neighbor is 1.1.1.2, remote AS 2, external link
  BGP version 4, remote router ID 2.4.0.1
  BGP state ESTABLISHED, in this state for 00:00:01
  Last read 00:00:00, last write 00:00:01
  Hold time is 90, keepalive interval is 30 seconds
  Received 1404 messages, 0 in queue
    3 opens, 1 notifications, 1394 updates
    6 keepalives, 0 route refresh requests
  Sent 48 messages, 0 in queue
    3 opens, 2 notifications, 0 updates
    43 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

Capabilities received from neighbor for IPv4 Unicast :
  MULTIPROTO_EXT(1)
  ROUTE_REFRESH(2)
  CISCO_ROUTE_REFRESH(128)

Capabilities advertised to neighbor for IPv4 Unicast :
  MULTIPROTO_EXT(1)
  ROUTE_REFRESH(2)
  CISCO_ROUTE_REFRESH(128)
```

```
For address family: IPv4 Unicast
BGP table version 1395, neighbor version 1394
Prefixes accepted 1 (consume 4 bytes), 0 withdrawn by peer
Prefixes advertised 0, rejected 0, 0 withdrawn from peer

Connections established 3; dropped 2
Last reset 00:00:12, due to Missing well known attribute

Notification History
  'UPDATE error/Missing well-known attr' Sent : 1 Recv: 0
  'Connection Reset' Sent : 1 Recv: 0

  Last notification (len 21) sent 00:26:02 ago
  ffffffff ffffffff ffffffff ffffffff 00160303 03010000
 Last notification (len 21) received 00:26:20 ago
  ffffffff ffffffff ffffffff ffffffff 00150306 00000000
 Last PDU (len 41) received 00:26:02 ago that caused notification to be issued
  ffffffff ffffffff ffffffff ffffffff 00290200 00000e01 02040201 00024003 04141414 0218c0a8
  01000000
Local host: 1.1.1.1, Local port: 179
Foreign host: 1.1.1.2, Foreign port: 41758
```

PDU Counters

Dell EMC Networking OS supports additional counters for various types of PDUs sent and received from neighbors.

These are seen in the output of the `show ip bgp neighbor` command.

10

Content Addressable Memory (CAM)

CAM is a type of memory that stores information in the form of a lookup table. On Dell EMC Networking systems, CAM stores Layer 2 (L2) and Layer 3 (L3) forwarding information, access-lists (ACLs), flows, and routing policies.

Topics:

- CAM Allocation
- Test CAM Usage
- View CAM Profiles
- View CAM-ACL Settings
- View CAM Usage
- Configuring CAM Threshold and Silence Period
- CAM Optimization
- Troubleshoot CAM Profiling

CAM Allocation

CAM Allocation for Ingress

To allocate the space for regions such has L2 ingress ACL, IPV4 ingress ACL, IPV6 ingress ACL, IPV4 QoS, L2 QoS, PBR, VRF ACL, and so forth, use the `cam-acl` command in CONFIGURATION mode.

The CAM space is allotted in field processor (FP) blocks. The total space allocated must equal FP blocks.

The following table lists the default CAM allocation settings.

NOTE: There are FP blocks, but the system flow requires three blocks that cannot be reallocated.

The following table displays the default CAM allocation settings. To display the default CAM allocation, enter the `show cam-acl` command.

Table 11. Default Cam Allocation Settings

CAM Allocation	Setting
L2Acl	6
IPV4Acl	4
Ipv6Acl	0
Ipv4Qos	2
L2Qos	1
L2PT	0
IpMacAcl	0
VmanQos	0
VmanDualQos	0
EcfmAcl	0
nlbclusteracl	0
FcoeAcl	0
iscsiOptAcl	0
ipv4pbr	0

Table 11. Default Cam Allocation Settings (continued)

CAM Allocation	Setting
vrfv4Acl	0
Openflow	0
fedgovacl	0

NOTE: When you reconfigure CAM allocation, use the **nlbclusteracl** *number* command to change the number of NLB ARP entries. The range is from 0 to 2. The default value is 0. At the default value of 0, eight NLB ARP entries are available for use. This platform supports upto 256 CAM entries. Select 1 to configure 128 entries. Select 2 to configure 256 entries. Even though you can perform CAM carving to allocate the maximum number of NLB entries, Dell EMC Networking recommends you to use a maximum of 64 NLB ARP entries.

The following additional CAM allocation settings are supported.

Table 12. Additional Default CAM Allocation Settings

Additional CAM Allocation	Setting
FCoE ACL (fcoeacl)	0
ISCSI Opt ACL (iscsioptacl)	0

You must enter the `ipv6acl` and `vman-dual-qos` allocations as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd numbered ranges.

NOTE: You can only have one odd number of blocks in the CLI configuration; the other blocks must be in factors of 2. For example, a CLI configuration of 5+4+2+1+1 blocks is not supported; a configuration of 6+4+2+1 blocks is supported.

For the new settings to take effect, you must save the new CAM settings to the startup-config (`write-mem` or `copy run start`) then reload the system for the new settings to take effect.

CAM Allocation for Egress

To allocate the space for egress L2, IPV4, and IPV6 ACL, use the `cam-acl-egress` command. The total number of available FP blocks is 4. Allocate at least one group of L2ACL and IPV4 ACL.

```
DellEMC(conf)#do show cam-acl-egress

-- Chassis Egress Cam ACL --
              Current Settings(in block sizes)
                    1 block = 256 entries
L2Acl         :          1
Ipv4Acl       :          1
Ipv6Acl       :          2

-- Stack unit 1 --
           Current Settings(in block sizes)
L2Acl         :          1
Ipv4Acl       :          1
Ipv6Acl       :          2

DellEMC(conf)#
```

1. Select a cam-acl action.
   CONFIGURATION mode
   `cam-acl [default | l2acl]`
   NOTE: Selecting `default` resets the CAM entries to the default settings. Select `l2acl` to allocate the desired space for all other regions.
2. Enter the number of FP blocks for each region.
   EXEC Privilege mode

```
cam-acl {default | l2acl number ipv4acl number ipv6acl number ipv4qos number l2qos number
l2pt number ipmacacl number vman-qos | vman-dual-qos number ecfmacl number nlbcluster
number ipv4pbr number openflow number | fcoe number iscsioptacl number [vrfv4acl number]
```

NOTE: If you do not enter the allocation values for the CAM regions, the value is 0.

3. Execute write memory and verify that the new settings are written to the CAM on the next boot.
   EXEC Privilege mode
   `show cam-acl`
4. Reload the system.
   EXEC Privilege mode
   `reload`

Test CAM Usage

To determine whether sufficient CAM space is available to enable a service-policy, use the `test-cam-usage` command. To verify the actual CAM space required, create a Class Map with all required ACL rules, then execute the `test cam-usage` command in Privilege mode. The Status column in the command output indicates whether or not you can enable the policy.

Example of the `test cam-usage` Command

```
DellEMC#test cam-usage service-policy input test-cam-usage stack-unit 1 po 0

Stack-Unit| Portpipe|CAM Partition|Available CAM|Estimated CAM per Port|Status
-------------------------------------------------------------------------------
   2      |   0     |IPv4Flow     |192          |3                     |Allowed (64)
DellEMC#
```

View CAM Profiles

To view the current CAM profile for the chassis and each component, use the `show cam-profile` command.

This command also shows the profile that is loaded after the next chassis or component reload.

Example of the `show cam-profile` Command

```
Dell#show cam-profile

-- Chassis CAM Profile --

CamSize         : 18-Meg
                : Current Settings : Next Boot
Profile Name    : Default          : Default
L2FIB           : 32K entries      : 32K entries
L2ACL           : 1K entries       : 1K entries
IPv4FIB         : 256K entries     : 256K entries
IPv4ACL         : 12K entries      : 12K entries
IPv4Flow        : 24K entries      : 24K entries
EgL2ACL         : 1K entries       : 1K entries
EgIPv4ACL       : 1K entries       : 1K entries
Reserved        : 8K entries       : 8K entries
FIB           : 0 entries      : 0 entries
ACL           : 0 entries      : 0 entries
Flow          : 0 entries      : 0 entries
EgACL         : 0 entries      : 0 entries
MicroCode Name : Default          : Default
--More--
```

To view brief output of the `show cam-profile` command, use the `summary` option.

The `show running-config cam-profile` command shows the current profile and microcode.

NOTE: If you select the CAM profile from CONFIGURATION mode, the output of this command does not reflect any changes until you save the running-configuration and reload the chassis.

Example of `show running-config cam-profile` Command

```
Dell#show running-config cam-profile
!
cam-profile default microcode default

Dell#
```

View CAM-ACL Settings

The `show cam-acl` command shows the cam-acl setting that will be loaded after the next reload.

Example of Viewing CAM-ACL Settings

```
DellEMC(conf)#do show cam-acl

-- Chassis Cam ACL --
Current Settings(in block sizes)   Next Boot(in block sizes)
                  1 block = 128 entries
L2Acl         :         6                              4
Ipv4Acl       :         4                              2
Ipv6Acl       :         0                              0
Ipv4Qos       :         2                              2
L2Qos         :         1                              1
L2PT          :         0                              0
IpMacAcl      :         0                              0
VmanQos       :         0                              0
VmanDualQos   :         0                              0
EcfmAcl       :         0                              0
FcoeAcl       :         0                              0
iscsiOptAcl   :         0                              0
ipv4pbr       :         0                              2
vrfv4Acl      :         0                              2
Openflow      :         0                              0
fedgovacl     :         0                              0

-- Stack unit 0 --
 Current Settings(in block sizes)  Next Boot(in block sizes)
                  1 block = 128 entries
L2Acl         :         6                              4
Ipv4Acl       :         4                              2
Ipv6Acl       :         0                              0
Ipv4Qos       :         2                              2
L2Qos         :         1                              1
L2PT          :         0                              0
IpMacAcl      :         0                              0
VmanQos       :         0                              0
VmanDualQos   :         0                              0
EcfmAcl       :         0                              0
FcoeAcl       :         0                              0
iscsiOptAcl   :         0                              0
ipv4pbr       :         0                              2
vrfv4Acl      :         0                              2
Openflow      :         0                              0
fedgovacl     :         0                              0

DellEMC(conf)#
```

NOTE: If you change the cam-acl setting from CONFIGURATION mode, the output of this command does not reflect any changes until you save the running-configuration and reload the chassis.

The default values for the `show cam-acl` command are:

Example of Viewing CAM-ACL Settings

```
DellEMC#show cam-acl

-- Chassis Cam ACL --
             Current Settings(in block sizes)
                   1 block = 128 entries
L2Acl        :         6
Ipv4Acl      :         4
Ipv6Acl      :         0
Ipv4Qos      :         2
L2Qos        :         1
L2PT         :         0
IpMacAcl     :         0
VmanQos      :         0
VmanDualQos  :         0
EcfmAcl      :         0
FcoeAcl      :         0
iscsiOptAcl  :         0
ipv4pbr      :         0
vrfv4Acl     :         0
Openflow     :         0
fedgovacl    :         0

-- Stack unit 0 --
           Current Settings(in block sizes)
                   1 block = 128 entries
L2Acl        :         6
Ipv4Acl      :         4
Ipv6Acl      :         0
Ipv4Qos      :         2
L2Qos        :         1
L2PT         :         0
IpMacAcl     :         0
VmanQos      :         0
VmanDualQos  :         0
EcfmAcl      :         0
FcoeAcl      :         0
iscsiOptAcl  :         0
ipv4pbr      :         0
vrfv4Acl     :         0
Openflow     :         0
fedgovacl    :         0

-- Stack unit 7 --
           Current Settings(in block sizes)
                   1 block = 128 entries
L2Acl        :         6
Ipv4Acl      :         4
Ipv6Acl      :         0
Ipv4Qos      :         2
L2Qos        :         1
L2PT         :         0
IpMacAcl     :         0
VmanQos      :         0
VmanDualQos  :         0
EcfmAcl      :         0
FcoeAcl      :         0
iscsiOptAcl  :         0
ipv4pbr      :         0
vrfv4Acl     :         0
Openflow     :         0
fedgovacl    :         0

DellEMC#
```

View CAM Usage

View the amount of CAM space available, used, and remaining in each partition (including IPv4 and IPv6 Flow and Layer 2 ACL sub-partitions) using the `show cam-usage` command in EXEC Privilege mode

The following output shows CAM blocks usage for Layer 2 and Layer 3 ACLs and other processes that use CAM space:

Example of the `show cam-usage` Command

```
DellEMC#show cam-usage
Stackunit|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available CAM
========|========|=================|=============|=============|==============
    0      0       | IN-L3 ACL     |       512   |         1   |       511
        |          |  IN-V6 ACL    |         0   |         0   |         0
        |          | IN-L2 ACL     |       768   |         0   |       768
        |          | OUT-L3 ACL    |       158   |         5   |       153
        |          | IN-L3 ECMP GRP|      1024   |         0   |      1024
        |          | OUT-V6 ACL    |       158   |          0  |       158
        |          | OUT-L2 ACL    |       206   |          7  |       199
    7   |   0      | IN-L3 ACL     |       512   |          1  |       511
        |          | IN-V6 ACL     |         0   |           0 |         0
        |          | IN-L2 ACL     |        768  |           0 |       768
        |          | OUT-L3 ACL    |       158   |           5 |       153
        |          | OUT-V6 ACL    |       158   |           0 |       158
        |          | OUT-L2 ACL    |       206   |           7 |       199
Codes: * - cam usage is above 90%.
DellEMC#
```

Configuring CAM Threshold and Silence Period

This section describes how to configure CAM threshold and silence period between CAM threshold syslog warnings.

The CAM threshold and silence period configuration is applicable only for Ingress L2, IPv4, IPv6 and Egress L2, IPv4, and IPv6 ACL CAM groups. For other ACL CAM regions, the CAM threshold and silence period is fixed and the values are 90 percent and 0 respectively.

You can assign CAM threshold value using `cam-threshold` command to receive syslog messages when the CAM usage reaches the configured CAM threshold. The configured CAM threshold is a value specific to FP based on CAM features such as Ingress and Egress L2, IPV4, IPV6. The system checks the CAM usage of the features with the set threshold to display a syslog message, which contains the CAM region, slot/port-pipe and pipeline information. By default, syslog warning appears when the CAM usage is **90** percent.

You can also configure the silence period for the syslog message on the CAM usage. A syslog warning appears when the CAM usage exceeds the configured CAM threshold. The silence period starts after the initial syslog warning. The syslog warning does not appear until the silence period is active. By default, the silence period is **0** seconds.

Setting CAM Threshold and Silence Period

To configure the CAM threshold and silence period, use the following commands.

- Assign the CAM threshold percentage

  CONFIGURATION mode

  `cam-threshold threshold {default | threshold-percent}`

  The range of threshold value is from 1 to 100. The default threshold value is **90** percent.

- Assign the silence period for syslog warning.

  CONFIGURATION mode

  `cam-threshold threshold {default | threshold-percent} silence-period {default | silence-period-value}`

  The range of silence period is from 0 to 65535. The default is **0** seconds.

NOTE:

If you delete a FP in a CAM region that is assigned with threshold, a syslog warning appears even during the silence period.

The system triggers syslog during the following events:

- Re-configure the CAM threshold
- Add or delete an ACL rule

Example of Syslog message on CAM usage

Following table shows few possible scenarios during which the syslog message appear on re-configuring the CAM usage threshold value.

Consider if the last CAM threshold was set to 90 percent and now you re-configure the CAM threshold to 80. And, if the current CAM usage is 85 percent, then the system displays the syslog message saying that the CAM usage is above the configured CAM threshold value.

Table 13. Possible Scenarios of Syslog Warning

Old CAM Threshold	New CAM Threshold	Current CAM Usage	Syslog
90	80	85	`DellEMC(conf)#Nov 5 19:55:12 %S6000:0 %ACL_AGENT-4-ACL_AGENT_CAM_USAGE_OVER_THE_THRESHOLD: The Ipv4Acl cam region on stack-unit 0 Portpipe 0 Pipeline 0 is more than 80% Full.`
90	95	91	`DellEMC(conf)#Nov 5 19:55:12 %S6000:0 %ACL_AGENT-4-ACL_AGENT_CAM_USAGE_BELOW_THE_THRESHOLD: The cam-usage of Ipv4Acl cam region on stack-unit 0 Portpipe 0 Pipeline 0 is below 95%.`
98	100	100	No syslog
95	80	10	No syslog
92	90	89	No syslog

CAM Optimization

When you enable the CAM optimization, if a Policy Map containing classification rules (ACL and/or DSCP/ ip-precedence rules) is applied to more than one physical interface on the same port-pipe, only a single copy of the policy is written (only one FP entry is used). When you disable this command, the system behaves as described in this chapter.

Troubleshoot CAM Profiling

The following section describes CAM profiling troubleshooting.

QoS CAM Region Limitation

To store QoS service policies, the default CAM profile allocates a partition within the IPv4Flow region.

If the QoS CAM space is exceeded, a message similar to the following displays.

```
     %EX2YD:12 %DIFFSERV-2-DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3
Cam(PolicyQos) for class 2 (Te 1/20) entries on portpipe 1 for linecard 1
     %EX2YD:12 %DIFFSERV-2-
     DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for class 5 (Te 1/
22) entries on portpipe 1 for linecard 1
```

If you exceed the QoS CAM space, follow these steps.

1. Verify that you have configured a CAM profile that allocates 24 K entries to the IPv4 system flow region.

2. Allocate more entries in the IPv4Flow region to QoS.

Dell EMC Networking OS supports the ability to view the actual CAM usage before applying a service-policy. The `test cam-usage service-policy` command provides this test framework. For more information, refer to Pre-Calculating Available QoS CAM Space.

11

Control Plane Policing (CoPP)

Control plane policing (CoPP) uses access control list (ACL) rules and quality of service (QoS) policies to create filters for a system's control plane. That filter prevents traffic not specifically identified as legitimate from reaching the system control plane, rate-limits, traffic to an acceptable level.

CoPP increases security on the system by protecting the routing processor from unnecessary or DoS traffic, giving priority to important control plane and management traffic. CoPP uses a dedicated control plane configuration through the ACL and QoS command line interfaces (CLIs) to provide filtering and rate-limiting capabilities for the control plane packets.

The following illustration shows an example of the difference between having CoPP implemented and not having CoPP implemented.

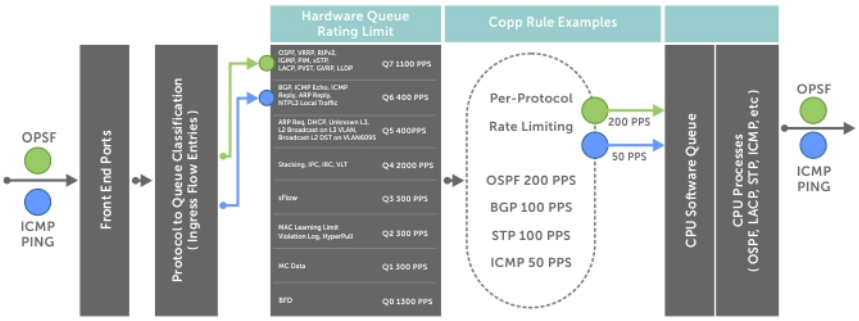


Figure 27. Control Plane Policing

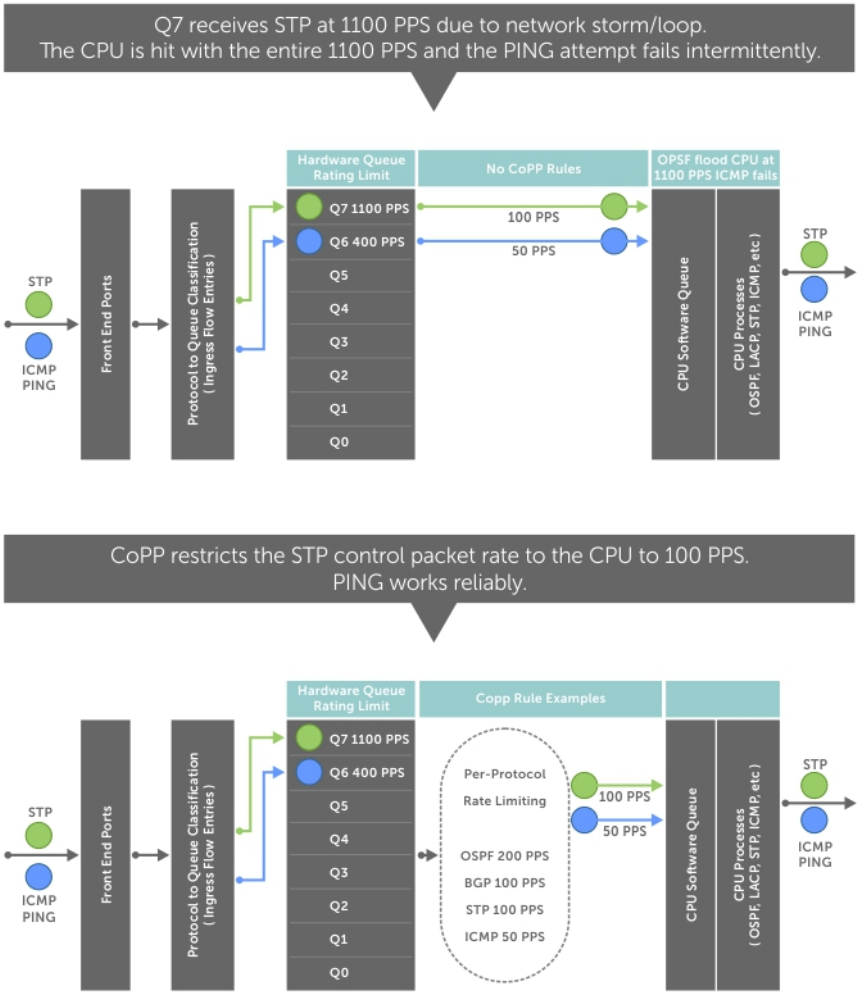


Figure 28. CoPP Implemented Versus CoPP Not Implemented

Topics:

- Configure Control Plane Policing

Configure Control Plane Policing

The system can process a maximum of 4200 packets per second (PPS). Protocols that share a single queue may experience flaps if one of the protocols receives a high rate of control traffic even though per protocol CoPP is applied. This happens because queue-based rate limiting is applied first.

For example, border gateway protocol (BGP) and internet control message protocol (ICMP) share same queue (Q6); Q6 has 400 PPS of bandwidth by default. The desired rate of ICMP is 100 PPS and the remaining 300 PPS is assigned to BGP. If ICMP packets come at 400 PPS, BGP packets may be dropped though ICMP packets are rate-limited to 100 PPS. You can solve this by increasing Q6 bandwidth to 700 PPS to allow both ICMP and BGP packets and then applying per-flow CoPP for ICMP and BGP packets. The setting of this Q6 bandwidth is dependent on the incoming traffic for the set of protocols sharing the same queue. If you are not aware of the incoming protocol traffic rate, you cannot set the required queue rate limit value. You must complete queue bandwidth tuning carefully because the system cannot open up to handle any rate, including traffic coming at the line rate.

CoPP policies are assigned on a per-protocol or a per-queue basis, and are assigned in CONTROL-PLANE mode to each port-pipe.

CoPP policies are configured by creating extended ACL rules and specifying rate-limits through QoS policies. The ACLs and QoS policies are assigned as service-policies.

Configuring CoPP for Protocols

This section lists the commands necessary to create and enable the service-policies for CoPP.

For complete information about creating ACLs and QoS rules, refer to Access Control Lists (ACLs) and Quality of Service (QoS).

The basics for creating a CoPP service policy are to create a Layer 2, Layer 3, and/or an IPv6 ACL rule for the desired protocol type. Then, create a QoS input policy to rate-limit the protocol traffics according to the ACL. The ACL and QoS policies are finally assigned to a control-plane service policy for each port-pipe.

1. Create a Layer 2 extended ACL for control-plane traffic policing for a particular protocol.
   CONFIGURATION mode
   `mac access-list extended name cpu-qos`
   `permit {arp | frrp | gvrp | isis | lacp | lldp | stp}`
2. Create a Layer 3 extended ACL for control-plane traffic policing for a particular protocol.
   CONFIGURATION mode
   `ip access-list extended name cpu-qos`
   `permit {bgp | dhcp | dhcp-relay | ftp | icmp | igmp | mcast-catch-all | msdp | ntp | ospf | pim | ip | ssh | telnet | vrrp}`
3. Create an IPv6 ACL for control-plane traffic policing for a particular protocol.
   CONFIGURATION mode
   `ipv6 access-list name cpu-qos`
   `permit {bgp | icmp | vrrp}`
4. Create a QoS input policy for the router and assign the policing.
   CONFIGURATION mode
   `qos-policy-input name`
   `cpu-qos rate-police rate-police-value`
5. Create a QoS class map to differentiate the control-plane traffic and assign to an ACL.
   CONFIGURATION mode
   `class-map match-any name`
   `cpu-qos match {ip | mac | ipv6} access-group name`
6. Create a QoS input policy map to match to the class-map and qos-policy for each desired protocol.
   CONFIGURATION mode
   `policy-map-input name`
   `cpu-qos class-map name qos-policy name`
7. Enter Control Plane mode.
   CONFIGURATION mode
   `control-plane-cpuqos`
8. Assign the protocol based the service policy on the control plane. Enabling this command on a port-pipe automatically enables the ACL and QoS rules creates with the `cpu-qos` keyword.
   CONTROL-PLANE mode
   `service-policy rate-limit-protocols`

The following example shows creating the IP/IPv6/MAC extended ACL.

```
DellEMC(conf)#ip access-list extended mcast cpu-qos
DellEMC(conf-ip-acl-cpuqos)#permit mcast-catch-all
DellEMC(conf-ip-acl-cpuqos)#exit

DellEMC(conf)#ip access-list extended ospf cpu-qos
DellEMC(conf-ip-acl-cpuqos)#permit ospf
DellEMC(conf-ip-acl-cpuqos)#exit

DellEMC(conf)#ip access-list extended bgp cpu-qos
DellEMC(conf-ip-acl-cpuqos)#permit bgp
DellEMC(conf-ip-acl-cpuqos)#exit
```

```
DellEMC(conf)#mac access-list extended lacp cpu-qos
DellEMC(conf-mac-acl-cpuqos)#permit lacp
DellEMC(conf-mac-acl-cpuqos)#exit

DellEMC(conf)#ipv6 access-list ipv6-icmp cpu-qos
DellEMC(conf-ipv6-acl-cpuqos)#permit icmp
DellEMC(conf-ipv6-acl-cpuqos)#exit

DellEMC(conf)#ipv6 access-list ipv6-vrrp cpu-qos
DellEMC(conf-ipv6-acl-cpuqos)#permit vrrp
DellEMC(conf-ipv6-acl-cpuqos)#exit
```

The following example shows creating the QoS input policy.

```
DellEMC(conf)#qos-policy-in rate_limit_200k cpu-qos
DellEMC(conf-in-qos-policy-cpuqos)#rate-police 200 40 peak 500 40
DellEMC(conf-in-qos-policy-cpuqos)#exit

DellEMC(conf)#qos-policy-in rate_limit_400k cpu-qos
DellEMC(conf-in-qos-policy-cpuqos)#rate-police 400 50 peak 600 50
DellEMC(conf-in-qos-policy-cpuqos)#exit

DellEMC(conf)#qos-policy-in rate_limit_500k cpu-qos
DellEMC(conf-in-qos-policy-cpuqos)#rate-police 500 50 peak 1000 50
DellEMC(conf-in-qos-policy-cpuqos)#exit
```

The following example shows creating the QoS class map.

```
DellEMC(conf)#class-map match-any class_ospf cpu-qos
DellEMC(conf-class-map-cpuqos)#match ip access-group ospf
DellEMC(conf-class-map-cpuqos)#exit

DellEMC(conf)#class-map match-any class_bgp cpu-qos
DellEMC(conf-class-map-cpuqos)#match ip access-group bgp
DellEMC(conf-class-map-cpuqos)#exit

DellEMC(conf)#class-map match-any class_lacp cpu-qos
DellEMC(conf-class-map-cpuqos)#match mac access-group lacp
DellEMC(conf-class-map-cpuqos)#exit

DellEMC(conf)#class-map match-any class-ipv6-icmp cpu-qos
DellEMC(conf-class-map-cpuqos)#match ipv6 access-group ipv6-icmp
DellEMC(conf-class-map-cpuqos)#exit
```

The following example shows matching the QoS class map to the QoS policy.

```
DellEMC(conf)#policy-map-input egressFP_rate_policy cpu-qos
DellEMC(conf-policy-map-in-cpuqos)#class-map class_ospf qos-policy rate_limit_500k
DellEMC(conf-policy-map-in-cpuqos)#class-map class_bgp qos-policy rate_limit_400k
DellEMC(conf-policy-map-in-cpuqos)#class-map class_lacp qos-policy rate_limit_200k
DellEMC(conf-policy-map-in-cpuqos)#class-map class-ipv6 qos-policy rate_limit_200k
DellEMC(conf-policy-map-in-cpuqos)#exit
```

The following example shows creating the control plane service policy.

```
DellEMC(conf)#control-plane-cpuqos
DellEMC(conf-control-cpuqos)#service-policy rate-limit-protocols egressFP_rate_policy
DellEMC(conf-control-cpuqos)#exit
```

Configuring CoPP for CPU Queues

Controlling traffic on the CPU queues does not require ACL rules, but does require QoS policies.

CoPP for CPU queues converts the input rate from kbps to pps, assuming 64 bytes is the average packet size, and applies that rate to the corresponding queue. Consequently, 1 kbps is roughly equivalent to 2 pps.

The basics for creating a CoPP service policy is to create QoS policies for the desired CPU bound queue and associate it with a particular rate-limit. The QoS policies are assigned to a control-plane service policy for each port-pipe.

1. Create a QoS input policy for the router and assign the policing.
   CONFIGURATION mode
   `qos-policy-input name cpu-qos`
2. Create an input policy-map to assign the QoS policy to the desired service queues.l.
   CONFIGURATION mode
   `policy-map--input name`
   `cpu-qos service-queue queue-number qos-policy name`
3. Enter Control Plane mode.
   CONFIGURATION mode
   `control-plane-cpuqos`
4. Assign a CPU queue-based service policy on the control plane in cpu-qos mode. Enabling this command sets the queue rates according to those configured.
   CONTROL-PLANE mode
   `service-policy rate-limit-cpu-queues input-policy-map`

The following example shows creating the QoS policy.

```
DellEMC#conf
DellEMC(conf)#qos-policy-input cpuq_1
DellEMC(conf-qos-policy-in)#rate-police 3000 40 peak 500 40
DellEMC(conf-qos-policy-in)#exit

DellEMC(conf)#qos-policy-input cpuq_2
DellEMC(conf-qos-policy-in)#rate-police 5000 80 peak 600 50
DellEMC(conf-qos-policy-in)#exit
```

The following example shows assigning the QoS policy to the queues.

```
DellEMC(conf)#policy-map-input cpuq_rate_policy cpu-qos
DellEMC(conf-qos-policy-in)#service-queue 5 qos-policy cpuq_1
DellEMC(conf-qos-policy-in)#service-queue 6 qos-policy cpuq_2
DellEMC(conf-qos-policy-in)#service-queue 7 qos-policy cpuq_1
```

The following example shows creating the control plane service policy.

```
DellEMC#conf
DellEMC(conf)#control-plane
DellEMC(conf-control-plane)#service-policy rate-limit-cpu-queues cpuq_rate_policy
```

CoPP for OSPFv3 Packets

You can create an IPv6 ACL for control-plane traffic policing for OSPFv3, in addition to the CoPP support for VRRP, BGP, and ICMP. You can use the `ipv6 access-list name cpu-qos permit ospfv3` command to allow CoPP traffic for OSPFv3. Control Plane Policing (CoPP) enables more number of CPU queues to be made available on ports for IPv6 and ICMPv6 packets.

CoPP enhancements are to enhance the capability of FTOS by utilizing more number of CPU queues on CMIC port and sending control packets to different queues that internally reduce limitation or contention of control protocols sharing the same queues (that is, before this functionality of CoPP for OSPV3 was introduced, OSPF might have caused the LACP flap because of both control traffic sent to same Q7 on CPU port). Non CPU port should have only 4 dedicated control queues and remaining shared for both data and traffic. Number of control queues is increased on the CPU port. When tunneling packets from non-master to master unit, high-gig queues are used.

Prior to the release 9.4.(0.0), all IPv6 packets are taken to same queues there is no priority between the ICMPv6 packets and unknown IPv6 packets. Due to this NS/NA/RS/RA packets not given high priority leads to the session establishment problem. To solve this issue, starting from release 9.4.(0.0), IPv6 NDP packets use different CPU queues when compared to the Generic IPv6 multicast traffic. These entries are installed in system when application is triggered..

CPU Processing of CoPP Traffic

The systems use FP rules to take the packets to control plane by CopyToCPU or redirect packet to CPU port. Only 8 CPU queues are used while sending the packet to CPU. The CPU Management Interface Controller (CMIC) interface on all the systems supports 48 queues in hardware. However, FTOS supports only 8 CMIC queues – 4 for data streams that are CPU bound – SFLOW packets, packet streams that are trapped to CPU for logging info on MAC learn limit exceeded and other violations, L3 packets with unknown destination for soft forwarding etc. Other 4 CMIC queues will carry the L2/L3 well-known protocol streams. However there are about 20 well known protocol streams that have to share these 4 CMIC queues. Before 9.4.(0.0)Dell EMC Networking OS used only 8 queues most of the queues are shared to multiple protocols. So, increasing the number of CMIC queues will reduce the contention among the protocols for the queue bandwidth.

Currently, there are 4 Queues for data and 4 for control in both front-end and back-plane ports. In stacked systems, the control streams that reach standby or slave units will be tunneled through the backplane ports across stack-units to reach the CPU of the master unit. In this case, the packets that reach slave unit's CMIC via queues 0 – 7 will take same queues 0 – 7 on the back-plane ports while traversing across units and finally on the master CMIC, they are queued on the same queues 0 – 7. In this case, the queue (4 – 7) taken by the well-known protocol streams are uniform across different queuing points, and the queue (0 – 3) taken by the CPU bound data streams are uniform. In back-plane ports, queue 0 – 3 will carry both the front-end bound data streams as well as the CPU bound data streams which is acceptable but the well-known protocol streams must not be mixed with the data streams on queues 0 – 3 in back-plane ports.

Increased CPU Queues for CoPP

FTOS classifies every packet ingress from the front end port to system as control traffic or data traffic by having the pre-defined rules based on protocol type or packets types like ttl, slow path etc. FP is used to classify the traffic to transmit the control traffic to CMIC port. Other major function performed by the FP rule is to decide to which CPU queue the packet must be sent. All other packets will be forwarded or dropped at the ingress.

All packet transmitted to CPU will transmit to local CPU by using the CPU queues and processed. But in stacked system only mater CPU is responsible for the control plane actions. So control packets received in master or slave units will be tunneled to master CPU to process.

As part of enhancements, CPU queues are increased from 8 to 12 on CPU port. However, the front-end port and the backplane ports support only 8 queues. As a result, when packets are transmitted to the local CPU, the CPU uses Q0-Q11 queues. The control packets that are tunneled to the master unit are isolated from the data queues and the control queues in the backplane links. Control traffic must be sent over the control queues Q4-Q7 on higig links. After reaching the master unit tunneled packets must be transmitted to the CPU using the Q0-Q11 queues.

The backplane ports can have a maximum of 4 control queues. So, when we have more than 'n' CMIC queues for well-known protocols and n > 4, then streams on 'n' CMIC queues must be multiplexed on 4 control queues on back-plane ports and on the Master unit, these streams must be de-multiplexed to 'n' CMIC queues on the Master CPU.

After control packets reach the CPU through the CMIC port, the software schedules to process traffic on each 12 CPU queues. This aspect must be ensured even in case of stand-alone systems and there is no dependency with stacking.

Policing provides a method for protecting CPU bound control plane packets by policing packets transmited to CPU with a specified rate and from undesired or malicious traffic. This is done at each CPU queue on each unit.

FP Entries for Distribution of NDP Packets to Various CPU Queues

- At present generic mac based entries in system flow region will take IPv6 packets to CPU.
  - OSPFv3 – 33:33:0:0:0:5 – Q7
  - - 33:33:0:0:0:6 – Q7
  - IPv6 Multicast – 33:33:0:0:0:0 – Q1
- Add/remove specific ICMPv6 NDP protocol entry when user configures the first ipv6 address in the front panel port
  - Distribute ICMPv6 NS/RS packets to Q5.
  - Distribute ICMPv6 NA/RA packets to Q6.

FP is installed for all Front panel ports.

NDP Packets

Neighbor discovery protocol has 4 types of packets NS, NA, RA, RS. These packets need to be taken to CPU for neighbor discovery.

- Unicast NDP packets:
  - Packets hitting the L3 host/route table and discovered as local terminated packets/CPU bound traffic. For CPU bound traffic route entry have CPU action. Below are packets are CPU bound traffic.
    - Packets destined to chassis.
    - Route with Unresolved Arp
    - Unknown traffic in IP Subnet range
    - Unknown traffic hitting the default route entry.
- Multicast NDP packets
  - NDP packets with destination MAC is multicast
    - DST MAC 33:33:XX:XX:XX:XX
- NDP Packets in VLT peer routing enable
  - VLT peer routing enable cases each VLT node will have route entry for link local address of both self and peer VLT node. Peer VLT link local entry will have egress port as ICL link. And Actual link local address will have entry to CopyToCpu. But NDP packets destined to peer VLT node needs to be taken to CPU and tunneled to the peer VLT node..
- NDP packets in VLT peer routing disable case
  - NDP packets intended to peer VLT chassis taken to CPU and tunnel to peer.

The following table describes the protocol to queue mapping with the CPU queues increased to be 12.

Table 14. Redirecting Control Traffic to 12 CPU queues

CPU Queue	Weights	Rate (pps)	Protocol
0	100	1300	BFD
1	1	300	MC
2	2	300	TTL0, TTL1, IP with options, Mac limit violation, Hyper pull, L3 with Bcast MacDA, Unknown L3, ARP unresolved, ACL Logging
3	4	400	sFlow, L3 MTU Fail frames
4	127	2000	IPC/IRC, VLT Control frames
5	16	300	ARP Request, NS, RS, iSCSI OPT Snooping
6	16	400	ICMP, ARP Reply, NTP, Local terminated L3, NA, RA,ICMPv6 (other Than NDP and MLD)
7	64	400	xSTP, FRRP, LACP, 802.1x,ECFM,L2PT,TRILL, Open flow
8	32	400	PVST, LLDP, GVRP, FCOE, FEFD, Trace flow
9	64	600	OSPF, ISIS, RIPv2, BGP
10	32	300	DHCP, VRRP
11	32	300	PIM, IGMP, MSDP, MLD

Catch-All Entry for IPv6 Packets

Dell EMC Networking OS currently supports configuration of IPv6 subnets greater than /64 mask length, but the agent writes it to the default LPM table where the key length is 64 bits. The device supports table to store up to 256 subnets of maximum of /128 mask lengths. This can be enabled and agent can be modified to update the /128 table for mask lengths greater than /64. This will restrict the subnet sizes to required optimal level which would avoid these NDP attacks. The IPv6 stack already supports handling of >/64 subnets and doesn't require any additional work. The default catch-all entry is put in the LPM table for IPv4 and IPv6. If this is included for IPv6, you can disable this capability by using the `no ipv6 unknown-unicast` command. Typically, the catch-all entry in LPM table is used for soft forwarding and generating ICMP unreachable messages to the source. If this is in place then irrespective of whether it is </64 subnet or >/64 subnet, it doesn't have any effect as there would always be LPM hit and traffic are sent to CPU.

Unknown unicast L3 packets are terminated to the CPU CoS queue which is also shared for other types of control-plane packets like ARP Request, Multicast traffic, L3 packets with Broadcast MAC address. The catch-all route poses a risk of overloading the CPU with unknown unicast packets. This CLI knob to turn off the catch-all route is of use in networks where the user does not want to generate Destination Unreachable messages and have the CPU queue's bandwidth available for higher priority control-plane traffic.

Configuring CoPP for OSPFv3

You can create an IPv6 ACL for control-plane traffic policing for OSPFv3, in addition to the CoPP support for VRRPv3, BGPv6, and ICMPv6. You can use the `ipv6 access-list` *name* `cpu-qos permit ospfv3` or the `ipv6 access-list` *name* `cpu-qos ospfv3` command to allow CoPP traffic for OSPFv3. The control plane management support for IPv6 ICMPv6 packets is enhanced to enable more number of CPU queues on port to be available and other COPP improvements have been implemented.

To configure control-plane policing, perform the following:

1. Create an IPv6 ACL for control-plane traffic policing for ospfv3.
   CONFIGURATION mode
   ```
   Dell(conf)#ipv6 access-list ospfv3 cpu-qos
   Dell(conf-ipv6-acl-cpuqos)#permit ospf
   ```
2. Create a QoS input policy for the router and assign the policing.
   CONFIGURATION mode
   ```
   Dell(conf)#qos-policy-input ospfv3_rate cpu-qos
   Dell(conf-in-qos-policy-cpuqos)#rate-police 1500 16 peak 1500 16
   ```
3. Create a QoS class map to differentiate the control-plane traffic and assign to the ACL.
   CONFIGURATION mode
   ```
   Dell(conf)#class-map match-any ospfv3 cpu-qos
   Dell(conf-class-map-cpuqos)#match ipv6 access-group ospfv3
   ```
4. Create a QoS input policy map to match to the class-map and qos-policy for each desired protocol.
   CONFIGURATION mode
   ```
   Dell(conf)#policy-map-input ospfv3_policy cpu-qos
   Dell(conf-policy-map-in-cpuqos)#class-map ospfv3 qos-policy ospfv3_rate
   ```
5. Enter Control Plane mode.
   CONFIGURATION mode
   ```
   Dell(conf)#control-plane-cpuqos
   ```
6. Assign the protocol based service policy on the control plane. Enabling this command on a port-pipe automatically enables the ACL and QoS rules created with the cpu-qos keyword.
   CONTROL-PLANE mode
   ```
   Dell(conf-control-cpuqos)#service-policy rate-limit-protocols ospfv3_policy
   ```

Displaying CoPP Configuration

The CLI provides show commands to display the protocol traffic assigned to each control-plane queue and the current rate-limit applied to each queue. Other show commands display statistical information for trouble shooting CoPP operation.

To view the rates for each queue, use the `show cpu-queue rate cp` command.

Viewing Queue Rates

Example of Viewing Queue Rates

```
DellEMC#show cpu-queue rate cp
Service-Queue  Rate (PPS)
-------------- -----------
Q0             1300
Q1             300
Q2             300
Q3             300
Q4             2000
Q5             400
Q6             400
Q7             1100
DellEMC#
```

To view the queue mapping for each configured protocol, use the `show ip protocol-queue-mapping` command.

Example of Viewing Queue Mapping

```
DellEMC#show ip protocol-queue-mapping
Protocol     Src-Port Dst-Port TcpFlag Queue EgPort Rate (kbps)
--------     -------- -------- ------- ----- ------ -----------
TCP (BGP)     any/179 179/any    _       Q6    CP      100
UDP (DHCP)    67/68   68/67      _       Q6/Q5 CP      _
UDP (DHCP-R) 67       67         _       Q6    CP      _
TCP (FTP)     any     21         _       Q6    CP      _
ICMP          any     any        _       Q6    CP      _
IGMP          any     any        _       Q7    CP      _
TCP (MSDP)    any/639 639/any    _       Q6    CP      _
UDP (NTP)     any     123        _       Q6    CP      _
OSPF          any     any        _       Q7    CP      _
PIM           any     any        _       Q7    CP      _
UDP (RIP)     any     520        _       Q7    CP      _
TCP (SSH)     any     22         _       Q6    CP      _
TCP (TELNET) any      23         _       Q6    CP      _
VRRP          any     any        _       Q7    CP      _
DellEMC#
```

To view the queue mapping for the MAC protocols, use the `show mac protocol-queue-mapping` command.

Example of Viewing Queue Mapping for MAC Protocols

```
DellEMC#show mac protocol-queue-mapping
Protocol Destination Mac      EtherType Queue EgPort Rate (kbps)
-------- ----------------   ----------- ----- ------ -----------
ARP      any                0x0806    Q5/Q6   CP     _
FRRP     01:01:e8:00:00:10/11 any       Q7      CP     _
LACP     01:80:c2:00:00:02    0x8809    Q7      CP     _
LLDP     any                  0x88cc    Q7      CP     _
GVRP     01:80:c2:00:00:21    any       Q7      CP     _
STP      01:80:c2:00:00:00    any       Q7      CP     _
ISIS     01:80:c2:00:00:14/15 any       Q7      CP     _
         09:00:2b:00:00:04/05 any       Q7      CP

DellEMC#
```

To view the queue mapping for IPv6 protocols, use the `show ipv6 protocol-queue-mapping` command.

Example of Viewing Queue Mapping for IPv6 Protocols

```
DellEMC#show ipv6 protocol-queue-mapping
Protocol   Src-Port Dst-Port TcpFlag Queue EgPort Rate (kbps)
--------   -------- -------- ------- ----- ------ -----------
TCP (BGP)  any/179  179/any   _       Q6     CP     _
ICMP       any      any       _       Q6     CP     _
VRRP       any      any       _       Q7     CP     _

DellEMC#
```

12

Data Center Bridging (DCB)

Data center bridging (DCB) refers to a set of enhancements to Ethernet local area networks used in data center environments, particularly with clustering and storage area networks.

NOTE:

DCB is not supported when you use 10GBaseT ports for stacking.

Topics:

- Ethernet Enhancements in Data Center Bridging
- Enabling Data Center Bridging
- Data Center Bridging: Default Configuration
- Configuring Priority-Based Flow Control
- Configuring PFC in a DCB Map
- Applying a DCB Map on a Port
- Configuring PFC without a DCB Map
- Priority-Based Flow Control Using Dynamic Buffer Method
- Behavior of Tagged Packets
- Configuration Example for DSCP and PFC Priorities
- Using PFC to Manage Converged Ethernet Traffic
- Configure Enhanced Transmission Selection
- Hierarchical Scheduling in ETS Output Policies
- Using ETS to Manage Converged Ethernet Traffic
- Applying DCB Policies in a Switch Stack
- Configure a DCBx Operation
- Verifying the DCB Configuration
- Sample DCB Configuration
- QoS dot1p Traffic Classification and Queue Assignment
- Configuring the Dynamic Buffer Method

Ethernet Enhancements in Data Center Bridging

The following section describes DCB.

The S4820T system supports loading two *DCB_Config* files: *FCoE_DCB_Config* and *iSCSI_DCB_Config*.

These files are located in the root directory flash:/CONFIG_TEMPLATE. After copying the configuration files to the startup config and reloading the system.

The device supports the following DCB features:

- Data center bridging exchange protocol (DCBx)
- Priority-based flow control (PFC)
- Enhanced transmission selection (ETS)

To configure PFC, ETS, and DCBx for DCB, refer to Sample DCB Configuration for the CLI configurations.

DCB refers to a set of IEEE Ethernet enhancements that provide data centers with a single, robust, converged network to support multiple traffic types, including local area network (LAN), server, and storage traffic. Through network consolidation, DCB results in reduced operational cost, simplified management, and easy scalability by avoiding the need to deploy separate application-specific networks.

For example, instead of deploying an Ethernet network for LAN traffic, include additional storage area networks (SANs) to ensure lossless Fibre Channel traffic, and a separate InfiniBand network for high-performance inter-processor computing within server clusters, only one DCB-enabled network is required in a data center. The Dell EMC Networking switches that support

a unified fabric and consolidate multiple network infrastructures use a single input/output (I/O) device called a converged network adapter (CNA).

A CNA is a computer input/output device that combines the functionality of a host bus adapter (HBA) with a network interface controller (NIC). Multiple adapters on different devices for several traffic types are no longer required.

Data center bridging satisfies the needs of the following types of data center traffic in a unified fabric:

Traffic	Description
LAN traffic	LAN traffic consists of many flows that are insensitive to latency requirements, while certain applications, such as streaming video, are more sensitive to latency. Ethernet functions as a best-effort network that may drop packets in the case of network congestion. IP networks rely on transport protocols (for example, TCP) for reliable data transmission with the associated cost of greater processing overhead and performance impact LAN traffic consists of a large number of flows that are generally insensitive to latency requirements, while certain applications, such as streaming video, are more sensitive to latency. Ethernet functions as a best-effort network that may drop packets in case of network congestion. IP networks rely on transport protocols (for example, TCP) for reliable data transmission with the associated cost of greater processing overhead and performance impact.
Storage traffic	Storage traffic based on Fibre Channel media uses the Small Computer System Interface (SCSI) protocol for data transfer. This traffic typically consists of large data packets with a payload of 2K bytes that cannot recover from frame loss. To successfully transport storage traffic, data center Ethernet must provide no-drop service with lossless links.
InterProcess Communication (IPC) traffic	InterProcess Communication (IPC) traffic within high-performance computing clusters to share information. Server traffic is extremely sensitive to latency requirements.

To ensure lossless delivery and latency-sensitive scheduling of storage and service traffic and I/O convergence of LAN, storage, and server traffic over a unified fabric, IEEE data center bridging adds the following extensions to a classical Ethernet network:

- 802.1Qbb — Priority-based Flow Control (PFC)
- 802.1Qaz — Enhanced Transmission Selection (ETS)
- 802.1Qau — Congestion Notification
- Data Center Bridging Exchange (DCBx) protocol

NOTE: Dell EMC Networking OS supports only the PFC, ETS, and DCBx features in data center bridging.

Priority-Based Flow Control

In a data center network, priority-based flow control (PFC) manages large bursts of one traffic type in multiprotocol links so that it does not affect other traffic types and no frames are lost due to congestion.

When PFC detects congestion on a queue for a specified priority, it sends a pause frame for the 802.1p priority traffic to the transmitting device. In this way, PFC ensures that PFC-enabled priority traffic is not dropped by the switch.

PFC enhances the existing 802.3x pause and 802.1p priority capabilities to enable flow control based on 802.1p priorities (classes of service). Instead of stopping all traffic on a link (as performed by the traditional Ethernet pause mechanism), PFC pauses traffic on a link according to the 802.1p priority set on a traffic type. You can create lossless flows for storage and server traffic while allowing for loss in case of LAN traffic congestion on the same physical interface.

The following illustration shows how PFC handles traffic congestion by pausing the transmission of incoming traffic with dot1p priority 4.

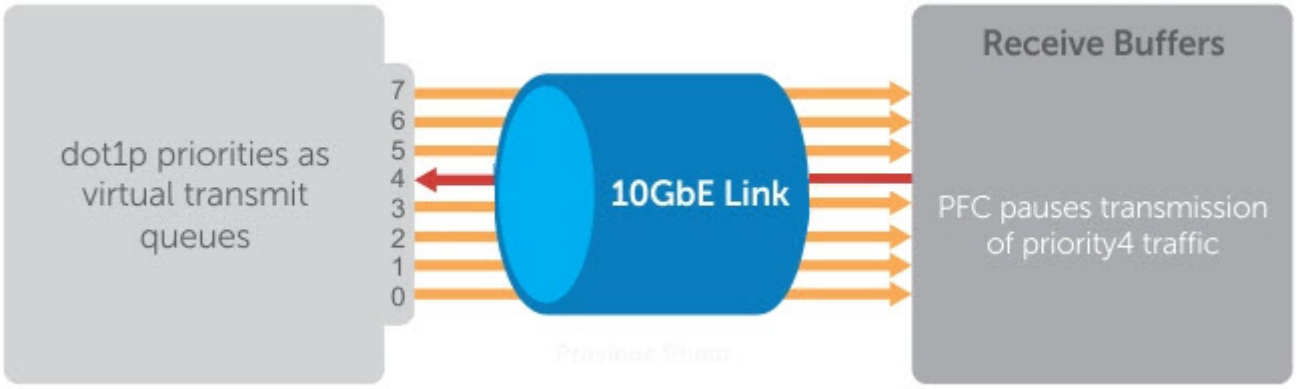


Figure 29. Illustration of Traffic Congestion

The system supports loading two DCB_Config files:

- FCoE converged traffic with priority 3.
- iSCSI storage traffic with priority 4.

In the Dell EMC Networking OS, PFC is implemented as follows:

- PFC is supported on specified 802.1p priority traffic (dot1p 0 to 7) and is configured per interface. However, only two lossless queues are supported on an interface: one for Fibre Channel over Ethernet (FCoE) converged traffic and one for Internet Small Computer System Interface (iSCSI) storage traffic. Configure the same lossless queues on all ports.
- PFC supports buffering to receive data that continues to arrive on an interface while the remote system reacts to the PFC operation.
- PFC uses DCB MIB IEEE 802.1azd2.5 and PFC MIB IEEE 802.1bb-d2.2.
- A dynamic threshold handles intermittent traffic bursts and varies based on the number of PFC priorities contending for buffers, while a static threshold places an upper limit on the transmit time of a queue after receiving a message to pause a specified priority. PFC traffic is paused only after surpassing both static and dynamic thresholds for the priority specified for the port.
- By default, PFC is enabled when you enable DCB. If you have not loaded FCoE_DCB_Config and iSCSI_DCB_Config, DCB is disabled. When you enable DCB globally, you cannot simultaneously enable link-level flow control.
- Buffer space is allocated and de-allocated only when you configure a PFC priority on the port.

Enhanced Transmission Selection

Enhanced transmission selection (ETS) supports optimized bandwidth allocation between traffic types in multiprotocol (Ethernet, FCoE, SCSI) links.

ETS allows you to divide traffic according to its 802.1p priority into different priority groups (traffic classes) and configure bandwidth allocation and queue scheduling for each group to ensure that each traffic type is correctly prioritized and receives its required bandwidth. For example, you can prioritize low-latency storage or server cluster traffic in a traffic class to receive more bandwidth and restrict best-effort LAN traffic assigned to a different traffic class.

NOTE: Use the following command to enable etsacl: `cam-acl l2acl 2 ipv4acl 2 ipv6acl 0 ipv4qos 0 l2qos 0 l2pt 0 ipmacacl 0 vman-qos 0 fcoeacl 2 etsacl 3`. After executing this command, you must save the configuration and then reload the system.

The following figure shows how ETS allows you to allocate bandwidth when different traffic types are classed according to 802.1p priority and mapped to priority groups.

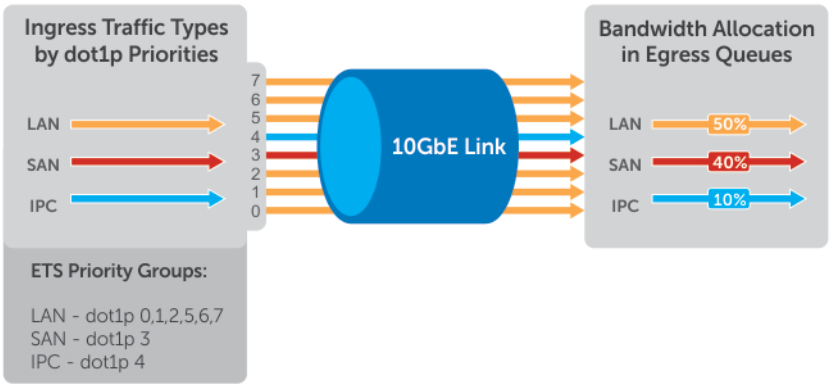


Figure 30. Enhanced Transmission Selection

The following table lists the traffic groupings ETS uses to select multiprotocol traffic for transmission.

Table 15. ETS Traffic Groupings

Traffic Groupings	Description
Group ID	A 4-bit identifier assigned to each priority group. The range is from 0 to 7 configurable; 8 - 14 reservation and 15.0 - 15.7 is strict priority group..
Group bandwidth	Percentage of available bandwidth allocated to a priority group.
Group transmission selection algorithm (TSA)	Type of queue scheduling a priority group uses.

In Dell EMC Networking OS, ETS is implemented as follows:

- ETS supports groups of 802.1p priorities that have:
  - PFC enabled or disabled
  - No bandwidth limit or no ETS processing
- ETS uses the DCB MIB IEEE 802.1azd2.5.

Data Center Bridging Exchange Protocol (DCBx)

DCBx allows a switch to automatically discover DCB-enabled peers and exchange configuration information. PFC and ETS use DCBx to exchange and negotiate parameters with peer devices. DCBx capabilities include:

- Discovery of DCB capabilities on peer-device connections.
- Determination of possible mismatch in DCB configuration on a peer link.
- Configuration of a peer device over a DCB link.

DCBx requires the link layer discovery protocol (LLDP) to provide the path to exchange DCB parameters with peer devices. Exchanged parameters are sent in organizationally specific TLVs in LLDP data units. The following LLDP TLVs are supported for DCB parameter exchange:

PFC parameters PFC Configuration TLV and Application Priority Configuration TLV.

ETS parameters ETS Configuration TLV and ETS Recommendation TLV.

Data Center Bridging in a Traffic Flow

The following figure shows how DCB handles a traffic flow on an interface.

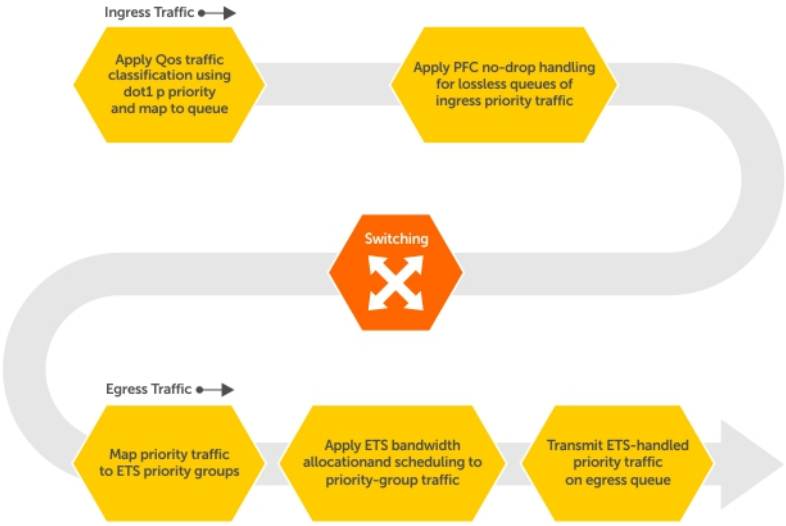


Figure 31. DCB PFC and ETS Traffic Handling

Enabling Data Center Bridging

DCB is automatically configured when you configure FCoE or iSCSI optimization.

Data center bridging supports converged enhanced Ethernet (CEE) in a data center network. DCB is disabled by default. It must be enabled to support CEE.

- Priority-based flow control
- Enhanced transmission selection
- Data center bridging exchange protocol
- FCoE initialization protocol (FIP) snooping

DCB processes virtual local area network (VLAN)-tagged packets and dot1p priority values. Untagged packets are treated with a dot1p priority of 0.

For DCB to operate effectively, you can classify ingress traffic according to its dot1p priority so that it maps to different data queues. The dot1p-queue assignments used are shown in the following table.

To enable DCB, enable either the iSCSI optimization configuration or the FCoE configuration.

To enable DCB with PFC buffers on a switch, enter the following commands, save the configuration, and reboot the system to allow the changes to take effect.

1. Enable DCB.

   CONFIGURATION mode

   ```
   dcb enable
   ```

2. Set PFC buffering on the DCB stack unit.

   CONFIGURATION mode

   ```
   DellEMC(conf)#dcb enable pfc-queues
   ```

   NOTE: To save the pfc buffering configuration changes, save the configuration and reboot the system.

NOTE: **Dell EMC Networking OS Behavior**: DCB is not supported if you enable link-level flow control on one or more interfaces. For more information, refer to Ethernet Pause Frames.

DCB Maps and its Attributes

This topic contains the following sections that describe how to configure a DCB map, apply the configured DCB map to a port, configure PFC without a DCB map, and configure lossless queues.

DCB Map: Configuration Procedure

A DCB map consists of PFC and ETS parameters. By default, PFC is not enabled on any 802.1p priority and ETS allocates equal bandwidth to each priority. To configure user-defined PFC and ETS settings, you must create a DCB map. The following is an overview of the steps involved in configuring DCB.

- Enter global configuration mode to create a DCB map or edit PFC and ETS settings.
- Configure the PFC setting (on or off) and the ETS bandwidth percentage allocated to traffic in each priority group, or whether the priority group traffic should be handled with strict priority scheduling. You can enable PFC on a maximum of two priority queues on an interface. Enabling PFC for dot1p priorities makes the corresponding port queue lossless. The sum of all allocated bandwidth percentages in all groups in the DCB map must be 100%. Strict-priority traffic is serviced first. Afterwards, you can configure either the peak rates or the committed rates. The bandwidth allocated to other priority groups is made available and allocated according to the specified percentages. If a priority group does not use its allocated bandwidth, the unused bandwidth is made available to other priority groups.
- Repeat the above procedure to configure PFC and ETS traffic handling for each priority group
- Specify the dot1p priority-to-priority group mapping for each priority. The priority group range is from 0 to 7. All priorities that map to the same queue must be in the same priority group.

  Leave a space between each priority group number. For example: **priority-pgid 0 0 0 1 2 4 4 4** in which priority group 0 maps to dot1p priorities 0, 1, and 2; priority group 1 maps to dot1p priority 3; priority group 2 maps to dot1p priority 4; priority group 4 maps to dot1p priorities 5, 6, and 7.

Important Points to Remember

- If you remove a dot1p priority-to-priority group mapping from a DCB map (`no priority pgid` command), the PFC and ETS parameters revert to their default values on the interfaces on which the DCB map is applied. By default, PFC is not applied on specific 802.1p priorities; ETS assigns equal bandwidth to each 802.1p priority.

  As a result, PFC and lossless port queues are disabled on 802.1p priorities, and all priorities are mapped to the same priority queue and equally share the port bandwidth.
- To change the ETS bandwidth allocation configured for a priority group in a DCB map, do not modify the existing DCB map configuration. Instead, first create a new DCB map with the desired PFC and ETS settings, and apply the new map to the interfaces to override the previous DCB map settings. Then, delete the original dot1p priority-priority group mapping.

  If you delete the dot1p priority-priority group mapping (`no priority pgid` command) before you apply the new DCB map, the default PFC and ETS parameters are applied on the interfaces. This change may create a DCB mismatch with peer DCB devices and interrupt network operation.

Data Center Bridging: Default Configuration

Before you configure PFC and ETS on a switch see the priority group setting taken into account the following default settings:

DCB is enabled.

PFC and ETS are globally enabled by default.

The default dot1p priority-queue assignments are applied as follows:

```
DellEMC(conf)#do show qos dot1p-queue-mapping
 Dot1p Priority : 0  1  2  3  4  5  6  7
          Queue : 0  0  0  1  2  3  3  3
DellEMC(conf)#
```

PFC is not applied on specific dot1p priorities.

ETS: Equal bandwidth is assigned to each port queue and each dot1p priority in a priority group.

To configure PFC and ETS parameters on an interface, you must specify the PFC mode, the ETS bandwidth allocation for a priority group, and the 802.1p priority-to-priority group mapping in a DCB map. No default PFC and ETS settings are applied to Ethernet interfaces.

Configuring Priority-Based Flow Control

Priority-Based Flow Control (PFC) provides a flow control mechanism based on the 802.1p priorities in converged Ethernet traffic received on an interface and is enabled by default when you enable DCB.

As an enhancement to the existing Ethernet pause mechanism, PFC stops traffic transmission for specified priorities (Class of Service (CoS) values) without impacting other priority classes. Different traffic types are assigned to different priority classes.

When traffic congestion occurs, PFC sends a pause frame to a peer device with the CoS priority values of the traffic that is to be stopped. Data Center Bridging Exchange protocol (DCBx) provides the link-level exchange of PFC parameters between peer devices. PFC allows network administrators to create zero-loss links for Storage Area Network (SAN) traffic that requires no-drop service, while retaining packet-drop congestion management for Local Area Network (LAN) traffic.

To configure PFC, follow these steps:

1. Create a DCB Map.

   CONFIGURATION mode

   `dcb-map` *`dcb-map-name`*

   The *`dcb-map-name`* variable can have a maximum of 32 characters.

2. Create a PFC group.

   CONFIGURATION mode

   `priority-group` *`group-num`* `{bandwidth` *`bandwidth`* `| strict-priority} pfc on`

   The range for priority group is from 0 to 7.

   Set the bandwidth in percentage. The percentage range is from 1 to 100% in units of 1%.

   Committed and peak bandwidth is in megabits per second. The range is from 0 to 40000.

   Committed and peak burst size is in kilobytes. Default is 50. The range is from 0 to 40000.

   The `pfc on` command enables priority-based flow control.

3. Specify the dot1p priority-to-priority group mapping for each priority.

   `priority-pgid` *`dot1p0_group_num dot1p1_group_num ...dot1p7_group_num`*

   Priority group range is from 0 to 7. All priorities that map to the same queue must be in the same priority group.

   Leave a space between each priority group number. For example: priority-pgid 0 0 0 1 2 4 4 4 in which priority group 0 maps to dot1p priorities 0, 1, and 2; priority group 1 maps to dot1p priority 3; priority group 2 maps to dot1p priority 4; priority group 4 maps to dot1p priorities 5, 6, and 7.

Dell EMC Networking OS Behavior: As soon as you apply a DCB policy with PFC enabled on an interface, DCBx starts exchanging information with PFC-enabled peers. The IEEE802.1Qbb, CEE, and CIN versions of PFC Type, Length, Value (TLV) are supported. DCBx also validates PFC configurations that are received in TLVs from peer devices.

NOTE: You cannot enable PFC and link-level flow control at the same time on an interface.

Configuring Lossless Queues

DCB also supports the manual configuration of lossless queues on an interface when PFC mode is turned off.

Prerequisite: A DCB with PFC configuration is applied to the interface with the following conditions:

- PFC mode is off (no pfc mode on).
- No PFC priority classes are configured (`no pfc priority priority-range`).

Example:

```
Port A –> Port B

Port C –> Port B
```

PFC no-drop queues are configured for queues 1, 2 on Port B. PFC capability is enabled on priorities 3, 4 on PORT A and C.

Port B acting as Egress

During the congestion, [traffic pump on priorities 3 and 4 from PORT A and PORT C is at full line rate], PORT A and C send out the PFCs to rate the traffic limit. Egress drops are not observed on Port B since traffic flow on priorities is mapped to loss less queues.

Port B acting as Ingress

If the traffic congestion is on PORT B , Egress DROP is on PORT A or C, as the PFC is not enabled on PORT B.

Refer the following configuration for queue to dot1p mapping:

NOTE: Although, each port on the S4810, S4820T, and S5000 devices support 8 QoS queues, you can configure only 4 QoS queues (0-3) to manage data traffic. The remaining 4 queues (4-7) are reserved for control traffic.

```
DellEMC(conf)#do show qos dot1p-queue-mapping
 Dot1p Priority : 0  1  2  3  4  5  6  7
          Queue : 0  0  0  1  2  3  3  3
DellEMC(conf)#
```

The configuration of no-drop queues provides flexibility for ports on which PFC is not needed but lossless traffic should egress from the interface.

Lossless traffic egresses out the no-drop queues. Ingress dot1p traffic from PFC-enabled interfaces is automatically mapped to the no-drop egress queues.

1. Enter INTERFACE Configuration mode.

   CONFIGURATION mode

   `interface` *interface-type*
2. Configure the port queues that will still function as no-drop queues for lossless traffic.

   INTERFACE mode

   `pfc no-drop queues` *queue-range*

   For the dot1p-queue assignments, refer to the dot1p Priority-Queue Assignment table.

   The maximum number of lossless queues globally supported on the switch is two.

   The range is from 0 to 3. Separate the queue values with a comma; specify a priority range with a dash; for example, `pfc no-drop queues 1,3` or `pfc no-drop queues 2-3`.

   The default: No lossless queues are configured.

NOTE: Dell EMC Networking OS Behavior: By default, no lossless queues are configured on a port.

A limit of two lossless queues is supported on a port. If the amount of priority traffic that you configure to be paused exceeds the two lossless queues, an error message displays.

Configuring PFC in a DCB Map

A switch supports the use of a DCB map in which you configure priority-based flow control (PFC) setting. To configure PFC parameters, you must apply a DCB map on an interface.

PFC Configuration Notes

PFC provides flow control based on the 802.1p priorities in a converged Ethernet traffic that is received on an interface and is enabled by default when you enable DCB. As an enhancement to the existing Ethernet pause functionality, PFC stops traffic transmission for specified priorities (CoS values) without impacting other priority classes. Different traffic types are assigned to different priority classes.

When traffic congestion occurs, PFC sends a pause frame to a peer device with the CoS priority values of the traffic that needs to be stopped. DCBx provides the link-level exchange of PFC parameters between peer devices. PFC allows network administrators to create zero-loss links for SAN traffic that requires no-drop service, while at the same time retaining packet-drop congestion management for LAN traffic.

On a switch, PFC is enabled by default on Ethernet ports (`pfc mode on` command). You can configure PFC parameters using a DCB map or the `pfc priority` command in Interface configuration mode. For more information, see Configuring Priority-Based Flow Control.

As soon as you apply a DCB map with PFC enabled on an interface, DCBx starts exchanging information with a peer. The IEEE802.1Qbb, CEE and CIN versions of PFC TLV are supported. DCBx also validates PFC configurations that are received in TLVs from peer devices. By applying a DCB map with PFC enabled, you enable PFC operations on ingress port traffic. To achieve complete lossless handling of traffic, configure PFC priorities on all DCB egress ports.

When you apply or remove a DCB input policy from an interface, one or two CRC errors are expected to be noticed on the ingress ports for each removal or attachment of the policy. This behavior occurs because the port is brought down when PFC is configured. When a DCB input policy with PFC profile is configured or unconfigured on an interface or a range of interfaces not receiving any traffic, interfaces with PFC settings that receive appropriate PFC-enabled traffic (unicast, mixed-frame-size traffic) display incremental values in the CRC and discards counters. (These ingress interfaces receiving pfc-enabled traffic have an egress interface that has a compatible PFC configuration).

NOTE: DCB maps are supported only on physical Ethernet interfaces.

- To remove a DCB map, including the PFC configuration it contains, use the `no dcb map` command in Interface configuration mode.
- To disable PFC operation on an interface, use the `no pfc mode on` command in DCB-Map configuration mode.
- Traffic may be interrupted when you reconfigure PFC no-drop priorities in a DCB map or re-apply the DCB map to an interface.
- For PFC to be applied, the configured priority traffic must be supported by a PFC peer (as detected by DCBx).
- If you apply a DCB map with PFC disabled (`pfc off`), you can enable link-level flow control on the interface using the `flowcontrol rx on tx on` command. To delete the DCB map, first disable link-level flow control. PFC is then automatically enabled on the interface because an interface is PFC-enabled by default.
- To ensure no-drop handling of lossless traffic, PFC allows you to configure lossless queues on a port (see Configuring Lossless Queues).
- When you configure a DCB map, an error message is displayed if the PFC dot1p priorities result in more than two lossless queues.
- When you apply a DCB map, an error message is displayed if link-level flow control is already enabled on an interface. You cannot enable PFC and link-level flow control at the same time on an interface.
- In a switch stack, configure all stacked ports with the same PFC configuration.
- Dell EMC Networking OS allows you to change the default dot1p priority-queue assignments only if the change satisfies the following requirements in DCB maps already applied to the interfaces:
- All 802.1p priorities mapped to the same queue must be in the same priority group.
- A maximum of two PFC-enabled, lossless queues are supported on an interface.

Otherwise, the reconfiguration of a default dot1p-queue assignment is rejected.

- To ensure complete no-drop service, apply the same PFC parameters on all PFC-enabled peers.

PFC Prerequisites and Restrictions

On a switch, PFC is globally enabled by default, but not applied on specific 802.1p priorities. To enable PFC on 802.1p priorities, create a DCB map.

The following prerequisites and restrictions apply when you configure PFC in a DCB map:

- You can enable PFC on a maximum of two priority queues on an interface. Enabling PFC for dot1p priorities configures the corresponding port queue as lossless.
- You cannot enable PFC and link-level flow control at the same time on an interface.

Applying a DCB Map on a Port

When you apply a DCB map with PFC enabled on a switch interface, a memory buffer for PFC-enabled priority traffic is automatically allocated. The buffer size is allocated according to the number of PFC-enabled priorities in the assigned map.

To apply a DCB map to an Ethernet port, follow these steps:

Table 16. DCB Map to an Ethernet Port

Step	Task	Command	Command Mode
1	Enter interface configuration mode on an Ethernet port.	`interface` *interface-type* }	CONFIGURATION
2	Apply the DCB map on the Ethernet port to configure it with the PFC and ETS settings in the map; for example: `DellEMC# interface tengigabitEthernet 1/1` `DellEMC(config-if-te-1/1)# dcb-map SAN_A_dcb_map1` Repeat Steps 1 and 2 to apply a DCB map to more than one port. You cannot apply a DCB map on an interface that has been already configured for PFC using the `pfc priority` command or which is already configured for lossless queues (`pfc no-drop queues` command).	`dcb-map` *name*	INTERFACE

Configuring PFC without a DCB Map

In a network topology that uses the default ETS bandwidth allocation (assigns equal bandwidth to each priority), you can also enable PFC for specific dot1p-priorities on individual interfaces without using a DCB map. This type of DCB configuration is useful on interfaces that require PFC for lossless traffic, but do not transmit converged Ethernet traffic.

Table 17. Configuring PFC without a DCB Map

Step	Task	Command	Command Mode
1	Enter interface configuration mode on an Ethernet port.	`interface` *interface-type*}	CONFIGURATION
2	Enable PFC on specified priorities. Range: 0-7. Default: None. Maximum number of lossless queues supported on an Ethernet port: 2. Separate priority values with a comma. Specify a priority range with a dash, for example: pfc priority 3,5-7 1. You cannot configure PFC using the `pfc priority` command on an interface on which a DCB map has been applied or which is already configured for lossless queues (`pfc no-drop queues` command).	`pfc priority` *priority-range*	INTERFACE

Configuring Lossless Queues

DCB also supports the manual configuration of lossless queues on an interface when PFC mode is disabled in a DCB map, apply the map on the interface. The configuration of no-drop queues provides flexibility for ports on which PFC is not needed, but lossless traffic should egress from the interface.

Configuring no-drop queues is applicable only on the interfaces which do not need PFC.

Example:

```
Port A –> Port B

Port C –> Port B
```

PFC no-drop queues are configured for queues 1, 2 on Port B. PFC capability is enabled on priorities 3, 4 on PORT A and C.

Port B acting as Egress

During the congestion, [traffic pump on priorities 3 and 4 from PORT A and PORT C is at full line rate], PORT A and C send out the PFCs to rate the traffic limit. Egress drops are not observed on Port B since traffic flow on priorities is mapped to loss less queues.

Port B acting as Ingress

If the traffic congestion is on PORT B , Egress DROP is on PORT A or C, as the PFC is not enabled on PORT B.

Refer the following configuration for queue to dot1p mapping:

```
DellEMC(conf)#do show qos dot1p-queue-mapping
 Dot1p Priority : 0  1  2  3  4  5  6  7  -> On ingress interfaces[Port A and C] we used
the PFC on priority level.
          Queue : 0  0  0  1  2  3  3  3  -> On Egress interface[Port B] we used no-drop
queues.
```

Lossless traffic egresses out the no-drop queues. Ingress 802.1p traffic from PFC-enabled peers is automatically mapped to the no-drop egress queues.

When configuring lossless queues on a port interface, consider the following points:

- By default, no lossless queues are configured on a port.
- A limit of two lossless queues is supported on a port. If the number of lossless queues configured exceeds the maximum supported limit per port (two), an error message is displayed. Reconfigure the value to a smaller number of queues.
- If you configure lossless queues on an interface that already has a DCB map with PFC enabled (**pfc on**), an error message is displayed.

Table 18. Configuring Lossless Queues on a Port Interface

Step	Task	Command	Command Mode
1	Enter INTERFACE Configuration mode.	`interface` *interface-type*	CONFIGURATION
2	Open a DCB map and enter DCB map configuration mode.	`dcb-map` *name*	INTERFACE
3	Disable PFC.	`no pfc mode on`	DCB MAP
4	Return to interface configuration mode.	`exit`	DCB MAP
5	Apply the DCB map, created to disable the PFC operation, on the interface	`dcb-map` {*name* | **default**}	INTERFACE
6	Configure the port queues that still function as no-drop queues for lossless traffic. For the dot1p-queue assignments.	`pfc no-drop queues`*queue-range*	INTERFACE

Table 18. Configuring Lossless Queues on a Port Interface (continued)

Step	Task	Command	Command Mode
	The maximum number of lossless queues globally supported on a port is 2. You cannot configure PFC no-drop queues on an interface on which a DCB map with PFC enabled has been applied, or which is already configured for PFC using the `pfc priority` command. Range: 0-3. Separate queue values with a comma; specify a priority range with a dash; for example: pfc no-drop queues 1,3 or pfc no-drop queues 2-3 Default: No lossless queues are configured.		

Priority-Based Flow Control Using Dynamic Buffer Method

In a data center network, priority-based flow control (PFC) manages large bursts of one traffic type in multiprotocol links so that it does not affect other traffic types and no frames are lost due to congestion. When PFC detects congestion on a queue for a specified priority, it sends a pause frame for the 802.1p priority traffic to the transmitting device.

Pause and Resume of Traffic

The pause message is used by the sending device to inform the receiving device about a congested, heavily-loaded traffic state that has been identified. When the interface of a sending device transmits a pause frame, the recipient acknowledges this frame by temporarily halting the transmission of data packets. The sending device requests the recipient to restart the transmission of data traffic when the congestion eases and reduces. The time period that is specified in the pause frame defines the duration for which the flow of data packets is halted. When the time period elapses, the transmission restarts.

When a device sends a pause frame to another device, the time for which the sending of packets from the other device must be stopped is contained in the pause frame. The device that sent the pause frame empties the buffer to be less than the threshold value and restarts the acceptance of data packets.

Dynamic ingress buffering enables the sending of pause frames at different thresholds based on the number of ports that experience congestion at a time. This behavior impacts the total buffer size used by a particular lossless priority on an interface. The pause and resume thresholds can also be configured dynamically. You can configure a buffer size, pause threshold, ingress shared threshold weight, and resume threshold to control and manage the total amount of buffers that are to be used in your network environment.

Buffer Sizes for Lossless or PFC Packets

You can configure up to a maximum of 4 lossless (PFC) queues. By configuring 4 lossless queues, you can configure 4 different priorities and assign a particular priority to each application that your network is used to process. For example, you can assign a higher priority for time-sensitive applications and a lower priority for other services, such as file transfers. You can configure the amount of buffer space to be allocated for each priority and the pause or resume thresholds for the buffer. This method of configuration enables you to effectively manage and administer the behavior of lossless queues.

Although the system contains 9 MB of space for shared buffers, a minimum guaranteed buffer is provided to all the internal and external ports in the system for both unicast and multicast traffic. This minimum guaranteed buffer reduces the total available shared buffer to 7,787 KB. This shared buffer can be used for lossy and lossless traffic.

The default behavior causes up to a maximum of 6.6 MB to be used for PFC-related traffic. The remaining approximate space of 1 MB can be used by lossy traffic. You can allocate all the remaining 1 MB to lossless PFC queues. If you allocate in such a way, the performance of lossy traffic is reduced and degraded. Although you can allocate a maximum buffer size, it is used only if a PFC priority is configured and applied on the interface.

The number of lossless queues supported on the system is dependent on the availability of total buffers for PFC. The default configuration in the system guarantees a minimum of 52 KB per queue if all the 128 queues are congested. However, modifying the buffer allocation per queue impacts this default behavior.

By default the total available buffer for PFC is 6.6 MB and when you configure dynamic ingress buffering, a minimum of least 52 KB per queue is used when all ports are congested. By default, the system enables a maximum of two lossless queues on the S4820T platform.

This default behavior is impacted if you modify the total buffer available for PFC or assign static buffer configurations to the individual PFC queues.

Behavior of Tagged Packets

The below is example for enabling PFC for priority 2 for tagged packets. Priority (Packet Dot1p) 2 will be mapped to PG6 on PRIO2PG setting. All other Priorities for which PFC is not enabled are mapped to default PG – PG7.

Classification rules on ingress (Ingress FP CAM region) matches incoming packet-dot1p and assigns an internal priority (to select queue as per Table 1 and Table 2).

The internal Priority assigned for the packet by Ingress FP is used by the memory management unit (MMU) to assign the packet to right queue by indexing the internal-priority to queue map table (TABLE 1) in hardware.

PRIO2COS setting for honoring the PFC protocol packets from the Peer switches is as per above Packet-Dot1p->queue table (Table 2).

The packets come in with packet-dot1p 2 alone are assign to PG6 on ingress.

The packets come in with packet-dot1p 2 alone use Q1 (as per dot1p to Queue classification – Table 2) on the egress port.

- When Peer sends a PFC message for Priority 2, based on above PRIO2COS table (TABLE 2), Queue 1 is halted.
- Queue 1 starts buffering the packets with Dot1p 2. This causes PG6 buffer counter to increase on the ingress, since P-dot1p 2 is mapped to PG6.
- As the PG6 watermark threshold is reached, PFC generates for dot1p 2.

Configuration Example for DSCP and PFC Priorities

Consider a scenario in which the following DSCP and PFC priorities are necessary:

DSCP	0 – 5, 10 - 15	20 – 25, 30 – 35
Expected PFC Priority	1	2

To configure the aforementioned DSCP and PFC priority values, perform the following tasks:

1. Create class-maps to group the DSCP subsets

```
class-map match-any dscp-pfc-1
 match ip dscp 0-5,10-15
!
class-map match-any dscp-pfc-2
 match ip dscp 20-25,30-35
```

2. Associate above class-maps to Queues Queue assignment as below.

NOTE: Although, each port on the S4810, S4820T, and S5000 devices support 8 QoS queues, you can configure only 4 QoS queues (0-3)to manage data traffic. The remaining 4 queues (4-7) are reserved for control traffic.

Table 19. Queue Assignments

Internal-priority	Queue
0	0
1	0
2	0
3	1

Table 19. Queue Assignments (continued)

Internal-priority	Queue
4	2
5	3
6	3
7	3

Table 19. Queue Assignments

Internal-priority	Queue
0	2
1	0
2	1
3	3
4	4
5	5
6	6
7	7

3. Dot1p->Queue Mapping Configuration is retained at the default value.

```
Default dot1p-queue mapping is,
DellEMC#show qos dot1p-queue-mapping
 Dot1p Priority : 0  1  2  3  4  5  6  7
          Queue : 0  0  0  1  2  3  3  3
```

```
Default dot1p-queue mapping is,
DellEMC#show qos dot1p-queue-mapping
 Dot1p Priority : 0  1  2  3  4  5  6  7
          Queue : 2  0  1  3  4  5  6  7
```

4. Interface Configurations on server connected ports.
   a. Enable DCB globally.

```
DellEMC(conf)#dcb enable
```

   b. Apply PFC Priority configuration. Configure priorities on which PFC is enabled.

```
DellEMC(conf-if-te-1/1)#pfc priority 1,2
```

Using PFC to Manage Converged Ethernet Traffic

To use PFC for managing converged Ethernet traffic, use the following command:

```
dcb-map stack-unit all dcb-map-name
```

Configure Enhanced Transmission Selection

ETS provides a way to optimize bandwidth allocation to outbound 802.1p classes of converged Ethernet traffic.

Different traffic types have different service needs. Using ETS, you can create groups within an 802.1p priority class to configure different treatment for traffic with different bandwidth, latency, and best-effort needs.

For example, storage traffic is sensitive to frame loss; interprocess communication (IPC) traffic is latency-sensitive. ETS allows different traffic types to coexist without interruption in the same converged link by:

- Allocating a guaranteed share of bandwidth to each priority group.
- Allowing each group to exceed its minimum guaranteed bandwidth if another group is not fully using its allotted bandwidth.

ETS Prerequisites and Restrictions

The following prerequisites and restrictions apply when you configure ETS bandwidth allocation or queue scheduling.

- Configuring ETS bandwidth allocation or a queue scheduler for dot1p priorities in a priority group is applicable if the DCBx version used on a port is CIN (refer to Configuring DCBx).
- When allocating bandwidth or configuring a queue scheduler for dot1p priorities in a priority group on a DCBx CIN interface, take into account the CIN bandwidth allocation (refer to Configuring Bandwidth Allocation for DCBx CIN) and dot1p-queue mapping.

NOTE: The IEEE 802.1Qaz, CEE, and CIN versions of ETS are supported.

Creating an ETS Priority Group

An ETS priority group specifies the range of 802.1p priority traffic to which a QoS output policy with ETS settings is applied on an egress interface.

1. Configure a DCB Map.

   CONFIGURATION mode

   dcb-map *dcb-map-name*

   The *dcb-map-name* variable can have a maximum of 32 characters.
2. Create an ETS priority group.

   CONFIGURATION mode

   `priority-group group-num {bandwidth bandwidth | strict-priority} pfc off`

   The range for priority group is from 0 to 7.

   Set the bandwidth in percentage. The percentage range is from 1 to 100% in units of 1%.

   Committed and peak bandwidth is in megabits per second. The range is from 0 to 40000.

   Committed and peak burst size is in kilobytes. Default is 50. The range is from 0 to 10000.
3. Configure the 802.1p priorities for the traffic on which you want to apply an ETS output policy.

   PRIORITY-GROUP mode

   `priority-list value`

   The range is from 0 to 7.

   The default is **none**.

   Separate priority values with a comma. Specify a priority range with a dash. For example, priority-list 3,5-7.
4. Exit priority-group configuration mode.

   PRIORITY-GROUP mode

   `exit`
5. Repeat Steps 1 to 4 to configure all remaining dot1p priorities in an ETS priority group.
6. Specify the dot1p priority-to-priority group mapping for each priority.

   `priority-pgid dot1p0_group_num dot1p1_group_num ...dot1p7_group_num`

   Priority group range is from 0 to 7. All priorities that map to the same queue must be in the same priority group.

Leave a space between each priority group number. For example: priority-pgid 0 0 0 1 2 4 4 4 in which priority group 0 maps to dot1p priorities 0, 1, and 2; priority group 1 maps to dot1p priority 3; priority group 2 maps to dot1p priority 4; priority group 4 maps to dot1p priorities 5, 6, and 7.

Dell EMC Networking OS Behavior: A priority group consists of 802.1p priority values that are grouped for similar bandwidth allocation and scheduling, and that share latency and loss requirements. All 802.1p priorities mapped to the same queue must be in the same priority group.

Configure all 802.1p priorities in priority groups associated with an ETS output policy. You can assign each dot1p priority to only one priority group.

By default, all 802.1p priorities are grouped in priority group 0 and 100% of the port bandwidth is assigned to priority group 0. The complete bandwidth is equally assigned to each priority class so that each class has 12 to 13%.

The maximum number of priority groups supported in ETS output policies on an interface is equal to the number of data queues (8)on the port. The 802.1p priorities in a priority group can map to multiple queues.

If you configure more than one priority queue as strict priority or more than one priority group as strict priority, the higher numbered priority queue is given preference when scheduling data traffic.

If multiple lossful priorities are mapped to a single priority group (PG1) and lossless priorities to another priority group (PG0), then bandwidth split across lossful priorities is not even.

ETS Operation with DCBx

The following section describes DCBx negotiation with peer ETS devices.

In DCBx negotiation with peer ETS devices, ETS configuration is handled as follows:

- ETS TLVs are supported in DCBx versions CIN, CEE, and IEEE2.5.
- The DCBx port-role configurations determine the ETS operational parameters (refer to Configure a DCBx Operation).
- ETS configurations received from TLVs from a peer are validated.
- If there is a hardware limitation or TLV error:
  - DCBx operation on an ETS port goes down.
  - New ETS configurations are ignored and existing ETS configurations are reset to the default ETS settings.
- ETS operates with legacy DCBx versions as follows:
  - In the CEE version, the priority group/traffic class group (TCG) ID 15 represents a non-ETS priority group. Any priority group configured with a scheduler type is treated as a strict-priority group and is given the priority-group (TCG) ID 15.
  - The CIN version supports two types of strict-priority scheduling:
    - Group strict priority: Use this to increase its bandwidth usage to the bandwidth total of the priority group and allow a single priority flow in a priority group. A single flow in a group can use all the bandwidth allocated to the group.
    - Link strict priority: Use this to increase to the maximum link bandwidth and allow a flow in any priority group.

    NOTE: CIN supports only the dot1p priority-queue assignment in a priority group. To configure a dot1p priority flow in a priority group to operate with link strict priority, you configure: The dot1p priority for strict-priority scheduling (`strict-priority` command). The priority group for strict-priority scheduling (`scheduler strict` command.

Configuring Bandwidth Allocation for DCBx CIN

After you apply an ETS output policy to an interface, if the DCBx version used in your data center network is CIN, you may need to configure a QoS output policy to overwrite the default CIN bandwidth allocation.

This default setting divides the bandwidth allocated to each port queue equally between the dot1p priority traffic assigned to the queue.

To create a QoS output policy that allocates different amounts of bandwidth to the different traffic types/ dot1p priorities assigned to a queue and apply the output policy to the interface, follow these steps.

1. Create a QoS output policy.

   CONFIGURATION mode

   ```
   Dell(conf)#qos-policy-output test12
   ```

   The maximum 32 alphanumeric characters.

2. Configure the percentage of bandwidth to allocate to the dot1p priority/queue traffic in the associated L2 class map.

   QoS OUTPUT POLICY mode

```
Dell(conf-qos-policy-out)#bandwidth-percentage 100
```

The default is **none**.

3. Repeat Step 2 to configure bandwidth percentages for other priority queues on the port.

   QoS OUTPUT POLICY mode

   ```
   Dell(conf-qos-policy-out)#bandwidth-percentage 100
   ```

4. Exit QoS Output Policy Configuration mode.

   QoS OUTPUT POLICY mode

   ```
   Dell(conf-if-te-0/1)#exit
   ```

5. Enter INTERFACE Configuration mode.

   CONFIGURATION mode

   ```
   interface type slot/port
   ```

6. Apply the QoS output policy with the bandwidth percentage for specified priority queues to an egress interface.

   INTERFACE mode

   ```
   Dell(conf-if-te-0/1)#service-policy output test12
   ```

Configuring ETS in a DCB Map

A switch supports the use of a DCB map in which you configure enhanced transmission selection (ETS) setting. To configure ETS parameters, you must apply a DCB map on an interface.

ETS Configuration Notes

ETS provides a way to optimize bandwidth allocation to outbound 802.1p classes of converged Ethernet traffic. Different traffic types have different service needs. Using ETS, you can create groups within an 802.1p priority class to configure different treatment for traffics with different bandwidth, latency, and best-effort needs.

When you configure ETS in a DCB map:

- The DCB map associates a priority group with a PFC operational mode (on or off) and an ETS scheduling and bandwidth allocation. You can apply a DCB map on multiple egress ports.
- Use the ETS configuration associated with 802.1p priority traffic in a DCB map in DCBx negotiation with ETS peers.
- Traffic in priority groups is assigned to strict-queue or weighted round-robin (WRR) scheduling in an ETS configuration and is managed using the ETS bandwidth-assignment algorithm. Dell EMC Networking OS de-queues all frames of strict-priority traffic before servicing any other queues. A queue with strict-priority traffic can starve other queues in the same port.
- ETS-assigned bandwidth allocation and strict-priority scheduling apply only to data queues, not to control queues.
- Dell EMC Networking OS supports hierarchical scheduling on an interface. The control traffic on Dell EMC Networking OS is redirected to control queues as higher priority traffic with strict priority scheduling. After the control queues drain out, the remaining data traffic is scheduled to queues according to the bandwidth and scheduler configuration in the DCB map. The available bandwidth calculated by the ETS algorithm is equal to the link bandwidth after scheduling non-ETS higher-priority traffic.
- The configuration of bandwidth allocation and strict-queue scheduling is not supported at the same time for a priority group.
- **Bandwidth assignment**: By default, equal bandwidth is assigned to each dot1p priority in a priority group. To configure the bandwidth assigned to the port queues associated with dot1p priorities in a priority group, use the **bandwidth** *percentage* parameter. The sum of the bandwidth allocated to all priority groups in a DCB map must be 100% of the bandwidth on the link. You must allocate at least 1% of the total bandwidth to each priority group.
- **Scheduling of priority traffic**: dot1p priority traffic on the switch is scheduled to the current queue mapping. dot1p priorities within the same queue must have the same traffic properties and scheduling method.
- **ETS configuration error**: If an error occurs in an ETS configuration, the configuration is ignored and the scheduler and bandwidth allocation settings are reset to the ETS default value: 100% of available bandwidth is allocated to priority group 0 and the bandwidth is equally assigned to each dot1p priority.

If an error occurs when a port receives a peer's ETS configuration, the port's configuration resets to the ETS configuration in the previously configured DCB map. If no DCB map was previously applied, the port resets to the default ETS parameters.

ETS Prerequisites and Restrictions

On a switch, ETS is enabled by default on Ethernet ports with equal bandwidth assigned to each 802.1p priority. You can change the default ETS configuration only by using a DCB map.

The following prerequisites and restrictions apply when you configure ETS bandwidth allocation or strict-priority queuing in a DCB map:

- When allocating bandwidth or configuring strict-priority queuing for dot1p priorities in a priority group on a DCBx CIN interface, take into account the CIN bandwidth allocation and dot1p-queue mapping.
- Because all the priorities mapped to a priority group is scheduled using a single queue, the priorities are treated with first come first served basis.
- Although ETS bandwidth allocation or strict-priority queuing does not support weighted random early detection (WRED), explicit congestion notification (ECN), rate shaping, and rate limiting because these parameters are not negotiated by DCBx with peer devices, you can apply a QoS output policy with WRED and/or rate shaping on a DCBx CIN-enabled interface. In this case, the WRED or rate shaping configuration in the QoS output policy must take into account the bandwidth allocation or queue scheduler configured in the DCB map.

Priority-Group Configuration Notes

When you configure priority groups in a DCB map:

- A priority group consists of 802.1p priority values that are grouped together for similar bandwidth allocation and scheduling, and that share the same latency and loss requirements. All 802.1p priorities mapped to the same queue must be in the same priority group.
- In a DCB map, each 802.1p priority must map to a priority group.
- The maximum number of priority groups supported in a DCB map on an interface is equal to the number of data queues (4) on the port. Each priority group can support more than one data queue.
- You can enable PFC on a maximum of two priority queues on an interface.
- If you configure more than one priority group as strict priority, the higher numbered priority queue is given preference when scheduling data traffic.

Hierarchical Scheduling in ETS Output Policies

ETS supports up to three levels of hierarchical scheduling.

For example, you can apply ETS output policies with the following configurations:

Priority group 1	Assigns traffic to one priority queue with 20% of the link bandwidth and strict-priority scheduling.
Priority group 2	Assigns traffic to one priority queue with 30% of the link bandwidth.
Priority group 3	Assigns traffic to two priority queues with 50% of the link bandwidth and strict-priority scheduling.

In this example, the configured ETS bandwidth allocation and scheduler behavior is as follows:

Unused bandwidth usage:	Normally, if there is no traffic or unused bandwidth for a priority group, the bandwidth allocated to the group is distributed to the other priority groups according to the bandwidth percentage allocated to each group. However, when three priority groups with different bandwidth allocations are used on an interface: • If priority group 3 has free bandwidth, it is distributed as follows: 20% of the free bandwidth to priority group 1 and 30% of the free bandwidth to priority group 2. • If priority group 1 or 2 has free bandwidth, (20 + 30)% of the free bandwidth is distributed to priority group 3. Priority groups 1 and 2 retain whatever free bandwidth remains up to the (20+ 30)%.
Strict-priority groups:	If two priority groups have strict-priority scheduling, traffic assigned from the priority group with the higher priority-queue number is scheduled first. However, when three priority groups are used and two groups have strict-priority scheduling (such as groups 1 and 3 in the example), the strict priority group whose traffic is mapped to one queue takes precedence over the strict priority group whose traffic is mapped to two queues.

Therefore, in this example, scheduling traffic to priority group 1 (mapped to one strict-priority queue) takes precedence over scheduling traffic to priority group 3 (mapped to two strict-priority queues).

Using ETS to Manage Converged Ethernet Traffic

To use ETS for managing converged Ethernet traffic, use the following command:

```
dcb-map stack-unit all dcb-map-name
```

Applying DCB Policies in a Switch Stack

You can apply DCB policies with PFC and ETS configurations to all stacked ports in a switch stack or on a stacked switch.

To apply DCB policies in a switch stack, follow this step.

NOTE: Use only 40G ports as stacking ports when you enable DCB. S4820T does not support DCB when you use 10GBaseT ports as stacking ports.

- Apply the specified DCB policy on all ports of the switch stack or a single stacked switch.

  CONFIGURATION mode

  ```
  dcb-map {stack-unit all | stack-ports all} dcb-map-name
  ```

Configure a DCBx Operation

DCB devices use data center bridging exchange protocol (DCBx) to exchange configuration information with directly connected peers using the link layer discovery protocol (LLDP) protocol.

DCBx can detect the misconfiguration of a peer DCB device, and optionally, configure peer DCB devices with DCB feature settings to ensure consistent operation in a data center network.

DCBx is a prerequisite for using DCB features, such as priority-based flow control (PFC) and enhanced traffic selection (ETS), to exchange link-level configurations in a converged Ethernet environment. DCBx is also deployed in topologies that support lossless operation for FCoE or iSCSI traffic. In these scenarios, all network devices are DCBx-enabled (DCBx is enabled end-to-end). For more information about how these features are implemented and used, refer to:

- Configure Enhanced Transmission Selection

DCBx supports the following versions: CIN, CEE, and IEEE2.5.

Prerequisite: For DCBx, enable LLDP on all DCB devices.

DCBx Operation

DCBx performs the following operations:

- Discovers DCB configuration (such as PFC and ETS) in a peer device.
- Detects DCB mis-configuration in a peer device; that is, when DCB features are not compatibly configured on a peer device and the local switch. Mis-configuration detection is feature-specific because some DCB features support asymmetric configuration.
- Reconfigures a peer device with the DCB configuration from its configuration source if the peer device is willing to accept configuration.
- Accepts the DCB configuration from a peer if a DCBx port is in “willing” mode to accept a peer’s DCB settings and then internally propagates the received DCB configuration to its peer ports.

DCBx Port Roles

To enable the auto-configuration of DCBx-enabled ports and propagate DCB configurations learned from peer DCBx devices internally to other switch ports, use the following DCBx port roles.

Auto-upstream The port advertises its own configuration to DCBx peers and is *willing* to receive peer configuration. The port also propagates its configuration to other ports on the switch.

The first auto-upstream that is capable of receiving a peer configuration is elected as the configuration source. The elected configuration source then internally propagates the configuration to other auto-upstream and auto-downstream ports. A port that receives an internally propagated configuration overwrites its local configuration with the new parameter values. When an auto-upstream port (besides the configuration source) receives and overwrites its configuration with internally propagated information, one of the following actions is taken:

- If the peer configuration received is compatible with the internally propagated port configuration, the link with the DCBx peer is enabled.
- If the received peer configuration is not compatible with the currently configured port configuration, the link with the DCBx peer port is disabled and a syslog message for an incompatible configuration is generated. The network administrator must then reconfigure the peer device so that it advertises a compatible DCB configuration.
  - The configuration received from a DCBx peer or from an internally propagated configuration is not stored in the switch's running configuration.
  - On a DCBx port in an auto-upstream role, the PFC and application priority TLVs are enabled. ETS recommend TLVs are disabled and ETS configuration TLVs are enabled.

Auto-downstream

The port advertises its own configuration to DCBx peers but is *not willing* to receive remote peer configuration. The port always accepts internally propagated configurations from a configuration source. An auto-downstream port that receives an internally propagated configuration overwrites its local configuration with the new parameter values.

When an auto-downstream port receives and overwrites its configuration with internally propagated information, one of the following actions is taken:

- If the peer configuration received is compatible with the internally propagated port configuration, the link with the DCBx peer is enabled.
- If the received peer configuration is not compatible with the currently configured port configuration, the link with the DCBx peer port is disabled and a syslog message for an incompatible configuration is generated. The network administrator must then reconfigure the peer device so that it advertises a compatible DCB configuration.
  - The internally propagated configuration is not stored in the switch's running configuration.
  - On a DCBx port in an auto-downstream role, all PFC, application priority, ETS recommend, and ETS configuration TLVs are enabled.

Configuration source

The port is configured to serve as a source of configuration information on the switch. Peer DCB configurations received on the port are propagated to other DCBx auto-configured ports. If the peer configuration is compatible with a port configuration, DCBx is enabled on the port.

On a configuration-source port, the link with a DCBx peer is enabled when the port receives a DCB configuration that can be internally propagated to other auto-configured ports. The configuration received from a DCBx peer is not stored in the switch's running configuration. On a DCBx port that is the configuration source, all PFC and application priority TLVs are enabled. ETS recommend TLVs are disabled and ETS configuration TLVs are enabled.

Manual

The port is configured to operate only with administrator-configured settings and does not auto-configure with DCB settings received from a DCBx peer or from an internally propagated configuration from the configuration source. If you enable DCBx, ports in Manual mode advertise their configurations to peer devices but do not accept or propagate internal or external configurations. Unlike other user-configured ports, the configuration of DCBx ports in Manual mode is saved in the running configuration.

On a DCBx port in a manual role, all PFC, application priority, ETS recommend, and ETS configuration TLVs are enabled.

When making a configuration change to a DCBx port in a Manual role, Dell EMC Networking recommends shutting down the interface using the `shutdown` command, change the configuration, then re-activate the interface using the `no shutdown` command.

The default for the DCBx port role is **manual**.

NOTE: On a DCBx port, application priority TLV advertisements are handled as follows:

- The application priority TLV is transmitted only if the priorities in the advertisement match the configured PFC priorities on the port.
- On auto-upstream and auto-downstream ports:

- If a configuration source is elected, the ports send an application priority TLV based on the application priority TLV received on the configuration-source port. When an application priority TLV is received on the configuration-source port, the auto-upstream and auto-downstream ports use the internally propagated PFC priorities to match against the received application priority. Otherwise, these ports use their locally configured PFC priorities in application priority TLVs.
- If no configuration source is configured, auto-upstream and auto-downstream ports check to see that the locally configured PFC priorities match the priorities in a received application priority TLV.

- On manual ports, an application priority TLV is advertised only if the priorities in the TLV match the PFC priorities configured on the port.

DCB Configuration Exchange

The DCBx protocol supports the exchange and propagation of configuration information for the enhanced transmission selection (ETS) and priority-based flow control (PFC) DCB features.

DCBx uses the following methods to exchange DCB configuration parameters:

Asymmetric	DCB parameters are exchanged between a DCBx-enabled port and a peer port without requiring that a peer port and the local port use the same configured values for the configurations to be compatible. For example, ETS uses an asymmetric exchange of parameters between DCBx peers.
Symmetric	DCB parameters are exchanged between a DCBx-enabled port and a peer port but requires that each configured parameter value be the same for the configurations in order to be compatible. For example, PFC uses an symmetric exchange of parameters between DCBx peers.

Configuration Source Election

When an auto-upstream or auto-downstream port receives a DCB configuration from a peer, the port first checks to see if there is an active configuration source on the switch.

- If a configuration source already exists, the received peer configuration is checked against the local port configuration. If the received configuration is compatible, the DCBx marks the port as DCBx-enabled. If the configuration received from the peer is not compatible, a warning message is logged and the DCBx frame error counter is incremented. Although DCBx is operationally disabled, the port keeps the peer link up and continues to exchange DCBx packets. If a compatible peer configuration is later received, DCBx is enabled on the port.
- If there is no configuration source, a port may elect itself as the configuration source. A port may become the configuration source if the following conditions exist:
  - No other port is the configuration source.
  - The port role is auto-upstream.
  - The port is enabled with link up and DCBx enabled.
  - The port has performed a DCBx exchange with a DCBx peer.
  - The switch is capable of supporting the received DCB configuration values through either a symmetric or asymmetric parameter exchange.

A newly elected configuration source propagates configuration changes received from a peer to the other auto-configuration ports. Ports receiving auto-configuration information from the configuration source ignore their current settings and use the configuration source information.

Propagation of DCB Information

When an auto-upstream or auto-downstream port receives a DCB configuration from a peer, the port acts as a DCBx client and checks if a DCBx configuration source exists on the switch.

- If a configuration source is found, the received configuration is checked against the currently configured values that are internally propagated by the configuration source. If the local configuration is compatible with the received configuration, the port is enabled for DCBx operation and synchronization.
- If the configuration received from the peer is not compatible with the internally propagated configuration used by the configuration source, the port is disabled as a client for DCBx operation and synchronization and a syslog error message is generated. The port keeps the peer link up and continues to exchange DCBx packets. If a compatible configuration is later received from the peer, the port is enabled for DCBx operation.

NOTE: DCB configurations internally propagated from a configuration source do not overwrite the configuration on a DCBx port in a manual role. When a configuration source is elected, all auto-upstream ports other than the configuration source are marked as *willing disabled*. The internally propagated DCB configuration is refreshed on all auto-configuration ports and each port may begin configuration negotiation with a DCBx peer again.

Auto-Detection and Manual Configuration of the DCBx Version

When operating in Auto-Detection mode (the `DCBx version auto` command), a DCBx port automatically detects the DCBx version on a peer port. Legacy CIN and CEE versions are supported in addition to the standard IEEE version 2.5 DCBx.

A DCBx port detects a peer version after receiving a valid frame for that version. The local DCBx port reconfigures to operate with the peer version and maintains the peer version on the link until one of the following conditions occurs:

- The switch reboots.
- The link is reset (goes down and up).
- User-configured CLI commands require the version negotiation to restart.
- The peer times out.
- Multiple peers are detected on the link.

If you configure a DCBx port to operate with a specific version (the `DCBx version {cee | cin | ieee-v2.5}` command in the Configuring DCBx), DCBx operations are performed according to the configured version, including fast and slow transmit timers and message formats. If a DCBx frame with a different version is received, a syslog message is generated and the peer version is recorded in the peer status table. If the frame cannot be processed, it is discarded and the discard counter is incremented.

NOTE: Because DCBx TLV processing is best effort, it is possible that CIN frames may be processed when DCBx is configured to operate in CEE mode and vice versa. In this case, the unrecognized TLVs cause the unrecognized TLV counter to increment, but the frame is processed and is not discarded.

Legacy DCBx (CIN and CEE) supports the DCBx control state machine that is defined to maintain the sequence number and acknowledge the number sent in the DCBx control TLVs.

DCBx Example

The following figure shows how to use DCBx.

The external 40GbE 40GbE ports on the base module (ports 33 and 37) of two switches are used for uplinks configured as DCBx auto-upstream ports. The device is connected to third-party, top-of-rack (ToR) switches through 40GbE uplinks. The ToR switches are part of a Fibre Channel storage network.

The internal ports (ports 1-32) connected to the 10GbE backplane are configured as auto-downstream ports.

On the S4820T, PFC and ETS use DCBx to exchange link-level configuration with DCBx peer devices.

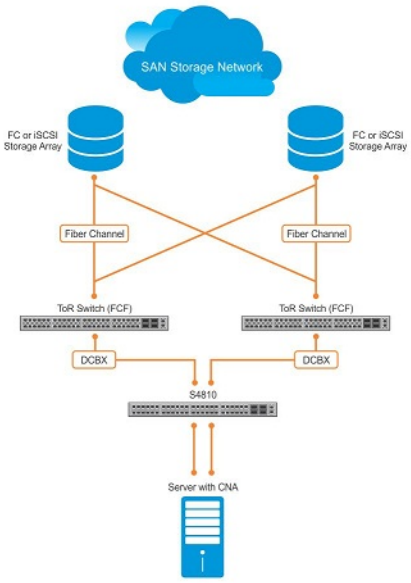


Figure 32. DCBx Sample Topology

DCBx Prerequisites and Restrictions

The following prerequisites and restrictions apply when you configure DCBx operation on a port:

- For DCBx, on a port interface, enable LLDP in both Send (TX) and Receive (RX) mode (the `protocol lldp mode` command; refer to the example in CONFIGURATION versus INTERFACE Configurations in the Link Layer Discovery Protocol (LLDP) chapter). If multiple DCBx peer ports are detected on a local DCBx interface, LLDP is shut down.
- The CIN version of DCBx supports only PFC, ETS, and FCOE; it does not support iSCSI, backward congestion management (BCN), logical link down (LLDF), and network interface virtualization (NIV).

Configuring DCBx

To configure DCBx, follow these steps.

For DCBx, to advertise DCBx TLVs to peers, enable LLDP. For more information, refer to Link Layer Discovery Protocol (LLDP).

Configure DCBx operation at the interface level on a switch or globally on the switch. To configure the S4820Tsystem for DCBx operation in a data center network, you must:

1. Configure ToR- and FCF-facing interfaces as auto-upstream ports.
2. Configure server-facing interfaces as auto-downstream ports.
3. Configure a port to operate in a configuration-source role.
4. Configure ports to operate in a manual role.

1. Enter INTERFACE Configuration mode.

   CONFIGURATION mode

   `interface type slot/port`

2. Enter LLDP Configuration mode to enable DCBx operation.

   INTERFACE mode

   `[no] protocol lldp`

3. Configure the DCBx version used on the interface, where: `auto` configures the port to operate using the DCBx version received from a peer.

   PROTOCOL LLDP mode

   `[no] DCBx version {auto | cee | cin | ieee-v2.5}`

   - `cee`: configures the port to use CEE (Intel 1.01).
   - `cin`: configures the port to use Cisco-Intel-Nuova (DCBx 1.0).

- `ieee-v2.5`: configures the port to use IEEE 802.1Qaz (Draft 2.5).

The default is **Auto**.

4. Configure the DCBx port role the interface uses to exchange DCB information.
   PROTOCOL LLDP mode
   `[no] DCBx port-role {config-source | auto-downstream | auto-upstream | manual}`
   - `auto-upstream`: configures the port to receive a peer configuration. The configuration source is elected from auto-upstream ports.
   - `auto-downstream`: configures the port to accept the internally propagated DCB configuration from a configuration source.
   - `config-source`: configures the port to serve as the configuration source on the switch.
   - `manual`: configures the port to operate only on administer-configured DCB parameters. The port does not accept a DCB configuration received from a peer or a local configuration source.

   The default is **Manual**.
5. **On manual ports only**: Configure the PFC and ETS TLVs advertised to DCBx peers.
   PROTOCOL LLDP mode
   `[no] advertise DCBx-tlv {ets-conf | ets-reco | pfc} [ets-conf | ets-reco | pfc] [ets-conf | ets-reco | pfc]`
   - `ets-conf`: enables the advertisement of ETS Configuration TLVs.
   - `ets-reco`: enables the advertisement of ETS Recommend TLVs.
   - `pfc enables`: the advertisement of PFC TLVs.

   The default is All PFC and ETS TLVs are advertised.

   NOTE: You can configure the transmission of more than one TLV type at a time; for example, advertise `DCBx-tlv ets-conf ets-reco`. You can enable ETS recommend TLVs (`ets-reco`) only if you enable ETS configuration TLVs (`ets-conf`).

   To disable TLV transmission, use the `no` form of the command; for example, `no advertise DCBx-tlv pfc ets-reco`.
6. **On manual ports only**: Configure the Application Priority TLVs advertised on the interface to DCBx peers.
   PROTOCOL LLDP mode
   `[no] advertise DCBx-appln-tlv {fcoe | iscsi}`
   - `fcoe`: enables the advertisement of FCoE in Application Priority TLVs.
   - `iscsi`: enables the advertisement of iSCSI in Application Priority TLVs.

   The default is Application Priority TLVs are enabled to advertise FCoE and iSCSI.

   NOTE: To disable TLV transmission, use the `no` form of the command; for example, `no advertise DCBx-appln-tlv iscsi`.

   For information about how to use iSCSI, refer to iSCSI Optimization

To verify the DCBx configuration on a port, use the `show interface DCBx detail` command.

Configuring DCBx Globally on the Switch

To globally configure the DCBx operation on a switch, follow these steps.

1. Enter Global Configuration mode.
   EXEC PRIVILEGE mode
   `configure`
2. Enter LLDP Configuration mode to enable DCBx operation.
   CONFIGURATION mode
   `[no] protocol lldp`
3. Configure the DCBx version used on all interfaces not already configured to exchange DCB information.
   PROTOCOL LLDP mode
   `[no] DCBx version {auto | cee | cin | ieee-v2.5}`

- `auto`: configures all ports to operate using the DCBx version received from a peer.
- `cee`: configures a port to use CEE (Intel 1.01). cin configures a port to use Cisco-Intel-Nuova (DCBx 1.0).
- `ieee-v2.5`: configures a port to use IEEE 802.1Qaz (Draft 2.5).

The default is **Auto**.

NOTE: To configure the DCBx port role the interfaces use to exchange DCB information, use the `DCBx port-role` command in INTERFACE Configuration mode (Step 3).

4. Configure the PFC and ETS TLVs that advertise on unconfigured interfaces with a manual port-role.
   PROTOCOL LLDP mode
   `[no] advertise DCBx-tlv {ets-conf | ets-reco | pfc} [ets-conf | ets-reco | pfc] [ets-conf | ets-reco | pfc]`
   - `ets-conf`: enables transmission of ETS Configuration TLVs.
   - `ets-reco`: enables transmission of ETS Recommend TLVs.
   - `pfc`: enables transmission of PFC TLVs.

   NOTE: You can configure the transmission of more than one TLV type at a time. You can only enable ETS recommend TLVs (`ets-reco`) if you enable ETS configuration TLVs (`ets-conf`). To disable TLV transmission, use the `no` form of the command; for example, `no advertise DCBx-tlv pfc ets-reco`.

   The default is All TLV types are enabled.
5. Configure the Application Priority TLVs that advertise on unconfigured interfaces with a manual port-role.
   PROTOCOL LLDP mode
   `[no] advertise DCBx-appln-tlv {fcoe | iscsi}`
   - `fcoe`: enables the advertisement of FCoE in Application Priority TLVs.
   - `iscsi`: enables the advertisement of iSCSI in Application Priority TLVs.

   The default is Application Priority TLVs are enabled and advertise FCoE and iSCSI.

   NOTE: To disable TLV transmission, use the `no` form of the command; for example, `no advertise DCBx-appln-tlv iscsi`.
6. Configure the FCoE priority advertised for the FCoE protocol in Application Priority TLVs.
   PROTOCOL LLDP mode
   `[no] fcoe priority-bits` *`priority-bitmap`*
   The priority-bitmap range is from 1 to FF.
   The default is **0x8**.
7. Configure the iSCSI priority advertised for the iSCSI protocol in Application Priority TLVs.
   PROTOCOL LLDP mode
   `[no] iscsi priority-bits` *`priority-bitmap`*
   The priority-bitmap range is from 1 to FF.
   The default is **0x10**.

DCBx Error Messages

The following syslog messages appear when an error in DCBx operation occurs.

```
LLDP_MULTIPLE_PEER_DETECTED: DCBx is operationally disabled after detecting more than
one DCBx
peer on the port interface.
```

```
LLDP_PEER_AGE_OUT: DCBx is disabled as a result of LLDP timing out on a DCBx peer
interface.
```

```
DSM_DCBx_PEER_VERSION_CONFLICT: A local port expected to receive the IEEE, CIN, or CEE
version
in a DCBx TLV from a remote peer but received a different, conflicting DCBx version.
```

```
DSM_DCBx_PFC_PARAMETERS_MATCH and DSM_DCBx_PFC_PARAMETERS_MISMATCH: A local DCBx port
received
a compatible (match) or incompatible (mismatch) PFC configuration from a peer.
```

```
DSM_DCBx_ETS_PARAMETERS_MATCH and DSM_DCBx_ETS_PARAMETERS_MISMATCH: A local DCBx port
received
a compatible (match) or incompatible (mismatch) ETS configuration from a peer.
```

```
LLDP_UNRECOGNISED_DCBx_TLV_RECEIVED: A local DCBx port received an unrecognized DCBx TLV
from
a peer.
```

Debugging DCBx on an Interface

To enable DCBx debug traces for all or a specific control paths, use the following command.

- Enable DCBx debugging.
  EXEC PRIVILEGE mode
  `debug DCBx {all | auto-detect-timer | config-exchng | fail | mgmt | resource | sem | tlv}`
  - `all`: enables all DCBx debugging operations.
  - `auto-detect-timer`: enables traces for DCBx auto-detect timers.
  - `config-exchng`: enables traces for DCBx configuration exchanges.
  - `fail`: enables traces for DCBx failures.
  - `mgmt`: enables traces for DCBx management frames.
  - `resource`: enables traces for DCBx system resource frames.
  - `sem`: enables traces for the DCBx state machine.
  - `tlv`: enables traces for DCBx TLVs.

Verifying the DCB Configuration

To display DCB configurations, use the following `show` commands.

Table 20. Displaying DCB Configurations

Command	Output
`show qos dot1p-queue mapping`	Displays the current 802.1p priority-queue mapping.

Table 20. Displaying DCB Configurations (continued)

Command	Output
`show dcb [stack-unit unit-number]`	Displays the data center bridging status, number of PFC-enabled ports, and number of PFC-enabled queues. On the master switch in a stack, you can specify a stack-unit number. The range is from 0 to 5.
`show qos priority-groups`	Displays the ETS priority groups configured on the switch, including the 802.1p priority classes and ID of each group.
`show interface port-type pfc {summary | detail}`	Displays the PFC configuration applied to ingress traffic on an interface, including priorities and link delay. To clear PFC TLV counters, use the `clear pfc counters interface port-type slot/port` command.
`show interface port-type pfc statistics`	Displays counters for the PFC frames received and transmitted (by dot1p priority class) on an interface. You can use the `show interface pfc statistics` command even without enabling DCB on the system.
`show interface port-type ets {summary | detail}`	Displays the ETS configuration applied to egress traffic on an interface, including priority groups with priorities and bandwidth allocation. To clear ETS TLV counters, enter the `clear ets counters interface port-type slot/port` command.
`show interface port-type DCBx detail`	Plays the DCBx configuration on an interface.
`show stack-unit {0-11 | all} stack ports all pfc details`	Displays the PFC configuration applied to ingress traffic, including priorities and link delay.
`show stack-unit {0-11 | all} stack ports all ets details`	Displays the ETS configuration applied to ingress traffic on stack-links, including priorities and link delay.

The following example shows the `show dot1p-queue mapping` command.

```
DellEMC(conf)# show qos dot1p-queue-mapping
Dot1p Priority: 0 1 2 3 4 5 6 7
Queue         : 0 0 0 1 2 3 3 3
```

The following example shows the `show dcb` command.

```
DellEMC# show dcb
stack-unit 0 port-set 0
     DCB Status : Enabled
 PFC Port Count : 56 (current), 56 (configured)
PFC Queue Count : 2 (current),  2 (configured)
```

The following example shows the `show qos priority-groups` command.

```
DellEMC#show qos priority-groups
priority-group ipc
  priority-list 4
  set-pgid 2
```

The following example shows the output of the `show qos dcb-map test` command.

```
DellEMC#show qos dcb-map test
-----------------------

State   :Complete
PfcMode:ON
--------------------
```

```
PG:0 TSA:ETS  BW:50  PFC:OFF
Priorities:0 1 2 5 6 7


PG:1 TSA:ETS  BW:50  PFC:ON
Priorities:3 4
```

The following example shows the `show interfaces pfc summary` command.

```
DellEMC# show interfaces tengigabitethernet 1/4 pfc summary
Interface TenGigabitEthernet 1/4
    Admin mode is on
    Admin is enabled
    Remote is enabled, Priority list is 4
    Remote Willing Status is enabled
    Local is enabled
    Oper status is Recommended
    PFC DCBx Oper status is Up
    State Machine Type is Feature
    TLV Tx Status is enabled
    PFC Link Delay 45556 pause quantams
    Application Priority TLV Parameters :
    ------------------------------------
    FCOE TLV Tx Status is disabled
    ISCSI TLV Tx Status is disabled
    Local FCOE PriorityMap is 0x8
    Local ISCSI PriorityMap is 0x10
    Remote FCOE PriorityMap is 0x8
    Remote ISCSI PriorityMap is 0x8

DellEMC# show interfaces tengigabitethernet 1/4 pfc detail
Interface TenGigabitEthernet 1/4
    Admin mode is on
    Admin is enabled
    Remote is enabled
    Remote Willing Status is enabled
    Local is enabled
    Oper status is recommended
    PFC DCBx Oper status is Up
    State Machine Type is Feature
    TLV Tx Status is enabled
    PFC Link Delay 45556 pause quanta
    Application Priority TLV Parameters :
    ------------------------------------
    FCOE TLV Tx Status is disabled
    ISCSI TLV Tx Status is disabled
    Local FCOE PriorityMap is 0x8
    Local ISCSI PriorityMap is 0x10
    Remote FCOE PriorityMap is 0x8
    Remote ISCSI PriorityMap is 0x8

0 Input TLV pkts, 1 Output TLV pkts, 0 Error pkts, 0 Pause Tx pkts, 0 Pause Rx pkts
```

The following table describes the `show interface pfc summary` command fields.

Table 21. `show interface pfc summary` Command Description

Fields	Description
Interface	Interface type with stack-unit and port number.
Admin mode is on; Admin is enabled	PFC Admin mode is on or off with a list of the configured PFC priorities . When PFC admin mode is on, PFC advertisements are enabled to be sent and received from peers; received PFC configuration takes effect. The admin operational status for a DCBx exchange of PFC configuration is enabled or disabled.
Remote is enabled; Priority list Remote Willing Status is enabled	Operational status (enabled or disabled) of peer device for DCBx exchange of PFC configuration with a list of the configured PFC priorities. Willing status of peer device for

Table 21. `show interface pfc summary` Command Description (continued)

Fields	Description
	DCBx exchange (Willing bit received in PFC TLV): enabled or disabled.
Local is enabled	DCBx operational status (enabled or disabled) with a list of the configured PFC priorities
Operational status (local port)	DCBx operational status (enabled or disabled) with a list of the configured PFC priorities.
	Port state for current operational PFC configuration: • Init: Local PFC configuration parameters were exchanged with peer. • Recommend: Remote PFC configuration parameters were received from peer. • Internally propagated: PFC configuration parameters were received from configuration source.
PFC DCBx Oper status	Operational status for exchange of PFC configuration on local port: match (up) or mismatch (down).
State Machine Type	Type of state machine used for DCBx exchanges of PFC parameters: • Feature: for legacy DCBx versions • Symmetric: for an IEEE version
TLV Tx Status	Status of PFC TLV advertisements: enabled or disabled.
PFC Link Delay	Link delay (in quanta) used to pause specified priority traffic.
Application Priority TLV: FCOE TLV Tx Status	Status of FCoE advertisements in application priority TLVs from local DCBx port: enabled or disabled.
Application Priority TLV: ISCSI TLV Tx Status	Status of ISCSI advertisements in application priority TLVs from local DCBx port: enabled or disabled.
Application Priority TLV: Local FCOE Priority Map	Priority bitmap used by local DCBx port in FCoE advertisements in application priority TLVs.
Application Priority TLV: Local ISCSI Priority Map	Priority bitmap used by local DCBx port in ISCSI advertisements in application priority TLVs.
Application Priority TLV: Remote FCOE Priority Map	Status of FCoE advertisements in application priority TLVs from remote peer port: enabled or disabled.
Application Priority TLV: Remote ISCSI Priority Map	Status of iSCSI advertisements in application priority TLVs from remote peer port: enabled or disabled.
PFC TLV Statistics: Input TLV pkts	Number of PFC TLVs received.
PFC TLV Statistics: Output TLV pkts	Number of PFC TLVs transmitted.
PFC TLV Statistics: Error pkts	Number of PFC error packets received.
PFC TLV Statistics: Pause Tx pkts	Number of PFC pause frames transmitted.
PFC TLV Statistics: Pause Rx pkts	Number of PFC pause frames received

The following example shows the `show interface pfc statistics` command.

```
DellEMC#show interfaces te 1/1 pfc statistics
Interface TenGigabitEthernet 1/1
Interface       Priority        Rx XOFF Frames          Rx Total Frames         Tx Total
Frames
-------------------------------------------------------------------------------------
------
Te 1/1          P0              0                       0                       0
Te 1/1          P1              0                       0                       0
Te 1/1          P2              0                       0                       0
```

```
Te 1/1          P3              0                       0                       0
Te 1/1          P4              0                       0                       0
Te 1/1          P5              0                       0                       0
Te 1/1          P6              0                       0                       0
Te 1/1          P7              0                       0                       0
```

The following example shows the `show interface ets summary` command.

```
DellEMC(conf)#do show interfaces te 1/1 ets summary

Interface TenGigabitEthernet 1/1
Max Supported TC is 4
Number of Traffic Classes is 8
Admin mode is on

Admin Parameters :
------------------
Admin is enabled

PG-grp     Priority#        BW-%      BW-COMMITTED           BW-PEAK          TSA
                             %     Rate(Mbps) Burst(KB) Rate(Mbps) Burst(KB)
-------------------------------------------------------------------------------
0          3                25    -          -         -          -         ETS
1          4                25    -          -         -          -         ETS
2          0,1,2,5,6,7      50    -          -         -          -         ETS
3                           -     -          -         -          -         -
4                           -     -          -         -          -         -
5                           -     -          -         -          -         -
6                           -     -          -         -          -         -
7                           -     -          -         -          -         -

Remote Parameters :
-------------------
Remote is disabled

Local Parameters :
------------------
Local is enabled

PG-grp     Priority#        BW-%      BW-COMMITTED           BW-PEAK          TSA
                             %     Rate(Mbps) Burst(KB) Rate(Mbps) Burst(KB)
-------------------------------------------------------------------------------
0          3                25    -          -         -          -         ETS
1          4                25    -          -         -          -         ETS
2          0,1,2,5,6,7      50    -          -         -          -         ETS
3                           -     -          -         -          -         -
4                           -     -          -         -          -         -
5                           -     -          -         -          -         -
6                           -     -          -         -          -         -
7                           -     -          -         -          -         -

Oper status is init
ETS DCBX Oper status is Down
Reason: Port Shutdown
State Machine Type is Asymmetric
Conf TLV Tx Status is enabled
Reco TLV Tx Status is enabled
```

The following example shows the `show interface ets detail` command.

```
DellEMC(conf)# show interfaces tengigabitethernet 1/1 ets detail
Interface TenGigabitEthernet 1/1
Max Supported TC Groups is 4
Number of Traffic Classes is 8
Admin mode is on
Admin Parameters :
------------------
Admin is enabled
TC-grp    Priority#         Bandwidth    TSA
0         0,1,2,3,4,5,6,7   100%         ETS
1                           0%           ETS
2                           0%           ETS
```

```
3                               0%              ETS
4                               0%              ETS
5                               0%              ETS
6                               0%              ETS
7                               0%              ETS

Priority# Bandwidth TSA
0                               13%             ETS
1                               13%             ETS
2                               13%             ETS
3                               13%             ETS
4                               12%             ETS
5                               12%             ETS
6                               12%             ETS
7                               12%             ETS
Remote Parameters:
-------------------
Remote is disabled

Local Parameters :
------------------
Local is enabled
TC-grp    Priority#             Bandwidth     TSA
0         0,1,2,3,4,5,6,7       100%          ETS
1                               0%            ETS
2                               0%            ETS
3                               0%            ETS
4                               0%            ETS
5                               0%            ETS
6                               0%            ETS
7                               0%            ETS

Priority#                       Bandwidth     TSA
0                               13%           ETS
1                               13%           ETS
2                               13%           ETS
3                               13%           ETS
4                               12%           ETS
5                               12%           ETS
6                               12%           ETS
7                               12%           ETS
Oper status is init
Conf TLV Tx Status is disabled
Traffic Class TLV Tx Status is disabled
0 Input Conf TLV Pkts, 0 Output Conf TLV Pkts, 0 Error Conf TLV Pkts
0 Input Traffic Class TLV Pkts, 0 Output Traffic Class TLV Pkts, 0 Error Traffic Class
TLV
Pkts
```

The following table describes the `show interface ets detail` command fields.

Table 22. `show interface ets detail` Command Description

Field	Description
Interface	Interface type with stack-unit and port number.
Maximum Supported TC Group	Maximum number of priority groups supported.
Number of Traffic Classes	Number of 802.1p priorities currently configured.
Admin mode	ETS mode: on or off.
Admin Parameters	ETS configuration on local port, including priority groups, assigned dot1p priorities, and bandwidth allocation.
Remote Parameters	ETS configuration on remote peer port, including Admin mode (enabled if a valid TLV was received or disabled), priority groups, assigned dot1p priorities, and bandwidth allocation. If the ETS Admin mode is enabled on the remote port for DCBx exchange, the Willing bit received in ETS TLVs from the remote peer is included.

Table 22. `show interface ets detail` Command Description (continued)

Field	Description
Local Parameters	ETS configuration on local port, including Admin mode (enabled when a valid TLV is received from a peer), priority groups, assigned dot1p priorities, and bandwidth allocation.
Operational status (local port)	Port state for current operational ETS configuration: • Init: Local ETS configuration parameters were exchanged with peer. • Recommend: Remote ETS configuration parameters were received from peer. • Internally propagated: ETS configuration parameters were received from configuration source.
ETS DCBx Oper status	Operational status of ETS configuration on local port: match or mismatch.
State Machine Type	Type of state machine used for DCBx exchanges of ETS parameters: • Feature: for legacy DCBx versions • Asymmetric: for an IEEE version
Conf TLV Tx Status	Status of ETS Configuration TLV advertisements: enabled or disabled.
ETS TLV Statistic: Input Conf TLV pkts	Number of ETS Configuration TLVs received.
ETS TLV Statistic: Output Conf TLV pkts	Number of ETS Configuration TLVs transmitted.
ETS TLV Statistic: Error Conf TLV pkts	Number of ETS Error Configuration TLVs received.

The following example shows the `show stack-unit all stack-ports all pfc details` command.

```
DellEMC(conf)# show stack-unit all stack-ports all pfc details

stack unit 0 stack-port all
    Admin mode is On
    Admin is enabled, Priority list is 4-5
    Local is enabled, Priority list is 4-5
    Link Delay 45556 pause quantum
    0 Pause Tx pkts, 0 Pause Rx pkts

stack unit 1 stack-port all
    Admin mode is On
    Admin is enabled, Priority list is 4-5
    Local is enabled, Priority list is 4-5
    Link Delay 45556 pause quantum
    0 Pause Tx pkts, 0 Pause Rx pkts
```

The following example shows the `show stack-unit all stack-ports all ets details` command.

```
DellEMC(conf)# show stack-unit all stack-ports all ets details
Stack unit 0 stack port all
Max Supported TC Groups is 4
Number of Traffic Classes is 1

Admin mode is on
Admin Parameters:
--------------------
Admin is enabled
TC-grp    Priority#        Bandwidth    TSA
-----------------------------------------------
0         0,1,2,3,4,5,6,7  100%         ETS
1                          -            -
2                          -            -
3                          -            -
4                          -            -
5                          -            -
```

```
6                           -           -
7                           -           -
8                           -           -

Stack unit 1 stack port all
Max Supported TC Groups is 4
Number of Traffic Classes is 1
Admin mode is on
Admin Parameters:
--------------------
Admin is enabled
TC-grp    Priority#        Bandwidth    TSA
-----------------------------------------------
0         0,1,2,3,4,5,6,7  100%         ETS
1                          -            -
2                          -            -
3                          -            -
4                          -            -
5                          -            -
6                          -            -
7                          -            -
8                          -            -
```

The following example shows the `show interface DCBx detail` command (IEEE).

```
DellEMC(conf-if-te-1/17-lldp)#do sho int te 2/12 dc d

E-ETS Configuration TLV enabled             e-ETS Configuration TLV disabled
R-ETS Recommendation TLV enabled            r-ETS Recommendation TLV disabled
P-PFC Configuration TLV enabled             p-PFC Configuration TLV disabled
F-Application priority for FCOE enabled     f-Application Priority for FCOE disabled
I-Application priority for iSCSI enabled    i-Application Priority for iSCSI disabled
-------------------------------------------------------------------------------------
-

Interface TenGigabitEthernet 2/12
  Remote Mac Address 00:01:e8:8a:df:a0
    Port Role is Manual
  DCBx Operational Status is Enabled
  Is Configuration Source? FALSE
  Local DCBx Compatibility mode is IEEEv2.5
  Local DCBx Configured mode is IEEEv2.5
  Peer Operating version is IEEEv2.5
  Local DCBx TLVs Transmitted: ERPFi
  1 Input PFC TLV pkts, 2 Output PFC TLV pkts, 0 Error PFC pkts
  0 PFC Pause Tx pkts, 0 Pause Rx pkts
  1 Input ETS Conf TLV Pkts, 1 Output ETS Conf TLV Pkts, 0 Error ETS Conf TLV Pkts
  1 Input ETS Reco TLV pkts, 1 Output ETS Reco TLV pkts, 0 Error ETS Reco TLV Pkts
```

The following example shows the `show interface DCBx detail` command (legacy CEE).

```
DellEMC(conf-if-te-1/17-lldp)#do sho int te 1/14 dc d

E-ETS Configuration TLV enabled             e-ETS Configuration TLV disabled
R-ETS Recommendation TLV enabled            r-ETS Recommendation TLV disabled
P-PFC Configuration TLV enabled             p-PFC Configuration TLV disabled
F-Application priority for FCOE enabled     f-Application Priority for FCOE disabled
I-Application priority for iSCSI enabled    i-Application Priority for iSCSI disabled
-------------------------------------------------------------------------

Interface TenGigabitEthernet 1/14
  Remote Mac Address 00:01:e8:8a:df:a0
    Port Role is Auto-Upstream
  DCBx Operational Status is Enabled
  Is Configuration Source? FALSE
  Local DCBx Compatibility mode is CEE
  Local DCBx Configured mode is CEE
  Peer Operating version is CEE
  Local DCBx TLVs Transmitted: ErPFi

Local DCBx Status
-----------------
```

```
   DCBx Operational Version is 0
   DCBx Max Version Supported is 0
   Sequence Number: 1
   Acknowledgment Number: 1
   Protocol State: In-Sync

 Peer DCBx Status:
 ----------------
   DCBx Operational Version is 0
   DCBx Max Version Supported is 0
   Sequence Number: 1
   Acknowledgment Number: 1
   Total DCBx Frames transmitted 994
   Total DCBx Frames received 646
   Total DCBx Frame errors 0
   Total DCBx Frames unrecognized 0
```

The following table describes the `show interface DCBx detail` command fields.

Table 23. `show interface DCBx detail` Command Description

Field	Description
Interface	Interface type with chassis slot and port number.
Port-Role	Configured DCBx port role: auto-upstream, auto-downstream, config-source, or manual.
DCBx Operational Status	Operational status (enabled or disabled) used to elect a configuration source and internally propagate a DCB configuration. The DCBx operational status is the combination of PFC and ETS operational status.
Configuration Source	Specifies whether the port serves as the DCBx configuration source on the switch: true (yes) or false (no).
Local DCBx Compatibility mode	DCBx version accepted in a DCB configuration as compatible. In auto-upstream mode, a port can only received a DCBx version supported on the remote peer.
Local DCBx Configured mode	DCBx version configured on the port: CEE, CIN, IEEE v2.5, or Auto (port auto-configures to use the DCBx version received from a peer).
Peer Operating version	DCBx version that the peer uses to exchange DCB parameters.
Local DCBx TLVs Transmitted	Transmission status (enabled or disabled) of advertised DCB TLVs (see TLV code at the top of the show command output).
Local DCBx Status: DCBx Operational Version	DCBx version advertised in Control TLVs.
Local DCBx Status: DCBx Max Version Supported	Highest DCBx version supported in Control TLVs.
Local DCBx Status: Sequence Number	Sequence number transmitted in Control TLVs.
Local DCBx Status: Acknowledgment Number	Acknowledgement number transmitted in Control TLVs.
Local DCBx Status: Protocol State	Current operational state of DCBx protocol: ACK or IN-SYNC.
Peer DCBx Status: DCBx Operational Version	DCBx version advertised in Control TLVs received from peer device.
Peer DCBx Status: DCBx Max Version Supported	Highest DCBx version supported in Control TLVs received from peer device.
Peer DCBx Status: Sequence Number	Sequence number transmitted in Control TLVs received from peer device.
Peer DCBx Status: Acknowledgment Number	Acknowledgement number transmitted in Control TLVs received from peer device.

Table 23. `show interface DCBx detail` Command Description (continued)

Field	Description
Total DCBx Frames transmitted	Number of DCBx frames sent from local port.
Total DCBx Frames received	Number of DCBx frames received from remote peer port.
Total DCBx Frame errors	Number of DCBx frames with errors received.
Total DCBx Frames unrecognized	Number of unrecognizable DCBx frames received.

Sample DCB Configuration

The following shows examples of using PFC and ETS to manage your data center traffic.

In the following example:

- Incoming SAN traffic is configured for priority-based flow control.
- Outbound LAN, IPC, and SAN traffic is mapped into three ETS priority groups and configured for enhanced traffic selection (bandwidth allocation and scheduling).
- One lossless queue is used.

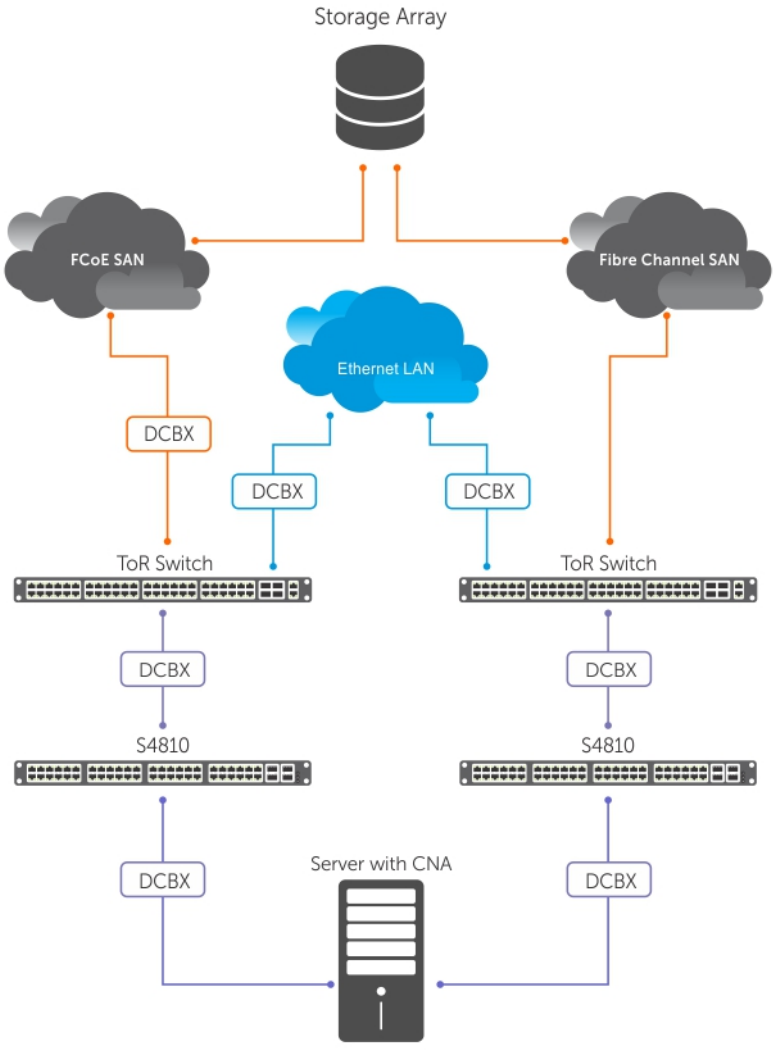


Figure 33. PFC and ETS Applied to LAN, IPC, and SAN Priority Traffic

QoS Traffic Classification: The `service-class dynamic dot1p` command has been used in Global Configuration mode to map ingress dot1p frames to the queues shown in the following table. For more information, refer to QoS dot1p Traffic Classification and Queue Assignment.

The following describes the dot1p-priority class group assignment

dot1p Value in the Incoming Frame	Priority Group Assignment
0	LAN
1	LAN
2	LAN
3	SAN
4	IPC
5	LAN
6	LAN
7	LAN

The following describes the priority group-bandwidth assignment.

Priority Group	Bandwidth Assignment
IPC	5%
SAN	50%
LAN	45%

PFC and ETS Configuration Command Examples

The following examples show PFC and ETS configuration commands to manage your data center traffic.

1. **Enabling DCB**

```
DellEMC(conf)#dcb enable
```

2. **Configure DCB map and enable PFC, and ETS**

```
DellEMC(conf)# service-class dynamic dot1p

Or

DellEMC(conf)# interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)# service-class dynamic dot1p
```

3. **Apply DCB map to relevant interface**

```
dcb-map test
priority-group 1 bandwidth 50 pfc on
priority-group 2 bandwidth 45 pfc off
priority group 3 bandwidth 5 pfc  on
priority-pgid 2 2 2 1 3 2 2 2
```

Example of Applying DCB Map to an Interface

```
DellEMC(conf)# int tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)# dcb-map test
```

QoS dot1p Traffic Classification and Queue Assignment

The following section describes QoS dot1P traffic classification and assignments.

DCB supports PFC, ETS, and DCBx to handle converged Ethernet traffic that is assigned to an egress queue according to the following QoS methods:

Honor dot1p	You can honor dot1p priorities in ingress traffic at the port or global switch level (refer to Default dot1p to Queue Mapping) using the `service-class dynamic dot1p` command in INTERFACE configuration mode.
Layer 2 class maps	You can use dot1p priorities to classify traffic in a class map and apply a service policy to an ingress port to map traffic to egress queues.

NOTE: Dell EMC Networking does not recommend mapping all ingress traffic to a single queue when using PFC and ETS. However, Dell EMC Networking does recommend using Ingress traffic classification using the `service-class dynamic dot1p` command (honor dot1p) on all DCB-enabled interfaces. If you use L2 class maps to map dot1p priority traffic to egress queues, take into account the default dot1p-queue assignments in the following table and the maximum number of two lossless queues supported on a port (refer to Configuring Lossless Queues).

Although Dell EMC Networking OS allows you to change the default dot1p priority-queue assignments (refer to Setting dot1p Priorities for Incoming Traffic), DCB policies applied to an interface may become invalid if you reconfigure dot1p-queue mapping. If the configured DCB policy remains valid, the change in the dot1p-queue assignment is allowed.

NOTE: Although, each port on the S4810, S4820T, and S5000 devices support 8 QoS queues, you can configure only 4 QoS queues (0-3) to manage data traffic. The remaining 4 queues (4-7) are reserved for control traffic.

dot1p Value in the Incoming Frame	Egress Queue Assignment
0	0
1	0
2	0
3	1
4	2
5	3
6	3
7	3

dot1p Value in the Incoming Frame	Egress Queue Assignment
0	2
1	0
2	1
3	3
4	4
5	5
6	6
7	7

Configuring the Dynamic Buffer Method

Priority-based flow control using dynamic buffer spaces is supported on the switch.

To configure the dynamic buffer capability, perform the following steps:

1. Enable the DCB application. By default, DCB is enabled and link-level flow control is disabled on all interfaces.
   CONFIGURATION mode
   ```
   dcb enable
   ```
2. Configure the shared PFC buffer size and the total buffer size. A maximum of 4 lossless queues are supported.
   CONFIGURATION mode
   ```
   dcb pfc-shared-buffer-size value
   dcb pfc-total-buffer-size value
   ```
   The buffer size range is from 0 to 3399. Default is 3088.
3. Configure the number of PFC queues.
   CONFIGURATION mode
   ```
   dcb enable pfc-queues pfc-queues
   ```
   The number of ports supported based on lossless queues configured depends on the buffer. The default number of PFC queues in the system is two.

   For each priority, you can specify the shared buffer threshold limit, the ingress buffer size, buffer limit for pausing the acceptance of packets, and the buffer offset limit for resuming the acceptance of received packets.
4. Configure the profile name for the DCB buffer threshold
   CONFIGURATION mode
   ```
   dcb-buffer-threshold dcb-buffer-threshold
   ```
5. DCB-BUFFER-THRESHOLD mode
   ```
   priority 0 buffer-size 52 pause-threshold 16 resume-offset 10 shared-threshold-weight 7
   ```
6. Assign the DCB policy to the DCB buffer threshold profile on stack ports.
   CONFIGURATION mode
   ```
   DellEMC(conf)# dcb-policy buffer-threshold stack-unit all stack-ports all dcb-policy-name
   ```
7. Assign the DCB policy to the DCB buffer threshold profile on interfaces. This setting takes precedence over the default buffer-threshold setting.
   INTERFACE mode
   ```
   dcb-policy buffer-threshold dcb-buffer-threshold
   ```
8. Configuring Global total buffer size on stack ports.
   CONFIGURATION mode
   ```
   dcb pfc-total-buffer-size buffer-size stack-unit all port-set {port-pipe |all}
   ```
   Port-set number range is from 0 to 3.

13

Dynamic Host Configuration Protocol (DHCP)

DHCP is an application layer protocol that dynamically assigns IP addresses and other configuration parameters to network end-stations (hosts) based on configuration policies determined by network administrators.

DHCP relieves network administrators of manually configuring hosts, which can be a tedious and error-prone process when hosts often join, leave, and change locations on the network and it reclaims IP addresses that are no longer in use to prevent address exhaustion.

DHCP is based on a client-server model. A host discovers the DHCP server and requests an IP address, and the server either leases or permanently assigns one. There are three types of devices that are involved in DHCP negotiation:

DHCP Server	This is a network device offering configuration parameters to the client.
DHCP Client	This is a network device requesting configuration parameters from the server.
Relay Agent	This is an intermediary network device that passes DHCP messages between the client and server when the server is not on the same subnet as the host.

Topics:

- DHCP Packet Format and Options
- Assign an IP Address using DHCP
- Implementation Information
- Configure the System to be a DHCP Server
- Configure the System to be a Relay Agent
- Configure the System to be a DHCP Client
- Configure the System for User Port Stacking (Option 230)
- Configuring DHCP relay source interface
- Configure Secure DHCP
- Source Address Validation

DHCP Packet Format and Options

DHCP uses the user datagram protocol (UDP) as its transport protocol.

The server listens on port 67 and transmits to port 68; the client listens on port 68 and transmits to port 67. The configuration parameters are carried as options in the DHCP packet in Type, Length, Value (TLV) format; many options are specified in RFC 2132. To limit the number of parameters that servers must provide, hosts specify the parameters that they require, and the server sends only those parameters. Some common options are shown in the following illustration.

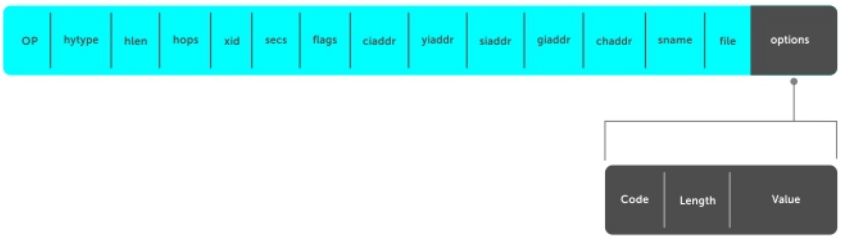


Figure 34. DHCP packet Format

The following table lists common DHCP options.

Option	Number and Description
Subnet Mask	Option 1 Specifies the client's subnet mask.
Router	Option 3 Specifies the router IP addresses that may serve as the client's default gateway.
Domain Name Server	Option 6 Specifies the domain name servers (DNSs) that are available to the client.
Domain Name	Option 15 Specifies the domain name that clients should use when resolving hostnames via DNS.
IP Address Lease Time	Option 51 Specifies the amount of time that the client is allowed to use an assigned IP address.
DHCP Message Type	Option 53 ● 1: DHCPDISCOVER ● 2: DHCPOFFER ● 3: DHCPREQUEST ● 4: DHCPDECLINE ● 5: DHCPACK ● 6: DHCPNACK ● 7: DHCPRELEASE ● 8: DHCPINFORM
Parameter Request List	Option 55 Clients use this option to tell the server which parameters it requires. It is a series of octets where each octet is DHCP option code.
Renewal Time	Option 58 Specifies the amount of time after the IP address is granted that the client attempts to renew its lease with the *original* server.
Rebinding Time	Option 59 Specifies the amount of time after the IP address is granted that the client attempts to renew its lease with *any* server, if the original server does not respond.
Vendor Class Identifer	Option 60 Identifies a user-defined string used by the Relay Agent to forward DHCP client packets to a specific server.
L2 DHCP Snooping	Option 82 Specifies IP addresses for DHCP messages received from the client that are to be monitored to build a DHCP snooping database.
User Port Stacking	Option 230 Set the stacking option variable to provide DHCP server stack-port detail when the DHCP offer is set.
End	Option 255 Signals the last option in the DHCP packet.

Assign an IP Address using DHCP

The following section describes DHCP and the client in a network.

When a client joins a network:

1. The client initially broadcasts a **DHCPDISCOVER** message on the subnet to discover available DHCP servers. This message includes the parameters that the client requires and might include suggested values for those parameters.
2. Servers unicast or broadcast a **DHCPOFFER** message in response to the DHCPDISCOVER that offers to the client values for the requested parameters. Multiple servers might respond to a single DHCPDISCOVER; the client might wait a period of time and then act on the most preferred offer.
3. The client broadcasts a **DHCPREQUEST** message in response to the offer, requesting the offered values.
4. After receiving a DHCPREQUEST, the server binds the clients' unique identifier (the hardware address plus IP address) to the accepted configuration parameters and stores the data in a database called a binding table. The server then broadcasts a **DHCPACK** message, which signals to the client that it may begin using the assigned parameters.
5. When the client leaves the network, or the lease time expires, returns its IP address to the server in a **DHCPRELEASE** message.

There are additional messages that are used in case the DHCP negotiation deviates from the process previously described and shown in the illustration below.

DHCPDECLINE	A client sends this message to the server in response to a DHCPACK if the configuration parameters are unacceptable; for example, if the offered address is already in use. In this case, the client starts the configuration process over by sending a DHCPDISCOVER.
DHCPINFORM	A client uses this message to request configuration parameters when it assigned an IP address manually rather than with DHCP. The server responds by unicast.
DHCPNAK	A server sends this message to the client if it is not able to fulfill a DHCPREQUEST; for example, if the requested address is already in use. In this case, the client starts the configuration process over by sending a DHCPDISCOVER.

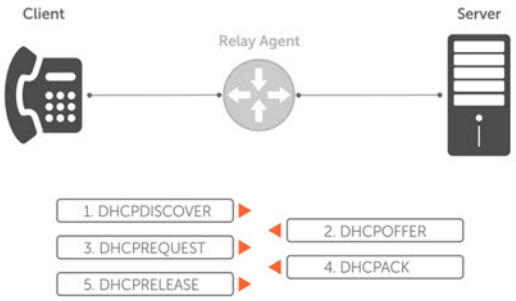


Figure 35. Client and Server Messaging

Implementation Information

The following describes DHCP implementation.

- Dell EMC Networking implements DHCP based on RFC 2131 and RFC 3046.
- IP source address validation is a sub-feature of DHCP Snooping; the Dell EMC Networking OS uses access control lists (ACLs) internally to implement this feature and as such, you cannot apply ACLs to an interface which has IP source address validation. If you configure IP source address validation on a member port of a virtual local area network (VLAN) and then to apply an access list to the VLAN, Dell EMC Networking OS displays the first line in the following message. If you first apply an ACL to a VLAN and then enable IP source address validation on one of its member ports, Dell EMC Networking OS displays the second line in the following message.

```
% Error: Vlan member has access-list configured.
% Error: Vlan has an access-list configured.
```

  NOTE: If you enable DHCP Snooping globally and you have any configured L2 ports, any IP ACL, MAC ACL, or DHCP source address validation ACL does not block DHCP packets.
- Dell EMC Networking OS provides 40000 entries that can be divided between leased addresses and excluded addresses. By extension, the maximum number of pools you can configure depends on the subnet mask that you give to each pool. For example, if all pools were configured for a /24 mask, the total would be 40000/253 (approximately 158). If the subnet is increased, more pools can be configured. The maximum subnet that can be configured for a single pool is /17. Dell EMC Networking OS displays an error message for configurations that exceed the allocated memory.
- This platform supports 4000 DHCP Snooping entries.

- All platforms support Dynamic ARP Inspection on 16 VLANs per system. For more information, refer to Dynamic ARP Inspection.

  NOTE: If the DHCP server is on the top of rack (ToR) and the VLTi (ICL) is down due to a failed link, when a VLT node is rebooted in BMP (Bare Metal Provisioning) mode, it is not able to reach the DHCP server, resulting in BMP failure.

Configure the System to be a DHCP Server

A DHCP server is a network device that has been programmed to provide network configuration parameters to clients upon request. Servers typically serve many clients, making host management much more organized and efficient.

NOTE: If the management port is associated with any non-default VRF, then the `ip address dhcp` command does not work.

The following table lists the key responsibilities of DHCP servers.

Table 24. DHCP Server Responsibilities

DHCP Server Responsibilities	Description
Address Storage and Management	DHCP servers are the owners of the addresses used by DHCP clients.The server stores the addresses and manages their use, keeping track of which addresses have been allocated and which are still available.
Configuration Parameter Storage and Management	DHCP servers also store and maintain other parameters that are sent to clients when requested. These parameters specify in detail how a client is to operate.
Lease Management	DHCP servers use leases to allocate addresses to clients for a limited time. The DHCP server maintains information about each of the leases, including lease length.
Responding To Client Requests	DHCP servers respond to different types of requests from clients, primarily, granting, renewing, and terminating leases.
Providing Administration Services	DHCP servers include functionality that allows an administrator to implement policies that govern how DHCP performs its other tasks.

Configuring the Server for Automatic Address Allocation

Automatic address allocation is an address assignment method by which the DHCP server leases an IP address to a client from a pool of available addresses.

An address pool is a range of IP addresses that the DHCP server may assign. The subnet number indexes the address pools.

To create an address pool, follow these steps.

1. Access the DHCP server CLI context.

   CONFIGURATION mode

   ```
   ip dhcp server
   ```

2. Create an address pool and give it a name.

   DHCP mode

   ```
   pool name
   ```

3. Specify the range of IP addresses from which the DHCP server may assign addresses.

   DHCP <POOL> mode

   ```
   network network/prefix-length
   ```

   - `network`: the subnet address.
   - `prefix-length`: specifies the number of bits used for the network portion of the address you specify.

   The prefix-length range is from 17 to 31.

4. Display the current pool configuration.

DHCP <POOL> mode

```
show config
```

After an IP address is leased to a client, only that client may release the address. Dell EMC Networking OS performs a IP + MAC source address validation to ensure that no client can release another clients address. This validation is a default behavior and is separate from IP+MAC source address validation.

Configuration Tasks

To configure DHCP, an administrator must first set up a DHCP server and provide it with configuration parameters and policy information including IP address ranges, lease length specifications, and configuration data that DHCP hosts need.

Configuring the Dell system to be a DHCP server is a three-step process:

1. Configuring the Server for Automatic Address Allocation
2. Specifying a Default Gateway

Related Configuration Tasks

- Configure a Method of Hostname Resolution
- Creating Manual Binding Entries
- Debugging the DHCP Server
- Using DHCP Clear Commands

Excluding Addresses from the Address Pool

The DHCP server assumes that all IP addresses in a DHCP address pool are available for assigning to DHCP clients.

You must specify the IP address that the DHCP server should not assign to clients.

To exclude an address, follow this step.

- Exclude an address range from DHCP assignment. The exclusion applies to all configured pools.
  DHCP mode
  ```
  excluded-address
  ```

Specifying an Address Lease Time

To specify an address lease time, use the following command.

- Specify an address lease time for the addresses in a pool.
  DHCP <POOL>
  ```
  lease {days [hours] [minutes] | infinite}
  ```
  The default is **24 hours**.

Specifying a Default Gateway

The IP address of the default router should be on the same subnet as the client.

To specify a default gateway, follow this step.

- Specify default gateway(s) for the clients on the subnet, in order of preference.
  DHCP <POOL>
  ```
  default-router address
  ```

Configure a Method of Hostname Resolution

Dell systems are capable of providing DHCP clients with parameters for two methods of hostname resolution—using DNS or NetBIOS WINS.

Using DNS for Address Resolution

A domain is a group of networks. DHCP clients query DNS IP servers when they need to correlate host names to IP addresses.

1. Create a domain.
   DHCP <POOL>
   `domain-name` *name*
2. Specify in order of preference the DNS servers that are available to a DHCP client.
   DHCP <POOL>
   `dns-server` *address*

Using NetBIOS WINS for Address Resolution

Windows internet naming service (WINS) is a name resolution service that Microsoft DHCP clients use to correlate host names to IP addresses within a group of networks. Microsoft DHCP clients can be one of four types of NetBIOS nodes: broadcast, peer-to-peer, mixed, or hybrid.

1. Specify the NetBIOS WINS name servers, in order of preference, that are available to Microsoft Dynamic Host Configuration Protocol (DHCP) clients.
   DHCP <POOL> mode
   `netbios-name-server` *address*
2. Specify the NetBIOS node type for a Microsoft DHCP client. Dell EMC Networking recommends specifying clients as hybrid.
   DHCP <POOL> mode
   `netbios-node-type` *type*

Creating Manual Binding Entries

An address binding is a mapping between the IP address and the media access control (MAC) address of a client.

The DHCP server assigns the client an available IP address automatically, and then creates an entry in the binding table. However, the administrator can manually create an entry for a client; manual bindings are useful when you want to guarantee that a particular network device receives a particular IP address. Manual bindings can be considered single-host address pools. There is no limit on the number of manual bindings, but you can only configure one manual binding per host.

NOTE: Dell EMC Networking OS does not prevent you from using a network IP as a host IP; be sure to not use a network IP as a host IP.

1. Create an address pool.
   DHCP mode
   `pool` *name*
2. Specify the client IP address.
   DHCP <POOL>
   `host–address` *address*
3. Specify the client hardware address.
   DHCP <POOL>
   `hardware-address` *hardware-address type*
   - `hardware-address`: the client MAC address.
   - `type`: the protocol of the hardware platform.

   The default protocol is **Ethernet**.

Debugging the DHCP Server

To debug the DHCP server, use the following command.

- Display debug information for DHCP server.
  EXEC Privilege mode
  `debug ip dhcp server [events | packets]`

Using DHCP Clear Commands

To clear DHCP binding entries, address conflicts, and server counters, use the following commands.

- Clear DHCP binding entries for the entire binding table.
  EXEC Privilege mode.
  `clear ip dhcp binding`
- Clear a DHCP binding entry for an individual IP address.
  EXEC Privilege mode.
  `clear ip dhcp binding` *`ip address`*

Configure the System to be a Relay Agent

DHCP clients and servers request and offer configuration information via broadcast DHCP messages. Routers do not forward broadcasts, so if there are no DHCP servers on the subnet, the client does not receive a response to its request and therefore cannot access the network.

You can configure an interface on the Dell EMC Networking system to relay the DHCP messages to a specific DHCP server using the `ip helper-address dhcp-address` command from INTERFACE mode, as shown in the following illustration. Specify multiple DHCP servers by using the `ip helper-address dhcp-address` command multiple times.

When you configure the `ip helper-address` command, the system listens for DHCP broadcast messages on port 67. The system rewrites packets received from the client and forwards them via unicast to the DHCP servers; the system rewrites the destination IP address and writes its own address as the relay device. Responses from the server are unicast back to the relay agent on port 67 and the relay agent rewrites the destination address and forwards the packet to the client subnet via broadcast or unicast, depending whether the client has set or cleared the BROADCAST flag in the DHCP Client PDUs.

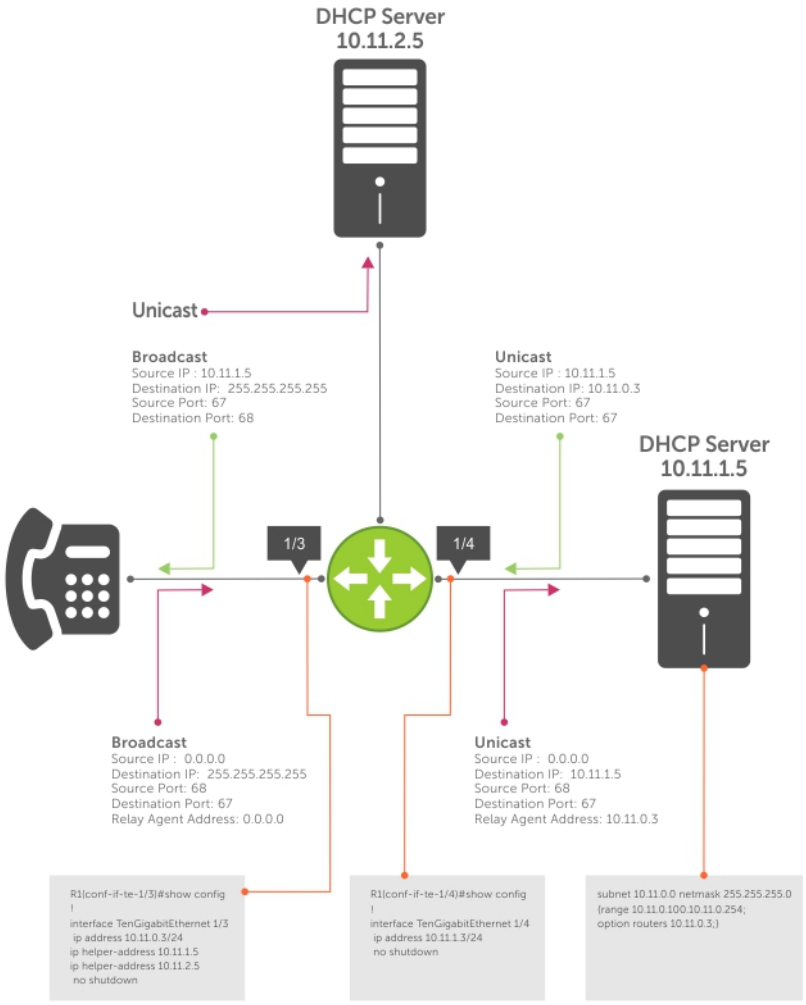


Figure 36. Configuring a Relay Agent

To view the `ip helper-address` configuration for an interface, use the `show ip interface` command from EXEC privilege mode.

Example of the `show ip interface` Command

```
R1_E600#show ip int tengigabitethernet 1/3
TenGigabitEthernet 1/3 is up, line protocol is down
Internet address is 10.11.0.1/24
Broadcast address is 10.11.0.255
Address determined by user input
IP MTU is 1500 bytes
Helper address is 192.168.0.1
                  192.168.0.2
Directed broadcast forwarding is disabled
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent
```

Configure the System to be a DHCP Client

A DHCP client is a network device that requests an IP address and configuration parameters from a DHCP server.

Implement the DHCP client functionality as follows:

- The switch can obtain a dynamically assigned IP address from a DHCP server. A start-up configuration is not received. Use bare metal provisioning (BMP) to receive configuration parameters (Dell EMC Networking OS version and a configuration file). BMP is enabled as a factory-default setting on a switch.

  A switch cannot operate with BMP and as a DHCP client simultaneously. To disable BMP in EXEC mode, use the `stop bmp` command. After BMP stops, the switch acts as a DHCP client.

- Acquire a dynamic IP address from a DHCP client is for a limited period or until the client releases the address.
- A DHCP server manages and assigns IP addresses to clients from an address pool stored on the server. For more information, refer to Configuring the Server for Automatic Address Allocation.
- Dynamically assigned IP addresses are supported on Ethernet, VLAN, and port-channel interfaces.
- The public out-of-band management interface and default VLAN 1 are configured by default as a DHCP client to acquire a dynamic IP address from a DHCP server.
- By default, the switch is configured to operate in BMP mode as a DHCP client that sends DHCP requests to a DHCP server to retrieve configuration information (IP address, boot-image filename, and configuration file). All ports and management interfaces are brought up in Layer 3 mode and pre-configured with `no shutdown` and `no ip address`. For this reason, you cannot enter configuration commands to set up the switch. To interrupt a BMP process, prevent a loop from occurring, and apply the Dell EMC Networking OS image and startup configuration stored in the local flash, enter the `stop bmp` command from the console. To reconfigure the switch so that it boots up in normal mode using the Dell EMC Networking OS image and startup configuration file in local flash, enter the `reload-type normal-reload` command and save it to the startup configuration:

  ```
  DellEMC# reload-type normal-reload
  DellEMC# write memory
  DellEMC# reload
  ```

- To re-enable BMP mode for the next reload, enter the `reload-type jump-start` command.
- If the management port is associated with any non-default VRF, then the `ip address dhcp` command does not work.

DHCP Client Operation with Other Features

The DHCP client operates with other Dell EMC Networking OS features, as the following describes.

Stacking

The DHCP client daemon runs only on the master unit and handles all DHCP packet transactions. It periodically synchronizes the lease file with the standby unit.

When a stack failover occurs, the new master requires the same DHCP server-assigned IP address on DHCP client interfaces. The new master reinitiates a DHCP packet transaction by sending a DHCP discovery packet on nonbound interfaces.

Virtual Link Trunking (VLT)

A DHCP client is not supported on VLT interfaces.

VLAN and Port Channels

DHCP client configuration and behavior are the same on Virtual LAN (VLAN) and port-channel (LAG) interfaces as on a physical interface.

DHCP Snooping

A DHCP client can run on a switch simultaneously with the DHCP snooping feature as follows:

- If you enable DHCP snooping globally on a switch and you enable a DHCP client on an interface, the trust port, source MAC address, and snooping table validations are not performed on the interface by DHCP snooping for packets destined to the DHCP client daemon.

  The following criteria determine packets destined for the DHCP client:
  - DHCP is enabled on the interface.
  - The user data protocol (UDP) destination port in the packet is 68.
  - The `chaddr` (change address) in the DHCP header of the packet is the same as the interface's MAC address.
- An entry in the DHCP snooping table is not added for a DHCP client interface.

DHCP Server

A switch can operate as a DHCP client and a DHCP server. DHCP client interfaces cannot acquire a dynamic IP address from the DHCP server running on the switch. Acquire a dynamic IP address from another DHCP server.

Virtual Router Redundancy Protocol (VRRP)

Do not enable the DHCP client on an interface and set the priority to 255 or assign the same DHCP interface IP address to a VRRP virtual group. Doing so guarantees that this router becomes the VRRP group owner.

To use the router as the VRRP owner, if you enable a DHCP client on an interface that is added to a VRRP group, assign a priority less than 255 but higher than any other priority assigned in the group.

DHCP Client on a Management Interface

These conditions apply when you enable a management interface to operate as a DHCP client.

- The management default route is added with the gateway as the router IP address received in the DHCP ACK packet. It is required to send and receive traffic to and from other subnets on the external network. The route is added irrespective when the DHCP client and server are in the same or different subnets. The management default route is deleted if the management IP address is released like other DHCP client management routes.
- *ip route for 0.0.0.0* takes precedence if it is present or added later.
- Management routes added by a DHCP client display with Route Source as **DHCP** in the `show ip management route` and `show ip management-route dynamic` command output.
- Management routes added by DHCP are automatically reinstalled if you configure a static IP route with the `ip route` command that replaces a management route added by the DHCP client. If you remove the statically configured IP route using the `no ip route` command, the management route is reinstalled. Manually delete management routes added by the DHCP client.
- To reinstall management routes added by the DHCP client that is removed or replaced by the same statically configured management routes, release the DHCP IP address and renew it on the management interface.
- Management routes added by the DHCP client have higher precedence over the same statically configured management route. Static routes are not removed from the running configuration if a dynamically acquired management route added by the DHCP client overwrites a static management route.
- Management routes added by the DHCP client are not added to the running configuration.

NOTE: Management routes added by the DHCP client include the specific routes to reach a DHCP server in a different subnet and the management route.

Configure the System for User Port Stacking (Option 230)

Set the stacking-option variable to provide stack-port detail on the DHCP server when you set the DHCP offer. A stack can be formed when the units are connected.

Option 230 is the option for user port stacking. Use it to create up to eight stack groups. Define the configuration parameters on the DHCP server for each chassis based on the chassis MAC address. Configure the following parameters:

- unit number
- priority

- stack group ID

The received stacking configuration is always applied on the master stack unit.

```
option #230 "unit-number:3#priority:2#stack-group:14"
```

Configuring DHCP relay source interface

The following section explains how to configure global and interface level DHCP relay source IPv4 or IPv6 configuration to forward all the DHCP packets from the DHCP client to DHCP server through the configured source interface. This feature is applicable only for L3 interface with relay configuration and L3 DHCP snooping enabled VLANs.

Global DHCP relay source IPv4 or IPv6 configuration

You can configure global DHCP relay source IPv4 or IPv6 configuration using the command `{ip | ipv6} dhcp-relay source-interface` *`interface`* command in the CONFIGURATION mode. DHCP relay uses the IPv4 or IPv6 global source address of the configured interface for relaying packets to the DHCP server. For example, if you configure DHCP relay source interface as a loopback interface, the DHCP relay uses the configured loopback interface to forward the packets to the DHCP server using the configured IPv4 or IPv6 address. The source interface can be a VLAN, LAG, physical and loopback interfaces.

- Specify the type of an interface and interface-number that should be used as a DHCP relay source interface.

  CONFIGURATION mode

  `{ip | ipv6} dhcp relay source-interface` *`interface`*

Following is the sample configuration to configure loopback interface with IPv4 and IPv6 address in CONFIGURATION MODE.

```
Dell(conf)# interface loopback 1
Dell(conf-if-lo-1)# ip vrf forwarding vrf1
Dell(conf-if-lo-1)# ip address 1.1.1.1/32
Dell(conf-if-lo-1)# ipv6 address 1::1/128
Dell(conf-if-lo-1)# no shutdown
```

To configure the loopback interface as IPv4 or IPv6 DHCP relay source interface, enter the following commands in the CONFIGURATION MODE.

```
Dell(conf)# ip dhcp relay source-interface loopback 1
Dell(conf)# ipv6 dhcp relay source-interface loopback 1
```

When you configure the above commands in the CONFIGURATION MODE, it will configure the loopback interface as the DHCP relay source interface for forwarding the DHCP packets from DHCP client to server. When this command is configured in the CONFIGURATION mode level, it is applied globally. So, all the DHCP packets will be relayed or forwarded through the configured L3 interface (loopback 1) using the IPv4 (1.1.1.1/32) and IPv6 addresses (1::1/128) of the loopback configuration.

Interface level DHCP relay source IPv4 or IPv6 configuration

You can configure interface specific DHCP relay source IPv4 or IPv6 configuration.

If the DHCP relay source interface is configured on the interface level, the DHCP relay forwards the packets from these interfaces to the DHCP server using the interface. In case, the DHCP relay source interface is not configured in the interface level (if configured), the DHCP relay chooses the configured global DHCP relay source interface for forwarding the packets.

NOTE:

The DHCP relay source IPv4 or IPv6 configuration at interface level takes precedence over the DHCP relay source IPv4 or IPv6 configuration at the global level.

- Specify the type of an interface and interface-number that should be used as a DHCP relay source interface at the interface level.

  INTERFACE mode

  `{ip | ipv6} dhcp relay source-interface` *`interface`*

Following are the steps to configure interface specific source IPv4 or IPv6 configuration for DHCP relay. The below example shows when the DHCP relay uses the interface specific configuration and global source interface configuration depending on the configuration.

1. Configuring L3 interface with IPv4 or IPv6 address.

   Following are the steps to configure a L3 interface (loopback) with IPv4 and IPv6 address in INTERFACE MODE.

```
Dell(conf)# interface loopback 2
Dell(conf-if-lo-1)# ip vrf forwarding vrf1
Dell(conf-if-lo-1)# ip address 2.2.2.2/32
Dell(conf-if-lo-1)# ipv6 address 2::2/128
Dell(conf-if-lo-1)# no shutdown
Dell(conf)# interface loopback 3
Dell(conf-if-lo-1)# ip vrf forwarding vrf2
Dell(conf-if-lo-1)# ip address 3.3.3.3/32
Dell(conf-if-lo-1)# ipv6 address 3::3/128
Dell(conf-if-lo-1)# no shutdown
```

2. Creating L3 interfaces with the DHCP helper configuration.

   Following are the steps to configure IPv4 or IPv6 interfaces with the DHCP helper configuration. The below example shows two VLAN interfaces (Vlan 2 and 4), DHCP helper (100.0.0.1 and 100::1) for the respective VLANs and the DHCP relay source IPv4 and IPv6 configuration, and two different loopback interfaces (loopback 2 and 3). DHCP relay forwards packets using the loopback 2 interface with IPv4 and IPv6 addresses ((2.2.2.2/32 and 2::2/128) from Vlan 2. The same way, the relay uses IPv4 and IPv6 addresses (3.3.3.3/32 and 3::3/128) of loopback 3 interface from Vlan 3.

```
Dell(conf)# interface Vlan 2
Dell(conf-if-vl-2)# ip vrf forwarding vrf1
Dell(conf-if-vl-2)# ip address 2.0.0.1/24
Dell(conf-if-vl-2)# ipv6 address 2::1/64
Dell(conf-if-vl-2)# tagged fortyGigE 0/0
Dell(conf-if-vl-2)# ip helper-address vrf vrf1 100.0.0.1
Dell(conf-if-vl-2)# ipv6 helper-address vrf vrf1 100::1
Dell(conf-if-vl-2)# ip dhcp relay source-interface loopback 2
Dell(conf-if-vl-2)# ipv6 dhcp relay source-interface loopback 2
Dell(conf)# interface Vlan 4
Dell(conf-if-vl-4)# ip vrf forwarding vrf1
Dell(conf-if-vl-4)# ip address 4.0.0.1/24
Dell(conf-if-vl-4)# ipv6 address 4::1/64
Dell(conf-if-vl-4)# tagged fortyGigE 0/4
Dell(conf-if-vl-4)# ip helper-address vrf vrf1 100.0.0.1
Dell(conf-if-vl-4)# ipv6 helper-address vrf vrf1 100::1
Dell(conf-if-vl-4)# ip dhcp relay source-interface loopback 3
Dell(conf-if-vl-4)# ipv6 dhcp relay source-interface loopback 3
```

3. In the below configuration, the DHCP relay source interface is not configured in the VLAN interface. So, the DHCP relay uses the configured global DHCP relay source interface to forward the packets from the DHCP client to server.

```
Dell(conf)# interface Vlan 5
Dell(conf-if-vl-4)# ip vrf forwarding vrf1
Dell(conf-if-vl-4)# ip address 4.0.0.1/24
Dell(conf-if-vl-4)# ipv6 address 4::1/64
Dell(conf-if-vl-4)# tagged TenGigE 1/4
Dell(conf-if-vl-4)# ip helper-address vrf vrf1 100.0.0.1
Dell(conf-if-vl-4)# ipv6 helper-address vrf vrf1 100::1
```

Configure Secure DHCP

DHCP as defined by RFC 2131 provides no authentication or security mechanisms. Secure DHCP is a suite of features that protects networks that use dynamic address allocation from spoofing and attacks.

- Option 82
- DHCP Snooping
- Dynamic ARP Inspection
- Source Address Validation

Option 82 (DHCPv4 relay options)

RFC 3046 (the relay agent information option, or Option 82) is used for class-based IP address assignment.

The code for the relay agent information option is 82, and includes two suboptions, circuit ID and remote ID.

Circuit ID	This is the interface on which the client-originated message is received.
Remote ID	This identifies the host from which the message is received. The value of this suboption is the MAC address of the relay agent that adds Option 82.

In DHCPv4 relay , the Option 82 is not added by default. When the `ip dhcp relay information-option` is configured, the Option 82 with suboptions 1 (Agent Circuit ID) and 2 (Agent Remote ID) are added to the relayed DHCP packet. The default values of the suboptions are as follows:

- Default Agent Circuit ID is constructed in the format `VLANID:LagID:SlotID:PortStr`. When the port is fanned-out, the `PortStr` is represented as `mainPort:subPort` (all in ASCII format).
- Default Agent Remote ID is the system MAC address (in binary format).

The following example shows the format of the Circuit ID - 723:0:1:1.

Table 25. Circuit ID Format

VLAN ID	LAG ID	Slot ID	Port Str
723	0	1	1

The DHCP relay agent inserts Option 82 before forwarding DHCP packets to the server. The server can use this information to:

- Track the number of address requests per relay agent. Restricting the number of addresses available per relay agent can harden a server against address exhaustion attacks.
- Associate client MAC addresses with a relay agent to prevent offering an IP address to a client spoofing the same MAC address on a different relay agent.
- Assign IP addresses according to the relay agent. This prevents generating DHCP offers in response to requests from an unauthorized relay agent.

The server echoes the option back to the relay agent in its response, and the relay agent uses the information in the option to forward a reply out the interface on which the request was received, rather than flooding it on the entire VLAN.

The relay agent strips Option 82 from DHCP responses before forwarding them to the client.

By default, Option 82 is not inserted in DHCP packets.

To insert Option 82 into DHCP packets, follow this step.

- Insert Option 82 into DHCP packets.

  CONFIGURATION mode

  ```
  ip dhcp relay information-option [trust-downstream]
  ```

  For routers between the relay agent and the DHCP server, enter the `trust-downstream` option.
- Manually reset the remote ID for Option 82.

  CONFIGURATION mode

  ```
  ip dhcp relay information-option remote-id
  ```

DHCPv6 relay agent options

The DHCPv6 relay agent inserts Options 18 and 37 before forwarding DHCPv6 packets to the server. These DHCPv6 options are enabled by default and are not configurable.

Interface ID (Option 18)	Interface on which the client-originated message is received. The interface-ID is 12 bytes long and is constructed using three ifindexes: Logical, Received, and Physical. Each of the ifindex is 4 bytes long.
Remote ID (Option 37)	Identifies the host from which the message is received.

The default values of the Options 18 and 37 are as follows:

- Default Agent Interface ID is constructed in the format `VLANID:LagID:SlotID:PortStr`. When the port is fanned-out, the `PortStr` is represented as `mainPort:subPort` (all in ASCII format).
- Default Agent Remote ID is the system MAC address of the relay agent that adds Option 37 (in binary format).

DHCP Snooping

DHCP snooping is a feature that protects networks from spoofing. It acts as a firewall between the DHCP server and DHCP clients.

DHCP snooping places the ports either in trusted or non-trusted mode. By default, all ports are set to the non-trusted mode. An attacker can not connect to the DHCP server through trusted ports. While configuring DHCP snooping, manually configure ports connected to legitimate servers and relay agents as trusted ports.

When you enable DHCP snooping, the relay agent builds a binding table — using DHCPACK messages — containing the client MAC address, IP addresses, IP address lease time, port, VLAN ID, and binding type. Every time the relay agent receives a DHCPACK on a trusted port, it adds an entry to the table.

The relay agent checks all subsequent DHCP client-originated IP traffic (DHCPRELEASE, DHCPNACK, and DHCPDECLINE) against the binding table to ensure that the MAC-IP address pair is legitimate and that the packet arrived on the correct port. Packets that do not pass this check are forwarded to the server for validation. This checkpoint prevents an attacker from spoofing a client and declining or releasing the real client's address. Server-originated packets (DHCPOFFER, DHCPACK, and DHCPNACK) that arrive on a not trusted port are also dropped. This checkpoint prevents an attacker from acting as an imposter as a DHCP server to facilitate a man-in-the-middle attack.

Binding table entries are deleted when a lease expires, or the relay agent encounters a DHCPRELEASE, DHCPNACK, or DHCPDECLINE.

DHCP snooping is supported on Layer 2 and Layer 3 traffic. DHCP snooping on Layer 2 interfaces does not require a relay agent.

NOTE: In DHCP relay agent, configure DHCP snooping such that the packet from DHCP client must not pass through DHCP snooping-enabled switches twice before reaching the DHCP server.

Binding table entries are deleted when a lease expires or when the relay agent encounters a DHCPRELEASE. Line cards maintain a list of snooped VLANs. When the binding table is exhausted, DHCP packets are dropped on snooped VLANs, while these packets are forwarded across non-snooped VLANs. Because DHCP packets are dropped, no new IP address assignments are made. However, DHCPRELEASE and DHCPDECLINE packets are allowed so that the DHCP snooping table can decrease in size. After the table usage falls below the maximum limit of 4000 entries, new IP address assignments are allowed.

NOTE: DHCP server packets are dropped on all non-trusted interfaces of a system configured for DHCP snooping. To prevent these packets from being dropped, configure `ip dhcp snooping trust` on the server-connected port.

DHCP Snooping for a Multi-Tenant Host

You can configure the DHCP snooping feature such that multiple IP addresses are expected for the same MAC address. You can use the `ip dhcp snooping` command multiple times to map the same MAC address with different IP addresses. This configuration is also used for dynamic ARP inspection (DAI) and source address validation (SAV). The DAI and SAV tables reflect the same entries in the DHCP snooping binding table.

NOTE: If you enable DHCP Option 82 using the `ip dhcp relay` command, by default, the remote-ID field contains the MAC address of the relay agent. If you configure the remote ID as the host name in a VLT setup, configure different host names on both VLT peers. If you configure the remote ID with your own string, ensure that your strings are different on both VLT peers.

DHCP Snooping in a VLT Setup

In a VLT setup, the DHCP snooping binding table synchronizes between the VLT nodes. Similarly, the DAI and SAV tables also synchronize between VLT nodes. For this feature to work in a VLT setup, make sure that DHCP relay, DHCP snooping, SAV, and DAI configurations are identical between the VLT peer nodes.

Enabling DHCP Snooping

To enable DHCP snooping, use the following commands.

1. Enable DHCP snooping globally.
   CONFIGURATION mode
   `ip dhcp snooping`
2. Specify ports connected to DHCP servers as trusted.
   INTERFACE mode
   INTERFACE PORT EXTENDER mode
   `ip dhcp snooping trust`
3. Enable DHCP snooping on a VLAN.
   CONFIGURATION mode
   `ip dhcp snooping vlan` *`name`*

Enabling IPv6 DHCP Snooping

To enable IPv6 DHCP snooping, use the following commands.

1. Enable IPv6 DHCP snooping globally.
   CONFIGURATION mode
   `ipv6 dhcp snooping`
2. Specify ports connected to IPv6 DHCP servers as trusted.
   INTERFACE mode
   `ipv6 dhcp snooping trust`
3. Enable IPv6 DHCP snooping on a VLAN or range of VLANs.
   CONFIGURATION mode
   `ipv6 dhcp snooping vlan` *`vlan-id`*

Adding a Static Entry in the Binding Table

To add a static entry in the binding table, use the following command.

- Add a static entry in the binding table.
  EXEC Privilege mode
  `ip dhcp snooping binding mac` *`mac-address`* `vlan-id` *`vlan-id`* `ip` *`ip-address`* `interface` *`interface-type interface-number`* `lease` *`lease-value`*
  If multiple IP addresses are expected for the same MAC address, repeat this step for all IP addresses.

Adding a Static IPV6 DHCP Snooping Binding Table

To add a static entry in the snooping database, use the following command.

- Add a static entry in the snooping binding table.
  EXEC Privilege mode
  `ipv6 dhcp snooping binding mac` *`address`* `vlan-id` *`vlan-id`* `ipv6` *`ipv6-address`* `interface` *`interface-type | interface-number`* `lease` *`value`*

Clearing the Binding Table

To clear the binding table, use the following command.

- Delete all of the entries in the binding table.
  EXEC Privilege mode

```
clear ip dhcp snooping binding
```

Clearing the DHCP IPv6 Binding Table

To clear the DHCP IPv6 binding table, use the following command.

- Delete all of the entries in the binding table.

  EXEC Privilege mode

  ```
  clear ipv6 dhcp snooping binding
  ```

```
DellEMC# clear ipv6 dhcp snooping?
binding Clear the snooping binding database
```

Displaying the Contents of the Binding Table

To display the contents of the binding table, use the following command.

- Display the DHCP snooping information.

  EXEC Privilege mode

  ```
  show ip dhcp snooping
  ```

- Display the contents of the binding table.

  EXEC Privilege mode

  ```
  show ip dhcp snooping binding
  ```

View the DHCP snooping statistics with the `show ip dhcp snooping` command.

```
DellEMC#show ip dhcp snooping

IP DHCP Snooping                                : Enabled.
IP DHCP Snooping Mac Verification               : Disabled.
IP DHCP Relay Information-option                : Disabled.
IP DHCP Relay Trust Downstream                  : Disabled.

Database write-delay (In minutes)               : 0

DHCP packets information
Relay Information-option packets                : 0
Relay Trust downstream packets                  : 0
Snooping packets                                : 0

Packets received on snooping disabled L3 Ports  : 0
Snooping packets processed on L2 vlans          : 142

DHCP Binding File Details
Invalid File                                    : 0
Invalid Binding Entry                           : 0
Binding Entry lease expired                     : 0
List of Trust Ports                             :Te 1/4
List of DHCP Snooping Enabled Vlans             :Vl 10
List of DAI Trust ports                         :Te 1/4
```

View the DHCP snooping binding table using the `show ip dhcp snooping binding` command.

```
DellEMC#show ip dhcp snooping binding
Codes :  S - Static D - Dynamic
IP Address        MAC Address        Expires(Sec)  Type  VLAN   Interface
=========================================================================
10.1.1.254       00:00:a0:00:00:02    162           D    Vl 200  Te 1/4
10.1.1.225       00:00:a0:00:00:03    163           D    Vl 100  Te 1/4
```

The following example output of the `show ip dhcp snooping binding` command displays that different IP addresses are mapped to the same MAC address:

```
DellEMC#show ip dhcp snooping binding
Codes :  S - Static D - Dynamic
IP Address        MAC Address        Expires(Sec)  Type  VLAN   Interface
=========================================================================
10.1.1.100       00:00:a0:00:00:00    39735         S    Vl 200  Te 1/4
10.1.1.101       00:00:a0:00:00:00    39736         S    Vl 200  Te 1/4
10.1.1.254       00:00:a0:00:00:00    162           D    Vl 200  Te 1/4
```

The following example shows a sample output of the `show ip dhcp snooping binding` command for a device connected to both the VLT peers. The Po 10 interface is the VLT port channel connected to a ToR switch or an end device.

```
DellEMC#show ip dhcp  snooping binding
Codes :  S - Static D - Dynamic
IP Address        MAC Address        Expires(Sec)  Type  VLAN   Interface
=========================================================================
10.1.1.10        00:00:a0:00:00:00    39735         S    Vl 200  Po 10
10.1.1.11        00:00:a0:00:00:00    39736         S    Vl 200  Po 10
10.1.1.25        00:00:a0:00:00:00    162           D    Vl 200  Po 10
```

The following example shows a sample output of the `show ip dhcp snooping binding` command for a device connected to one of the VLT peers only (orphaned). The physical interface is the one that is directly connected to the VLT peer.

```
DellEMC#show ip dhcp snooping binding
Codes :  S - Static D - Dynamic
IP Address        MAC Address        Expires(Sec)  Type  VLAN   Interface
=========================================================================
10.1.1.100       00:00:a0:00:00:00    39735         S    Vl 200  Te 1/4
10.1.1.101       00:00:a0:00:00:00    39736         S    Vl 200  Te 1/4
10.1.1.254       00:00:a0:00:00:00    162           D    Vl 200  Te 1/4
```

The following example shows a sample output of the `show ip dhcp snooping binding` command for a device connected to the peer VLT node, but not to itself. The Po 10 interface is the VLTi link between the VLT peers.

```
DellEMC#show ip dhcp  snooping binding
Codes :  S - Static D - Dynamic
IP Address        MAC Address        Expires(Sec)  Type  VLAN   Interface
=========================================================================
10.1.1.10        00:00:a0:00:00:00    39735         S    Vl 200  Po 10
10.1.1.11        00:00:a0:00:00:00    39736         S    Vl 200  Po 10
10.1.1.25        00:00:a0:00:00:00    162           D    Vl 200  Po 10
```

Displaying the Contents of the DHCPv6 Binding Table

To display the contents of the DHCP IPv6 binding table, use the following command.

- Display the contents of the binding table.

  EXEC Privilege mode

  `show ipv6 dhcp snooping biniding`

View the DHCP snooping statistics with the `show ipv6 dhcp snooping` command.

```
DellEMC#show ipv6 dhcp snooping binding
Codes :  S - Static D - Dynamic
IPv6 Address    MAC Address       Expires(Sec)  Type VLAN    Interface
=========================================================================
11:11::22      11:22:11:22:11:22    120331      S    Vl 100   Te 1/1
33::22         11:22:11:22:11:23    120331      S    Vl 200   Te 1/1
333:22::22     11:22:11:22:11:24    120331      D    Vl 300   Te 1/2
```

Debugging the IPv6 DHCP

To debug the IPv6 DHCP, use the following command.

- Display debug information for IPV6 DHCP.

  EXEC Privilege mode

  ```
  debug ipv6 dhcp
  ```

IPv6 DHCP Snooping MAC-Address Verification

Configure to enable verify source mac-address in the DHCP packet against the mac address stored in the snooping binding table.

- Enable IPV6 DHCP snooping .

  CONFIGURATION mode

  ```
  ipv6 dhcp snooping verify mac-address
  ```

Configuring the DHCP secondary-subnet

DHCP Secondary subnet feature is an extended feature of DHCP relay agent. After enabling DHCP Relay Secondary Subnet, relay agent tries to get IP Address from secondary IP Address subnet pool, only when relay agent fails to get IP Address through primary IP Address.

1. To enable all the DHCP relay secondary — subnet interfaces.

   CONFIGURATION mode

   ```
   ip dhcp relay secondary-subnet
   ```

2. Sample Configuration:

   ```
   Dell(conf)#ip dhcp relay secondary-subnet
   Dell(conf)#interface TenGigabitEthernet 0/0
   Dell(conf-if-te-0/0)#ip address 10.1.1.1/24
   Dell(conf-if-te-0/0)#ip address 11.1.1.1/24 secondary
   Dell(conf-if-te-0/0)#ip helper-address 2.1.1.1
   Dell(conf-if-te-0/0)#no shutdown
   Dell(conf-if-te-0/0)#
   ```

   DHCP relay tries to get ip address for the client, through configured primary address filling giaddr (relay address) 10.1.1.1. If there is no OFFER from DHCP Server, for three DISCOVER retries. Since secondary-subnet enabled, DHCP Relay Agent tries to get ip address through secondary ip address filling giaddr (relay address) 11.1.1.1.

3. To disable the DHCP relay secondary-subnet:

   ```
   Dell(conf)# no ip dhcp relay secondary-subnet
   ```

Drop DHCP Packets on Snooped VLANs Only

Binding table entries are deleted when a lease expires or the relay agent encounters a DHCPRELEASE.

Line cards maintain a list of snooped VLANs. When the binding table fills, DHCP packets are dropped only on snooped VLANs, while such packets are forwarded across non-snooped VLANs. Because DHCP packets are dropped, no new IP address assignments are made. However, DHCP release and decline packets are allowed so that the DHCP snooping table can decrease in size. After the table usage falls below the maximum limit of 4000 entries, new IP address assignments are allowed.

To view the number of entries in the table, use the `show ip dhcp snooping binding` command. This output displays the snooping binding table created using the ACK packets from the trusted port.

```
DellEMC#show ip dhcp snooping binding

Codes : S - Static D - Dynamic
```

```
IP Address  MAC Address         Expires(Sec) Type VLAN   Interface
============================================================
10.1.1.251  00:00:4d:57:f2:50   172800       D    Vl 10   Te 1/2
10.1.1.252  00:00:4d:57:e6:f6   172800       D    Vl 10   Te 1/1
10.1.1.253  00:00:4d:57:f8:e8   172740       D    Vl 10   Te 1/3
10.1.1.254  00:00:4d:69:e8:f2   172740       D    Vl 10   Te 1/5

Total number of Entries in the table : 4
```

Dynamic ARP Inspection

Dynamic address resolution protocol (ARP) inspection prevents ARP spoofing by forwarding only ARP frames that have been validated against the DHCP binding table.

ARP is a stateless protocol that provides no authentication mechanism. Network devices accept ARP requests and replies from any device. ARP replies are accepted even when no request was sent. If a client receives an ARP message for which a relevant entry already exists in its ARP cache, it overwrites the existing entry with the new information.

The lack of authentication in ARP makes it vulnerable to spoofing. ARP spoofing is a technique attackers use to inject false IP-to-MAC mappings into the ARP cache of a network device. It is used to launch man-in-the-middle (MITM), and denial-of-service (DoS) attacks, among others.

A spoofed ARP message is one in which the MAC address in the sender hardware address field and the IP address in the sender protocol field are strategically chosen by the attacker. For example, in an MITM attack, the attacker sends a client an ARP message containing the attacker's MAC address and the gateway's IP address. The client then thinks that the attacker is the gateway, and sends all internet-bound packets to it. Likewise, the attacker sends the gateway an ARP message containing the attacker's MAC address and the client's IP address. The gateway then thinks that the attacker is the client and forwards all packets addressed to the client to it. As a result, the attacker is able to sniff all packets to and from the client.

Other attacks using ARP spoofing include:

Broadcast	An attacker can broadcast an ARP reply that specifies FF:FF:FF:FF:FF:FF as the gateway's MAC address, resulting in all clients broadcasting all internet-bound packets.
MAC flooding	An attacker can send fraudulent ARP messages to the gateway until the ARP cache is exhausted, after which, traffic from the gateway is broadcast.
Denial of service	An attacker can send a fraudulent ARP messages to a client to associate a false MAC address with the gateway address, which would blackhole all internet-bound packets from the client.

NOTE: Dynamic ARP inspection (DAI) uses entries in the L2SysFlow CAM region, a sub-region of SystemFlow. One CAM entry is required for every DAI-enabled VLAN. You can enable DAI on up to 16 VLANs on a system.

Configuring Dynamic ARP Inspection

To enable dynamic ARP inspection, use the following commands.

1. Enable DHCP snooping.
2. Validate ARP frames against the DHCP snooping binding table.

   INTERFACE VLAN mode

   `arp inspection`

To view entries in the ARP database, use the `show arp inspection database` command.

```
DellEMC#show arp inspection database

Protocol  Address     Age(min) Hardware Address   Interface VLAN   CPU
-------------------------------------------------------------------------
Internet  10.1.1.251  -        00:00:4d:57:f2:50  Te 1/2    Vl 10  CP
Internet  10.1.1.252  -        00:00:4d:57:e6:f6  Te 1/1    Vl 10  CP
Internet  10.1.1.253  -        00:00:4d:57:f8:e8  Te 1/3    Vl 10  CP
Internet  10.1.1.254  -        00:00:4d:69:e8:f2  Te 1/5   Vl 10  CP
DellEMC#
```

To see how many valid and invalid ARP packets have been processed, use the `show arp inspection statistics` command.

```
DellEMC#show arp inspection statistics

Dynamic ARP Inspection (DAI) Statistics
---------------------------------------
Valid ARP Requests           : 0
Valid ARP Replies            : 1000
Invalid ARP Requests         : 1000
Invalid ARP Replies          : 0
DellEMC#
```

Configuring dynamic ARP inspection-limit

To configure dynamic ARP inspection rate limit on a port, perform the following task.

1. Enter into global configuration mode.
   EXEC Privilege mode
   `configure terminal`
2. Select the interface to be configured.
   CONFIGURATION mode
   `interface interface-name`
3. Configure ARP packet inspection rate limiting.
   INTERFACE CONFIGURATION mode
   `arp inspection-limit {rate pps [interval seconds]}`
   The rate packet per second (pps) range is from 1 to 2048. The default is 15.
   The rate burst interval range is from 1 to 15 seconds. The default is 1.

Bypassing the ARP Inspection

You can configure a port to skip ARP inspection by defining the interface as trusted, which is useful in multi-switch environments.

ARPs received on trusted ports bypass validation against the binding table. All ports are untrusted by default.

To bypass the ARP inspection, use the following command.

- Specify an interface as trusted so that ARPs are not validated against the binding table.
  INTERFACE mode
  `arp inspection-trust`

Dynamic ARP inspection is supported on Layer 2 and Layer 3.

Source Address Validation

Using the DHCP binding table, Dell EMC Networking OS can perform three types of source address validation (SAV).

Table 26. Three Types of Source Address Validation

Source Address Validation	Description
IP Source Address Validation	Prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table.
DHCP MAC Source Address Validation	Verifies a DHCP packet's source hardware address matches the client hardware address field (CHADDR) in the payload.
IP+MAC Source Address Validation	Verifies that the IP source address and MAC source address are a legitimate pair.

Enabling IP Source Address Validation

IP source address validation (SAV) prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table.

A spoofed IP packet is one in which the IP source address is strategically chosen to disguise the attacker. For example, using ARP spoofing, an attacker can assume a legitimate client's identity and receive traffic addressed to it. Then the attacker can spoof the client's IP address to interact with other clients.

The DHCP binding table associates addresses the DHCP servers assign with the port or the port channel interface on which the requesting client is attached and the VLAN the client belongs to. When you enable IP source address validation on a port, the system verifies that the source IP address is one that is associated with the incoming port and optionally that the client belongs to the permissible VLAN. If an attacker is impostering as a legitimate client, the source address appears on the wrong ingress port and the system drops the packet. If the IP address is fake, the address is not on the list of permissible addresses for the port and the packet is dropped. Similarly, if the IP address does not belong to the permissible VLAN, the packet is dropped.

To enable IP source address validation, use the following command.

NOTE: If you enable IP source guard using the `ip dhcp source-address-validation` command and if there are more entries in the current DHCP snooping binding table than the available CAM space, SAV may not be applied to all entries. To ensure that SAV is applied correctly to all entries, enable the `ip dhcp source-address-validation` command before adding entries to the binding table.

- Enable IP source address validation.

  INTERFACE mode

  ```
  ip dhcp source-address-validation
  ```
- Enable IP source address validation with VLAN option.

  INTERFACE mode

  ```
  ip dhcp source-address-validation vlan vlan-id
  ```

  NOTE:

  Before enabling SAV With VLAN option, allocate at least one FP block to the ipmacacl CAM region.

DHCP MAC Source Address Validation

DHCP MAC source address validation (SAV) validates a DHCP packet's source hardware address against the client hardware address field (CHADDR) in the payload.

Dell EMC Networking OS ensures that the packet's source MAC address is checked against the CHADDR field in the DHCP header only for packets from snooped VLANs.

- Enable DHCP MAC SAV.

  CONFIGURATION mode

  ```
  ip dhcp snooping verify mac-address
  ```

Enabling IP+MAC Source Address Validation

IP source address validation (SAV) validates the IP source address of an incoming packet and optionally the VLAN ID of the client against the DHCP snooping binding table. IP+MAC SAV ensures that the IP source address and MAC source address are a legitimate pair, rather than validating each attribute individually. You cannot configure IP+MAC SAV with IP SAV.

1. Allocate at least one FP block to the ipmacacl CAM region.

   CONFIGURATION mode

   ```
   cam-acl l2acl
   ```
2. Save the running-config to the startup-config.

   EXEC Privilege mode

   ```
   copy running-config startup-config
   ```
3. Reload the system.

   EXEC Privilege

```
reload
```

4. Do one of the following.
   - Enable IP+MAC SAV.

     INTERFACE mode

     ```
     ip dhcp source-address-validation ipmac
     ```

   - Enable IP+MAC SAV with VLAN option.

     INTERFACE mode

     ```
     ip dhcp source-address-validation ipmac vlan vlan-id
     ```

Dell EMC Networking OS creates an ACL entry for each IP+MAC address pair and optionally with its VLAN ID in the binding table and applies it to the interface.

To display the IP+MAC ACL for an interface for the entire system, use the `show ip dhcp snooping source-address-validation [interface]` command in EXEC Privilege mode.

Viewing the Number of SAV Dropped Packets

The following output of the `show ip dhcp snooping source-address-validation discard-counters` command displays the number of SAV dropped packets.

```
DellEMC>show ip dhcp snooping source-address-validation discard-counters
deny access-list on TenGigabitEthernet 1/1
Total cam count 1
deny count (0 packets)
deny access-list on TenGigabitEthernet 1/2
Total cam count 2
deny vlan 10 count (0 packets)
deny vlan 20 count (0 packets)
```

The following output of the `show ip dhcp snooping source-address-validation discard-counters interface interface` command displays the number of SAV dropped packets on a particular interface.

```
DellEMC>show ip dhcp snooping source-address-validation discard-counters interface
TenGigabitEthernet 1/1
deny access-list on TenGigabitEthernet 1/1
Total cam count 2
deny vlan 10 count (0 packets)
deny vlan 20 count (0 packets)
```

Clearing the Number of SAV Dropped Packets

To clear the number of SAV dropped packets, use the `clear ip dhcp snooping source-address-validation discard-counters` command.

```
DellEMC>clear ip dhcp snooping source-address-validation discard-counters
```

To clear the number of SAV dropped packets on a particular interface, use the `clear ip dhcp snooping source-address-validation discard-counters interface interface` command.

```
DellEMC>clear ip dhcp snooping source-address-validation discard-counters interface
TenGigabitEthernet 1/1
```

14

Equal Cost Multi-Path (ECMP)

This chapter describes configuring ECMP. This chapter describes configuring ECMP.

Topics:

- ECMP for Flow-Based Affinity
- Link Bundle Monitoring
- RTAG7
- Flow-based Hashing for ECMP

ECMP for Flow-Based Affinity

ECMP for flow-based affinity includes link bundle monitoring.

Configuring the Hash Algorithm

TeraScale has one algorithm that is used for link aggregation groups (LAGs), ECMP, and NH-ECMP, and ExaScale can use three different algorithms for each of these features.

To adjust the ExaScale behavior to match TeraScale, use the following command.

- Change the ExaScale hash-algorithm for LAG, ECMP, and NH-ECMP to match TeraScale.

  CONFIGURATION mode.

  ```
  hash-algorithm ecmp checksum 0 lag checksum 0 nh-ecmp checksum 0
  ```

Dell EMC Networking OS Behavior: In the Dell EMC Networking OS versions prior to 8.2.1.2, the ExaScale default hash-algorithm is 0. Beginning with Dell EMC Networking OS version 8.2.1.2, the default hash-algorithm is 24.

Enabling Deterministic ECMP Next Hop

Deterministic ECMP next hop arranges all ECMPs in order before writing them into the content addressable memory (CAM).

For example, suppose the RTM learns eight ECMPs in the order that the protocols and interfaces came up. In this case, the forwarding information base (FIB) and CAM sorts them so that the ECMPs are always arranged. This implementation ensures that every chassis having the same prefixes orders the ECMPs the same.

With eight or less ECMPs, the ordering is lexicographic and deterministic. With more than eight ECMPs, ordering is deterministic, but it is not in lexicographic order.

To enable deterministic ECMP next hop, use the appropriate command.

NOTE: Packet loss might occur when you enable `ip/ipv6 ecmp-deterministic` for the first-time only.

- Enable IPv4 Deterministic ECMP next hop.

  CONFIGURATION mode.

  ```
  ip ecmp-deterministic
  ```

- Enable IPv6 Deterministic ECMP next hop.

  CONFIGURATION mode.

  ```
  ipv6 ecmp-deterministic
  ```

Configuring the Hash Algorithm Seed

Deterministic ECMP sorts ECMPs in order even though RTM provides them in a random order. However, the hash algorithm uses as a seed the lower 12 bits of the chassis MAC, which yields a different hash result for every chassis.

This behavior means that for a given flow, even though the prefixes are sorted, two unrelated chassis can select different hops.

Dell EMC Networking OS provides a command line interface (CLI)-based solution for modifying the hash seed to ensure that on each configured system, the ECMP selection is same. When configured, the same seed is set for ECMP, LAG, and NH, and is used for incoming traffic only.

NOTE: While the seed is stored separately on each port-pipe, the same seed is used across all CAMs.

NOTE: You cannot separate LAG and ECMP, but you can use different algorithms across the chassis with the same seed. If LAG member ports span multiple port-pipes and line cards, set the seed to the same value on each port-pipe to achieve deterministic behavior.

NOTE: If you remove the hash algorithm configuration, the hash seed does not return to the original factory default setting.

To configure the hash algorithm seed, use the following command.

- Specify the hash algorithm seed.

  CONFIGURATION mode.

  `hash-algorithm seed` *value* `[stack–unit` *stack–unit-number*`] [port-set` *number*`]`

  The range is from 0 to 4095.

Link Bundle Monitoring

Link bundle monitoring allows the system to monitor the use of multiple links for an uneven distribution.

Monitoring linked ECMP bundles allows traffic distribution amounts in a link to be monitored for unfair distribution at any given time. A global default threshold of 60% is Link bundle monitoring allows the system to monitor the use of multiple links for an uneven distribution. . Links are monitored in 15-second intervals for three consecutive instances. Any deviation within that time causes a syslog to be sent and an alarm event generate. When the deviation clears, another syslog is sent and a clear alarm event generates. For example, link bundle monitoring percent threshold: `%STKUNIT0-M:CP %IFMGR-5-BUNDLE_UNEVEN_DISTRIBUTION: Found uneven distribution in LAG bundle 11.`

Link bundle utilization is calculated as the total bandwidth of all links divided by the total bytes-per-second of all links. Within each ECMP group, you can specify interfaces. If you enable monitoring for the ECMP group, utilization calculation performs when the utilization of the link-bundle (not a link within a bundle) exceeds 60%.

Enable link bundle monitoring using the `ecmp-group` command.

NOTE: An ecmp-group index is generated automatically for each unique ecmp-group when you configure multipath routes to the same network. The system can generate a maximum of 512 unique ecmp-groups. The ecmp-group indexes are generated in even numbers (0, 2, 4, 6... 1022) and are for information only.

For link bundle monitoring with ECMP, to enable the link bundle monitoring feature, use the `ecmp-group` command. In the following example, the ecmp-group with id 2, enabled for link bundle monitoring is user configured. This is different from the ecmp-group index 2 that is created by configuring routes and is automatically generated.

These two ecmp-groups are not related in any way.

Example of Viewing Link Bundle Monitoring

```
DellEMC# show link-bundle-distribution ecmp-group 1
Link-bundle trigger threshold - 60
ECMP bundle - 1 Utilization[In Percent] - 44 Alarm State - Active
Interface  Line Protocol  Utilization[In Percent]
Te 1/1     Up             36
Te 1/1     Up             52
```

Managing ECMP Group Paths

To avoid path degeneration, configure the maximum number of paths for an ECMP route that the L3 CAM can hold. When you do not configure the maximum number of routes, the CAM can hold a maximum ECMP per route.

To configure the maximum number of paths, use the following command.

NOTE: For the new settings to take effect, save the new ECMP settings to the startup-config (`write-mem`) then reload the system.

- Configure the maximum number of paths per ECMP group.

  CONFIGURATION mode.

  `ip ecmp-group maximum-paths {2-64}`
- Enable ECMP group path management.

  CONFIGURATION mode.

  `ip ecmp-group path-fallback`

```
DellEMC(conf)#ip ecmp-group maximum-paths 3
User configuration has been changed. Save the configuration and reload to take effect
DellEMC(conf)#
```

Creating an ECMP Group Bundle

Within each ECMP group, you can specify an interface.

If you enable monitoring for the ECMP group, the utilization calculation is performed when the average utilization of the link-bundle (as opposed to a single link within the bundle) exceeds 60%.

1. Create a user-defined ECMP group bundle.

   CONFIGURATION mode

   `ecmp-group ecmp-group-id`

   The range is from 1 to 64.
2. Add interfaces to the ECMP group bundle.

   CONFIGURATION ECMP-GROUP mode

   `interface interface`
3. Enable monitoring for the bundle.

   CONFIGURATION ECMP-GROUP mode

   `link-bundle-monitor enable`

Modifying the ECMP Group Threshold

You can customize the threshold percentage for monitoring ECMP group bundles.

To customize the ECMP group bundle threshold and to view the changes, use the following commands.

- Modify the threshold for monitoring ECMP group bundles.

  CONFIGURATION mode

  `link-bundle-distribution trigger-threshold {percent}`

  The range is from 1 to 90%.

  The default is **60%**.
- Display details for an ECMP group bundle.

  EXEC mode

  `show link-bundle-distribution ecmp-group ecmp-group-id`

  The range is from 1 to 64.

NOTE: An ecmp-group index is generated automatically for each unique ecmp-group when you configure multipath routes to the same network. The system can generate a maximum of 512 unique ecmp-groups. The ecmp-group indices are generated in even numbers (0, 2, 4, 6... 1022) and are for information only.

You can configure ecmp-group with *id 2* for link bundle monitoring. This ecmp-group is different from the ecmp-group *index 2* that is created by configuring routes and is automatically generated. These two ecmp-groups are not related in any way.

```
DellEMC(conf-ecmp-group-5)#show config
!
ecmp-group 5
  interface tengigabitethernet 1/2
  interface tengigabitethernet 1/3
  link-bundle-monitor enable
DellEMC(conf-ecmp-group-5)#
```

RTAG7

RTAG7 is a hashing algorithm that load balances the traffic within a trunk group in a controlled manner. In order to effectively increase the bandwidth of the LAG/Equal Cost Multiple Path routes, traffic is balanced across the member links. The balancing is performed by using the RTAG7 hashing, which is designed to have the member links used efficiently as the traffic profile gets more diverse.Hashing-based load balancing is used in the following applications:

- L3 ECMP
- LAGs
- HiGig trunking

The RTAG7 hash scheme generates a hash that consists of the following two portions:

- The first portion is primarily generated from packet headers to identify micro-flows in the traffic. The Parameters that are considered for hash computations by default in RTAG7 hashing is shown in the given example:

```
DellEMC#show load-balance
Load-Balancing Configuration For LAG & ECMP:
-------------------------------------------
IPV4 Load Balancing Enabled
IPV4 FIELDS : source-ipv4 dest-ipv4 vlan protocol L4-source-port L4-dest-port

IPV6 Load Balancing Enabled
IPV6 FIELDS : source-ipv6 dest-ipv6 vlan protocol L4-source-port L4-dest-port

Mac Load Balancing Enabled
MAC FIELDS : source-mac dest-mac vlan ethertype

Load Balancing Configuration for tunnels
ipv4-over-ipv4 Payload header
ipv4-over-ipv6 Payload header
ipv6-over-ipv6 Payload header
ipv6-over-ipv4 Payload header
ipv4-over-gre-ipv4 Payload header
ipv6-over-gre-ipv4 Payload header
ipv4-over-gre-ipv6 Payload header
ipv6-over-gre-ipv6 Payload header

mac-in-mac header based hashing is disabled
TcpUdp Load Balancing  Enabled
DellEMC(conf)#
```

- Packet Header parameters for the first portion of the RTAG7 hash can be controlled. By default, all the listed parameters from the Packet header are considered for hash computation. Few parameters [on demand] can be removed using the given CLIs.

```
DellEMC(conf)#load-balance ?
flexhash                Enable flexhash based on IP Protocol
ingress-port            Option to Source Port Id for ECMP/LAG hashing
ip-selection            Set the IPV4 key fields to use in hash computation(default =
source-ip dest-ip vlan protocol L4-source-port L4-dest-port)
ipv6-selection          Set the IPV6 key fields to use in hash computation(default =
source-ipv6 dest-ipv6 vlan protocol L4-source-port L4-dest-port)
```

```
mac                     Set the mac key fields to use in hash computation(default =
source-mac dest-mac vlan ethertype)
tcp-udp                 Option to use TCP/UDP ports in packet for ECMP/LAG hashing
tunnel                  Set the tunnel key fields to use in hash computation(default
= Hash-computation based on Inner Header)]
```

- The second portion comes from static physical configuration such as ingress and egress port numbers.
- RTAG7 hashing also provides options to select between multiple hash algorithms that would result in balanced traffic distribution for various traffic patterns.

```
DellEMC(conf)#hash-algorithm ecmp ?
crc16                   CRC16_BISYNC - 16 bit CRC16-bisync polynomial
crc16cc                 CRC16_CCITT - 16 bit CRC16 using CRC16-CCITT polynomial
crc32LSB                CRC32_LOWER - LSB 16 bits of computed CRC32
crc32MSB                CRC32_UPPER - MSB 16 bits of computed CRC32(default)
crc-upper               Use Upper 32 bits of key for hash computation
flow-based-hashing      Enable flow based hashing
dest-ip                 Use Destination IP for ECMP hashing
lsb                     Always return the LSB of the key as the hash
xor1                    CRC16_BISYNC_AND_XOR1 - Upper 8 bits of CRC16-BISYNC and
lower 8 bits of xor1
xor2                    CRC16_BISYNC_AND_XOR2 - Upper 8 bits of CRC16-BISYNC and
lower 8 bits of xor2
xor4                    CRC16_BISYNC_AND_XOR4 - Upper 8 bits of CRC16-BISYNC and
lower 8 bits of xor4
xor8                    CRC16_BISYNC_AND_XOR8 - Upper 8 bits of CRC16-BISYNC and
lower 8 bits of xor8
xor16                   CR16 - 16 bit XOR]
```

Flow-based Hashing for ECMP

Flow-based hashing is one of RTAG7 hashing techniques to cater to ECMP routing in multi-tier networks. It addresses traffic polarization issues by ensuring proper flow distribution between ECMP members in the higher layers of a multi-tier network. It facilitates a dynamic hash function selection across different nodes in the network on a macro flow basis, by reducing route starvation and the unfair distribution of bandwidth between members.

Polarization

Multipath routing is a method that is often used to address data forwarding issues during network failures so that the network traffic reaches its desired destination. Multipath routing in IP networks is typically implemented using Equal-Cost Multipath (ECMP) routing, which employs load balancing algorithms to distribute the traffic over multiple paths towards its destination. In a multi-tier network where load balancing is performed at each tier, static hash algorithms polarize the traffic where load balancing is ineffective in the higher tiers. The polarization effect is exaggerated if all the nodes in the network have to choose from the same set of ECMP paths. Traffic polarization results in packet reordering and route flapping. The following figure explains the traffic polarization effect. Router B performs the same hash as router A and all the traffic goes through the same path to router D, while no traffic is redirected to router E. The following figure explains the traffic polarization effect:

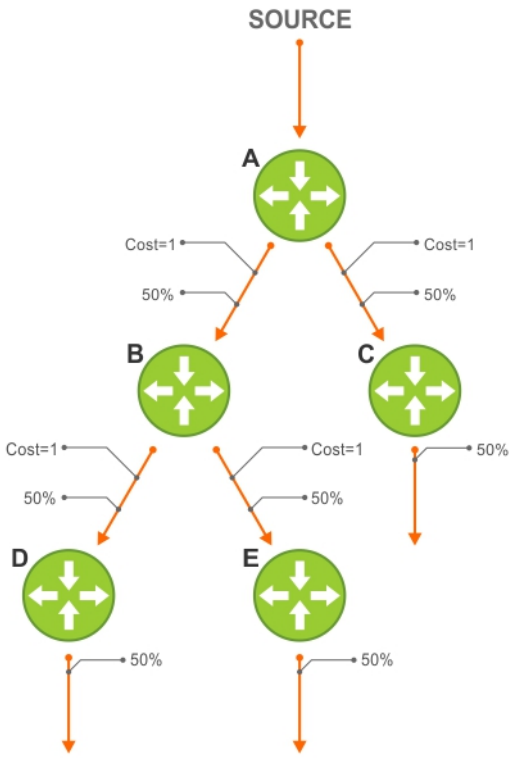


Figure 37. Before Polarization Effect

Router B performs the same hash as router A and all the traffic goes through the same path to router D, while no traffic is redirected to router E.

Some of the anti-polarization techniques used generally to mitigate unequal traffic distribution in LAG/ECMP as follows:

1. Configuring different hash-seed values at each node - Hash seed is the primary parameter in hash computations that determine distribution of traffic among the ECMP paths. The ECMP path can be configured different in each of the nodes "*hash-algorithm seed-value*" would result in better traffic distribution for a given flow, by reducing Polarization effect.
2. Configuring Ingress port as an additional load-balancing parameters [using "load-balance ingress-port enable"] would reduce the polarization effect.
3. Configuring different load-balancing parameters at each tier. In Router A, the hash fields for load balancing could be source-ip, dest-ip, vlan, protocol, L4-source-port and L4-dest-port, whereas on Router B, the hash fields use only source-ip, dest-ip, and protocol
4. Configuring different hash algorithms at different tiers. For example, Router A could use crc16 as the hash algorithm while router B can use XOR16 as the hash algorithm.

Configuration and Benefits

The preceding anti-polarization techniques require some coordinated configuration of network nodes to solve the problem and these techniques are not scalable when the number of tiers in the network is high. Flow based hashing specifically addresses this using **Macro flow-based Hash** function. It facilitates a dynamic hash function selection across different nodes in a network on a macro flow basis, thus reducing unfair distribution of bandwidth between members and starvation.

Selection of Algorithms is available under flow-based-hashing enabling another level of randomness in hash selection. CLI to enable flow-based hashing is shown in the given example:

```
Dell_GW1(conf)#hash-algorithm ecmp flow-based-hashing ?
crc16                  CRC16_BISYNC - 16 bit CRC16-bisync polynomial(default)
crc16cc                CRC16_CCITT - 16 bit CRC16 using CRC16-CCITT polynomial
crc32MSB               CRC32_UPPER - MSB 16 bits of computed CRC32
crc32LSB               CRC32_LOWER - LSB 16 bits of computed CRC32
xor1                   CRC16_BISYNC_AND_XOR1 - Upper 8 bits of CRC16-BISYNC and lower 8
bits of xor1
xor2                   CRC16_BISYNC_AND_XOR2 - Upper 8 bits of CRC16-BISYNC and lower 8
bits of xor2
xor4                   CRC16_BISYNC_AND_XOR4 - Upper 8 bits of CRC16-BISYNC and lower 8
```

```
bits of xor4
xor8                    CRC16_BISYNC_AND_XOR8 - Upper 8 bits of CRC16-BISYNC and lower 8
bits of xor8
xor16                   CR16 - 16 bit XOR]
```

Example to view show hash-algorithm:

```
DellEMC(conf)#hash-algorithm ecmp flow-based-hashing crc16
DellEMC(conf)#end
DellEMC#show hash-algorithm

Hash-Algorithm

linecard 0 Port-Set 0
Seed 185270328
Hg-Seed 185282673
EcmpFlowBasedHashingAlgo- crc16
EcmpAlgo- crc32MSB       LagAlgo- crc32LSB  HgAlgo- crc16
```

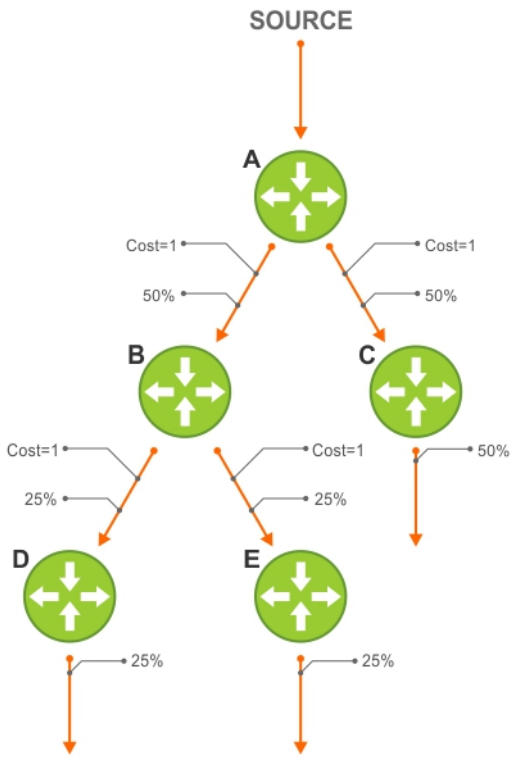


Figure 38. After Polarization Effect

Traffic flow after enabling flow-based hashing

When the flow-based hashing is enabled at all the nodes in the multi-tier network, traffic distribution is balanced at all tiers of the network nullifying the polarization effect. Traffic occurs by the randomness for the flow-based hashing algorithm across multiple nodes in a given network.

15

FIP Snooping

The Fibre Channel over Ethernet (FCoE) Transit feature is supported on Ethernet interfaces. When you enable the switch for FCoE transit, the switch functions as a FIP snooping bridge.

NOTE: FIP snooping is not supported on Fibre Channel interfaces or in a switch stack.

Topics:

- Fibre Channel over Ethernet
- Ensure Robustness in a Converged Ethernet Network
- FIP Snooping on Ethernet Bridges
- FIP Snooping in a Switch Stack
- Using FIP Snooping
- Displaying FIP Snooping Information
- FCoE Transit Configuration Example

Fibre Channel over Ethernet

FCoE provides a converged Ethernet network that allows the combination of storage-area network (SAN) and LAN traffic on a Layer 2 link by encapsulating Fibre Channel data into Ethernet frames.

FCoE works with the Ethernet enhancements provided in data center bridging (DCB) to support lossless (no-drop) SAN and LAN traffic. In addition, DCB provides flexible bandwidth sharing for different traffic types, such as LAN and SAN, according to 802.1p priority classes of service. DCBx should be enabled on the system before the FIP snooping feature is enabled. For more information, refer to the Data Center Bridging (DCB) chapter.

Ensure Robustness in a Converged Ethernet Network

Fibre Channel networks used for SAN traffic employ switches that operate as trusted devices. To communicate with other end devices attached to the Fibre Channel network, end devices log into the switch to which they are attached.

Because Fibre Channel links are point-to-point, a Fibre Channel switch controls all storage traffic that an end device sends and receives over the network. As a result, the switch can enforce zoning configurations, ensure that end devices use their assigned addresses, and secure the network from unauthorized access and denial-of-service (DoS) attacks.

To ensure similar Fibre Channel robustness and security with FCoE in an Ethernet cloud network, FIP establishes virtual point-to-point links between FCoE end-devices (server ENodes and target storage devices) and FCoE forwarders (FCFs) over transit FCoE-enabled bridges.

Ethernet bridges commonly provide ACLs that can emulate a point-to-point link by providing the traffic enforcement required to create a Fibre Channel-level of robustness. You can configure ACLs to emulate point-to-point links, providing control over the traffic received or transmitted into the switch. To automatically generate ACLs, use FIP snooping. In addition, FIP serves as a Layer 2 protocol to:

- Operate between FCoE end-devices and FCFs over intermediate Ethernet bridges to prevent unauthorized access to the network and achieve the required security.
- Allow transit Ethernet bridges to efficiently monitor FIP frames passing between FCoE end-devices and an FCF. To dynamically configure ACLs on the bridge to only permit traffic authorized by the FCF, use the FIP snooping data.

FIP enables FCoE devices to discover one another, initialize and maintain virtual links over an Ethernet network, and access storage devices in a storage area network (SAN). FIP satisfies the Fibre Channel requirement for point-to-point connections by creating a unique virtual link for each connection between an FCoE end-device and an FCF via a transit switch.

FIP provides functionality for discovering and logging into an FCF. After discovering and logging in, FIP allows FCoE traffic to be sent and received between FCoE end-devices (ENodes) and the FCF. FIP uses its own EtherType and frame format. The following illustration shows the communication that occurs between an ENode server and an FCoE switch (FCF).

The following table lists the FIP functions.

Table 27. FIP Functions

FIP Function	Description
FIP VLAN discovery	FCoE devices (ENodes) discover the FCoE VLANs on which to transmit and receive FIP and FCoE traffic.
FIP discovery	FCoE end-devices and FCFs are automatically discovered.
Initialization	FCoE devices learn ENodes from the FLOGI and FDISC to allow immediate login and create a virtual link with an FCoE switch.
Maintenance	A valid virtual link between an FCoE device and an FCoE switch is maintained and the LOGO functions properly.
Logout	On receiving a FLOGI packet, FSB deletes all existing sessions from the ENode to the FCF.

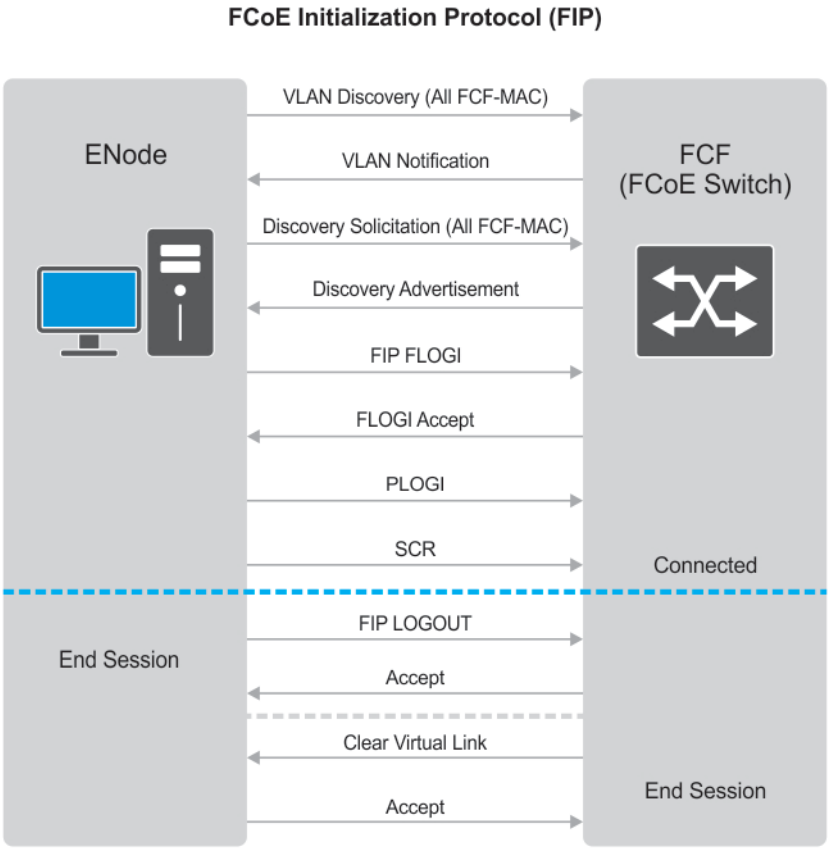


Figure 39. FIP Discovery and Login Between an ENode and an FCF

FIP Snooping on Ethernet Bridges

In a converged Ethernet network, intermediate Ethernet bridges can snoop on FIP packets during the login process on an FCF. Then, using ACLs, a transit bridge can permit only authorized FCoE traffic to be transmitted between an FCoE end-device and an FCF. An Ethernet bridge that provides these functions is called a FIP snooping bridge (FSB).

On a FIP snooping bridge, ACLs are created dynamically as FIP login frames are processed. The ACLs are installed on switch ports configured for ENode mode for server-facing ports and FCF mode for a trusted port directly connected to an FCF.

Enable FIP snooping on the switch, configure the FIP snooping parameters, and configure CAM allocation for FCoE. When you enable FIP snooping, all ports on the switch by default become ENode ports.

Dynamic ACL generation on the switch operating as a FIP snooping bridge function as follows:

Port-based ACLs	These ACLs are applied on all three port modes: on ports directly connected to an FCF, server-facing ENode ports, and bridge-to-bridge links. Port-based ACLs take precedence over global ACLs.
FCoE-generated ACLs	These take precedence over user-configured ACLs. A user-configured ACL entry cannot deny FCoE and FIP snooping frames.

The following illustration shows a switch used as a FIP snooping bridge in a converged Ethernet network. The top-of-rack (ToR) switch operates as an FCF for FCoE traffic. Converged LAN and SAN traffic is transmitted between the ToR switch and an S4820T switch. The switch operates as a lossless FIP snooping bridge to transparently forward FCoE frames between the ENode servers and the FCF switch.

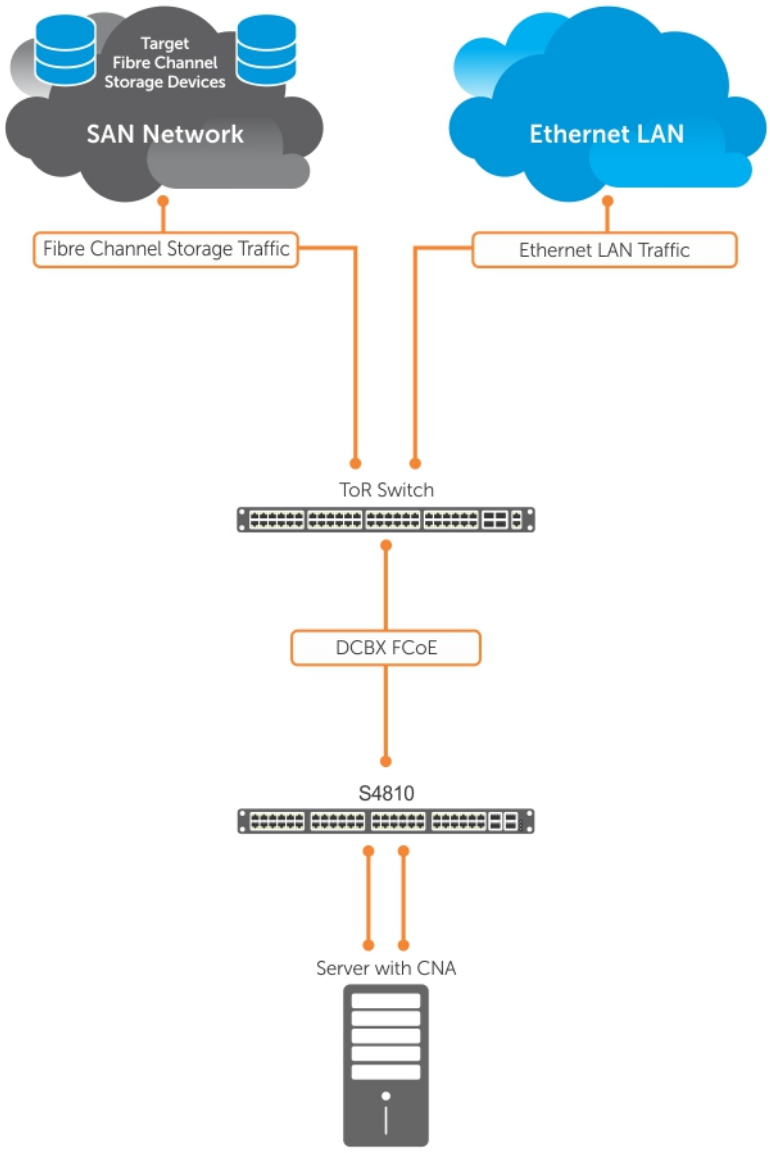


Figure 40. FIP Snooping on a Dell EMC Networking Switch

The following sections describe how to configure the FIP snooping feature on a switch:

- Allocate CAM resources for FCoE.
- Perform FIP snooping (allowing and parsing FIP frames) globally on all VLANs or on a per-VLAN basis.
- To assign a MAC address to an FCoE end-device (server ENode or storage device) after a server successfully logs in, set the FCoE MAC address prefix (FC-MAP) value an FCF uses. The FC-MAP value is used in the ACLs installed in bridge-to-bridge links on the switch.

- To provide more port security on ports that are directly connected to an FCF and have links to other FIP snooping bridges, set the FCF or Bridge-to-Bridge Port modes.
- To ensure that they are operationally active, check FIP snooping-enabled VLANs.
- Process FIP VLAN discovery requests and responses, advertisements, solicitations, FLOGI/FDISC requests and responses, FLOGO requests and responses, keep-alive packets, and clear virtual-link messages.

FIP Snooping in a Switch Stack

FIP snooping supports switch stacking as follows:

- A switch stack configuration is synchronized with the standby stack unit.
- Dynamic population of the FCoE database (ENode, Session, and FCF tables) is synchronized with the standby stack unit. The FCoE database is maintained by snooping FIP keep-alive messages.
- In case of a failover, the new master switch starts the required timers for the FCoE database tables. Timers run only on the master stack unit.

Using FIP Snooping

There are four steps to configure FCoE transit.

1. Enable the FCoE transit feature on a switch.
2. Enable FIP snooping globally on all Virtual Local Area Networks (VLANs) or individual VLANs on a FIP snooping bridge.
3. Configure the FC-Map value applied globally by the switch on all VLANs or an individual VLAN.
4. Configure FCF mode for a FIP snooping bridge-to-FCF link.

For a sample FIP snooping configuration, refer to FIP Snooping Configuration Example.

Statistical information is available for FIP Snooping-related information. For available commands, refer to the *FCoE Transit* chapter in the *Dell EMC Networking OS Command Line Reference Guide*.

FIP Snooping Prerequisites

Before you enable FCoE transit and configure FIP snooping on a switch, ensure that certain conditions are met.

A FIP snooping bridge requires data center bridging exchange protocol (DCBx) and priority-based flow control (PFC) to be enabled on the switch for lossless Ethernet connections (refer to the Data Center Bridging (DCB) chapter). Dell EMC Networking recommends also enabling enhanced transmission selection (ETS); however, ETS is recommended but not required.

If you enable DCBx and PFC mode is on (PFC is operationally up) in a port configuration, FIP snooping is operational on the port. If the PFC parameters in a DCBx exchange with a peer are not synchronized, FIP and FCoE frames are dropped on the port after you enable the FIP snooping feature.

For VLAN membership, you must:

- create the VLANs on the switch which handles FCoE traffic (use the `interface vlan` command).
- configure each FIP snooping port to operate in Hybrid mode so that it accepts both tagged and untagged VLAN frames (use the `portmode hybrid` command).
- configure tagged VLAN membership on each FIP snooping port that sends and receives FCoE traffic and has links with an FCF, ENode server, or another FIP snooping bridge (use the `tagged port-type slot/port` command).

The default VLAN membership of the port must continue to operate with untagged frames. FIP snooping is not supported on a port that is configured for non-default untagged VLAN membership.

Important Points to Remember

- Enable DCBx on the switch before enabling the FIP Snooping feature.
- To enable the feature on the switch, configure FIP Snooping.
- To allow FIP frames to pass through the switch on all VLANs, enable FIP snooping globally on a switch.
- A switch can support a maximum eight VLANs. Configure at least one FCF/bridge-to-bridge port mode interface for any FIP snooping-enabled VLAN.
- You can configure multiple FCF-trusted interfaces in a VLAN.

- When you disable FIP snooping:
  - ACLs are not installed, FIP and FCoE traffic is not blocked, and FIP packets are not processed.
  - The existing per-VLAN and FIP snooping configuration is stored. The configuration is re-applied the next time you enable the FIP snooping feature.
- You must apply the CAM-ACL space for the FCoE region before enabling the FIP-Snooping feature. If you do not apply CAM-ACL space, the following error message is displayed:

```
DellEMC(conf)#feature fip-snooping
% Error: Cannot enable fip snooping. CAM Region not allocated for Fcoe.
DellEMC(conf)#
```

NOTE: Manually add the CAM-ACL space to the FCoE region as it is not applied by default.

To support FIP-Snooping and set CAM-ACL, use`cam-acl l2acl 4 ipv4acl 4 ipv6acl 0 ipv4qos 2 l2qos 1 l2pt 0 ipmacacl 0 vman-qos 0 ecfmacl 0 fcoeacl 2` command.

CAM ACL Table

```
-- Chassis Cam ACL --
             Current Settings(in block sizes)
                   1 block = 128 entries
L2Acl         :          4
Ipv4Acl       :          4
Ipv6Acl       :          0
Ipv4Qos       :          2
L2Qos         :          1
L2PT          :          0
IpMacAcl      :          0
VmanQos       :          0
VmanDualQos   :          0
EcfmAcl       :          0
FcoeAcl       :          2
iscsiOptAcl   :          0
ipv4pbr       :          0
vrfv4Acl      :          0
Openflow      :          0
fedgovacl     :          0
nlbclusteracl:           0

st-sjc-s5000-29#
```

Enabling the FCoE Transit Feature

The following sections describe how to enable FCoE transit.

NOTE: FCoE transit is disabled by default. To enable this feature, you must follow the Configure FIP Snooping.

As soon as you enable the FCoE transit feature on a switch-bridge, existing VLAN-specific and FIP snooping configurations are applied. The FCoE database is populated when the switch connects to a converged network adapter (CNA) or FCF port and compatible DCB configurations are synchronized. By default, all FCoE and FIP frames are dropped unless specifically permitted by existing FIP snooping-generated ACLs. You can reconfigure any of the FIP snooping settings.

If you disable FCoE transit, FIP and FCoE traffic are handled as normal Ethernet frames and no FIP snooping ACLs are generated. The VLAN-specific and FIP snooping configuration is disabled and stored until you re-enable FCoE transit and the configurations are re-applied.

Enable FIP Snooping on VLANs

You can enable FIP snooping globally on a switch on all VLANs or on a specified VLAN.

When you enable FIP snooping on VLANs:

- FIP frames are allowed to pass through the switch on the enabled VLANs and are processed to generate FIP snooping ACLs.

- FCoE traffic is allowed on VLANs only after a successful virtual-link initialization (fabric login FLOGI) between an ENode and an FCF. All other FCoE traffic is dropped.
- You must configure at least one interface for FCF (FCoE Forwarder) mode on a FIP snooping-enabled VLAN. You can configure multiple FCF trusted interfaces in a VLAN.
- A maximum of eight VLANS are supported for FIP snooping on the switch. When enabled globally, FIP snooping processes FIP packets in traffic only from the first eight incoming VLANs. When enabled on a per-VLAN basis, FIP snooping is supported on up to eight VLANs.

Configure the FC-MAP Value

You can configure the FC-MAP value to be applied globally by the switch on all or individual FCoE VLANs to authorize FCoE traffic.

The configured FC-MAP value is used to check the FC-MAP value for the MAC address assigned to ENodes in incoming FCoE frames. If the FC-MAP value does not match, FCoE frames are dropped. A session between an ENode and an FCF is established by the switch-bridge only when the FC-MAP value on the FCF matches the FC-MAP value on the FIP snooping bridge.

Configure a Port for a Bridge-to-Bridge Link

If a switch port is connected to another FIP snooping bridge, configure the FCoE-Trusted Port mode for bridge-bridge links.

Initially, all FCoE traffic is blocked. Only FIP frames with the ALL_FCF_MAC and ALL_ENODE_MAC values in their headers are allowed to pass. After the switch learns the MAC address of a connected FCF, it allows FIP frames destined to or received from the FCF MAC address.

FCoE traffic is allowed on the port only after the switch learns the FC-MAP value associated with the specified FCF MAC address and verifies that it matches the configured FC-MAP value for the FCoE VLAN.

Configure a Port for a Bridge-to-FCF Link

If a port is directly connected to an FCF, configure the port mode as FCF. Initially, all FCoE traffic is blocked; only FIP frames are allowed to pass.

FCoE traffic is allowed on the port only after a successful fabric login (FLOGI) request/response and confirmed use of the configured FC-MAP value for the VLAN.

FLOGI and fabric discovery (FDISC) request/response packets are trapped to the CPU. They are forwarded after the necessary ACLs are installed.

Impact on Other Software Features

When you enable FIP snooping on a switch, other software features are impacted. The following table lists the impact of FIP snooping.

Table 28. Impact of Enabling FIP Snooping

Impact	Description
MAC address learning	MAC address learning is not performed on FIP and FCoE frames, which are denied by ACLs dynamically created by FIP snooping on server-facing ports in ENode mode.
MTU auto-configuration	MTU size is set to mini-jumbo (2500 bytes) when a port is in Switchport mode, the FIP snooping feature is enabled on the switch, and FIP snooping is enabled on all or individual VLANs.
Link aggregation group (LAG)	FIP snooping is supported on port channels on ports on which PFC mode is on (PFC is operationally up).
STP	If you enable an STP protocol (STP, RSTP, PVSTP, or MSTP) on the switch and ports enter a blocking state, when the state change occurs, the corresponding port-based ACLs are

Table 28. Impact of Enabling FIP Snooping (continued)

Impact	Description
	deleted. If a port is enabled for FIP snooping in ENode or FCF mode, the ENode/FCF MAC-based ACLs are deleted.

FIP Snooping Restrictions

The following restrictions apply when you configure FIP snooping.

- The maximum number of FCoE VLANs supported on the switch is eight.
- The maximum number of FIP snooping sessions supported per ENode server is 32. To increase the maximum number of sessions to 64, use the `fip-snooping max-sessions-per-enodemac` command.
- The maximum number of FCFs supported per FIP snooping-enabled VLAN is twelve.
- When FCoE is configured on fanned-out ports or unusable 100G ports, traffic outage occurs for about 45 seconds.

Configuring FIP Snooping

You can enable FIP snooping globally on all FCoE VLANs on a switch or on an individual FCoE VLAN.

By default, FIP snooping is disabled.

To enable FCoE transit on the switch and configure the FCoE transit parameters on ports, follow these steps.

1. Configure FCoE.
   FCoE configuration:
   ```
   copy flash:/ CONFIG_TEMPLATE/ FCoE_DCB_Config running-config
   ```
   The configuration files are stored in the flash memory in the CONFIG_TEMPLATE file.
   NOTE: DCB/DCBx is enabled when either of these configurations is applied.
2. Save the configuration on the switch.
   EXEC Privilege mode.
   ```
   write memory
   ```
3. Reload the switch to enable the configuration.
   EXEC Privilege mode.
   ```
   reload
   ```
   After the switch is reloaded, DCB/DCBx is enabled.
4. Enable the FCoE transit feature on a switch.
   CONFIGURATION mode.
   ```
   feature fip-snooping
   ```
5. Enable FIP snooping on all VLANs or on a specified VLAN.
   CONFIGURATION mode or VLAN INTERFACE mode.
   ```
   fip-snooping enable
   ```
6. Configure the port for bridge-to-FCF links.
   INTERFACE mode or CONFIGURATION mode
   ```
   fip-snooping port-mode fcf
   ```

NOTE: To disable the FCoE transit feature or FIP snooping on VLANs, use the `no` version of a command; for example, `no feature fip-snooping` or `no fip-snooping enable`.

Displaying FIP Snooping Information

Use the following `show` commands to display information on FIP snooping.

Table 29. Displaying FIP Snooping Information

Command	Output
`show fip-snooping sessions [interface vlan vlan-id]`	Displays information on FIP-snooped sessions on all VLANs or a specified VLAN, including the ENode interface and MAC address, the FCF interface and MAC address, VLAN ID, FCoE MAC address and FCoE session ID number (FC-ID), worldwide node name (WWNN) and the worldwide port name (WWPN).
`show fip-snooping config`	Displays the FIP snooping status and configured FC-MAP values.
`show fip-snooping enode [enode-mac-address]`	Displays information on the ENodes in FIP-snooped sessions, including the ENode interface and MAC address, FCF MAC address, VLAN ID and FC-ID.
`show fip-snooping fcf [fcf-mac-address]`	Displays information on the FCFs in FIP-snooped sessions, including the FCF interface and MAC address, FCF interface, VLAN ID, FC-MAP value, FKA advertisement period, and number of ENodes connected.
`clear fip-snooping database interface vlan vlan-id {fcoe-mac-address | enode-mac-address | fcf-mac-address}`	Clears FIP snooping information on a VLAN for a specified FCoE MAC address, ENode MAC address, or FCF MAC address, and removes the corresponding ACLs generated by FIP snooping.
`show fip-snooping statistics [interface vlan vlan-id| interface port-type port/slot | interface port-channel port-channel-number]`	Displays statistics on the FIP packets snooped on all interfaces, including VLANs, physical ports, and port channels.
`clear fip-snooping statistics [interface vlan vlan-id | interface port-type port/slot | interface port-channel port-channel-number]`	Clears the statistics on the FIP packets snooped on all VLANs, a specified VLAN, or a specified port interface.
`show fip-snooping system`	Displays information on the status of FIP snooping on the switch (enabled or disabled), including the number of FCoE VLANs, FCFs, ENodes, and currently active sessions.
`show fip-snooping vlan`	Displays information on the FCoE VLANs on which FIP snooping is enabled.

The following example shows the `show fip-snooping sessions` command.

```
DellEMC#show fip-snooping sessions
Enode MAC          Enode Intf  FCF MAC            FCF Intf  VLAN
aa:bb:cc:00:00:00  Te 1/42     aa:bb:cd:00:00:00  Te 1/43   100
aa:bb:cc:00:00:00  Te 1/42     aa:bb:cd:00:00:00  Te 1/43   100
aa:bb:cc:00:00:00  Te 1/42     aa:bb:cd:00:00:00  Te 1/43   100
aa:bb:cc:00:00:00  Te 1/42     aa:bb:cd:00:00:00  Te 1/43   100
aa:bb:cc:00:00:00  Te 1/42     aa:bb:cd:00:00:00  Te 1/43   100

FCoE MAC          FC-ID          Port WWPN                 Port WWNN
0e:fc:00:01:00:01 01:00:01 31:00:0e:fc:00:00:00:00 21:00:0e:fc:00:00:00:00
0e:fc:00:01:00:02 01:00:02 41:00:0e:fc:00:00:00:00 21:00:0e:fc:00:00:00:00
0e:fc:00:01:00:03 01:00:03 41:00:0e:fc:00:00:00:01 21:00:0e:fc:00:00:00:00
0e:fc:00:01:00:04 01:00:04 41:00:0e:fc:00:00:00:02 21:00:0e:fc:00:00:00:00
0e:fc:00:01:00:05 01:00:05 41:00:0e:fc:00:00:00:03 21:00:0e:fc:00:00:00:00
```

The following table describes the `show fip-snooping sessions` command fields.

Table 30. `show fip-snooping sessions` Command Description

Field	Description
ENode MAC	MAC address of the ENode .
ENode Interface	Slot/port number of the interface connected to the ENode.
FCF MAC	MAC address of the FCF.
FCF Interface	Slot/port number of the interface to which the FCF is connected.
VLAN	VLAN ID number used by the session.
FCoE MAC	MAC address of the FCoE session assigned by the FCF.
FC-ID	Fibre Channel ID assigned by the FCF.
Port WWPN	Worldwide port name of the CNA port.
Port WWNN	Worldwide node name of the CNA port.

The following example shows the `show fip-snooping config` command.

```
DellEMC# show fip-snooping config
FIP Snooping Feature enabled Status: Enabled
FIP Snooping Global enabled Status: Enabled
Global FC-MAP Value: 0X0EFC00

FIP Snooping enabled VLANs
VLAN   Enabled  FC-MAP
----   -------  --------
100    TRUE     0X0EFC00
```

The following example shows the `show fip-snooping enode` command.

```
DellEMC# show fip-snooping enode
Enode MAC          Enode Interface FCF MAC          VLAN  FC-ID
---------          --------------- -------          ----  -----
d4:ae:52:1b:e3:cd  Te 1/11         54:7f:ee:37:34:40 100   62:00:11
```

The following table describes the `show fip-snooping enode` command fields.

Table 31. `show fip-snooping enode` Command Description

Field	Description
ENode MAC	MAC address of the ENode.
ENode Interface	Slot/port number of the interface connected to the ENode.
FCF MAC	MAC address of the FCF.
VLAN	VLAN ID number used by the session.
FC-ID	Fibre Channel session ID assigned by the FCF.

The following example shows the `show fip-snooping fcf` command.

```
DellEMC# show fip-snooping fcf
FCF MAC           FCF Interface VLAN FC-MAP   FKA_ADV_PERIOD No. of Enodes
-------           ------------- ---- ------   -------------- -------------
54:7f:ee:37:34:40 Po 22         100  0e:fc:00 4000           2
```

The following table describes the `show fip-snooping fcf` command fields.

Table 32. `show fip-snooping fcf` Command Description

Field	Description
FCF MAC	MAC address of the FCF.

Table 32. `show fip-snooping fcf` Command Description (continued)

Field	Description
FCF Interface	Slot/port number of the interface to which the FCF is connected.
VLAN	VLAN ID number used by the session.
FC-MAP	FC-Map value advertised by the FCF.
ENode Interface	Slot/port number of the interface connected to the ENode.
FKA_ADV_PERIOD	Period of time (in milliseconds) during which FIP keep-alive advertisements are transmitted.
No of ENodes	Number of ENodes connected to the FCF.
FC-ID	Fibre Channel session ID assigned by the FCF.

The following example shows the `show fip-snooping statistics interface vlan` command (VLAN and port).

```
DellEMC# show fip-snooping statistics interface vlan 100
Number of Vlan Requests                          :0
Number of Vlan Notifications                     :0
Number of Multicast Discovery Solicits           :2
Number of Unicast Discovery Solicits             :0
Number of FLOGI                                  :2
Number of FDISC                                  :16
Number of FLOGO                                  :0
Number of Enode Keep Alive                       :9021
Number of VN Port Keep Alive                     :3349
Number of Multicast Discovery Advertisement      :4437
Number of Unicast Discovery Advertisement        :2
Number of FLOGI Accepts                          :2
Number of FLOGI Rejects                          :0
Number of FDISC Accepts                          :16
Number of FDISC Rejects                          :0
Number of FLOGO Accepts                          :0
Number of FLOGO Rejects                          :0
Number of CVL                                    :0
Number of FCF Discovery Timeouts                 :0
Number of VN Port Session Timeouts               :0
Number of Session failures due to Hardware Config :0
DellEMC(conf)#

DellEMC# show fip-snooping statistics int tengigabitethernet 1/11
Number of Vlan Requests                          :1
Number of Vlan Notifications                     :0
Number of Multicast Discovery Solicits           :1
Number of Unicast Discovery Solicits             :0
Number of FLOGI                                  :1
Number of FDISC                                  :16
Number of FLOGO                                  :0
Number of Enode Keep Alive                       :4416
Number of VN Port Keep Alive                     :3136
Number of Multicast Discovery Advertisement      :0
Number of Unicast Discovery Advertisement        :0
Number of FLOGI Accepts                          :0
Number of FLOGI Rejects                          :0
Number of FDISC Accepts                          :0
Number of FDISC Rejects                          :0
Number of FLOGO Accepts                          :0
Number of FLOGO Rejects                          :0
Number of CVL                                    :0
Number of FCF Discovery Timeouts                 :0
Number of VN Port Session Timeouts               :0
Number of Session failures due to Hardware Config :0
```

The following example shows the `show fip-snooping statistics port-channel` command.

```
DellEMC# show fip-snooping statistics interface port-channel 22
Number of Vlan Requests                          :0
Number of Vlan Notifications                     :2
Number of Multicast Discovery Solicits           :0
Number of Unicast Discovery Solicits             :0
Number of FLOGI                                  :0
Number of FDISC                                  :0
Number of FLOGO                                  :0
Number of Enode Keep Alive                       :0
Number of VN Port Keep Alive                     :0
Number of Multicast Discovery Advertisement      :4451
Number of Unicast Discovery Advertisement        :2
Number of FLOGI Accepts                          :2
Number of FLOGI Rejects                          :0
Number of FDISC Accepts                          :16
Number of FDISC Rejects                          :0
Number of FLOGO Accepts                          :0
Number of FLOGO Rejects                          :0
Number of CVL                                    :0
Number of FCF Discovery Timeouts                 :0
Number of VN Port Session Timeouts               :0
Number of Session failures due to Hardware Config :0
```

The following table describes the `show fip-snooping statistics` command fields.

Table 33. `show fip-snooping statistics` Command Descriptions

Field	Description
Number of VLAN Requests	Number of FIP-snooped VLAN request frames received on the interface.
Number of VLAN Notifications	Number of FIP-snooped VLAN notification frames received on the interface.
Number of Multicast Discovery Solicits	Number of FIP-snooped multicast discovery solicit frames received on the interface.
Number of Unicast Discovery Solicits	Number of FIP-snooped unicast discovery solicit frames received on the interface.
Number of FLOGI	Number of FIP-snooped FLOGI request frames received on the interface.
Number of FDISC	Number of FIP-snooped FDISC request frames received on the interface.
Number of FLOGO	Number of FIP-snooped FLOGO frames received on the interface.
Number of ENode Keep Alives	Number of FIP-snooped ENode keep-alive frames received on the interface.
Number of VN Port Keep Alives	Number of FIP-snooped VN port keep-alive frames received on the interface.
Number of Multicast Discovery Advertisements	Number of FIP-snooped multicast discovery advertisements received on the interface.
Number of Unicast Discovery Advertisements	Number of FIP-snooped unicast discovery advertisements received on the interface.
Number of FLOGI Accepts	Number of FIP FLOGI accept frames received on the interface.
Number of FLOGI Rejects	Number of FIP FLOGI reject frames received on the interface.
Number of FDISC Accepts	Number of FIP FDISC accept frames received on the interface.

Table 33. `show fip-snooping statistics` Command Descriptions (continued)

Field	Description
Number of FDISC Rejects	Number of FIP FDISC reject frames received on the interface.
Number of FLOGO Accepts	Number of FIP FLOGO accept frames received on the interface.
Number of FLOGO Rejects	Number of FIP FLOGO reject frames received on the interface.
Number of CVLs	Number of FIP clear virtual link frames received on the interface.
Number of FCF Discovery Timeouts	Number of FCF discovery timeouts that occurred on the interface.
Number of VN Port Session Timeouts	Number of VN port session timeouts that occurred on the interface.
Number of Session failures due to Hardware Config	Number of session failures due to hardware configuration that occurred on the interface.

The following example shows the `show fip-snooping system` command.

```
DellEMC# show fip-snooping system
Global Mode                   : Enabled
FCOE VLAN List (Operational) : 1, 100
FCFs                          : 1
Enodes                        : 2
Sessions                      : 17
```

The following example shows the `show fip-snooping vlan` command.

```
DellEMC# show fip-snooping vlan
* = Default VLAN

VLAN  FC-MAP    FCFs  Enodes  Sessions
----  ------    ----  ------  --------
*1    -         -     -       -
100   0X0EFC00  1     2       17
```

FCoE Transit Configuration Example

The following illustration shows a switch used as a FIP snooping bridge for FCoE traffic between an ENode (server blade) and an FCF (ToR switch). The ToR switch operates as an FCF and FCoE gateway.

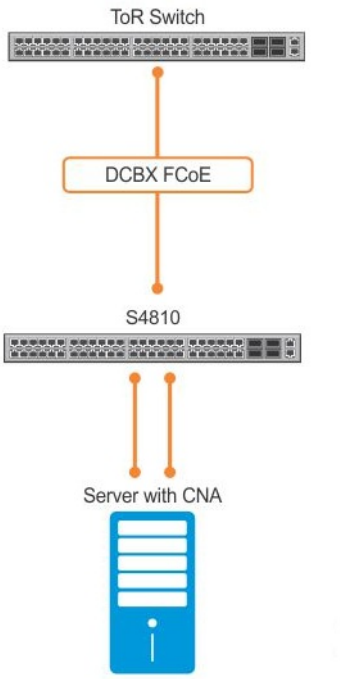


Figure 41. Configuration Example: FIP Snooping on a Switch

In this example, DCBx and PFC are enabled on the FIP snooping bridge and on the FCF ToR switch. On the FIP snooping bridge, DCBx is configured as follows:

- A server-facing port is configured for DCBx in an auto-downstream role.
- An FCF-facing port is configured for DCBx in an auto-upstream or configuration-source role.

The DCBx configuration on the FCF-facing port is detected by the server-facing port and the DCB PFC configuration on both ports is synchronized. For more information about how to configure DCBx and PFC on a port, refer to the Data Center Bridging (DCB) chapter.

The following example shows how to configure FIP snooping on FCoE VLAN 10, on an FCF-facing port (1/5), on an ENode server-facing port (1/1), and to configure the FIP snooping ports as tagged members of the FCoE VLAN enabled for FIP snooping.

Example of Enabling the FIP Snooping Feature on the Switch (FIP Snooping Bridge)

```
DellEMC(conf)# feature fip-snooping
```

Example of Enabling FIP Snooping on the FCoE VLAN

```
DellEMC(conf)# interface vlan 10
DellEMC(conf-if-vl-10)# fip-snooping enable
```

Example of Enabling an FC-MAP Value on a VLAN

```
DellEMC(conf-if-vl-10)# fip-snooping fc-map 0xOEFC01
```

NOTE: Configuring an FC-MAP value is only required if you do not use the default FC-MAP value (0x0EFC00).

Example of Configuring the ENode Server-Facing Port

```
DellEMC(conf)# interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)# portmode hybrid
DellEMC(conf-if-te-1/1)# switchport
DellEMC(conf-if-te-1/1)# protocol lldp
DellEMC(conf-if-te-1/1-lldp)# dcbx port-role auto-downstream
```

NOTE: A port is enabled by default for bridge-ENode links.

Example of Configuring the FCF-Facing Port

```
DellEMC(conf)# interface tengigabitethernet 1/5
DellEMC(conf-if-te-1/5)# portmode hybrid
DellEMC(conf-if-te-1/5)# switchport
DellEMC(conf-if-te-1/5)# fip-snooping port-mode fcf
DellEMC(conf-if-te-1/5)# protocol lldp
DellEMC(conf-if-te-1/5-lldp)# dcbx port-role auto-upstream
```

Example of Configuring FIP Snooping Ports as Tagged Members of the FCoE VLAN

```
DellEMC(conf)# interface vlan 10
DellEMC(conf-if-vl-10)# tagged tengigabitethernet 1/1
DellEMC(conf-if-vl-10)# tagged tengigabitethernet 1/5
DellEMC(conf-if-te-1/1)# no shut
DellEMC(conf-if-te-1/5)# no shut
DellEMC(conf-if-vl-10)# no shut
```

After FIP packets are exchanged between the ENode and the switch, a FIP snooping session is established. ACLs are dynamically generated for FIP snooping on the FIP snooping bridge/switch.

16

FIPS Cryptography

Federal information processing standard (FIPS) cryptography provides cryptographic algorithms conforming to various FIPS standards published by the National Institute of Standards and Technology (NIST), a non-regulatory agency of the US Department of Commerce. FIPS mode is also validated for numerous platforms to meet the FIPS-140-2 standard for a software-based cryptographic module.

This chapter describes how to enable FIPS cryptography requirements on Dell EMC Networking platforms.

NOTE: The Dell EMC Networking OS uses an embedded FIPS 140-2-validated cryptography module (Certificate #1747) running on NetBSD 5.1 per FIPS 140-2 Implementation Guidance section G.5 guidelines.

NOTE: Only the following features use the embedded FIPS 140-2-validated cryptography module:

- SSH Client
- SSH Server
- RSA Host Key Generation
- SCP File Transfers

Currently, other features using cryptography do not use the embedded FIPS 140-2-validated cryptography module.

Topics:

- Configuration Tasks
- Preparing the System
- Enabling FIPS Mode
- Generating Host-Keys
- Monitoring FIPS Mode Status
- Disabling FIPS Mode

Configuration Tasks

To enable FIPS cryptography, complete the following configuration tasks.

- Preparing the System
- Enabling FIPS Mode
- Generating Host-Keys
- Monitoring FIPS Mode Status
- Disabling FIPS Mode

Preparing the System

Before you enable FIPS mode, Dell EMC Networking recommends making the following changes to your system.

1. Disable the Telnet server (only use secure shell [SSH] to access the system).
2. Disable the FTP server (only use secure copy [SCP] to transfer files to and from the system).
3. Attach a secure, standalone host to the console port for the FIPS configuration to use.

Enabling FIPS Mode

To enable or disable FIPS mode, use the console port.

Secure the host attached to the console port against unauthorized access. Any attempts to enable or disable FIPS mode from a virtual terminal session are denied.

When you enable FIPS mode, the following actions are taken:

- If enabled, the SSH server is disabled.
- All open SSH and Telnet sessions, as well as all SCP and FTP file transfers, are closed.
- Any existing host keys (both RSA and RSA1) are deleted from system memory and NVRAM storage.
- FIPS mode is enabled.
  - If you enable the SSH server when you enter the `fips mode enable` command, it is re-enabled for version 2 *only*.
  - If you re-enable the SSH server, a new RSA host key-pair is generated automatically. You can also manually create this key-pair using the `crypto key generate` command.

NOTE: Under certain unusual circumstances, it is possible for the `fips enable` command to indicate a failure.

- This failure occurs if any of the self-tests fail when you enable FIPS mode.
- This failure occurs if there were existing SSH/Telnet sessions that could not be closed successfully in a reasonable amount of time. In general, this failure can occur if a user at a remote host is in the process of establishing an SSH session to the local system, and has been prompted to accept a new host key or to enter a password, but is not responding to the request. Assuming this failure is a transient condition, attempting to enable FIPS mode again should be successful.

To enable FIPS mode, use the following command.

- Enable FIPS mode from a console port.

  CONFIGURATION

  ```
  fips mode enable
  ```

  The following warning message displays:

  ```
  WARNING: Enabling FIPS mode will close all SSH/Telnet connections, restart those servers,
  and destroy all configured host keys. Proceed (y/n) ?
  ```

Generating Host-Keys

The following describes hot-key generation.

When you enable or disable FIPS mode, the system deletes the current public/private host-key pair, terminates any SSH sessions that are in progress (deleting all the per-session encryption key information), actually enables/tests FIPS mode, generates new host-keys, and re-enables the SSH server (assuming it was enabled before enabling FIPS).

For more information, refer to the *SSH Server and SCP Commands* section in the *Security* chapter of the *Dell EMC Networking OS Command Line Reference Guide*.

Monitoring FIPS Mode Status

To view the status of the current FIPS mode (enabled/disabled), use the following commands.

- Use either command to view the status of the current FIPS mode.

  ```
  show fips status
  ```

  ```
  show system
  ```

The following example shows the `show fips status` command.

```
DellEMC#show fips status
FIPS Mode : Enabled
for the system using the show system command.
```

The following example shows the `show system` command.

Disabling FIPS Mode

When you disable FIPS mode, the following changes occur:

- The SSH server disables.
- All open SSH and Telnet sessions, as well as all SCP and FTP file transfers, close.

- Any existing host keys (both RSA and RSA1) are deleted from system memory and NVRAM storage.
- FIPS mode disables.
- The SSH server re-enables.
- The Telnet server re-enables (if it is present in the configuration).
- New 1024–bit RSA and RSA1 host key-pairs are created.

To disable FIPS mode, use the following command.

- To disable FIPS mode from a console port.

  CONFIGURATION mode

  ```
  no fips mode enable
  ```

  The following Warning message displays:

  ```
  WARNING: Disabling FIPS mode will close all SSH/Telnet connections, restart those
  servers, and destroy
  all configured host keys.
  Proceed (y/n) ?
  ```

17

Force10 Resilient Ring Protocol (FRRP)

FRRP provides fast network convergence to Layer 2 switches interconnected in a ring topology, such as a metropolitan area network (MAN) or large campuses. FRRP is similar to what can be achieved with the spanning tree protocol (STP), though even with optimizations, STP can take up to 50 seconds to converge (depending on the size of network and node of failure) and may require 4 to 5 seconds to reconverge. FRRP can converge within 150ms to 1500ms when a link in the ring breaks (depending on network configuration).

To operate a deterministic network, a network administrator must run a protocol that converges independently of the network size or node of failure. FRRP is a proprietary protocol that provides this flexibility, while preventing Layer 2 loops. FRRP provides sub-second ring-failure detection and convergence/re-convergence in a Layer 2 network while eliminating the need for running spanning-tree protocol. With its two-way path to destination configuration, FRRP provides protection against any single link/switch failure and thus provides for greater network uptime.

Topics:

- Protocol Overview
- Implementing FRRP
- FRRP Configuration
- Troubleshooting FRRP
- Sample Configuration and Topology
- FRRP Support on VLT

Protocol Overview

FRRP is built on a ring topology.

You can configure up to 255 rings on a system. FRRP uses one Master node and multiple Transit nodes in each ring. There is no limit to the number of nodes on a ring. The Master node is responsible for the intelligence of the Ring and monitors the status of the Ring. The Master node checks the status of the Ring by sending ring health frames (RHF) around the Ring from its Primary port and returning on its Secondary port. If the Master node misses three consecutive RHFs, the Master node determines the ring to be in a failed state. The Master then sends a Topology Change RHF to the Transit Nodes informing them that the ring has changed. This causes the Transit Nodes to flush their forwarding tables, and re-converge to the new network structure.

One port of the Master node is designated the Primary port (P) to the ring; another port is designated as the Secondary port (S) to the ring. In normal operation, the Master node blocks the Secondary port for all non-control traffic belonging to this FRRP group, thereby avoiding a loop in the ring, like STP. Layer 2 switching and learning mechanisms operate per existing standards on this ring.

Each Transit node is also configured with a Primary port and a Secondary port on the ring, but the port distinction is ignored as long as the node is configured as a Transit node. If the ring is complete, the Master node logically blocks all data traffic in the transmit and receive directions on the Secondary port to prevent a loop. If the Master node detects a break in the ring, it unblocks its Secondary port and allows data traffic to be transmitted and received through it. Refer to the following illustration for a simple example of this FRRP topology. Note that ring direction is determined by the Master node's Primary and Secondary ports.

A virtual LAN (VLAN) is configured on all node ports in the ring. All ring ports must be members of the Member VLAN and the Control VLAN.

The Member VLAN is the VLAN used to transmit data as described earlier.

The Control VLAN is used to perform the health checks on the ring. The Control VLAN can always pass through all ports in the ring, including the secondary port of the Master node.

Ring Status

The ring failure notification and the ring status checks provide two ways to ensure the ring remains up and active in the event of a switch or port failure.

Ring Checking

At specified intervals, the Master node sends a ring health frame (RHF) through the ring. If the ring is complete, the frame is received on its secondary port and the Master node resets its fail-period timer and continues normal operation.

If the Master node does not receive the RHF before the fail-period timer expires (a configurable timer), the Master node moves from the Normal state to the Ring-Fault state and unblocks its Secondary port. The Master node also clears its forwarding table and sends a control frame to all other nodes, instructing them to also clear their forwarding tables. Immediately after clearing its forwarding table, each node starts learning the new topology.

Ring Failure

If a Transit node detects a link down on any of its ports on the FRRP ring, it immediately sends a link-down control frame on the Control VLAN to the Master node.

When the Master node receives this control frame, the Master node moves from the Normal state to the Ring-Fault state and unblocks its Secondary port. The Master node clears its routing table and sends a control frame to all other ring nodes, instructing them to clear their routing tables as well. Immediately after clearing its routing table, each node begins learning the new topology.

Ring Restoration

The Master node continues sending ring health frames out its primary port even when operating in the Ring-Fault state.

After the ring is restored, the next status check frame is received on the Master node's Secondary port. This causes the Master node to transition back to the Normal state. The Master node then logically blocks non-control frames on the Secondary port, clears its own forwarding table, and sends a control frame to the Transit nodes, instructing them to clear their forwarding tables and re-learn the topology.

During the time between the Transit node detecting that its link is restored and the Master node detecting that the ring is restored, the Master node's Secondary port is still forwarding traffic. This can create a temporary loop in the topology. To prevent this, the Transit node places all the ring ports transiting the newly restored port into a temporary blocked state. The Transit node remembers which port has been temporarily blocked and places it into a pre- forwarding state. When the Transit node in the pre-forwarding state receives the control frame instructing it to clear its routing table, it does so and unblocks the previously blocked ring ports on the newly restored port. Then the Transit node returns to the Normal state.

Multiple FRRP Rings

Up to 255 rings are allowed per system and multiple rings can be run on one system.

More than the recommended number of rings may cause interface instability. You can configure multiple rings with a single switch connection; a single ring can have multiple FRRP groups; multiple rings can be connected with a common link.

The platform supports up to 32 rings on a system (including stacked units).

Member VLAN Spanning Two Rings Connected by One Switch

A member VLAN can span two rings interconnected by a common switch, in a figure-eight style topology.

A switch can act as a Master node for one FRRP group and a Transit for another FRRP group, or it can be a Transit node for both rings.

In the following example, FRRP 101 is a ring with its own Control VLAN, and FRRP 202 has its own Control VLAN running on another ring. A Member VLAN that spans both rings is added as a Member VLAN to both FRRP groups. Switch R3 has two instances of FRRP running on it: one for each ring. The example topology that follows shows R3 assuming the role of a Transit node for both FRRP 101 and FRRP 202.

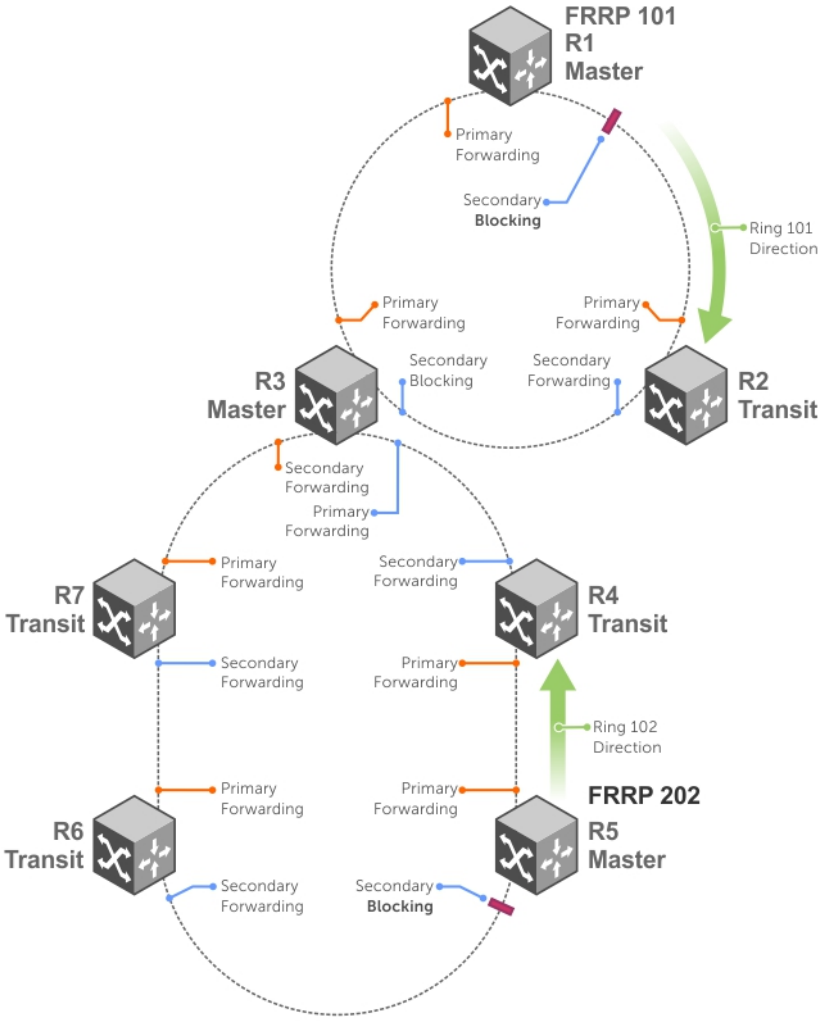


Figure 42. Example of Multiple Rings Connected by Single Switch

Important FRRP Points

FRRP provides a convergence time that can generally range between 150ms and 1500ms for Layer 2 networks.

The Master node originates a high-speed frame that circulates around the ring. This frame, appropriately, sets up or breaks down the ring.

- The Master node transmits ring status check frames at specified intervals.
- You can run multiple physical rings on the same switch.
- One Master node per ring — all other nodes are Transit.
- Each node has two member interfaces — primary and secondary.
- There is no limit to the number of nodes on a ring.
- Master node ring port states — blocking, pre-forwarding, forwarding, and disabled.
- Transit node ring port states — blocking, pre-forwarding, forwarding, and disabled.
- STP disabled on ring interfaces.
- Master node secondary port is in blocking state during Normal operation.
- Ring health frames (RHF)
  - Hello RHF: sent at 500ms (hello interval); Only the Master node transmits and processes these.
  - Topology Change RHF: triggered updates; processed at all nodes.

Important FRRP Concepts

The following table lists some important FRRP concepts.

Concept	Explanation
Ring ID	Each *ring* has a unique 8-bit ring ID through which the ring is identified (for example, FRRP 101 and FRRP 202, as shown in the illustration in Member VLAN Spanning Two Rings Connected by One Switch.
Control VLAN	Each *ring* has a unique Control VLAN through which tagged ring health frames (RHF) are sent. Control VLANs are used only for sending RHF, and cannot be used for any other purpose.
Member VLAN	Each *ring* maintains a list of member VLANs. Member VLANs must be consistent across the entire ring.
Port Role	Each *node* has two ports for each ring: Primary and Secondary. The Master node Primary port generates RHFs. The Master node Secondary port receives the RHFs. On Transit nodes, there is no distinction between a Primary and Secondary interface when operating in the Normal state.
Ring Interface State	Each interface (port) that is part of the ring maintains one of four states" • **Blocking State** — Accepts ring protocol packets but blocks data packets. LLDP, FEFD, or other Layer 2 control packets are accepted. Only the Master node Secondary port can enter this state. • **Pre-Forwarding State** — A transition state before moving to the Forward state. Control traffic is forwarded but data traffic is blocked. The Master node Secondary port transitions through this state during ring bring-up. All ports transition through this state when a port comes up. • **Pre-Forwarding State** — A transition state before moving to the Forward state. Control traffic is forwarded but data traffic is blocked. The Master node Secondary port transitions through this state during ring bring-up. All ports transition through this state when a port comes up. • **Disabled State** — When the port is disabled or down, or is not on the VLAN.
Ring Protocol Timers	• **Hello Interval** — The interval when ring frames are generated from the Master node's Primary interface (default **500 ms**). The Hello interval is configurable in 50 ms increments from 50 ms to 2000 ms. • **Dead Interval** — The interval when data traffic is blocked on a port. The default is three times the Hello interval rate. The dead interval is configurable in 50 ms increments from 50 ms to 6000 ms.
Ring Status	The state of the FRRP ring. During initialization/configuration, the default ring status is Ring-down (disabled). The Primary and Secondary interfaces, control VLAN, and Master and Transit node information must be configured for the ring to be up. • **Ring-Up** — Ring is up and operational. • **Ring-Down** — Ring is broken or not set up.
Ring Health-Check Frame (RHF)	The Master node generates two types of RHFs. RHFs never loop the ring because they terminate at the Master node's secondary port. • **Hello RHF (HRHF)** — These frames are processed only on the Master node's Secondary port. The Transit nodes pass the HRHF through without processing it. An HRHF is sent at every Hello interval. • **Topology Change RHF (TCRHF)** — These frames contains ring status, keepalive, and the control and member VLAN hash. The TCRHF is processed at each node of the ring. TCRHFs are sent out the Master Node's Primary and Secondary interface when the ring is declared in a Failed state with the same sequence number, on any topology change to ensure that all Transit nodes receive it. There is no periodic transmission of TCRHFs. The TCRHFs are sent on triggered events of ring failure or ring restoration only.

Implementing FRRP

- FRRP is media and speed independent.
- FRRP is a Dell proprietary protocol that does not interoperate with any other vendor.
- You must disable the spanning tree protocol (STP) on both the Primary and Secondary interfaces before you can enable FRRP.
- All ring ports must be Layer 2 ports. This is required for both Master and Transit nodes.
- A VLAN configured as a control VLAN for a ring cannot be configured as a control or member VLAN for any other ring.
- The control VLAN is not used to carry any data traffic; it carries only RHFs.
- The control VLAN cannot have members that are not ring ports.

- If multiple rings share one or more member VLANs, they cannot share any links between them.
- Member VLANs across multiple rings are not supported in Master nodes.
- Each ring has only one Master node; all others are transit nodes.

FRRP Configuration

These are the tasks to configure FRRP.

- Creating the FRRP Group
- Configuring the Control VLAN
  - Configure Primary and Secondary ports
- Configuring and Adding the Member VLANs
  - Configure Primary and Secondary ports

Other FRRP related commands are:

- Clearing the FRRP Counters
- Viewing the FRRP Configuration
- Viewing the FRRP Information

Creating the FRRP Group

Create the FRRP group on each switch in the ring.

To create the FRRP group, use the command.

- Create the FRRP group with this Ring ID.

  CONFIGURATION mode

  `protocol frrp ring-id`

  Ring ID: the range is from 1 to 255.

Configuring the Control VLAN

Control and member VLANS are configured normally for Layer 2. Their status as control or member is determined at the FRRP group commands.

For more information about configuring VLANS in Layer 2 mode, refer to Layer 2.

Be sure to follow these guidelines:

- All VLANS must be in Layer 2 mode.
- You can only add ring nodes to the VLAN.
- A control VLAN can belong to one FRRP group only.
- Tag control VLAN ports.
- All ports on the ring must use the same VLAN ID for the control VLAN.
- You cannot configure a VLAN as both a control VLAN and member VLAN on the same ring.
- Only two interfaces can be members of a control VLAN (the Master Primary and Secondary ports).
- Member VLANs across multiple rings are not supported in Master nodes.

To create the control VLAN for this FRRP group, use the following commands on the switch that is to act as the Master node.

1. Create a VLAN with this ID number.

   CONFIGURATION mode.

   `interface vlan vlan-id`

   VLAN ID: from 1 to 4094.
2. Tag the specified interface or range of interfaces to this VLAN.

   CONFIG-INT-VLAN mode.

   `tagged interface {range}`

   Interface:

   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

3. Assign the Primary and Secondary ports and the control VLAN for the ports on the ring.
   CONFIG-FRRP mode.
   `interface primary interface secondary interface control-vlan vlan id`
   Interface:
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
4. Configure the Master node.
   CONFIG-FRRP mode.
   `mode master`
5. Identify the Member VLANs for this FRRP group.
   CONFIG-FRRP mode.
   `member-vlan vlan-id {range}`
   VLAN-ID, Range: VLAN IDs for the ring's member VLANS.
6. Enable FRRP.
   CONFIG-FRRP mode.
   `no disable`

Configuring and Adding the Member VLANs

Control and member VLANS are configured normally for Layer 2. Their status as Control or Member is determined at the FRRP group commands.

For more information about configuring VLANS in Layer 2 mode, refer to the Layer 2 chapter.

Be sure to follow these guidelines:

- All VLANS must be in Layer 2 mode.
- Tag control VLAN ports. Member VLAN ports, except the Primary/Secondary interface, can be tagged or untagged.
- The control VLAN must be the same for all nodes on the ring.

To create the Members VLANs for this FRRP group, use the following commands on all of the Transit switches in the ring.

1. Create a VLAN with this ID number.
   CONFIGURATION mode.
   `interface vlan vlan-id`
   VLAN ID: the range is from 1 to 4094.
2. Tag the specified interface or range of interfaces to this VLAN.
   CONFIG-INT-VLAN mode.
   `tagged interface {range}`
   Interface:
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
3. Assign the Primary and Secondary ports and the Control VLAN for the ports on the ring.
   CONFIG-FRRP mode.
   `interface primary interface secondary interface control-vlan vlan id`
   Interface:
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

   VLAN ID: Identification number of the Control VLAN.
4. Configure a Transit node.
   CONFIG-FRRP mode.
   `mode transit`

5. Identify the Member VLANs for this FRRP group.
   CONFIG-FRRP mode.
   `member-vlan vlan-id {range}`
   VLAN-ID, Range: VLAN IDs for the ring's Member VLANs.
6. Enable this FRRP group on this switch.
   CONFIG-FRRP mode.
   `no disable`

Setting the FRRP Timers

To set the FRRP timers, use the following command.

NOTE: Set the Dead-Interval time 3 times the Hello-Interval.

- Enter the desired intervals for Hello-Interval or Dead-Interval times.
  CONFIG-FRRP mode.
  `timer {hello-interval|dead-interval} milliseconds`
  - *Hello-Interval*: the range is from 50 to 2000, in increments of 50 (default is **500**).
  - *Dead-Interval*: the range is from 50 to 6000, in increments of 50 (default is **1500**).

Clearing the FRRP Counters

To clear the FRRP counters, use one of the following commands.

- Clear the counters associated with this Ring ID.
  EXEC PRIVELEGED mode.
  `clear frrp ring-id`
  Ring ID: the range is from 1 to 255.
- Clear the counters associated with all FRRP groups.
  EXEC PRIVELEGED mode.
  `clear frrp`

Viewing the FRRP Configuration

To view the configuration for the FRRP group, use the following command.

- Show the configuration for this FRRP group.
  CONFIG-FRRP mode.
  `show configuration`

Viewing the FRRP Information

To view general FRRP information, use one of the following commands.

- Show the information for the identified FRRP group.
  EXEC or EXEC PRIVELEGED mode.
  `show frrp ring-id`
  Ring ID: the range is from 1 to 255.
- Show the state of all FRRP groups.
  EXEC or EXEC PRIVELEGED mode.
  `show frrp summary`

Ring ID: the range is from 1 to 255.

Troubleshooting FRRP

To troubleshoot FRRP, use the following information.

Configuration Checks

- Each Control Ring must use a unique VLAN ID.
- Only two interfaces on a switch can be Members of the same control VLAN.
- There can be only one Master node for any FRRP group.
- You can configure FRRP on Layer 2 interfaces only.
- Spanning Tree (if you enable it globally) must be disabled on both Primary and Secondary interfaces when you enable FRRP.
  - When the interface ceases to be a part of any FRRP process, if you enable Spanning Tree globally, also enable it explicitly for the interface.
- The maximum number of rings allowed on a chassis is 255.

Sample Configuration and Topology

The following example shows a basic FRRP topology.

Example of R1 MASTER

```
interface TenGigabitEthernet 1/24
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 1/34
  no ip address
  switchport
  no shutdown
!
interface Vlan 101
  no ip address
  tagged TenGigabitEthernet 1/24,34
  no shutdown
!
interface Vlan 201
  no ip address
  tagged TenGigabitEthernet 1/24,34
  no shutdown

!
protocol frrp 101
  interface primary TenGigabitEthernet 1/24
secondary TenGigabitEthernet 1/34 control-vlan 101
  member-vlan 201
  mode master
  no disable
```

Example of R2 TRANSIT

```
interface TenGigabitEthernet 2/14
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 2/31
  no ip address
  switchport
  no shutdown
!
```

```
interface Vlan 101
  no ip address
  tagged TenGigabitEthernet 2/14,31
  no shutdown
!
interface Vlan 201
  no ip address
  tagged TenGigabitEthernet 2/14,31
  no shutdown
!
protocol frrp 101
  interface primary TenGigabitEthernet 2/14 secondary TenGigabitEthernet 2/31 control-
vlan 101
  member-vlan 201
  mode transit
  no disable
```

Example of R3 TRANSIT

```
interface TenGigabitEthernet 3/14
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 3/21
  no ip address
  switchport
  no shutdown
!
interface Vlan 101
  no ip address
  tagged TenGigabitEthernet 3/14,21
  no shutdown
!
interface Vlan 201
  no ip address
  tagged TenGigabitEthernet 3/14,21
  no shutdown

!
protocol frrp 101
  interface primary TenGigabitEthernet 3/21
secondary TenGigabitEthernet 3/14 control-vlan 101
  member-vlan 201
  mode transit
  no disable
```

FRRP Support on VLT

Using FRRP rings, you can inter-connect VLT domains across data centers.

These FRRP rings make use of Layer2 VLANs that spawn across Data Centers and provide resiliency by detecting node or link level failures.

You can configure a simple FRRP ring that connects a VLT device in one data center to a VLT devices in two or more Data Centers.

NOTE: This configuration connects VLT devices across Data Centers using FRRP; however, the VLTi may or may not participate as a ring interface of any FRRP ring.

Following figure shows a simple FRRP ring inter-connecting VLT device:

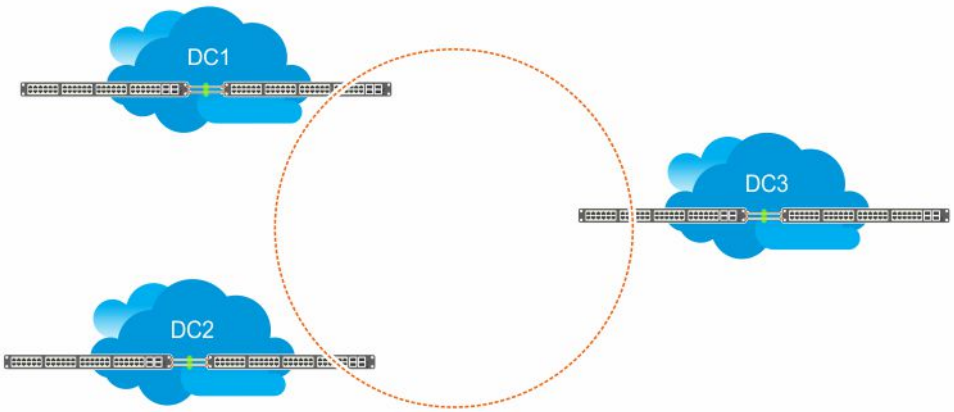


Figure 43. FRRP Ring Connecting VLT Devices

You can also configure an FRRP ring where both the VLT peers are connected to the FRRP ring and the VLTi acts as the primary interface for the FRRP Master and transit nodes.

This active-active FRRP configuration blocks the FRRP ring on a per VLAN or VLAN group basis enabling the configuration to spawn across different set of VLANs. The FRRP configuration where VLTi nodes act as the primary or secondary interfaces ensure that all the optics used to connect VLT domains across data centers are fully utilized.

The primary requirement for the active-active FRRP configuration to work is that the VLTi between two VLT peers must act as the primary interface of the Master and transit nodes of the FRRP ring.

NOTE: As the secondary interface of the FRRP master node is blocked for member VLAN traffic, VLTi cannot act as the secondary interface in an FRRP ring topology.

Example Scenario

Following example scenario describes an Active-Active FRRP ring topology where the ring is blocked on a per VLAN or VLAN group basis allowing active-active FRRP ring for different set of VLANs.

In this scenario, an FRRP ring named R1 is configured with VLT Node1 acting as the Master node and VLT Node2 as the transit node. Simillary, an FRRP ring named R2 is configured with VLT Node2 as the master node and VLT node1 as the trasit node.

In the FRRP ring R1, the primary interface for VLT Node1 is the VLTi. P1 is the secondary interface, which is an orphan port that is participating in the FRRP ring topology. V1 is the control VLAN through which the RFHs are exchanged indicating the health of the nodes and the FRRP ring itself. In addition to the control VLAN, multiple member VLANS are configured (for example, M1 through M10) that carry the data traffic across the FRRP rings. The secondary port P1 is tagged to the control VLAN (V1). VLTi is implicitly tagged to the member VLANs when these VLANs are configured in the VLT peer.

As a result of the VLT Node1 configuration, the FRRP ring R1 becomes active by blocking the secondary interface P1 for the member VLANs (M1 to M10).

VLT Node2 is the transit node. The primary interface for VLT Node2 is VLTi. P2 is the secondary interface, which is one of the orphan port participating in the FRRP ring. V1 is the control VLAN through which the RFHs are exchanged. In addition to the control VLAN, multiple member VLANS are configured (for example, M1 to M10) that carry the data traffic across the FRRP rings. The secondary port P2 is tagged to the control VLAN (V1). VLTi is implicitly tagged to the member VLANs when these VLANs are configured in the VLT peer.

As a result of the VLT Node2 configuration on R2, the primary interface VLTi and the secondary interface P1 act as forwarding ports for the member VLANs (M1 to M10).

In the FRRP ring R2, the primary interface for VLT Node1 (transit node) is the VLTi. P1 is the secondary interface, which is an orphan port that is participating in the FRRP ring topology. V1 is the control VLAN through which the RFHs are exchanged indicating the health of the nodes and the FRRP ring itself. In addition to the control VLAN, multiple member VLANS are configured (for example, M11 through Mn) that carry the data traffic across the FRRP rings. The secondary port P1 is tagged to the control VLAN (V1). VLTi is implicitly tagged to the member VLANs when these VLANs are configured in the VLT peer.

As a result of the VLT Node1 configuration on R2, the FRRP ring R2 becomes active. The primary interface VLTi and the secondary interface P1 act as forwarding ports for the member VLANs (M11 to Mn).

VLT Node2 is the master node. The primary interface for VLT Node2 is VLTi. P2 is the secondary interface, which is one of the orphan port participating in the FRRP ring. V1 is the control VLAN through which the RFHs are exchanged. In addition to the

control VLAN, multiple member VLANS are configured (for example, M1 to M10) that carry the data traffic across the FRRP rings. The secondary port P2 is tagged to the control VLAN (V1). VLTi is implicitly tagged to the member VLANs when these VLANs are configured in the VLT peer.

As a result of the VLT Node2 configuration on R2, the secondary interface P2 is blocked for the member VLANs (M11 to Mn).

Following figure illustrated the FRRP Ring R1 topology:

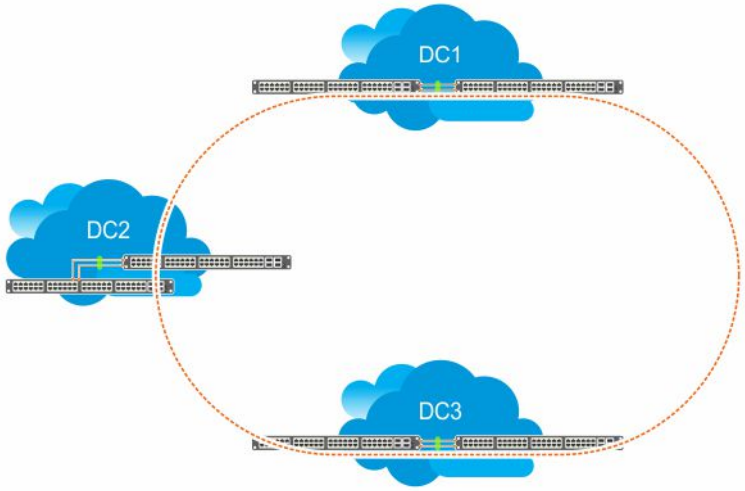


Figure 44. FRRP Ring using VLTi links

Important Points to Remember

- VLTi can be configured only as the primary interface for the primary interface of any FRRP ring.
- Only RSTP and PVST are supported in the VLT environment. Enabling either RSTP or PVST effects FRRP functionality even though these features are disabled on FRRP enabled interfaces.
- Dell EMC Networking OS does not support coexistence of xSTP and FRRP configurations. Meaning, if there is any active FRRP ring in the system, then you cannot enable xSTP in the system globally or at the interface level. Similarly, if xSTP is enabled, then you cannot configure FRRP in the system.
- You cannot configure VLT LAG interfaces as FRRP ring interfaces.
- When ICL is configured as an FRRP ring interface, you cannot remove ICL and VLT domain configurations.
- When FRRP is enabled on a VLT domain, you cannot enable any flavor of the spanning tree protocol (STP) concurrently on the nodes corresponding to that VLT domain. Meaning, FRRP and xSTP cannot coexist in a VLT environment.

18

GARP VLAN Registration Protocol (GVRP)

The generic attribute registration protocol (GARP) VLAN registration protocol (GVRP), defined by the IEEE 802.1q specification, is a Layer 2 network protocol that provides for automatic VLAN configuration of switches. GVRP-compliant switches use GARP to register and de-register attribute values, such as VLAN IDs, with each other.

Typical virtual local area network (VLAN) implementation involves manually configuring each Layer 2 switch that participates in a given VLAN. GVRP, defined by the IEEE 802.1q specification, is a Layer 2 network protocol that provides for automatic VLAN configuration of switches. GVRP-compliant switches use GARP to register and de-register attribute values, such as VLAN IDs, with each other.

GVRP exchanges network VLAN information to allow switches to dynamically forward frames for one or more VLANs. Therefore, GVRP spreads this information and configures the needed VLANs on any additional switches in the network. Data propagates via the exchange of GVRP protocol data units (PDUs).

The purpose of GVRP is to simplify (but not eliminate) static configuration. The idea is to configure switches at the edge and have the information dynamically propagate into the core. As such, the edge ports must still be statically configured with VLAN membership information, and they do not run GVRP. It is this information that is propagated to create dynamic VLAN membership in the core of the network.

Important Points to Remember

- GVRP propagates VLAN membership throughout a network. GVRP allows end stations and switches to issue and revoke declarations relating to VLAN membership.
- VLAN registration is made in the context of the port that receives the GARP PDU and is propagated to the other active ports.
- GVRP is disabled by default; enable GVRP for the switch and then for individual ports.
- Dynamic VLANs are aged out after the LeaveAll timer expires three times without receipt of a Join message. To display status, use the `show gvrp statistics {interface interface | summary}` command.
- If spanning tree and GVRP are both required, implement the rapid spanning tree protocol (RSTP). The device does not support enabling GVRP and MSTP at the same time.

```
DellEMC(conf)#protocol spanning-tree pvst
DellEMC(conf-pvst)#no disable
% Error: GVRP running. Cannot enable PVST.

.........
DellEMC(conf)#protocol spanning-tree mstp
DellEMC(conf-mstp)#no disable
% Error: GVRP running. Cannot enable MSTP.

.........

DellEMC(conf)#protocol gvrp
DellEMC(conf-gvrp)#no disable
% Error: PVST running. Cannot enable GVRP.
% Error: MSTP running. Cannot enable GVRP.
```

Topics:

- Configure GVRP
- Enabling GVRP Globally
- Enabling GVRP on a Layer 2 Interface
- Configure GVRP Registration
- Configure a GARP Timer
- RPM Redundancy

Configure GVRP

To begin, enable GVRP.

To facilitate GVRP communications, enable GVRP globally on each switch. Then, GVRP configuration is per interface on a switch-by-switch basis. Enable GVRP on each port that connects to a switch where you want GVRP information exchanged. In the following example, GVRP is configured on VLAN trunk ports.

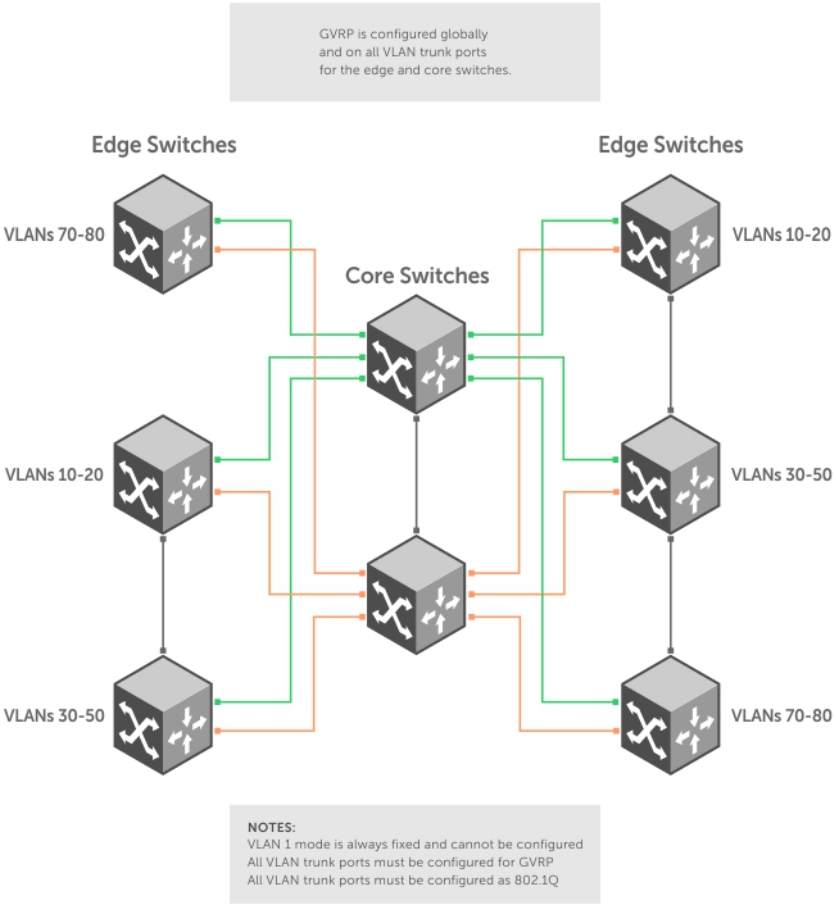


Figure 45. Global GVRP Configuration Example

Basic GVRP configuration is a two-step process:

1. Enabling GVRP Globally
2. Enabling GVRP on a Layer 2 Interface

Related Configuration Tasks

- Configure GVRP Registration
- Configure a GARP Timer

Enabling GVRP Globally

To configure GVRP globally, use the following command.

- Enable GVRP for the entire switch.
  CONFIGURATION mode

gvrp enable

```
DellEMC(conf)#protocol gvrp
DellEMC(config-gvrp)#no disable
DellEMC(config-gvrp)#show config
!
protocol gvrp
no disable
DellEMC(config-gvrp)#
```

To inspect the global configuration, use the `show gvrp brief` command.

Enabling GVRP on a Layer 2 Interface

To enable GVRP on a Layer 2 interface, use the following command.

- Enable GVRP on a Layer 2 interface.

  INTERFACE mode

  `gvrp enable`

```
DellEMC(conf-if-te-1/21)#switchport
DellEMC(conf-if-te-1/21)#gvrp enable
DellEMC(conf-if-te-1/21)#no shutdown
DellEMC(conf-if-te-1/21)#show config
!
interface TenGigabitEthernet 1/21
no ip address
switchport
gvrp enable
no shutdown
```

To inspect the interface configuration, use the `show config` command from INTERFACE mode or use the `show gvrp interface` command in EXEC or EXEC Privilege mode.

Configure GVRP Registration

Configure GVRP registration.

There are two GVRP registration modes:

- **Fixed Registration Mode** — figuring a port in fixed registration mode allows for manual creation and registration of VLANs, prevents VLAN deregistration, and registers all VLANs known on other ports on the port. For example, if an interface is statically configured via the CLI to belong to a VLAN, it should not be unconfigured when it receives a Leave PDU. Therefore, the registration mode on that interface is FIXED.
- **Forbidden Mode** — Disables the port to dynamically register VLANs and to propagate VLAN information except information about VLAN 1. A port with forbidden registration type thus allows only VLAN 1 to pass through even though the PDU carries information for more VLANs. Therefore, if you do not want the interface to advertise or learn about particular VLANS, set the interface to the registration mode of FORBIDDEN.

Based on the configuration in the following example, the interface is not removed from VLAN 34 or VLAN 35 despite receiving a GVRP Leave message. Additionally, the interface is not dynamically added to VLAN 45 or VLAN 46, even if a GVRP Join message is received.

Example of the `gvrp registration` Command

```
DellEMC(conf-if-te-1/21)#gvrp registration fixed 34,35
DellEMC(conf-if-te-1/21)#gvrp registration forbidden 45,46
DellEMC(conf-if-te-1/21)#show conf
!
interface TenGigabitEthernet 1/21
  no ip address
  switchport
  gvrp enable
  gvrp registration fixed 34-35
  gvrp registration forbidden 45-46
```

```
  no shutdown
DellEMC(conf-if-te-1/21)#
```

Configure a GARP Timer

Set GARP timers to the same values on all devices that are exchanging information using GVRP.

There are three GARP timer settings.

- **Join** — A GARP device reliably transmits Join messages to other devices by sending each Join message two times. To define the interval between the two sending operations of each Join message, use this parameter. The Dell EMC Networking OS default is **200ms**.
- **Leave** — When a GARP device expects to de-register a piece of attribute information, it sends out a Leave message and starts this timer. If a Join message does not arrive before the timer expires, the information is de-registered. The Leave timer must be greater than or equal to 3x the Join timer. The Dell EMC Networking OS default is **600ms**.
- **LeaveAll** — After startup, a GARP device globally starts a LeaveAll timer. After expiration of this interval, it sends out a LeaveAll message so that other GARP devices can re-register all relevant attribute information. The device then restarts the LeaveAll timer to begin a new cycle. The LeaveAll timer must be greater than or equal to 5x of the Leave timer. The Dell EMC Networking OS default is **10000ms**.

Example of the `garp timer` Command

```
DellEMC(conf)#garp timer leav 1000
DellEMC(conf)#garp timers leave-all 5000
DellEMC(conf)#garp timer join 300

Verification:

DellEMC(conf)#do show garp timer
GARP Timers Value (milliseconds)
----------------------------------------
Join Timer      300
Leave Timer     1000
LeaveAll Timer  5000
DellEMC(conf)#
```

Dell EMC Networking OS displays this message if an attempt is made to configure an invalid GARP timer:

```
DellEMC(conf)#garp timers join 300 % Error: Leave timer should be >= 3*Join timer.
```

RPM Redundancy

The current version of Dell EMC Networking OS supports 1+1 hitless route processor module (RPM) redundancy.

The primary RPM performs all routing, switching, and control operations while the standby RPM monitors the primary RPM. In the event that the primary RPM fails, the standby RPM can assume control of the system without requiring a chassis reboot.

This section contains the following sub-sections:

- Boot the Chassis with a Single RPM
- Boot the Chassis with Dual RPMs
- Automatic and Manual RPM Failover
- Support for RPM Redundancy by Dell EMC Networking OS Version
- RPM Synchronization

19

High Availability (HA)

High availability (HA) is supported on Dell EMC Networking OS.

HA is a collection of features that preserves system continuity by maximizing uptime and minimizing packet loss during system disruptions.

To support all the features within the HA collection, you should have the latest boot code. The following table lists the boot code requirements as of this Dell EMC Networking OS release.

Table 34. Boot Code Requirements

Component	Boot Code
S4820T	1 2.0.3

The features in this collection are:

- Component Redundancy
- Online Insertion and Removal
- Hitless Behavior
- Graceful Restart
- Software Resiliency
- Hot-Lock Behavior

Topics:

- Component Redundancy
- Online Insertion and Removal
- Hitless Behavior
- Graceful Restart
- Software Resiliency
- Hot-Lock Behavior
- Process Restartability

Component Redundancy

Dell EMC Networking systems eliminate single points of failure by providing dedicated or load-balanced redundancy for each component.

RPM Redundancy

The current version of Dell EMC Networking OS supports 1+1 hitless route processor module (RPM) redundancy.

The primary RPM performs all routing, switching, and control operations while the standby RPM monitors the primary RPM. In the event that the primary RPM fails, the standby RPM can assume control of the system without requiring a chassis reboot.

This section contains the following sub-sections:

- Boot the Chassis with a Single RPM
- Boot the Chassis with Dual RPMs
- Automatic and Manual RPM Failover
- Support for RPM Redundancy by Dell EMC Networking OS Version
- RPM Synchronization

Boot the Chassis with a Single RPM

You can boot the chassis with one RPM and later add a second RPM, which automatically becomes the standby RPM.

Dell Networking OS displays the following message when the standby RPM is online.

```
%RPM-2-MSG:CP0 %POLLMGR-2-ALT_RPM_STATE: Alternate RPM is present
%IRC-6-IRC_COMMUP: Link to peer RPM is up
%RAM-6-RAM_TASK: RPM1 is in Standby State.
```

Boot the Chassis with Dual RPMs

When you boot the system with two RPMs installed, the RPM in slot R0 is the primary RPM by default. Both RPMs must be running the same version of Dell Networking OS.

To configure either RPM to be the primary after the next chassis reboot, use the `redundancy primary` command from CONFIGURATION mode.

Version Compatibility Between RPMs

In general, the two RPMs should have the same Dell Networking OS version. However, Dell Networking OS tolerates some degree of difference between the two versions, as described in the following table. To view the configuration loaded on each RPM, use the `show redundancy` command, as shown in the example in Automatic and Manual RPM Failover.

Table 35. Mismatch Condition Behaviors

Mismatch Condition	Example	Behavior
Different Dell Networking OS versions with only the first two digits matching.	Primary: 7.4.2.0 Standby: 7.4.1.0	The link to the standby RPM is up and Dell Networking OS block syncs only the startup-config. The failover type is warm upgrade. Dell Networking OS displays the following message.
Different Dell Networking OS versions with the first two digits not matching.	Primary: 7.6.1.0 Standby: 7.5.1.0	The link to the standby RPM is down and the standby RPM is in a boot loop. Dell Networking OS displays the second message and a boot fail prompt.
Different Dell Networking OS versions with only the first three digits matching.	Primary: 7.4.2.0 Standby: 7.4.2.1	The link to the peer RPM is up and Dell Networking OS performs a complete block sync. The failover type is hot failover. Dell Networking OS displays the following message.

Example of Dell Networking OS Version Incompatibility Error

```
*********************************************
*
* Warning !!! Warning !!! Warning !!!
*
* -------------------------------------------
*
* Incompatible SW Version detected !!
*
* This RPM -> 7.4.2.0
* Peer RPM -> 7.4.1.0
*
*********************************************

00:00:12: %RPM0-U:CP %IRC-4-IRC_VERSION: Current RPM 7.4.2.0 Peer RPM 7.4.1.0 -
Different software version detected
00:00:12: %RPM0-U:CP %IRC-6-IRC_COMMUP: Link to peer RPM is up
00:00:14: %RPM0-U:CP %RAM-6-ELECTION_ROLE: RPM0 is transitioning to Primary RPM.
```

Example of Boot Failure on Standby RPM

```
System failed to boot up. Please reboot the chassis !!!
00:12:46: %RPM1-U:CP %TME-0-RPM BRINGUP FAIL: FTOS failed to bring up the system
    Communication between RPMs is not up, check the software version and reboot chassis.

Dell(standby)(bootfail)#
```

Automatic and Manual Stack Unit Failover

Stack unit failover is the process of the standby unit becoming a management unit.

Dell EMC Networking OS fails over to the standby stack unit when:

1. Communication is lost between the standby and primary stack unit.
2. You request a failover via the CLI.

To display the reason for the last failover, use the `show redundancy` command from EXEC Privilege mode.

Example of the `show redundancy` Command

```
DellEMC#show redundancy

-- Stack-unit Status --
-------------------------------------------------
  Mgmt ID:                                     0
  Stack-unit ID:                               0
  Stack-unit Redundancy Role:     Primary
  Stack-unit State:                    Active
  Stack-unit SW Version:             9.6(0.0)
  Link to Peer:                           Down
    Peer Stack-unit:                      not present

-- Stack-unit Redundancy Configuration --
-------------------------------------------------
  Primary Stack-unit:                   mgmt-id        0
  Auto Data Sync:                       Full
  Failover Type:                          Hot Failover
  Auto reboot Stack-unit:            Enabled
  Auto failover limit:                  3 times in 60 minutes

-- Stack-unit Failover Record --
-------------------------------------------------
  Failover Count:                       0
  Last failover timestamp:            None
  Last failover Reason:                 None
  Last failover type:                   None

-- Last Data Block Sync Record: --
-------------------------------------------------
  Stack Unit Config:                    no block sync done
    Start-up Config:          no block sync done
  Runtime Event Log:          no block sync done
     Running Config:          no block sync done
             ACL Mgr:         no block sync done
                LACP:         no block sync done
                 STP:         no block sync done
                SPAN:         no block sync done

DellEMC#
```

Support for RPM Redundancy by Dell EMC Networking OS Version

Dell EMC Networking OS supports increasing levels of RPM redundancy (warm and hot) as described in the table below.

Table 36. Support for RPM Redundancy by Dell EMC Networking OS Version

Platform	Failover Type	Failover Behavior
S4820T	Hot Failover	Only the failed RPM reboots. All the line cards and SFMs remain online. All application tasks are spawned on the secondary RPM before failover. The running configuration is synchronized at runtime so it does not need to be reapplied during failover.

Synchronization between Management and Standby Units

Data between the Management and Standby units is synchronized immediately after bootup.

After the Management and Standby units have done an initial full synchronization (block sync), Dell EMC Networking OS only updates changed data (incremental sync). The data that is synchronized consists of configuration data, operational data, state and status, and statistics depending on the Dell EMC Networking OS version.

Configuring RPM Redundancy

To select a primary RPM, use the following command. The RPM in slot 0 is the primary RPM by default.

- Manually select the primary RPM.
  CONFIGURATION mode
  `redundancy primary`

```
Dell#show running-config redundancy
!
redundancy auto-failover-limit count 3 period 60
redundancy auto-synchronize full
redundancy primary rpm0
Dell#
```

To view which RPM is the primary, use the `show running-config redundancy` command from EXEC Privilege mode, as shown in the example in the Forcing an RPM Failover.

Forcing a Stack Unit Failover

To force a Stack unit failover, use the following command.

Use this feature when you are replacing a stack unit and when you are performing a warm upgrade.

- To trigger a stack unit failover.
  EXEC Privilege mode
  `redundancy force-failover stack-unit`

```
Dell#redundancy force-failover stack-unit
System configuration has been modified. Save? [yes/no]: yes
Proceed with Stack-unit hot failover [confirm yes/no]:yes
Dell#
```

Disabling Auto-Reboot

To disable auto-reboot, use the following command.

- Prevent a failed stack unit from rebooting after a failover.
  CONFIGURATION mode
  `redundancy disable-auto-reboot`

Manually Synchronizing Management and Standby Units

To manually synchronize Management and Standby units at any time, use the following command.

- Manually synchronize Management and Standby units.
  EXEC Privilege mode
  `redundancy synchronize full`

Online Insertion and Removal

You can add, replace, or remove chassis components while the chassis is operating.

This section contains the following sub-sections:

- RPM Online Insertion and Removal
- Linecard Online Insertion and Removal

RPM Online Insertion and Removal

Dell Networking systems are functional with only one RPM.

If you insert a second RPM, it comes online as the standby RPM.

To see SFM status information, use the `show sfm all` command.

Example of the `show rpm all` Command

```
Dell#show rpm all

-- Route Processor Modules --
Slot   Status   NxtBoot   Version
-----------------------------------------------------------
0      active   online    7-5-1-71
1      not present
%RPM0-P:CP %POLLMGR-2-ALT_RPM_STATE: Alternate RPM is present
%RPM0-P:CP %IRC-6-IRC_COMMUP: Link to peer RPM is up
%RPM1-S:CP %RAM-5-RPM_STATE: RPM1 is in Standby State

Dell#show rpm all
-- Route Processor Modules --
Slot   Status    NxtBoot   Version
-----------------------------------------------------------
0      active    online    7-5-1-71
1      standby   online    7-5-1-71
```

Linecard Online Insertion and Removal

Dell Networking OS detects the line card type when you insert a line card into a online chassis.

Dell Networking OS writes the line card type to the running-config and maintains this information as a logical configuration if you remove the card (or the card fails).

Example of Viewing Line Card Configuration

```
Dell(conf)#do show linecard all

-- Line cards --
Slot Status NxtBoot ReqTyp CurTyp  Version    Ports
------------------------------------------------------------------------------
0    not present
[output omitted]

Dell(conf)# %RPM0-P:CP %CHMGR-5-CARDDETECTED: Line card 0 present
Dell(conf)# do show linecard all

-- Line cards --
Slot Status  NxtBoot  ReqTyp CurTyp  Version   Ports
------------------------------------------------------------------------------
0    online  online   E48VB  E48VB   7-5-1-71  48
[output omitted]

Dell(conf)#%RPM0-P:CP %CHMGR-2-CARD_DOWN: Line card 0 down - card removed
Dell(conf)#do show linecard all
-- Line cards --
Slot Status  NxtBoot  ReqTyp CurTyp  Version   Ports
------------------------------------------------------------------------------
0   not present       E48VB
[output omitted]
```

Pre-Configuring a Stack Unit Slot

You may also pre-configure an empty stack unit slot with a logical stack unit.

To pre-configure an empty stack unit slot, use the following command.

- Pre-configure an empty stack unit slot with a logical stack unit.
  CONFIGURATION mode
  `stack-unit` *`unit_id`* `provision S4820T`

After creating the logical stack unit, you can configure the interfaces on the stack unit as if it is present.

```
Dell#show system stack-unit 1

--  Unit 1 --
Unit Type                  : Member Unit
Status                     : not present

Dell#con
Dell(conf)#stack-unit 1 provision S4820T
Dell(conf)#end

Dell#show system stack-unit 1

--  Unit 1 --
Unit Type                  : Member Unit
Status                     : not present
Required Type              : S4820T - 52-port GE/TE/FG (SE)

Dell#

Dell(conf)#interface tengigabitethernet 1/0
Dell(conf-if-te-1/0)#
```

Removing a Provisioned Logical Stack Unit

To remove the line card configuration, use the following command.

- To remove a logical stack-unit configuration, use the following command:
  CONFIGURATION mode

```
no stack-unit unit_id provision
```

Hitless Behavior

Hitless is a protocol-based system behavior that makes a stack unit failover on the local system transparent to remote systems. The system synchronizes protocol information on the Management and Standby stack units such that, in the event of a stack unit failover, it is not necessary to notify the remote systems of a local state change.

Hitless behavior is defined in the context of a stack unit failover only.

- Only failovers via the CLI are hitless. The system is not hitless in any other scenario.

Hitless protocols are compatible with other hitless and graceful restart protocols. For example, if hitless open shortest path first (OSPF) is configured over hitless the link aggregation control protocol (LACP) link aggregation groups (LAGs), both features work seamlessly to deliver a hitless OSPF-LACP result. However, to achieve a hitless end result, if the hitless behavior involves multiple protocols, all protocols must be hitless. For example, if OSPF is hitless but bidirectional forwarding detection (BFD) is not, OSPF operates hitlessly and BFD flaps upon an RPM failover.

The following protocols are hitless:

- Link aggregation control protocol.
- Spanning tree protocol. Refer to Configuring Spanning Trees as Hitless.

Graceful Restart

Graceful restart (also known as non-stop forwarding) is a protocol-based mechanism that preserves the forwarding table of the restarting router and its neighbors for a specified period to minimize the loss of packets. A graceful-restart router does not immediately assume that a neighbor is permanently down and so does not trigger a topology change. Packet loss is non-zero, but trivial, and so is still called hitless.

Dell EMC Networking OS supports graceful restart for the following protocols:

- Border gateway
- Open shortest path first
- Protocol independent multicast — sparse mode
- Intermediate system to intermediate system

Software Resiliency

During normal operations, Dell EMC Networking OS monitors the health of both hardware and software components in the background to identify potential failures, even before these failures manifest.

Software Component Health Monitoring

On each of the line cards and the stack unit, there are a number of software components. Dell EMC Networking OS performs a periodic health check on each of these components by querying the status of a flag, which the corresponding component resets within a specified time.

If any health checks on the stack unit fail, the Dell EMC Networking OS fails over to standby stack unit. If any health checks on a line card fail, Dell EMC Networking OS resets the card to bring it back to the correct state.

System Health Monitoring

Dell EMC Networking OS also monitors the overall health of the system.

Key parameters such as CPU utilization, free memory, and error counters (for example, CRC failures and packet loss) are measured, and after exceeding a threshold can be used to initiate recovery mechanism.

Failure and Event Logging

Dell EMC Networking systems provide multiple options for logging failures and events.

Trace Log

Developers interlace messages with software code to track the execution of a program.

These messages are called trace messages and are primarily used for debugging and to provide lower-level information then event messages, which system administrators primarily use. Dell EMC Networking OS retains executed trace messages for hardware and software and stores them in files (logs) on the internal flash.

- **NV Trace Log** — contains line card bootup trace messages that Dell EMC Networking OS never overwrites and is stored in internal flash under the directory NVTRACE_LOG_DIR.
- **Trace Log** — contains trace messages related to software and hardware events, state, and errors. Trace Logs are stored in internal flash under the directory TRACE_LOG_DIR.
- **Crash Log** — contains trace messages related to IPC and IRC timeouts and task crashes on line cards and is stored under the directory CRASH_LOG_DIR.

For more information about trace logs and configuration options, refer to S-Series Debugging and Diagnostics.

Core Dumps

A core dump is the contents of RAM a program uses at the time of a software exception and is used to identify the cause of the exception.

There are two types of core dumps: application and kernel.

- Application core dump is the contents of the memory allocated to a failed application at the time of an exception.
- Kernel core dump is the central component of an operating system that manages system processors and memory allocation and makes these facilities available to applications. A kernel core dump is the contents of the memory in use by the kernel at the time of an exception.

System Log

Event messages provide system administrators diagnostics and auditing information.

Dell EMC Networking OS sends event messages to the internal buffer, all terminal lines, the console, and optionally to a syslog server. For more information about event messages and configurable options, refer to Management.

Hot-Lock Behavior

Dell EMC Networking OS hot-lock features allow you to append and delete their corresponding content addressable memory (CAM) entries dynamically without disrupting traffic. Existing entries are simply shuffled to accommodate new entries.

Hot-Lock IP ACLs allows you to append rules to and delete rules from an access control list (ACL) that is already written to CAM. This behavior is enabled by default and is available for both standard and extended ACLs on ingress and egress. For information about configuring ACLs, refer to Access Control Lists (ACLs).

Process Restartability

Process restartability is an extension to the Dell Networking OS high availability system component that enables application processes and system protocol tasks to be restarted.

This extension increases system reliability and uptime by attempting to restart the crashed process on primary RPM before executing the failover procedure as a last resort.

Currently, if a software exception occurs, Dell Networking OS executes a failover procedure. In a single-RPM system, the system generates a coredump and reboots; in a dual-RPM system, the system generates a coredump and fails over to the standby RPM.

With a system reload, the system must read and apply the entire startup-config file, which might take some time if the startup-config is large. Restarting a process saves time because only a portion of the configuration related to the crashed process is read and reapplied.

For a dual-RPMs system, restarting a process also precludes launching the failover process on the primary and standby RPMs. Recovery is attempted first locally on the primary RPM, which involves less CPU overhead, increasing the systems availability for other activities.

However, in both single and dual-RPM systems, even when you configure process restart, the coredump portion of failover is still executed.

The processes that you can restart fall under three categories:

- **Interface-related processes** — TACACS+, RADIUS, CLI, and SSH, and so on.
- **Protocol tasks** — OSPF, RIP, and ACL, and so on. Process restart is not currently available for protocol tasks; the failover procedure is executed immediately after software exception.
- **Line card processes** — IPC, Event Log Agent, Line Card Manager, and so on. Process restart is not currently available for line card processes; the failover procedure is executed immediately after software exception.

Enabling Process Restartability

The restart time varies by process.

In general, interface-related processes are hitless and can be restarted in seconds; if a restart is successful, traffic is not interrupted. Protocol tasks and line card processes are not hitless and take longer to restart. You can select which process may attempt to restart and the number of consecutive restart attempts before failover, but by default, every process fails over.

- Enable process restartability for a process or task.

  CONFIGURATION mode

  `process restartable [process] [try number] [timestamp hours]`
- Display the processes and tasks configured for restart.

  EXEC Privilege

When a process restarts, FTOS displays a message similar to the following message.

```
[9/18 23:22:21] TME-(tme): Starting to restart the failed process tacplus
[9/18 23:22:41] TME-(tme): Finishing restarting the failed process tacplus
```

You can specify the timestamp in hour(s) so that if the number of attempts to restart exceeds the maximum allowed within this timestamp, Restart mode is changed into Failover mode from that moment forward. This means that the next time the crashed process does NOT restart but failover to the standby RPM if it is on a dual RPM environment and rebooted if it is on a single RPM.

20

Internet Group Management Protocol (IGMP)

Internet group management protocol (IGMP) is a Layer 3 multicast protocol that hosts use to join or leave a multicast group.

Multicast is premised on identifying many hosts by a single destination IP address; hosts represented by the same IP address are a multicast group. Multicast routing protocols (such as protocol-independent multicast [PIM]) use the information in IGMP messages to discover which groups are active and to populate the multicast routing table.

IGMP Implementation Information

- Dell EMC Networking Operating System (OS) supports IGMP versions 1, 2, and 3 based on RFCs 1112, 2236, and 3376, respectively.
- Dell EMC Networking OS does not support IGMP version 3 and versions 1 or 2 on the same subnet.
- IGMP on Dell EMC Networking OS supports 95 interfaces.
- Dell EMC Networking systems cannot serve as an IGMP host or an IGMP version 1 IGMP Querier.
- Dell EMC Networking OS automatically enables IGMP on interfaces on which you enable a multicast routing protocol.

Topics:

- IGMP Protocol Overview
- Configure IGMP
- Viewing IGMP Enabled Interfaces
- Selecting an IGMP Version
- Viewing IGMP Groups
- Adjusting Timers
- Enabling IGMP Immediate-Leave
- IGMP Snooping
- Fast Convergence after MSTP Topology Changes
- Egress Interface Selection (EIS) for HTTP and IGMP Applications
- Designating a Multicast Router Interface

IGMP Protocol Overview

IGMP has three versions. Version 3 obsoletes and is backwards-compatible with version 2; version 2 obsoletes version 1.

IGMP Version 2

IGMP version 2 improves on version 1 by specifying IGMP Leave messages, which allows hosts to notify routers that they no longer care about traffic for a particular group.

Leave messages reduce the amount of time that the router takes to stop forwarding traffic for a group to a subnet (leave latency) after the last host leaves the group. In version 1 hosts quietly leave groups, and the router waits for a query response timer several times the value of the query interval to expire before it stops forwarding traffic.

To receive multicast traffic from a particular source, a host must join the multicast group to which the source is sending traffic. A host that is a member of a group is called a receiver. A host may join many groups, and may join or leave any group at any time. A host joins and leaves a multicast group by sending an IGMP message to its IGMP Querier. The querier is the router that surveys a subnet for multicast receivers and processes survey responses to populate the multicast routing table.

IGMP messages are encapsulated in IP packets, as shown in the following illustration.

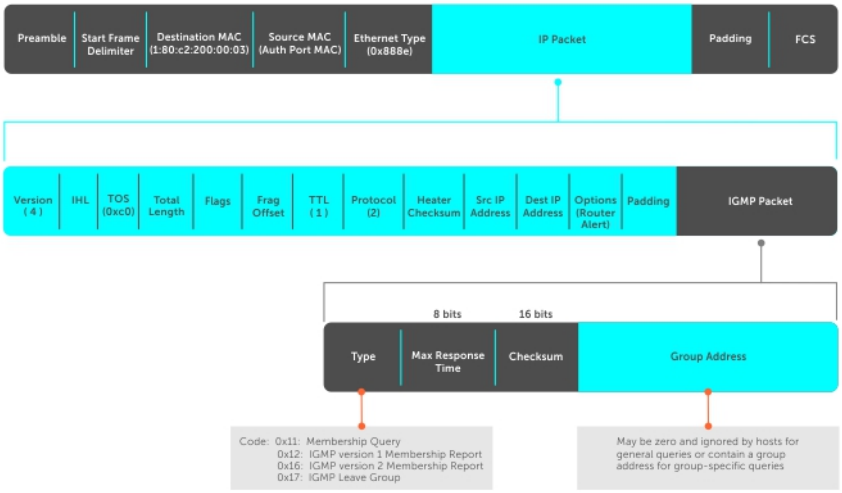


Figure 46. IGMP Messages in IP Packets

Join a Multicast Group

There are two ways that a host may join a multicast group: it may respond to a general query from its querier or it may send an unsolicited report to its querier.

Responding to an IGMP Query

The following describes how a host can join a multicast group.

1. One router on a subnet is elected as the querier. The querier periodically multicasts (to all-multicast-systems address 224.0.0.1) a general query to all hosts on the subnet.
2. A host that wants to join a multicast group responds with an IGMP Membership Report that contains the multicast address of the group it wants to join (the packet is addressed to the same group). If multiple hosts want to join the same multicast group, only the report from the first host to respond reaches the querier and the remaining hosts suppress their responses (For how the delay timer mechanism works, refer to Adjusting Query and Response Timers).
3. The querier receives the report for a group and adds the group to the list of multicast groups associated with its outgoing port to the subnet. Multicast traffic for the group is then forwarded to that subnet.

Sending an Unsolicited IGMP Report

A host does not have to wait for a general query to join a group. It may send an unsolicited IGMP Membership Report, also called an IGMP Join message, to the querier.

Leaving a Multicast Group

The following describes how a host can leave a multicast group.

1. A host sends a membership report of type 0x17 (IGMP Leave message) to the all routers multicast address 224.0.0.2 when it no longer cares about multicast traffic for a particular group.
2. The querier sends a Group-Specific Query to determine whether there are any remaining hosts in the group. There must be at least one receiver in a group on a subnet for a router to forward multicast traffic for that group to the subnet.
3. Any remaining hosts respond to the query according to the delay timer mechanism (refer to Adjusting Query and Response Timers). If no hosts respond (because there are none remaining in the group), the querier waits a specified period and sends another query. If it still receives no response, the querier removes the group from the list associated with forwarding port and stops forwarding traffic for that group to the subnet.

IGMP Version 3

Conceptually, IGMP version 3 behaves the same as version 2. However, there are differences.

- Version 3 adds the ability to filter by multicast source, which helps multicast routing protocols avoid forwarding traffic to subnets where there are no interested receivers.
- To enable filtering, routers must keep track of more state information, that is, the list of sources that must be filtered. An additional query type, the Group-and-Source-Specific Query, keeps track of state changes, while the Group-Specific and General queries still refresh the existing state.
- Reporting is more efficient and robust: hosts do not suppress query responses (non-suppression helps track state and enables the immediate-leave and IGMP snooping features), state-change reports are retransmitted to insure delivery, and a single membership report bundles multiple statements from a single host, rather than sending an individual packet for each statement.

The version 3 packet structure is different from version 2 to accommodate these protocol enhancements. Queries are still sent to the all-systems address 224.0.0.1, as shown in the following illustration, but reports are sent to the all IGMP version 3-capable multicast routers address 244.0.0.22, as shown in the second illustration.

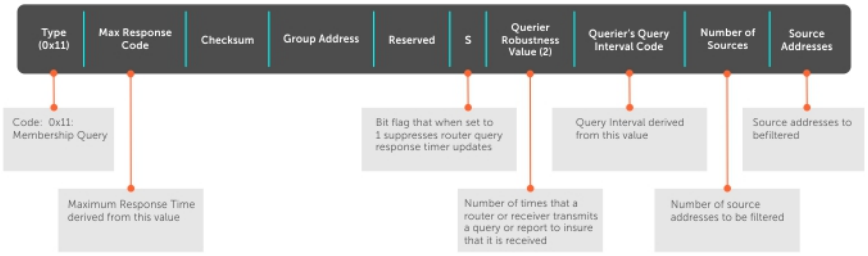


Figure 47. IGMP Version 3 Packet Structure

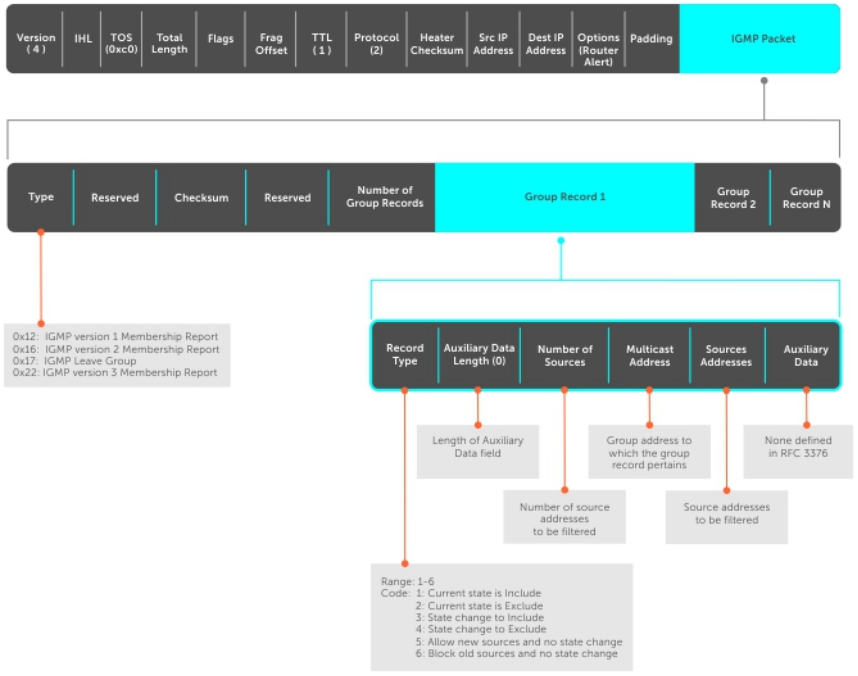


Figure 48. IGMP Version 3–Capable Multicast Routers Address Structure

Joining and Filtering Groups and Sources

The following illustration shows how multicast routers maintain the group and source information from unsolicited reports.

1. The first unsolicited report from the host indicates that it wants to receive traffic for group 224.1.1.1.
2. The host's second report indicates that it is only interested in traffic from group 224.1.1.1, source 10.11.1.1. Include messages prevents traffic from all other sources in the group from reaching the subnet. Before recording this request, the querier sends a group-and-source query to verify that there are no hosts interested in any other sources. The multicast router must satisfy all hosts if they have conflicting requests. For example, if another host on the subnet is interested in traffic from 10.11.1.3, the router cannot record the include request. There are no other interested hosts, so the request is recorded. At this point, the multicast routing protocol prunes the tree to all but the specified sources.
3. The host's third message indicates that it is only interested in traffic from sources 10.11.1.1 and 10.11.1.2. Because this request again prevents all other sources from reaching the subnet, the router sends another group-and-source query so that it can satisfy all other hosts. There are no other interested hosts so the request is recorded.

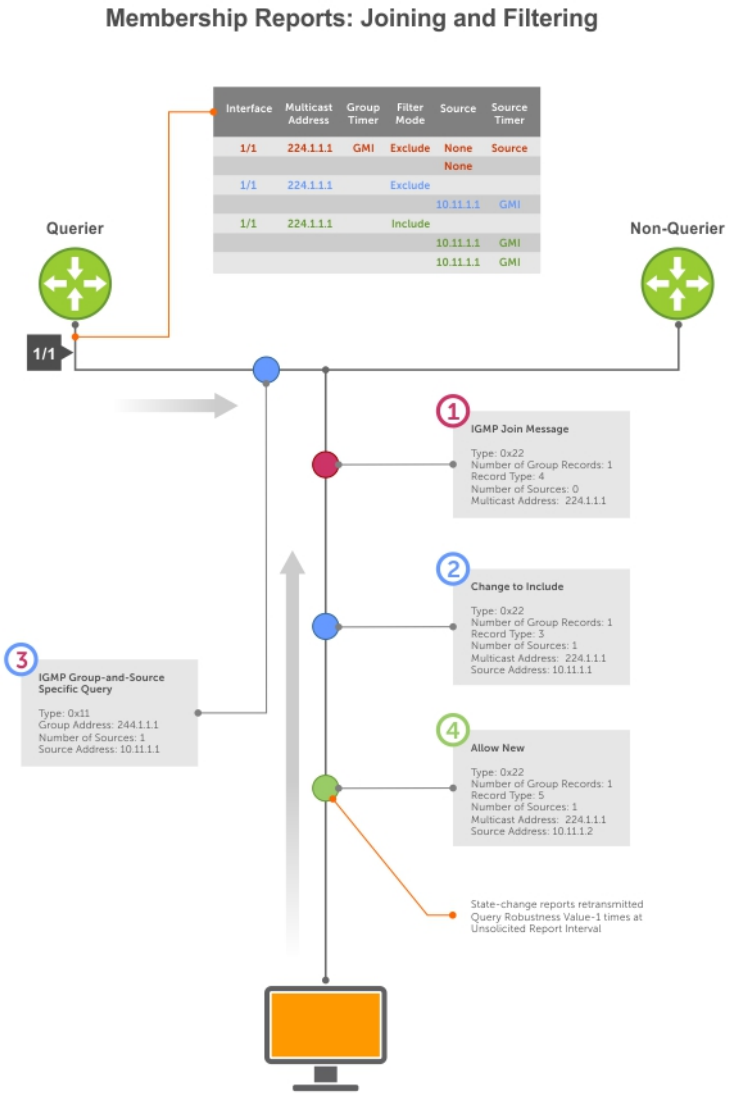


Figure 49. Membership Reports: Joining and Filtering

Leaving and Staying in Groups

The following illustration shows how multicast routers track and refresh state changes in response to group-and-specific and general queries.

1. Host 1 sends a message indicating it is leaving group 224.1.1.1 and that the included filter for 10.11.1.1 and 10.11.1.2 are no longer necessary.
2. The querier, before making any state changes, sends a group-and-source query to see if any other host is interested in these two sources; queries for state-changes are retransmitted multiple times. If any are, they respond with their current state information and the querier refreshes the relevant state information.
3. Separately in the following illustration, the querier sends a general query to 224.0.0.1.
4. Host 2 responds to the periodic general query so the querier refreshes the state information for that group.

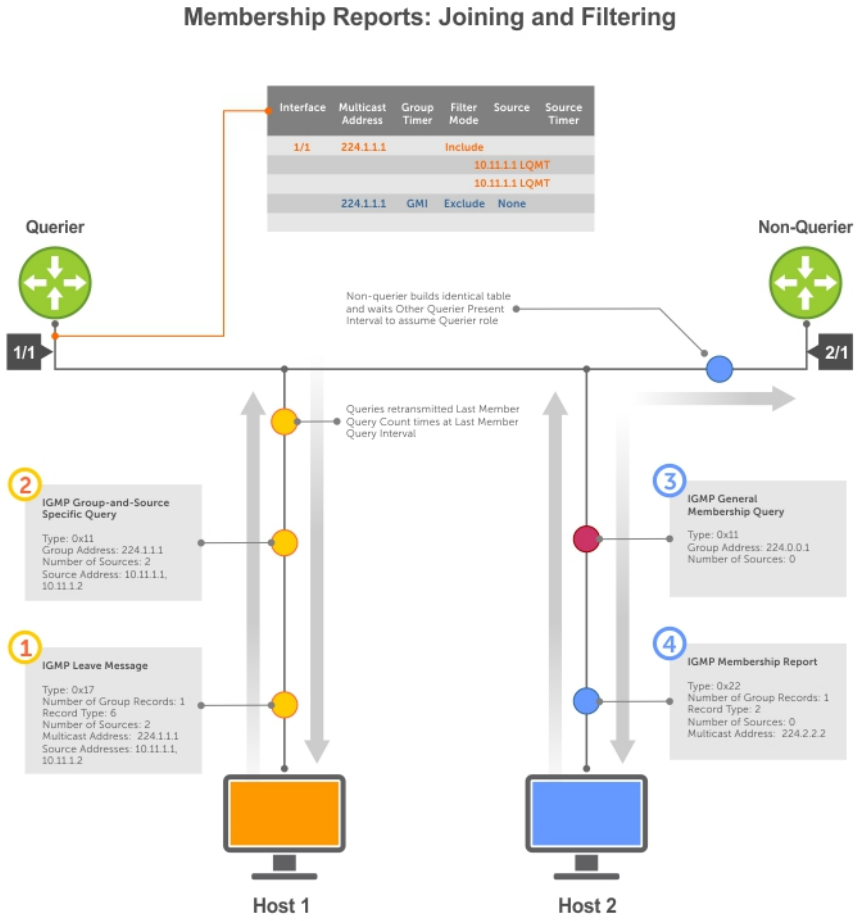


Figure 50. Membership Queries: Leaving and Staying

Configure IGMP

Configuring IGMP is a two-step process.

1. Enable multicast routing using the `ip multicast-routing` command.
2. Enable a multicast routing protocol.

Related Configuration Tasks

- Viewing IGMP Enabled Interfaces
- Selecting an IGMP Version
- Viewing IGMP Groups

- Adjusting Timers
- Preventing a Host from Joining a Group
- Enabling IGMP Immediate-Leave
- IGMP Snooping
- Fast Convergence after MSTP Topology Changes
- Designating a Multicast Router Interface

Viewing IGMP Enabled Interfaces

Interfaces that are enabled with PIM-SM are automatically enabled with IGMP.

To view IGMP-enabled interfaces, use the following command.

- View IGMP-enabled IPv4 interfaces.

  EXEC Privilege mode

  ```
  show ip igmp interface
  ```

- View IGMP-enabled IPv6 interfaces.

  EXEC Privilege mode

  ```
  show ipv6 mld interface
  ```

```
DellEMC#show ip igmp interface TenGigabitEthernet 3/10
  Inbound IGMP access group is not set
  Internet address is 165.87.34.5/24
  IGMP is up on the interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 0 seconds
  IGMP max query response time is 10 seconds
  IGMP last member query response interval is 1000 ms
  IGMP immediate-leave is disabled
  IGMP activity: 2 joins
  IGMP querying router is 165.87.34.5 (this system)
  IGMP version is 2
DellEMC#
```

Selecting an IGMP Version

Dell EMC Networking OS enables IGMP version 2 by default, which supports version 1 and 2 hosts, but is not compatible with version 3 on the same subnet.

If hosts require IGMP version 3, you can switch to IGMP version 3.

To switch to version 3, use the following command.

- Switch to a different IGMP version.

  INTERFACE mode

  ```
  ip igmp version
  ```

```
DellEMC(conf-if-te-1/13)#ip igmp version 3
DellEMC(conf-if-te-1/13)#do show ip igmp interface
TenGigabitEthernet 1/13 is up, line protocol is down
  Inbound IGMP access group is not set
  Interface IGMP group join rate limit is not set
  Internet address is 1.1.1.1/24
  IGMP is enabled on interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 125 seconds
  IGMP max query response time is 10 seconds
  IGMP last member query response interval is 1000 ms
  IGMP immediate-leave is disabled
  IGMP activity: 0 joins, 0 leaves, 0 channel joins, 0 channel leaves
  IGMP querying router is 1.1.1.1 (this system)
  IGMP version is 3
```

Viewing IGMP Groups

To view both learned and statically configured IGMP groups, use the following command.

- View both learned and statically configured IGMP groups.

  EXEC Privilege mode

  ```
  show ip igmp groups
  ```

  ```
  show ipv6 mld groups
  ```

```
DellEMC# show ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address Interface                Mode    Uptime    Expires   Last Reporter
225.1.1.1     TenGigabitEthernet 1/1   IGMPV2  00:11:19  00:01:50  165.87.34.100
225.1.2.1     TenGigabitEthernet 1/1   IGMPV2  00:10:19  00:01:50  165.87.31.100
```

Adjusting Timers

The following sections describe viewing and adjusting timers.

To view the current value of all IGMP timers, use the following command.

- View the current value of all IGMP timers.

  EXEC Privilege mode

  ```
  show ip igmp interface
  ```

For more information, refer to the example shown in Viewing IGMP Enabled Interfaces.

Adjusting Query and Response Timers

The querier periodically sends a general query to discover which multicast groups are active. A group must have at least one host to be active.

When a host receives a query, it does not respond immediately, but rather starts a delay timer. The delay time is set to a random value between 0 and the maximum response time. The host sends a response when the timer expires; in version 2, if another host responds before the timer expires, the timer is nullified, and no response is sent.

The maximum response time is the amount of time that the querier waits for a response to a query before taking further action. The querier advertises this value in the query (refer to the illustration in IGMP Version 2). Lowering this value decreases leave latency but increases response burstiness because all host membership reports must be sent before the maximum response time expires. Inversely, increasing this value decreases burstiness at the expense of leave latency.

When the querier receives a leave message from a host, it sends a group-specific query to the subnet. If no response is received, it sends another. The amount of time that the querier waits to receive a response to the initial query before sending a second one is the last member query interval (LMQI). The switch waits one LMQI after the second query before removing the group from the state table.

- Adjust the period between queries.

  INTERFACE mode

  ```
  ip igmp query-interval
  ```

- Adjust the time period between queries used to discover IPv6 multicast groups.

  Interface mode

  ```
  ipv6 mld query-interval
  ```

- Adjust the maximum response time.

  INTERFACE mode

  ```
  ip igmp query-max-resp-time
  ```

- Adjust the maximum amount of time that the querier waits, for an IPv6 query response, before taking further action.

  Interface mode

  ```
  ipv6 mld query-max-response-time
  ```

- Adjust the last member query interval.

  INTERFACE mode

  ```
  ip igmp last-member-query-interval
  ```
- Adjust the amount of time the querier waits, for the initial query response, before sending the next IPv6 query.

  Interface mode

  ```
  ipv6 mld last-member-query-interval
  ```

Enabling IGMP Immediate-Leave

If the querier does not receive a response to a group-specific or group-and-source query, it sends another (querier robustness value). Then, after no response, it removes the group from the outgoing interface for the subnet.

IGMP immediate leave reduces leave latency by enabling a router to immediately delete the group membership on an interface after receiving a Leave message (it does not send any group-specific or group-and-source queries before deleting the entry).

- Configure the system for IGMP immediate leave.

  ```
  ip igmp immediate-leave
  ```
- Configure the system to immediately delete the IPv6 group membership on an interface after receiving a leave message.

  ```
  ipv6 mld immediate-leave
  ```
- View the enable status of the IGMP immediate leave feature.

  EXEC Privilege mode

  ```
  show ip igmp interface
  ```

  ```
  show ipv6 mld interface
  ```

View the enable status of this feature using the command from EXEC Privilege mode, as shown in the example in Selecting an IGMP Version.

IGMP Snooping

IGMP snooping enables switches to use information in IGMP packets to generate a forwarding table that associates ports with multicast groups so that when they receive multicast frames, they can forward them only to interested receivers.

Multicast packets are addressed with multicast MAC addresses, which represent a group of devices, rather than one unique device. Switches forward multicast frames out of all ports in a virtual local area network (VLAN) by default, even though there may be only some interested hosts, which is a waste of bandwidth.

If you enable IGMP snooping on a VLT unit, IGMP snooping dynamically learned groups and multicast router ports are made to learn on the peer by explicitly tunneling the received IGMP control packets.

IGMP Snooping Implementation Information

- IGMP snooping on Dell EMC Networking OS uses IP multicast addresses not MAC addresses.
- IGMP snooping is supported on all stack members.
- IGMP snooping reacts to spanning tree protocol (STP) and multiple spanning tree protocol (MSTP) topology changes by sending a general query on the interface that transitions to the forwarding state.
- If IGMP snooping is enabled on a PIM-enabled VLAN interface, data packets using the router as an Layer 2 hop may be dropped. To avoid this scenario, Dell EMC Networking recommends that users enable IGMP snooping on server-facing end-point VLANs only.

Configuring IGMP Snooping

Configuring IGMP snooping is a one-step process. To enable, view, or disable IGMP snooping, use the following commands.

There is no specific configuration needed for IGMP snooping with virtual link trunking (VLT). For information about VLT configurations, refer to Virtual Link Trunking (VLT).

- Enable IGMP snooping on a switch.

  CONFIGURATION mode

```
ip igmp snooping enable
```

- View the configuration.

  CONFIGURATION mode

  ```
  show running-config
  ```

- Disable snooping on a VLAN.

  INTERFACE VLAN mode

  ```
  no ip igmp snooping
  ```

Related Configuration Tasks

- Removing a Group-Port Association
- Disabling Multicast Flooding
- Specifying a Port as Connected to a Multicast Router
- Configuring the Switch as Querier

```
DellEMC(conf)#ip igmp snooping enable
DellEMC(conf)#do show running-config igmp
ip igmp snooping enable
DellEMC(conf)#
```

Removing a Group-Port Association

To configure or view the remove a group-port association feature, use the following commands.

- Configure the switch to remove a group-port association after receiving an IGMP Leave message.

  INTERFACE VLAN mode

  ```
  ip igmp fast-leave
  ```

- View the configuration.

  INTERFACE VLAN mode

  ```
  show config
  ```

```
DellEMC(conf-if-vl-100)#show config
!
interface Vlan 100
  no ip address
  ip igmp snooping fast-leave
  shutdown
DellEMC(conf-if-vl-100)#
```

Disabling Multicast Flooding

If the switch receives a multicast packet that has an IP address of a group it has not learned (unregistered frame), the switch floods that packet out of all ports on the VLAN.

When you configure the `no ip igmp snooping flood` command, the system drops the packets immediately. The system does not forward the frames on mrouter ports, even if they are present. Disable Layer 3 multicast (`no ip multicast-routing`) in order to disable multicast flooding.

- Configure the switch to only forward unregistered packets to ports on a VLAN that are connected to mrouter ports.

  CONFIGURATION mode

  ```
  no ip igmp snooping flood
  ```

Specifying a Port as Connected to a Multicast Router

To statically specify or view a port in a VLAN, use the following commands.

- Statically specify a port in a VLAN as connected to a multicast router.

  INTERFACE VLAN mode

```
ip igmp snooping mrouter
```

- View the ports that are connected to multicast routers.

  EXEC Privilege mode.

  ```
  show ip igmp snooping mrouter
  ```

Configuring the Switch as Querier

To configure the switch as a querier, use the following command.

Hosts that do not support unsolicited reporting wait for a general query before sending a membership report. When the multicast source and receivers are in the same VLAN, multicast traffic is not routed and so there is no querier. Configure the switch to be the querier for a VLAN so that hosts send membership reports and the switch can generate a forwarding table by snooping.

- Configure the switch to be the querier for a VLAN by first assigning an IP address to the VLAN interface.

  INTERFACE VLAN mode

  ```
  ip igmp snooping querier
  ```

  IGMP snooping querier does not start if there is a statically configured multicast router interface in the VLAN.

  The switch may lose the querier election if it does not have the lowest IP address of all potential queriers on the subnet.

  When enabled, IGMP snooping querier starts after one query interval in case no IGMP general query (with IP SA lower than its VLAN IP address) is received on any of its VLAN members.

Adjusting the Last Member Query Interval

To adjust the last member query interval, use the following command.

When the querier receives a Leave message from a receiver, it sends a group-specific query out of the ports specified in the forwarding table. If no response is received, it sends another. The amount of time that the querier waits to receive a response to the initial query before sending a second one is the last member query interval (LMQI). The switch waits one LMQI after the second query before removing the group-port entry from the forwarding table.

- Adjust the last member query interval.

  INTERFACE VLAN mode

  ```
  ip igmp snooping last-member-query-interval
  ```

Fast Convergence after MSTP Topology Changes

When a port transitions to the Forwarding state as a result of an STP or MSTP topology change, Dell EMC Networking OS sends a general query out of all ports except the multicast router ports. The host sends a response to the general query and the forwarding database is updated without having to wait for the query interval to expire.

When an IGMP snooping switch is not acting as a querier, it sends out the general query in response to the MSTP triggered link-layer topology change, with the source IP address of 0.0.0.0 to avoid triggering querier election.

Egress Interface Selection (EIS) for HTTP and IGMP Applications

You can use the Egress Interface Selection (EIS) feature to isolate the management and front-end port domains for HTTP and IGMP traffic. Also, EIS enables you to configure the responses to switch-destined traffic by using the management port IP address as the source IP address. This information is sent out of the switch through the management port instead of the front-end port.

The management EIS feature is applicable only for the out-of-band (OOB) management port. References in this section to the management default route or static route denote the routes configured using the `management route` command. The management default route can be either configured statically or returned dynamically by the DHCP client. A static route points to the management interface or a forwarding router.

Transit traffic (destination IP not configured in the switch) that is received on the front-end port with destination on the management port is dropped and received in the management port with destination on the front-end port is dropped.

Switch-destined traffic (destination IP configured in the switch) is:

- Received in the front-end port with destination IP equal to management port IP address or management port subnet broadcast address is dropped.
- Received in the management port with destination IP not equal to management IP address or management subnet broadcast address is dropped.

Traffic (switch initiated management traffic or responses to switch-destined traffic with management port IP address as the source IP address) for user-specified management protocols must exit out of the management port. In this chapter, all the references to traffic indicate switch-initiated traffic and responses to switch-destined traffic with management port IP address as the source IP address.

In customer deployment topologies, it might be required that the traffic for certain management applications needs to exit out of the management port only. You can use EIS to control and the traffic can exit out of any port based on the route lookup in the IP stack.

One typical example is an SSH session to an unknown destination or an SSH connection that is destined to the management port IP address. The management default route can coexist with front-end default routes. If SSH is specified as a management application, SSH links to and from an unknown destination uses the management default route.

Protocol Separation

When you configure the `application` *`application-type`* command to configure a set of management applications with TCP/UDP port numbers to the OS, the following table describes the association between applications and their port numbers.

Table 37. Association Between Applications and Port Numbers

Application Name	Port Number	Client	Server
SSH	22	Supported	Supported
Sflow-Collector	6343	Supported	
SNMP	162 for SNMP Traps (client), 161 for SNMP MIB response (server)	Supported	
NTP	123	Supported	
DNS	53	Supported	
FTP	20/21	Supported	Supported
Syslog	514	Supported	
Telnet	23	Supported	Supported
TFTP	69	Supported	
Radius	1812,1813	Supported	
Tacacs	49	Supported	
HTTP	80 for httpd 443 for secure httpd 8008 HTTP server port for confd application 8888 secure HTTP server port for confd application		Supported

If you configure a source interface is for any EIS management application, EIS might not coexist with that interface and the behavior is undefined in such a case. You can configure the source interface for the following applications: FTP, ICMP (ping and traceroute utilities), NTP, RADIUS, TACACS, Telnet, TFTP, syslog, and SNMP traps. Out of these applications, EIS can coexist with only syslog and SNMP traps because these applications do not require a response after a packet is sent.

The switch also processes user-specified port numbers for applications such as RADIUS, TACACS, SSH, and sFlow. The OS maintains a list of configured management applications and their port numbers. You can configure two default routes, one configured on the management port and the other on the front-end port.

Two tables, namely, Egress Interface Selection routing table and default routing table, are maintained. In the preceding table, the columns Client and Server indicate that the applications can act as both a client and a server within the switch. The Management Egress Interface Selection table contains all management routes (connected, static and default route). The default routing table contains all management routes (connected, static and default route) and all front-end port routes.

Enabling and Disabling Management Egress Interface Selection

You can enable or disable egress-interface-selection using the `management egress-interface-selection` command.

NOTE: Egress Interface Selection (EIS) works only with IPv4 routing.

When the feature is enabled using the `management egress-interface-selection` command, the following events are performed:

- The CLI prompt changes to the EIS mode.
- In this mode, you can run the `application` and `no application` commands
- Applications can be configured or unconfigured as management applications using the `application` or `no application` command. All configured applications are considered as management applications and the rest of them as non-management applications.
- All the management routes (connected, static and default) are duplicated and added to the management EIS routing table.
- Any new management route added is installed to both the EIS routing table and default routing table.
- For management applications, route lookup is preferentially done in the management EIS routing table for all traffic. management port is the preferred egress port. For example, if SSH is a management application, an SSH session to a front-panel port IP on the peer box is initiated via management port only, if the management port is UP and management route is available.
- If SSH request is received on the management port destined to the management port IP address, the response to the request is sent out of the management port by performing a route lookup in the EIS routing table
- If the SSH request is received on the front-end port destined for the front-end IP address, the response traffic is sent by doing a route lookup in the default routing table only.
- If the management port is down or route lookup fails in the management EIS routing table, packets are dropped.
- For all non-management applications, traffic exits out of either front-end data port or management port based on route lookup in default routing table.
- Ping and traceroute are always non-management applications and route lookup for these applications is done in the default routing table only.
- For ping and traceroute utilities that are initiated from the switch, if reachability needs to be tested through routes in the management EIS routing table, you must configure ICMP as a management application.
- If ping and traceroute are destined to the management port IP address, the response traffic for these packets is sent by doing route lookup in the EIS routing table.

When the feature is disabled using the `no management egress-interface-selection` command, the following operations are performed:

- All management application configuration is removed.
- All routes installed in the management EIS routing table are removed.

Handling of Management Route Configuration

When the EIS feature is enabled, the following processing occurs:

- All existing management routes (connected, static and default) are duplicated and added to the management EIS routing table.
- Any management static route newly added using the `management route` CLI is installed to both the management EIS routing table and default routing table.
- As per existing behavior, for routes in the default routing table, conflicting front-end port routes if configured has higher precedence over management routes. So there can be scenarios where the same management route is present in the EIS routing table but not in the default routing table.
- Routes in the EIS routing table are displayed using the `show ip management-eis-route` command.
- In the netstat output, the prefix “mgmt” is added to routes in the EIS table so that the user can distinguish between routes in the EIS Routing table and default routing table.
- If the management port IP address is removed, the corresponding connected route is removed from both the EIS routing table and default routing table.
- If a management route is deleted, then the route is removed from both the EIS routing table and default routing table.

Handling of Switch-Initiated Traffic

When the control processor (CP) initiates a control packet, the following processing occurs:

- TCP/UDP port number is extracted from the sockaddr structure in the in_selectsrc call which is called as part of the connect system call or in the ip_output function. If the destination TCP/UDP port number belongs to a configured management application, then sin_port of destination sockaddr structure is set to Management EIS ID 2 so that route lookup can be done in the management EIS routing table.
- To ensure that protocol separation is done only for switch initiated traffic where the application acts as client, only the destination TCP/UDP port is compared and not the source TCP/UDP port. The source TCP/UDP port becomes a known port number when the box acts as server.
- TFTP is an exception to the preceding logic.
- For TFTP, data transfer is initiated on port 69, but the data transfer ports are chosen independently by the sender and receiver during initialization of the connection. The ports are chosen at random according to the parameters of the networking stack, typically from the range of temporary ports.
- If route lookup in EIS routing table succeeds, the application-specific packet count is incremented. This counter is viewed using the `show management application pkt-cntr` command. This counter is cleared using `clear management application pkt-cntr` command.
- If the route lookup in the EIS routing table fails or if management port is down, then packets are dropped. The application-specific count of the dropped packets is incremented and is viewed using the `show management application pkt-drop-cntr` command. This counter is cleared using `clear management application pkt-drop-cntr` command.
- Packets whose destination TCP/UDP port does not match a configured management application, take the regular route lookup flow in the IP stack.
- In the ARP layer, for all ARP packets received through the management interface, a double route lookup is done, one in the default routing table and another in the management EIS routing table. This is because in the ARP layer, we do not have TCP/UDP port information to decide the table in which the route lookup should be done.
- The `show arp` command is enhanced to show the routing table type for the ARP entry.
- For the `clear arp-cache` command, upon receiving the ARP delete request, the route corresponding to the destination IP is identified. The ARP entries learned in the management EIS routing table are also cleared.
- Therefore, a separate control over clearing the ARP entries learned via routes in the EIS table is not present. If the ARP entry for a destination is cleared in the default routing table, then if an ARP entry for the destination exists in the EIS table, that entry is also cleared.

- Because fallback support is removed, if the management port is down or the route lookup in EIS table fails packets are dropped. Therefore, switch-initiated traffic sessions that used to work previously via fallback may not work now.

Handling of Switch-Destined Traffic

- The switch processes all traffic received on the management port destined to the management port IP address or the front-end port destined to the front-end IP address.
- If the source TCP/UDP port number matches a configured EIS or non-EIS management application and the source IP address is a management Port IP address, then the EIS route lookup is done for the response traffic and hence is sent out of the management port. In this case, the source IP address is a management port IP address only if the traffic was originally destined to the management port IP.
- ICMP-based applications like ping and traceroute are exceptions to the preceding logic since we do not have TCP/UDP port number. So if source IP address of the packet matches the management port IP address EIS route lookup is done.
- Management application packet counter is incremented if EIS route lookup succeeds and packet is sent out of the management port.
- If route lookup in the EIS routing table fails or if the management port is down, then packets are dropped. The management application drop counter is incremented.
- Whenever IP address is assigned to the management port, it is stored in a global variable in the IP stack, which is used for comparison with the source IP address of the packet.
- Rest of the response traffic is handled as per existing behavior by doing route lookup in the default routing table. So if the traffic is destined to the front-end port IP address, the response is sent out by doing a route lookup in the default routing table, which is an existing behavior.

Consider a sample topology in which ip1 is an address assigned to the management port and ip2 is an address assigned to any of the front panel port. A and B are end users on the management and front-panel port networks. The OS-initiated traffic for management applications takes a preference for ip1 as source IP and uses the management network to reach the destination. If the management port is down or the route lookup in EIS routing table fails, ip2 is the source IP and the front-panel port is used to reach the destination. The fallback route between the management and data networks is used in such a case. At any given time, end users can access Dell EMC Networking OS applications using either ip1 or ip2. Return traffic for such end-user-originated sessions destined to management port ip1 is handled using the EIS route lookup.

Handling of Transit Traffic (Traffic Separation)

This is forwarded traffic where destination IP is not an IP address configured in the switch.

- Packets received on the management port with destination on the front-end port is dropped.
- Packets received on the front-end port with destination on the management port is dropped.
- A separate drop counter is incremented for this case. This counter is viewed using the `netstat` command, like all other IP layer counters.

Consider a scenario in which ip1 is an address assigned to the management port and ip2 is an address assigned to any of the front panel port of a switch. End users on the management and front panel port networks are connected. In such an environment, traffic received in the management port destined on the data port network is dropped and traffic received in the front-end port destined on the management network is dropped.

Mapping of Management Applications and Traffic Type

The following table summarizes the behavior of applications for various types of traffic when the management egress interface selection feature is enabled.

Table 38. Mapping of Management Applications and Traffic Type

Traffic type / Application type	Switch initiated traffic	Switch-destined traffic	Transit Traffic
EIS Management Application	Management is the preferred egress port selected based on route lookup in EIS table. If the management port is down or the route lookup fails, packets are dropped.	If source TCP/UDP port matches a management application and source IP address is management port IP address, management port is the preferred egress port selected based on route lookup in EIS table. If management port is down or route lookup fails, packets are dropped	Traffic from management port to data port and from data port to management port is blocked
Non-EIS management application	Front-end default route will take higher precedence over management default route and SSH session to an unknown destination uses the front-end default route only. No change in the existing behavior.	If source TCP/UDP port matches a management application and the source IP address is a management port IP address, the management port is the preferred egress port selected based on route lookup in EIS table. If the management port is down or the route lookup fails, packets are dropped	Traffic from management port to data port and from data port to management port is blocked

- EIS is enabled implies that EIS feature is enabled and the application might or might not be configured as a management application
- EIS is disabled implies that either EIS feature itself is disabled or that the application is not configured as a management application

Transit Traffic

This phenomenon occurs where traffic is transiting the switch. Traffic has not originated from the switch and is not terminating on the switch.

- Drop the packets that are received on the front-end data port with destination on the management port.
- Drop the packets that received on the management port with destination as the front-end data port.

Switch-Destined Traffic

This phenomenon occurs where traffic is terminated on the switch. Traffic has not originated from the switch and is not transiting the switch.

The switch accepts all traffic destined to the switch, which is received on management or front-end data port. Response traffic with management port IP address as source IP address is handled in the same manner as switch originated traffic.

Switch-Originated Traffic

This phenomenon occurs where traffic is originating from the switch.

1. Management Applications (Applications that are configured as management applications):

   The management port is an egress port for management applications. If the management port is down or the destination is not reachable through the management port (next hop ARP is not resolved, and so on), and if the destination is reachable through a data port, then the management application traffic is sent out through the front-end data port. This fallback mechanism is required.

2. Non-Management Applications (Applications that are not configured as management applications as defined by this feature):

   Non-management application traffic exits out of either front-end data port or management port based on routing table. If there is a default route on both the management and front-end data port, the default for the data port is preferred route.

Behavior of Various Applications for Switch-Initiated Traffic

This section describes the different system behaviors that occur when traffic is originating from the switch:

EIS Behavior: If the destination TCP/UDP port matches a configured management application, a route lookup is done in the EIS table and the management port gets selected as the egress port. If management port is down or the route lookup fails, packets are dropped.

EIS Behavior for ICMP: ICMP packets do not have TCP/UDP ports. To do an EIS route lookup for ICMP-based applications (ping and traceroute) using the source ip option, the management port IP address should be specified as the source IP address. If management port is down or route lookup fails, packets are dropped.

Default Behavior: Route lookup is done in the default routing table and appropriate egress port is selected.

Table 39. Behavior of Various Applications for Switch-Initiated Traffic

Protocol	Behavior when EIS is Enabled	Behavior when EIS is Disabled
dns	EIS Behavior	Default Behavior
ftp	EIS Behavior	Default Behavior
ntp	EIS Behavior	Default Behavior
radius	EIS Behavior	Default Behavior
Sflow-collector		Default Behavior
Snmp (SNMP Mib response and SNMP Traps)	EIS Behavior	Default Behavior
ssh	EIS Behavior	Default Behavior
syslog	EIS Behavior	Default Behavior
tacacs	EIS Behavior	Default Behavior
telnet	EIS Behavior	Default Behavior
tftp	EIS Behavior	Default Behavior
icmp (ping and traceroute)	EIS Behavior for ICMP	Default Behavior

Behavior of Various Applications for Switch-Destined Traffic

This section describes the different system behaviors that occur when traffic is terminated on the switch. Traffic has not originated from the switch and is not transiting the switch. Switch-destined traffic is applicable only for applications which act as server for the TCP session and also for ICMP-based applications like ping and traceroute. FTP, SSH, and Telnet are the applications that can function as servers for the TCP session.

EIS Behavior: If source TCP or UDP port matches an EIS management or a non-EIS management application and source IP address is management port IP address, management port is the preferred egress port selected based on route lookup in EIS table. If the management port is down or the route lookup fails, packets are dropped.

If the source TCP/UDP port or source IP address does not match the management port IP address, a route lookup is done in the default routing table.

EIS behavior for ICMP: ICMP packets do not have TCP/UDP ports. In this case, to perform an EIS route lookup for ICMP-based applications (ping and traceroute), you must configure ICMP as a management application. If the management port is down or the route lookup fails, packets are dropped.

If source IP address does not match the management port IP address route lookup is done in the default routing table.

Default Behavior: Route lookup is done in the default routing table and appropriate egress port is selected.

Table 40. Behavior of Various Applications for Switch-Destined Traffic

Protocol	Behavior when EIS is Enabled	Behavior when EIS is Disabled
ftp	EIS Behavior	Default Behavior
http	EIS Behavior	Default Behavior

Table 40. Behavior of Various Applications for Switch-Destined Traffic (continued)

Protocol	Behavior when EIS is Enabled	Behavior when EIS is Disabled
ssh	EIS Behavior	Default Behavior
Snmp (snmp mib response)	EIS Behavior	Default Behavior
telnet	EIS Behavior	Default Behavior
icmp (ping and traceroute)	EIS Behavior for ICMP	Default Behavior

Interworking of EIS With Various Applications

Stacking

- The management EIS is enabled on the master and the standby unit.
- Because traffic can be initiated from the Master unit only, the preference to management EIS table for switch-initiated traffic and all its related ARP processing is done in the Master unit only.
- ARP-related processing for switch-destined traffic is done by both master and standby units.

VLT

VLT feature is for the front-end port only. Because this feature is specific to the management port, this feature can coexist with VLT and nothing specific needs to be done in this feature to handle VLT scenario.

DHCP

- If DHCP Client is enabled on the management port, a management default route is installed to the switch.
- If management EIS is enabled, this default route is added to the management EIS routing table and the default routing table.

ARP learn enable

- When ARP learn enable is enabled, the switch learns ARP entries for ARP Request packets even if the packet is not destined to an IP configured in the box.
- The ARP learn enable feature is not applicable to the EIS routing table. It is applicable to the default routing table only to avoid unnecessary double ARP entries

Sflow

sFlow management application is supported only in standalone boxes and switch shall throw error message if sFlow is configured in stacking environment

Designating a Multicast Router Interface

To designate an interface as a multicast router interface, use the following command.

Dell EMC Networking OS also has the capability of listening in on the incoming IGMP general queries and designate those interfaces as the multicast router interface when the frames have a non-zero IP source address. All IGMP control packets and IP multicast data traffic originating from receivers is forwarded to multicast router interfaces.

- Designate an interface as a multicast router interface.

  ```
  ip igmp snooping mrouter interface
  ```

21

Interfaces

This chapter describes interface types, both physical and logical, and how to configure them with Dell EMC Networking Operating System (OS).

The system supports 10 Gigabit Ethernet and 40 Gigabit Ethernet interfaces.

NOTE: Only Dell-qualified optics are supported on these interfaces. Non-Dell 40G optics are set to error-disabled state.

Basic Interface Configuration

- Interface Types
- View Basic Interface Information
- Enabling a Physical Interface
- Physical Interfaces
- Management Interfaces
- VLAN Interfaces
- Loopback Interfaces
- Null Interfaces
- Port Channel Interfaces

Advanced Interface Configuration

- Bulk Configuration
- Defining Interface Range Macros
- Monitoring and Maintaining Interfaces
- Link Dampening
- Link Bundle Monitoring
- Ethernet Pause Frames
- Configure the MTU Size on an Interface
- Port-pipes
- Auto-Negotiation on Ethernet Interfaces
- View Advanced Interface Information

Topics:

- Interface Types
- View Basic Interface Information
- Resetting an Interface to its Factory Default State
- Enabling Energy Efficient Ethernet
- View EEE Information
- Clear EEE Counters
- Enabling a Physical Interface
- Physical Interfaces
- Egress Interface Selection (EIS)
- Management Interfaces
- VLAN Interfaces
- Loopback Interfaces
- Null Interfaces
- Port Channel Interfaces
- Bulk Configuration

- Defining Interface Range Macros
- Monitoring and Maintaining Interfaces
- Configuring wavelength for 10–Gigabit SFP+ optics
- Link Dampening
- Link Bundle Monitoring
- Using Ethernet Pause Frames for Flow Control
- Configure the MTU Size on an Interface
- Port-Pipes
- Auto-Negotiation on Ethernet Interfaces
- View Advanced Interface Information
- Configuring the Traffic Sampling Size Globally
- Dynamic Counters
- Compressing Configuration Files

Interface Types

The following table describes different interface types.

Table 41. Different Types of Interfaces

Interface Type	Modes Possible	Default Mode	Requires Creation	Default State
Physical	L2, L3	Unset	No	Shutdown (disabled)
Management	N/A	N/A	No	No Shutdown (enabled)
Loopback	L3	L3	Yes	No Shutdown (enabled)
Null	N/A	N/A	No	Enabled
Port Channel	L2, L3	L3	Yes	Shutdown (disabled)
VLAN	L2, L3	L2	Yes (except default)	L2 - Shutdown (disabled) L3 - No Shutdown (enabled)

View Basic Interface Information

To view basic interface information, use the following command.

You have several options for viewing interface status and configuration parameters.

- Lists all configurable interfaces on the chassis.

  EXEC mode

  `show interfaces`

  This command has options to display the interface status, IP and MAC addresses, and multiple counters for the amount and type of traffic passing through the interface.

  If you configured a port channel interface, this command lists the interfaces configured in the port channel.

  NOTE: To end output from the system, such as the output from the `show interfaces` command, enter `CTRL+C` and Dell EMC Networking OS returns to the command prompt.

  NOTE: The CLI output may be incorrectly displayed as 0 (zero) for the Rx/Tx power values. To obtain the correct power information, perform a simple network management protocol (SNMP) query.

The following example shows the configuration and status information for one interface.

```
DellEMC#show interfaces tengigabitethernet 1/1
TenGigabitEthernet 1/1 is up, line protocol is up
```

```
Hardware is Force10Eth, address is 00:01:e8:05:f3:6a
  Current address is 00:01:e8:05:f3:6a
Pluggable media present, XFP type is 10GBASE-LR.
  Medium is MultiRate, Wavelength is 1310nm
  XFP receive power reading is -3.7685
Interface index is 67436603
Internet address is 65.113.24.238/28
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit, Mode full duplex, Master
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:09:54
Queueing strategy: fifo
Input Statistics:
  0 packets, 0 bytes
  0 Vlans
  0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  0 Multicasts, 0 Broadcasts
  0 runts, 0 giants, 0 throttles
  0 CRC, 0 overrun, 0 discarded
Output Statistics:
  3 packets, 192 bytes, 0 underruns
  3 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  0 Multicasts, 3 Broadcasts, 0 Unicasts
  0 Vlans, 0 throttles, 0 discarded, 0 collisions
Rate info (interval 299 seconds):
  Input 00.00 Mbits/sec, 0 packets/sec, 0.00% of line-rate
  Output 00.00 Mbits/sec, 0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:00:31
DellEMC#
```

To view which interfaces are enabled for Layer 3 data transmission, use the `show ip interfaces brief` command in EXEC Privilege mode. In the following example, TenGigabitEthernet interface 1/6 is in Layer 3 mode because an IP address has been assigned to it and the interface's status is operationally up.

```
DellEMC#show ip interface brief
Interface                IP-Address  OK? Method  Status                  Protocol
TenGigabitEthernet 1/1   unassigned  NO  Manual  administratively down  down
TenGigabitEthernet 1/2   unassigned  NO  Manual  administratively down  down
TenGigabitEthernet 1/3   unassigned  YES Manual  up                      up
TenGigabitEthernet 1/4   unassigned  YES Manual  up                      up
TenGigabitEthernet 1/5   unassigned  YES Manual  up                      up
TenGigabitEthernet 1/6   10.10.10.1  YES Manual  up                      up
TenGigabitEthernet 1/7   unassigned  NO  Manual  administratively down  down
TenGigabitEthernet 1/8   unassigned  NO  Manual  administratively down  down
TenGigabitEthernet 1/9   unassigned  NO  Manual  administratively down  down
```

To view only configured interfaces, use the `show interfaces configured` command in the EXEC Privilege mode.

To determine which physical interfaces are available, use the `show running-config` command in EXEC mode. This command displays all physical interfaces available on the system.

```
DellEMC#show running
Current Configuration ...
!
interface TenGigabitEthernet 2/6
  no ip address
  shutdown
!
interface TenGigabitEthernet 2/7
  no ip address
  shutdown
!
interface TenGigabitEthernet 2/8
  no ip address
  shutdown
!
interface TenGigabitEthernet 2/9
  no ip address
  shutdown
```

Resetting an Interface to its Factory Default State

You can reset the configurations applied on an interface to its factory default state. To reset the configuration, perform the following steps:

1. View the configurations applied on an interface.

   INTERFACE mode

   `show config`

   ```
   DellEMC(conf-if-te-1/5)#show config
   !
   interface TenGigabitEthernet 1/5
   no ip address
   portmode hybrid
   switchport
   rate-interval 8
   mac learning-limit 10 no-station-move
   no shutdown
   ```

2. Reset an interface to its factory default state.

   CONFIGURATION mode

   `default interface` *interface-type*]

   ```
   DellEMC(conf)#default interface tengigabitethernet 1/5
   ```

3. Verify the configuration.

   INTERFACE mode

   `show config`

   ```
   DellEMC(conf-if-te-1/5)#show config
   !
   interface TenGigabitEthernet 1/5
    no ip address
    shutdown
   ```

   All the applied configurations are removed and the interface is set to the factory default state.

Enabling Energy Efficient Ethernet

Energy Efficient Ethernet (EEE) is an IEEE 802.3 az standard that reduces power consumptions on Ethernet ports. EEE stops the transmission when there is no data to be transmitted and resumes the transmission at the arrival of new packets. You can enable EEE only on ten Gigabit native or optional module copper ports.

1. To enable EEE, use the `eee` command.

   INTERFACE mode

   ```
   Dell(conf)# interface tengigabitethernet 1/1
   Dell(conf-if-te-1/1)# eee
   ```

2. To disable EEE, use the `no eee` command.

   INTERFACE mode

   ```
   Dell(conf)# interface tengigabitethernet 1/1
   Dell(conf-if-te-1/1)# no eee
   ```

View EEE Information

To view the details of Energy Efficient Ethernet (EEE), you can use the following `show` commands.

You have several options for viewing the details of EEE on interfaces.

- List all the interfaces.

  EXEC mode

  EXEC PRIVILEGE mode

  `show interfaces`

  This command displays the status of each interface with various details along with the information whether EEE is enabled on the interfaces..
- List the status of eee on all the interfaces, on a specified port, or on a range of ports.

  EXEC mode

  EXEC PRIVILEGE mode

  `show interfaces eee`

  `show interfaces` *`interface-type slot/port`* `eee`

  `show interfaces` *`interface-type slot/port-range`* `eee`
- List the statistical information of eee on all the interfaces, on a specified port, or on a range of ports.

  EXEC mode

  EXEC PRIVILEGE mode

  `show interfaces eee statistics`

  `show interfaces` *`interface-type slot/port`* `eee statistics`

  `show interfaces` *`interface-type slot/port-range`* `eee statistics`
- List the hardware counters on a specified interface or a specified stack unit.

  EXEC mode

  EXEC PRIVILEGE mode

  `show hardware counters` *`interface-type slot/port`*

  `show hardware stack-unit` *`stack-unit-number`* `unit` *`unit-number`* `counters`

The following example shows the status information for all the interfaces.

S4820T

```
Dell# show interfaces
TenGigabitEthernet 2/1 is down, line protocol is down
Hardware is DellForce10Eth, address is 00:01:e8:8b:3d:e7
Current address is 00:01:e8:8b:3d:e7
Pluggable media present, Media type is unknown
Wavelength unknown
Interface index is 100992002
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
Flowcontrol rx on tx off
Energy Efficient Ethernet : Yes
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 3d17h53m
Queueing strategy: fifo
Input Statistics:
  0 packets, 0 bytes
  0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  0 Multicasts, 0 Broadcasts
  0 runts, 0 giants, 0 throttles
  0 CRC, 0 overrun, 0 discarded
Output Statistics:
  0 packets, 0 bytes, 0 underruns
  0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  0 Multicasts, 0 Broadcasts, 0 Unicasts
```

```
  0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
  Input 00.00 Mbits/sec, 0 packets/sec, 0.00% of line-rate
  Output 00.00 Mbits/sec, 0 packets/sec, 0.00% of line-rate
Time since last interface status change: 3d17h51m
```

The following example shows the status of EEE on a specific interface.

S4820T

```
Dell#show interfaces tengigabitethernet 0/5 eee
Interface                      EEE Status      Speed       Duplex
TenGigabitEthernet 0/5         Yes up          1000 Mbit   Full
```

The following example shows the statistical information of EEE on a specific interface.

S4820T

```
Dell#show interfaces tengigabitethernet 0/5 eee statistics
Port            EEE    LPI    TxLPIEventCount TxLPIDuration  RxLPIEventCount
RxLPIDuration
Te 0/5          Yes    Yes    12              32546138       0               0
```

The following example shows the hardware counters on a specified interface.

```
Dell#show hardware counters interface gigabitethernet 10/1
unit: 0 port: 1 (interface Gi 10/1)
Description                                    Value

RX - IPV4 L3 Unicast Frame Counter               0
RX - IPV4 L3 Routed Multicast Packets            0
RX - IPV6 L3 Unicast Frame Counter               0
RX - IPV6 L3 Routed Multicast Packets            0
RX - Unicast Packet Counter                      0
RX - 64 Byte Frame Counter                       0
RX - 65 to 127 Byte Frame Counter                0
RX - 128 to 255 Byte Frame Counter               0
RX - 256 to 511 Byte Frame Counter               0
RX - 512 to 1023 Byte Frame Counter              0
RX - 1024 to 1518 Byte Frame Counter             0
RX - 1519 to 1522 Byte Good VLAN Frame Counter   0
RX - 1519 to 2047 Byte Frame Counter             0
RX - 2048 to 4095 Byte Frame Counter             0
RX - 4096 to 9216 Byte Frame Counter             0
RX - Good Packet Counter                         0
RX - Packet/Frame Counter                        0
RX - Unicast Frame Counter                       0
RX - Multicast Frame Counter                     0
RX - Broadcast Frame Counter                     0
RX - Byte Counter                                0
RX - Control Frame Counter                       0
RX - Pause Control Frame Counter                 0
RX - Oversized Frame Counter                     0
RX - Jabber Frame Counter                        0
RX - VLAN Tag Frame Counter                      0
RX - Double VLAN Tag Frame Counter               0
RX - RUNT Frame Counter                          0
RX - Fragment Counter                            0
RX - VLAN Tagged Packets                         0
RX - Ingress Dropped Packet                      0
RX - MTU Check Error Frame Counter               0
RX - PFC Frame Priority 0                        0
RX - PFC Frame Priority 1                        0
RX - PFC Frame Priority 2                        0
RX - PFC Frame Priority 3                        0
RX - PFC Frame Priority 4                        0
RX - PFC Frame Priority 5                        0
RX - PFC Frame Priority 6                        0
RX - PFC Frame Priority 7                        0
RX - Debug Counter 0                             0
RX - Debug Counter 1                             0
RX - Debug Counter 2                             0
```

```
RX - Debug Counter 3                                0
RX - Debug Counter 4                                0
RX - Debug Counter 5                                0
RX - Debug Counter 6                                0
RX - Debug Counter 7                                0
RX - Debug Counter 8                                0
RX - EEE LPI Event Counter                          0
RX - EEE LPI Duration Counter                       0
TX - 64 Byte Frame Counter                          0
TX - 65 to 127 Byte Frame Counter                   0
TX - 128 to 255 Byte Frame Counter                  0
TX - 256 to 511 Byte Frame Counter                  0
TX - 512 to 1023 Byte Frame Counter                 0
TX - 1024 to 1518 Byte Frame Counter                0
TX - 1519 to 1522 Byte Good VLAN Frame Counter      0
TX - 1519 to 2047 Byte Frame Counter                0
TX - 2048 to 4095 Byte Frame Counter                0
TX - 4096 to 9216 Byte Frame Counter                0
TX - Good Packet Counter                            0
TX - Packet/Frame Counter                           0
TX - Unicast Frame Counter                          0
TX - Multicast Frame Counter                        0
TX - Broadcast Frame Counter                        0
TX - Byte Counter                                   0
TX - Control Frame Counter                          0
TX - Pause Control Frame Counter                    0
TX - Oversized Frame Counter                        0
TX - Jabber Counter                                 0
TX - VLAN Tag Frame Counter                         0
TX - Double VLAN Tag Frame Counter                  0
TX - RUNT Frame Counter                             0
TX - Fragment Counter                               0
TX - PFC Frame Priority 0                           0
TX - PFC Frame Priority 1                           0
TX - PFC Frame Priority 2                           0
TX - PFC Frame Priority 3                           0
TX - PFC Frame Priority 4                           0
TX - PFC Frame Priority 5                           0
TX - PFC Frame Priority 6                           0
TX - PFC Frame Priority 7                           0
TX - Debug Counter 0                                0
TX - Debug Counter 1                                0
TX - Debug Counter 2                                0
TX - Debug Counter 3                                0
TX - Debug Counter 4                                0
TX - Debug Counter 5                                0
TX - Debug Counter 6                                0
TX - Debug Counter 7                                0
TX - Debug Counter 8                                0
TX - Debug Counter 9                                0
TX - Debug Counter 10                               0
TX - Debug Counter 11                               0
TX - EEE LPI Event Counter                          0
TX - EEE LPI Duration Counter                       0
<<Output Truncated>>
```

The following example shows the hardware counters on specified stack unit.

```
Dell#show hardware stack-unit 10 unit 0 counters
unit: 0 port: 1 (interface Gi 10/1)
Description                                    Value

RX - IPV4 L3 Unicast Frame Counter               0
RX - IPV4 L3 Routed Multicast Packets            0
RX - IPV6 L3 Unicast Frame Counter               0
RX - IPV6 L3 Routed Multicast Packets            0
RX - Unicast Packet Counter                      0
RX - 64 Byte Frame Counter                       0
RX - 65 to 127 Byte Frame Counter                0
RX - 128 to 255 Byte Frame Counter               0
RX - 256 to 511 Byte Frame Counter               0
RX - 512 to 1023 Byte Frame Counter              0
RX - 1024 to 1518 Byte Frame Counter             0
```

```
RX - 1519 to 1522 Byte Good VLAN Frame Counter     0
RX - 1519 to 2047 Byte Frame Counter               0
RX - 2048 to 4095 Byte Frame Counter               0
RX - 4096 to 9216 Byte Frame Counter               0
RX - Good Packet Counter                           0
RX - Packet/Frame Counter                          0
RX - Unicast Frame Counter                         0
RX - Multicast Frame Counter                       0
RX - Broadcast Frame Counter                       0
RX - Byte Counter                                  0
RX - Control Frame Counter                         0
RX - Pause Control Frame Counter                   0
RX - Oversized Frame Counter                       0
RX - Jabber Frame Counter                          0
RX - VLAN Tag Frame Counter                        0
RX - Double VLAN Tag Frame Counter                 0
RX - RUNT Frame Counter                            0
RX - Fragment Counter                              0
RX - VLAN Tagged Packets                           0
RX - Ingress Dropped Packet                        0
RX - MTU Check Error Frame Counter                 0
RX - PFC Frame Priority 0                          0
RX - PFC Frame Priority 1                          0
RX - PFC Frame Priority 2                          0
RX - PFC Frame Priority 3                          0
RX - PFC Frame Priority 4                          0
RX - PFC Frame Priority 5                          0
RX - PFC Frame Priority 6                          0
RX - PFC Frame Priority 7                          0
RX - Debug Counter 0                               0
RX - Debug Counter 1                               0
RX - Debug Counter 2                               0
RX - Debug Counter 3                               0
RX - Debug Counter 4                               0
RX - Debug Counter 5                               0
RX - Debug Counter 6                               0
RX - Debug Counter 7                               0
RX - Debug Counter 8                               0
RX - EEE LPI Event Counter                         0
RX - EEE LPI Duration Counter                      0
TX - 64 Byte Frame Counter                         0
TX - 65 to 127 Byte Frame Counter                  0
TX - 128 to 255 Byte Frame Counter                 0
TX - 256 to 511 Byte Frame Counter                 0
TX - 512 to 1023 Byte Frame Counter                0
TX - 1024 to 1518 Byte Frame Counter               0
TX - 1519 to 1522 Byte Good VLAN Frame Counter     0
TX - 1519 to 2047 Byte Frame Counter               0
TX - 2048 to 4095 Byte Frame Counter               0
TX - 4096 to 9216 Byte Frame Counter               0
TX - Good Packet Counter                           0
TX - Packet/Frame Counter                          0
TX - Unicast Frame Counter                         0
TX - Multicast Frame Counter                       0
TX - Broadcast Frame Counter                       0
TX - Byte Counter                                  0
TX - Control Frame Counter                         0
TX - Pause Control Frame Counter                   0
TX - Oversized Frame Counter                       0
TX - Jabber Counter                                0
TX - VLAN Tag Frame Counter                        0
TX - Double VLAN Tag Frame Counter                 0
TX - RUNT Frame Counter                            0
TX - Fragment Counter                              0
TX - PFC Frame Priority 0                          0
TX - PFC Frame Priority 1                          0
TX - PFC Frame Priority 2                          0
TX - PFC Frame Priority 3                          0
TX - PFC Frame Priority 4                          0
TX - PFC Frame Priority 5                          0
TX - PFC Frame Priority 6                          0
TX - PFC Frame Priority 7                          0
TX - Debug Counter 0                               0
```

```
TX - Debug Counter 1                                0
TX - Debug Counter 2                                0
TX - Debug Counter 3                                0
TX - Debug Counter 4                                0
TX - Debug Counter 5                                0
TX - Debug Counter 6                                0
TX - Debug Counter 7                                0
TX - Debug Counter 8                                0
TX - Debug Counter 9                                0
TX - Debug Counter 10                               0
TX - Debug Counter 11                               0
TX - EEE LPI Event Counter                          0
TX - EEE LPI Duration Counter                       0

<<Output Truncated>>
```

Clear EEE Counters

To clear the EEE counters, use the following the commands.

- Clear the EEE counters on all the copper ports.

  EXEC Privilege mode

  `clear counters eee`
- Clear the EEE counters on the specified port.

  EXEC Privilege mode

  `clear counters` *`interface-type slot/port`* `eee`
- Clear the EEE counters on the specified range of ports.

  EXEC Privilege mode

  `clear counters` *`interface-type slot/port-range`* `eee`

When you use this command, confirm that you want Dell EMC Networking OS to clear the EEE counters.

```
DellEMC# clear counters eee
Clear eee counters on all Copper Ports [confirm] y
DellEMC#

DellEMC# clear counters tengigabitethernet 1/1 eee
Clear eee counters on Te 1/1 [confirm] y
DellEMC#

DellEMC# clear counters tengigabitethernet 1/1-10 eee
Clear eee counters on Te 1/1-10 [confirm] y
DellEMC#
```

Enabling a Physical Interface

After determining the type of physical interfaces available, to enable and configure the interfaces, enter INTERFACE mode by using the `interface` *`interface`* command.

1. Enter the keyword `interface` then the type of interface and slot/port information.

   CONFIGURATION mode

   `interface` *`interface`*

   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
2. Enable the interface.

   INTERFACE mode

   `no shutdown`

To confirm that the interface is enabled, use the `show config` command in INTERFACE mode. To leave INTERFACE mode, use the `exit` command or `end` command. You cannot delete a physical interface.

Physical Interfaces

The *Management Ethernet interface* is a single RJ-45 Fast Ethernet port on a switch.

The interface provides dedicated management access to the system.

Stack-unit interfaces support Layer 2 and Layer 3 traffic over the and 40-Gigabit Ethernet interfaces. These interfaces can also become part of virtual interfaces such as virtual local area networks (VLANs) or port channels.

For more information about VLANs, refer to Bulk Configuration. For more information on port channels, refer to Port Channel Interfaces.

Dell EMC Networking OS Behavior: The system uses a single MAC address for all physical interfaces.

Configuration Task List for Physical Interfaces

By default, all interfaces are operationally disabled and traffic does not pass through them.

The following section includes information about optional configurations for physical interfaces:

- Overview of Layer Modes
- Configuring Layer 2 (Data Link) Mode
- Configuring Layer 2 (Interface) Mode
- Management Interfaces
- Auto-Negotiation on Ethernet Interfaces
- Adjusting the Keepalive Timer
- Clearing Interface Counters

Overview of Layer Modes

On all systems running Dell EMC Networking OS, you can place physical interfaces, port channels, and VLANs in Layer 2 mode or Layer 3 mode.

By default, VLANs are in Layer 2 mode.

Table 42. Layer Modes

Type of Interface	Possible Modes	Requires Creation	Default State
10 Gigabit Ethernet and 40 Gigabit Ethernet	Layer 2 Layer 3	No	Shutdown (disabled)
Management	N/A	No	Shutdown (disabled)
Loopback	Layer 3	Yes	No shutdown (enabled)
Null interface	N/A	No	Enabled
Port Channel	Layer 2 Layer 3	Yes	Shutdown (disabled)
VLAN	Layer 2 Layer 3	Yes, except for the default VLAN.	No shutdown (active for Layer 2) Shutdown (disabled for Layer 3)

Configuring Layer 2 (Data Link) Mode

Do not configure switching or Layer 2 protocols such as spanning tree protocol (STP) on an interface unless the interface has been set to Layer 2 mode.

To set Layer 2 data transmissions through an individual interface, use the following command.

- Enable Layer 2 data transmissions through an individual interface.

  INTERFACE mode

  `switchport`

```
DellEMC(conf-if)#show config
!
interface Port-channel 1
  no ip address
  switchport
  no shutdown
DellEMC(conf-if)#
```

Configuring Layer 2 (Interface) Mode

To configure an interface in Layer 2 mode, use the following commands.

- Enable the interface.

  INTERFACE mode

  `no shutdown`
- Place the interface in Layer 2 (switching) mode.

  INTERFACE mode

  `switchport`

To view the interfaces in Layer 2 mode, use the `show interfaces switchport` command in EXEC mode.

Configuring Layer 3 (Network) Mode

When you assign an IP address to a physical interface, you place it in Layer 3 mode.

To enable Layer 3 mode on an individual interface, use the following commands. In all interface types except VLANs, the `shutdown` command prevents all traffic from passing through the interface. In VLANs, the `shutdown` command prevents Layer 3 traffic from passing through the interface. Layer 2 traffic is unaffected by the `shutdown` command. One of the interfaces in the system must be in Layer 3 mode before you configure or enter a Layer 3 protocol mode (for example, OSPF).

- Enable Layer 3 on an individual interface

  INTERFACE mode

  `ip address` *`ip-address`*
- Enable the interface.

  INTERFACE mode

  `no shutdown`

If an interface is in the incorrect layer mode for a given command, an error message is displayed (shown in bold). In the following example, the `ip address` command triggered an error message because the interface is in Layer 2 mode and the `ip address` command is a Layer 3 command only.

```
DellEMC(conf-if)#show config
!
interface TenGigabitEthernet 1/2
  no ip address
  switchport
  no shutdown
DellEMC(conf-if)#ip address 10.10.1.1 /24
% Error: Port is in Layer 2 mode Te 1/2.
DellEMC(conf-if)#
```

To determine the configuration of an interface, use the `show config` command in INTERFACE mode or the various `show interface` commands in EXEC mode.

Configuring Layer 3 (Interface) Mode

To assign an IP address, use the following commands.

- Enable the interface.

  INTERFACE mode

  `no shutdown`
- Configure a primary IP address and mask on the interface.

  INTERFACE mode

  `ip address` *`ip-address mask`* `[secondary]`

  The *`ip-address`* must be in dotted-decimal format (A.B.C.D) and the mask must be in slash format (/xx).

  Add the keyword `secondary` if the IP address is the interface's backup IP address.

You can only configure one primary IP address per interface. You can configure up to 255 secondary IP addresses on a single interface.

To view all interfaces to see with an IP address assigned, use the `show ip interfaces brief` command in EXEC mode as shown in View Basic Interface Information.

To view IP information on an interface in Layer 3 mode, use the `show ip interface` command in EXEC Privilege mode.

```
DellEMC>show ip interface vlan 58
Vlan 58 is up, line protocol is up
Internet address is 1.1.49.1/24
Broadcast address is 1.1.49.255
Address determined by config file
MTU is 1554 bytes
Inbound access list is not set
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent
IP unicast RPF check is not supported
```

Egress Interface Selection (EIS)

EIS allows you to isolate the management and front-end port domains by preventing switch-initiated traffic routing between the two domains. This feature provides additional security by preventing flooding attacks on front-end ports. The following protocols support EIS: DNS, FTP, NTP, RADIUS, sFlow, SNMP, SSH, Syslog, TACACS, Telnet, and TFTP. This feature does not support sFlow on stacked units.

When you enable this feature, all management routes (connected, static, and default) are copied to the management EIS routing table. Use the `management route` command to add new management routes to the default and EIS routing tables. Use the `show ip management-eis-route` command to view the EIS routes.

Important Points to Remember

- Deleting a management route removes the route from both the EIS routing table and the default routing table.
- If the management port is down or route lookup fails in the management EIS routing table, the outgoing interface is selected based on route lookup from the default routing table.
- If a route in the EIS table conflicts with a front-end port route, the front-end port route has precedence.
- Due to protocol, ARP packets received through the management port create two ARP entries (one for the lookup in the EIS table and one for the default routing table).

Configuring EIS

EIS is compatible with the following protocols: DNS, FTP, NTP, RADIUS, sFlow, SNMP, SSH, Syslog, TACACS, Telnet, and TFTP.

To enable and configure EIS, use the following commands:

1. Enter EIS mode.
   CONFIGURATION mode
   ```
   management egress-interface-selection
   ```
2. Configure which applications uses EIS.
   EIS mode
   ```
   application {all | application-type}
   ```
   NOTE: If you configure SNMP as the management application for EIS and you add a default management route, when you perform an SNMP walk and check the debugging logs for the source and destination IPs, the SNMP agent uses the destination address of incoming SNMP packets as the source address for outgoing SNMP responses for security.

Management Interfaces

The system supports the Management Ethernet interface as well as the standard interface on any port. You can use either method to connect to the system.

Configuring Management Interfaces

The dedicated Management interface provides management access to the system.

You can configure this interface using the CLI, but the configuration options on this interface are limited. You cannot configure Gateway addresses and IP addresses if it appears in the main routing table of Dell EMC Networking OS. In addition, proxy ARP is not supported on this interface.

To configure a management interface, use the following commands.

- Enter the slot and the port (0) to configure a Management interface.
  CONFIGURATION mode
  ```
  interface managementethernet interface
  ```
  The slot range is 0.
- Configure an IP address and mask on a Management interface.
  INTERFACE mode
  ```
  ip address ip-address mask
  ```
  - *ip-address mask*: enter an address in dotted-decimal format (A.B.C.D). The mask must be in /prefix format (/x).

You can configure two global IPv6 addresses on the system in EXEC Privilege mode. To view the addresses, use the `show interface managementethernet` command, as shown in the following example. If you try to configure a third IPv6 address, an error message displays. If you enable auto-configuration, all IPv6 addresses on that management interface are auto-configured. The first IPv6 address that you configure on the management interface is the primary address. If deleted, you must re-add it; the secondary address is not promoted.

The following rules apply to having two IPv6 addresses on a management interface:

- IPv6 addresses on a single management interface cannot be in the same subnet.
- IPv6 secondary addresses on management interfaces:
  - across a platform *must* be in the same subnet.
  - must not match the virtual IP address and must not be in the same subnet as the virtual IP.

```
DellEMC#show interfaces managementethernet 0/0
ManagementEthernet 0/0 is up, line protocol is up
Hardware is DellForce10Eth, address is 00:01:e8:a0:bf:f3
    Current address is 00:01:e8:a0:bf:f3
Pluggable media not present
Interface index is 302006472
Internet address is 10.16.130.5/16
Link local IPv6 address: fe80::201:e8ff:fea0:bff3/64
Global IPv6 address: 1::1/
Global IPv6 address: 2::1/64
Virtual-IP is not set
```

```
Virtual-IP IPv6 address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode full duplex
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:06:14
Queueing strategy: fifo
    Input 791 packets, 62913 bytes, 775 multicast
    Received 0 errors, 0 discarded
    Output 21 packets, 3300 bytes, 20 multicast
    Output 0 errors, 0 invalid protocol
Time since last interface status change: 00:06:03
```

If there are two RPMs on the system, configure each Management interface with a different IP address. Unless you configure the `management route` command, you can only access the Management interface from the local LAN. To access the Management interface from another LAN, configure the `management route` command to point to the Management interface.

Alternatively, you can use the `virtual-ip` command to manage a system with one or two RPMs. A virtual IP is an IP address assigned to the system (not to any management interfaces) and is a CONFIGURATION mode command. When a virtual IP address is assigned to the system, the active management interface of the RPM is recognized by the virtual IP address — not by the actual interface IP address assigned to it. During an RPM failover, you do not have to remember the IP address of the new RPM's management interface — the system still recognizes the virtual-IP address.

- `virtual-ip` is a CONFIGURATION mode command.
- When applied, the management port on the primary RPM assumes the virtual IP address. Executing the `show interfaces` and `show ip interface brief` commands on the primary RPM management interface displays the virtual IP address and not the actual IP address assigned on that interface.
- A duplicate IP address message is printed for the management port's virtual IP address on an RPM failover. This behavior is a harmless error that is generated due to a brief transitory moment during failover when both RPMs' management ports own the virtual IP address, but have different MAC addresses.
- The primary management interface uses only the virtual IP address if it is configured. The system cannot be accessed through the native IP address of the primary RPM's management interface.
- After the virtual IP address is removed, the system is accessible through the native IP address of the primary RPM's management interface.
- Primary and secondary management interface IP and virtual IP must be in the same subnet.

To view the Primary RPM Management port, use the `show interface Managementethernet` command in EXEC Privilege mode. If there are two RPMs, you cannot view information on that interface.

Configuring a Management Interface on an Ethernet Port

You can manage the system through any port using remote access such as Telnet.

To configure an IP address for the port, use the following commands. There is no separate management routing table, so configure all routes in the IP routing table (the `ip route` command).

- Configure an IP address.

  INTERFACE mode

  `ip address` *`ip-address mask`*
- Enable the interface.

  INTERFACE mode

  `no shutdown`
- The interface is the management interface.

  INTEFACE mode

  `description`

To display the configuration for a given port, use the `show interface` command in EXEC Privilege mode, as shown in the following example. To display the routing table, use the `show ip route` command in EXEC Privilege mode.

```
DellEMC#show int TenGigabitEthernet 1/1
TenGigabitEthernet 1/1 is up, line protocol is up
Description: This is the Managment Interface
Hardware is DellEMCEth, address is 00:01:e8:cc:cc:ce
  Current address is 00:01:e8:cc:cc:ce
Pluggable media not present
Interface index is 46449666
```

```
Internet address is 10.11.131.240/23
[output omitted]
DellEMC#show ip route
Codes: C - connected, S - static, R - RIP,
  B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
  O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
  N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
  E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
  L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
  > - non-active route, + - summary route

Gateway of last resort is 10.11.131.254 to network 0.0.0.0

    Destination     Gateway                     Dist/Metric Last Change
    -----------     -------                     ----------- -----------
*S  0.0.0.0/0       via 10.11.131.254, Te 1/1   1/0 1d2h
C   10.11.130.0/23  Direct, Te 1/1              0/0 1d2h
DellEMC#
```

VLAN Interfaces

VLANs are logical interfaces and are, by default, in Layer 2 mode. Physical interfaces and port channels can be members of VLANs.

For more information about VLANs and Layer 2, see Layer 2 and Virtual LANs (VLANs).

NOTE: To monitor VLAN interfaces, use Management Information Base for Network Management of TCP/IP-based internets: MIB-II (RFC 1213).

NOTE: You cannot simultaneously use egress rate shaping and ingress rate policing on the same VLAN.

Dell EMC Networking OS supports Inter-VLAN routing (Layer 3 routing in VLANs). You can add IP addresses to VLANs and use them in routing protocols in the same manner that physical interfaces are used. For more information about configuring different routing protocols, refer to the chapters on the specific protocol.

A consideration for including VLANs in routing protocols is that you must configure the `no shutdown` command. (For routing traffic to flow, you must enable the VLAN.)

NOTE: You cannot assign an IP address to the default VLAN, which is VLAN 1 (by default). To assign another VLAN ID to the default VLAN, use the `default vlan-id vlan-id` command.

To assign an IP address to an interface, use the following command.

- Configure an IP address and mask on the interface.

  INTERFACE mode

  `ip address` *`ip-address mask`* `[secondary]`
  - *`ip-address mask`*: enter an address in dotted-decimal format (A.B.C.D). The mask must be in slash format (/24).
  - `secondary`: the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

```
interface Vlan 10
  ip address 1.1.1.2/24
  tagged TenGigabitEthernet 2/2-13
  tagged TenGigabitEthernet 5/1
  ip ospf authentication-key force10
  ip ospf cost 1
  ip ospf dead-interval 60
  ip ospf hello-interval 15
  no shutdown
!
```

Loopback Interfaces

A Loopback interface is a virtual interface in which the software emulates an interface. Packets routed to it are processed locally.

Because this interface is not a physical interface, you can configure routing protocols on this interface to provide protocol stability. You can place Loopback interfaces in default Layer 3 mode.

To configure, view, or delete a Loopback interface, use the following commands.

- Enter a number as the Loopback interface.

  CONFIGURATION mode

  `interface loopback` *`number`*

  The range is from 0 to 16383.
- View Loopback interface configurations.

  EXEC mode

  `show interface loopback` *`number`*
- Delete a Loopback interface.

  CONFIGURATION mode

  `no interface loopback` *`number`*

Many of the commands supported on physical interfaces are also supported on a Loopback interface.

Null Interfaces

The Null interface is another virtual interface. There is only one Null interface. It is always up, but no traffic is transmitted through this interface.

To enter INTERFACE mode of the Null interface, use the following command.

- Enter INTERFACE mode of the Null interface.

  CONFIGURATION mode

  `interface null 0`

The only configurable command in INTERFACE mode of the Null interface is the `ip unreachable` command.

Port Channel Interfaces

Port channel interfaces support link aggregation, as described in IEEE Standard 802.3ad.

This section covers the following topics:

- Port Channel Definition and Standards
- Port Channel Benefits
- Port Channel Implementation
- Configuration Tasks for Port Channel Interfaces

Port Channel Definition and Standards

Link aggregation is defined by IEEE 802.3ad as a method of grouping multiple physical interfaces into a single logical interface—a link aggregation group (LAG) or port channel.

A LAG is “a group of links that appear to a MAC client as if they were a single link” according to IEEE 802.3ad. In Dell EMC Networking OS, a LAG is referred to as a port channel interface.

A port channel provides redundancy by aggregating physical interfaces into one logical interface. If one physical interface goes down in the port channel, another physical interface carries the traffic.

Port Channel Benefits

A port channel interface provides many benefits, including easy management, link redundancy, and sharing.

Port channels are transparent to network configurations and can be modified and managed as one interface. For example, you configure one IP address for the group and that IP address is used for all routed traffic on the port channel.

With this feature, you can create larger-capacity interfaces by utilizing a group of lower-speed links. For example, you can build a 50-Gigabit interface by aggregating five 10-Gigabit Ethernet interfaces together. If one of the five interfaces fails, traffic is redistributed across the remaining interfaces.

Port Channel Implementation

Dell EMC Networking OS supports static and dynamic port channels.

- **Static** — Port channels that are statically configured.
- **Dynamic** — Port channels that are dynamically configured using the link aggregation control protocol (LACP). For details, see Link Aggregation Control Protocol (LACP).

The port channel ID ranges from 1 to 128.

As soon as you configure a port channel, Dell EMC Networking OS treats it like a physical interface. For example, IEEE 802.1Q tagging is maintained while the physical interface is in the port channel.

Member ports of a LAG are added and programmed into the hardware in a predictable order based on the port ID, instead of in the order in which the ports come up. With this implementation, load balancing yields predictable results across device reloads.

A physical interface can belong to only one port channel at a time.

Each port channel must contain interfaces of the same interface type/speed.

Port channels can contain a mix of 1G/10G/40G. The interface speed that the port channel uses is determined by the first port channel member that is physically up. Dell EMC Networking OS disables the interfaces that do not match the interface speed that the first channel member sets. That first interface may be either the interface that is physically brought up first or was physically operating when interfaces were added to the port channel. For example, if the first operational interface in the port channel is a Tengigabit Ethernet interface, all interfaces at 10000 Mbps are kept up, and all other interfaces that are not set to 10G speed or auto negotiate are disabled.

Dell EMC Networking OS brings up the interfaces that are set to auto negotiate so that their speed is identical to the speed of the first channel member in the port channel.

Interfaces in Port Channels

When interfaces are added to a port channel, the interfaces must share a common speed. When interfaces have a configured speed different from the port channel speed, the software disables those interfaces.

The common speed is determined when the port channel is first enabled. Then, the software checks the first interface listed in the port channel configuration. If you enabled that interface, its speed configuration becomes the common speed of the port channel. If the other interfaces configured in that port channel are configured with a different speed, Dell EMC Networking OS disables them.

Port channels can contain a mix of 1G/10G/40G. The interface speed that the port channel uses is determined by the first port channel member that is physically up. Dell EMC Networking OS disables the interfaces that do not match the interface speed that the first channel member sets. That first interface may be either the interface that is physically brought up first or was physically operating when interfaces were added to the port channel. For example, if the first operational interface in the port channel is a Tengigabit Ethernet interface, all interfaces at 10000 Mbps are kept up, and all other interfaces that are not set to 10G speed or auto negotiate are disabled.

Dell EMC Networking OS brings up the interfaces that are set to auto negotiate so that their speed is identical to the speed of the first channel member in the port channel.

Configuration Tasks for Port Channel Interfaces

To configure a port channel (LAG), use the commands similar to those found in physical interfaces. By default, no port channels are configured in the startup configuration.

These are the mandatory and optional configuration tasks:

- Creating a Port Channel (mandatory)
- Adding a Physical Interface to a Port Channel (mandatory)
- Reassigning an Interface to a New Port Channel (optional)
- Configuring the Minimum Oper Up Links in a Port Channel (optional)
- Adding or Removing a Port Channel from a VLAN (optional)
- Assigning an IP Address to a Port Channel (optional)
- Deleting or Disabling a Port Channel (optional)
- Load Balancing Through Port Channels (optional)

Creating a Port Channel

You can create up to 128 port channels with up to 16 port members per group on the platform.

To configure a port channel, use the following commands.

1. Create a port channel.
   CONFIGURATION mode
   `interface port-channel id-number`
2. Ensure that the port channel is active.
   INTERFACE PORT-CHANNEL mode
   `no shutdown`

After you enable the port channel, you can place it in Layer 2 or Layer 3 mode. To place the port channel in Layer 2 mode or configure an IP address to place the port channel in Layer 3 mode, use the `switchport` command.

You can configure a port channel as you would a physical interface by enabling or configuring protocols or assigning access control lists.

Adding a Physical Interface to a Port Channel

The physical interfaces in a port channel can be on any line card in the chassis, but must be the same physical type.

NOTE: Port channels can contain a mix of Ethernet interfaces, but Dell EMC Networking OS disables the interfaces that are not the same speed of the first channel member in the port channel (refer to 10/100/1000 Mbps Interfaces in Port Channels).

You can add any physical interface to a port channel if the interface configuration is minimal. You can configure only the following commands on an interface if it is a member of a port channel:

- `description`
- `shutdown/no shutdown`
- `mtu`

NOTE: A logical port channel interface cannot have flow control. Flow control can only be present on the physical interfaces if they are part of a port channel.

NOTE: The system supports jumbo frames by default (the default maximum transmission unit (MTU) is 1554 bytes). To configure the MTU, use the `mtu` command from INTERFACE mode.

To view the interface's configuration, enter INTERFACE mode for that interface and use the `show config` command or from EXEC Privilege mode, use the `show running-config interface interface` command.

When an interface is added to a port channel, Dell EMC Networking OS recalculates the hash algorithm.

To add a physical interface to a port, use the following commands.

1. Add the interface to a port channel.

INTERFACE PORT-CHANNEL mode

`channel-member` *`interface`*

The *`interface`* variable is the physical interface type and slot/port information.

2. Double check that the interface was added to the port channel.

   INTERFACE PORT-CHANNEL mode

   `show config`

To view the port channel's status and channel members in a tabular format, use the `show interfaces port-channel brief` command in EXEC Privilege mode, as shown in the following example.

```
DellEMC#show int port brief

LAG Mode Status Uptime   Ports
1 L2L3   up     00:06:03 Te 1/6 (Up) *
                          Te 1/12 (Up)
2 L2L3   up     00:06:03 Te 1/7 (Up) *
                          Te 1/8 (Up)
                          Te 1/13 (Up)
                          Te 1/14 (Up)
DellEMC#
```

The following example shows the port channel's mode (L2 for Layer 2 and L3 for Layer 3 and L2L3 for a Layer 2-port channel assigned to a routed VLAN), the status, and the number of interfaces belonging to the port channel.

```
DellEMC>show interface port-channel 20
Port-channel 20 is up, line protocol is up
Hardware address is 00:01:e8:01:46:fa
Internet address is 1.1.120.1/24
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 2000 Mbit
Members in this channel: Te 1/10 Te 1/17
ARP type: ARPA, ARP timeout 04:00:00
Last clearing of "show interface" counters 00:00:00
Queueing strategy: fifo
  1212627 packets input, 1539872850 bytes
  Input 1212448 IP Packets, 0 Vlans 0 MPLS
  4857 64-byte pkts, 17570 over 64-byte pkts, 35209 over 127-byte pkts
  69164 over 255-byte pkts, 143346 over 511-byte pkts, 942523 over 1023-byte pkts
  Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
  42 CRC, 0 IP Checksum, 0 overrun, 0 discarded
  2456590833 packets output, 203958235255 bytes, 0 underruns
  Output 1640 Multicasts, 56612 Broadcasts, 2456532581 Unicasts
  2456590654 IP Packets, 0 Vlans, 0 MPLS
  0 throttles, 0 discarded
Rate info (interval 5 minutes):
  Input 00.01Mbits/sec, 2 packets/sec
  Output 81.60Mbits/sec, 133658 packets/sec
Time since last interface status change: 04:31:57

DellEMC>
```

When more than one interface is added to a Layer 2-port channel, Dell EMC Networking OS selects one of the active interfaces in the port channel to be the primary port. The primary port replies to flooding and sends protocol data units (PDUs). An asterisk in the `show interfaces port-channel brief` command indicates the primary port.

As soon as a physical interface is added to a port channel, the properties of the port channel determine the properties of the physical interface. The configuration and status of the port channel are also applied to the physical interfaces within the port channel. For example, if the port channel is in Layer 2 mode, you cannot add an IP address or a static MAC address to an interface that is part of that port channel. In the following example, interface 1/6 is part of port channel 5, which is in Layer 2 mode, and an error message appeared when an IP address was configured.

```
DellEMC(conf-if-portch)#show config
!
interface Port-channel 5
  no ip address
  switchport
  channel-member TenGigabitEthernet 1/6
DellEMC(conf-if-portch)#int Te 1/6
```

```
DellEMC(conf-if)#ip address 10.56.4.4 /24
% Error: Port is part of a LAG Te 1/6.
DellEMC(conf-if)#
```

Reassigning an Interface to a New Port Channel

An interface can be a member of only one port channel. If the interface is a member of a port channel, remove it from the first port channel and then add it to the second port channel.

Each time you add or remove a channel member from a port channel, Dell EMC Networking OS recalculates the hash algorithm for the port channel.

To reassign an interface to a new port channel, use the following commands.

1. Remove the interface from the first port channel.

   INTERFACE PORT-CHANNEL mode

   `no channel-member interface`
2. Change to the second port channel INTERFACE mode.

   INTERFACE PORT-CHANNEL mode

   `interface port-channel id number`
3. Add the interface to the second port channel.

   INTERFACE PORT-CHANNEL mode

   `channel-member interface`

The following example shows moving an interface from port channel 4 to port channel 3.

```
DellEMC(conf-if-po-4)#show config
!
interface Port-channel 4
  no ip address
  channel-member TenGigabitEthernet 1/8
  no shutdown
DellEMC(conf-if-po-4)#no chann tengi 1/8
DellEMC(conf-if-po-4)#int port 3
DellEMC(conf-if-po-3)#channel tengi 1/8
DellEMC(conf-if-po-3)#sho conf
!
interface Port-channel 3
  no ip address
  channel-member TenGigabitEthernet 1/8
  shutdown
DellEMC(conf-if-po-3)#
```

Configuring the Minimum Oper Up Links in a Port Channel

You can configure the minimum links in a port channel (LAG) that must be in “oper up” status to consider the port channel to be in “oper up” status.

To set the “oper up” status of your links, use the following command.

- Enter the number of links in a LAG that must be in “oper up” status.

  INTERFACE mode

  `minimum-links number`

  The default is **1**.

  Example of Configuring the Minimum Oper Up Links in a Port Channel

  ```
  DellEMC#config t
  DellEMC(conf)#int po 1
  DellEMC(conf-if-po-1)#minimum-links 5
  DellEMC(conf-if-po-1)#
  ```

Adding or Removing a Port Channel from a VLAN

As with other interfaces, you can add Layer 2 port channel interfaces to VLANs. To add a port channel to a VLAN, place the port channel in Layer 2 mode (by using the `switchport` command).

To add or remove a VLAN port channel and to view VLAN port channel members, use the following commands.

- Add the port channel to the VLAN as a tagged interface.

  INTERFACE VLAN mode

  `tagged port-channel id number`

  An interface with tagging enabled can belong to multiple VLANs.
- Add the port channel to the VLAN as an untagged interface.

  INTERFACE VLAN mode

  `untagged port-channel id number`

  An interface without tagging enabled can belong to only one VLAN.
- Remove the port channel with tagging enabled from the VLAN.

  INTERFACE VLAN mode

  `no tagged port-channel id number`

  or

  `no untagged port-channel id number`
- Identify which port channels are members of VLANs.

  EXEC Privilege mode

  `show vlan`

Configuring VLAN Tags for Member Interfaces

To configure and verify VLAN tags for individual members of a port channel, perform the following:

1. Configure VLAN membership on individual ports

   INTERFACE mode

   `DellEMC(conf-if)#vlan tagged 2,3-4`
2. Use the `switchport` command in INTERFACE mode to enable Layer 2 data transmissions through an individual interface

   INTERFACE mode

   `DellEMC(conf-if)#switchport`
3. Verify the manually configured VLAN membership (**show interfaces switchport** *interface* command).

   EXEC mode

```
DellEMC(conf)# interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#switchport
DellEMC(conf-if-te-1/1)# vlan tagged 2-5,100,4010
DellEMC#show interfaces switchport te 1/1

Codes:  U - Untagged, T - Tagged
        x - Dot1x untagged, X - Dot1x tagged
        G - GVRP tagged, M - Trunk, H - VSN tagged
        i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

Name: TenGigabitEthernet 1/1
802.1QTagged: True
Vlan membership:
Q       Vlans
T       2-5,100,4010

DellEMC#
```

Assigning an IP Address to a Port Channel

You can assign an IP address to a port channel and use port channels in Layer 3 routing protocols.

To assign an IP address, use the following command.

- Configure an IP address and mask on the interface.

  INTERFACE mode

  ```
  ip address ip-address mask [secondary]
  ```

  - *ip-address mask*: enter an address in dotted-decimal format (A.B.C.D). The mask must be in slash format (/24).
  - `secondary`: the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

Deleting or Disabling a Port Channel

To delete or disable a port channel, use the following commands.

- Delete a port channel.

  CONFIGURATION mode

  ```
  no interface portchannel channel-number
  ```

- Disable a port channel.

  ```
  shutdown
  ```

  When you disable a port channel, all interfaces within the port channel are operationally down also.

Load Balancing Through Port Channels

Dell EMC Networking OS uses hash algorithms for distributing traffic evenly over channel members in a port channel (LAG).

The hash algorithm distributes traffic among Equal Cost Multi-path (ECMP) paths and LAG members. The distribution is based on a flow, except for packet-based hashing. A flow is identified by the hash and is assigned to one link. In packet-based hashing, a single flow can be distributed on the LAG and uses one link.

Packet based hashing is used to load balance traffic across a port-channel based on the IP Identifier field within the packet. Load balancing uses source and destination packet information to get the greatest advantage of resources by distributing traffic over multiple paths when transferring data to a destination.

Dell EMC Networking OS allows you to modify the hashing algorithms used for flows and for fragments. The load-balance and hash-algorithm commands are available for modifying the distribution algorithms.

Changing the Hash Algorithm

The `load-balance` command selects the hash criteria applied to port channels.

If you do not obtain even distribution with the `load-balance` command, you can use the `hash-algorithm` command to select the hash scheme for LAG, ECMP and NH-ECMP. You can rotate or shift the 12–bit Lag Hash until the desired hash is achieved.

The `nh-ecmp` option allows you to change the hash value for recursive ECMP routes independently of non-recursive ECMP routes. This option provides for better traffic distribution over available equal cost links that involve a recursive next hop lookup.

To change to another algorithm, use the second command.

- Change the default (0) to another algorithm and apply it to ECMP, LAG hashing, or a particular line card.

  CONFIGURATION mode

  ```
  hash-algorithm {ecmp {crc16 | crc16cc | crc32MSB | crc32LSB | crc-upper | dest-ip |
  flow-based-hashing {crc16|crc16cc|crc32MSB|crc32LSB|xor1|xor2|xor4|xor8|xor16}|lsb | xor1
  | xor2 | xor4 | xor8 | xor16}[[hg {crc16 | crc16cc | crc32MSB | crc32LSB | xor1 | xor2 |
  xor4 | xor8 | xor16}]| [lag {crc16 | crc16cc | crc32MSB | crc32LSB | xor1 | xor2 | xor4
  | xor8 | xor16 }][stack-unit|linecard number | port-set number | [hg—seed seed-value |
  seedseed-value
  ```

For more information about algorithm choices, refer to the command details in the *IP Routing* chapter of the *Dell EMC Networking OS Command Reference Guide*.

- Change to another algorithm.

  CONFIGURATION mode

```
DellEMC(conf)#hash-algorithm ecmp xor 26 lag crc 26 nh-ecmp checksum 26
DellEMC(conf)#
```

The `hash-algorithm` command is specific to ECMP group. The default ECMP hash configuration is **crc-lower**. This command takes the lower 32 bits of the hash key to compute the egress port. Other options for ECMP hash-algorithms are:

- `crc16` — uses 16 bit CRC16-bisync polynomial
- `crc16cc` — uses 16 bit CRC16 using CRC16-CCITT polynomial
- `crc32LSB` — uses LSB 16 bits of computed CRC32
- `crc32MSB` — uses MSB 16 bits of computed CRC32(default)
- `crc-upper` — uses the upper 32 bits of the hash key to compute the egress port.
- `dest-ip` — uses destination IP address as part of the hash key.
- `flow-based-hashing` — uses flow-based-hashing followed by the algorithm *crc16* |*crc16cc* |*crc32MSB* |*crc32LSB* |*xor1* |*xor2* |*xor4* |*xor8* | *xor16*.
- `lsb` — uses the least significant bit of the hash key to compute the egress port.
- `xor1` — uses Upper 8 bits of CRC16-BISYNC and lower 8 bits of xor1
- `xor2` — Upper 8 bits of CRC16-BISYNC and lower 8 bits of xor2
- `xor4` —Upper 8 bits of CRC16-BISYNC and lower 8 bits of xor4
- `xor8` — Upper 8 bits of CRC16-BISYNC and lower 8 bits of xor8
- `xor16` — uses 16 bit XOR.

Bulk Configuration

Bulk configuration allows you to determine if interfaces are present for physical interfaces or configured for logical interfaces.

Interface Range

An interface range is a set of interfaces to which other commands may be applied and may be created if there is at least one valid interface within the range.

Bulk configuration excludes from configuration any non-existing interfaces from an interface range. A default VLAN may be configured only if the interface range being configured consists of only VLAN ports.

The `interface range` command allows you to create an interface range allowing other commands to be applied to that range of interfaces.

The interface range prompt offers the interface (with slot and port information) for valid interfaces. The maximum size of an interface range prompt is 32. If the prompt size exceeds this maximum, it displays (...) at the end of the output.

NOTE: Non-existing interfaces are excluded from the interface range prompt.

NOTE: When creating an interface range, interfaces appear in the order they were entered and are not sorted.

The `show range` command is available under Interface Range mode. This command allows you to display all interfaces that have been validated under the interface range context.

The `show configuration` command is also available under Interface Range mode. This command allows you to display the running configuration only for interfaces that are part of interface range.

You can avoid specifying spaces between the range of interfaces, separated by commas, that you configure by using the `interface range` command. For example, if you enter a list of interface ranges, such as `interface range fo 2/0-1,te 10/0,te 3/0,fo 0/0`, this configuration is considered valid. The comma-separated list is not required to be separated by spaces in between the ranges. You can associate multicast MAC or hardware addresses to an interface range and VLANs by using the `mac-address-table static` *`multicast-mac-address`* `vlan` *`vlan-id`* `output-range` *`interface`* command.

Bulk Configuration Examples

Use the `interface range` command for bulk configuration.

- Create a Single-Range
- Create a Multiple-Range
- Exclude Duplicate Entries
- Exclude a Smaller Port Range
- Overlap Port Ranges
- Commas
- Add Ranges

Create a Single-Range

The following is an example of a single range.

Example of the `interface range` Command (Single Range)

```
DellEMC(config)# interface range tengigabitethernet 1/1 - 23
DellEMC(config-if-range-te-1/1-23)# no shutdown
DellEMC(config-if-range-te-1/1-23)#
```

Create a Multiple-Range

The following is an example of multiple range.

Example of the `interface range` Command (Multiple Ranges)

```
DellEMC(conf)#interface range tengigabitethernet 1/5 - 10 , tengigabitethernet 1/1 ,
vlan 1
DellEMC(conf-if-range-te-1/1,te-1/5-10,vl-1)#
```

Exclude Duplicate Entries

The following is an example showing how duplicate entries are omitted from the interface-range prompt.

Example of the Interface-Range Prompt for Duplicate Interfaces

```
DellEMC(conf)#interface range vlan 1 , vlan 1 , vlan 3 , vlan 3
DellEMC(conf-if-range-vl-1,vl-3)#
DellEMC(conf)#interface range tengigabitethernet 2/1 - 23 , tengigabitethernet 2/1 - 23
, tengigab 2/1 - 23
DellEMC(conf-if-range-te-2/1-23)#
```

Exclude a Smaller Port Range

The following is an example show how the smaller of two port ranges is omitted in the interface-range prompt.

Example of the Interface-Range Prompt for Multiple Port Ranges

```
DellEMC(conf)#interface range tengigabitethernet 2/1 - 23 , tengigabitethernet 2/1 - 10
DellEMC(conf-if-range-te-2/1-23)#
```

Overlap Port Ranges

The following is an example showing how the interface-range prompt extends a port range from the smallest start port number to the largest end port number when port ranges overlap. handles overlapping port ranges.

Example of the Interface-Range Prompt for Overlapping Port Ranges

```
DellEMC(conf)#interface range tengigabitethernet 2/1 - 11 , tengigabitethernet 2/1 - 23
DellEMC(conf-if-range-te-2/1-23)#
```

Commas

The following is an example of how to use commas to add different interface types to a range of interfaces.

Example of Adding Interface Ranges

```
DellEMC(config-if)# interface range tengigabitethernet 5/1 - 23, tengigabitethernet 1/1
- 2
DellEMC(config-if-range-te-5/1-23,te1/1-2)# no shutdown
DellEMC(config-if-range-te-5/1-23,te1/1-2)#
```

Add Ranges

The following example shows how to use commas to add VLAN and port-channel interfaces to the range.

Example of Adding VLAN and Port-Channel Interface Ranges

```
DellEMC(config-if-range-te-1/1-2)# interface range Vlan 2 - 100 , Port 1 - 25
DellEMC(config-if-range-te-1/1-2-vl-2-100-po-1-25)# no shutdown
```

Defining Interface Range Macros

You can define an interface-range macro to automatically select a range of interfaces for configuration.

Before you can use the `macro` keyword in the `interface-range macro` command string, define the macro.

To define an interface-range macro, use the following command.

- Defines the interface-range macro and saves it in the running configuration file.
  CONFIGURATION mode

Define the Interface Range

The following example shows how to define an interface-range macro named “test” to select Ten Gigabit Ethernet interfaces 5/1 through 5/4.

Example of the `define interface-range` Command for Macros

```
DellEMC(config)# define interface-range test tengigabitethernet 5/1 - 4
```

Choosing an Interface-Range Macro

To use an interface-range macro, use the following command.

- Selects the interfaces range to be configured using the values saved in a named interface-range macro.
  CONFIGURATION mode
  `interface range macro` *`name`*

The following example shows how to change to the interface-range configuration mode using the interface-range macro named "test."

```
DellEMC(config)# interface range macro test
DellEMC(config-if)#
```

Monitoring and Maintaining Interfaces

Monitor interface statistics with the `monitor interface` command. This command displays an ongoing list of the interface status (up/down), number of packets, traffic statistics, and so on.

To view the interface's statistics, use the following command.

- View the interface's statistics.

  EXEC Privilege mode

  Enter the type of interface and the interface information:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

The information displays in a continuous run, refreshing every 2 seconds by default. To manage the output, use the following keys.

- `m` — Change mode
- `l` — Page up
- `T` — Increase refresh interval (by 1 second)
- `t` — Decrease refresh interval (by 1 second)
- `c` — Clear screen
- `a` — Page down
- `q` — Quit

```
DellEMC#monitor interface Te 3/1

Dell uptime is 1 day(s), 4 hour(s), 31 minute(s)
  Monitor time: 00:00:00 Refresh Intvl.: 2s

Interface: Te 3/1, Disabled, Link is Down, Linespeed is 1000 Mbit

Traffic statistics: Current    Rate    Delta
       Input bytes:        0    0 Bps       0
      Output bytes:        0    0 Bps       0
     Input packets:        0    0 pps       0
    Output packets:        0    0 pps       0
       64B packets:        0    0 pps       0
  Over 64B packets:        0    0 pps       0
 Over 127B packets:        0    0 pps       0
 Over 255B packets:        0    0 pps       0
 Over 511B packets:        0    0 pps       0
Over 1023B packets:        0    0 pps       0
  Error statistics:
   Input underruns:        0    0 pps       0
      Input giants:        0    0 pps       0
   Input throttles:        0    0 pps       0
         Input CRC:        0    0 pps       0
 Input IP checksum:        0    0 pps       0
     Input overrun:        0    0 pps       0
  Output underruns:        0    0 pps       0
  Output throttles:        0    0 pps       0

m - Change mode                 c - Clear screen
l - Page up                     a - Page down
T - Increase refresh interval   t - Decrease refresh interval
q - Quit

q
DellEMC#
```

Maintenance Using TDR

The time domain reflectometer (TDR) is supported on all Dell EMC Networking switches.

TDR is an assistance tool to resolve link issues that helps detect obvious open or short conditions within any of the four copper pairs. TDR sends a signal onto the physical cable and examines the reflection of the signal that returns. By examining the reflection, TDR is able to indicate whether there is a cable fault (when the cable is broken, becomes unterminated, or if a transceiver is unplugged).

TDR is useful for troubleshooting an interface that is not establishing a link; that is, when the link is flapping or not coming up. TDR is not intended to be used on an interface that is passing traffic. When a TDR test is run on a physical cable, it is important to shut down the port on the far end of the cable. Otherwise, it may lead to incorrect test results.

NOTE: TDR is an intrusive test. Do not run TDR on a link that is up and passing traffic.

To test and display TDR results, use the following commands.

1. To test for cable faults on the TenGigabitEthernet cable.
   EXEC Privilege mode
   `tdr-cable-test tengigabitethernet slot/port`
   Between two ports, do not start the test on both ends of the cable.
   Enable the interface before starting the test.
   Enable the port to run the test or the test prints an error message.
2. Displays TDR test results.
   EXEC Privilege mode
   `show tdr tengigabitethernet slot/port`

Configuring wavelength for 10–Gigabit SFP+ optics

You can set the wavelength for tunable 10–Gigabit SFP+ optics using the `wavelength` command. To set the wavelength, follow these steps:

- Enter the interface mode and set the wavelength.
  INTERFACE mode
  `wavelength 1529.0`
  The wavelength range is from 1528.3 nm to 1568.77nm.
- Verify configuration changes.
  INTERFACE mode
  `show config`

Link Dampening

Interface state changes occur when interfaces are administratively brought up or down or if an interface state changes.

Every time an interface changes a state or flaps, routing protocols are notified of the status of the routes that are affected by the change in state. These protocols go through the momentous task of re-converging. Flapping; therefore, puts the status of entire network at risk of transient loops and black holes. Dampening limits the notification of status to the routing protocols. Link dampening minimizes the risk created by flapping by imposing a penalty (1024) for each interface flap and decaying the penalty exponentially based on the half-time. When the accumulated penalty exceeds a certain threshold (suppress threshold), the interface is put in an Error-Disabled state and for all practical purposes of routing, the interface is deemed to be “down.” After the interface becomes stable and the penalty decays below a certain threshold (reuse threshold), the interface comes up again and the routing protocols re-converge.

You configure link dampening using the `dampening [[[[half-life] [reuse-threshold]] [suppress-threshold]] [max-suppress-time]]` command on the interface.

Following is the detailed explanation of interface state change events:

- *suppress-threshold*— The suppress threshold is a value that triggers a flapping interface to dampen. The system adds penalty when the interface state goes up and down. When the accumulated penalty reaches the default or configured suppress threshold, the interface state changes to Error-Disabled state. The range of suppress threshold is from 1 to 20000. The default is **2500**.
- *half-life*— The accumulated penalty decays exponentially based on the half-life period. The accumulated penalty decreases half after each half-life period. The range of half-life is from 1 to 30 seconds. The default is **5 seconds**.
- *reuse-threshold*— After exponential decay, the penalty reaches the default or configured reuse threshold. The interface is unsuppressed and the state changes to "up". The range of reuse threshold is from 1 to 20000. The default is **750**.
- *max-suppress-time*— The maximum amount of time during which the interface remain suppressed. The range is from 1 to 86400. The default is **20 seconds** or four times the default half-life period (5 seconds).

NOTE:

`suppress-threshold` should be greater than `reuse-threshold`.

`max-suppress-time` should be at least 4 times `half-life`.

Link dampening:

- reduces processing on the CPUs by reducing excessive interface flapping.
- improves network stability by penalizing misbehaving interfaces and redirecting traffic.
- improves convergence times and stability throughout the network by isolating failures so that disturbances are not propagated.

Important Points to Remember

- Link dampening is not supported on VLAN interfaces.
- Link dampening is disabled when the interface is configured for port monitoring.
- You can apply link dampening to Layer 2 and Layer 3 interfaces.
- You can configure link dampening on individual interfaces in a LAG.

Configuration Example of Link Dampening

The figure shows a how link dampening works in a sample scenario when an interface is configured with dampening. The following figure shows the interface state change, accumulation and decay of penalty, and the interface advertised state based on the set dampening parameters. The dampening parameters are set as follows:

- *half-life*— 10 seconds
- *reuse-threshold*—300
- *suppress-threshold*—2400
- *max-suppress-time*—60 seconds

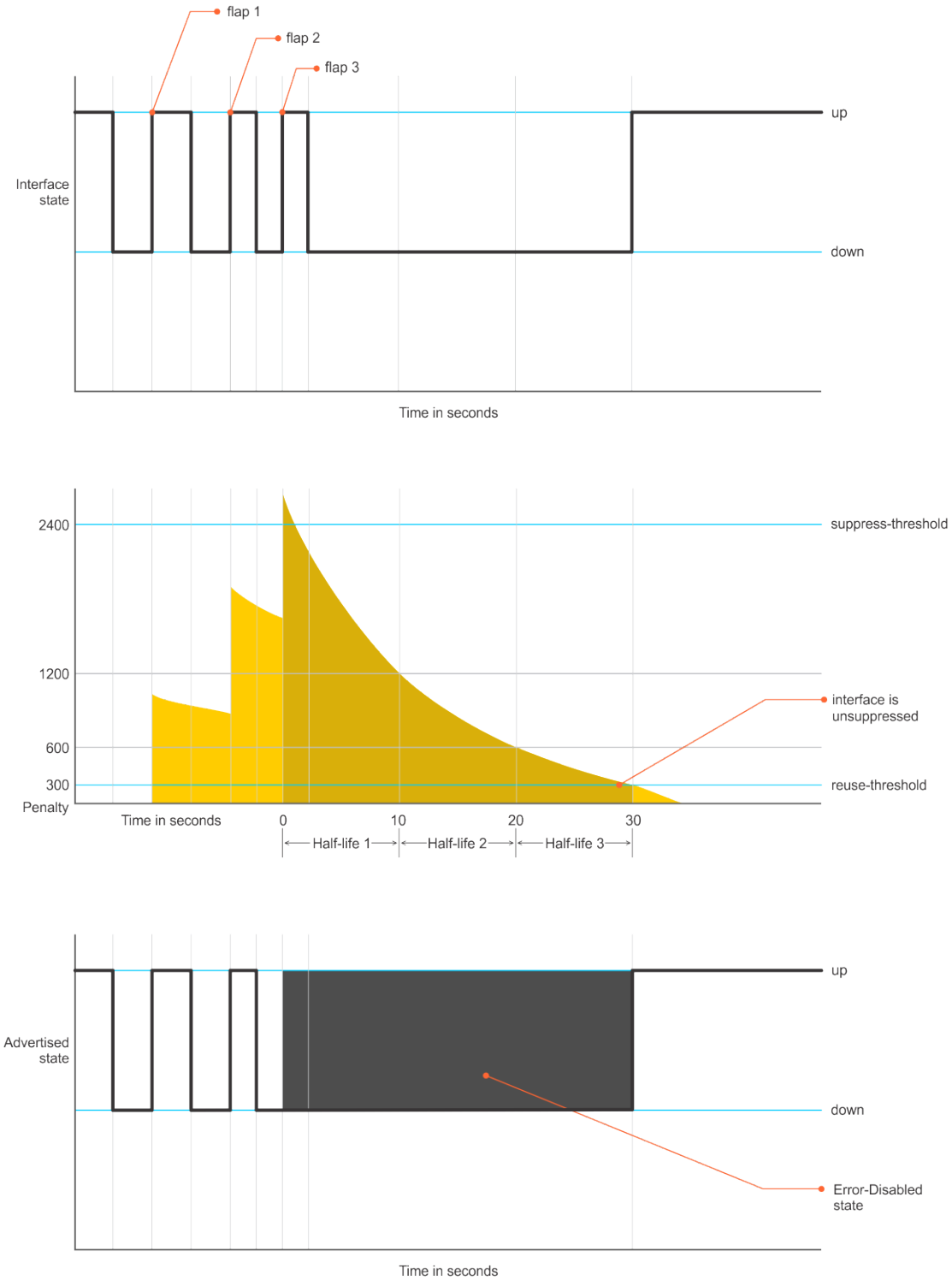


Figure 51. Interface State Change

Consider an interface periodically flaps as shown above. Every time the interface goes down, a penalty (1024) is added. In the above example, during the first interface flap (flap 1), the penalty is added to 1024. And, the accumulated penalty will exponentially decay based on the set half-life, which is set as 10 seconds in the above example. During the second interface flap (flap 2), again the penalty (1024) is accumulated. When the accumulated penalty exceeds the configured suppress threshold (2400), the interface state is set to Error-Disabled state. After the flap (flap 3), the interface flap stops. Then, the accumulated penalty decays exponentially and when it reaches below the set reuse threshold (300), the interface is unsuppressed and the interface state changes to “up” state.

Enabling Link Dampening

To enable link dampening, use the following command.

- Enable link dampening.
  INTERFACE mode
  `dampening`

To view the link dampening configuration on an interface, use the `show config` command.

```
R1(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  ip address 10.10.19.1/24
  dampening 1 2 3 4
  no shutdown
```

To view dampening information on all or specific dampened interfaces, use the `show interfaces dampening` command from EXEC Privilege mode.

```
DellEMC#show interfaces dampening
Interface   Supp    Flaps   Penalty Half-Life   Reuse   Suppress    Max-Sup
            State
Te 1/1      Up      0       0       1           2       3           4
Te 1/2      Up      0       0       1           2       3           4
Te 1/2      Up      0       0       1           2       3           4
DellEMC#
```

To view a dampening summary for the entire system, use the `show interfaces dampening summary` command from EXEC Privilege mode.

```
DellEMC# show interfaces dampening summary
20 interfaces are configured with dampening. 3 interfaces are currently suppressed.
Following interfaces are currently suppressed:
Te 1/2
Te 3/1
Te 4/2
DellEMC#
```

Clearing Dampening Counters

To clear dampening counters and accumulated penalties, use the following command.

- Clear dampening counters.
  `clear dampening`

```
DellEMC# clear dampening interface Te 1/1

DellEMC#show interfaces dampening Tengigabitethernet 1/1
Interface   Supp    Flaps   Penalty Half-Life   Reuse   Suppress    Max-Sup
            State
Te 1/1      Up      0       0       1           2       3           4
DellEMC#
```

Link Dampening Support for XML

View the output of the following `show` commands in XML by adding `| display xml` to the end of the command.

- `show interfaces dampening`
- `show interfaces dampening summary`
- `show interfaces interface slot/port`

Configure MTU Size on an Interface

In Dell EMC Networking OS, Maximum Transmission Unit (MTU) is defined as the entire Ethernet packet (Ethernet header + FCS + payload).

The following table lists the range for each transmission media.

Transmission Media	MTU Range (in bytes)
Ethernet	594-12000 = link MTU 576-9234 = IP MTU

Link Bundle Monitoring

Monitoring linked LAG bundles allows traffic distribution amounts in a link to be monitored for unfair distribution at any given time. A threshold of 60% is defined as an acceptable amount of traffic on a member link. Links are monitored in 15-second intervals for three consecutive instances. Any deviation within that time sends Syslog and an alarm event generates. When the deviation clears, another Syslog sends and a clear alarm event generates.

The link bundle utilization is calculated as the total bandwidth of all links divided by the total bytes-per-second of all links. If you enable monitoring, the utilization calculation is performed when the utilization of the link-bundle (not a link within a bundle) exceeds 60%.

To enable and view link bundle monitoring, use the following commands.

- Enable link bundle monitoring.

  `ecmp-group`

- View all LAG link bundles being monitored.

  `show running-config ecmp-group`

- Enable link bundle monitoring on port channel interfaces.

  `link-bundle-monitor enable`

  ```
  DellEMC(conf-if-po-10)#link-bundle-monitor enable
  ```

- Configure threshold level for link bundle monitoring.

  `link-bundle-distribution trigger-threshold`

  ```
  DellEMC(conf)#link-bundle-distribution trigger-threshold
  ```

- View the link bundle monitoring status.

  `show link-bundle-distribution`

Using Ethernet Pause Frames for Flow Control

Ethernet pause frames and threshold settings are supported on the Dell EMC Networking OS.

Ethernet Pause Frames allow for a temporary stop in data transmission. A situation may arise where a sending device may transmit data faster than a destination device can accept it. The destination sends a PAUSE frame back to the source, stopping the sender's transmission for a period of time.

An Ethernet interface starts to send pause frames to a sending device when the transmission rate of ingress traffic exceeds the egress port speed. The interface stops sending pause frames when the ingress rate falls to less than or equal to egress port speed.

The globally assigned 48-bit Multicast address 01-80-C2-00-00-01 is used to send and receive pause frames. To allow full-duplex flow control, stations implementing the pause operation instruct the MAC to enable reception of frames with destination address equal to this multicast address.

The PAUSE frame is defined by IEEE 802.3x and uses MAC Control frames to carry the PAUSE commands. Ethernet pause frames are supported on full duplex only.

If a port is over-subscribed, Ethernet Pause Frame flow control does not ensure no-loss behavior.

Restriction: Ethernet Pause Frame flow control is not supported if PFC is enabled on an interface.

Control how the system responds to and generates 802.3x pause frames on Ethernet interfaces. The default is rx off tx off. INTERFACE mode. `flowcontrol rx [off | on] tx [off | on]| [monitor session-ID]`

Where:

`rx on`: Processes the received flow control frames on this port.

`rx off`: Ignores the received flow control frames on this port.

`tx on`: Sends control frames from this port to the connected device when a higher rate of traffic is received.

`tx off`: Flow control frames are not sent from this port to the connected device when a higher rate of traffic is received.

`monitor session-ID`: Enables mirror flow control frames on this port.

Changes in the flow-control values may not be reflected automatically in the `show interface` output. To display the change, apply the new flow-control setting, shutdown the interface using the `shutdown` command, enable the interface using the `no shutdown` command, and use the `show interface` command to verify the changes.

Enabling Pause Frames

Enable Ethernet pause frames flow control on all ports on a chassis or a line card. If not, the system may exhibit unpredictable behavior.

NOTE: Changes in the flow-control values may not be reflected automatically in the `show interface` output. As a workaround, apply the new settings, execute `shut` then `no shut` on the interface, and then check the running-config of the port.

NOTE: If you disable `rx flow control`, Dell EMC Networking recommends rebooting the system.

The flow control sender and receiver must be on the same port-pipe. Flow control is not supported across different port-pipes.

To enable pause frames, use the following command.

- Control how the system responds to and generates 802.3x pause frames on the Ethernet ports.
  INTERFACE mode
  `flowcontrol {rx [off | on] tx [off | on] [negotiate] }`
  - `rx on`: enter the keywords `rx on` to process the received flow control frames on this port.
  - `rx off`: enter the keywords `rx off` to ignore the received flow control frames on this port.
  - `tx on`: enter the keywords `tx on` to send control frames from this port to the connected device when a higher rate of traffic is received.
  - `tx off`: enter the keywords `tx off` so that flow control frames are not sent from this port to the connected device when a higher rate of traffic is received.
  - `negotiate`: enable pause-negotiation with the egress port of the peer device. If the `negotiate` command is not used, pause-negotiation is disabled. 40 gigabit Ethernet interfaces do not support pause-negotiation.

Configure the MTU Size on an Interface

If a packet includes a Layer 2 header, the difference in bytes between the link MTU and IP MTU must be enough to include the Layer 2 header.

For example, for VLAN packets, if the IP MTU is 1400, the Link MTU must be no less than 1422:

`1400-byte IP MTU + 22-byte VLAN Tag = 1422-byte link MTU`

The following table lists the various Layer 2 overheads found in the Dell EMC Networking OS and the number of bytes.

Table 43. Layer 2 Overhead

Layer 2 Overhead	Difference Between Link MTU and IP MTU
Ethernet (untagged)	18 bytes
VLAN Tag	22 bytes
Untagged Packet with VLAN-Stack Header	22 bytes
Tagged Packet with VLAN-Stack Header	26 bytes

Link MTU and IP MTU considerations for port channels and VLANs are as follows.

Port Channels:

- All members must have the same link MTU value and the same IP MTU value.
- The port channel link MTU and IP MTU must be less than or equal to the link MTU and IP MTU values configured on the channel members.

For example, if the members have a link MTU of 2100 and an IP MTU 2000, the port channel's MTU values cannot be higher than 2100 for link MTU or 2000 bytes for IP MTU.

VLANs:

- All members of a VLAN must have the same IP MTU value.
- Members can have different Link MTU values. Tagged members must have a link MTU 4–bytes higher than untagged members to account for the packet tag.
- The VLAN link MTU and IP MTU must be less than or equal to the link MTU and IP MTU values configured on the VLAN members.

For example, the VLAN contains tagged members with Link MTU of 1522 and IP MTU of 1500 and untagged members with Link MTU of 1518 and IP MTU of 1500. The VLAN's Link MTU cannot be higher than 1518 bytes and its IP MTU cannot be higher than 1500 bytes.

NOTE: When configuring IP MTU on an interface, make sure to set the same or maximum IP MTU value on both ingress and egress interfaces to make IPv4 traffic flow.

Port-Pipes

A port pipe is a Dell EMC Networking-specific term for the hardware packet-processing elements that handle network traffic to and from a set of front-end I/O ports. The physical, front-end I/O ports are referred to as a `port-set`. In the command-line interface, a port pipe is entered as `port-set` *`port-pipe-number`*.

Auto-Negotiation on Ethernet Interfaces

By default, auto-negotiation of speed and full duplex mode is enabled on 10/100/1000 Base-T Ethernet interfaces. Only 10GE interfaces do not support auto-negotiation.

When using 10GE interfaces, verify that the settings on the connecting devices are set to no auto-negotiation.

The local interface and the directly connected remote interface must have the same setting, and auto-negotiation is the easiest way to accomplish that, as long as the remote interface is capable of auto-negotiation.

NOTE: As a best practice, Dell EMC Networking recommends keeping auto-negotiation enabled. Only disable auto-negotiation on switch ports that attach to devices not capable of supporting negotiation or where connectivity issues arise from interoperability issues.

Setting the Speed of Ethernet Interfaces

To discover whether the remote and local interface requires manual speed synchronization, and to manually synchronize them if necessary, use the following command sequence.

1. Determine the local interface status. Refer to the following example.

   EXEC Privilege mode

```
show interfaces [interface | stack-unit stack-unit-number] status
```

2. Determine the remote interface status.
   EXEC mode or EXEC Privilege mode
   [Use the command on the remote system that is equivalent to the first command.]
3. Access CONFIGURATION mode.
   EXEC Privilege mode
   ```
   config
   ```
4. Access the port.
   CONFIGURATION mode
   ```
   interface interface-type
   ```
5. Set the local port speed.
   INTERFACE mode
   ```
   speed {10 | 100 | 1000 | 10000 | auto}
   ```
   NOTE: If you use an active optical cable (AOC), you can convert the QSFP+ port to a 10 Gigabit SFP+ port or 1 Gigabit SFP port. You can use the `speed` command to enable the required speed.
6. Optionally, set full- or half-duplex.
   INTERFACE mode
   ```
   duplex {half | full}
   ```
7. Disable auto-negotiation on the port.
   INTERFACE mode
   ```
   no negotiation auto
   ```
   If the speed was set to 1000, do not disable auto-negotiation.
8. Verify configuration changes.
   INTERFACE mode
   ```
   show config
   ```

NOTE: The `show interfaces status` command displays link status, but not administrative status. For both link and administrative status, use the `show ip interface` command.

```
DellEMC#show interfaces status
Port Description Status Speed      Duplex Vlan
Te 1/1           Down   1000 Mbit  Auto   1
Te 1/2           Down   Auto       Auto   1
Te 1/3           Down   Auto       Auto   --
Te 1/4 Force10Port Up   1000 Mbit  Auto   30-130
Te 1/5           Down   Auto       Auto   --
Te 1/6           Down   Auto       Auto   --
Te 1/7           Up     1000 Mbit  Auto   1502,1504,1506-1508,1602
Te 1/8           Down   Auto       Auto   --
Te 1/9           Down   Auto       Auto   --
Te 1/10          Down   Auto       Auto   --
Te 1/11          Down   Auto       Auto   --
Te 1/12          Down   Auto       Auto   --
[output omitted]
```

In the previous example, several ports display “Auto” in the Speed field. In the following example, the speed of port 1/1 is set to 100Mb and then its auto-negotiation is disabled.

```
DellEMC#configure
DellEMC(config)#interface tengig 1/1
DellEMC(conf-if-te-1/1)#speed 100
DellEMC(conf-if-te-1/1)#duplex full
DellEMC(conf-if-te-1/1)#no negotiation auto
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
no ip address
speed 100
```

```
duplex full
no shutdown
```

Set Auto-Negotiation Options

The `negotiation auto` command provides a mode option for configuring an individual port to forced master/ forced slave once auto-negotiation is enabled.

CAUTION: Ensure that only one end of the node is configured as forced-master and the other is configured as forced-slave. If both are configured the same (that is, both as forced-master or both as forced-slave), the `show interface` command flaps between an auto-neg-error and forced-master/slave states.

Example of the `negotiation auto` Command

```
DellEMC(conf)# int tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#neg auto
DellEMC(conf-if-te-1/1-autoneg)# ?

end              Exit from configuration mode
exit             Exit from autoneg configuration mode
mode      Specify autoneg mode
no             Negate a command or set its defaults
show             Show autoneg configuration information
DellEMC(conf-if-te-1/1-autoneg)#mode ?
forced-master Force port to master mode
forced-slave     Force port to slave mode
DellEMC(conf-if-te-1/1-autoneg)#
```

For details about the `speed`, `duplex`, and `negotiation auto` commands, refer to the *Interfaces* chapter of the *Dell EMC Networking OS Command Reference Guide*.

NOTE: While using 10GBASE-T, auto-negotiation is enabled on the external PHY by default, and auto-negotiation should be enabled on the peer for the link to come up.

Adjusting the Keepalive Timer

To change the time interval between keepalive messages on the interfaces, use the `keepalive` command. The interface sends keepalive messages to itself to test network connectivity on the interface.

To change the default time interval between keepalive messages, use the following command.

- Change the default interval between keepalive messages.

  INTERFACE mode

  `keepalive [seconds]`
- View the new setting.

  INTERFACE mode

  `show config`

View Advanced Interface Information

The following options have been implemented for the `show [ip | running-config] interfaces` commands for (only) stack-unit interfaces.

When you use the `configured` keyword, only interfaces that have non-default configurations are displayed. Dummy stack-unit interfaces (created with the `stack-unit` command) are treated like any other physical interface.

The following example lists the possible `show` commands that have the configured keyword available:

```
DellEMC#show interfaces configured
DellEMC#show interfaces stack-unit 0 configured
DellEMC#show interfaces tengigabitEthernet 0 configured
DellEMC#show ip interface configured
```

```
DellEMC#show ip interface stack-unit 1 configured
DellEMC#show ip interface tengigabitEthernet 1 configured
DellEMC#show ip interface br configured
DellEMC#show ip interface br stack-unit 1 configured
DellEMC#show ip interface br tengigabitEthernet 1 configured
DellEMC#show running-config interfaces configured
DellEMC#show running-config interface tengigabitEthernet 1 configured
```

In EXEC mode, the `show interfaces switchport` command displays only interfaces in Layer 2 mode and their relevant configuration information. The `show interfaces switchport` command displays the interface, whether it supports IEEE 802.1Q tagging or not, and the VLANs to which the interface belongs.

Configuring the Interface Sampling Size

Although you can enter any value between 30 and 299 seconds (the default), software polling is done once every 15 seconds. So, for example, if you enter “19”, you actually get a sample of the past 15 seconds.

All LAG members inherit the rate interval configuration from the LAG.

The following example shows how to configure rate interval when changing the default value.

To configure the number of seconds of traffic statistics to display in the show interfaces output, use the following command.

- Configure the number of seconds of traffic statistics to display in the show interfaces output.

  INTERFACE mode

  `rate-interval`

The bold lines shows the default value of 299 seconds, the change-rate interval of 100, and the new rate interval set to 100.

```
DellEMC#show interfaces
TenGigabitEthernet 1/1 is down, line protocol is down
Hardware is DellEMCEth, address is 00:01:e8:01:9e:d9
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 1d23h44m
Queueing strategy: fifo
  0 packets input, 0 bytes
  Input 0 IP Packets, 0 Vlans 0 MPLS
  0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
  0 CRC, 0 IP Checksum, 0 overrun, 0 discarded
  0 packets output, 0 bytes, 0 underruns
  Output 0 Multicasts, 0 Broadcasts, 0 Unicasts
  0 IP Packets, 0 Vlans, 0 MPLS
  0 throttles, 0 discarded
Rate info (interval 299 seconds):
  Input 00.00 Mbits/sec,  0 packets/sec, 0.00% of line-rate
  Output 00.00 Mbits/sec, 0 packets/sec, 0.00% of line-rate
Time since last interface status change: 1d23h40m

DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#rate-interval 100

DellEMC#show interfaces
TenGigabitEthernet 1/1 is down, line protocol is down
Hardware is DellEMCEth, address is 00:01:e8:01:9e:d9
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 1d23h45m
Queueing strategy: fifo
  0 packets input, 0 bytes
  Input 0 IP Packets, 0 Vlans 0 MPLS
  0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
```

```
  0 CRC, 0 IP Checksum, 0 overrun, 0 discarded
  0 packets output, 0 bytes, 0 underruns
  Output 0 Multicasts, 0 Broadcasts, 0 Unicasts
  0 IP Packets, 0 Vlans, 0 MPLS
  0 throttles, 0 discarded
Rate info (interval 100 seconds):
  Input 00.00 Mbits/sec,  0 packets/sec, 0.00% of line-rate
  Output 00.00 Mbits/sec, 0 packets/sec, 0.00% of line-rate
Time since last interface status change: 1d23h42m
```

Configuring the Traffic Sampling Size Globally

You can configure the traffic sampling size for an interface in the global configuration mode.

All LAG members inherit the rate interval configuration from the LAG.

Although you can enter any value between 30 and 299 seconds (the default), software polling is done once every 15 seconds. So, for example, if you enter “19”, you actually get a sample of the past 15 seconds.

The following example shows how to configure rate interval when changing the default value.

To configure the number of seconds of traffic statistics to display in the show interfaces output, use the following command.

- Configure the number of seconds of traffic statistics to display in the show interfaces output.
  CONFIGURATION Mode
  `rate-interval`

The bold lines shows the default value of 299 seconds, the change-rate interval of 100, and the new rate interval set to 100.

```
DellEMC#configure terminal
DellEMC(Conf)#rate-interval 150
```

```
DellEMC#show interface TenGigabitEthernet 10/0
TenGigabitEthernet 10/0 is up, line protocol is up
Description: interface tengig 10/0
Hardware is DellEMCEth, address is 34:17:eb:01:20:f3
    Current address is 34:17:eb:01:20:f3
Pluggable media present, SFP+ type is 10GBASE-SR
    Medium is MultiRate, Wavelength is 850nm
    SFP+ receive power reading is -36.9897dBm
Interface index is 11534340
Internet address is not set
Mode of IPv4 Address Assignment : NONE
DHCP Client-ID :3417eb0120f3
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit
Flowcontrol rx off tx off
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 2w6d21h
Queueing strategy: fifo
Input Statistics:
     3106 packets, 226755 bytes
     133 64-byte pkts, 2973 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     406 Multicasts, 0 Broadcasts, 2700 Unicasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     3106 packets, 226755 bytes, 0 underruns
     133 64-byte pkts, 2973 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     406 Multicasts, 0 Broadcasts, 2700 Unicasts
     0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 150 seconds):
     Input 300.00 Mbits/sec,          1534517 packets/sec, 30.00% of line-rate
     Output 100.00 Mbits/sec,          4636111 packets/sec, 10.00% of line-rate
Time since last interface status change: 01:07:44
```

```
DellEMC#show int po 20
Port-channel 20 is up, line protocol is up
Hardware address is 4c:76:25:f4:ab:02, Current address is 4c:76:25:f4:ab:02
Interface index is 1258301440
Minimum number of links to bring Port-channel up is 1
Internet address is not set
Mode of IPv4 Address Assignment : NONE
DHCP Client-ID :4c7625f4ab02
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 80000 Mbit
Members in this channel:  Fo 1/1/7/1(U) Fo 1/1/8/1(U)
ARP type: ARPA, ARP Timeout 04:00:00
Queueing strategy: fifo
Input Statistics:
     13932 packets, 1111970 bytes
     5588 64-byte pkts, 8254 over 64-byte pkts, 89 over 127-byte pkts
     1 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     13761 Multicasts, 9 Broadcasts, 162 Unicasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     13908 packets, 1114396 bytes, 0 underruns
     5555 64-byte pkts, 8213 over 64-byte pkts, 140 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     13727 Multicasts, 5 Broadcasts, 176 Unicasts
     0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 150 seconds):
     Input 300.00 Mbits/sec,          1534517 packets/sec, 30.00% of line-rate
     Output 100.00 Mbits/sec,          4636111 packets/sec, 10.00% of line-rate
Time since last interface status change: 21:00:43
```

Dynamic Counters

By default, counting is enabled for IPFLOW, IPACL, L2ACL, L2FIB.

For the remaining applications, Dell EMC Networking OS automatically turns on counting when you enable the application, and is turned off when you disable the application.

NOTE: If you enable more than four counter-dependent applications on a port pipe, there is an impact on line rate performance.

The following counter-dependent applications are supported by Dell EMC Networking OS:

- Egress VLAN
- Ingress VLAN
- Next Hop 2
- Next Hop 1
- Egress ACLs
- ILM
- IP FLOW
- IP ACL
- IP FIB
- L2 ACL
- L2 FIB

Clearing Interface Counters

The counters in the `show interfaces` command are reset by the `clear counters` command. This command does not clear the counters any SNMP program captures.

To clear the counters, use the following the command.

- Clear the counters used in the `show interface` commands for all VRRP groups, VLANs, and physical interfaces or selected ones. Without an interface specified, the command clears all interface counters.

EXEC Privilege mode

`clear counters [interface] [vrrp [vrid] | learning-limit]`

(OPTIONAL) Enter the following interface keywords and slot/port or number information:

- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
- For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
- For the Management interface on the stack-unit, enter the keyword `ManagementEthernet` then the slot/port information.
- For a port channel interface, enter the keywords `port-channel` then a number.
- For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.
- (OPTIONAL) To clear statistics for all VRRP groups configured, enter the keyword `vrrp`. Enter a number from 1 to 255 as the *vrid*.
- (OPTIONAL) To clear unknown source address (SA) drop counters when you configure the MAC learning limit on the interface, enter the keywords `learning-limit`.

When you enter this command, confirm that you want Dell EMC Networking OS to clear the interface counters for that interface.

```
DellEMC#clear counters te 1/1
Clear counters on TenGigabitEthernet 1/1 [confirm]
DellEMC#
```

Compressing Configuration Files

You can optimize and reduce the sizes of the configuration files.

You can compress the running configuration by grouping all the VLANs and the physical interfaces with the same property. Support to store the operating configuration to the startup config in the compressed mode and to perform an image downgrade without any configuration loss are provided.

You can create groups of VLANs using the `interface group` command. This command will create nonexistent VLANs specified in a range. On successful command execution, the CLI switches to the interface group context. The configuration commands inside the group context will be the similar to that of the existing range command.

Two existing exec mode CLIs are enhanced to display and store the running configuration in the compressed mode.

`show running-config compressed and write memory compressed`

The compressed configuration will group all the similar looking configuration thereby reducing the size of the configuration. For this release, the compression will be done only for interface related configuration (VLAN & physical interfaces)

The following table describes how the standard and the compressed configuration differ:

Table 44. Standard and Compressed Configurations

Uncompressed	Compressed
DellEMC# show running-config <snip> ! interface TenGigabitEthernet 1/1 no ip address switchport shutdown ! interface TenGigabitEthernet 1/2 no ip address shutdown	DellEMC# show running-config compressed <snip> ! interface TenGigabitEthernet 1/1 no ip address switchport shutdown ! Interface group TenGigabitEthernet 1/2 – 4 , TenGigabitEthernet 1/10 no ip address shutdown

Table 44. Standard and Compressed Configurations

Uncompressed	Compressed
!	!
interface TenGigabitEthernet 1/3	interface TenGigabitEthernet 1/34
no ip address	ip address 2.1.1.1/16
shutdown	shutdown
!	!
interface TenGigabitEthernet 1/4	interface group Vlan 2 , Vlan 100
no ip address	no ip address
shutdown	no shutdown
!	!
interface TenGigabitEthernet 1/10	interface group Vlan 3 – 5
no ip address	tagged te 1/1
shutdown	no ip address
!	shutdown
interface TenGigabitEthernet 1/34	!
ip address 2.1.1.1/16	interface Vlan 1000
shutdown	ip address 1.1.1.1/16
!	no shutdown
interface Vlan 2	!
no ip address	<snip>
no shutdown	Compressed config size – 27 lines.
!	
interface Vlan 3	
tagged te 1/1	
no ip address	
shutdown	
!	
interface Vlan 4	
tagged te 1/1	
no ip address	
shutdown	
!	
interface Vlan 5	
tagged te 1/1	
no ip address	
shutdown	
!	
interface Vlan 100	
no ip address	
no shutdown	
!	
interface Vlan 1000	

Table 44. Standard and Compressed Configurations (continued)

Uncompressed	Compressed
ip address 1.1.1.1/16 no shutdown Uncompressed config size – 52 lines	

`write memory compressed`

The write memory compressed CLI will write the operating configuration to the startup-config file in the compressed mode. In stacking scenario, it will also take care of syncing it to all the standby and member units.

The following is the sample output:

```
DellEMC#write memory compressed
!
Jul 30 08:50:26: %STKUNIT0-M:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-
config in flash by default
```

copy compressed-config

Copy one file, after optimizing and reducing the size of the configuration file, to another location. Dell EMC Networking OS supports IPv4 and IPv6 addressing for FTP, TFTP, and SCP (in the hostip field).

22

Internet Protocol Security (IPSec)

Internet protocol security (IPSec) is an end-to-end security scheme for protecting IP communications by authenticating and encrypting all packets in a communication session. Use IPSec between hosts, between gateways, or between hosts and gateways.

IPSec is compatible with Telnet and FTP protocols. It supports two operational modes: Transport and Tunnel.

- Transport mode — (default) Use to encrypt only the payload of the packet. Routing information is unchanged.
- Tunnel mode — Use to encrypt the entire packet including the routing information of the IP header. Typically used when creating virtual private networks (VPNs).

NOTE: The Dell EMC Networking OS supports IPSec only for FTP and telnet protocols (ports 20, 21, and 23). The system rejects if you configure IPSec for other protocols.

IPSec uses the following protocols:

- **Authentication Headers (AH)** — Disconnected integrity and origin authentication for IP packets
- **Encapsulating Security Payload (ESP)** — Confidentiality, authentication, and data integrity for IP packets
- **Security Associations (SA)** — Necessary algorithmic parameters for AH and ESP functionality

IPSec supports the following authentication and encryption algorithms:

- Authentication only:
  - MD5
  - SHA1
- Encryption only:
  - 3DES
  - CBC
  - DES
- ESP Authentication and Encryption:
  - MD5 & 3DES
  - MD5 & CBC
  - MD5 & DES
  - SHA1 & 3DES
  - SHA1 & CBC
  - SHA1 & DES

Topics:

- Configuring IPSec

Configuring IPSec

The following sample configuration shows how to configure FTP and telnet for IPSec.

1. Define the transform set.
   CONFIGURATION mode
   ```
   crypto ipsec transform-set myXform-seta esp-authentication md5 esp-encryption des
   ```
2. Define the crypto policy.
   CONFIGURATION mode
   ```
   crypto ipsec policy myCryptoPolicy 10 ipsec-manual
   transform-set myXform-set
   session-key inbound esp 256 auth <key>
   encrypt <key>
   session-key outbound esp 257 auth <key> encrypt <key>
   ```

```
match 0 tcp a::1 /128 0 a::2 /128 23
match 1 tcp a::1 /128 23 a::2 /128 0
match 2 tcp a::1 /128 0 a::2 /128 21
match 3 tcp a::1 /128 21 a::2 /128 0
match 4 tcp 1.1.1.1 /32 0 1.1.1.2 /32 23
match 5 tcp 1.1.1.1 /32 23 1.1.1.2 /32 0
match 6 tcp 1.1.1.1 /32 0 1.1.1.2 /32 21
match 7 tcp 1.1.1.1 /32 21 1.1.1.2 /32 0
```

3. Apply the crypto policy to management traffic.
   CONFIGURATION mode

```
management crypto-policy myCryptoPolicy
```

23

IPv4 Routing

The Dell EMC Networking Operating System (OS) supports various IP addressing features. This chapter describes the basics of domain name service (DNS), address resolution protocol (ARP), and routing principles and their implementation in the Dell EMC Networking OS.

IP Feature	Default
DNS	Disabled
Directed Broadcast	Disabled
Proxy ARP	Enabled
ICMP Unreachable	Disabled
ICMP Redirect	Disabled

Topics:

- IP Addresses
- Configuration Tasks for IP Addresses
- Assigning IP Addresses to an Interface
- Configuring Static Routes
- Configure Static Routes for the Management Interface
- IPv4 Path MTU Discovery Overview
- Using the Configured Source IP Address in ICMP Messages
- Configuring the Duration to Establish a TCP Connection
- Enabling Directed Broadcast
- Resolution of Host Names
- Enabling Dynamic Resolution of Host Names
- Specifying the Local System Domain and a List of Domains
- Configuring DNS with Traceroute
- ARP
- Configuration Tasks for ARP
- Configuring Static ARP Entries
- Enabling Proxy ARP
- Clearing ARP Cache
- ARP Learning via Gratuitous ARP
- Enabling ARP Learning via Gratuitous ARP
- ARP Learning via ARP Request
- Configuring ARP Retries
- ICMP
- Configuration Tasks for ICMP
- Enabling ICMP Unreachable Messages
- ICMP Redirects
- UDP Helper
- Enabling UDP Helper
- Configuring a Broadcast Address
- Configurations Using UDP Helper
- UDP Helper with Broadcast-All Addresses
- UDP Helper with Subnet Broadcast Addresses
- UDP Helper with Configured Broadcast Addresses
- UDP Helper with No Configured Broadcast Addresses

- Troubleshooting UDP Helper

IP Addresses

Dell EMC Networking OS supports IP version 4 (as described in RFC 791), classful routing, and variable length subnet masks (VLSM).

With VLSM, you can configure one network with different masks. Supernetting, which increases the number of subnets, is also supported. To subnet, you add a mask to the IP address to separate the network and host portions of the IP address.

At its most basic level, an IP address is 32-bits composed of network and host portions and represented in dotted decimal format. For example, 00001010110101100101011110000011 is represented as 10.214.87.131.

For more information about IP addressing, refer to RFC 791, Internet Protocol.

Implementation Information

You can configure any IP address as a static route except IP addresses already assigned to interfaces.

NOTE: Dell EMC Networking OS supports 31-bit subnet masks (/31, or 255.255.255.254) as defined by RFC 3021. This feature allows you to save two more IP addresses on point-to-point links than 30-bit masks. Dell EMC Networking OS supports RFC 3021 with ARP.

NOTE: Even though Dell EMC Networking OS listens to all ports, you can only use the ports starting from 35001 for IPv4 traffic. Ports starting from 0 to 35000 are reserved for internal use and you cannot use them for IPv4 traffic.

Configuration Tasks for IP Addresses

The following describes the tasks associated with IP address configuration.

Configuration tasks for IP addresses includes:

- Assigning IP Addresses to an Interface (mandatory)
- Configuring Static Routes (optional)
- Configure Static Routes for the Management Interface (optional)

For a complete listing of all commands related to IP addressing, refer to the *Dell EMC Networking OS Command Line Interface Reference Guide*.

Assigning IP Addresses to an Interface

Assign primary and secondary IP addresses to physical or logical (for example, [virtual local area network [VLAN] or port channel) interfaces to enable IP communication between the system and hosts connected to that interface.

You can assign one primary address and up to 255 secondary IP addresses to each interface.

1. Enter the keyword `interface` then the type of interface and slot/port information.

   CONFIGURATION mode

   `interface` *`slot/port`*

   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
   - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
   - For the Management interface on the stack-unit, enter the keyword `ManagementEthernet` then the slot/port information.
   - For a port channel interface, enter the keywords `port-channel` then a number.
   - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

2. Enable the interface.

   INTERFACE mode

```
no shutdown
```

3. Configure a primary IP address and mask on the interface.

   INTERFACE mode

   `ip address ip-address mask [secondary]`

   - *ip-address mask*: the IP address must be in dotted decimal format (A.B.C.D). The mask must be in slash prefix-length format (/24).
   - `secondary`: add the keyword `secondary` if the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

To view the configuration, use the `show config` command in INTERFACE mode or use the `show ip interface` command in EXEC privilege mode, as shown in the second example.

```
DellEMC(conf-if)#show conf
!
interface TenGigabitEthernet 1/1
  ip address 10.11.1.1/24
  no shutdown
!
```

Configuring Static Routes

A static route is an IP address that you manually configure and that the routing protocol does not learn, such as open shortest path first (OSPF). Often, static routes are used as backup routes in case other dynamically learned routes are unreachable.

You can enter as many static IP addresses as necessary.

To configure a static route, use the following command.

- Configure a static IP address.

  CONFIGURATION mode

  `ip route [vrf vrf-name] ip-address mask {ip-address | interface [ip-address]} [distance] [name description] [permanent] [tag tag-value] [vrf vrf-name]`

  Use the following required and optional parameters:

  - `vrf vrf-name` : use the VRF option after the `ip route` keyword to configure a static route on that particular VRF, use the VRF option after the next hop to specify which VRF the next hop belongs to. This will be used in route leaking cases.

    NOTE: For more information on route leaking, see the Route Leaking Between VRFs section.

  - *ip-address*: enter an address in dotted decimal format (A.B.C.D).
  - *mask*: enter a mask in slash prefix-length format (/X).
  - *interface*: enter an interface type then the slot/port information.
  - *distance*: the range is from 1 to 255. (optional)
  - `name`: enter the keyword `name` and the description of the IPv4 static route configuration. (optional)
  - `permanent`: keep the static route in the routing table (if you use the `interface` option) even if you disable the interface with the route. (optional)
  - `tag tag-value`: the range is from 1 to 4294967295. (optional)

To view the configured routes, use the `show ip route static` command.

```
DellEMC#show ip route static
  Destination  Gateway                Dist/Metric Last Change
-----------    -------                ----------- -----------
S 2.1.2.0/24   Direct, Nu 0              0/0       00:02:30
S 6.1.2.0/24   via 6.1.20.2, Te 5/1       1/0       00:02:30
S 6.1.2.2/32   via 6.1.20.2, Te 5/1       1/0       00:02:30
S 6.1.2.3/32   via 6.1.20.2, Te 5/1       1/0       00:02:30
S 6.1.2.4/32   via 6.1.20.2, Te 5/1       1/0       00:02:30
S 6.1.2.5/32   via 6.1.20.2, Te 5/1       1/0       00:02:30
```

```
S 6.1.2.6/32   via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.7/32   via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.8/32   via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.9/32   via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.10/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.11/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.12/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.13/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.14/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.15/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.16/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 6.1.2.17/32  via 6.1.20.2, Te 5/1      1/0      00:02:30
S 11.1.1.0/24  Direct, Nu 0              0/0      00:02:30
Direct, Lo 0
--More--
```

Dell EMC Networking OS installs a next hop that is on the directly connected subnet of current IP address on the interface.

Dell EMC Networking OS also installs a next hop that is not on the directly connected subnet but which recursively resolves to a next hop on the interface's configured subnet.

- When the interface goes down, Dell EMC Networking OS withdraws the route.
- When the interface comes up, Dell EMC Networking OS re-installs the route.
- When the recursive resolution is “broken,” Dell EMC Networking OS withdraws the route.
- When the recursive resolution is satisfied, Dell EMC Networking OS re-installs the route.

Adding description for IPv4 and IPv6 static routes

Dell EMC Networking OS provides support to add a description for the IPv4 or IPv6 static route configurations. A `name` option is introduced to provide the description about the static route configured. This feature enables you to segregate or distinguish between the configured IPv4 or IPv6 static routes.

Following is the syntax of the `{ip | ipv6} route` command with the name option:

`{ip | ipv6} route [vrf` *`vrf-name`*`] ip-address mask {ip-address | interface [ip-address]} [distance] [name` *`description`*`] [permanent] [tag` *`tag-value`*`] [vrf` *`vrf-name`*`] [weight` *`weight-value`*`]`

You can enter up to 32 characters as the name or description for this route. These description can be a combination of numbers, special characters, and alphabets. To add multiple strings using space, use double quotes.

Following is the sample configuration steps to provide the descriptions for the IPv4 and IPv6 static route configuration:

```
DellEMC(conf)# ip route 199.1.1.0 /24 vlan 100 name "Uplink To NewYork"
DellEMC(conf)# ipv6 route 1001:1001::/64 GigabitEthernet 1/42 2001:2001::1 name
ipv6_link_going_to_europe_centre
DellEMC(conf)# ip route 19.1.1.0/24 19.1.1.1 name This_link_goes_to_London_Central
```

To view the description for the IPv4 or IPv6 static routes, use the `show running-config static` command. Following is the sample `show running-config static` output:

```
DellEMC#show running-config static
!
ipv6 route 1::/32 GigabitEthernet 2/3 11::1 name Stack-2
ipv6 route 2::/32 GigabitEthernet 2/48 11::1 name Stack-2
ip route 2.2.2.0/24 GigabitEthernet 2/47 name Stack-2
ipv6 route 1001:1001::/64 GigabitEthernet 1/42 2001:2001::1 name
ipv6_link_going_to_europe_centre
ip route 19.1.1.0/24 19.1.1.1 name This_link_goes_to_London_Central
ipv6 route 1500:1500::/96 2500:2500::1 name Under-Sea_Links_to_Asia-Pacific
ip route 50.50.1.32/30 Vlan 210 name AZ!#$_Link_to_Asia_$#
ip route 199.1.1.0 /24 vlan 100 name "Uplink To NewYork"
ip route 100.1.1.0/24 Vlan 100 name Any_Description_upto_32_characte
ip route 199.1.1.0/24 GigabitEthernet 1/41 name Added_Description_for_StaticRoute
```

NOTE: You can view the description of the configured static routes only using the `show running-config static` command.

Configure Static Routes for the Management Interface

When an IP address that a protocol uses and a static management route exists for the same prefix, the protocol route takes precedence over the static management route.

To configure a static route for the management port, use the following command.

- Assign a static route to point to the management interface or forwarding router.

  CONFIGURATION mode

  `management route` *ip-address mask* {*forwarding-router-address* | `ManagementEthernet` *slot/port*}

To view the configured static routes for the management port, use the `show ip management-route` command in EXEC privilege mode.

```
DellEMC#show ip management-route

Destination         Gateway                   State      Route Source
-----------         -------                   -----      ------------
10.16.0.0/16        ManagementEthernet 1/1    Connected  Connected
172.16.1.0/24       10.16.151.4               Active     Static
```

IPv4 Path MTU Discovery Overview

The size of the packet that can be sent across each hop in the network path without being fragmented is called the path maximum transmission unit (PMTU). This value might vary for the same route between two devices, mainly over a public network, depending on the network load and speed, and it is not a consistent value. The MTU size can also be different for various types of traffic sent from one host to the same endpoint.

Path MTU discovery (PMTD) identifies the path MTU value between the sender and the receiver, and uses the determined value to transmit packets across the network. PMTD, as described in RFC 1191, denotes that the default byte size of an IP packet is 576. This packet size is called the maximum transmission unit (MTU) for IPv4 frames. PMTD operates by containing the do not fragment (DF) bit set in the IP headers of outgoing packets. When any device along the network path contains an MTU that is smaller than the size of the packet that it receives, the device drops the packet and sends an Internet Control Message Protocol (ICMP) Fragmentation Needed (Type 3, Code 4) message with its MTU value to the source or the sending device. This message enables the source to identify that the transmitted packet size must be reduced. The packet is retransmitted with a lower size than the previous value. This process is repeated in an interactive way until the MTU of the transmitted packet is lower or equal to the MTU of the receiving device for it to obtain the packet without fragmentation. If the ICMP message from the receiving device, which is sent to the originating device, contains the next-hop MTU, then the sending device lowers the packet size accordingly and resends the packet. Otherwise, the iterative method is followed until the packet can traverse without being fragmented.

To use the PMTD in the physical interface, you must allocate and activate the fedgov CAM ACL space using the `cam-acl` command. The fedgov CAM ACL space is defined as a value (0-8) and you can select the required value to define the space. When the space is activated, you can send the ICMP type 3 messages through the physical port.

To use the PMTD functionality, you must enter the `ip unreachables` command on a VLAN interface to enable the generation of ICMP unreachable messages in the intermediate nodes. The PMTD functionality is based on the ICMPv4 destination unreachable message, which is generated by the intermediate device only when the `ip unreachables` command is configured. PMTD is supported on all the layer 3 VLAN interfaces. Since all of the Layer 3 interfaces are mapped to the VLAN ID of 4095, you cannot configure unique layer 3 MTU values for each of the layer 3 interfaces. If a VLAN interface contains both IPv4 and IPv6 addresses configured in the system, both the IPv4 and IPv6 traffic are applied to the same MTU size; you cannot specify different MTU values for IPv4 and IPv6 packets.

Packet handling during MTU mismatch

When you configure the MTU size on an interface, ensure that the MTU size of both ingress and egress interfaces are set to the same value for IPv4 traffic to work correctly. If there is an MTU mismatch between the ingress and egress interface, there may be a high CPU usage. If egress interface MTU size is smaller than the ingress interface, packets may get fragmented.

Using the Configured Source IP Address in ICMP Messages

ICMP error or unreachable messages are now sent with the configured IP address of the source interface instead of the front-end port IP address as the source IP address. Enable the generation of ICMP unreachable messages through the `ip unreachable` command in Interface mode. When a ping or traceroute packet from an endpoint or a device arrives at the null 0 interface configured with a static route, it is discarded. In such cases, you can configure Internet Control Message Protocol (ICMP) unreachable messages to be sent to the transmitting device.

Configuring the ICMP Source Interface

You can enable the ICMP error and unreachable messages to contain the configured IP address of the source device instead of the previous hop's IP address. This configuration helps identify the devices along the path because the DNS server maps the loopback IP address to the host name, and does not translate the IP address of every interface of the switch to the host name.

Configure the source to send the configured source interface IP address instead of using its front-end IP address in the ICMP unreachable messages and in the `traceroute` command output. Use the `ip icmp source-interface` *`interface`* or the `ipv6 icmp source-interface` *`interface`* commands in Configuration mode to enable the ICMP error messages to be sent with the source interface IP address. This functionality is supported on loopback, VLAN, port channel, and physical interfaces for IPv4 and IPv6 messages. feature is not supported on tunnel interfaces. ICMP error relay, PATH MTU transmission, and fragmented packets are not supported for tunnel interfaces. The traceroute utilities for IPv4 and IPv6 list the IP addresses of the devices in the hops of the path for which ICMP source interface is configured.

Configuring the Duration to Establish a TCP Connection

You can configure the duration for which the device must wait before it attempts to establish a TCP connection. Using this capability, you can limit the wait times for TCP connection requests. Upon responding to the initial SYN packet that requests a connection to the router for a specific service (such as SSH or BGP) with a SYN ACK, the router waits for a period of time for the ACK packet to be sent from the requesting host that will establish the TCP connection.

You can set this duration or interval for which the TCP connection waits to be established to a significantly high value to prevent the device from moving into an out-of-service condition or becoming unresponsive during a SYN flood attack that occurs on the device. You can set the wait time to be 10 seconds or lower. If the device does not contain any BGP connections with the BGP neighbors across WAN links, you must set this interval to a higher value, depending on the complexity of your network and the configuration attributes.

To configure the duration for which the device waits for the ACK packet to be sent from the requesting host to establish the TCP connection, perform the following steps:

1. Define the wait duration in seconds for the TCP connection to be established.

   CONFIGURATION mode

   ```
   DellEMC(conf)#ip tcp reduced-syn-ack-wait <9-75>
   ```

   You can use the `no ip tcp reduced-syn-ack-wait` command to restore the default behavior, which causes the wait period to be set as 8 seconds.
2. View the interval that you configured for the device to wait before the TCP connection is attempted to be established.

   EXEC mode

   ```
   DellEMC>show ip tcp reduced-syn-ack-wait
   ```

Enabling Directed Broadcast

By default, Dell EMC Networking OS drops directed broadcast packets destined for an interface. This default setting provides some protection against denial of service (DoS) attacks.

To enable Dell EMC Networking OS to receive directed broadcasts, use the following command.

- Enable directed broadcast.

INTERFACE mode

```
ip directed-broadcast
```

To view the configuration, use the `show config` command in INTERFACE mode.

Resolution of Host Names

Domain name service (DNS) maps host names to IP addresses. This feature simplifies commands such as Telnet and FTP by allowing you to enter a name instead of an IP address.

Dynamic resolution of host names is disabled by default. Unless you enable the feature, the system resolves only host names entered into the host table with the `ip host` command.

In a dual stack setup, the system sends both A (for IPv4 — RFC 1035) and AAAA (for IPv6 — RFC 3596) record requests to a DNS server even if you configure only the `ip name-server` command.

The following sections describe DNS and the resolution of host names.

- Enabling Dynamic Resolution of Host Names
- Specifying the Local System Domain and a List of Domains
- Configuring DNS with Traceroute

Name server, Domain name, and Domain list are VRF specific. The maximum number of Name servers and Domain lists per VRF is six.

Enabling Dynamic Resolution of Host Names

By default, dynamic resolution of host names (DNS) is disabled.

To enable DNS, use the following commands.

- Enable dynamic resolution of host names.

  CONFIGURATION mode

  ```
  ip domain-lookup
  ```

- Specify up to six name servers.

  CONFIGURATION mode

  ```
  ip name-server ip-address [ip-address2 ... ip-address6]
  ```

  The order you entered the servers determines the order of their use.

To view current bindings, use the `show hosts` command.

```
DellEMC>show host
Default domain is force10networks.com
Name/address lookup uses domain service
Name servers are not set
Host      Flags TTL     Type Address
--------- ----- ----    ---- -------
ks        (perm, OK) -  IP   2.2.2.2
patch1    (perm, OK) -  IP   192.68.69.2
tomm-3    (perm, OK) -  IP   192.68.99.2
gxr       (perm, OK) -  IP   192.71.18.2
f00-3     (perm, OK) -  IP   192.71.23.1
DellEMC>
```

To view the current configuration, use the `show running-config resolve` command.

Specifying the Local System Domain and a List of Domains

If you enter a partial domain, Dell EMC Networking OS can search different domains to finish or fully qualify that partial domain.

A fully qualified domain name (FQDN) is any name that is terminated with a period/dot. Dell EMC Networking OS searches the host table first to resolve the partial domain. The host table contains both statically configured and dynamically learnt host and IP addresses. If Dell EMC Networking OS cannot resolve the domain, it tries the domain name assigned to the local system. If that does not resolve the partial domain, Dell EMC Networking OS searches the list of domains configured.

To configure a domain name or a list of domain names, use the following commands.

- Enter up to 63 characters to configure one domain name.

  CONFIGURATION mode

  `ip domain-name` *name*
- Enter up to 63 characters to configure names to complete unqualified host names.

  CONFIGURATION mode

  `ip domain-list` *name*

  Configure this command up to six times to specify a list of possible domain names. Dell EMC Networking OS searches the domain names in the order they were configured until a match is found or the list is exhausted.

Configuring DNS with Traceroute

To configure your switch to perform DNS with traceroute, use the following commands.

- Enable dynamic resolution of host names.

  CONFIGURATION mode

  `ip domain-lookup`
- Specify up to six name servers.

  CONFIGURATION mode

  `ip name-server` *ip-address* [*ip-address2 ... ip-address6*]

  The order you entered the servers determines the order of their use.
- When you enter the `traceroute` command without specifying an IP address (`Extended Traceroute`), you are prompted for a target and source IP address, timeout in seconds (default is **5**), a probe count (default is **3**), minimum TTL (default is **1**), maximum TTL (default is **30**), and port number (default is **33434**).

  CONFIGURATION mode

  `traceroute` [*host* | *ip-address*]

  To keep the default setting for these parameters, press the ENTER key.

The following text is example output of DNS using the `traceroute` command.

```
DellEMC#traceroute www.force10networks.com

Translating "www.force10networks.com"...domain server (10.11.0.1) [OK]
Type Ctrl-C to abort.

------------------------------------------------------------------------
Tracing the route to www.force10networks.com (10.11.84.18), 30 hops max, 40 byte packets
------------------------------------------------------------------------

TTL Hostname       Probe1     Probe2     Probe3
1   10.11.199.190  001.000 ms 001.000 ms 002.000 ms
2   gwegress-sjc-02.force10networks.com (10.11.30.126) 005.000 ms 001.000 ms 001.000 ms
3   fw-sjc-01.force10networks.com (10.11.127.254) 000.000 ms 000.000 ms 000.000 ms
4   www.dell.com (10.11.84.18) 000.000 ms 000.000 ms 000.000 ms
DellEMC#
```

ARP

Dell EMC Networking OS uses two forms of address resolution: address resolution protocol (ARP) and Proxy ARP.

ARP runs over Ethernet and enables endstations to learn the MAC addresses of neighbors on an IP network. Over time, Dell EMC Networking OS creates a forwarding table mapping the MAC addresses to their corresponding IP address. This table is called the ARP Cache and dynamically learned addresses are removed after a defined period of time.

For more information about ARP, refer to RFC 826, *An Ethernet Address Resolution Protocol*.

In Dell EMC Networking OS, Proxy ARP enables hosts with knowledge of the network to accept and forward packets from hosts that contain no knowledge of the network. Proxy ARP makes it possible for hosts to be ignorant of the network, including subnetting.

For more information about Proxy ARP, refer to RFC 925, *Multi-LAN Address Resolution*, and RFC 1027, *Using ARP to Implement Transparent Subnet Gateways*.

Configuration Tasks for ARP

For a complete listing of all ARP-related commands, refer to the *Dell EMC Networking OS Command Line Reference Guide*.

Configuration tasks for ARP include:

- Configuring Static ARP Entries (optional)
- Enabling Proxy ARP (optional)
- Clearing ARP Cache (optional)
- ARP Learning via Gratuitous ARP
- ARP Learning via ARP Request
- Configuring ARP Retries

Configuring Static ARP Entries

ARP dynamically maps the MAC and IP addresses, and while most network host support dynamic mapping, you can configure an ARP entry (called a static ARP) for the ARP cache.

To configure a static ARP entry, use the following command.

- Configure an IP address and MAC address mapping for an interface.

  CONFIGURATION mode

  `arp vrf vrf-name ip-address mac-address interface`

  - `vrf vrf-name`: use the VRF option to configure a static ARP on that particular VRF.
  - *`ip-address`*: IP address in dotted decimal format (A.B.C.D).
  - *`mac-address`*: MAC address in nnnn.nnnn.nnnn format.
  - *`interface`*: enter the interface type slot/port information.

These entries do not age and can only be removed manually. To remove a static ARP entry, use the `no arp ip-address` command.

To view the static entries in the ARP cache, use the `show arp static` command in EXEC privilege mode.

```
DellEMC#show arp

Protocol  Address  Age(min) Hardware Address   Interface VLAN CPU
--------------------------------------------------------------------------------
Internet  10.1.2.4  17      08:00:20:b7:bd:32  Ma 1/1     -   CP
DellEMC#
```

Enabling Proxy ARP

By default, Proxy ARP is enabled. To disable Proxy ARP, use the `no proxy-arp` command in the interface mode.

To re-enable Proxy ARP, use the following command.

- Re-enable Proxy ARP.

  INTERFACE mode

  ```
  ip proxy-arp
  ```

To view if Proxy ARP is enabled on the interface, use the `show config` command in INTERFACE mode. If it is not listed in the `show config` command output, it is enabled. Only non-default information is displayed in the `show config` command output.

Clearing ARP Cache

To clear the ARP cache of dynamically learnt ARP information, use the following command.

- Clear the ARP caches for all interfaces or for a specific interface by entering the following information.

  EXEC privilege

  ```
  clear arp-cache [interface | ip ip-address] [no-refresh]
  ```

  - `ip ip-address` (OPTIONAL): enter the keyword `ip` then the IP address of the ARP entry you wish to clear.
  - `no-refresh` (OPTIONAL): enter the keywords `no-refresh` to delete the ARP entry from CAM. Or to specify which dynamic ARP entries you want to delete, use this option with *`interface`* or *`ip ip-address`*.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

NOTE: Transit traffic may not be forwarded during the period when deleted ARP entries are resolved again and re-installed in CAM. Use this option with extreme caution.

ARP Learning via Gratuitous ARP

Gratuitous ARP can mean an ARP request or reply.

In the context of ARP learning via gratuitous ARP on Dell EMC Networking OS, the gratuitous ARP is a request. A gratuitous ARP request is an ARP request that is not needed according to the ARP specification, but one that hosts may send to:

- detect IP address conflicts
- inform switches of their presence on a port so that packets can be forwarded
- update the ARP table of other nodes on the network in case of an address change

In the request, the host uses its own IP address in the Sender Protocol Address and Target Protocol Address fields.

Enabling ARP Learning via Gratuitous ARP

To enable ARP learning via gratuitous ARP, use the following command.

- Enable ARP learning via gratuitous ARP.

  CONFIGURATION mode

  ```
  arp learn-enable
  ```

ARP Learning via ARP Request

In Dell EMC Networking OS versions prior to 8.3.1.0, Dell EMC Networking OS learns via ARP requests only if the target IP specified in the packet matches the IP address of the receiving router interface. This is the case when a host is attempting to resolve the gateway address.

If the target IP does not match the incoming interface, the packet is dropped. If there is an existing entry for the requesting host, it is updated.

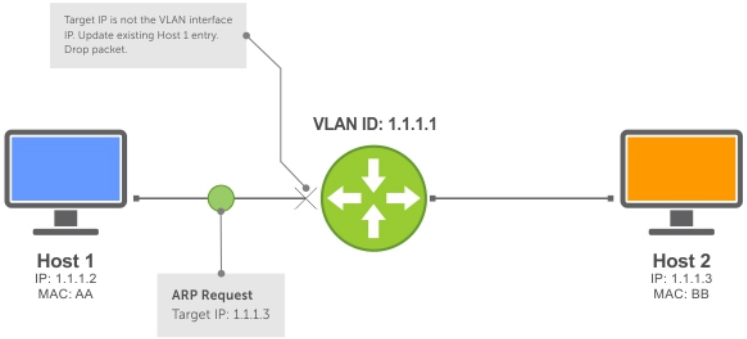


Figure 52. ARP Learning via ARP Request

Beginning with Dell EMC Networking OS version 8.3.1.0, when you enable ARP learning via gratuitous ARP, the system installs a new ARP entry, or updates an existing entry for all received ARP requests.

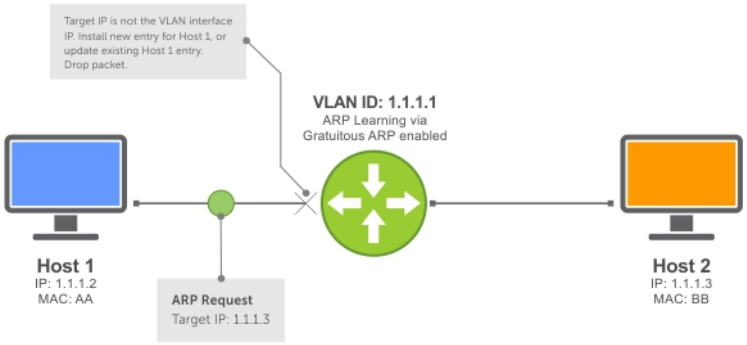


Figure 53. ARP Learning via ARP Request with ARP Learning via Gratuitous ARP Enabled

Whether you enable or disable ARP learning via gratuitous ARP, the system does not look up the target IP. It only updates the ARP entry for the Layer 3 interface with the source IP of the request.

Configuring ARP Retries

You can configure the number of ARP retries.

The default backoff interval remains at 20 seconds. On the device, the time between ARP resend is configurable. This timer is an exponential backoff timer. Over the specified period, the time between ARP requests increases. This time increase reduces the potential for the system to slow down while waiting for a multitude of ARP responses.

To set and display ARP retries, use the following commands.

- Set the number of ARP retries.
  CONFIGURATION mode
  `arp retries` *number*

The default is **5**.

The range is from 1 to 20.

- Set the exponential timer for resending unresolved ARPs.

  CONFIGURATION mode

  ```
  arp backoff-time
  ```

  The default is **30**.

  The range is from 1 to 3600.
- Display all ARP entries learned via gratuitous ARP.

  EXEC Privilege mode

  ```
  show arp retries
  ```

ICMP

For diagnostics, the internet control message protocol (ICMP) provides routing information to end stations by choosing the best route (ICMP redirect messages) or determining if a router is reachable (ICMP Echo or Echo Reply).

ICMP error messages inform the router of problems in a particular packet. These messages are sent only on unicast traffic.

Configuration Tasks for ICMP

The following lists the configuration tasks for ICMP.

- Enabling ICMP Unreachable Messages

For a complete listing of all commands related to ICMP, refer to the *Dell EMC Networking OS Command Line Reference Guide*.

Enabling ICMP Unreachable Messages

By default, ICMP unreachable messages are disabled.

When enabled, ICMP unreachable messages are created and sent out all interfaces.

To disable and re-enable ICMP unreachable messages, use the following commands.

- To disable ICMP unreachable messages.

  INTERFACE mode

  ```
  no ip unreachable
  ```
- Set Dell EMC Networking OS to create and send ICMP unreachable messages on the interface.

  INTERFACE mode

  ```
  ip unreachable
  ```

To view if ICMP unreachable messages are sent on the interface, use the `show config` command in INTERFACE mode. If it is not listed in the `show config` command output, it is enabled. Only non-default information is displayed in the `show config` command output.

ICMP Redirects

When a host sends a packet to a destination, it sends the packet to the configured default gateway. If the gateway router finds that a better route is available through a different router in the same network, that is, the same data link, the gateway router sends the source host an ICMP redirect message with the better route. The gateway router routes the packet to its destination and the host sends subsequent packets to that particular destination through the correct router.

Dell EMC Networking OS supports both ICMP and ICMP6 redirect messages. The following diagram depicts a topology in which ICMP redirect messages are useful.

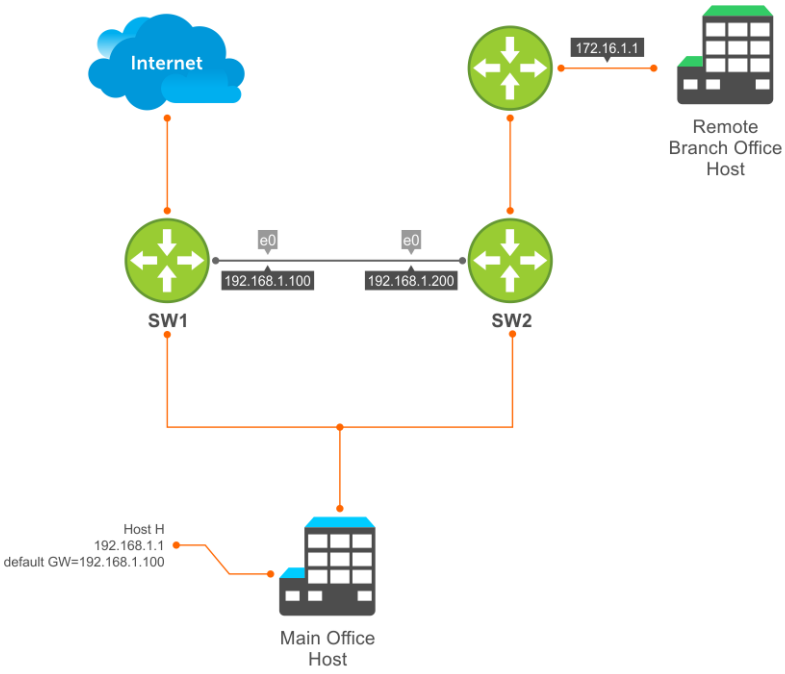


Figure 54. ICMP Redirect

Host H is connected to the same Ethernet segment as SW1 and SW2. SW1 and SW2 are multi-layer switches which can route packets. The default gateway of Host H is configured as SW1. Although the best route to the remote branch office host may be through SW2, Host H sends a packet destined for Host R to its default gateway — SW1. After SW1 finds that the route to Host R is through SW2, and that SW1 must forward the packet out the same Ethernet interface on which packet was received, it forwards the packet to SW2 and sends an ICMP redirect message to Host H. Host H learns of the new route using the ICMP redirect message and sends all future packets to the remote branch office host to SW2.

To enable ICMP or ICMP6 redirect messages, use the `icmp6-redirect enable` command.

NOTE: The `icmp6-redirect enable` command is applicable for both ICMP and ICMP6 redirects.

By default, Dell EMC Networking OS supports redirects on VLAN interfaces. For physical ports and port channel interfaces, carve the `fedgovacl` CAM region.

UDP Helper

User datagram protocol (UDP) helper allows you to direct the forwarding IP/UDP broadcast traffic by creating special broadcast addresses and rewriting the destination IP address of packets to match those addresses.

Configure UDP Helper

To configure Dell EMC Networking OS to direct UDP broadcast, enable UDP helper and specify the UDP ports for which traffic is forwarded. See Enabling UDP Helper

Important Points to Remember

- The existing `ip directed broadcast` command is rendered meaningless if you enable UDP helper on the same interface.
- The broadcast traffic rate should not exceed 200 packets per second when you enable UDP helper.
- You may specify a maximum of 16 UDP ports.
- UDP helper is compatible with IP helper (`ip helper-address`):

- UDP broadcast traffic with port number 67 or 68 are unicast to the dynamic host configuration protocol (DHCP) server per the `ip helper-address` configuration whether or not the UDP port list contains those ports.
- If the UDP port list contains ports 67 or 68, UDP broadcast traffic is forwarded on those ports.

Enabling UDP Helper

To enable UDP helper, use the following command.

- Enable UPD helper.
  `ip udp-helper udp-ports`

```
DellEMC(conf-if-te-1/1)#ip udp-helper udp-port 1000
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  ip address 2.1.1.1/24
  ip udp-helper udp-port 1000
  no shutdown
```

To view the interfaces and ports on which you enabled UDP helper, use the `show ip udp-helper` command from EXEC Privilege mode.

```
DellEMC#show ip udp-helper
-------------------------------------------------
Port UDP port list
-------------------------------------------------
te 1/1 1000
```

Configuring a Broadcast Address

To configure a broadcast address, use the following command.

- Configure a broadcast address on an interface.
  `ip udp-broadcast-address`

```
DellEMC(conf-if-vl-100)#ip udp-broadcast-address 1.1.255.255
DellEMC(conf-if-vl-100)#show config
!
interface Vlan 100
ip address 1.1.0.1/24
ip udp-broadcast-address 1.1.255.255
untagged TenGigabitEthernet 1/2
no shutdown
```

To view the configured broadcast address for an interface, use `show interfaces` command.

```
DellEMC#show interfaces vlan 100
Vlan 100 is up, line protocol is down
Address is 00:01:e8:0d:b9:7a, Current address is 00:01:e8:0d:b9:7a
Interface index is 1107787876
Internet address is 1.1.0.1/24
IP UDP-Broadcast address is 1.1.255.255
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:07:44
Queueing strategy: fifo
Input Statistics:
      0 packets, 0 bytes
Time since last interface status change: 00:07:44
```

Configurations Using UDP Helper

When you enable UDP helper and the destination IP address of an incoming packet is a broadcast address, Dell EMC Networking OS suppresses the destination address of the packet.

The following sections describe various configurations that employ UDP helper to direct broadcasts.

- UDP Helper with Broadcast-All Addresses
- UDP Helper with Subnet Broadcast Addresses
- UDP Helper with Configured Broadcast Addresses
- UDP Helper with No Configured Broadcast Addresses

UDP Helper with Broadcast-All Addresses

When the destination IP address of an incoming packet is the IP broadcast address, Dell EMC Networking OS rewrites the address to match the configured broadcast address.

In the following illustration:

1. Packet 1 is dropped at ingress if you did not configure UDP helper address.
2. If you enable UDP helper (using the `ip udp-helper udp-port` command), and the UDP destination port of the packet matches the UDP port configured, the system changes the destination address to the configured broadcast 1.1.255.255 and routes the packet to VLANs 100 and 101. If you do not configure an IP broadcast address (using the `ip udp-broadcast-address` command) on VLANs 100 or 101, the packet is forwarded using the original destination IP address 255.255.255.255.

Packet 2, sent from a host on VLAN 101 has a broadcast MAC address and IP address. In this case:

1. It is flooded on VLAN 101 without changing the destination address because the forwarding process is Layer 2.
2. If you enabled UDP helper, the system changes the destination IP address to the configured broadcast address 1.1.255.255 and forwards the packet to VLAN 100.
3. Packet 2 is also forwarded to the ingress interface with an unchanged destination address because it does not have broadcast address configured.

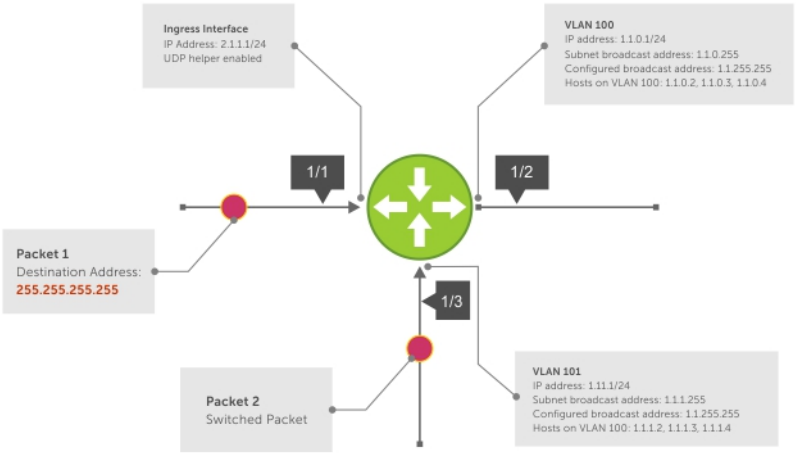


Figure 55. UDP Helper with Broadcast-All Addresses

UDP Helper with Subnet Broadcast Addresses

When the destination IP address of an incoming packet matches the subnet broadcast address of any interface, the system changes the address to the configured broadcast address and sends it to matching interface.

In the following illustration, Packet 1 has the destination IP address 1.1.1.255, which matches the subnet broadcast address of VLAN 101. If you configured UDP helper and the packet matches the specified UDP port, the system changes the address to the configured IP broadcast address and floods the packet on VLAN 101.

Packet 2 is sent from the host on VLAN 101. It has a broadcast MAC address and a destination IP address of 1.1.1.255. In this case, it is flooded on VLAN 101 in its original condition as the forwarding process is Layer 2.

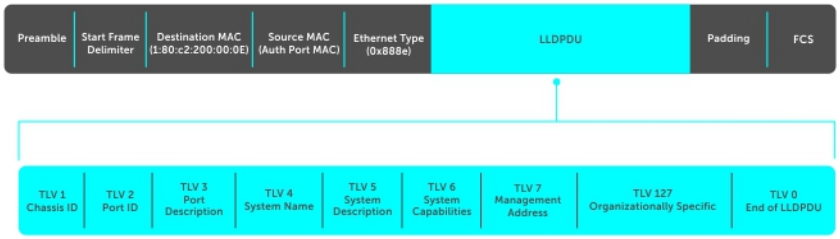


Figure 56. UDP Helper with Subnet Broadcast Addresses

UDP Helper with Configured Broadcast Addresses

Incoming packets with a destination IP address matching the configured broadcast address of any interface are forwarded to the matching interfaces.

In the following illustration, Packet 1 has a destination IP address that matches the configured broadcast address of VLAN 100 and 101. If you enabled UDP helper and the UDP port number matches, the packet is flooded on both VLANs with an unchanged destination address.

Packet 2 is sent from a host on VLAN 101. It has broadcast MAC address and a destination IP address that matches the configured broadcast address on VLAN 101. In this case, Packet 2 is flooded on VLAN 101 with the destination address unchanged because the forwarding process is Layer 2. If you enabled UDP helper, the packet is flooded on VLAN 100 as well.

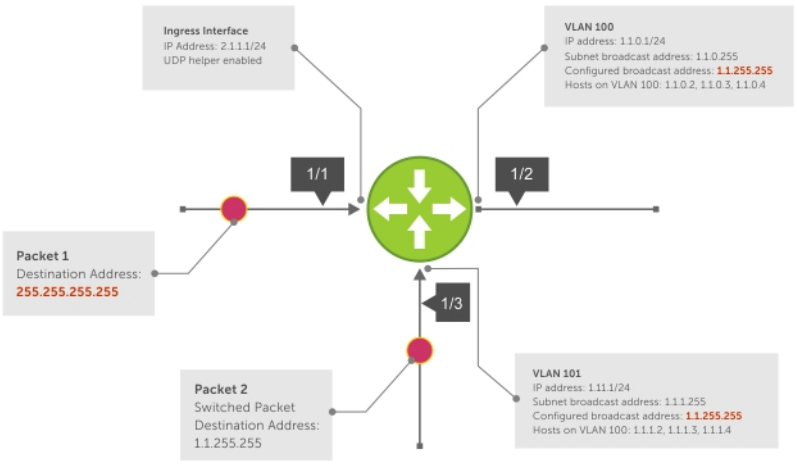


Figure 57. UDP Helper with Configured Broadcast Addresses

UDP Helper with No Configured Broadcast Addresses

The following describes UDP helper with no broadcast addresses configured.

- If the incoming packet has a broadcast destination IP address, the unaltered packet is routed to all Layer 3 interfaces.
- If the Incoming packet has a destination IP address that matches the subnet broadcast address of any interface, the unaltered packet is routed to the matching interfaces.

Troubleshooting UDP Helper

To display debugging information for troubleshooting, use the `debug ip udp-helper` command.

Example of the `debug ip udp-helper` Command

```
DellEMC(conf)# debug ip udp-helper
01:20:22: Pkt rcvd on Te 5/1 with IP DA (0xffffffff) will be sent on Te 5/2 Te 5/3 Vlan 3
01:44:54: Pkt rcvd on Te 7/1 is handed over for DHCP processing.
```

When using the IP helper and UDP helper on the same interface, use the `debug ip dhcp` command.

Example Output from the `debug ip dhcp` Command

```
Packet 0.0.0.0:68 -> 255.255.255.255:67 TTL 128

2017-08-05 11:59:35 %RELAY-I-PACKET, BOOTP REQUEST (Unicast) received at interface
172.21.50.193 BOOTP Request, XID = 0x9265f901, secs = 0 hwaddr = 00:02:2D:8D:46:DC,
giaddr = 0.0.0.0, hops = 2

2017-08-05 11:59:35 %RELAY-I-BOOTREQUEST, Forwarded BOOTREQUEST for 00:02:2D:8D:46:DC
to 137.138.17.6

2017-08-05 11:59:36 %RELAY-I-PACKET, BOOTP REPLY (Unicast) received at interface
194.12.129.98 BOOTP Reply, XID = 0x9265f901, secs = 0 hwaddr = 00:02:2D:8D:46:DC,
giaddr = 172.21.50.193, hops = 2

2017-08-05 11:59:36 %RELAY-I-BOOTREPLY, Forwarded BOOTREPLY for 00:02:2D:8D:46:DC to
128.141.128.90 Packet 0.0.0.0:68 -> 255.255.255.255:67 TTL 128
```

24

IPv6 Routing

Internet protocol version 6 (IPv6) routing is the successor to IPv4. Due to the rapid growth in internet users and IP addresses, IPv4 is reaching its maximum usage. IPv6 will eventually replace IPv4 usage to allow for the constant expansion.

This chapter provides a brief description of the differences between IPv4 and IPv6, and the Dell EMC Networking support of IPv6. This chapter is not intended to be a comprehensive description of IPv6.

NOTE: The IPv6 basic commands are supported on all platforms. However, not all features are supported on all platforms, nor for all releases. To determine the Dell EMC Networking Operating System (OS) version supporting which features and platforms, refer to Implementing IPv6 with Dell EMC Networking OS.

NOTE: Even though Dell EMC Networking OS listens to all ports, you can only use the ports starting from 1024 for IPv6 traffic. Ports from 0 to 1023 are reserved for internal use and you cannot use them for IPv6 traffic.

Topics:

- Protocol Overview
- Implementing IPv6 with Dell EMC Networking OS
- ICMPv6
- Path MTU discovery
- IPv6 Neighbor Discovery
- Configuration Task List for IPv6 RDNSS
- Secure Shell (SSH) Over an IPv6 Transport
- Configuration Tasks for IPv6
- Configuring IPv6 RA Guard

Protocol Overview

IPv6 is an evolution of IPv4. IPv6 is generally installed as an upgrade in devices and operating systems. Most new devices and operating systems support both IPv4 and IPv6.

Some key changes in IPv6 are:

- Extended address space
- Stateless autoconfiguration
- Header format simplification
- Improved support for options and extensions

Extended Address Space

The address format is extended from 32 bits to 128 bits. This not only provides room for all anticipated needs, it allows for the use of a hierarchical address space structure to optimize global addressing.

Stateless Autoconfiguration

When a booting device comes up in IPv6 and asks for its network prefix, the device can get the prefix (or prefixes) from an IPv6 router on its link. It can then autoconfigure one or more global IPv6 addresses by using either the MAC address or a private random number to build its unique IPv6 address.

Stateless autoconfiguration uses three mechanisms for IPv6 address configuration:

- **Prefix Advertisement** — Routers use “Router Advertisement” messages to announce the network prefix. Hosts then use their interface-identifier MAC address to generate their own valid IPv6 address.

- **Duplicate Address Detection (DAD)** — Before configuring its IPv6 address, an IPv6 host node device checks whether that address is used anywhere on the network using this mechanism.
- **Prefix Renumbering** — Useful in transparent renumbering of hosts in the network when an organization changes its service provider.

NOTE: As an alternative to stateless autoconfiguration, network hosts can obtain their IPv6 addresses using the dynamic host control protocol (DHCP) servers via stateful auto-configuration.

NOTE: Dell EMC Networking OS provides the flexibility to add prefixes on Router Advertisements (RA) to advertise responses to Router Solicitations (RS). By default, RA response messages are sent when an RS message is received.

Dell EMC Networking OS manipulation of IPv6 stateless autoconfiguration supports the router side only. Neighbor discovery (ND) messages are advertised so the neighbor can use this information to auto-configure its address. However, received ND messages are not used to create an IPv6 address.

NOTE: Inconsistencies in router advertisement values between routers are logged per RFC 4861. The values checked for consistency include:

- Cur Hop limit
- M and O flags
- Reachable time
- Retrans timer
- MTU options
- Preferred and valid lifetime values for the same prefix

Only management ports support stateless auto-configuration as a host.

The router redirect functionality in the neighbor discovery protocol (NDP) is similar to IPv4 router redirect messages. NDP uses ICMPv6 redirect messages (Type 137) to inform nodes that a better router exists on the link.

IPv6 Headers

The IPv6 header has a fixed length of 40 bytes. This fixed length provides 16 bytes each for source and destination information and 8 bytes for general header information.

The IPv6 header includes the following fields:

- Version (4 bits)
- Traffic Class (8 bits)
- Flow Label (20 bits)
- Payload Length (16 bits)
- Next Header (8 bits)
- Hop Limit (8 bits)
- Source Address (128 bits)
- Destination Address (128 bits)

IPv6 provides for extension headers. Extension headers are used only if necessary. There can be no extension headers, one extension header or more than one extension header in an IPv6 packet. Extension headers are defined in the Next Header field of the preceding IPv6 header.

IPv6 Header Fields

The 40 bytes of the IPv6 header are ordered, as shown in the following illustration.

Figure 58. IPv6 Header Fields

Version (4 bits)

The Version field always contains the number 6, referring to the packet's IP version.

Traffic Class (8 bits)

The Traffic Class field deals with any data that needs special handling. These bits define the packet priority and are defined by the packet Source. Sending and forwarding routers use this field to identify different IPv6 classes and priorities. Routers understand the priority settings and handle them appropriately during conditions of congestion.

Flow Label (20 bits)

The Flow Label field identifies packets requiring special treatment in order to manage real-time data traffic.

The sending router can label sequences of IPv6 packets so that forwarding routers can process packets within the same flow without needing to reprocess each packet's header separately.

NOTE: All packets in the flow must have the same source and destination addresses.

Payload Length (16 bits)

The Payload Length field specifies the packet payload. This is the length of the data following the IPv6 header. IPv6 Payload Length only includes the data following the header, not the header itself.

The Payload Length limit of 2 bytes requires that the maximum packet payload be 64 KB. However, the Jumbogram option type Extension header supports larger packet sizes when required.

Next Header (8 bits)

The Next Header field identifies the next header's type. If an Extension header is used, this field contains the type of Extension header (as shown in the following table). If the next header is a transmission control protocol (TCP) or user datagram protocol (UDP) header, the value in this field is the same as for IPv4. The Extension header is located between the IP header and the TCP or UDP header.

The following lists the Next Header field values.

Value	Description
0	Hop-by-Hop option header
4	IPv4
6	TCP
8	Exterior Gateway Protocol (EGP)
41	IPv6
43	Routing header
44	Fragmentation header
50	Encrypted Security
51	Authentication header
59	No Next Header
60	Destinations option header

NOTE: This table is not a comprehensive list of Next Header field values. For a complete and current listing, refer to the Internet Assigned Numbers Authority (IANA) web page at .

Hop Limit (8 bits)

The Hop Limit field shows the number of hops remaining for packet processing. In IPv4, this is known as the Time to Live (TTL) field and uses seconds rather than hops.

Each time the packet moves through a forwarding router, this field decrements by 1. If a router receives a packet with a Hop Limit of 1, it decrements it to 0 (zero). The router discards the packet and sends an ICMPv6 message back to the sending router indicating that the Hop Limit was exceeded in transit.

Source Address (128 bits)

The Source Address field contains the IPv6 address for the packet originator.

Destination Address (128 bits)

The Destination Address field contains the intended recipient's IPv6 address. This can be either the ultimate destination or the address of the next hop router.

Extension Header Fields

Extension headers are used only when necessary. Due to the streamlined nature of the IPv6 header, adding extension headers do not severely impact performance. Each Extension headers's lengths vary, but they are always a multiple of 8 bytes.

Each extension header is identified by the Next Header field in the IPv6 header that precedes it. Extension headers are viewed only by the destination router identified in the Destination Address field. If the Destination Address is a multicast address, the Extension headers are examined by all the routers in that multicast group.

However, if the Destination Address is a Hop-by-Hop options header, the Extension header is examined by every forwarding router along the packet's route. The Hop-by-Hop options header must immediately follow the IPv6 header, and is noted by the value 0 (zero) in the Next Header field.

Extension headers are processed in the order in which they appear in the packet header.

Hop-by-Hop Options Header

The Hop-by-Hop options header contains information that is examined by every router along the packet's path. It follows the IPv6 header and is designated by the Next Header value 0 (zero).

When a Hop-by-Hop Options header is not included, the router knows that it does not have to process any router specific information and immediately processes the packet to its final destination.

When a Hop-by-Hop Options header is present, the router only needs this extension header and does not need to take the time to view further into the packet.

The Hop-by-Hop Options header contains:

- Next Header (1 byte)

This field identifies the type of header following the Hop-by-Hop Options header and uses the same values.

- Header Extension Length (1 byte)

This field identifies the length of the Hop-by-Hop Options header in 8-byte units, but does not include the first 8 bytes. Consequently, if the header is less than 8 bytes, the value is 0 (zero).

- Options (size varies)

This field can contain one or more options. The first byte if the field identifies the Option type, and directs the router how to handle the option.

00	Skip and continue processing.
01	Discard the packet.
10	Discard the packet and send an ICMP Parameter Problem Code 2 message to the packet's Source IP Address identifying the unknown option type.

11	Discard the packet and send an ICMP Parameter Problem, Code 2 message to the packet's Source IP Address only if the Destination IP Address is not a multicast address.

The second byte contains the Option Data Length.

The third byte specifies whether the information can change en route to the destination. The value is 1 if it can change; the value is 0 if it cannot change.

Addressing

IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:).

For example, 2001:0db8:0000:0000:0000:0000:1428:57ab is a valid IPv6 address. If one or more four-digit group(s) is 0000, the zeros may be omitted and replaced with two colons(::). For example, 2001:0db8:0000:0000:0000:0000:1428:57ab can be shortened to 2001:0db8::1428:57ab. Only one set of double colons is supported in a single address. Any number of consecutive 0000 groups may be reduced to two colons, as long as there is only one double colon used in an address. Leading and/or trailing zeros in a group can also be omitted (as in ::1 for localhost, 1:: for network addresses and :: for unspecified addresses).

All the addresses in the following list are all valid and equivalent.

- 2001:0db8:0000:0000:0000:0000:1428:57ab
- 2001:0db8:0000:0000:0000::1428:57ab
- 2001:0db8:0:0:0:0:1428:57ab
- 2001:0db8:0:0::1428:57ab
- 2001:0db8::1428:57ab
- 2001:db8::1428:57ab

IPv6 networks are written using classless inter-domain routing (CIDR) notation. An IPv6 network (or subnet) is a contiguous group of IPv6 addresses the size of which must be a power of two; the initial bits of addresses, which are identical for all hosts in the network, are called the network's prefix.

A network is denoted by the first address in the network and the size in bits of the prefix (in decimal), separated with a slash. Because a single host is seen as a network with a 128-bit prefix, host addresses may be written with a following /128.

For example, 2001:0db8:1234::/48 stands for the network with addresses 2001:0db8:1234:0000:0000:0000:0000:0000 through 2001:0db8:1234:ffff:ffff:ffff:ffff:ffff.

Link-local Addresses

Link-local addresses, starting with fe80:, are assigned only in the local link area.

The addresses are generated usually automatically by the operating system's IP layer for each network interface. This provides instant automatic network connectivity for any IPv6 host and means that if several hosts connect to a common hub or switch, they have an instant communication path via their link-local IPv6 address.

Link-local addresses cannot be routed to the public Internet.

Static and Dynamic Addressing

Static IPv6 addresses are manually assigned to a computer by an administrator.

Dynamic IPv6 addresses are assigned either randomly or by a server using dynamic host configuration protocol (DHCP). Even though IPv6 addresses assigned using DHCP may stay the same for long periods of time, they can change. In some cases, a network administrator may implement dynamically assigned static IPv6 addresses. In this case, a DHCP server is used, but it is specifically configured to always assign the same IPv6 address to a particular computer, and never to assign that IP address to another computer. This allows static IPv6 addresses to be configured in one place, without having to specifically configure each computer on the network in a different way.

In IPv6, every interface, whether using static or dynamic address assignments, also receives a local-link address automatically in the fe80::/64 subnet.

Implementing IPv6 with Dell EMC Networking OS

Dell EMC Networking OS supports both IPv4 and IPv6 and both may be used simultaneously in your system.

The following table lists the Dell EMC Networking OS version in which an IPv6 feature became available for each platform. The sections following the table give greater detail about the feature.

Table 45. Dell EMC Networking OS versions and supported platforms with IPv6 support

Feature and Functionality		Dell EMC Networking OS Release Introduction	Documentation and Chapter Location
		S4820T	
Basic IPv6 Commands		8.3.19	IPv6 Basic Commands in the *Dell EMC Networking OS Command Line Interface Reference Guide*.
IPv6 Basic Addressing			
IPv6 address types: Unicast		8.3.19	Extended Address Space
IPv6 neighbor discovery		8.3.19	IPv6 Neighbor Discovery
IPv6 stateless autoconfiguration		8.3.19	Stateless Autoconfiguration
IPv6 MTU path discovery		8.3.19	Path MTU Discovery
IPv6 ICMPv6		8.3.19	ICMPv6
IPv6 ping		8.3.19	ICMPv6
IPv6 traceroute		8.3.19	ICMPv6
IPv6 SNMP		8.3.19	
IPv6 Routing			
Static routing		8.3.19	Assigning a Static IPv6 Route
Route redistribution		8.3.19	OSPF, IS-IS, and IPv6 BGP chapters in the *Dell EMC Networking OS Command Line Reference Guide*.
Multiprotocol BGP extensions for IPv6		8.3.19	IPv6 BGP in the *Dell EMC Networking OS Command Line Reference Guide*.
IPv6 BGP MD5 Authentication		8.3.19	IPv6 BGP in the *Dell EMC Networking OS Command Line Reference Guide*.
IS-IS for IPv6		8.3.19	Intermediate System to Intermediate System IPv6 IS-IS in the *Dell EMC Networking OS Command Line Reference Guide*.
IS-IS for IPv6 support for redistribution		8.3.19	Intermediate System to Intermediate System IPv6 IS-IS in the *Dell EMC Networking OS Command Line Reference Guide*.
ISIS for IPv6 support for distribute lists and administrative distance		8.3.19	Intermediate System to Intermediate System

Table 45. Dell EMC Networking OS versions and supported platforms with IPv6 support (continued)

Feature and Functionality		Dell EMC Networking OS Release Introduction	Documentation and Chapter Location
		S4820T	
			IPv6 IS-IS in the *Dell EMC Networking OS Command Line Reference Guide*.
OSPF for IPv6 (OSPFv3)		9.1(0.0)	OSPFv3 in the *Dell EMC Networking OS Command Line Reference Guide*.
Equal Cost Multipath for IPv6		8.3.19	
IPv6 Services and Management			
Telnet client over IPv6 (outbound Telnet)		8.3.19	Configuring Telnet with IPv6 Control and Monitoring in the *Dell EMC Networking OS Command Line Reference Guide*.
Telnet server over IPv6 (inbound Telnet)		8.3.19	Configuring Telnet with IPv6 Control and Monitoring in the *Dell EMC Networking OS Command Line Reference Guide*.
Secure Shell (SSH) client support over IPv6 (outbound SSH) Layer 3 only		8.3.19	Secure Shell (SSH) Over an IPv6 Transport
Secure Shell (SSH) server support over IPv6 (inbound SSH) Layer 3 only		8.3.19	Secure Shell (SSH) Over an IPv6 Transport
IPv6 Access Control Lists		8.3.19	IPv6 Access Control Lists in the *Dell EMC Networking OS Command Line Reference Guide*.
IPv6 Multicast			
MLDv1/v2		8.3.19	IPv6 PIM in the *Dell EMC Networking OS Command Line Reference Guide*.

ICMPv6

ICMP for IPv6 combines the roles of ICMP, IGMP and ARP in IPv4. Like IPv4, it provides functions for reporting delivery and forwarding errors, and provides a simple echo service for troubleshooting. The Dell EMC Networking OS implementation of ICMPv6 is based on RFC 4443.

Generally, ICMPv6 uses two message types:

- Error reporting messages indicate when the forwarding or delivery of the packet failed at the destination or intermediate node. These messages include Destination Unreachable, Packet Too Big, Time Exceeded and Parameter Problem messages.
- Informational messages provide diagnostic functions and additional host functions, such as Neighbor Discovery and Multicast Listener Discovery. These messages also include Echo Request and Echo Reply messages.

The Dell EMC Networking OS `ping` and `traceroute` commands extend to support IPv6 addresses. These commands use ICMPv6 Type-2 messages.

Path MTU discovery

The size of the packet that can be sent across each hop in the network path without being fragmented is called the path maximum transmission unit (PMTU). The PMTU value might differ for the same route between two devices, mainly over a public network, depending on the network load and speed, and it is not a consistent value. The MTU size can also be different for various types of traffic sent from one host to the same endpoint.

Path MTU discovery (PMTD) identifies the path MTU value between the sender and the receiver, and uses the determined value to transmit the packets across the network. Path MTU, in accordance with RFC 1981, defines the largest packet size that can traverse a transmission path without suffering fragmentation. Path MTU for IPv6 uses ICMPv6 Type-2 messages to discover the largest MTU along the path from source to destination and avoid the need to fragment the packet. The recommended MTU for IPv6 is 1280. Greater MTU settings increase the processing efficiency because each packet carries more data while protocol overheads (for example, headers) or underlying per-packet delays remain fixed.

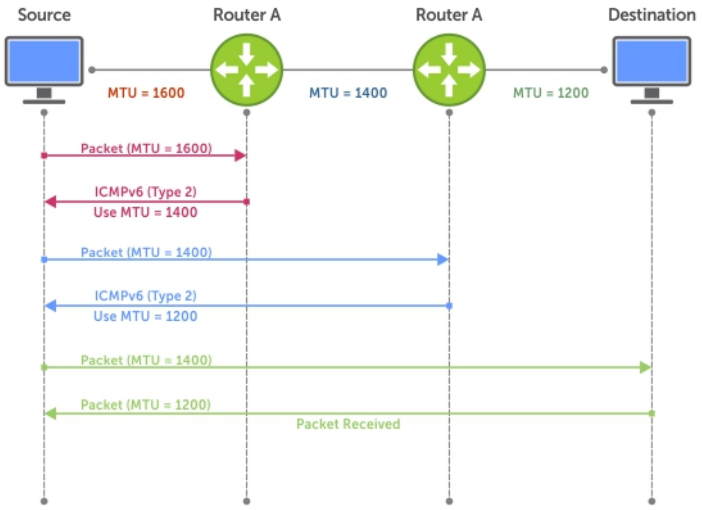


Figure 59. Path MTU discovery process

IPv6 Neighbor Discovery

The IPv6 neighbor discovery protocol (NDP) is a top-level protocol for neighbor discovery on an IPv6 network.

In place of address resolution protocol (ARP), NDP uses “Neighbor Solicitation” and “Neighbor Advertisement” ICMPv6 messages for determining relationships between neighboring nodes. Using these messages, an IPv6 device learns the link-layer addresses for neighbors known to reside on attached links, quickly purging cached values that become invalid.

NOTE: If a neighboring node does not have an IPv6 address assigned, it must be manually pinged to allow the IPv6 device to determine the relationship of the neighboring node.

NOTE: To avoid problems with network discovery, Dell EMC Networking recommends configuring the static route last or assigning an IPv6 address to the interface and assigning an address to the peer (the forwarding router’s address) less than 10 seconds apart.

With ARP, each node broadcasts ARP requests on the entire link. This approach causes unnecessary processing by uninterested nodes. With NDP, each node sends a request only to the intended destination via a multicast address with the unicast address used as the last 24 bits. Other hosts on the link do not participate in the process, greatly increasing network bandwidth efficiency.

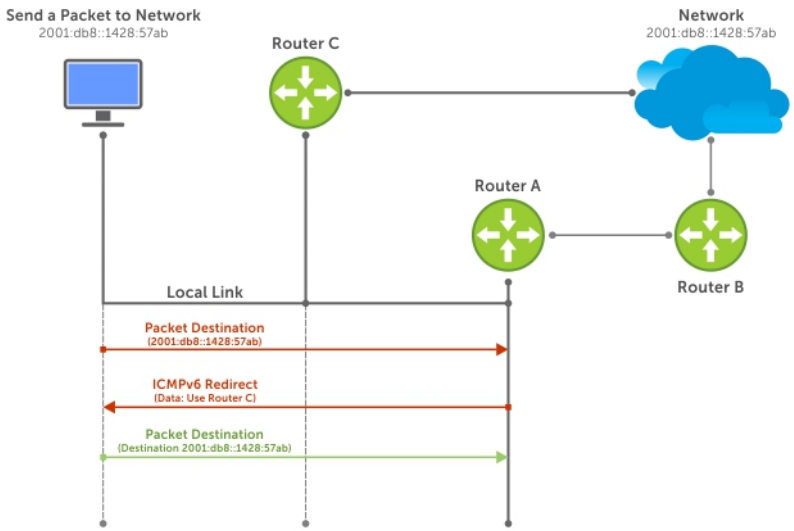


Figure 60. NDP Router Redirect

IPv6 Neighbor Discovery of MTU Packets

You can set the MTU advertised through the RA packets to incoming routers, without altering the actual MTU setting on the interface.

The `ipv6 nd mtu` command sets the value advertised to routers. It does not set the actual MTU rate. For example, if you set `ipv6 nd mtu` to 1280, the interface still passes 1500-byte packets, if that is what is set with the `mtu` command.

Configuration Task List for IPv6 RDNSS

This section describes how to configure the IPv6 Recursive DNS Server.

This sections contains the following configuration task list for IPv6 RDNSS:

- Configuring the IPv6 Recursive DNS Server
- Debugging IPv6 RDNSS Information Sent to the Host
- Displaying IPv6 RDNSS Information

Configuring the IPv6 Recursive DNS Server

You can configure up to four Recursive DNS Server (RDNSS) addresses to be distributed via IPv6 router advertisements to an IPv6 device, using the `ipv6 nd dns-server` *`ipv6-RDNSS-address`* `{`*`lifetime`* `|` `infinite}` command in INTERFACE CONFIG mode.

The lifetime parameter configures the amount of time the IPv6 host can use the IPv6 RDNSS address for name resolution. The lifetime range is `0` to `4294967295` seconds. When the maximum lifetime value, `4294967295`, or the `infinite` keyword is specified, the lifetime to use the RDNSS address does not expire. A value of `0` indicates to the host that the RDNSS address should not be used. You must specify a lifetime using the lifetime or infinite parameter.

The DNS server address does not allow the following:

- link local addresses
- loopback addresses
- prefix addresses
- multicast addresses

- invalid host addresses

If you specify this information in the IPv6 RDNSS configuration, a DNS error is displayed.

Example for Configuring an IPv6 Recursive DNS Server

The following example configures a RDNNS server with an IPv6 address of `1000::1` and a lifetime of `1` second.

```
DellEMC(conf-if-te-1/1)#ipv6 nd dns-server ?
X:X:X:X::X          Recursive DNS Server's (RDNSS) IPv6 address
DellEMC(conf-if-te-1/1)#ipv6 nd dns-server 1000::1 ?
<0-4294967295>       Max lifetime (sec) which RDNSS address may be used for name
resolution
infinite             Infinite lifetime (sec) which RDNSS address may be used for name
resolution

DellEMC(conf-if-te-1/1)#ipv6 nd dns-server 1000::1 1
```

Debugging IPv6 RDNSS Information Sent to the Host

To verify that the IPv6 RDNSS information sent to the host is configured correctly, use the `debug ipv6 nd` command in EXEC Privilege mode.

Example of Debugging IPv6 RDNSS Information Sent to the Host

The following example debugs IPv6 RDNSS information sent to the host.

```
DellEMC(conf-if-te-1/1)#do debug ipv6 nd tengigabitethernet 1/1
ICMPv6 Neighbor Discovery packet debugging is on for tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#00:13:02 : : cp-ICMPV6-ND: Sending RA on Te 1/1
            current hop limit=64, flags: M-, O-,
           router lifetime=1800 sec, reachable time=0 ms, retransmit time=0 ms
            SLLA=00:01:e8:8b:75:70
            prefix=1212::/64 on-link autoconfig
            valid lifetime=2592000 sec, preferred lifetime=604800 sec
            dns-server=1000::0001, lifetime=1 sec
            dns-server=3000::0001, lifetime=1 sec
            dns-server=2000::0001, lifetime=0 sec
```

The last 3 lines indicate that the IPv6 RDNSS information was configured correctly.

```
dns-server=1000::0001, lifetime=1 sec
dns-server=3000::0001, lifetime=1 sec
dns-server=2000::0001, lifetime=0 sec
```

If the DNS server information is not displayed, verify that the IPv6 recursive DNS server configuration was configured on the correct interface.

Displaying IPv6 RDNSS Information

To display IPv6 interface information, including IPv6 RDNSS information, use the `show ipv6 interface` command in EXEC or EXEC Privilege mode.

Examples of Displaying IPv6 RDNSS Information

The following example displays IPv6 RDNSS information. The output in the last 3 lines indicates that the IPv6 RDNSS was correctly configured on interface `te 1/1`.

```
DellEMC#show ipv6 interface te 1/1
TenGigabitEthernet 1/1 is up, line protocol is up
  IPV6 is enabled
  Link Local address: fe80::201:e8ff:fe8b:7570
  Global Unicast address(es):
    1212::12, subnet is 1212::/64 (MANUAL)
     Remaining lifetime: infinite
  Global Anycast address(es):
  Joined Group address(es):
```

```
    ff02::1
    ff02::2
    ff02::1:ff00:12
    ff02::1:ff8b:7570
  ND MTU is 0
  ICMP redirects are not sent
  DAD is enabled, number of DAD attempts: 3
  ND reachable time is 20120 milliseconds
  ND base reachable time is 30000 milliseconds
  ND advertised reachable time is 0 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 198 to 600 seconds
  ND router advertisements live for 1800 seconds
  ND advertised hop limit is 64
  IPv6 hop limit for originated packets is 64
  ND dns-server address is 1000::1 with lifetime of 1 seconds
  ND dns-server address is 3000::1 with lifetime of 1 seconds
  ND dns-server address is 2000::1 with lifetime of 0 seconds
  IP unicast RPF check is not supported
```

To display IPv6 RDNSS information, use the `show configuration` command in INTERFACE CONFIG mode.

```
DellEMC(conf-if-te-1/1)#show configuration
```

The following example uses the `show configuration` command to display IPv6 RDNSS information.

```
!
interface TenGigabitEthernet 1/1
no ip address
ipv6 address 1212::12/64
ipv6 nd dns-server 1000::1 1
ipv6 nd dns-server 3000::1 1
ipv6 nd dns-server 2000::1 0
no shutdown
```

Secure Shell (SSH) Over an IPv6 Transport

Dell EMC Networking OS supports both inbound and outbound SSH sessions using IPv6 addressing. Inbound SSH supports accessing the system through the management interface as well as through a physical Layer 3 interface.

For SSH configuration details, refer to the *Security* chapter in the *Dell EMC Networking OS Command Line Interface Reference Guide*.

Configuration Tasks for IPv6

The following are configuration tasks for the IPv6 protocol.

- Adjusting Your CAM-Profile
- Assigning an IPv6 Address to an Interface
- Assigning a Static IPv6 Route
- Configuring Telnet with IPv6
- SNMP over IPv6
- Showing IPv6 Information
- Clearing IPv6 Routes

Adjusting Your CAM-Profile

Although adjusting your CAM-profile is not a mandatory step, if you plan to implement IPv6 ACLs, adjust your CAM settings.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. There are 16 FP blocks, but the System Flow requires three blocks that cannot be reallocated.

You must enter the `ipv6acl` allocation as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd-numbered ranges.

The default option sets the CAM Profile as follows:

- L3 ACL (ipv4acl): 6
- L2 ACL(l2acl): 5
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1

To have the changes take effect, save the new CAM settings to the startup-config (`write-mem` or `copy run start`) then reload the system for the new settings.

- Allocate space for IPV6 ACLs. Enter the CAM profile name then the allocated amount.

  CONFIGURATION mode

  ```
  cam-acl { ipv6acl }
  ```

  When not selecting the default option, enter all of the profiles listed and a range for each.

  The total space allocated must equal 13.

  The `ipv6acl` range must be a factor of 2.
- Show the current CAM settings.

  EXEC mode or EXEC Privilege mode

  ```
  show cam-acl
  ```
- Provides information on FP groups allocated for the egress acl.

  CONFIGURATION mode

  ```
  show cam-acl-egress
  ```

  Allocate at least one group for L2ACL and IPv4 ACL.

  The total number of groups is 4.

Assigning an IPv6 Address to an Interface

Essentially, IPv6 is enabled in Dell EMC Networking OS simply by assigning IPv6 addresses to individual router interfaces. You can use IPv6 and IPv4 together on a system, but be sure to differentiate that usage carefully. To assign an IPv6 address to an interface, use the `ipv6 address` command.

NOTE: If you assign an IPv6 address to an interface enabled with DAD, Dell EMC Networking OS consider the assigned IPv6 address as tentative and will not use the address until the DAD verification is done. The time for an IPv6 address to become active increases as the number of DAD attempts increase.

You can configure up to two IPv6 addresses on management interfaces, allowing required default router support on the management port that is acting as host, per RFC 4861. Data ports support more than two IPv6 addresses.

When you configure IPv6 addresses on multiple interfaces (the `ipv6 address` command) and verify the configuration (the `show ipv6 interfaces` command), the same link local (fe80) address is displayed for each IPv6 interface.

- Enter the IPv6 Address for the device.

  CONFIG-INTERFACE mode

  ```
  ipv6 address ipv6 address/mask
  ```
  - *ipv6 address*: x:x:x:x::x
  - *mask*: The prefix length is from 0 to 128

  NOTE: IPv6 addresses are normally written as eight groups of four hexadecimal digits. Separate each group by a colon (:). Omitting zeros is accepted as described in Addressing.

Assigning a Static IPv6 Route

To configure IPv6 static routes, use the `ipv6 route` command.

NOTE: After you configure a static IPv6 route (the `ipv6 route` command) and configure the forwarding router's address (specified in the `ipv6 route` command) on a neighbor's interface, the IPv6 neighbor does not display in the `show ipv6 route` command output.

- Set up IPv6 static routes.
  CONFIGURATION mode
  `ipv6 route [vrf vrf-name] prefix interface-type slot/port forwarding router tag`
  - `vrf vrf-name`:(OPTIONAL) name of the VRF.
  - *`prefix`*: IPv6 route prefix
  - *slot/port* : interface type and slot/port
  - *`forwarding router`*: forwarding router's address
  - *`tag`*: route tag

  Enter the keyword `interface` then the type of interface and slot/port information:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a Null interface, enter the keyword null then the Null interface number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

Configuring Telnet with IPv6

The Telnet client and server in Dell EMC Networking OS supports IPv6 connections. You can establish a Telnet session directly to the router using an IPv6 Telnet client, or you can initiate an IPv6 Telnet connection from the router.

NOTE: Telnet to link local addresses is supported on the system.

- Enter the IPv6 Address for the device.
  EXEC mode or EXEC Privileged mode
  `telnet [vrf vrf-name] ipv6 address`
  - *`ipv6 address`*: x:x:x:x::x
  - *`mask`*: prefix length is from 0 to 128.

  NOTE: IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). Omitting zeros is accepted as described in Addressing.

SNMP over IPv6

You can configure SNMP over IPv6 transport so that an IPv6 host can perform SNMP queries and receive SNMP notifications from a device running Dell EMC Networking OS IPv6. The Dell EMC Networking OS SNMP-server commands for IPv6 have been extended to support IPv6. For more information regarding SNMP commands, refer to the *SNMP* and *SYSLOG* chapters in the *Dell EMC Networking OS Command Line Interface Reference Guide*.

- `snmp-server host`
- `snmp-server user ipv6`
- `snmp-server community ipv6`
- `snmp-server community access-list-name ipv6`
- `snmp-server group ipv6`
- `snmp-server group access-list-name ipv6`

Displaying IPv6 Information

View specific IPv6 configuration with the following commands.

- List the IPv6 show options.

  EXEC mode or EXEC Privileged mode

  `show ipv6 ?`

```
DellEMC#show ipv6 ?
accounting    IPv6 accounting information
cam           IPv6 CAM Entries
fib           IPv6 FIB Entries
interface     IPv6 interface information
mbgproutes    MBGP routing table
mld           MLD information
mroute        IPv6 multicast-routing table
neighbors     IPv6 neighbor information
ospf          OSPF information
pim           PIM V6 information
prefix-list   List IPv6 prefix lists
route         IPv6 routing information
rpf           RPF table
DellEMC#
```

Displaying an IPv6 Interface Information

To view the IPv6 configuration for a specific interface, use the following command.

- Show the currently running configuration for the specified interface.

  EXEC mode

  `show ipv6 interface` *`interface {slot/port}`*

  Enter the keyword `interface` then the type of interface and slot/port information:
  - For all brief summary of IPv6 status and configuration, enter the keyword `brief`.
  - For all IPv6 configured interfaces, enter the keyword `configured`.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

```
DellEMC#show ipv6 int ManagementEthernet 1/1
ManagementEthernet 1/1 is up, line protocol is up
  IPV6 is enabled
  Stateless address autoconfiguration is enabled
  Link Local address: fe80::201:e8ff:fe8b:386e
  Global Unicast address(es):
    Actual address is 400::201:e8ff:fe8b:386e, subnet is 400::/64
    Actual address is 412::201:e8ff:fe8b:386e, subnet is 412::/64
     Virtual-IP IPv6 address is not set
  Received Prefix(es):
    400::/64 onlink autoconfig
     Valid lifetime: 2592000, Preferred lifetime: 604800
     Advertised by: fe80::201:e8ff:fe8b:3166
  412::/64 onlink autoconfig
     Valid lifetime: 2592000, Preferred lifetime: 604800
     Advertised by: fe80::201:e8ff:fe8b:3166
  Global Anycast address(es):
  Joined Group address(es):
     ff02::1
     ff02::1:ff8b:386e
ND MTU is 0
ICMP redirects are not sent
DAD is enabled, number of DAD attempts: 3
ND reachable time is 32000 milliseconds
```

```
ND base reachable time is 30000 milliseconds
ND retransmit interval is 1000 milliseconds
ND hop limit is 64
```

Showing IPv6 Routes

To view the global IPv6 routing information, use the following command.

- Show IPv6 routing information for the specified route type.
  EXEC mode
  `show ipv6 route [vrf vrf-name] type`

  The following keywords are available:
  - To display information about a network, enter `ipv6 address` (X:X:X:X::X).
  - To display information about a host, enter `hostname`.
  - To display information about all IPv6 routes (including non-active routes), enter `all`.
  - To display information about all connected IPv6 routes, enter `connected`.
  - To display information about brief summary of all IPv6 routes, enter `summary`.
  - To display information about Border Gateway Protocol (BGP) routes, enter `bgp`.
  - To display information about ISO IS-IS routes, enter `isis`.
  - To display information about Open Shortest Path First (OSPF) routes, enter `ospf`.
  - To display information about Routing Information Protocol (RIP), enter `rip`.
  - To display information about static IPv6 routes, enter `static`.
  - To display information about an IPv6 Prefix lists, enter `list` and the prefix-list name.

The following example shows the `show ipv6 route summary` command.

```
DellEMC#show ipv6 route summary

Route Source Active Routes Non-active Routes
connected 5 0
static 0 0
Total 5 0
```

The following example shows the `show ipv6 route` command.

```
DellEMC#show ipv6 route
Codes: C - connected, L - local, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
Gateway of last resort is not set

   Destination  Dist/Metric,        Gateway, Last Change
------------------------------------------------------
C  600::/64 [0/0]
      Direct, Te 1/24, 00:34:42
C  601::/64 [0/0]
      Direct, Te 1/24, 00:34:18
C  912::/64 [0/0]
      Direct, Lo 2, 00:02:33
O  IA 999::1/128 [110/2]
      via fe80::201:e8ff:fe8b:3166, Te 1/24, 00:01:30
L  fe80::/10 [0/0]
      Direct, Nu 0, 00:34:42
DellEMC#
```

The following example shows the `show ipv6 route static` command.

```
DellEMC#show ipv6 route static
Destination Dist/Metric, Gateway, Last Change
---------------------------------------------------
  S         8888:9999:5555:6666:1111:2222::/96 [1/0]
               via         2222:2222:3333:3333::1, Te 9/1, 00:03:16
  S         9999:9999:9999:9999::/64 [1/0]
               via 8888:9999:5555:6666:1111:2222:3333:4444, 00:03:16
```

Showing the Running-Configuration for an Interface

To view the configuration for any interface, use the following command.

- Show the currently running configuration for the specified interface.

  EXEC mode

  `show running-config interface type {slot/port}`

  Enter the keyword `interface` then the type of interface and slot/port information:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For the Management interface on the stack-unit, enter the keyword `ManagementEthernet` then the slot/port information.

```
DellEMC#show run int Te 2/2
!
interface TenGigabitEthernet 2/2
  no ip address
  ipv6 address 3:4:5:6::8/24
  shutdown
DellEMC#
```

Clearing IPv6 Routes

To clear routes from the IPv6 routing table, use the following command.

- Clear (refresh) all or a specific route from the IPv6 routing table.

  EXEC mode

  `clear ipv6 route [vrf vrf-name] {* | ipv6 address prefix-length}`
  - `vrf vrf-name`:(OPTIONAL) name of the VRF.
  - `*`: all routes.
  - `ipv6 address`: the format is x:x:x:x::x.
  - `mask`: the prefix length is from 0 to 128.

  NOTE: IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). Omitting zeros is accepted as described in Addressing.

Disabling ND Entry Timeout

When a peer system warmboots or performs an ISSU, the ND entries in the local system may time out resulting in traffic loss. You can configure the system to keep the learnt neighbor discovery entries stateless so that the ND entries do not time out. To configure the system to keep the learnt neighbor discovery entries stateless, follow these steps:

- Disable the ND timer:

  INTERFACE

  `ipv6 nd disable-reachable-timer`

- To reenable the ND timer, use the no form of the command:
  INTERFACE
  `no ipv6 nd disable-reachable-timer`

The following example shows how to disable the ND timer.

```
DellEMC(conf-if-fo-1/1/1)#ipv6 nd disable-reachable-timer
```

Configuring IPv6 RA Guard

The IPv6 Router Advertisement (RA) guard allows you to block or reject the unwanted router advertisement guard messages that arrive at the network device platform.

To configure the IPv6 RA guard, perform the following steps:

1. Configure the terminal to enter the Global Configuration mode.
   EXEC Privilege mode
   `configure terminal`
2. Enable the IPv6 RA guard.
   CONFIGURATION mode
   `ipv6 nd ra-guard enable`
3. Create the policy.
   POLICY LIST CONFIGURATION mode
   `ipv6 nd ra-guard policy` *`policy-name`*
4. Define the role of the device attached to the port.
   POLICY LIST CONFIGURATION mode
   `device-role {host | router}`
   Use the keyword `host` to set the device role as host.
   Use the keyword `router` to set the device role as router.
5. Set the hop count limit.
   POLICY LIST CONFIGURATION mode
   `hop-limit {maximum | minimum` *`limit`*`}`
   The hop limit range is from 0 to 254.
6. Set the managed address configuration flag.
   POLICY LIST CONFIGURATION mode
   `managed-config-flag {on | off}`
7. Enable verification of the sender IPv6 address in inspected messages from the authorized device source access list.
   POLICY LIST CONFIGURATION mode
   `match ra{ipv6-access-list` *`name`* `| ipv6-prefix-list` *`name`* `| mac-access-list` *`name`*`}`
8. Enable verification of the advertised other configuration parameter.
   POLICY LIST CONFIGURATION mode
   `other-config-flag {on | off}`
9. Enable verification of the advertised default router preference value. The preference value must be less than or equal to the specified limit.
   POLICY LIST CONFIGURATION mode
   `router-preference maximum {high | low | medium}`
10. Set the router lifetime.
    POLICY LIST CONFIGURATION mode
    `router—lifetime` *`value`*
    The router lifetime range is from 0 to 9,000 seconds.
11. Apply the policy to trusted ports.

POLICY LIST CONFIGURATION mode

```
trusted-port
```

12. Set the maximum transmission unit (MTU) value.
    POLICY LIST CONFIGURATION mode

    ```
    mtu value
    ```

13. Set the advertised reachability time.
    POLICY LIST CONFIGURATION mode

    ```
    reachable—time value
    ```

    The reachability time range is from 0 to 3,600,000 milliseconds.
14. Set the advertised retransmission time.
    POLICY LIST CONFIGURATION mode

    ```
    retrans—timer value
    ```

    The retransmission time range is from 100 to 4,294,967,295 milliseconds.
15. Display the configurations applied on the RA guard policy mode.
    POLICY LIST CONFIGURATION mode

    ```
    show config
    ```

Example of the `show config` Command

```
DellEMC(conf-ra_guard_policy_list)#show config
!
ipv6 nd ra-guard policy test
 device-role router
 hop-limit maximum 251
 mtu 1350
 other-config-flag on
 reachable-time 540
 retrans-timer 101
 router-preference maximum medium
 trusted-port
DellEMC(conf-ra_guard_policy_list)#
```

Configuring IPv6 RA Guard on an Interface

To configure the IPv6 Router Advertisement (RA) guard on an interface, perform the following steps:

1. Configure the terminal to enter the Interface mode.
   CONFIGURATION mode

   ```
   interface interface-type slot/port
   ```

2. Apply the IPv6 RA guard to a specific interface.
   INTERFACE mode

   ```
   ipv6 nd ra-guard attach policy policy-name [vlan [vlan 1, vland 2, vlan 3.....]]
   ```

3. Display the configurations applied on all the RA guard policies or a specific RA guard policy.
   EXEC Privilege mode

   ```
   show ipv6 nd ra-guard policy policy-name
   ```

   The policy name string can be up to 140 characters.

Example of the `show ipv6 nd ra-guard policy` Command

```
DellEMC#show ipv6 nd ra-guard policy test

ipv6 nd ra-guard policy test
  device-role router
  hop-limit maximum    1
  match  ra  ipv6-access-list  access
  other-config-flag on
  router-preference maximum medium
  trusted-port
```

```
   Interfaces :
   Te 1/1
DellEMC#
```

Monitoring IPv6 RA Guard

To debug IPv6 RA guard, use the following command.

EXEC Privilege mode

`debug ipv6 nd ra-guard [interface slot/port | count value]`

The count range is from 1 to 65534. The default is infinity.

For a complete listing of all commands related to IPv6 RA Guard, see the *Dell EMC Networking OS Command Line Reference Guide*.

25

iSCSI Optimization

This chapter describes how to configure internet small computer system interface (iSCSI) optimization, which enables quality-of-service (QoS) treatment for iSCSI traffic. The topics covered in this chapter include:

- iSCSI Optimization Overview
- Default iSCSI Optimization Values
- iSCSI Optimization Prerequisites
- Configuring iSCSI Optimization
- Displaying iSCSI Optimization Information

Topics:

- iSCSI Optimization Overview
- Default iSCSI Optimization Values
- iSCSI Optimization Prerequisites
- Configuring iSCSI Optimization
- Displaying iSCSI Optimization Information

iSCSI Optimization Overview

iSCSI is a TCP/IP-based protocol for establishing and managing connections between IP-based storage devices and initiators in a storage area network (SAN).

iSCSI optimization enables the network switch to auto-detect Dell's iSCSI storage arrays and triggers a self-configuration of several key network configurations that enables optimization of the network for better storage traffic throughput. iSCSI is disabled by default.

iSCSI optimization also provides a means of monitoring iSCSI sessions and applying quality of service (QoS) policies on iSCSI traffic. When enabled, iSCSI optimization allows a switch to monitor (snoop) the establishment and termination of iSCSI connections. The switch uses the snooped information to detect iSCSI sessions and connections established through the switch.

iSCSI optimization allows you to reduce deployment time and management complexity in data centers. In a data center network, Dell EqualLogic and Compellent iSCSI storage arrays are connected to a converged Ethernet network using the data center bridging exchange protocol (DCBx) through stacked and/or non-stacked Ethernet switches.

iSCSI session monitoring over virtual link trunking (VLT) synchronizes the iSCSI session information between the VLT peers, allowing session information to be available in both the VLT peers. You can enable or disable iSCSI when you configure VLT.

iSCSI optimization functions as follows:

- Auto-detection of EqualLogic storage arrays — the switch detects any active EqualLogic array directly attached to its ports.
- Manual configuration to detect Compellent storage arrays where auto-detection is not supported.
- Automatic configuration of switch ports after detection of storage arrays.
- If you configure flow-control, iSCSI uses the current configuration. If you do not configure flow-control, iSCSI auto-configures flow control settings so that receive-only is enabled and transmit-only is disabled. .
- iSCSI monitoring sessions — the switch monitors and tracks active iSCSI sessions in connections on the switch, including port information and iSCSI session information.
- iSCSI QoS — A user-configured iSCSI class of service (CoS) profile is applied to all iSCSI traffic. Classifier rules are used to direct the iSCSI data traffic to queues that can be given preferential QoS treatment over other data passing through the switch. Preferential treatment helps to avoid session interruptions during times of congestion that would otherwise cause dropped iSCSI packets.
- iSCSI DCBx TLVs are supported.

NOTE: After a switch is reloaded, powercycled, or upgraded, any information exchanged during the initial handshake is not available. If the switch establishes communication after reloading, it detects that a session was in progress but could not obtain complete information for it. Any incomplete information is not available in the show commands.

NOTE: After a switch is reloaded, powercycled, or upgraded, the system may display the `ACL_AGENT-3-ISCSI_OPT_MAX_SESS_LIMIT_REACHED: Monitored iSCSI sessions reached maximum limit` log message. This cannot be inferred as the maximum supported iSCSI sessions are reached. Also, number of iSCSI sessions displayed on the system may show any number equal to or less than the maximum.

The following illustration shows iSCSI optimization between servers and a storage array in which a stack of three switches connect installed servers (iSCSI initiators) to a storage array (iSCSI targets) in a SAN network. iSCSI optimization running on the master switch is configured to use dot1p priority-queue assignments to ensure that iSCSI traffic in these sessions receives priority treatment when forwarded on stacked switch hardware.

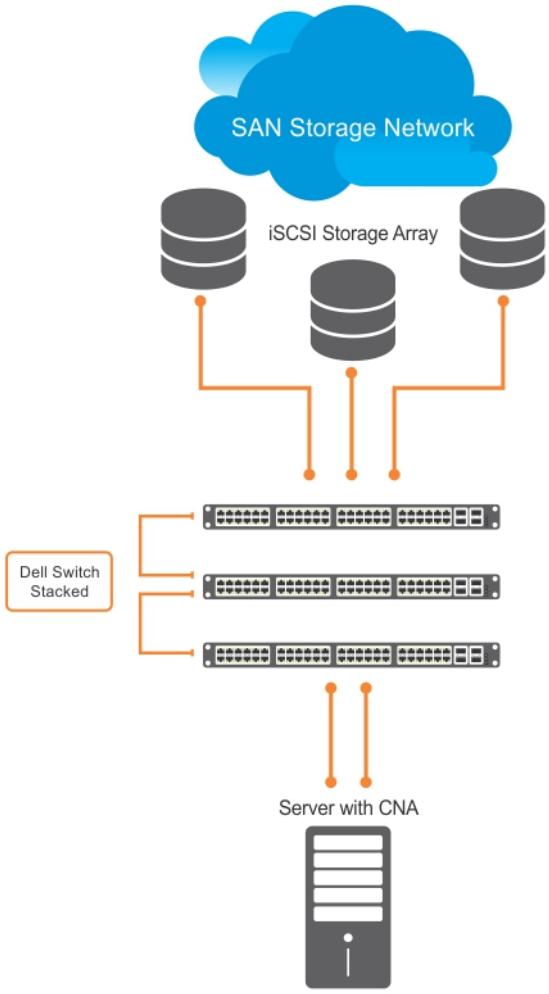


Figure 61. iSCSI Optimization Example

Monitoring iSCSI Traffic Flows

The switch snoops iSCSI session-establishment and termination packets by installing classifier rules that trap iSCSI protocol packets to the CPU for examination.

Devices that initiate iSCSI sessions usually use well-known TCP ports 3260 or 860 to contact targets. When you enable iSCSI optimization, by default the switch identifies IP packets to or from these ports as iSCSI traffic.

You can configure the switch to monitor traffic for additional port numbers or a combination of port number and target IP address, and you can remove the well-known port numbers from monitoring.

Application of Quality of Service to iSCSI Traffic Flows

You can configure iSCSI CoS mode. This mode controls whether CoS (dot1p priority) queue assignment and/or packet marking is performed on iSCSI traffic.

When you enable iSCSI CoS mode, the CoS policy is applied to iSCSI traffic. When you disable iSCSI CoS mode, iSCSI sessions and connections are still detected and displayed in the status tables, but no CoS policy is applied to iSCSI traffic.

You can configure whether the iSCSI optimization feature uses the VLAN priority or IP DSCP mapping to determine the traffic class queue. By default, iSCSI flows are assigned to dot1p priority 4. To map incoming iSCSI traffic on an interface to a dot1p priority-queue other than 4, use the `QoS dot1p-priority` command (refer to QoS dot1p Traffic Classification and Queue Assignment). Dell EMC Networking recommends setting the CoS dot1p priority-queue to 0 (zero).

You can configure whether iSCSI frames are re-marked to contain the configured VLAN priority tag or IP DSCP when forwarded through the switch.

NOTE: On a switch in which a large proportion of traffic is iSCSI, CoS queue assignments may interfere with other network control-plane traffic, such as ARP or LACP. Balance preferential treatment of iSCSI traffic against the needs of other critical data in the network.

Information Monitored in iSCSI Traffic Flows

iSCSI optimization examines the following data in packets and uses the data to track the session and create the classifier entries that enable QoS treatment.

- Initiator's IP Address
- Target's IP Address
- ISID (Initiator defined session identifier)
- Initiator's IQN (iSCSI qualified name)
- Target's IQN
- Initiator's TCP Port
- Target's TCP Port
- Connection ID
- Aging
- Up Time

If no iSCSI traffic is detected for a session during a user-configurable aging period, the session data is cleared.

If more than 256 simultaneous sessions are logged continuously, the following message displays indicating the queue rate limit has been reached:

```
%STKUNIT2-M:CP %iSCSI-5-ISCSI_OPT_MAX_SESS_EXCEEDED: New iSCSI Session Ignored: ISID -
400001370000 InitiatorName - iqn.1991-05.com.microsoft:dt-brcd-cna-2 TargetName -
iqn.2001-05.com.equallogic:4-52aed6-b90d9446c-162466364804fa49-wj-v1 TSIH - 0"
```

NOTE: If you are using EqualLogic or Compellent storage arrays, more than 256 simultaneous iSCSI sessions are possible. However, iSCSI session monitoring is not capable of monitoring more than 256 simultaneous iSCSI sessions. If this number is exceeded, sessions may display as `unknown` in session monitoring output. Dell EMC Networking recommends that you disable iSCSI session monitoring for EqualLogic and Compellent storage arrays or for installations with more than 256 simultaneous iSCSI sessions.

Only sessions the switch observes are learned; sessions flowing through an adjacent switch are not learned. Session monitoring learns sessions that actually flow through the switch, it does not learn all sessions in the entire topology.

After a switch is reloaded, any information exchanged during the initial handshake is not available. If the switch picks up the communication after reloading, it would detect a session was in progress but could not obtain complete information for it. Any incomplete information of this type would not be available in the `show` commands.

Detection and Auto-Configuration for Dell EqualLogic Arrays

The iSCSI optimization feature includes auto-provisioning support with the ability to detect directly connected Dell EqualLogic storage arrays and automatically reconfigure the switch to enhance storage traffic flows.

The switch uses the link layer discovery protocol (LLDP) to discover Dell EqualLogic devices on the network. LLDP is enabled by default. For more information about LLDP, refer to Link Layer Discovery Protocol (LLDP).

The following message displays the first time a Dell EqualLogic array is detected and describes the configuration changes that are automatically performed:

```
%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_AUTO_CONFIG: This switch is being configured for
optimal
conditions to support iSCSI traffic which will cause some automatic configuration to
occur
including jumbo frames and flow-control on all ports; no storm control and spanning-tree
port
fast to be enabled on the port of detection.
```

The following syslog message is generated the first time an EqualLogic array is detected:

```
%STKUNIT0-M:CP %LLDP-5-LLDP_EQL_DETECTED: EqualLogic Storage Array detected on interface
Te 1/
43
```

- At the first detection of an EqualLogic array, the maximum supported MTU is enabled on all ports and port-channels (if it has not already been enabled).
- Spanning-tree portfast is enabled on the interface LLDP identifies.
- Unicast storm control is disabled on the interface LLDP identifies.

Configuring Detection and Ports for Dell Compellent Arrays

To configure a port connected to a Dell Compellent storage array, use the following command.

- Configure a port connected to a Dell Compellent storage array.

  INTERFACE Configuration mode

  `iscsi profile-compellent`

  The command configures a port for the best iSCSI traffic conditions.

The following message displays the first time you use the `iscsi profile-compellent` command to configure a port connected to a Dell Compellent storage array and describes the configuration changes that are automatically performed:

```
%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_AUTO_CONFIG: This switch is being configured for
optimal
conditions to support iSCSI traffic which will cause some automatic configuration to
occur
including jumbo frames and flow-control on all ports; no storm control and spanning-tree
port
fast to be enabled on the port of detection.
```

After you execute the `iscsi profile-compellent` command, the following actions occur:

- Jumbo frame size is set to the maximum for all interfaces on all ports and port-channels, if it is not already enabled.
- Spanning-tree portfast is enabled on the interface.
- Unicast storm control is disabled on the interface.

Enter the `iscsi profile-compellent` command in INTERFACE Configuration mode.

Synchronizing iSCSI Sessions Learned on VLT-Lags with VLT-Peer

The following behavior occurs during synchronization of iSCSI sessions.

- If the iSCSI login request packet is received on a port belonging to a VLT lag, the information is synced to the VLT peer and the connection is associated with this interface.

- Additional updates to connections (including aging updates) that are learnt on VLT lag members are synced to the peer.
- When receiving an iSCSI login request on a non-VLT interface followed by a response from a VLT interface, the session is not synced since it is initially learnt on a non-VLT interface through the request packet.
- The peer generates a new connection log that sees the login response packet. If the login response packet uses the ICL path, it is seen by both the peers, which in turn generate logs for this connection.

Enable and Disable iSCSI Optimization

The following describes enabling and disabling iSCSI optimizaiton.

NOTE: iSCSI monitoring is disabled by default. iSCSI auto-configuration and auto-detection is enabled by default.

If you enable iSCSI, flow control is automatically enabled on all interfaces. To disable flow control on all interfaces, use the `no flow control rx on tx off` command and save the configuration. To disable iSCSI optimization, which can turn on flow control again on reboot, use the `no iscsi enable` command and save the configuration.

When you enable iSCSI on the switch, the following actions occur:

- Link-level flow control is globally enabled, if it is not already enabled, and PFC is disabled.
- iSCSI session snooping is enabled.
- iSCSI LLDP monitoring starts to automatically detect EqualLogic arrays.

The following message displays when you enable iSCSI on a switch and describes the configuration changes that are automatically performed:

```
%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_ENABLE: iSCSI has been enabled causing flow control to
be
enabled on all interfaces. EQL detection and enabling iscsi profile-compellent on an
interface
may cause some automatic configurations to occur like jumbo frames on all ports and no
storm
control and spanning tree port-fast on the port of detection.
```

You can reconfigure any of the auto-provisioned configuration settings that result when you enable iSCSI on a switch.

When you disable the iSCSI feature, iSCSI resources are released and the detection of EqualLogic arrays using LLDP is disabled. Disabling iSCSI does not remove the MTU, flow control, portfast, or storm control configuration applied as a result of enabling iSCSI.

NOTE: By default, CAM allocation for iSCSI is set to 0. This disables session monitoring.

Default iSCSI Optimization Values

The following table lists the default values for the iSCSI optimization feature.

Table 46. iSCSI Optimization Defaults

Parameter	Default Value
iSCSI Optimization global setting	Disabled on the S4810, S4820T, S3048–ON, S4048–ON, and S3100 series.
iSCSI CoS mode (802.1p priority queue mapping)	dot1p priority 4 without the `remark` setting when you enable iSCSI. If you do not enable iSCSI, this feature is disabled.
iSCSI CoS Packet classification	When you enable iSCSI, iSCSI packets are queued based on dot1p, instead of DSCP values.
VLAN priority tag	iSCSI flows are assigned by default to dot1p priority 4 without the `remark` setting.
DSCP	None: user-configurable.
Remark	Not configured.
iSCSI session aging time	10 minutes

Table 46. iSCSI Optimization Defaults (continued)

Parameter	Default Value
iSCSI optimization target ports	iSCSI well-known ports 3260 and 860 are configured as default (with no IP address or name) but can be removed as any other configured target.
iSCSI session monitoring	Disabled. The CAM allocation for iSCSI is set to zero (0).

iSCSI Optimization Prerequisites

The following are iSCSI optimization prerequisites.

- iSCSI optimization requires LLDP on the switch. LLDP is enabled by default (refer to Link Layer Discovery Protocol (LLDP)).
- iSCSI optimization requires configuring two ingress ACL groups The ACL groups are allocated after iSCSI Optimization is configured.

Configuring iSCSI Optimization

To configure iSCSI optimization, use the following commands.

1. **For a non-DCB environment:** Enable session monitoring.

   CONFIGURATION mode

   ```
   cam-acl l2acl 4 ipv4acl 4 ipv6acl 0 ipv4qos 2 l2qos 1 l2pt 0 ipmacacl 0 vman-qos 0
   ecfmacl 0 fcoeacl 0 iscsioptacl 2
   ```

   NOTE: Content addressable memory (CAM) allocation is optional. If CAM is not allocated, the following features are disabled:

   - session monitoring
   - aging
   - class of service

   You can enable iSCSI even when allocated with zero (0) CAM blocks. However, if no CAM blocks are allocated, session monitoring is disabled and this information the `show iscsi` command displays this information.

2. **For a non-DCB environment:** Enable iSCSI.

   CONFIGURATION mode

   ```
   iscsi enable
   ```

3. **For a DCB environment:** Configure iSCSI Optimization.

   EXEC Privilege mode

   iSCSI configuration: `copy CONFIG_TEMPLATE/iSCSI_DCB_Config running-config`.

   The configuration files are stored in the flash memory in the CONFIG_TEMPLATE file.

   NOTE: DCB/DCBx is enabled when you apply the iSCSI configuration in step 3. If you manually apply the iSCSI configuration by following steps 1 and 2, enable link layer discovery protocol (LLDP) before enabling iSCSI in step 2. You cannot disable LLDP if you enable iSCSI.

4. Save the configuration on the switch.

   EXEC Privilege mode

   ```
   write memory
   ```

5. Reload the switch.

   EXEC Privilege mode

   ```
   reload
   ```

   After the switch is reloaded, DCB/ DCBx and iSCSI monitoring are enabled.

6. (Optional) Configure the iSCSI target ports and optionally the IP addresses on which iSCSI communication is monitored.

   CONFIGURATION mode

`[no] iscsi target port tcp-port-1 [tcp-port-2...tcp-port-16] [ip-address address]`

- *tcp-port-n* is the TCP port number or a list of TCP port numbers on which the iSCSI target listens to requests. You can configure up to 16 target TCP ports on the switch in one command or multiple commands. The default is **860**, **3260**.

  Separate port numbers with a comma. If multiple IP addresses are mapped to a single TCP port, use the `no iscsi target port tcp-port-n` command to remove all IP addresses assigned to the TCP number.

  To delete a specific IP address from the TCP port, use the `no iscsi target port tcp-port-n ip-address address` command to specify the address to be deleted.

- *ip-address* specifies the IP address of the iSCSI target. When you enter the `no` form of the command, and the TCP port you want to delete is one bound to a specific IP address, include the IP address value in the command.

  If multiple IP addresses are mapped to a single TCP port, use the `no iscsi target port` command to remove all IP addresses assigned to the TCP port number.

  To remove a single IP address from the TCP port, use the `no iscsi target port ip-address` command.

7. (Optional) Set the QoS policy that is applied to the iSCSI flows.

   CONFIGURATION mode

   `[no] iscsi cos {enable | disable | dot1p vlan-priority-value [remark] | dscp dscp-value [remark]}`

   - `enable`: enables the application of preferential QoS treatment to iSCSI traffic so that iSCSI packets are scheduled in the switch with a dot1p priority 4 regardless of the VLAN priority tag in the packet. The default is: iSCSI packets are handled with dotp1 priority 4 without `remark`.
   - `disable`: disables the application of preferential QoS treatment to iSCSI frames.
   - `dot1p vlan-priority-value`: specifies the virtual local area network (VLAN) priority tag assigned to incoming packets in an iSCSI session. The range is from 0 to 7. The default is: the dot1p value in ingress iSCSI frames is not changed and the same priority is used in iSCSI TLV advertisements if you do not enter the `iscsi priority-bits` command (Step 10).
   - `dscp dscp-value`: specifies the DSCP value assigned to incoming packets in an iSCSI session. The range is from 0 to 63. The default is: the DSCP value in ingress packets is not changed.
   - `remark`: marks incoming iSCSI packets with the configured dot1p or DSCP value when they egress the switch. The default is: the dot1 and DSCP values in egress packets are not changed.

8. (Optional) Set the aging time for iSCSI session monitoring.

   CONFIGURATION mode

   `[no] iscsi aging time time`.

   The range is from 5 to 43,200 minutes.

   The default is **10 minutes**.

9. (Optional) Configures DCBX to send iSCSI TLV advertisements.

   LLDP CONFIGURATION mode or INTERFACE LLDP CONFIGURATION mode

   `[no] advertise dcbx-app-tlv iscsi`.

   You can send iSCSI TLVs either globally or on a specified interface. The interface configuration takes priority over global configuration.

   The default is **Enabled**.

10. (Optional) Configures the advertised priority bitmap in iSCSI application TLVs.

    LLDP CONFIGURATION mode

    `[no] iscsi priority-bits`.

    The default is **4** (0x10 in the bitmap).

11. (Optional) Configures the auto-detection of Compellent arrays on a port.

    INTERFACE mode

    `[no] iscsi profile-compellent`.

    The default is: Compellent disk arrays are not detected.

Displaying iSCSI Optimization Information

To display information on iSCSI optimization, use the following `show` commands.

- Display the currently configured iSCSI settings.
  `show iscsi`
- Display information on active iSCSI sessions on the switch.
  `show iscsi sessions`
- Display detailed information on active iSCSI sessions on the switch . To display detailed information on specified iSCSI session, enter the session's iSCSI ID.
  `show iscsi sessions detailed [session isid]`
- Display all globally configured non-default iSCSI settings in the current Dell EMC Networking OS session.
  `show run iscsi`

The following example shows the `show iscsi` command.

```
DellEMC#show iscsi
iSCSI is enabled
iSCSI session monitoring is disabled
iSCSI COS : dot1p is 4 no-remark
Session aging time: 10
Maximum number of connections is 256
-----------------------------------------------
iSCSI Targets and TCP Ports:
-----------------------------------------------
TCP Port Target IP Address
3260
860
```

The following example shows the `show iscsi session` command.

```
VLT PEER1

DellEMC#show iscsi session
Session 0:
-------------------------------------------------------------------------------
Target: iqn.2001-05.com.equallogic:0-8a0906-0e70c2002-10a0018426a48c94-iom010 Initiator:
iqn.1991-05.com.microsoft:win-x9l8v27yajg
ISID: 400001370000
VLT PEER2

Session 0:
-------------------------------------------------------------------------------
Target: iqn.2001-05.com.equallogic:0-8a0906-0f60c2002-0360018428d48c94-iom011
iqn.1991-05.com.microsoft:win-x9l8v27yajg
ISID: 400001370000
```

The following example shows the `show iscsi session detailed` command.

```
VLT PEER1

DellEMC# show iscsi session detailed
Session 0:
-----------------------------------------------------------
Target:iqn.2010-11.com.ixia:ixload:iscsi-TG1
Initiator:iqn.2010-11.com.ixia.ixload:initiator-iscsi-2c
Up Time:00:00:01:28(DD:HH:MM:SS)
Time for aging out:00:00:09:34(DD:HH:MM:SS)
ISID:806978696102
Initiator Initiator Target Target Connection
IP Address TCP Port IP Address TCPPort ID
10.10.0.44 33345 10.10.0.101 3260 0

VLT PEER2
Session 0:
-----------------------------------------------------------
Target:iqn.2010-11.com.ixia:ixload:iscsi-TG1
Initiator:iqn.2010-11.com.ixia.ixload:initiator-iscsi-2c
```

```
Up Time:00:00:01:28(DD:HH:MM:SS)
Time for aging out:00:00:09:34(DD:HH:MM:SS)
ISID:806978696102
Initiator   Initiator  Target       Target   Connection
IP Address  TCP Port   IP Address   TCPPort  ID
10.10.0.53  33432      10.10.0.101  3260     0
```

26

Intermediate System to Intermediate System

The intermediate system to intermediate system (IS-IS) protocol that uses a shortest-path-first algorithm. Dell EMC Networking supports both IPv4 and IPv6 versions of IS-IS.

Topics:

- IS-IS Protocol Overview
- IS-IS Addressing
- Multi-Topology IS-IS
- Graceful Restart
- Implementation Information
- Configuration Information
- IS-IS Metric Styles
- Configure Metric Values
- Sample Configurations

IS-IS Protocol Overview

The IS-IS protocol, developed by the International Organization for Standardization (ISO), is an interior gateway protocol (IGP) that uses a shortest-path-first algorithm.

NOTE: This protocol supports routers passing both IP and OSI traffic, though the Dell EMC Networking implementation supports only IP traffic.

IS-IS is organized hierarchically into routing domains and each router or system resides in at least one area. In IS-IS, routers are designated as Level 1, Level 2 or Level 1-2 systems. Level 1 routers only route traffic within an area, while Level 2 routers route traffic between areas. At its most basic, Level 1 systems route traffic within the area and any traffic destined for outside the area is sent to a Level 1-2 system. Level 2 systems manage destination paths for external routers. Only Level 2 routers can exchange data packets or routing information directly with external routers located outside of the routing domains. Level 1-2 systems manage both inter-area and intra-area traffic by maintaining two separate link databases; one for Level 1 routes and one for Level 2 routes. A Level 1-2 router does not advertise Level 2 routes to a Level 1 router.

To establish adjacencies, each IS-IS router sends different protocol data units (PDU). For IP traffic, the IP addressing information is included in the IS-IS hello PDUs and the link state PDUs (LSPs).

This brief overview is not intended to provide a complete understanding of IS-IS; for that, consult the documents listed in Multi-Topology IS-IS.

IS-IS Addressing

IS-IS PDUs require ISO-style addressing called network entity title (NET).

For those familiar with name-to-network service mapping point (NSAP) addresses, the composition of the NET is identical to an NSAP address, except the last byte is always 0. The NET is composed of the IS-IS area address, system ID, and N-selector. The last byte is the N-selector. All routers within an area have the same area portion. Level 1 routers route based on the system address portion of the address, while the Level 2 routers route based on the area address.

The NET length is variable, with a maximum of 20 bytes and a minimum of 8 bytes. It is composed of the following:

- **area address** — within your routing domain or area, each area must have a unique area value. The first byte is called the authority and format indicator (AFI).
- **system address** — the router's MAC address.
- **N-selector** — this is always 0.

The following illustration is an example of the ISO-style address to show the address format IS-IS uses. In this example, the first five bytes (47.0005.0001) are the area address. The system portion is 000c.000a.4321 and the last byte is always 0.

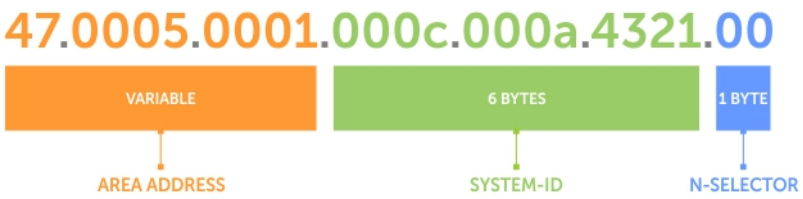


Figure 62. ISO Address Format

Multi-Topology IS-IS

Multi-topology IS-IS (MT IS-IS) allows you to create multiple IS-IS topologies on a single router with separate databases. Use this feature to place a virtual physical topology into logical routing domains, which can each support different routing and security policies.

All routers on a LAN or point-to-point must have at least one common supported topology when operating in Multi-Topology IS-IS mode. If IPv4 is the common supported topology between those two routers, adjacency can be formed. All topologies must share the same set of L1-L2 boundaries.

You must implement a wide metric-style globally on the autonomous system (AS) to run multi-topology IS-IS for IPv6 because the Type, Length, Value (TLVs) used to advertise IPv6 information in link-state packets (LSPs) are defined to use only extended metrics.

The multi-topology ID is shown in the first octet of the IS-IS packet. Certain MT topologies are assigned to serve predetermined purposes:

- MT ID #0: Equivalent to the “standard” topology.
- MT ID #1: Reserved for IPv4 in-band management purposes.
- MT ID #2: Reserved for IPv6 routing topology.
- MT ID #3: Reserved for IPv4 multicast routing topology.
- MT ID #4: Reserved for IPv6 multicast routing topology.
- MT ID #5: Reserved for IPv6 in-band management purposes.

Transition Mode

All routers in the area or domain must use the same type of IPv6 support, either single-topology or multi-topology. A router operating in multi-topology mode does not recognize the ability of the single-topology mode router to support IPv6 traffic, which leads to holes in the IPv6 topology.

While in Transition mode, both types of TLVs (single-topology and multi-topology) are sent in LSPs for all configured IPv6 addresses, but the router continues to operate in single-topology mode (that is, the topological restrictions of the single-topology mode remain in effect). Transition mode stops after all routers in the area or domain have been upgraded to support multi-topology IPv6. After all routers in the area or domain are operating in multi-topology IPv6 mode, the topological restrictions of single-topology mode are no longer in effect.

Interface Support

MT IS-IS is supported on physical Ethernet interfaces, physical synchronous optical network technologies (SONET) interfaces, port-channel interfaces (static and dynamic using LACP), and virtual local area network (VLAN) interfaces.

Adjacencies

Adjacencies on point-to-point interfaces are formed as usual, where IS-IS routers do not implement MT extensions.

If a local router does not participate in certain MTs, it does not advertise those MT IDs in its IS-IS hellos (IIHs) and so does not include that neighbor within its LSPs. If an MT ID is not detected in the remote side’s IIHs, the local router does not include that neighbor within its LSPs. The local router does not form an adjacency if both routers do not have at least one common MT over the interface.

Graceful Restart

Graceful restart is a protocol-based mechanism that preserves the forwarding table of the restarting router and its neighbors for a specified period to minimize the loss of packets. A graceful-restart router does not immediately assume that a neighbor is permanently down and so does not trigger a topology change.

Normally, when an IS-IS router is restarted, temporary disruption of routing occurs due to events in both the restarting router and the neighbors of the restarting router. When a router goes down without a graceful restart, there is a potential to lose access to parts of the network due to the necessity of network topology changes.

IS-IS graceful restart recognizes that in a modern router, the control plane and data plane are functionally separate. Restarting the control plane functionality (such as the failover of the active route processor module (RPM) to the backup in a redundant configuration) should not necessarily interrupt data packet forwarding. This behavior is supported because the forwarding tables previously computed by an active RPM have been downloaded into the forwarding information base (FIB) on the line cards (the data plane). For packets that have existing FIB/content addressable memory (CAM) entries, forwarding between ingress and egress ports can continue uninterrupted while the control plane IS-IS process comes back to full functionality and rebuilds its routing tables.

A new TLV (the Restart TLV) is introduced in the IIH PDUs, indicating that the router supports graceful restart.

Timers

Three timers are used to support IS-IS graceful restart functionality. After you enable graceful restart, these timers manage the graceful restart process.

There are three times, T1, T2, and T3.

- The T1 timer specifies the wait time before unacknowledged restart requests are generated. This is the interval before the system sends a Restart Request (an IIH with the RR bit set in Restart TLV) until the complete sequence number PDU (CSNP) is received from the helping router. You can set the duration to a specific amount of time (seconds) or a number of attempts.
- The T2 timer is the maximum time that the system waits for LSP database synchronization. This timer applies to the database type (level-1, level-2, or both).
- The T3 timer sets the overall wait time after which the router determines that it has failed to achieve database synchronization (by setting the overload bit in its own LSP). You can base this timer on adjacency settings with the value derived from adjacent routers that are engaged in graceful restart recovery (the minimum of all the Remaining Time values advertised by the neighbors) or by setting a specific amount of time manually.

Implementation Information

IS-IS implementation supports one instance of IS-IS and six areas.

You can configure the system as a Level 1 router, a Level 2 router, or a Level 1-2 router. For IPv6, the IPv4 implementation has been expanded to include two new type, length, values (TLVs) in the PDU that carry information required for IPv6 routing. The new TLVs are *IPv6 Reachability* and *IPv6 Interface Address*. Also, a new IPv6 protocol identifier has also been included in the supported TLVs. The new TLVs use the extended metrics and up/down bit semantics.

Multi-topology IS-IS adds TLVs:

- **MT TLV** — contains one or more Multi-Topology IDs in which the router participates. This TLV is included in IIH and the first fragment of an LSP.
- **MT Intermediate Systems TLV** — appears for every topology a node supports. An MT ID is added to the extended IS reachability TLV type 22.
- **MT Reachable IPv4 Prefixes TLV** — appears for each IPv4 an IS announces for a given MT ID. Its structure is aligned with the extended IS Reachability TLV Type 236 and it adds an MT ID.
- **MT Reachable IPv6 Prefixes TLV** — appears for each IPv6 an IS announces for a given MT ID. Its structure is aligned with the extended IS Reachability TLV Type 236 and add an MT ID.

By default, Dell EMC Networking OS supports dynamic host name exchange to assist with troubleshooting and configuration. By assigning a name to an IS-IS NET address, you can track IS-IS information on that address easier. Dell EMC Networking OS does not support ISO CLNS routing; however, the ISO NET format is supported for addressing.

To support IPv6, the Dell EMC Networking implementation of IS-IS performs the following tasks:

- Advertises IPv6 information in the PDUs.
- Processes IPv6 information received in the PDUs.

- Computes routes to IPv6 destinations.
- Downloads IPv6 routes to the RTM for installing in the FIB.
- Accepts external IPv6 information and advertises this information in the PDUs.

The following table lists the default IS-IS values.

Table 47. IS-IS Default Values

IS-IS Parameter	Default Value
Complete sequence number PDU (CSNP) interval	10 seconds
IS-to-IS hello PDU interval	10 seconds
IS-IS interface metric	10
Metric style	Narrow
Designated Router priority	64
Circuit Type	Level 1 and Level 2
IS Type	Level 1 and Level 2
Equal Cost Multi Paths	16

Configuration Information

To use IS-IS, you must configure and enable IS-IS in two or three modes: CONFIGURATION ROUTER ISIS, CONFIGURATION INTERFACE, and (when configuring for IPv6) ADDRESS-FAMILY mode. Commands in ROUTER ISIS mode configure IS-IS globally, while commands executed in INTERFACE mode enable and configure IS-IS features on that interface only. Commands in the ADDRESS-FAMILY mode are specific to IPv6.

NOTE: When using the IS-IS routing protocol to exchange IPv6 routing information and to determine destination reachability, you can route IPv6 along with IPv4 while using a single intra-domain routing protocol. The configuration commands allow you to enable and disable IPv6 routing and to configure or remove IPv6 prefixes on links.

Except where identified, the commands described in this chapter apply to both IPv4 and IPv6 versions of IS-IS.

Configuration Tasks for IS-IS

The following describes the configuration tasks for IS-IS.

- Enabling IS-IS
- Configure Multi-Topology IS-IS (MT IS-IS)
- Configuring IS-IS Graceful Restart
- Changing LSP Attributes
- Configuring the IS-IS Metric Style
- Configuring IS-IS Cost
- Changing the IS-Type
- Controlling Routing Updates
- Configuring Authentication Passwords
- Setting the Overload Bit
- Debugging IS-IS

Enabling IS-IS

By default, IS-IS is not enabled.

The system supports one instance of IS-IS. To enable IS-IS globally, create an IS-IS routing process and assign a NET address. To exchange protocol information with neighbors, enable IS-IS on an interface, instead of on a network as with other routing protocols.

In IS-IS, neighbors form adjacencies only when they are same IS type. For example, a Level 1 router never forms an adjacency with a Level 2 router. A Level 1-2 router forms Level 1 adjacencies with a neighboring Level 1 router and forms Level 2 adjacencies with a neighboring Level 2 router.

NOTE: Even though you enable IS-IS globally, enable the IS-IS process on an interface for the IS-IS process to exchange protocol information and form adjacencies.

To configure IS-IS globally, use the following commands.

1. Create an IS-IS routing process.

   CONFIGURATION mode

   `router isis [tag]`

   tag: (optional) identifies the name of the IS-IS process.

2. Configure an IS-IS network entity title (NET) for a routing process.

   ROUTER ISIS mode

   `net network-entity-title`

   Specify the area address and system ID for an IS-IS routing process. The last byte must be 00.

   For more information about configuring a NET, refer to IS-IS Addressing.

3. Enter the interface configuration mode.

   CONFIGURATION mode

   `interface interface`

   Enter the keyword `interface` then the type of interface and slot/port information:

   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
   - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
   - For a port channel interface, enter the keywords `port-channel` then a number.
   - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

4. Enter an IPv4 Address.

   INTERFACE mode

   `ip address ip-address mask`

   Assign an IP address and mask to the interface.

   The IP address must be on the same subnet as other IS-IS neighbors, but the IP address does not need to relate to the NET address.

5. Enter an IPv6 Address.

   INTERFACE mode

   `ipv6 address ipv6-address mask`

   - *ipv6 address*: x:x:x:x::x
   - *mask*: The prefix length is from 0 to 128.

   The IPv6 address must be on the same subnet as other IS-IS neighbors, but the IP address does not need to relate to the NET address.

6. Enable IS-IS on the IPv4 interface.

   ROUTER ISIS mode

   `ip router isis [tag]`

   If you configure a tag variable, it must be the same as the *tag* variable assigned in step 1.

7. Enable IS-IS on the IPv6 interface.

   ROUTER ISIS mode

   `ipv6 router isis [tag]`

   If you configure a tag variable, it must be the same as the *tag* variable assigned in step 1.

The default IS type is **level-1-2**. To change the IS type to Level 1 only or Level 2 only, use the `is-type` command in ROUTER ISIS mode.

To view the IS-IS configuration, enter the `show isis protocol` command in EXEC Privilege mode or the `show config` command in ROUTER ISIS mode.

```
DellEMC#show isis protocol
IS-IS Router: <Null Tag>
System Id: EEEE.EEEE.EEEE IS-Type: level-1-2
Manual area address(es):
  47.0004.004d.0001
Routing for area address(es):
  21.2223.2425.2627.2829.3031.3233
  47.0004.004d.0001
Interfaces supported by IS-IS:
  Vlan 2
  TenGigabitEthernet 4/22
  Loopback 0
Redistributing:
Distance: 115
Generate narrow metrics: level-1-2
Accept narrow metrics:   level-1-2
Generate wide metrics:   none
Accept wide metrics:     none
DellEMC#
```

To view IS-IS protocol statistics, use the `show isis traffic` command in EXEC Privilege mode.

```
DellEMC#show isis traffic
  IS-IS: Level-1 Hellos (sent/rcvd) : 4272/1538
  IS-IS: Level-2 Hellos (sent/rcvd) : 4272/1538
  IS-IS: PTP Hellos (sent/rcvd) : 0/0
  IS-IS: Level-1 LSPs sourced (new/refresh) : 0/0
  IS-IS: Level-2 LSPs sourced (new/refresh) : 0/0
  IS-IS: Level-1 LSPs flooded (sent/rcvd) : 32/19
  IS-IS: Level-2 LSPs flooded (sent/rcvd) : 32/17
  IS-IS: Level-1 LSPs CSNPs (sent/rcvd) : 1538/0
  IS-IS: Level-2 LSPs CSNPs (sent/rcvd) : 1534/0
  IS-IS: Level-1 LSPs PSNPs (sent/rcvd) : 0/0
  IS-IS: Level-2 LSPs PSNPs (sent/rcvd) : 0/0
  IS-IS: Level-1 DR Elections : 2
  IS-IS: Level-2 DR Elections : 2
  IS-IS: Level-1 SPF Calculations : 29
  IS-IS: Level-2 SPF Calculations : 29
  IS-IS: LSP checksum errors received : 0
  IS-IS: LSP authentication failures : 0
DellEMC#
```

You can assign more NET addresses, but the System ID portion of the NET address must remain the same. Dell EMC Networking OS supports up to six area addresses.

Some address considerations are:

- In order to be neighbors, configure Level 1 routers with at least one common area address.
- A Level 2 router becomes a neighbor with another Level 2 router regardless of the area address configured. However, if the area addresses are different, the link between the Level 2 routers is only at Level 2.

Configuring Multi-Topology IS-IS (MT IS-IS)

To configure multi-topology IS-IS (MT IS-IS), use the following commands.

1. Enable multi-topology IS-IS for IPv6.

   ROUTER ISIS AF IPV6 mode

   `multi-topology [transition]`

   Enter the keyword *transition* to allow an IS-IS IPv6 user to continue to use single-topology mode while upgrading to multi-topology mode. After every router has been configured with the *transition* keyword, and all the routers are in MT IS-IS IPv6 mode, you can remove the *transition* keyword on each router.

   NOTE: When you do not enable transition mode, you do not have IPv6 connectivity between routers operating in single-topology mode and routers operating in multi-topology mode.

2. Exclude this router from other router's SPF calculations.

ROUTER ISIS AF IPV6 mode

```
set-overload-bit
```

3. Set the minimum interval between SPF calculations.

   ROUTER ISIS AF IPV6 mode

   ```
   spf-interval [level-1 | level-2 | interval] [initial_wait_interval
   [second_wait_interval]]
   ```

   Use this command for IPv6 route computation only when you enable multi-topology. If using single-topology mode, to apply to both IPv4 and IPv6 route computations, use the `spf-interval` command in CONFIG ROUTER ISIS mode.

4. Implement a *wide metric-style* globally.

   ROUTER ISIS AF IPV6 mode

   ```
   isis ipv6 metric metric-value [level-1 | level-2 | level-1-2]
   ```

   To configure wide or wide transition metric style, the cost can be between 0 and 16,777,215.

Configuring IS-IS Graceful Restart

To enable IS-IS graceful restart globally, use the following commands. Additionally, you can implement optional commands to enable the graceful restart settings.

- Enable graceful restart on ISIS processes.

  ROUTER-ISIS mode

  ```
  graceful-restart ietf
  ```

- Configure the time during which the graceful restart attempt is prevented.

  ROUTER-ISIS mode

  ```
  graceful-restart interval minutes
  ```

  The range is from 1 to 120 minutes.

  The default is **5 minutes**.

- Enable the graceful restart maximum wait time before a restarting peer comes up.

  ROUTER-ISIS mode

  ```
  graceful-restart restart-wait seconds
  ```

  When implementing this command, be sure to set the t3 timer to adjacency on the restarting router.

  The range is from 1 to 120 minutes.

  The default is **30 seconds**.

- Configure the time that the graceful restart timer T1 defines for a restarting router to use for each interface, as an interval before regenerating Restart Request (an IIH with RR bit set in Restart TLV) after waiting for an acknowledgement.

  ROUTER-ISIS mode

  ```
  graceful-restart t1 {interval seconds | retry-times value}
  ```

  - `interval`: wait time (the range is from 5 to 120. The default is **5**.)
  - `retry-times`: number of times an unacknowledged restart request is sent before the restarting router gives up the graceful restart engagement with the neighbor. (The range is from 1 to 10 attempts. The default is **1**.)

- Configure the time for the graceful restart timer T2 that a restarting router uses as the wait time for each database to synchronize.

  ROUTER-ISIS mode

  ```
  graceful-restart t2 {level-1 | level-2} seconds
  ```

  - `level-1`, `level-2`: identifies the database instance type to which the wait interval applies.

  The range is from 5 to 120 seconds.

  The default is **30 seconds**.

- Configure graceful restart timer T3 to set the time used by the restarting router as an overall maximum time to wait for database synchronization to complete.

  ROUTER-ISIS mode

  ```
  graceful-restart t3 {adjacency | manual seconds}
  ```

- `adjacency`: the restarting router receives the remaining time value from its peer and adjusts its T3 value so if user has configured this option.
- `manual`: allows you to specify a fixed value that the restarting router should use.

The range is from 50 to 120 seconds.

The default is **30 seconds**.

NOTE: If this timer expires before the synchronization has completed, the restarting router sends the overload bit in the LSP. The 'overload' bit is an indication to the receiving router that database synchronization did not complete at the restarting router.

To view all graceful restart-related configurations, use the `show isis graceful-restart detail` command in EXEC Privilege mode.

```
DellEMC#show isis graceful-restart detail
Configured Timer Value
======================
Graceful Restart       : Enabled
Interval/Blackout time : 1 min
T3 Timer               : Manual
T3 Timeout Value       : 30
T2 Timeout Value       : 30 (level-1), 30 (level-2)
T1 Timeout Value       : 5, retry count: 1
Adjacency wait time    : 30

Operational Timer Value
=======================
Current Mode/State     : Normal/RUNNING
T3 Time left           : 0
T2 Time left           : 0 (level-1), 0 (level-2)
Restart ACK rcv count  : 0 (level-1), 0 (level-2)
Restart Req rcv count  : 0 (level-1), 0 (level-2)
Suppress Adj rcv count : 0 (level-1), 0 (level-2)
Restart CSNP rcv count : 0 (level-1), 0 (level-2)
Database Sync count    : 0 (level-1), 0 (level-2)

Circuit TenGigabitEthernet 2/10:
  Mode: Normal L1-State:NORMAL, L2-State: NORMAL

  L1: Send/Receive: RR:0/0, RA: 0/0, SA:0/0
    T1 time left: 0, retry count left:0

  L2: Send/Receive: RR:0/0, RA: 0/0, SA:0/0
    T1 time left: 0, retry count left:0
DellEMC#
```

Example of the `show isis interface` Command

To view all interfaces configured with IS-IS routing along with the defaults, use the `show isis interface` command in EXEC Privilege mode.

```
DellEMC#show isis interface TenGigabitEthernet 1/34
TenGigabitEthernet 1/34 is up, line protocol is up
  MTU 1497, Encapsulation SAP
  Routing Protocol: IS-IS
    Circuit Type: Level-1-2
    Interface Index 0x62cc03a, Local circuit ID 1
    Level-1 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
        Hello Interval: 10, Hello Multiplier: 3, CSNP Interval: 10
    Number of active level-1 adjacencies: 1
    Level-2 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
        Hello Interval: 10, Hello Multiplier: 3, CSNP Interval: 10
    Number of active level-2 adjacencies: 1
    Next IS-IS LAN Level-1 Hello in 4 seconds
    Next IS-IS LAN Level-2 Hello in 6 seconds
    LSP Interval: 33 Next IS-IS LAN Level-1 Hello in 4 seconds
    Next IS-IS LAN Level-2 Hello in 6 seconds
    LSP Interval: 33
Restart Capable Neighbors: 2, In Start: 0, In Restart: 0
DellEMC#
```

Changing LSP Attributes

IS-IS routers flood link state PDUs (LSPs) to exchange routing information. LSP attributes include the generation interval, maximum transmission unit (MTU) or size, and the refresh interval.

You can modify the LSP attribute defaults, but it is not necessary.

To change the defaults, use any or all of the following commands.

- Set interval between LSP generation.

  ROUTER ISIS mode

  `lsp-gen-interval [level-1 | level-2] seconds`

  - *seconds*: the range is from 0 to 120.

  The default is **5 seconds**.

  The default level is **Level 1**.
- Set the LSP size.

  ROUTER ISIS mode

  `lsp-mtu size`

  - *size*: the range is from 128 to 9195.

  The default is **1497**.
- Set the LSP refresh interval.

  ROUTER ISIS mode

  `lsp-refresh-interval seconds`

  - *seconds*: the range is from 1 to 65535.

  The default is **900 seconds**.
- Set the maximum time LSPs lifetime.

  ROUTER ISIS mode

  `max-lsp-lifetime seconds`

  - *seconds*: the range is from 1 to 65535.

  The default is **1200 seconds**.

To view the configuration, use the `show config` command in ROUTER ISIS mode or the `show running-config isis` command in EXEC Privilege mode.

```
DellEMC#show running-config isis
!
router isis
  lsp-refresh-interval 902
  net 47.0005.0001.000C.000A.4321.00
  net 51.0005.0001.000C.000A.4321.00
DellEMC#
```

Configuring the IS-IS Metric Style

All IS-IS links or interfaces are associated with a cost that is used in the shortest path first (SPF) calculations. The possible cost varies depending on the metric style supported.

If you configure narrow, transition, or narrow transition metric style, the cost can be a number between 0 and 63. If you configure wide or wide transition metric style, the cost can be a number between 0 and 16,777,215. Dell EMC Networking OS supports five different metric styles: narrow, wide, transition, narrow transition, and wide transition.

By default, Dell EMC Networking OS generates and receives narrow metric values. Matrixes or costs higher than 63 are not supported. To accept or generate routes with a higher metric, you must change the metric style of the IS-IS process. For example, if you configure the metric as narrow, and a link state PDU (LSP) with wide metrics is received, the route is not installed.

Dell EMC Networking OS supports the following IS-IS metric styles.

Table 48. Metric Styles

Metric Style	Characteristics	Cost Range Supported on IS-IS Interfaces
narrow	Sends and accepts narrow or old TLVs (Type, Length, Value).	0 to 63
wide	Sends and accepts wide or new TLVs.	0 to 16777215
transition	Sends both wide (new) and narrow (old) TLVs.	0 to 63
narrow transition	Sends narrow (old) TLVs and accepts both narrow (old) and wide (new) TLVs.	0 to 63
wide transition	Sends wide (new) TLVs and accepts both narrow (old) and wide (new) TLVs.	0 to 16777215

To change the IS-IS metric style of the IS-IS process, use the following command.

- Set the metric style for the IS-IS process.

  ROUTER ISIS mode

  ```
  metric-style {narrow [transition] | transition | wide [transition]} [level-1 | level-2]
  ```

  The default is **narrow**.

  The default is Level 1 and Level 2 (**level-1–2**)

To view which metric types are generated and received, use the `show isis protocol` command in EXEC Privilege mode. The IS-IS matrixes settings are in bold.

Example of Viewing IS-IS Metric Types

```
DellEMC#show isis protocol
IS-IS Router: <Null Tag>
  System Id: EEEE.EEEE.EEEE IS-Type: level-1-2
  Manual area address(es):
    47.0004.004d.0001
  Routing for area address(es):
    21.2223.2425.2627.2829.3031.3233
    47.0004.004d.0001
  Interfaces supported by IS-IS:
    Vlan 2
    TenGigabitEthernet 4/22
    Loopback 0
  Redistributing:
  Distance: 115
  Generate narrow metrics: level-1-2
  Accept narrow metrics:   level-1-2
  Generate wide metrics:   none
  Accept wide metrics:     none
DellEMC#
```

Configuring the IS-IS Cost

When you change from one IS-IS metric style to another, the IS-IS metric value could be affected. For each interface with IS-IS enabled, you can assign a cost or metric that is used in the link state calculation.

To change the metric or cost of the interface, use the following commands.

- Assign an IS-IS metric.

  INTERFACE mode

  ```
  isis metric default-metric [level-1 | level-2]
  ```

  - *default-metric*: the range is from 0 to 63 if the metric-style is narrow, narrow-transition, or transition.

  The range is from 0 to 16777215 if the metric style is wide or wide transition.

- Assign a metric for an IPv6 link or interface.

  INTERFACE mode

`isis ipv6 metric default-metric [level-1 | level-2]`

- *default-metric*: the range is from 0 to 63 for narrow and transition metric styles. The range is from 0 to 16777215 for wide metric styles.

The default is **10**.

The default level is **level-1**.

For more information about this command, refer to Configuring the IS-IS Metric Style.

The following table describes the correct value range for the `isis metric` command.

Metric Sytle	Correct Value Range
wide	0 to 16777215
narrow	0 to 63
wide transition	0 to 16777215
narrow transition	0 to 63
transition	0 to 63

To view the interface's current metric, use the `show config` command in INTERFACE mode or the `show isis interface` command in EXEC Privilege mode.

Configuring the Distance of a Route

To configure the distance for a route, use the following command.

- Configure the distance for a route.
  ROUTER ISIS mode
  `distance`

Changing the IS-Type

To change the IS-type, use the following commands.

You can configure the system to act as a Level 1 router, a Level 1-2 router, or a Level 2 router.

To change the IS-type for the router, use the following commands.

- Configure IS-IS operating level for a router.
  ROUTER ISIS mode
  `is-type {level-1 | level-1-2 | level-2-only}`
  Default is **level-1-2**.
- Change the IS-type for the IS-IS process.
  ROUTER ISIS mode
  `is-type {level-1 | level-1-2 | level-2}`

To view which IS-type is configured, use the `show isis protocol` command in EXEC Privilege mode. The `show config` command in ROUTER ISIS mode displays only non-default information. If you do not change the IS-type, the default value (**level-1-2**) is not displayed.

The default is Level 1-2 router. When the IS-type is Level 1-2, the software maintains two Link State databases, one for each level. To view the Link State databases, use the `show isis database` command.

```
DellEMC#show isis database
IS-IS Level-1 Link State Database
LSPID           LSP Seq Num  LSP Checksum LSP Holdtime ATT/P/OL
B233.00-00      0x00000003   0x07BF       1088         0/0/0
eljefe.00-00 *  0x00000009   0xF76A       1126         0/0/0
eljefe.01-00 *  0x00000001   0x68DF       1122         0/0/0
eljefe.02-00 *  0x00000001   0x2E7F       1113         0/0/0
Force10.00-00   0x00000002   0xD1A7       1102         0/0/0
IS-IS Level-2 Link State Database
```

```
LSPID           LSP Seq Num  LSP Checksum LSP Holdtime ATT/P/OL
B233.00-00      0x00000006   0xC38A       1124         0/0/0
eljefe.00-00 *  0x0000000D   0x51C6       1129         0/0/0
eljefe.01-00 *  0x00000001   0x68DF       1122         0/0/0
eljefe.02-00 *  0x00000001   0x2E7F       1113         0/0/0
Force10.00-00   0x00000004   0xCDA9       1107         0/0/0

DellEMC#
```

Controlling Routing Updates

To control the source of IS-IS route information, use the following command.

- Disable a specific interface from sending or receiving IS-IS routing information.
  ROUTER ISIS mode
  `passive-interface interface`
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

Distribute Routes

Another method of controlling routing information is to filter the information through a prefix list.

Prefix lists are applied to incoming or outgoing routes and routes must meet the conditions of the prefix lists or Dell EMC Networking OS does not install the route in the routing table. The prefix lists are globally applied on all interfaces running IS-IS.

Configure the prefix list in PREFIX LIST mode prior to assigning it to the IS-IS process. For configuration information on prefix lists, refer to Access Control Lists (ACLs).

Applying IPv4 Routes

To apply prefix lists to incoming or outgoing IPv4 routes, use the following commands.

NOTE: These commands apply to IPv4 IS-IS only. To apply prefix lists to IPv6 routes, use ADDRESS-FAMILY IPV6 mode, shown later.

- Apply a configured prefix list to all incoming IPv4 IS-IS routes.
  ROUTER ISIS mode
  `distribute-list prefix-list-name in [interface]`
  Enter the type of interface and the interface information:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.
- Apply a configured prefix list to all outgoing IPv4 IS-IS routes.
  ROUTER ISIS mode
  `distribute-list prefix-list-name out [bgp as-number | connected | ospf process-id | rip | static]`
  You can configure one of the optional parameters:
  - `connected`: for directly connected routes.
  - `ospf process-id`: for OSPF routes only.
  - `rip`: for RIP routes only.

- `static`: for user-configured routes.
- `bgp`: for BGP routes only.

- Deny RTM download for pre-existing redistributed IPv4 routes.
  ROUTER ISIS mode
  ```
  distribute-list redistributed-override in
  ```

Applying IPv6 Routes

To apply prefix lists to incoming or outgoing IPv6 routes, use the following commands.

NOTE: These commands apply to IPv6 IS-IS only. To apply prefix lists to IPv4 routes, use ROUTER ISIS mode, previously shown.

- Apply a configured prefix list to all incoming IPv6 IS-IS routes.
  ROUTER ISIS-AF IPV6 mode
  ```
  distribute-list prefix-list-name in [interface]
  ```
  Enter the type of interface and the interface information:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.
- Apply a configured prefix list to all outgoing IPv6 IS-IS routes.
  ROUTER ISIS-AF IPV6 mode
  ```
  distribute-list prefix-list-name out [bgp as-number | connected | ospf process-id | rip | static]
  ```
  You can configure one of the optional parameters:
  - `connected`: for directly connected routes.
  - `ospf process-id`: for OSPF routes only.
  - `rip`: for RIP routes only.
  - `static`: for user-configured routes.
  - `bgp`: for BGP routes only.
- Deny RTM download for pre-existing redistributed IPv6 routes.
  ROUTER ISIS-AF IPV6 mode
  ```
  distribute-list redistributed-override in
  ```

Redistributing IPv4 Routes

In addition to filtering routes, you can add routes from other routing instances or protocols to the IS-IS process. With the `redistribute` command syntax, you can include BGP, OSPF, RIP, static, or directly connected routes in the IS-IS process.

NOTE: Do not route iBGP routes to IS-IS unless there are route-maps associated with the IS-IS redistribution.

To add routes from other routing instances or protocols, use the following commands.

NOTE: These commands apply to IPv4 IS-IS only. To apply prefix lists to IPv6 routes, use ADDRESS-FAMILY IPV6 mode, shown later.

- Include BGP, directly connected, RIP, or user-configured (static) routes in IS-IS.
  ROUTER ISIS mode
  ```
  redistribute {bgp as-number | connected | rip | static} [level-1 level-1-2 | level-2] [metric metric-value] [metric-type {external | internal}] [route-map map-name]
  ```

Configure the following parameters:

- `level-1`, `level-1-2`, or `level-2`: assign all redistributed routes to a level. The default is **level-2**.
- *`metric-value`* the range is from 0 to 16777215. The default is **0**.
- *`metric-type`*: choose either external or internal. The default is **internal**.
- *`map-name`*: enter the name of a configured route map.

- Include specific OSPF routes in IS-IS.

  ROUTER ISIS mode

  `redistribute ospf` *`process-id`* `[level-1| level-1-2 | level-2] [metric` *`value`*`] [match external {1 | 2} | match internal] [metric-type {external | internal}] [route-map` *`map-name`*`]`

  Configure the following parameters:

  - *`process-id`* the range is from 1 to 65535.
  - `level-1`, `level-1-2`, or `level-2`: assign all redistributed routes to a level. The default is **level-2**.
  - `metric` *`value`* the range is from 0 to 16777215. The default is **0**.
  - `match external` the range is from 1 or 2.
  - `match internal`
  - `metric-type`: external or internal.
  - *`map-name`*: enter the name of a configured route map.

Redistributing IPv6 Routes

To add routes from other routing instances or protocols, use the following commands.

NOTE: These commands apply to IPv6 IS-IS only. To apply prefix lists to IPv4 routes, use the ROUTER ISIS mode previously shown.

- Include BGP, directly connected, RIP, or user-configured (static) routes in IS-IS.

  ROUTER ISIS mode

  `redistribute {bgp` *`as-number`* `| connected | rip | static} [level-1 level-1-2 | level-2] [metric` *`metric-value`*`] [metric-type {external | internal}] [route-map` *`map-name`*`]`

  Configure the following parameters:

  - `level-1`, `level-1-2`, or `level-2`: assign all redistributed routes to a level. The default is **level-2**.
  - *`metric-value`*: the range is from 0 to 16777215. The default is **0**.
  - *`metric-type`*: choose either external or internal. The default is **internal**.
  - *`map-name`*: enter the name of a configured route map.

- Include specific OSPF routes in IS-IS.ROUTER ISIS mode

  `redistribute ospf` *`process-id`* `[level-1| level-1-2 | level-2] [metric` *`value`*`] [match external {1 | 2} | match internal] [metric-type {external | internal}] [route-map` *`map-name`*`]`

  Configure the following parameters:

  - *`process-id`*: the range is from 1 to 65535.
  - `level-1`, `level-1-2`, or `level-2`: assign all redistributed routes to a level. The default is **level-2**.
  - `metric` *`value`*: the range is from 0 to 16777215. The default is **0**.
  - `metric` *`value`*: the range is from 0 to 16777215. The default is **0**.
  - `match external`: the range is 1 or 2.
  - `match internal`
  - `metric-type`: external or internal.
  - *`map-name`*: name of a configured route map.

To view the IS-IS configuration globally (including both IPv4 and IPv6 settings), use the `show running-config isis` command in EXEC Privilege mode. To view the current IPv4 IS-IS configuration, use the `show config` command in ROUTER ISIS mode. To view the current IPv6 IS-IS configuration, use the `show config` command in ROUTER ISIS-ADDRESS FAMILY IPV6 mode.

Configuring Authentication Passwords

You can assign an authentication password for routers in Level 1 and for routers in Level 2.

Because Level 1 and Level 2 routers do not communicate with each other, you can assign different passwords for Level 1 routers and for Level 2 routers. However, if you want the routers in the level to communicate with each other, configure them with the same password.

To configure a simple text password, use the following commands.

- Configure authentication password for an area.

  ROUTER ISIS mode

  `area-password [hmac-md5] password`

  The Dell OS supports HMAC-MD5 authentication.

  This password is inserted in Level 1 LSPs, Complete SNPs, and Partial SNPs.
- Set the authentication password for a routing domain.

  ROUTER ISIS mode

  `domain-password [encryption-type | hmac-md5] password`

  The Dell OS supports both DES and HMAC-MD5 authentication methods.

  This password is inserted in Level 2 LSPs, Complete SNPs, and Partial SNPs.

To view the passwords, use the `show config` command in ROUTER ISIS mode or the `show running-config isis` command in EXEC Privilege mode.

To remove a password, use either the `no area-password` or `no domain-password` commands in ROUTER ISIS mode.

Setting the Overload Bit

Another use for the overload bit is to prevent other routers from using this router as an intermediate hop in their shortest path first (SPF) calculations. For example, if the IS-IS routing database is out of memory and cannot accept new LSPs, Dell EMC Networking OS sets the overload bit and IS-IS traffic continues to transit the system.

To set or remove the overload bit manually, use the following commands.

- Set the overload bit in LSPs.

  ROUTER ISIS mode

  `set-overload-bit`

  This setting prevents other routers from using it as an intermediate hop in their shortest path first (SPF) calculations.
- Remove the overload bit.

  ROUTER ISIS mode

  `no set-overload-bit`

When the bit is set, a 1 is placed in the *OL* column in the `show isis database` command output. The overload bit is set in both the Level-1 and Level-2 database because the IS type for the router is Level-1-2.

```
DellEMC#show isis database
IS-IS Level-1 Link State Database
LSPID            LSP Seq Num LSP Checksum LSP Holdtime ATT/P/OL
B233.00-00     0x00000003   0x07BF       1074          0/0/0
eljefe.00-00 * 0x0000000A   0xF963       1196          0/0/1
eljefe.01-00 * 0x00000001   0x68DF       1108          0/0/0
eljefe.02-00 * 0x00000001   0x2E7F       1099          0/0/0
Force10.00-00  0x00000002   0xD1A7       1088          0/0/0
IS-IS Level-2 Link State Database
LSPID            LSP Seq Num LSP Checksum LSP Holdtime ATT/P/OL
B233.00-00     0x00000006   0xC38A       1110          0/0/0
eljefe.00-00 * 0x0000000E   0x53BF       1196          0/0/1
eljefe.01-00 * 0x00000001   0x68DF       1108          0/0/0
eljefe.02-00 * 0x00000001   0x2E7F       1099          0/0/0
Force10.00-00  0x00000004   0xCDA9       1093          0/0/0
DellEMC#
```

Debugging IS-IS

To debug IS-IS processes, use the following commands.

- View all IS-IS information.
  EXEC Privilege mode
  `debug isis`
- View information on all adjacency-related activity (for example, hello packets that are sent and received).
  EXEC Privilege mode
  `debug isis adj-packets [interface]`
  To view specific information, enter the following optional parameter:
  - *interface*: Enter the type of interface and slot/port information to view IS-IS information on that interface only.
- View information about IS-IS local update packets.
  EXEC Privilege mode
  `debug isis local-updates [interface]`
  To view specific information, enter the following optional parameter:
  - *interface*: Enter the type of interface and slot/port information to view IS-IS information on that interface only.
- View IS-IS SNP packets, include CSNPs and PSNPs.
  EXEC Privilege mode
  `debug isis snp-packets [interface]`
  To view specific information, enter the following optional parameter:
  - *interface*: Enter the type of interface and slot/port information to view IS-IS information on that interface only.
- View the events that triggered IS-IS shortest path first (SPF) events for debugging purposes.
  EXEC Privilege mode
  `debug isis spf-triggers`
- View sent and received LSPs.
  EXEC Privilege mode
  `debug isis update-packets [interface]`
  To view specific information, enter the following optional parameter:
  - *interface*: Enter the type of interface and slot/port information to view IS-IS information on that interface only.

Dell EMC Networking OS displays debug messages on the console. To view which debugging commands are enabled, use the `show debugging` command in EXEC Privilege mode.

To disable a specific debug command, enter the keyword `no` then the `debug` command. For example, to disable debugging of IS-IS updates, use the `no debug isis updates-packets` command.

To disable all IS-IS debugging, use the `no debug isis` command.

To disable all debugging, use the `undebug all` command.

IS-IS Metric Styles

The following sections provide additional information about the IS-IS metric styles.

- Configuring the IS-IS Metric Style
- Configure Metric Values

Dell EMC Networking OS supports the following IS-IS metric styles:

- narrow (supports only type, length, and value [TLV] up to 63)
- wide (supports TLV up to 16777215)
- transition (supports both narrow and wide and uses a TLV up to 63)
- narrow transition (accepts both narrow and wide and sends only narrow or old-style TLV)
- wide transition (accepts both narrow and wide and sends only wide or new-style TLV)

Configure Metric Values

For any level (Level-1, Level-2, or Level-1-2), the value range possible in the `isis metric` command in INTERFACE mode changes depending on the metric style.

The following describes the correct value range for the `isis metric` command.

Metric Style	Correct Value Range for the isis metric Command
wide	0 to 16777215
narrow	0 to 63
wide transition	0 to 16777215
narrow transition	0 to 63
transition	0 to 63

Maximum Values in the Routing Table

IS-IS metric styles support different cost ranges for the route. The cost range for the narrow metric style is 0 to 1023, while all other metric styles support a range of 0 to 0xFE000000.

Change the IS-IS Metric Style in One Level Only

By default, the IS-IS metric style is narrow. When you change from one IS-IS metric style to another, the IS-IS metric value (configured with the `isis metric` command) could be affected.

In the following scenarios, the IS-type is either Level-1 or Level-2 or Level-1-2 and the metric style changes.

Table 49. Metric Value When the Metric Style Changes

Beginning Metric Style	Final Metric Style	Resulting IS-IS Metric Value
wide	narrow	default value (10) if the original value is greater than 63. A message is sent to the console.
wide	transition	truncated value (the truncated value appears in the LSP only). The original `isis metric` value is displayed in the `show config` and `show running-config` commands and is used if you change back to transition metric style. **NOTE:** A truncated value is a value that is higher than 63, but set back to 63 because the higher value is not supported.
wide	narrow transition	default value (10) if the original value is greater than 63. A message is sent to the console.
wide	wide transition	original value
narrow	wide	original value
narrow	transition	original value
narrow	narrow transition	original value
narrow	wide transition	original value
transition	wide	original value

Table 49. Metric Value When the Metric Style Changes (continued)

Beginning Metric Style	Final Metric Style	Resulting IS-IS Metric Value
transition	narrow	original value
transition	narrow	original value
transition	wide transition	original value
narrow transition	wide	original value
narrow transition	narrow	original value
narrow transition	wide transition	original value
narrow transition	transition	original value
wide transition	wide	original value
wide transition	narrow	default value (10) if the original value is greater than 63. A message is sent to the console.
wide transition	narrow transition	default value (10) if the original value is greater than 63. A message is sent to the console.
wide transition	transition	truncated value (the truncated value appears in the LSP only). The original `isis metric` value is displayed in the `show config` and `show running-config` commands and is used if you change back to transition metric style.

Moving to transition and then to another metric style produces different results.

Table 50. Metric Value when the Metric Style Changes Multiple Times

Beginning Metric Style	Next Metric Style	Resulting Metric Value	Next Metric Style	Final Metric Value
wide	transition	truncated value	wide	original value is recovered
wide transition	transition	truncated value	wide transition	original value is recovered
wide	transition	truncated value	narrow	default value (10). A message is sent to the logging buffer
wide transition	transition	truncated value	narrow transition	default value (10). A message is sent to the logging buffer

Leaks from One Level to Another

In the following scenarios, each IS-IS level is configured with a different metric style.

Table 51. Metric Value with Different Levels Configured with Different Metric Styles

Level-1 Metric Style	Level-2 Metric Style	Resulting Metric Value
narrow	wide	original value
narrow	wide transition	original value
narrow	narrow transition	original value

Table 51. Metric Value with Different Levels Configured with Different Metric Styles (continued)

Level-1 Metric Style	Level-2 Metric Style	Resulting Metric Value
narrow	transition	original value
wide	narrow	truncated value
wide	narrow transition	truncated value
wide	wide transition	original value
wide	transition	truncated value
narrow transition	wide	original value
narrow transition	narrow	original value
narrow transition	wide transition	original value
narrow transition	transition	original value
transition	wide	original value
transition	narrow	original value
transition	wide transition	original value
transition	narrow transition	original value
wide transition	wide	original value
wide transition	narrow	truncated value
wide transition	narrow transition	truncated value
wide transition	transition	truncated value

Sample Configurations

The following configurations are examples for enabling IPv6 IS-IS. These examples are not comprehensive directions. They are intended to give you some guidance with typical configurations.

NOTE: Only one IS-IS process can run on the router, even if both IPv4 and IPv6 routing is being used.

You can copy and paste from these examples to your CLI. To support your own IP addresses, interfaces, names, and so on, be sure that you make the necessary changes.

NOTE: Whenever you make IS-IS configuration changes, clear the IS-IS process (re-started) using the `clear isis` command. The `clear isis` command must include the tag for the ISIS process. The following example shows the response from the router:

```
DellEMC#clear isis *
% ISIS not enabled.
DellEMC#clear isis 9999 *
```

You can configure IPv6 IS-IS routes in one of the following three different methods:

- **Congruent Topology** — You *must* configure both IPv4 and IPv6 addresses on the interface. Enable the `ip router isis` and `ipv6 router isis` commands on the interface. Enable the `wide-metrics` parameter in router isis configuration mode.
- **Multi-topology** — You *must* configure the IPv6 address. Configuring the IPv4 address is optional. You *must* enable the `ipv6 router isis` command on the interface. If you configure IPv4, also enable the `router isis` command. In router isis configuration mode, enable `multi-topology` under address-family ipv6 unicast.
- **Multi-topology Transition** — You *must* configure the IPv6 address. Configuring the IPv4 address is optional. You *must* enable the `ipv6 router isis` command on the interface. If you configure IPv4, also enable the `ip router isis` command. In router isis configuration mode, enable `multi-topology transition` under address-family ipv6 unicast.

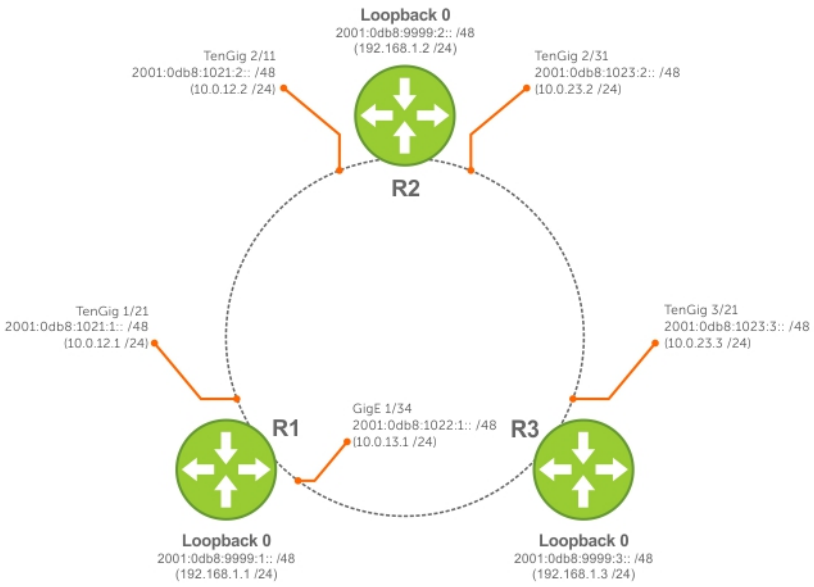


Figure 63. IPv6 IS-IS Sample Topography

The following is a sample configuration for enabling IPv6 IS-IS.

IS-IS Sample Configuration — Congruent Topology

```
DellEMC(conf-if-te-3/17)#show config
!
interface TenGigabitEthernet 3/17
ip address 24.3.1.1/24
ipv6 address 24:3::1/76
ip router isis
ipv6 router isis
no shutdown
DellEMC(conf-if-te-3/17)#

DellEMC(conf-router_isis)#show config
!
router isis
metric-style wide level-1
metric-style wide level-2
net 34.0000.0000.AAAA.00
DellEMC(conf-router_isis)#
```

IS-IS Sample Configuration — Multi-topology

```
DellEMC(conf-if-te-3/17)#show config
!
interface TenGigabitEthernet 3/17
ipv6 address 24:3::1/76
ipv6 router isis
no shutdown
DellEMC(conf-if-te-3/17)#

DellEMC(conf-router_isis)#show config
!
router isis
net 34.0000.0000.AAAA.00
!
address-family ipv6 unicast
multi-topology
exit-address-family
DellEMC(conf-router_isis)#
```

IS-IS Sample Configuration — Multi-topology Transition

```
DellEMC(conf-if-te-3/17)#show config
!
interface TenGigabitEthernet 3/17
ipv6 address 24:3::1/76
ipv6 router isis
no shutdown
DellEMC(conf-if-te-3/17)#

DellEMC(conf-router_isis)#show config
!
router isis
net 34.0000.0000.AAAA.00
!
address-family ipv6 unicast
multi-topology transition
exit-address-family
DellEMC(conf-router_isis)#
```

27

Link Aggregation Control Protocol (LACP)

A link aggregation group (LAG), referred to as a *port channel* by the Dell EMC Networking OS, can provide both load-sharing and port redundancy across line cards. You can enable LAGs as static or dynamic.

Topics:

- Introduction to Dynamic LAGs and LACP
- LACP Configuration Tasks
- Shared LAG State Tracking
- LACP Basic Configuration Example

Introduction to Dynamic LAGs and LACP

A link aggregation group (LAG), referred to as a *port channel* by Dell EMC Networking OS, can provide both load-sharing and port redundancy across line cards. You can enable LAGs as static or dynamic.

The benefits and constraints are basically the same, as described in *Port Channel Interfaces* in the Interfaces chapter.

The unique benefit of a dynamic LAG is that its ports can toggle between participating in the LAG or acting as dedicated ports, whereas ports in a static LAG must be removed from the LAG in order to act alone.

The Dell EMC Networking OS uses LACP to create dynamic LAGs. LACP provides a standardized means of exchanging information between two systems (also called Partner Systems) and automatically establishes the LAG between the systems. LACP permits the exchange of messages on a link to allow their LACP instances to:

- Reach an agreement on the identity of the LAG to which the link belongs.
- Move the link to that LAG.
- Enable the transmission and reception functions in an orderly manner.

The Dell EMC Networking OS implementation of LACP is based on the standards specified in the IEEE 802.3: “Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications.”

LACP functions by constantly exchanging custom MAC protocol data units (PDUs) across local area network (LAN) Ethernet links. The protocol packets are only exchanged between ports that are configured as LACP capable.

Important Points to Remember

- LACP allows you to add members to a port channel (LAG) as long as it has no static members. Conversely, if the LAG already contains a statically defined member (the `channel-member` command), the `port-channel mode` command is not permitted.
- A static LAG cannot be created if a dynamic LAG using the selected number exists.
- No dual membership in static and dynamic LAGs:
  - If a physical interface is a part of a static LAG, the `port-channel-protocol lacp` command is rejected on that interface.
  - If a physical interface is a part of a dynamic LAG, it cannot be added as a member of a static LAG. The `channel-member tengigabitethernet` command is rejected in the static LAG interface for that physical interface.
- A dynamic LAG can be created with any type of configuration.
- There is a difference between the `shutdown` and `no interface port-channel` commands:
  - The `shutdown` command on LAG “xyz” disables the LAG and retains the user commands. However, the system does not allow the channel number “xyz” to be statically created.
  - The `no interface port-channel` *`channel-number`* command deletes the specified LAG, including a dynamically created LAG. This command removes all LACP-specific commands on the member interfaces. The interfaces are restored to a state that is ready to be configured.

    NOTE: There is no configuration on the interface because that condition is required for an interface to be part of a LAG.

- You can configure link dampening on individual members of a LAG.

LACP Modes

Dell EMC Networking OS provides three modes for configuration of LACP — Off, Active, and Passive.

- **Off** — In this state, an interface is not capable of being part of a dynamic LAG. LACP does not run on any port that is configured to be in this state.
- **Active** — In this state, the interface is said to be in the "active negotiating state." LACP runs on any link that is configured to be in this state. A port in Active state also automatically initiates negotiations with other ports by initiating LACP packets.
- **Passive** — In this state, the interface is not in an active negotiating state, but LACP runs on the link. A port in Passive state also responds to negotiation requests (from ports in Active state). Ports in Passive state respond to LACP packets.

Dell EMC Networking OS supports LAGs in the following cases:

- A port in Active state can set up a port channel (LAG) with another port in Active state.
- A port in Active state can set up a LAG with another port in Passive state.

A port in Passive state cannot set up a LAG with another port in Passive state.

Configuring LACP Commands

If you configure aggregated ports with compatible LACP modes (Off, Active, Passive), LACP can automatically link them, as defined in IEEE 802.3, Section 43.

To configure LACP, use the following commands.

- Configure the system priority.

  CONFIGURATION mode

  `[no] lacp system-priority` *`priority-value`*

  The range is from 1 to 65535 (the higher the number, the lower the priority).

  The default is **32768**.
- Enable or disable LACP on any LAN port.

  INTERFACE mode

  `[no] port-channel-protocol lacp`

  The default is **LACP disabled**.

  This command creates context.
- Configure LACP mode.

  LACP mode

  `[no] port-channel` *`number`* `mode [active | passive | off]`

  - *`number`*: cannot statically contain any links.

  The default is **LACP active**.
- Configure port priority.

  LACP mode

  `[no] lacp port-priority` *`priority-value`*

  The range is from 1 to 65535 (the higher the number, the lower the priority).

  The default is **32768**.

LACP Configuration Tasks

The following configuration tasks apply to LACP.

- Creating a LAG
- Configuring the LAG Interfaces as Dynamic
- Setting the LACP Long Timeout
- Monitoring and Debugging LACP
- Configuring Shared LAG State Tracking

Creating a LAG

To create a dynamic port channel (LAG), use the following command. First you define the LAG and then the LAG interfaces.

- Create a dynamic port channel (LAG).
  CONFIGURATION mode
  `interface port-channel`
- Create a dynamic port channel (LAG).
  CONFIGURATION mode
  `switchport`

```
DellEMC(conf)#interface port-channel 32
DellEMC(conf-if-po-32)#no shutdown
DellEMC(conf-if-po-32)#switchport
```

The LAG is in the default VLAN. To place the LAG into a non-default VLAN, use the `tagged` command on the LAG.

```
DellEMC(conf)#interface vlan 10
DellEMC(conf-if-vl-10)#tagged port-channel 32
```

Configuring the LAG Interfaces as Dynamic

After creating a LAG, configure the dynamic LAG interfaces.

To configure the dynamic LAG interfaces, use the following command.

- Configure the dynamic LAG interfaces.
  CONFIGURATION mode
  `port-channel-protocol lacp`

```
DellEMC(conf)#interface TenGigabitethernet 3/15
DellEMC(conf-if-te-3/15)#no shutdown
DellEMC(conf-if-te-3/15)#port-channel-protocol lacp
DellEMC(conf-if-te-3/15-lacp)#port-channel 32 mode active
...
DellEMC(conf)#interface TenGigabitethernet 3/16
DellEMC(conf-if-te-3/16)#no shutdown
DellEMC(conf-if-te-3/16)#port-channel-protocol lacp
DellEMC(conf-if-te-3/16-lacp)#port-channel 32 mode active
...
DellEMC(conf)#interface TenGigabitethernet 4/15
DellEMC(conf-if-te-4/15)#no shutdown
DellEMC(conf-if-te-4/15)#port-channel-protocol lacp
DellEMC(conf-if-te-4/15-lacp)#port-channel 32 mode active
...
DellEMC(conf)#interface TenGigabitethernet 4/16
DellEMC(conf-if-te-4/16)#no shutdown
DellEMC(conf-if-te-4/16)#port-channel-protocol lacp
DellEMC(conf-if-te-4/16-lacp)#port-channel 32 mode active
```

The port-channel 32 mode active command shown here may be successfully issued as long as there is no existing static channel-member configuration in LAG 32.

Setting the LACP Long Timeout

PDUs are exchanged between port channel (LAG) interfaces to maintain LACP sessions.

PDUs are transmitted at either a slow or fast transmission rate, depending upon the LACP timeout value. The timeout value is the amount of time that a LAG interface waits for a PDU from the remote system before bringing the LACP session down. The default timeout value is **1 second**. You can configure the default timeout value to be **30 seconds**. Invoking the longer timeout might prevent the LAG from flapping if the remote system is up but temporarily unable to transmit PDUs due to a system interruption.

NOTE: The 30-second timeout is available for dynamic LAG interfaces only. You can enter the `lacp long-timeout` command for static LAGs, but it has no effect.

To configure LACP long timeout, use the following command.

- Set the LACP timeout value to 30 seconds.

  CONFIG-INT-PO mode

  `lacp long-timeout`

```
DellEMC(conf)# interface port-channel 32
DellEMC(conf-if-po-32)#no shutdown
DellEMC(conf-if-po-32)#switchport
DellEMC(conf-if-po-32)#lacp long-timeout
DellEMC(conf-if-po-32)#end
DellEMC# show lacp 32
Port-channel 32 admin up, oper up, mode lacp
Actor System ID: Priority 32768, Address 0001.e800.a12b
Partner System ID: Priority 32768, Address 0001.e801.45a5
Actor Admin Key 1, Oper Key 1, Partner Oper Key 1
LACP LAG 1 is an aggregatable link
A - Active LACP, B - Passive LACP, C - Short Timeout, D - Long Timeout
E - Aggregatable Link, F - Individual Link, G - IN_SYNC, H - OUT_OF_SYNC
I - Collection enabled, J - Collection disabled, K - Distribution enabled L -
Distribution disabled,
M - Partner Defaulted, N - Partner Non-defaulted, O - Receiver is in expired
state,
P - Receiver is not in expired state
Port Te 3/6 is enabled, LACP is enabled and mode is lacp
Actor Admin: State ADEHJLMP Key 1 Priority 128
```

To view the PDU exchanges and the timeout value, use the `debug lacp` command. For more information, refer to Monitoring and Debugging LACP.

Monitoring and Debugging LACP

The system log (syslog) records faulty LACP actions.

To debug LACP, use the following command.

- Debug LACP, including configuration and events.

  EXEC mode

  `[no] debug lacp [config | events | pdu [in | out | [interface [in | out]]]]`

Shared LAG State Tracking

Shared LAG state tracking provides the flexibility to bring down a port channel (LAG) based on the operational state of another LAG.

At any time, only two LAGs can be a part of a group such that the fate (status) of one LAG depends on the other LAG.

As shown in the following illustration, the line-rate traffic from R1 destined for R4 follows the lowest-cost route via R2. Traffic is equally distributed between LAGs 1 and 2. If LAG 1 fails, all traffic from R1 to R4 flows across LAG 2 only. This condition over-subscribes the link and packets are dropped.

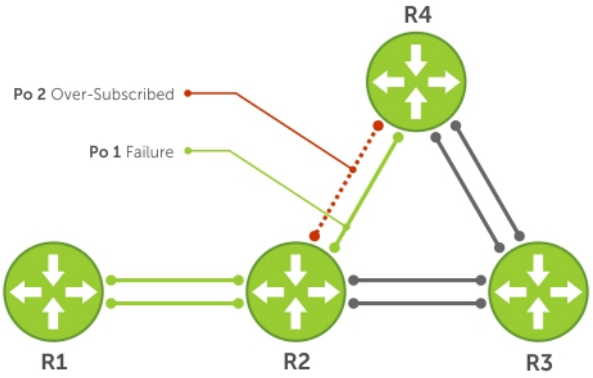


Figure 64. Shared LAG State Tracking

To avoid packet loss, redirect traffic through the next lowest-cost link (R3 to R4). Dell EMC Networking OS has the ability to bring LAG 2 down if LAG 1 fails, so that traffic can be redirected. This redirection is what is meant by shared LAG state tracking. To achieve this functionality, you must group LAG 1 and LAG 2 into a single entity, called a failover group.

Configuring Shared LAG State Tracking

To configure shared LAG state tracking, you configure a failover group.

NOTE: If a LAG interface is part of a redundant pair, you cannot use it as a member of a failover group created for shared LAG state tracking.

1. Enter port-channel failover group mode.
   CONFIGURATION mode
   `port-channel failover-group`
2. Create a failover group and specify the two port-channels that will be members of the group.
   CONFIG-PO-FAILOVER-GRP mode
   `group` *`number`* `port-channel` *`number`* `port-channel` *`number`*

In the following example, LAGs 1 and 2 have been placed into to the same failover group.

Example of LAGs in the Same Failover Group

```
DellEMC#config
DellEMC(conf)#port-channel failover-group
DellEMC(conf-po-failover-grp)#group 1 port-channel 1 port-channel 2
```

To view the failover group configuration, use the `show running-configuration po-failover-group` command.

```
DellEMC#show running-config po-failover-group
!
port-channel failover-group
group 1 port-channel 1 port-channel 2
```

As shown in the following illustration, LAGs 1 and 2 are members of a failover group. LAG 1 fails and LAG 2 is brought down after the failure. This effect is logged by Message 1, in which a console message declares both LAGs down at the same time.

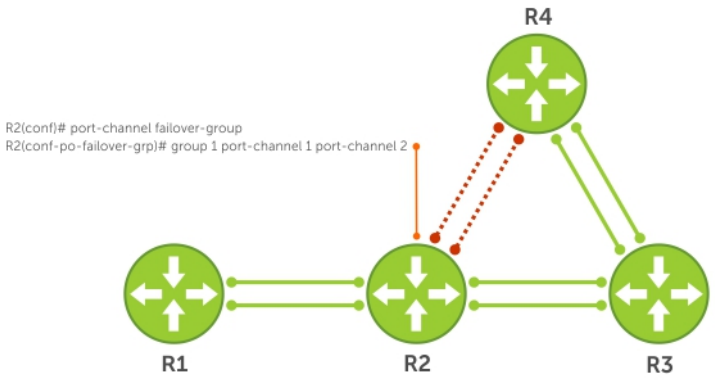


Figure 65. Configuring Shared LAG State Tracking

The following are shared LAG state tracking console messages:

- `2d1h45m: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 1`
- `2d1h45m: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 2`

To view the status of a failover group member, use the `show interface port-channel` command.

```
DellEMC#show interface port-channel 2
Port-channel 2 is up, line protocol is down (Failover-group 1 is down)
Hardware address is 00:01:e8:05:e8:4c, Current address is 00:01:e8:05:e8:4c
Interface index is 1107755010
Minimum number of links to bring Port-channel up is 1
Port-channel is part of failover-group 1
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit
Members in this channel: Te 1/17(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:01:28
Queueing strategy: fifo
```

NOTE: The set of console messages shown above appear only if you configure shared LAG state tracking on that router (you can configure the feature on one or both sides of a link). For example, as previously shown, if you configured shared LAG state tracking on R2 only, no messages appear on R4 regarding the state of LAGs in a failover group.

Important Points about Shared LAG State Tracking

The following is more information about shared LAG state tracking.

- This feature is available for static and dynamic LAGs.
- Only a LAG can be a member of a failover group.
- You can configure shared LAG state tracking on one side of a link or on both sides.
- If a LAG that is part of a failover group is deleted, the failover group is deleted.
- If a LAG moves to the Down state due to this feature, its members may still be in the Up state.

LACP Basic Configuration Example

The screenshots in this section are based on the following example topology. Two routers are named ALPHA and BRAVO, and their hostname prompts reflect those names.

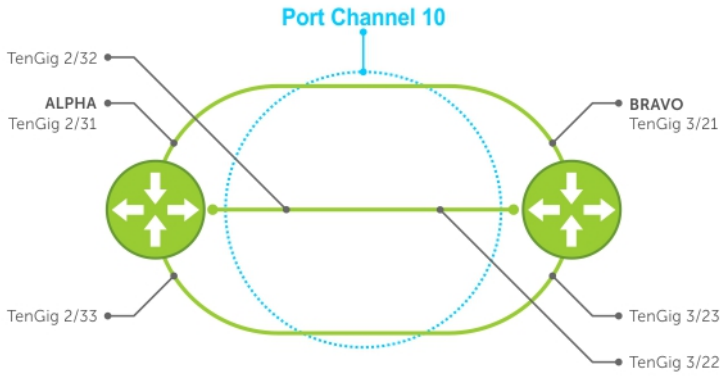


Figure 66. LACP Basic Configuration Example

Configure a LAG on ALPHA

The following example creates a LAG on ALPHA.

Example of Configuring a LAG

```
Alpha(conf)#interface port-channel 10
Alpha(conf-if-po-10)#no ip address
Alpha(conf-if-po-10)#switchport
Alpha(conf-if-po-10)#no shutdown
Alpha(conf-if-po-10)#show config
!
interface Port-channel 10
  no ip address
  switchport
  no shutdown
!
Alpha(conf-if-po-10)#
```

The following example inspects a LAG port configuration on ALPHA.

Example of Viewing a LAG Port Configuration

```
Alpha#sh int TenGigabitEthernet 2/31
TenGigabitEthernet 2/31 is up, line protocol is up
Port is part of Port-channel 10
Hardware is DellEMCEth, address is 00:01:e8:06:95:c0
     Current address is 00:01:e8:06:95:c0
Interface Index is 109101113
Port will not be disabled on partial SFM failure
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit, Mode full duplex, Slave
Flowcontrol rx on tx on
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:02:11
Queueing strategy: fifo
Input statistics:
     132 packets, 163668 bytes
     0 Vlans
     0 64-byte pkts, 12 over 64-byte pkts, 120 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     132 Multicasts, 0 Broadcasts
```

```
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics
     136 packets, 16718 bytes, 0 underruns
     0 64-byte pkts, 15 over 64-byte pkts, 121 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     136 Multicasts, 0 Broadcasts, 0 Unicasts
     0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
     Input 00.00 Mbits/sec,0 packets/sec, 0.00% of line-rate
     Output 00.00 Mbits/sec,0 packets/sec, 0.00% of line-rate
     Time since last interface status change: 00:02:14
```

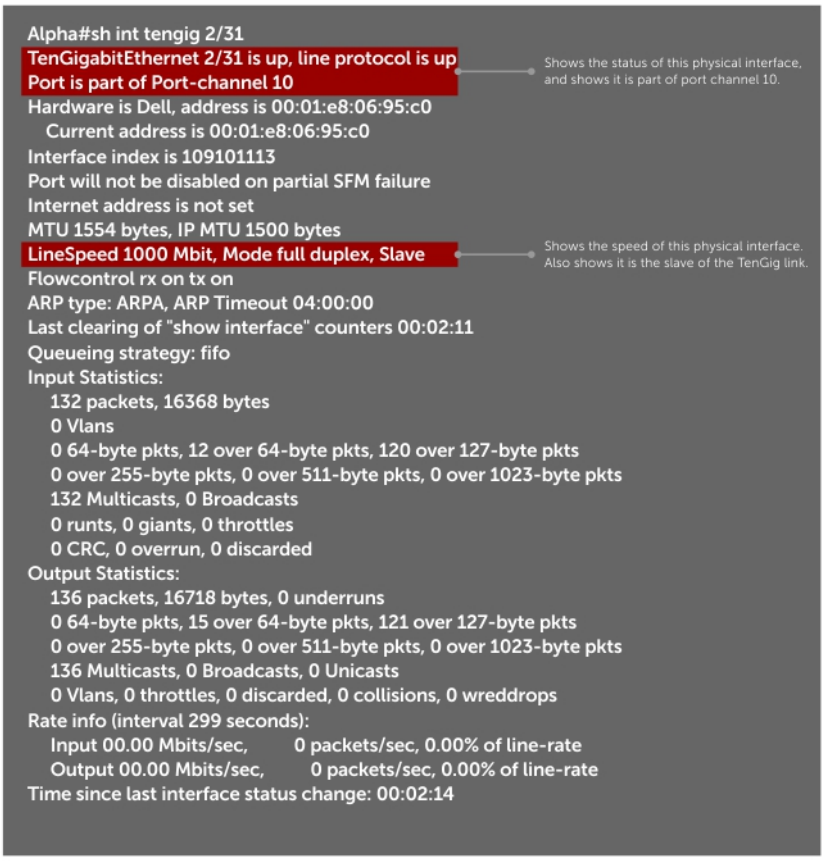


Figure 67. Inspecting the LAG Configuration

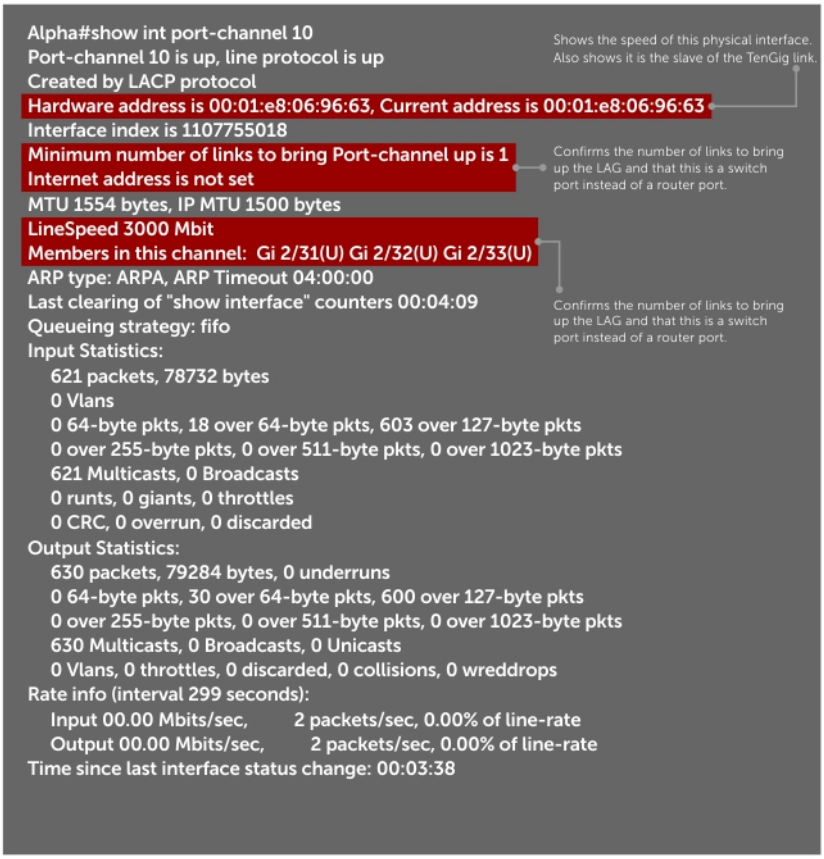


Figure 68. Inspecting Configuration of LAG 10 on ALPHA

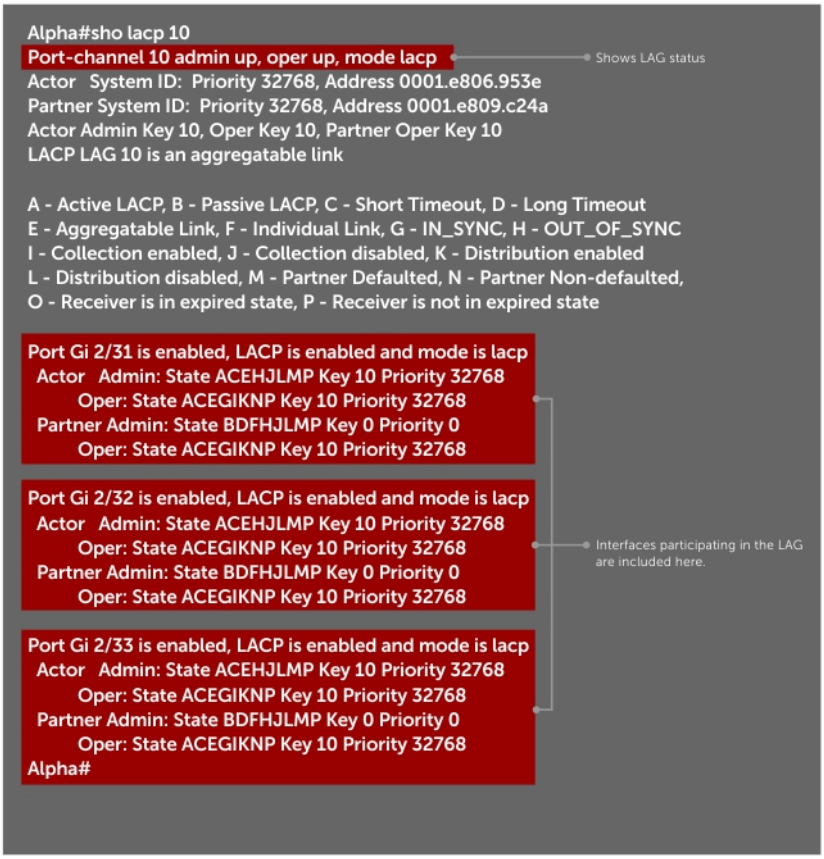


Figure 69. Verifying LAG 10 Status on ALPHA Using the show lacp Command

Summary of the LAG Configuration on Alpha

```
Alpha(conf-if-po-10)#int tengig 2/31
Alpha(conf-if-te-2/31)#no ip address
Alpha(conf-if-te-2/31)#no switchport
Alpha(conf-if-te-2/31)#shutdown
Alpha(conf-if-te-2/31)#port-channel-protocol lacp
Alpha(conf-if-te-2/31-lacp)#port-channel 10 mode active
Alpha(conf-if-te-2/31-lacp)#no shut
Alpha(conf-if-te-2/31)#show config

!
interface GigabitEthernet 2/31
  no ip address
!
  port-channel-protocol LACP
   port-channel 10 mode active
  no shutdown
!
Alpha(conf-if-te-2/31)#

interface Port-channel 10
no ip address
switchport
no shutdown

interface TenGigabitEthernet 2/31
no ip address
```

Summary of the LAG Configuration on Bravo

```
Bravo(conf-if-te-3/21)#int port-channel 10
Bravo(conf-if-po-10)#no ip add
```

```
Bravo(conf-if-po-10)#switch
Bravo(conf-if-po-10)#no shut
Bravo(conf-if-po-10)#show config
!
interface Port-channel 10
  no ip address
  switchport
  no shutdown
!
Bravo(conf-if-po-10)#exit

Bravo(conf)#int tengig 3/21
Bravo(conf)#no ip address
Bravo(conf)#no switchport
Bravo(conf)#shutdown
Bravo(conf-if-te-3/21)#port-channel-protocol lacp
Bravo(conf-if-te-3/21-lacp)#port-channel 10 mode active
Bravo(conf-if-te-3/21-lacp)#no shut
Bravo(conf-if-te-3/21)#end

!
interface TenGigabitEthernet 3/21
  no ip address
!
  port-channel-protocol LACP
   port-channel 10 mode active
  no shutdown
Bravo(conf-if-te-3/21)#end

int port-channel 10
no ip address
switchport
no shutdown
show config

int tengig 3/21
no ip address
```

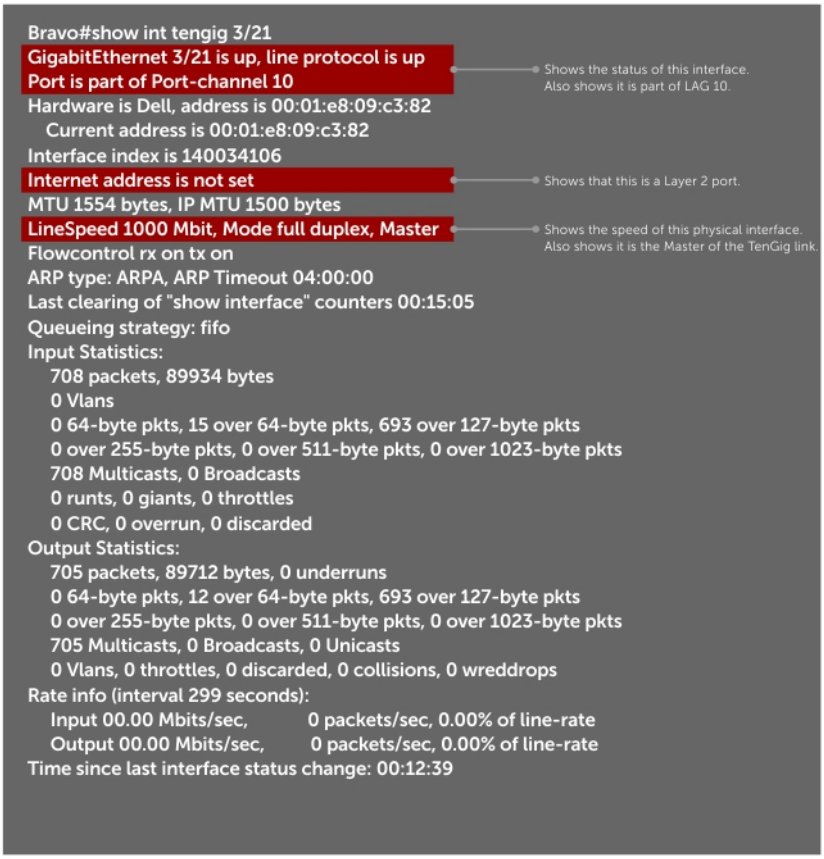


Figure 70. Inspecting a LAG Port on BRAVO Using the show interface Command

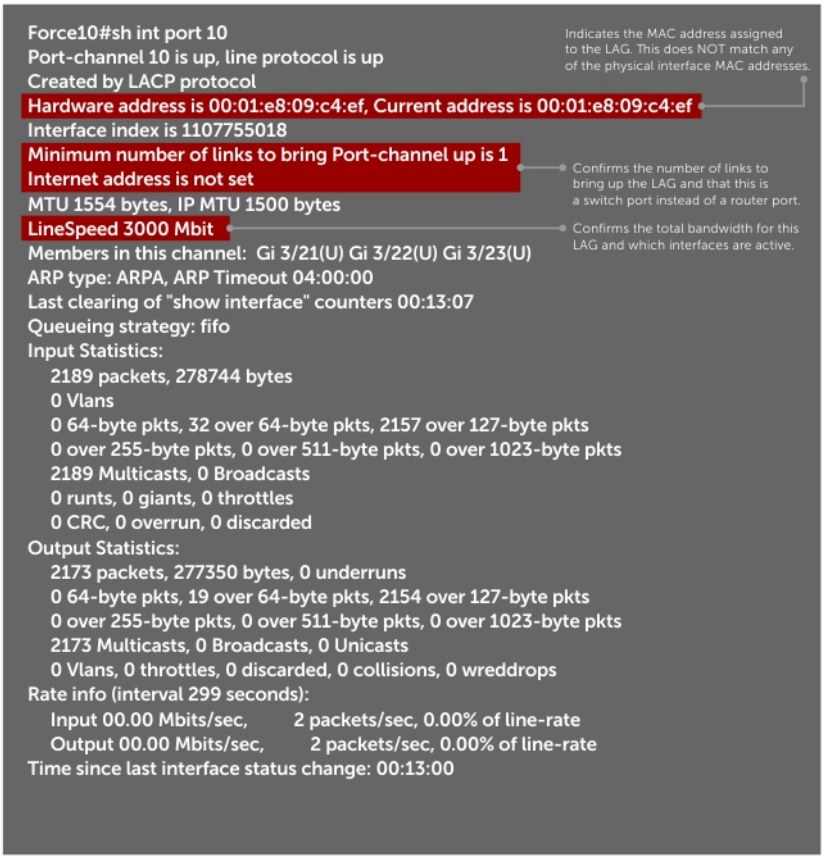


Figure 71. Inspecting LAG 10 Using the show interfaces port-channel Command

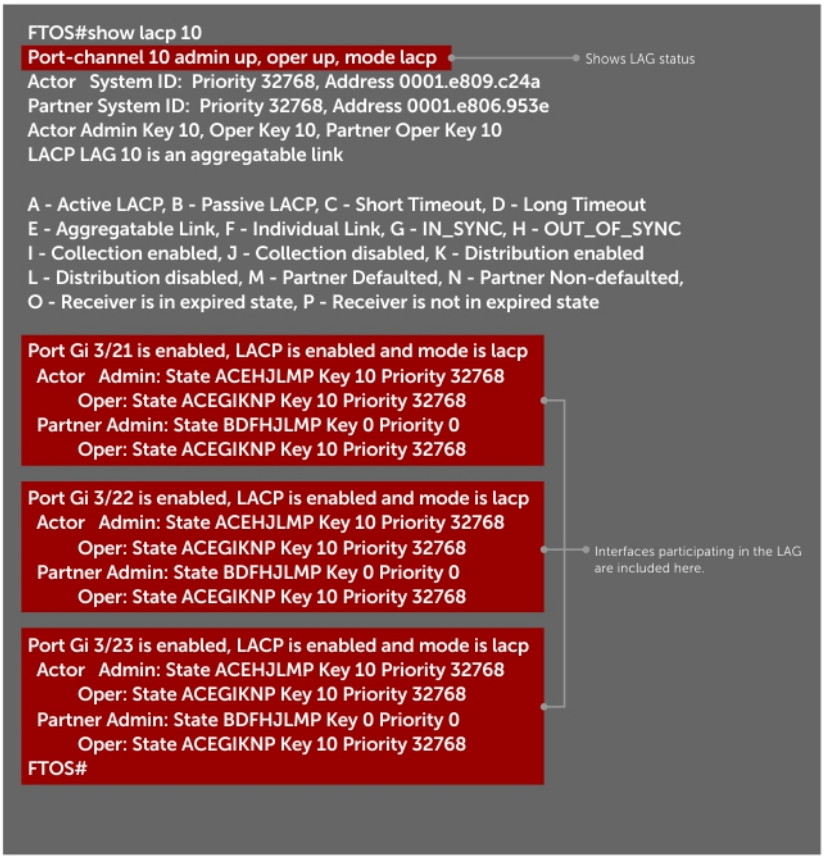


Figure 72. Inspecting the LAG Status Using the show lacp command

The point-to-point protocol (PPP) is a connection-oriented protocol that enables layer two links over various different physical layer connections. It is supported on both synchronous and asynchronous lines, and can operate in Half-Duplex or Full-Duplex mode. It was designed to carry IP traffic but is general enough to allow any type of network layer datagram to be sent over a PPP connection. As its name implies, it is for point-to-point connections between exactly two devices, and assumes that frames are sent and received in the same order.

28

Layer 2

This chapter describes the Layer 2 features supported on the device.

Topics:

- Manage the MAC Address Table
- MAC Learning Limit
- NIC Teaming
- Configure Redundant Pairs
- Far-End Failure Detection

Manage the MAC Address Table

You can perform the following management tasks in the MAC address table.

- Clearing the MAC Address Table
- Setting the Aging Time for Dynamic Entries
- Configuring a Static MAC Address
- Displaying the MAC Address Table

Clearing the MAC Address Table

You may clear the MAC address table of dynamic entries.

To clear a MAC address table, use the following command.

- Clear a MAC address table of dynamic entries.

  EXEC Privilege mode

  ```
  clear mac-address-table {dynamic | sticky} {address | all | interface | vlan}
  ```

  - *address*: deletes the specified entry.
  - `all`: deletes all dynamic entries.
  - `interface`: deletes all entries for the specified interface.
  - `vlan`: deletes all entries for the specified VLAN.

Setting the Aging Time for Dynamic Entries

Learned MAC addresses are entered in the table as dynamic entries, which means that they are subject to aging.

For any dynamic entry, if no packet arrives on the switch with the MAC address as the source or destination address within the timer period, the address is removed from the table. The default aging time is **1800 seconds**.

To disable a MAC address and specify an aging time, use the following commands.

- Disable MAC address aging for all dynamic entries.

  CONFIGURATION mode

  ```
  mac-address-table aging-time 0
  ```

- Specify an aging time.

  CONFIGURATION mode

  ```
  mac-address-table aging-time seconds
  ```

  The range is from 10 to 1000000.

Configuring a Static MAC Address

A static entry is one that is not subject to aging. Enter static entries manually.

To create a static MAC address entry, use the following command.

- Create a static MAC address entry in the MAC address table.

  CONFIGURATION mode

  `mac-address-table static`

Displaying the MAC Address Table

To display the MAC address table, use the following command.

- Display the contents of the MAC address table.

  EXEC Privilege mode

  `show mac-address-table [address | aging-time [vlan` *`vlan-id`*`]| count | dynamic | interface | static | vlan]`

  - `address`: displays the specified entry.
  - `aging-time`: displays the configured aging-time.
  - `count`: displays the number of dynamic and static entries for all VLANs, and the total number of entries.
  - `dynamic`: displays only dynamic entries.
  - `interface`: displays only entries for the specified interface.
  - `static`: displays only static entries.
  - `vlan`: displays only entries for the specified VLAN.

MAC Learning Limit

MAC address learning limit is a method of port security on Layer 2 port-channel and physical interfaces, and VLANs. It allows you to set an upper limit on the number of MAC addresses that learned on an interface/VLAN. After the limit is reached, the system drops all traffic from a device with an unlearned MAC address.

This section describes the following:

- Setting the MAC Learning Limit
- mac learning-limit Dynamic
- mac learning-limit mac-address-sticky
- mac learning-limit station-move
- Learning Limit Violation Actions
- Setting Station Move Violation Actions
- Recovering from Learning Limit and Station Move Violations

Dell EMC Networking OS Behavior: When configuring the MAC learning limit on a port or VLAN, the configuration is accepted (becomes part of `running-config` and `show mac learning-limit interface`) before the system verifies that sufficient CAM space exists. If the CAM check fails, a message is displayed:

```
%E90MH:5 %ACL_AGENT-2-ACL_AGENT_LIST_ERROR: Unable to apply access-list Mac-Limit on
TenGigabitEthernet 1/24
```

In this case, the configuration is still present in the `running-config` and `show` output. Remove the configuration before re-applying a MAC learning limit with a lower value. Also, ensure that you can view the Syslog messages on your session.

NOTE: The CAM-check failure message beginning in Dell EMC Networking OS version 8.3.1.0 is different from versions 8.2.1.1 and earlier, which read:

```
% Error: ACL returned error
% Error: Remove existing limit configuration if it was configured before
```

Setting the MAC Learning Limit

To set a MAC learning limit on an interface, use the following command.

- Specify the number of MAC addresses that the system can learn off a Layer 2 interface.
  INTERFACE mode
  `mac learning-limit address_limit`

  Three options are available with the `mac learning-limit` command:
  - `dynamic`
  - `no-station-move`
  - `station-move`

  NOTE: An SNMP trap is available for `mac learning-limit station-move`. No other SNMP traps are available for MAC Learning Limit, including limit violations.

mac learning-limit Dynamic

The MAC address table is stored on the Layer 2 forwarding information base (FIB) region of the CAM.

The Layer 2 FIB region allocates space for static MAC address entries and dynamic MAC address entries. When you enable MAC learning limit, entries created on this port are static by default. When you configure the `dynamic` option, learned MAC addresses are stored in the dynamic region and are subject to aging. Entries created before this option is set are not affected.

Dell EMC Networking OS Behavior: If you do not configure the `dynamic` option, the system does not detect station moves in which a MAC address learned from a MAC-limited port is learned on another port on the same system. Therefore, any configured violation response to detected station moves is not performed. When a MAC address is relearned on any other line card (any line card except the one to which the original MAC-limited port belongs), the station-move is detected and the system takes the configured the violation action.

mac learning-limit mac-address-sticky

Using sticky MAC addresses allows you to associate a specific port with MAC addresses from trusted devices. If you enable sticky MAC, the specified port retains any dynamically-learned addresses and prevents them from being transferred or learned on other ports.

If you configure `mac-learning-limit` and you enabled sticky MAC, all dynamically-learned addresses are converted to sticky MAC addresses for the selected port. Any new MAC addresses learned on this port is converted to sticky MAC addresses.

To save all sticky MAC addresses into a configuration file that can be used as a startup configuration file, use the `write config` command. If the number of existing MAC addresses is fewer than the configured mac learn limit, any additional MAC addresses are converted to sticky MACs on that interface. To remove all sticky MAC addresses from the running config file, disable sticky MAC and use the `write config` command.

When you enable sticky mac on an interface, dynamically-learned MAC addresses do not age, even if you enabled `mac-learning-limit dynamic`. If you configured `mac-learning-limit` and `mac-learning-limit dynamic` and you disabled sticky MAC, any dynamically-learned MAC addresses ages.

mac learning-limit station-move

The `mac learning-limit station-move` command allows a MAC address already in the table to be learned from another interface.

For example, if you disconnect a network device from one interface and reconnect it to another interface, the MAC address is learned on the new interface. When the system detects this “station move,” the system clears the entry learned on the original interface and installs a new entry on the new interface.

mac learning-limit no-station-move

The `no-station-move` option, also known as “sticky MAC,” provides additional port security by preventing a station move.

When you configure this option, the first entry in the table is maintained instead of creating an entry on the new interface. `no-station-move` is the default behavior. Entries created before you set this option are not affected.

To display a list of all interfaces with a MAC learning limit, use the following command.

Display a list of all interfaces with a MAC learning limit.

EXEC Privilege mode

```
show mac learning-limit
```

Dell EMC Networking OS Behavior: The systems do not generate a station-move violation log entry for physical interfaces or port-channels when you configure `mac learning-limit` or when you configure `mac learning-limit station-move-violation log`. Dell EMC Networking OS detects a station-move violation only when you configure `mac learning-limit dynamic` and logs the violation only when you configure the `mac learning-limit station-move-violation log`, as shown in the following example.

```
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  mac learning-limit 1 dynamic no-station-move
  mac learning-limit station-move-violation log
  no shutdown
```

Learning Limit Violation Actions

To configure the system to take an action when the MAC learning limit is reached on an interface and a new address is received using one the following options with the `mac learning-limit` command, use the following commands.

- Generate a system log message when the MAC learning limit is exceeded.

  INTERFACE mode

  ```
  learn-limit-violation log
  ```
- Shut down the interface and generate a system log message when the MAC learning limit is exceeded.

  INTERFACE mode

  ```
  learn-limit-violation shutdown
  ```

Setting Station Move Violation Actions

`no-station-move` is the default behavior. You can configure the system to take an action if a station move occurs using one the following options with the `mac learning-limit` command.

To display a list of interfaces configured with MAC learning limit or station move violation actions, use the following commands.

- Generate a system log message indicating a station move.

  INTERFACE mode

  ```
  station-move-violation log
  ```
- Shut down the first port to learn the MAC address.

  INTERFACE mode

  ```
  station-move-violation shutdown-original
  ```
- Shut down the second port to learn the MAC address.

  INTERFACE mode

  ```
  station-move-violation shutdown-offending
  ```
- Shut down both the first and second port to learn the MAC address.

  INTERFACE mode

  ```
  station-move-violation shutdown-both
  ```

- Display a list of all of the interfaces configured with MAC learning limit or station move violation.
  CONFIGURATION mode
  ```
  show mac learning-limit violate-action
  ```
  NOTE: When the MAC learning limit (MLL) is configured as `no-station-move`, the MLL will be processed as static entries internally. For static entries, the MAC address will be installed in all port-pipes, irrespective of the VLAN membership.

Recovering from Learning Limit and Station Move Violations

After a learning-limit or station-move violation shuts down an interface, you must manually reset it.

To reset the learning limit, use the following commands.

NOTE: Alternatively, you can reset the interface by shutting it down using the `shutdown` command and then re-enabling it using the `no shutdown` command.

- Reset interfaces in the ERR_Disabled state caused by a learning limit violation or station move violation.
  EXEC Privilege mode
  ```
  mac learning-limit reset
  ```
- Reset interfaces in the ERR_Disabled state caused by a learning limit violation.
  EXEC Privilege mode
  ```
  mac learning-limit reset learn-limit-violation [interface | all]
  ```
- Reset interfaces in the ERR_Disabled state caused by a station move violation.
  EXEC Privilege mode
  ```
  mac learning-limit reset station-move-violation [interface | all]
  ```

Disabling MAC Address Learning on the System

You can configure the system to not learn MAC addresses from LACP and LLDP BPDUs.

To disable source MAC address learning from LACP and LLDP BPDUs, follow this procedure:

- Disable source MAC address learning from LACP BPDUs.
  CONFIGURATION mode
  ```
  mac-address-table disable-learning lacp
  ```
- Disable source MAC address learning from LLDP BPDUs.
  CONFIGURATION mode
  ```
  mac-address-table disable-learning lldp
  ```
- Disable source MAC address learning from LACP and LLDP BPDUs.
  CONFIGURATION mode
  ```
  mac-address-table disable-learning
  ```
  If you don't use any option, the `mac-address-table disable-learning` command disables source MAC address learning from both LACP and LLDP BPDUs.

Enabling port security

You can enable or disable port security feature globally on the Dell EMC Networking OS.

You can configure all the MAC address learning limit configurations, only if the port security is enabled on the Dell EMC Networking OS. If the port security feature is disabled, all the interface level configurations are reset and all dynamically learnt MAC addresses on the interfaces configured with MAC address learning limit are cleared.

To enable the port security feature globally in the system, use the following command.

- Enable the port security feature.
  CONFIGURATION mode

```
mac port-security
```

NIC Teaming

NIC teaming is a feature that allows multiple network interface cards in a server to be represented by one MAC address and one IP address in order to provide transparent redundancy, balancing, and to fully utilize network adapter resources.

The following illustration shows a topology where two NICs have been teamed together. In this case, if the primary NIC fails, traffic switches to the secondary NIC because they are represented by the same set of addresses.

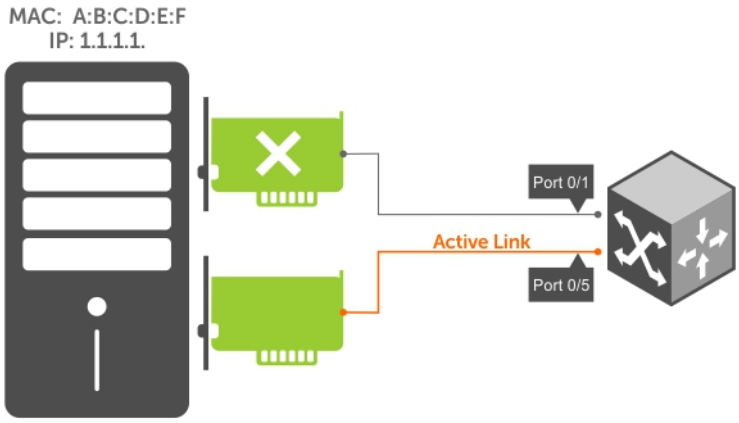


Figure 73. Redundant NICs with NIC Teaming

When you use NIC teaming, consider that the server MAC address is originally learned on Port 0/1 of the switch (shown in the following) and Port 0/5 is the failover port. When the NIC fails, the system automatically sends an ARP request for the gateway or host NIC to resolve the ARP and refresh the egress interface. When the ARP is resolved, the same MAC address is learned on the same port where the ARP is resolved (in the previous example, this location is Port 0/5 of the switch). To ensure that the MAC address is disassociated with one port and reassociated with another port in the ARP table, the `no mac-address-table station-move refresh-arp` command should not be configured on the Dell EMC Networking switch at the time that NIC teaming is being configured on the server.

NOTE: If you have configured the `no mac-address-table station-move refresh-arp` command, traffic continues to be forwarded to the failed NIC until the ARP entry on the switch times out.

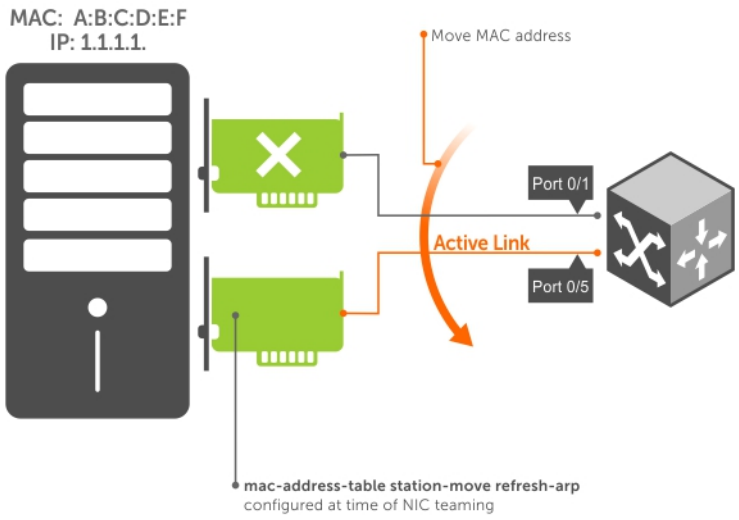


Figure 74. Configuring the mac-address-table station-move refresh-arp Command

Configure Redundant Pairs

Networks that employ switches that do not support the spanning tree protocol (STP) — for example, networks with digital subscriber line access multiplexers (DSLAM) — cannot have redundant links between switches because they create switching loops (as shown in the following illustration). The redundant pairs feature allows you to create redundant links in networks that do not use STP by configuring backup interfaces for the interfaces on either side of the primary link.

NOTE: For more information about STP, refer to Spanning Tree Protocol (STP).

Assign a backup interface to an interface using the `switchport backup` command. The backup interface remains in a Down state until the primary fails, at which point it transitions to Up state. If the primary interface fails, and later comes up, it becomes the backup interface for the redundant pair. Dell EMC Networking OS supports Gigabit, 10 Gigabit, and 40-Gigabit interfaces as backup interfaces.

Apply all other configurations to each interface in the redundant pair such that their configurations are *identical*, so that transition to the backup interface in the event of a failure is transparent to rest of the network.

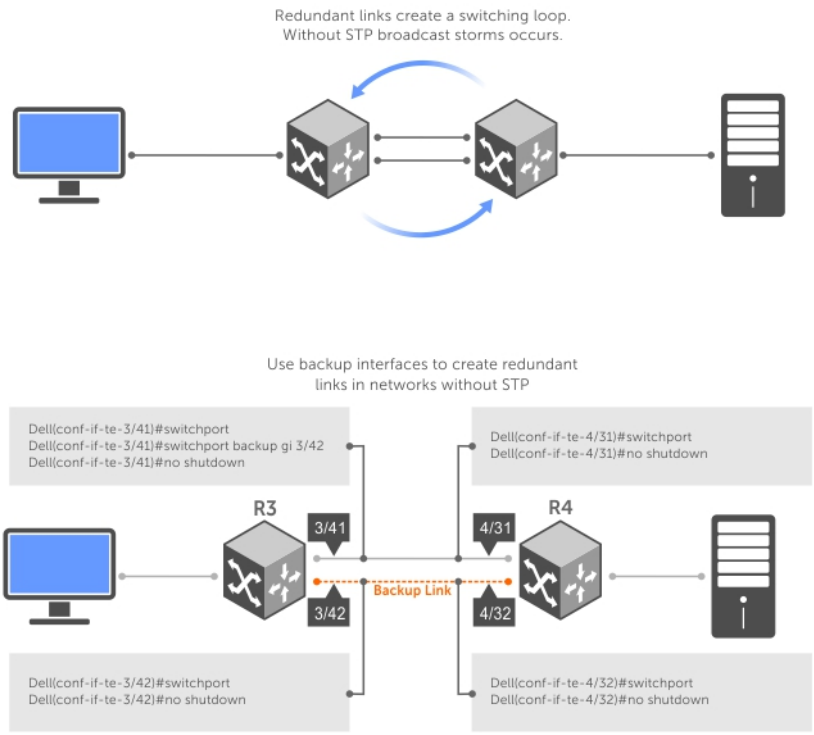


Figure 75. Configuring Redundant Layer 2 Pairs without Spanning Tree

You configure a redundant pair by assigning a backup interface to a primary interface with the `switchport backup interface` command. Initially, the primary interface is active and transmits traffic and the backup interface remains down. If the primary fails for any reason, the backup transitions to an active Up state. If the primary interface fails and later comes back up, it remains as the backup interface for the redundant pair.

Following primary/backup interfaces are also supported:

- primary interface is a physical interface, the backup interface can be a physical interface
- primary interface is a physical interface, the backup interface can be a static or dynamic LAG
- primary interface is a static or dynamic LAG, the backup interface can be a physical interface
- primary interface is a static or dynamic LAG, the backup interface can be a static or dynamic LAG

In a redundant pair, any combination of physical and port-channel interfaces is supported as the two interfaces in a redundant pair. For example, you can configure a static (without LACP) or dynamic (with LACP) port-channel interface as either the primary or backup link in a redundant pair with a physical interface.

To ensure that existing network applications see no difference when a primary interface in a redundant pair transitions to the backup interface, be sure to apply identical configurations of other traffic parameters to each interface.

If you remove an interface in a redundant link (remove the line card of a physical interface or delete a port channel with the `no interface port-channel` command), the redundant pair configuration is also removed.

Important Points about Configuring Redundant Pairs

- You may not configure any interface to be a backup for more than one interface, no interface can have more than one backup, and a backup interface may not have a backup interface.
- The active or backup interface can be a LAG, but it cannot be a member port of a LAG.
- The active and standby do not have to be of the same type (1G, 10G, and so on).
- You may not enable any Layer 2 protocol on any interface of a redundant pair or to ports connected to them.

As shown in the above illustration, interface 3/41 is a backup interface for 3/42, and 3/42 is in the Down state. If 3/41 fails, 3/42 transitions to the Up state, which makes the backup link active. A message similar to the following message appears whenever you configure a backup port.

```
02:28:04: %RPM0-P:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on Te 3/41
and Te 3/42
02:28:04: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 3/42
02:28:04: %RPM0-P:CP %IFMGR-5-STATE_ACT_STBY: Changed interface state to standby: te 3/42
```

Example of Configuring Redundant Layer 2 Pairs

```
DellEMC(conf-if-range-te-3/41-42)#switchport backup interface TenGigabitEthernet 3/42
DellEMC(conf-if-range-te-3/41-42)#show config
!
interface TenGigabitEthernet 3/41
  no ip address
  switchport
  switchport backup interface TenGigabitEthernet 3/42
  no shutdown
!
interface TenGigabitEthernet 3/42
  no ip address
  switchport
  no shutdown
DellEMC(conf-if-range-te-3/41-42)#
DellEMC(conf-if-range-te-3/41-42)#do show ip int brief | find 3/41
TenGigabitEthernet 3/41   unassigned         YES Manual up       up
TenGigabitEthernet 3/42   unassigned      NO Manual up     down
[output omitted]
DellEMC(conf-if-range-te-3/41-42)#interface tengig 3/41
DellEMC(conf-if-te-3/41)#shutdown
00:24:53: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 3/41
DellEMC(conf-if-te-3/41)#00:24:55: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface
state to
down: Te 3/41
00:24:55: %RPM0-P:CP %IFMGR-5-INACTIVE: Changed Vlan interface state to inactive: Vl 1
00:24:55: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 3/42
00:24:55: %RPM0-P:CP %IFMGR-5-ACTIVE: Changed Vlan interface state to active: Vl 1
00:24:55: %RPM0-P:CP %IFMGR-5-STATE_STBY_ACT: Changed interface state from standby to active:
Te 3/42

DellEMC(conf-if-te-3/41)#do show ip int brief | find 3/41
TenGigabitEthernet 3/41   unassigned      NO Manual administratively down down
TenGigabitEthernet 3/42   unassigned    YES Manual up          up
[output omitted]
```

Example of Configuring Redundant Pairs on a Port-Channel

```
DellEMC#show interfaces port-channel brief
Codes: L - LACP Port-channel

  LAG  Mode  Status  Uptime     Ports
  1    L2    up      00:08:33   Te 1/1 (Up)
  2    L2    up      00:00:02   Te 2/1 (Up)
DellEMC#configure
DellEMC(conf)#interface port-channel 1
DellEMC(conf-if-po-1)#switchport backup interface port-channel 2
Apr 9 00:15:13: %STKUNIT0-M:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on
Po 1 and Po 2
Apr 9 00:15:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 2
Apr 9 00:15:13: %STKUNIT0-M:CP %IFMGR-5-STATE_ACT_STBY: Changed interface state to
standby: Po 2
DellEMC(conf-if-po-1)#
DellEMC#
DellEMC#show interfaces switchport backup
Interface        Status   Paired Interface     Status
Port-channel 1   Active   Port-chato mannel 2  Standby
Port-channel 2   Standby  Port-channel 1       Active
DellEMC#

DellEMC(conf-if-po-1)#switchport backup interface tengigabitethernet 1/2
```

```
Apr 9 00:16:29: %STKUNIT0-M:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on
Po 1 and Te 1/2
DellEMC(conf-if-po-1)#
```

Far-End Failure Detection

Far-end failure detection (FEFD) is a protocol that senses remote data link errors in a network. FEFD responds by sending a unidirectional report that triggers an echoed response after a specified time interval. You can enable FEFD globally or locally on an interface basis. Disabling the global FEFD configuration does not disable the interface configuration.

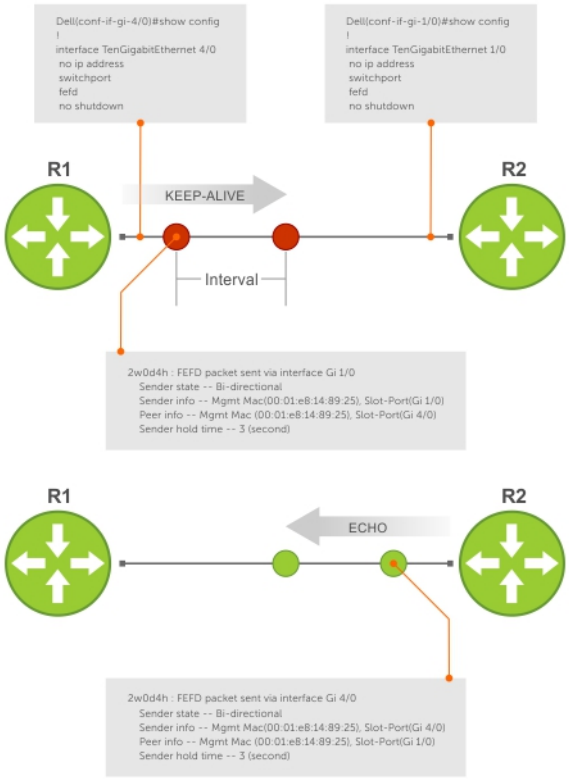


Figure 76. Configuring Far-End Failure Detection

The report consists of several packets in SNAP format that are sent to the nearest known MAC address.

In the event of a far-end failure, the device stops receiving frames and, after the specified time interval, assumes that the far-end is not available. The connecting line protocol is brought down so that upper layer protocols can detect the neighbor unavailability faster.

FEFD State Changes

FEFD has two operational modes, Normal and Aggressive.

When you enable Normal mode on an interface and a far-end failure is detected, no intervention is required to reset the interface to bring it back to an FEFD operational state. When you enable Aggressive mode on an interface in the same state, manual intervention is required to reset the interface.

FEFD enabled systems (comprised of one or more interfaces) automatically switches between four different states: Idle, Unknown, Bi-directional, and Err-disabled.

1. An interface on which FEFD is not configured is in Normal mode by default.
2. After you enable FEFD on an interface, it transitions to the Unknown state and sends an FEFD packet to the remote end of the link.

3. When the local interface receives the echoed packet from the remote end, the local interface transitions to the Bi-directional state.
4. If the FEFD enabled system is configured to use FEFD in Normal mode and neighboring echoes are not received after three intervals, (you can set each interval can be set between 3 and 300 seconds) the state changes to unknown.
5. If the FEFD system has been set to Aggressive mode and neighboring echoes are not received after three intervals, the state changes to Err-disabled. You must manually reset all interfaces in the Err-disabled state using the `fefd reset [interface]` command in EXEC privilege mode (it can be done globally or one interface at a time) before the FEFD enabled system can become operational again.

Table 52. State Change When Configuring FEFD

Local Event	Mode	Local State	Remote State	Local Admin Status	Local Protocol Status	Remote Admin Status	Remote Protocol Status
Shutdown	Normal	Admin Shutdown	Unknown	Down	Down	Up	Down
Shutdown	Aggressive	Admin Shutdown	Err-disabled	Down	Down	Up	Down
FEFD enable	Normal	Bi-directional	Bi-directional	Up	Up	Up	Up
FEFD enable	Aggressive	Bi-directional	Bi-directional	Up	Up	Up	Up
FEFD + FEFD disable	Normal	Locally disabled	Unknown	Up	Down	Up	Down
FEFD + FEFD disable	Aggressive	Locally disabled	Err-disabled	Up	Down	Up	Down
Link Failure	Normal	Unknown	Unknown	Up	Down	Up	Down
Link Failure	Aggressive	Unknown	Unknown	Up	Down	Up	Down

Important Points to Remember

- FEFD enabled ports are subject to an 8 to 10 second delay during an RPM failover before becoming operational.
- You can enable FEFD globally or on a per-interface basis. Interface FEFD configurations override global FEFD configurations.
- Dell EMC Networking OS supports FEFD on physical Ethernet interfaces only, excluding the management interface.
- FEFD is not supported on Fibre Channel and copper Ethernet ports.

Configuring FEFD

You can configure FEFD for all interfaces from CONFIGURATION mode, or on individual interfaces from INTERFACE mode.

To enable FEFD globally on all interfaces, use the following command.

- Enable FEFD globally on all interfaces.

  CONFIGURATION mode

  ```
  fefd-global
  ```

To report interval frequency and mode adjustments, use the following commands.

1. Setup two or more connected interfaces for Layer 2 or Layer 3.

   INTERFACE mode

   ```
   ip address ip address, switchport
   ```
2. Enable the necessary ports administratively.

   INTERFACE mode

   ```
   no shutdown
   ```
3. Enable fefd globally.

   CONFIGURATION mode

```
fefd-global {interval | mode}
```

To display information about the state of each interface, use the `show fefd` command in EXEC privilege mode.

```
DellEMC#show fefd
FEFD is globally 'ON', interval is 3 seconds, mode is 'Normal'.

INTERFACE  MODE    INTERVAL    STATE
                   (second)
Te 1/1     Normal  3           Bi-directional
Te 1/2     Normal  3           Admin Shutdown
Te 1/3     Normal  3           Admin Shutdown
Te 1/4     Normal  3           Admin Shutdown

DellEMC#show run fefd
!
fefd-global mode normal
fefd-global interval 3
```

Enabling FEFD on an Interface

To enable, change, or disable FEFD on an interface, use the following commands.

- Enable FEFD on a per interface basis.

  INTERFACE mode

  ```
  fefd
  ```

- Change the FEFD mode.

  INTERFACE mode

  ```
  fefd [mode {aggressive | normal}]
  ```

- Disable FEFD protocol on one interface.

  INTERFACE mode

  ```
  fefd disable
  ```

  Disabling an interface shuts down all protocols working on that interface's connected line. It does not delete your previous FEFD configuration which you can enable again at any time.

To set up and activate two or more connected interfaces, use the following commands.

1. Setup two or more connected interfaces for Layer 2 or Layer 3.

   INTERFACE mode

   ```
   ip address ip address, switchport
   ```

2. Activate the necessary ports administratively.

   INTERFACE mode

   ```
   no shutdown
   ```

3. INTERFACE mode

   ```
   fefd {disable | interval | mode}
   ```

```
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  fefd mode normal
  no shutdown

DellEMC(conf-if-te-1/1)#do show fefd | grep 1/1
Te 1/1          Normal        3           Unknown
```

Debugging FEFD

To debug FEFD, use the first command. To provide output for each packet transmission over the FEFD enabled connection, use the second command.

- Display output whenever events occur that initiate or disrupt an FEFD enabled connection.

  EXEC Privilege mode

  ```
  debug fefd events
  ```

- Provide output for each packet transmission over the FEFD enabled connection.

  EXEC Privilege mode

  ```
  debug fefd packets
  ```

```
DellEMC#debug fefd events
DellEMC#config
DellEMC(conf)#int te 1/1
DellEMC(conf-if-te-1/1)#shutdown
2w1d22h: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 1/1
DellEMC(conf-if-te-1/1)#2w1d22h : FEFD state on Te 1/1 changed from ANY to Unknown
2w1d22h: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 1/1
2w1d22h: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 4/1
2w1d22h: %RPM0-P:CP %IFMGR-5-INACTIVE: Changed Vlan interface state to inactive: Vl 1
2w1d22h : FEFD state on Te 4/1 changed from Bi-directional to Unknown
```

```
DellEMC#debug fefd packets
DellEMC#2w1d22h : FEFD packet sent via interface Te 1/1
   Sender state -- Bi-directional
   Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Te 1/1)
   Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Te 4/1)
   Sender hold time -- 3 (second)

2w1d22h : FEFD packet received on interface Te 4/1
   Sender state -- Bi-directional
   Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Te 1/1)
   Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Te 4/1)
   Sender hold time -- 3 (second)
```

An RPM Failover

In the event that an RPM failover occurs, FEFD becomes operationally down on all enabled ports for approximately 8-10 seconds before automatically becoming operational again.

```
02-05-2009    12:40:38    Local7.Debug    10.16.151.12    Feb 5 07:06:09:
%RPM1-S:CP %RAM-6-FAILOVER_REQ: RPM failover request from active peer: User request.
02-05-2009    12:40:38    Local7.Debug    10.16.151.12    Feb 5 07:06:19:
%RPM1-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 1/45
02-05-2009    12:40:38    Local7.Debug    10.16.151.12    Feb 5 07:06:19:
%RPM1-P:CP %FEFD-5-FEFD-BIDIRECTION-LINK-DETECTED: Interface Te 1/45 has bidirectional
link with its
peer
```

29

Link Layer Discovery Protocol (LLDP)

This chapter describes how to configure and use the link layer discovery protocol (LLDP).

Topics:

- 802.1AB (LLDP) Overview
- Optional TLVs
- TIA-1057 (LLDP-MED) Overview
- Configure LLDP
- CONFIGURATION versus INTERFACE Configurations
- Enabling LLDP
- Advertising TLVs
- Viewing the LLDP Configuration
- Viewing Information Advertised by Adjacent LLDP Neighbors
- Configuring LLDPDU Intervals
- Configuring Transmit and Receive Mode
- Configuring the Time to Live Value
- Debugging LLDP
- Relevant Management Objects

802.1AB (LLDP) Overview

LLDP — defined by IEEE 802.1AB — is a protocol that enables a local area network (LAN) device to advertise its configuration and receive configuration information from adjacent LLDP-enabled LAN infrastructure devices.

The collected information is stored in a management information base (MIB) on each device, and is accessible via simple network management protocol (SNMP).

Protocol Data Units

Configuration information is exchanged in the form of Type, Length, Value (TLV) segments.

- Type — The kind of information included in the TLV.
- Length — The value, in octets, of the TLV after the Length field.
- Value — The configuration information that the agent is advertising.

The chassis ID TLV is shown in the following illustration.

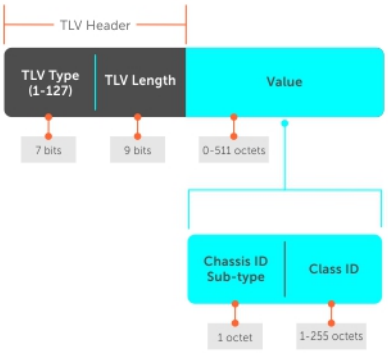


Figure 77. Type, Length, Value (TLV) Segment

TLVs are encapsulated in a frame called an LLDP data unit (LLDPDU) (shown in the following table), which is transmitted from one LLDP-enabled device to its LLDP-enabled neighbors. LLDP is a one-way protocol. LLDP-enabled devices (LLDP agents) can transmit and/or receive advertisements, but they cannot solicit and do not respond to advertisements.

There are five types of TLVs. All types are mandatory in the construction of an LLDPDU except Optional TLVs. You can configure the inclusion of individual Optional TLVs.

Table 53. Type, Length, Value (TLV) Types

Type	TLV	Description
0	End of LLDPDU	Marks the end of a LLDPDU.
1	Chassis ID	An administratively assigned name that identifies the LLDP agent.
2	Port ID	An administratively assigned name that identifies a port through which TLVs are sent and received.
3	Time to Live	An administratively assigned name that identifies a port through which TLVs are sent and received.
—	Optional	Includes sub-types of TLVs that advertise specific configuration information. These sub-types are Management TLVs, IEEE 802.1, IEEE 802.3, and TIA-1057 Organizationally Specific TLVs.

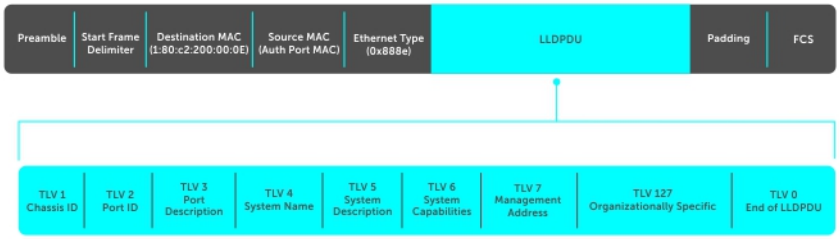


Figure 78. LLDPDU Frame

Optional TLVs

The Dell EMC Networking OS supports these optional TLVs: management TLVs, IEEE 802.1 and 802.3 organizationally specific TLVs, and TIA-1057 organizationally specific TLVs.

Management TLVs

A management TLV is an optional TLVs sub-type. This kind of TLV contains essential management information about the sender.

Organizationally Specific TLVs

A professional organization or a vendor can define organizationally specific TLVs. They have two mandatory fields (as shown in the following illustration) in addition to the basic TLV fields.

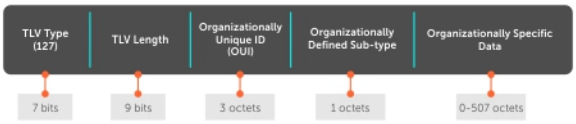


Figure 79. Organizationally Specific TLV

IEEE Organizationally Specific TLVs

Eight TLV types have been defined by the IEEE 802.1 and 802.3 working groups as a basic part of LLDP; the IEEE OUI is 00-80-C2. You can configure the Dell EMC Networking system to advertise any or all of these TLVs.

Table 54. Optional TLV Types

Type	TLV	Description
Optional TLVs		
4	Port description	A user-defined alphanumeric string that describes the port. Dell EMC Networking OS does not currently support this TLV.
5	System name	A user-defined alphanumeric string that identifies the system.
6	System description	A user-defined alphanumeric string that identifies the system.
7	System capabilities	Identifies the chassis as one or more of the following: repeater, bridge, WLAN Access Point, Router, Telephone, DOCSIS cable device, end station only, or other.
8	Management address	Indicates the network address of the management interface. Dell EMC Networking OS does not currently support this TLV.
IEEE 802.1 Organizationally Specific TLVs		
127	Port-VLAN ID	On Dell EMC Networking systems, indicates the untagged VLAN to which a port belongs.
127	Port and Protocol VLAN ID	On Dell EMC Networking systems, indicates the tagged VLAN to which a port belongs (and the untagged VLAN to which a port belongs if the port is in Hybrid mode).
127	Protocol Identity	Indicates the protocols that the port can process. Dell EMC Networking OS does not currently support this TLV.
IEEE 802.3 Organizationally Specific TLVs		
127	MAC/PHY Configuration/Status	Indicates the capability and current setting of the duplex status and bit rate, and whether the current settings are the result of auto-negotiation. This TLV is not available in the Dell EMC Networking OS implementation of LLDP, but is available and mandatory

Table 54. Optional TLV Types (continued)

Type	TLV	Description
		(non-configurable) in the LLDP-MED implementation.
127	Power via MDI	Dell EMC Networking supports the LLDP-MED protocol, which recommends that Power via MDI TLV be not implemented, and therefore Dell EMC Networking implements Extended Power via MDI TLV only.
127	Link Aggregation	Indicates whether the link is capable of being aggregated, whether it is currently in a LAG, and the port identification of the LAG. Dell EMC Networking OS does not currently support this TLV.
127	Maximum Frame Size	Indicates the maximum frame size capability of the MAC and PHY.

TIA-1057 (LLDP-MED) Overview

Link layer discovery protocol — media endpoint discovery (LLDP-MED) as defined by ANSI/ TIA-1057— provides additional organizationally specific TLVs so that endpoint devices and network connectivity devices can advertise their characteristics and configuration information; the OUI for the Telecommunications Industry Association (TIA) is 00-12-BB.

- **LLDP-MED Endpoint Device** — any device that is on an IEEE 802 LAN network edge can communicate using IP and uses the LLDP-MED framework.
- **LLDP-MED Network Connectivity Device** — any device that provides access to an IEEE 802 LAN to an LLDP-MED endpoint device and supports IEEE 802.1AB (LLDP) and TIA-1057 (LLDP-MED). The Dell EMC Networking system is an LLDP-MED network connectivity device.

Regarding connected endpoint devices, LLDP-MED provides network connectivity devices with the ability to:

- manage inventory
- manage Power over Ethernet (PoE)
- identify physical location
- identify network policy

LLDP-MED is designed for, but not limited to, VoIP endpoints.

TIA Organizationally Specific TLVs

The Dell EMC Networking system is an LLDP-MED Network Connectivity Device (Device Type 4).

Network connectivity devices are responsible for:

- transmitting an LLDP-MED capability TLV to endpoint devices
- storing the information that endpoint devices advertise

The following table describes the five types of TIA-1057 Organizationally Specific TLVs.

Table 55. TIA-1057 (LLDP-MED) Organizationally Specific TLVs

Type	SubType	TLV	Description
127	1	LLDP-MED Capabilities	Indicates: • whether the transmitting device supports LLDP-MED • what LLDP-MED TLVs it supports • LLDP device class

Table 55. TIA-1057 (LLDP-MED) Organizationally Specific TLVs (continued)

Type	SubType	TLV	Description
127	2	Network Policy	Indicates the application type, VLAN ID, Layer 2 Priority, and DSCP value.
127	3	Location Identification	Indicates that the physical location of the device expressed in one of three possible formats: • Coordinate Based LCI • Civic Address LCI • Emergency Call Services ELIN
127	4	Location Identification	Indicates power requirements, priority, and power status.
Inventory Management TLVs	Implementation of this set of TLVs is optional in LLDP-MED devices. None or all TLVs must be supported. Dell EMC Networking OS does not currently support these TLVs.		
127	5	Inventory — Hardware Revision	Indicates the hardware revision of the LLDP-MED device.
127	6	Inventory — Firmware Revision	Indicates the firmware revision of the LLDP-MED device.
127	7	Inventory — Software Revision	Indicates the software revision of the LLDP-MED device.
127	8	Inventory — Serial Number	Indicates the device serial number of the LLDP-MED device.
127	9	Inventory — Manufacturer Name	Indicates the manufacturer of the LLDP-MED device.
127	10	Inventory — Model Name	Indicates the model of the LLDP-MED device.
127	11	Inventory — Asset ID	Indicates a user specified device number to manage inventory.
127	12–255	Reserved	—

LLDP-MED Capabilities TLV

The LLDP-MED capabilities TLV communicates the types of TLVs that the endpoint device and the network connectivity device support. LLDP-MED network connectivity devices must transmit the Network Policies TLV.

- The value of the LLDP-MED capabilities field in the TLV is a 2–octet bitmap, each bit represents an LLDP-MED capability (as shown in the following table).
- The possible values of the LLDP-MED device type are shown in the following. The Dell EMC Networking system is a network connectivity device, which is Type 4.

When you enable LLDP-MED in Dell EMC Networking OS (using the `advertise med` command), the system begins transmitting this TLV.

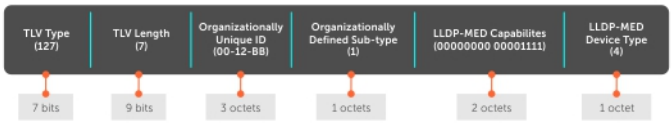


Figure 80. LLDP-MED Capabilities TLV

Table 56. Dell EMC Networking OS LLDP-MED Capabilities

Bit Position	TLV	Dell EMC Networking OS Support
0	LLDP-MED Capabilities	Yes
1	Network Policy	Yes
2	Location Identification	Yes
3	Extended Power via MDI-PSE	Yes
4	Extended Power via MDI-PD	No
5	Inventory	No
6–15	reserved	No

Table 57. LLDP-MED Device Types

Value	Device Type
0	Type Not Defined
1	Endpoint Class 1
2	Endpoint Class 2
3	Endpoint Class 3
4	Network Connectivity
5–255	Reserved

LLDP-MED Network Policies TLV

A network policy in the context of LLDP-MED is a device's VLAN configuration and associated Layer 2 and Layer 3 configurations.

LLDP-MED network policies TLV include:

- VLAN ID
- VLAN tagged or untagged status
- Layer 2 priority
- DSCP value

An integer represents the application type (the Type integer shown in the following table), which indicates a device function for which a unique network policy is defined. An individual LLDP-MED network policy TLV is generated for each application type that you specify with the Dell EMC Networking OS CLI (Advertising TLVs).

NOTE: As shown in the following table, signaling is a series of control packets that are exchanged between an endpoint device and a network connectivity device to establish and maintain a connection. These signal packets might require a different network policy than the media packets for which a connection is made. In this case, configure the signaling application.

Table 58. Network Policy Applications

Type	Application	Description
0	Reserved	—

Table 58. Network Policy Applications (continued)

Type	Application	Description
1	Voice	Specify this application type for dedicated IP telephony handsets and other appliances supporting interactive voice services.
2	Voice Signaling	Specify this application type only if voice control packets use a separate network policy than voice data.
3	Guest Voice	Specify this application type for a separate limited voice service for guest users with their own IP telephony handsets and other appliances supporting interactive voice services.
4	Guest Voice Signaling	Specify this application type only if guest voice control packets use a separate network policy than voice data.
5	Softphone Voice	Specify this application type only if guest voice control packets use a separate network policy than voice data.
6	Video Conferencing	Specify this application type for dedicated video conferencing and other similar appliances supporting real-time interactive video.
7	Streaming Video	Specify this application type for dedicated video conferencing and other similar appliances supporting real-time interactive video.
8	Video Signaling	Specify this application type only if video control packets use a separate network policy than video data.
9–255	Reserved	—

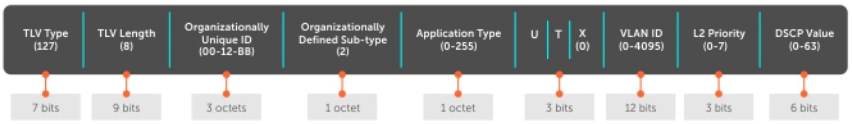


Figure 81. LLDP-MED Policies TLV

Extended Power via MDI TLV

The extended power via MDI TLV enables advanced PoE management between LLDP-MED endpoints and network connectivity devices.

Advertise the extended power via MDI on all ports that are connected to an 802.3af powered, LLDP-MED endpoint device.

- **Power Type** — there are two possible power types: power source entity (PSE) or power device (PD). The Dell EMC Networking system is a PSE, which corresponds to a value of 0, based on the TIA-1057 specification.
- **Power Source** — there are two possible power sources: primary and backup. The Dell EMC Networking system is a primary power source, which corresponds to a value of 1, based on the TIA-1057 specification.
- **Power Priority** — there are three possible priorities: Low, High, and Critical. On Dell EMC Networking systems, the default power priority is **High**, which corresponds to a value of 2 based on the TIA-1057 specification. You can configure a different power priority through the CLI. Dell EMC Networking also honors the power priority value the powered device sends; however, the CLI configuration takes precedence.
- **Power Value** — Dell EMC Networking advertises the maximum amount of power that can be supplied on the port. By default the power is **15.4W**, which corresponds to a power value of 130, based on the TIA-1057 specification. You can advertise a different power value using the `max-milliwatts` option with the `power inline auto | static` command. Dell EMC Networking also honors the power value (power requirement) the powered device sends when the port is configured for `power inline auto`.

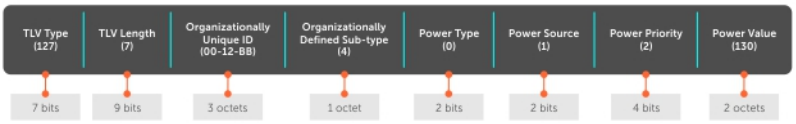


Figure 82. Extended Power via MDI TLV

Configure LLDP

Configuring LLDP is a two-step process.

1. Enable LLDP globally.
2. Advertise TLVs out of an interface.

Related Configuration Tasks

- Viewing the LLDP Configuration
- Viewing Information Advertised by Adjacent LLDP Agents
- Configuring LLDPDU Intervals
- Configuring Transmit and Receive Mode
- Configuring a Time to Live
- Debugging LLDP

Important Points to Remember

- LLDP is enabled by default.
- Dell EMC Networking systems support up to eight neighbors per interface.
- Dell EMC Networking systems support a maximum of 8000 total neighbors per system. If the number of interfaces multiplied by eight exceeds the maximum, the system does not configure more than 8000.
- INTERFACE level configurations override all CONFIGURATION level configurations.
- LLDP is not hitless.

LLDP Compatibility

- Spanning tree and force10 ring protocol “blocked” ports allow LLDPDUs.
- 802.1X controlled ports do not allow LLDPDUs until the connected device is authenticated.

CONFIGURATION versus INTERFACE Configurations

All LLDP configuration commands are available in PROTOCOL LLDP mode, which is a sub-mode of the CONFIGURATION mode and INTERFACE mode.

- Configurations made at the CONFIGURATION level are global; that is, they affect all interfaces on the system.
- Configurations made at the INTERFACE level affect only the specific interface; they override CONFIGURATION level configurations.

Example of the `protocol lldp` Command (CONFIGURATION Level)

```
R1(conf)#protocol lldp
R1(conf-lldp)#?
advertise        Advertise TLVs
disable          Disable LLDP protocol globally
end              Exit from configuration mode
exit             Exit from LLDP configuration mode
hello            LLDP hello configuration
```

```
mode              LLDP mode configuration (default = rx and tx)
multiplier        LLDP multiplier configuration
no                Negate a command or set its defaults
show              Show LLDP configuration


DellEMC(conf-lldp)#exit
DellEMC(conf)#interface tengigabitethernet 1/3
DellEMC(conf-if-te-1/3)#protocol lldp
DellEMC(conf-if-te-1/3-lldp)#?
advertise         Advertise TLVs
disable           Disable LLDP protocol on this interface
end               Exit from configuration mode
exit              Exit from LLDP configuration mode
hello             LLDP hello configuration
mode              LLDP mode configuration (default = rx and tx)
multiplier        LLDP multiplier configuration
no                Negate a command or set its defaults
show              Show LLDP configuration
DellEMC(conf-if-te-1/3-lldp)#
```

Enabling LLDP

LLDP is enabled by default. Enable and disable LLDP globally or per interface. If you enable LLDP globally, all UP interfaces send periodic LLDPDUs.

To enable LLDP, use the following command.

1. Enter Protocol LLDP mode.

   CONFIGURATION or INTERFACE mode

   `protocol lldp`

2. Enable LLDP.

   PROTOCOL LLDP mode

   `no disable`

Disabling and Undoing LLDP

To disable or undo LLDP, use the following command.

- Disable LLDP globally or for an interface.

  `disable`

To undo an LLDP configuration, precede the relevant command with the keyword `no`.

Advertising TLVs

You can configure the system to advertise TLVs out of all interfaces or out of specific interfaces.

- If you configure the system globally, all interfaces send LLDPDUs with the specified TLVs.
- If you configure an interface, only the interface sends LLDPDUs with the specified TLVs.
- If you configure LLDP both globally and at interface level, the interface level configuration overrides the global configuration.

To advertise TLVs, use the following commands.

1. Enter LLDP mode.

   CONFIGURATION or INTERFACE mode

   `protocol lldp`

2. Advertise one or more TLVs.

   PROTOCOL LLDP mode

   `advertise {dcbx-appln-tlv | dcbx-tlv | dot3-tlv | interface-port-desc | management-tlv | med }`

Include the keyword for each TLV you want to advertise.

- For management TLVs: `system-capabilities, system-description`.
- For 802.1 TLVs: `port-protocol-vlan-id, port-vlan-id vlan-name`.
- For 802.3 TLVs: `max-frame-size`.
- For TIA-1057 TLVs:
  - `guest-voice`
  - `guest-voice-signaling`
  - `location-identification`
  - `power-via-mdi`
  - `softphone-voice`
  - `streaming-video`
  - `video-conferencing`
  - `video-signaling`
  - `voice`
  - `voice-signaling`

In the following example, LLDP is enabled globally. R1 and R2 are transmitting periodic LLDPDUs that contain management, 802.1, and 802.3 TLVs.

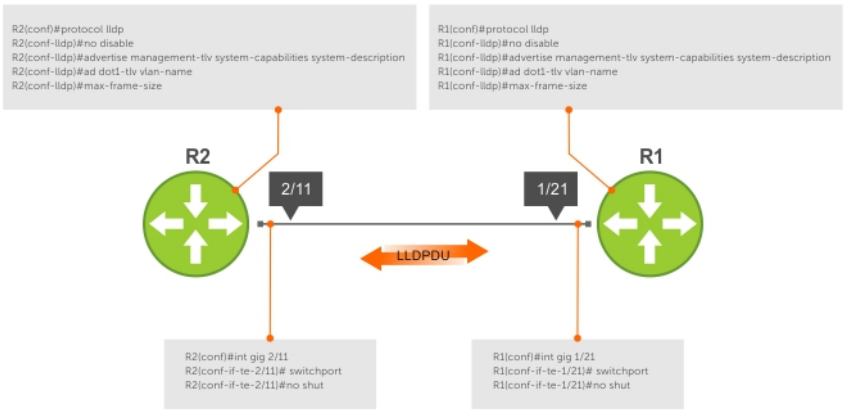


Figure 83. Configuring LLDP

Viewing the LLDP Configuration

To view the LLDP configuration, use the following command.

- Display the LLDP configuration.

  CONFIGURATION or INTERFACE mode

  `show config`

The following example shows viewing an LLDP global configuration.

```
DellEMC(conf)#protocol lldp
DellEMC(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  hello 10
  no disable
DellEMC(conf-lldp)#
```

```
DellEMC(conf-lldp)#exit
DellEMC(conf)#interface tengigabitethernet 1/31
DellEMC(conf-if-te-1/31)#show config
```

```
!
interface TenGigabitEthernet 1/31
  no ip address
  switchport
  no shutdown
DellEMC(conf-if-te-1/31)#protocol lldp
DellEMC(conf-if-te-1/31-lldp)#show config
!
  protocol lldp
DellEMC(conf-if-te-1/31-lldp)#
```

Viewing Information Advertised by Adjacent LLDP Neighbors

To view brief information about adjacent devices or to view all the information that neighbors are advertising, use the following commands.

- Display brief information about adjacent devices.

  `show lldp neighbors`
- Display all of the information that neighbors are advertising.

  `show lldp neighbors detail`

Examples of Viewing Information Advertised by Neighbors

Example of Viewing Brief Information Advertised by Neighbors

```
DellEMC(conf-if-te-1/3-lldp)#end
DellEMC(conf-if-te-1/3)#do show lldp neighbors
Loc PortID  Rem Host Name  Rem Port Id           Rem Chassis Id
-------------------------------------------------------------------
Te 1/1      -            TenGigabitEthernet 1/5  00:01:e8:05:40:46
Te 1/2      -            TenGigabitEthernet 1/6  00:01:e8:05:40:46
DellEMC(conf-if-te-1/3)#
```

Example of Viewing Detailed Information Advertised by Neighbors

```
DellEMC(conf)#do show lldp neighbors detail
================================================================================
Local Interface TenGigabitEthernet 1/1 has 2 neighbors
  Total Frames Out: 3
  Total Frames In: 8
  Total Neighbor information Age outs: 0
  Total Multiple Neighbors Detected: 0
  Total Frames Discarded: 0
  Total In Error Frames: 0
  Total Unrecognized TLVs: 960
  Total TLVs Discarded: 16
  Next packet will be sent after 9 seconds
  The neighbors are given below:
  ------------------------------------------------------------------------------

    Remote Chassis ID Subtype: Mac address (4)
    Remote Chassis ID:  00:00:00:00:00:01
    Remote Port Subtype:  Interface name (5)
    Remote Port ID:  TenGigabitEthernEt 1/40
    Local Port ID: TenGigabitEthernet 1/1
    Locally assigned remote Neighbor Index: 1
    Remote TTL:  120
    Information valid for next 44 seconds
    Time since last information change of this neighbor:  00:01:16
    UnknownTLVList:
          (  9,  4) ( 10,  4) ( 11,  4) ( 12,  4) ( 13,  4) ( 14,  4) ( 15,  4) ( 16,  4) ( 17,  4) ( 18,  4)
          ( 19,  4) ( 20,  4) ( 21,  4) ( 22,  4) ( 23,  4) ( 24,  4) ( 25,  4) ( 26,  4) ( 27,  4) ( 28,  4)
          ( 29,  4) ( 30,  4) ( 31,  4) ( 32,  4) ( 33,  4) ( 34,  4) ( 35,  4) ( 36,  4) ( 37,  4) ( 38,  4)
          ( 39,  4) ( 40,  4) ( 41,  4) ( 42,  4) ( 43,  4) ( 44,  4) ( 45,  4) ( 46,  4) ( 47,  4) ( 48,  4)
          ( 49,  4) ( 50,  4) ( 51,  4) ( 52,  4) ( 53,  4) ( 54,  4) ( 55,  4) ( 56,  4) ( 57,  4) ( 58,  4)
          ( 59,  4) ( 60,  4) ( 61,  4) ( 62,  4) ( 63,  4) ( 64,  4) ( 65,  4) ( 66,  4) ( 67,  4) ( 68,  4)
          ( 69,  4) ( 70,  4) ( 71,  4) ( 72,  4) ( 73,  4) ( 74,  4) ( 75,  4) ( 76,  4) ( 77,  4) ( 78,  4)
          ( 79,  4) ( 80,  4) ( 81,  4) ( 82,  4) ( 83,  4) ( 84,  4) ( 85,  4) ( 86,  4) ( 87,  4) ( 88,  4)
          ( 89,  4) ( 90,  4) ( 91,  4) ( 92,  4) ( 93,  4) ( 94,  4) ( 95,  4) ( 96,  4) ( 97,  4) ( 98,  4)
          ( 99,  4) (100,  4) (101,  4) (102,  4) (103,  4) (104,  4) (105,  4) (106,  4) (107,  4) (108,  4)
```

```
          (109,  4) (110,  4) (111,  4) (112,  4) (113,  4) (114,  4) (115,  4) (116,  4) (117,  4) (118,  4)
          (119,  4) (120,  4) (121,  4) (122,  4) (123,  4) (124,  4) (125,  4) (126,  4)
    OrgUnknownTLVList:
   ---------------------------------------------------------------------------

    Remote Chassis ID Subtype: Mac address (4)
    Remote Chassis ID:  00:00:00:00:00:02
    Remote Port Subtype:  Interface name (5)
    Remote Port ID:  TenGigabitEthernEt 1/40
    Local Port ID: TenGigabitEthernet 1/1
    Locally assigned remote Neighbor Index: 2
    Remote TTL:  120
    Information valid for next 43 seconds
    Time since last information change of this neighbor:  00:01:17
    UnknownTLVList:
          (  9,  4) ( 10,  4) ( 11,  4) ( 12,  4) ( 13,  4) ( 14,  4) ( 15,  4) ( 16,  4) ( 17,  4) ( 18,  4)
          ( 19,  4) ( 20,  4) ( 21,  4) ( 22,  4) ( 23,  4) ( 24,  4) ( 25,  4) ( 26,  4) ( 27,  4) ( 28,  4)
          ( 29,  4) ( 30,  4) ( 31,  4) ( 32,  4) ( 33,  4) ( 34,  4) ( 35,  4) ( 36,  4) ( 37,  4) ( 38,  4)
          ( 39,  4) ( 40,  4) ( 41,  4) ( 42,  4) ( 43,  4) ( 44,  4) ( 45,  4) ( 46,  4) ( 47,  4) ( 48,  4)
          ( 49,  4) ( 50,  4) ( 51,  4) ( 52,  4) ( 53,  4) ( 54,  4) ( 55,  4) ( 56,  4) ( 57,  4) ( 58,  4)
          ( 59,  4) ( 60,  4) ( 61,  4) ( 62,  4) ( 63,  4) ( 64,  4) ( 65,  4) ( 66,  4) ( 67,  4) ( 68,  4)
          ( 69,  4) ( 70,  4) ( 71,  4) ( 72,  4) ( 73,  4) ( 74,  4) ( 75,  4) ( 76,  4) ( 77,  4) ( 78,  4)
          ( 79,  4) ( 80,  4) ( 81,  4) ( 82,  4) ( 83,  4) ( 84,  4) ( 85,  4) ( 86,  4) ( 87,  4) ( 88,  4)
          ( 89,  4) ( 90,  4) ( 91,  4) ( 92,  4) ( 93,  4) ( 94,  4) ( 95,  4) ( 96,  4) ( 97,  4) ( 98,  4)
          ( 99,  4) (100,  4) (101,  4) (102,  4) (103,  4) (104,  4) (105,  4) (106,  4) (107,  4) (108,  4)
          (109,  4) (110,  4) (111,  4) (112,  4) (113,  4) (114,  4) (115,  4) (116,  4) (117,  4) (118,  4)
          (119,  4) (120,  4) (121,  4) (122,  4) (123,  4) (124,  4) (125,  4) (126,  4)
    OrgUnknownTLVList:
   ---------------------------------------------------------------------------

===============================================================================
Local Interface TenGigabitEthernet 1/2 has 3 neighbors
  Total Frames Out: 4
  Total Frames In: 8
  Total Neighbor information Age outs: 0
  Total Multiple Neighbors Detected: 0
  Total Frames Discarded: 0
  Total In Error Frames: 0
  Total Unrecognized TLVs: 1056
  Total TLVs Discarded: 0
  Next packet will be sent after 16 seconds
  The neighbors are given below:
  ---------------------------------------------------------------------------

    Remote Chassis ID Subtype: Mac address (4)
    Remote Chassis ID:  4c:76:25:f4:ab:01
    Remote Port Subtype:  Interface name (5)
    Remote Port ID:  fortyGigE 1/2/8/1
    Local Port ID: TenGigabitEthernet 1/2
    Locally assigned remote Neighbor Index: 1
    Remote TTL:  300
    Information valid for next 201 seconds
    Time since last information change of this neighbor:  00:01:39
    UnknownTLVList:
    OrgUnknownTLVList:
          ((00-01-66),127,  4) ((00-01-66),126,  4) ((00-01-66),125,  4) ((00-01-66),124,  4) ((00-01-66),123,
 4)
          ((00-01-66),122,  4) ((00-01-66),121,  4) ((00-01-66),120,  4) ((00-01-66),119,  4) ((00-01-66),118,
 4)
   ---------------------------------------------------------------------------

    Remote Chassis ID Subtype: Mac address (4)
    Remote Chassis ID:  4c:76:25:f4:ab:02
    Remote Port Subtype:  Interface name (5)
    Remote Port ID:  fortyGigE 1/2/8/1
    Local Port ID: TenGigabitEthernet 1/2
    Locally assigned remote Neighbor Index: 2
    Remote TTL:  300
    Information valid for next 201 seconds
    Time since last information change of this neighbor:  00:01:39
    UnknownTLVList:
    OrgUnknownTLVList:
          ((00-01-66),127,  4) ((00-01-66),126,  4) ((00-01-66),125,  4) ((00-01-66),124,  4) ((00-01-66),123,
 4)
          ((00-01-66),122,  4) ((00-01-66),121,  4) ((00-01-66),120,  4) ((00-01-66),119,  4) ((00-01-66),118,
 4)
   ---------------------------------------------------------------------------

    Remote Chassis ID Subtype: Mac address (4)
    Remote Chassis ID:  4c:76:25:f4:ab:03
    Remote Port Subtype:  Interface name (5)
    Remote Port ID:  fortyGigE 1/2/8/1
    Local Port ID: TenGigabitEthernet 1/2
    Locally assigned remote Neighbor Index: 3
    Remote TTL:  300
    Information valid for next 199 seconds
    Time since last information change of this neighbor:  00:01:41
    UnknownTLVList:
    OrgUnknownTLVList:
          ((00-01-66),127,  4) ((00-01-66),126,  4) ((00-01-66),125,  4) ((00-01-66),124,  4) ((00-01-66),123,
 4)
          ((00-01-66),122,  4) ((00-01-66),121,  4) ((00-01-66),120,  4) ((00-01-66),119,  4) ((00-01-66),118,
```

```
 4)
   ------------------------------------------------------------------------
```

Configuring LLDPDU Intervals

LLDPDUs are transmitted periodically; the default interval is **30 seconds**.

To configure LLDPDU intervals, use the following command.

- Configure a non-default transmit interval.

  CONFIGURATION mode or INTERFACE mode

  `hello`

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  no disable
R1(conf-lldp)#hello ?
<5-180>                 Hello interval in seconds (default=30)
R1(conf-lldp)#hello 25
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  hello 25
  no disable
R1(conf-lldp)#no hello
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  no disable
R1(conf-lldp)#
```

Configuring Transmit and Receive Mode

After you enable LLDP, the system transmits *and* receives LLDPDUs by default.

To configure the system to transmit or receive only and return to the default, use the following commands.

- Transmit only.

  CONFIGURATION mode or INTERFACE mode

  `mode tx`

- Receive only.

  CONFIGURATION mode or INTERFACE mode

  `mode rx`

- Return to the default setting.

  CONFIGURATION mode or INTERFACE mode

  `no mode`

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
```

```
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  no disable
R1(conf-lldp)#mode ?
rx              Rx only
tx              Tx only
R1(conf-lldp)#mode tx
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  mode tx
  no disable
R1(conf-lldp)#no mode
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  no disable
R1(conf-lldp)#
```

Configuring the Time to Live Value

The information received from a neighbor expires after a specific amount of time (measured in seconds) called a time to live (TTL).

The TTL is the product of the LLDPDU transmit interval (hello) and an integer called a multiplier. The default multiplier is **4**, which results in a default TTL of 120 seconds.

- Adjust the TTL value.

  CONFIGURATION mode or INTERFACE mode.

  `multiplier`
- Return to the default multiplier value.

  CONFIGURATION mode or INTERFACE mode.

  `no multiplier`

```
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  no disable
R1(conf-lldp)#multiplier ?
<2-10>            Multiplier (default=4)
R1(conf-lldp)#multiplier 5
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  multiplier 5
  no disable
R1(conf-lldp)#no multiplier
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot1-tlv port-protocol-vlan-id port-vlan-id
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  no disable
R1(conf-lldp)#
```

Debugging LLDP

You can view the TLVs that your system is sending and receiving.

To view the TLVs, use the following commands.

- View a readable version of the TLVs.
  `debug lldp brief`
- View a readable version of the TLVs plus a hexadecimal version of the entire LLDPDU, including unrecognized TLVs.
  `debug lldp detail`

  To stop viewing the LLDP TLVs sent and received by the system, use the `no debug lldp` command.

Figure 84. The debug lldp detail Command — LLDPDU Packet Dissection

Example of `debug lldp` Command Output with Unrecognized Reserved and Organizational Specific LLDP TLVs

The following is an example of LLDPDU with both (Reserved and Organizational specific) unrecognized TLVs.

```
DellEMC#Dec  4 22:38:27 : Received LLDP pkt on TenGigabitEthernet 1/1 of length 204 :
Dec  4 22:38:27 : Packet dump:
Dec  4 22:38:27 :  01 80 c2 00 00 0e 00 a0 c9 00 00 03 81 00 00 00
Dec  4 22:38:27 :  88 cc 02 07 04 00 a0 c9 00 00 01 04 02 05 54 06
Dec  4 22:38:27 :  02 01 2c fe 05 aa bb cc 04 61 fa 01 40 00 00 00
Dec  4 22:38:28 :  00 00 00 00 00 00 00 00 c6 0f ba 27
Dec  4 22:38:28 : TLV: Chassis ID, Len: 7, Subtype: Mac address (4) Value:
00:a0:c9:00:00:01
Dec  4 22:38:29 : TLV: Port ID, Len: 2, Subtype: Interface name (5) Value: T
Dec  4 22:38:29 : TLV: TTL, Len: 2, Value: 300
Dec  4 22:38:29 : TLV: UNKNOWN TLV, ORG_SPEC[aa-bb-cc, 4], Len: 1, Value:a
Dec  4 22:38:29 :  aa bb cc 04 61
Dec  4 22:38:29 :  40
Dec  4 22:38:29 : TLV: UNKNOWN TLV, Type: 125 Len: 1, Value: @
Dec  4 22:38:29 : TLV: ENDOFPDU, Len: 0
```

Relevant Management Objects

Dell EMC Networking OS supports all IEEE 802.1AB MIB objects.

The following tables list the objects associated with:

- received and transmitted TLVs
- the LLDP configuration on the local agent
- IEEE 802.1AB Organizationally Specific TLVs
- received and transmitted LLDP-MED TLVs

Table 59. LLDP Configuration MIB Objects

MIB Object Category	LLDP Variable	LLDP MIB Object	Description
LLDP Configuration	adminStatus	lldpPortConfigAdminStatus	Whether you enable the local LLDP agent for transmit, receive, or both.
	msgTxHold	lldpMessageTxHoldMultiplier	Multiplier value.
	msgTxInterval	lldpMessageTxInterval	Transmit Interval value.
	rxInfoTTL	lldpRxInfoTTL	Time to live for received TLVs.
	txInfoTTL	lldpTxInfoTTL	Time to live for transmitted TLVs.
Basic TLV Selection	mibBasicTLVsTxEnable	lldpPortConfigTLVsTxEnable	Indicates which management TLVs are enabled for system ports.
	mibMgmtAddrInstanceTxEnable	lldpManAddrPortsTxEnable	The management addresses defined for the system and the ports through which they are enabled for transmission.
LLDP Statistics	statsAgeoutsTotal	lldpStatsRxPortAgeoutsTotal	Total number of times that a neighbor's information is deleted on the local system due to an rxInfoTTL timer expiration.
	statsFramesDiscardedTotal	lldpStatsRxPortFramesDiscardedTotal	Total number of LLDP frames received then discarded.
	statsFramesInErrorsTotal	lldpStatsRxPortFramesErrors	Total number of LLDP frames received on a port with errors.
	statsFramesInTotal	lldpStatsRxPortFramesTotal	Total number of LLDP frames received through the port.
	statsFramesOutTotal	lldpStatsTxPortFramesTotal	Total number of LLDP frames transmitted through the port.
	statsTLVsDiscardedTotal	lldpStatsRxPortTLVsDiscardedTotal	Total number of TLVs received then discarded.
	statsTLVsUnrecognizedTotal	lldpStatsRxPortTLVsUnrecognizedTotal	Total number of all TLVs the local agent does not recognize.

Table 60. LLDP System MIB Objects

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
1	Chassis ID	chassis ID subtype	Local	lldpLocChassisIdSubtype
			Remote	lldpRemChassisIdSubtype
		chassid ID	Local	lldpLocChassisId
			Remote	lldpRemChassisId

Table 60. LLDP System MIB Objects (continued)

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
2	Port ID	port subtype	Local	lldpLocPortIdSubtype
			Remote	lldpRemPortIdSubtype
		port ID	Local	lldpLocPortId
			Remote	lldpRemPortId
4	Port Description	port description	Local	lldpLocPortDesc
			Remote	lldpRemPortDesc
5	System Name	system name	Local	lldpLocSysName
			Remote	lldpRemSysName
6	System Description	system description	Local	lldpLocSysDesc
			Remote	lldpRemSysDesc
7	System Capabilities	system capabilities	Local	lldpLocSysCapSupported
			Remote	lldpRemSysCapSupported
8	Management Address	enabled capabilities	Local	lldpLocSysCapEnabled
			Remote	lldpRemSysCapEnabled
		management address length	Local	lldpLocManAddrLen
			Remote	lldpRemManAddrLen
		management address subtype	Local	lldpLocManAddrSubtype
			Remote	lldpRemManAddrSubtype
		management address	Local	lldpLocManAddr
			Remote	lldpRemManAddr
		interface numbering subtype	Local	lldpLocManAddrIfSubtype
			Remote	lldpRemManAddrIfSubtype
		interface number	Local	lldpLocManAddrIfId
			Remote	lldpRemManAddrIfId
		OID	Local	lldpLocManAddrOID
			Remote	lldpRemManAddrOID

Table 61. LLDP 802.1 Organizationally specific TLV MIB Objects

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
127	Port-VLAN ID	PVID	Local	lldpXdot1LocPortVlanId
			Remote	lldpXdot1RemPortVlanId
127	Port and Protocol VLAN ID	port and protocol VLAN supported	Local	lldpXdot1LocProtoVlanSupported
			Remote	lldpXdot1RemProtoVlanSupported

Table 61. LLDP 802.1 Organizationally specific TLV MIB Objects (continued)

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
		port and protocol VLAN enabled	Local	lldpXdot1LocProtoVlanEnabled
			Remote	lldpXdot1RemProtoVlanEnabled
		PPVID	Local	lldpXdot1LocProtoVlanId
			Remote	lldpXdot1RemProtoVlanId
127	VLAN Name	VID	Local	lldpXdot1LocVlanId
			Remote	lldpXdot1RemVlanId
		VLAN name length	Local	lldpXdot1LocVlanName
			Remote	lldpXdot1RemVlanName
		VLAN name	Local	lldpXdot1LocVlanName
			Remote	lldpXdot1RemVlanName

Table 62. LLDP-MED System MIB Objects

TLV Sub-Type	TLV Name	TLV Variable	System	LLDP-MED MIB Object
1	LLDP-MED Capabilities	LLDP-MED Capabilities	Local	lldpXMedPortCapSupported lldpXMedPortConfigTLVsTx Enable
			Remote	lldpXMedRemCapSupported lldpXMedRemConfigTLVsTxEnable
		LLDP-MED Class Type	Local	lldpXMedLocDeviceClass
			Remote	lldpXMedRemDeviceClass
2	Network Policy	Application Type	Local	lldpXMedLocMediaPolicyAppType
			Remote	lldpXMedRemMediaPolicyAppType
		Unknown Policy Flag	Local	lldpXMedLocMediaPolicyUnknown
			Remote	lldpXMedLocMediaPolicyUnknown
		Tagged Flag	Local	lldpXMedLocMediaPolicyTagged
			Remote	lldpXMedLocMediaPolicyTagged
		VLAN ID	Local	lldpXMedLocMediaPolicyVlanID
			Remote	lldpXMedRemMediaPolicyVlanID
		L2 Priority	Local	lldpXMedLocMediaPolicyPriority

Table 62. LLDP-MED System MIB Objects (continued)

TLV Sub-Type	TLV Name	TLV Variable	System	LLDP-MED MIB Object
			Remote	lldpXMedRemMediaPolicyPriority
		DSCP Value	Local	lldpXMedLocMediaPolicyDscp
			Remote	lldpXMedRemMediaPolicyDscp
3	Location Identifier	Location Data Format	Local	lldpXMedLocLocationSubtype
			Remote	lldpXMedRemLocationSubtype
		Location ID Data	Local	lldpXMedLocLocationInfo
			Remote	lldpXMedRemLocationInfo
4	Extended Power via MDI	Power Device Type	Local	lldpXMedLocXPoEDeviceType
			Remote	lldpXMedRemXPoEDeviceType
		Power Source	Local	lldpXMedLocXPoEPSEPowerSource lldpXMedLocXPoEPDPowerSource
			Remote	lldpXMedRemXPoEPSEPowerSource lldpXMedRemXPoEPDPowerSource
		Power Priority	Local	lldpXMedLocXPoEPDPowerPriority lldpXMedLocXPoEPSEPortPDPriority
			Remote	lldpXMedRemXPoEPSEPowerPriority lldpXMedRemXPoEPDPowerPriority
		Power Value	Local	lldpXMedLocXPoEPSEPortPowerAv lldpXMedLocXPoEPDPowerReq
			Remote	lldpXMedRemXPoEPSEPowerAv lldpXMedRemXPoEPDPowerReq

30

Microsoft Network Load Balancing

Network load balancing (NLB) is a clustering functionality that is implemented by Microsoft on Windows 2000 Server and Windows Server 2003 operating systems (OSs). NLB uses a distributed methodology or pattern to equally split and balance the network traffic load across a set of servers that are part of the cluster or group. NLB combines the servers into a single multicast group and attempts to use the standard multicast IP or unicast IP addresses and MAC addresses to transmit of network traffic. At the same time, NLB also uses a single virtual IP address for all clients as the destination IP address, which enables servers to join the same multicast group that is transparent to the clients (the clients do not notice the addition of new servers to the group). The clients use a cluster IP address to connect to the server. To avoid overloading and effective performance of the servers for optimal processing of data packets, NLB enables flooding of traffic over the virtual local area network (VLAN) ports (for Unicast mode) or a subset of ports in a VLAN (for Multicast mode.)

NLB functions in two modes, Unicast mode and Multicast mode. Configure the cluster IP address and the associated cluster MAC address in the NLB application running on the Windows Server.

- In Unicast mode, when the server IP address attempts to be resolved to the MAC address using the address resolution protocol (ARP), the switch determines whether the ARP reply, obtained from the server, is an NLB type. The switch then maps the IP address (cluster IP) with the MAC address (cluster MAC address).
- In Multicast mode, the cluster IP address maps to a cluster multicast MAC address you configured using a `static ARP` command. After the NLB entry is learned, the traffic forwards to all the servers in the VLAN corresponding to the cluster virtual IP address.

NLB Unicast Mode Scenario

Consider a topology in which you configure four servers, S1 through S4, as a cluster or a farm. This set of servers connects to a Layer 3 switch, which connects to the end-clients. The servers contain a single IP address (IP-cluster address of 172.16.2.20) and a single unicast MAC address (MAC-Cluster address of 00-bf-ac-10-00-01) for load-balancing. Because multiple ports on a switch cannot learn a single MAC address, the servers are assigned MAC addresseses of MAC-s1 to MAC-s4), respectively, on S1 through S4 in addition to the MAC cluster address. All the servers of the cluster belong to VLAN1.

In Unicast NLB mode, the following sequence of events occurs:

- The switch sends an ARP request to resolve the IP address to the cluster MAC address.
- The ARP servers send an ARP response with the MAC cluster address in the ARP header and a MAC address of MAC-s1/s2/s3/s4 (for servers S1 through S4) in the Ethernet header.
- The switch associates the IP address with the MAC cluster address with the last ARP response it obtains. Assume that the last ARP reply is obtained from MAC-s4 (assuming that the ARP response with MAC-s4 is received as the last one). The interface associated with server, S4, is added to the ARP table.
- With NLB enabled, after learning the NLB ARP entry, all the subsequent traffic floods on all ports in VLAN1.

To perform load-balancing, NLB, forwards the data frame to all the servers.

NLB Multicast Mode Scenario

Consider a topology in which you configure four servers, S1 through S4, as a cluster or a farm. This set of servers connects to a Layer 3 switch, which connects to the end-clients. They contain a single multicast MAC address (MAC-Cluster: 03-00-5E-11-11-11).

In Multicast NLB mode, configure a `static ARP configuration` command to associate the cluster IP address with a multicast cluster MAC address.

With Multicast NLB mode, the data forwards to all the servers based on the port specified using the following Layer 2 multicast command in CONFIGURATION MODE:

```
mac-address-table static <multicast_mac> multicast vlan <vlan_id> output-range <port1>,
<port2>
```

Limitations of the NLB Feature

The following limitations apply to switches on which you configure NLB:

- The NLB Unicast mode uses switch flooding to transmit all packets to all the servers that are part of the VLAN. When a large volume of traffic is processed, the clustering performance might be impacted in a small way. This limitation is applicable to switches that perform unicast flooding in the software.
- The `ip vlan-flooding` command applies globally across the system and for all VLANs. In cases where NLB is applicable and ARP replies contain a discrepancy in the Ethernet SHA and ARP header SHA frames, a flooding of packets over the relevant VLAN occurs.
- The maximum number of concurrent clusters that is supported is eight.

Microsoft Clustering

To provide transparent failover or balancing, Microsoft clustering allows multiple servers using Microsoft Windows to be represented by one MAC address and IP address. The Dell EMC Networking OS does not recognize server clusters by default; you must configure it to do so. When an ARP request is sent to a server cluster, either the active server or all the servers send a reply, depending on the cluster configuration. If the active server sends a reply, the Dell EMC switch learns the active server's MAC address. If all servers reply, the switch registers only the last received ARP reply and the switch learns one server's actual MAC address; the virtual MAC address is never learned. Because the virtual MAC address is never learned, traffic is forwarded to only one server rather than the entire cluster, and failover and balancing are not preserved.

To preserve failover and balancing, the switch forwards the traffic destined for the server cluster to all member ports in the VLAN connected to the cluster. To ensure that this happens, use the `ip vlan-flooding` command on the Dell EMC switch when you configure the Microsoft cluster. The server MAC address is given in the Ethernet frame header of the ARP reply, while the virtual MAC address representing the cluster is given in the payload. Then, all the traffic destined for the cluster is flooded out of all member ports. Because all the servers in the cluster receive traffic, failover and balancing are preserved.

Enable and Disable VLAN Flooding

- The older ARP entries are overwritten whenever newer NLB entries are learned.
- All ARP entries, learned after you enable VLAN flooding, are deleted when you disable VLAN flooding, and RP2 triggers an ARP resolution. Disable VLAN flooding with the `no ip vlan-flooding` command.
- When you add a port to the VLAN, the port automatically receives traffic if you enabled VLAN flooding. Old ARP entries are not deleted or updated.
- When you delete a member port, its ARP entries are also deleted from the content addressable memory (CAM).
- Port channels in the VLAN also receive traffic.
- There is no impact on the configuration from saving the configuration.
- If you enable VLAN flooding, it displays in the `show running-config` command output that displays the `ip vlan-flooding` CLI configuration. This is the only output where you see the VLAN flooding status (enabled or disabled).

Topics:

- Configuring a Switch for NLB

Configuring a Switch for NLB

To enable a switch for Unicast NLB mode, perform the following steps:

- Enter the `ip vlan-flooding` command to specify that all Layer 3 unicast routed data traffic going through a VLAN member port floods across all the member ports of that VLAN.

  CONFIGURATION mode

  ```
  ip vlan-flooding
  ```

There might be some ARP table entries that are resolved through ARP packets, which had the Ethernet MAC SA different from the MAC information inside the ARP packet. This unicast data traffic flooding occurs only for those packets that use these ARP entries.

Enabling a Switch for Multicast NLB

To enable a switch for Multicast NLB mode, perform the following steps:

1. Add a static ARP entry by entering the `arp ip-address multicast-mac-address` command in the Global configuration mode to associate an IP address with a multicast MAC address in the switch.

   CONFIGURATION mode

   `arp ip-address multicast-mac-address interface`

   This setting causes the multicast MAC address to be mapped to the Cluster IP address for the NLB mode of operation of the switch.

   NOTE: While configuring static ARP for the Cluster IP, provide any one of the interfaces that is used in the static multicast MAC configuration, where the Cluster host is connected. As the switch does not accept only one ARP-interface pair, if you configure static ARP with each egress interface, the switch overwrites the previous egress-interface configuration.

2. Associate specific MAC or hardware addresses to VLANs.

   CONFIGURATION mode

   `mac-address-table static multicast-mac-address vlan vlan-id output-range interface`

   NOTE: When you use the `mac-address-table static multicast-mac-address` command in a VLT setup, Dell EMC Networking OS recommends to add VLTi as one of the egress interfaces along with other cluster facing interfaces.

31

Multicast Source Discovery Protocol (MSDP)

Multicast source discovery protocol (MSDP) is supported on Dell EMC Networking OS.

Protocol Overview

MSDP is a Layer 3 protocol that connects IPv4 protocol-independent multicast-sparse mode (PIM-SM) domains. A domain in the context of MSDP is a contiguous set of routers operating PIM within a common boundary defined by an exterior gateway protocol, such as border gateway protocol (BGP).

Each rendezvous point (RP) peers with every other RP via the transmission control protocol (TCP). Through this connection, peers advertise the sources in their domain.

1. When an RP in a PIM-SM domain receives a PIM register message from a source, it sends a source-active (SA) message to MSDP peers, as shown in the following illustration.
2. Each MSDP peer receives and forwards the message to its peers away from the originating RP.
3. When an MSDP peer receives an SA message, it determines if there are any group members within the domain interested in any of the advertised sources. If there are, the receiving RP sends a join message to the originating RP, creating a shortest path tree (SPT) to the source.

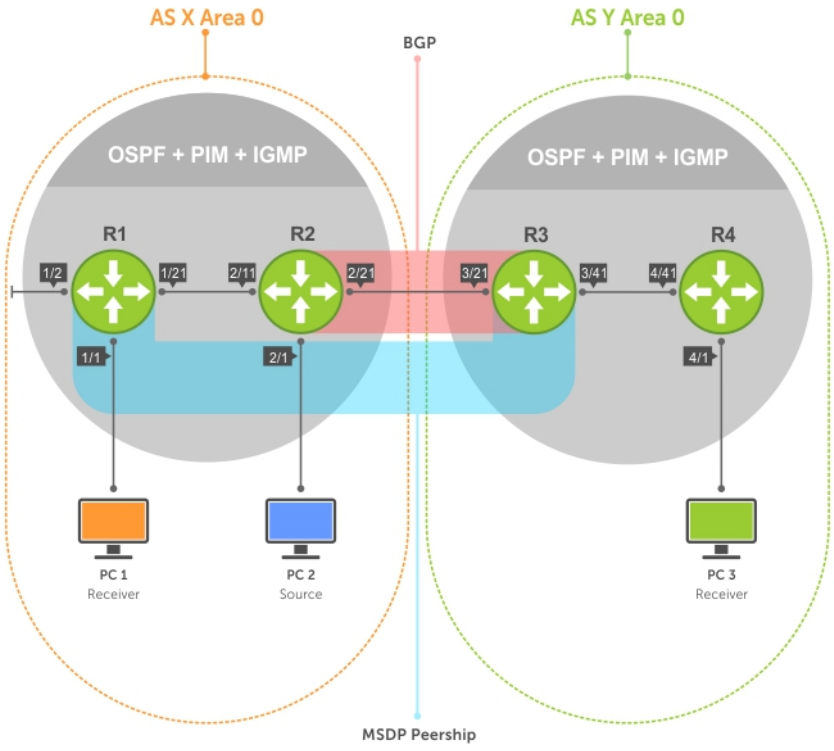


Figure 85. Multicast Source Discovery Protocol (MSDP)

RPs advertise each (S,G) in its domain in type, length, value (TLV) format. The total number of TLVs contained in the SA is indicated in the “Entry Count” field. SA messages are transmitted every 60 seconds, and immediately when a new source is detected.

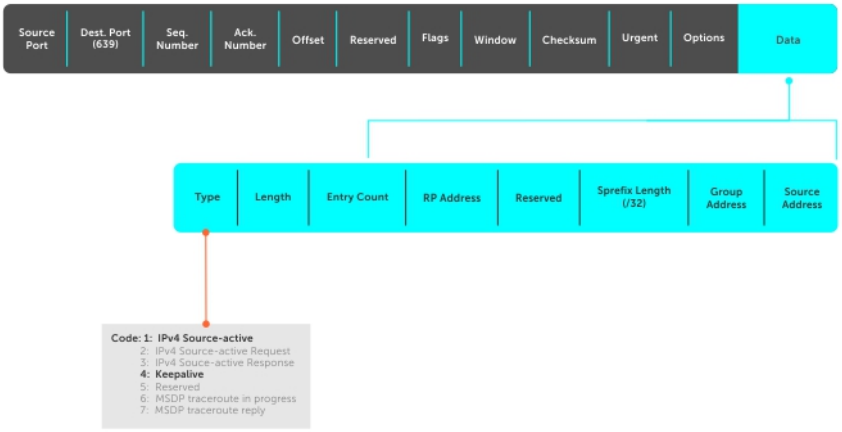


Figure 86. MSDP SA Message Format

Topics:

- Anycast RP
- Implementation Information
- Configure Multicast Source Discovery Protocol
- Enable MSDP
- Manage the Source-Active Cache
- Accept Source-Active Messages that Fail the RFP Check
- Specifying Source-Active Messages
- Limiting the Source-Active Messages from a Peer
- Preventing MSDP from Caching a Local Source
- Preventing MSDP from Caching a Remote Source
- Preventing MSDP from Advertising a Local Source
- Logging Changes in Peership States
- Terminating a Peership
- Clearing Peer Statistics
- Debugging MSDP
- MSDP with Anycast RP
- Configuring Anycast RP
- MSDP Sample Configurations

Anycast RP

Using MSDP, anycast RP provides load sharing and redundancy in PIM-SM networks. Anycast RP allows two or more rendezvous points (RPs) to share the load for source registration and the ability to act as hot backup routers for each other.

Anycast RP allows you to configure two or more RPs with the same IP address on Loopback interfaces. The Anycast RP Loopback address are configured with a 32-bit mask, making it a host address. All downstream routers are configured to know that the Anycast RP Loopback address is the IP address of their local RP. IP routing automatically selects the closest RP for each source and receiver. Assuming that the sources are evenly spaced around the network, an equal number of sources register with each RP. Consequently, all the RPs in the network share the process of registering the sources equally. Because a source may register with one RP and receivers may join to a different RP, a method is needed for the RPs to exchange information about active sources. This information exchange is done with MSDP.

With Anycast RP, all the RPs are configured to be MSDP peers of each other. When a source registers with one RP, an SA message is sent to the other RPs informing them that there is an active source for a particular multicast group. The result is that each RP is aware of the active sources in the area of the other RPs. If any of the RPs fail, IP routing converges and one of the RPs becomes the active RP in more than one area. New sources register with the backup RP. Receivers join toward the new RP and connectivity is maintained.

Implementation Information

The Dell EMC Networking OS implementation of MSDP is in accordance with RFC 3618 and Anycast RP is in accordance with RFC 3446.

Configure Multicast Source Discovery Protocol

Configuring MSDP is a four-step process.

1. Enable an exterior gateway protocol (EGP) with at least two routing domains.

   Refer to the following figures.

   The MSDP Sample Configurations show the OSPF-BGP configuration used in this chapter for MSDP. Also, refer to Open Shortest Path First (OSPFv2) and Border Gateway Protocol IPv4 (BGPv4).

2. Configure PIM-SM within each EGP routing domain.

   Refer to the following figures.

   The MSDP Sample Configurations show the PIM-SM configuration in this chapter for MSDP. Also, refer to PIM Sparse-Mode (PIM-SM).

3. Enable MSDP.
4. Peer the RPs in each routing domain with each other. Refer to Enable MSDP.

Related Configuration Tasks

The following lists related MSDP configuration tasks.

- Enable MSDP
- Manage the Source-Active Cache
- Accept Source-Active Messages that Fail the RFP Check
- Specifying Source-Active Messages
- Limiting the Source-Active Cache
- Preventing MSDP from Caching a Local Source
- Preventing MSDP from Caching a Remote Source
- Preventing MSDP from Advertising a Local Source
- Terminating a Peership
- Clearing Peer Statistics
- Debugging MSDP
- MSDP with Anycast RP
- MSDP Sample Configurations

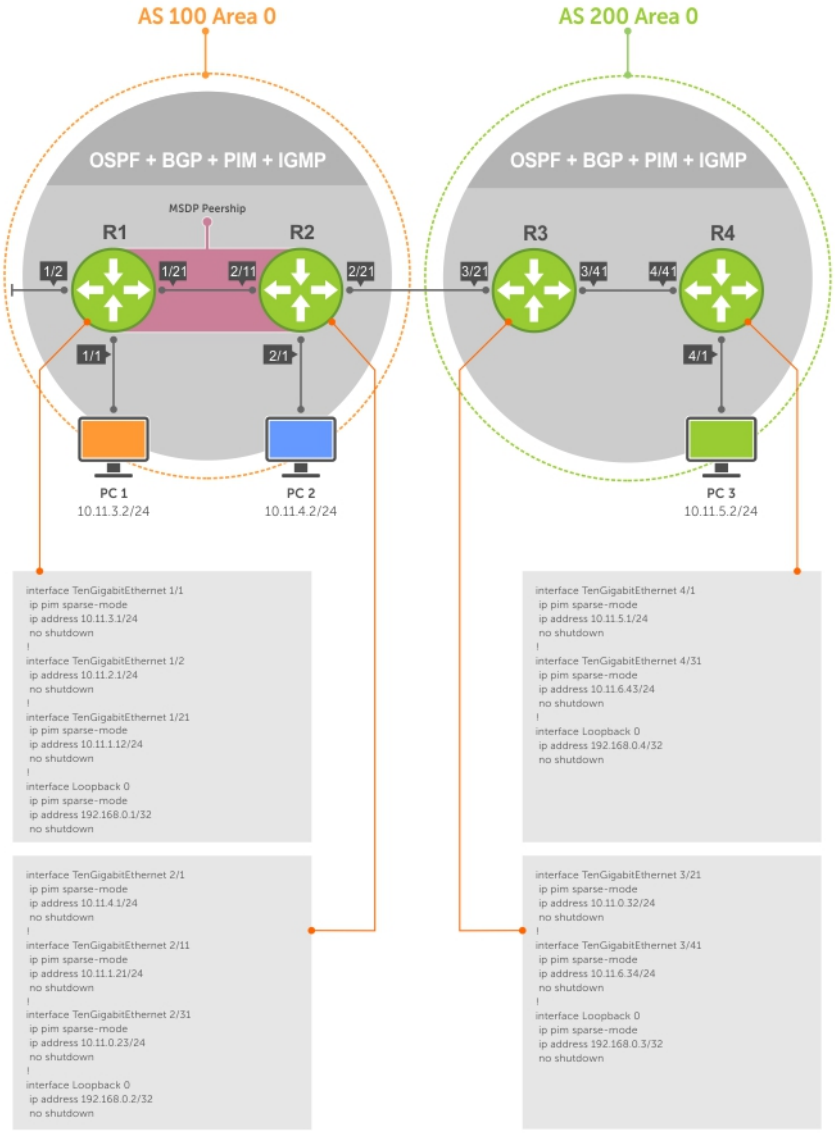


Figure 87. Configuring Interfaces for MSDP

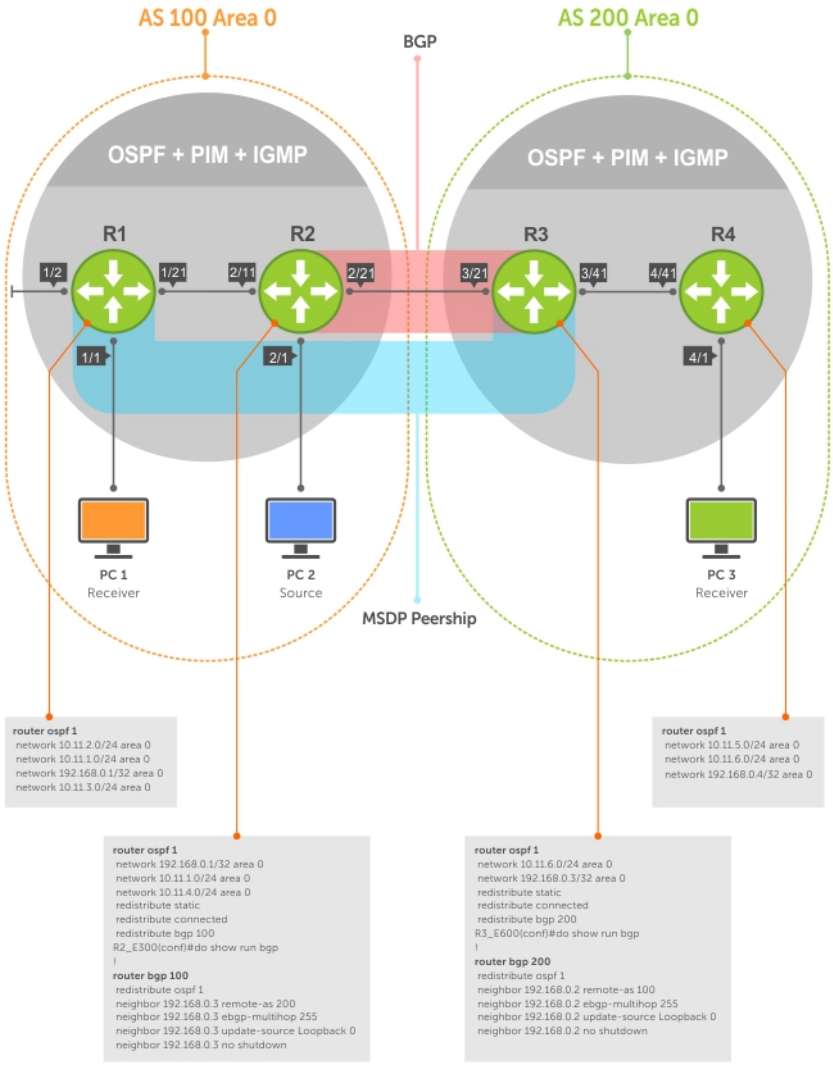


Figure 88. Configuring OSPF and BGP for MSDP

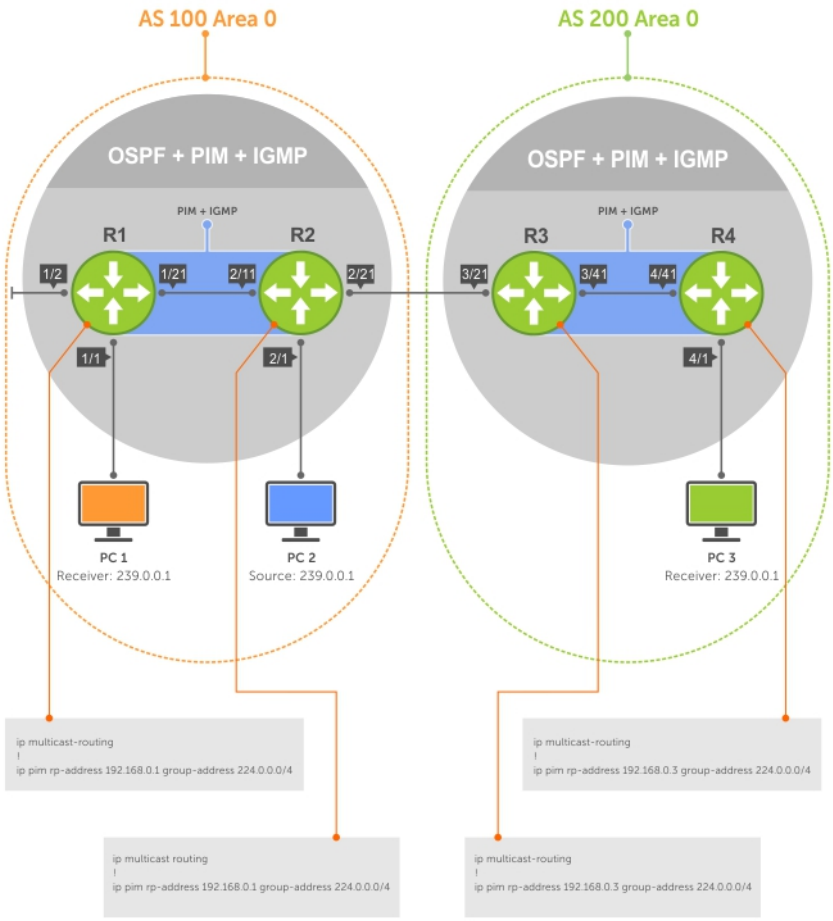


Figure 89. Configuring PIM in Multiple Routing Domains

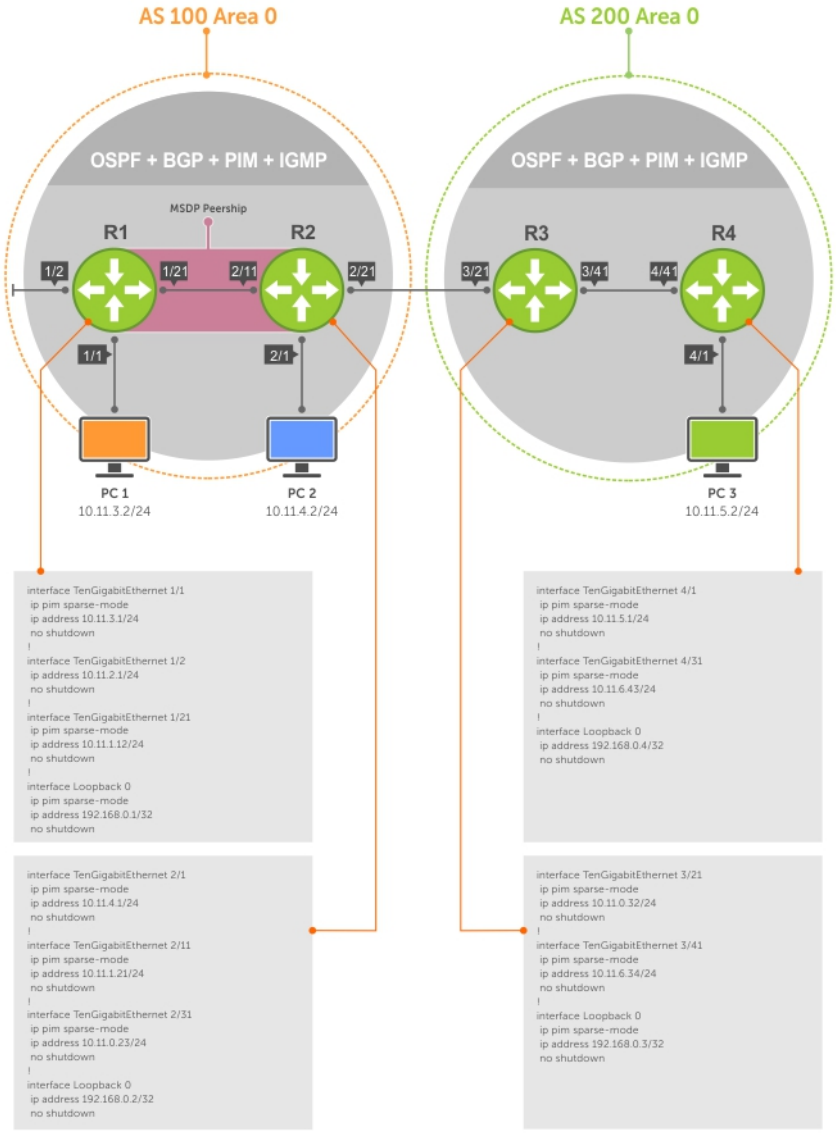


Figure 90. Configuring MSDP

Enable MSDP

Enable MSDP by peering RPs in different administrative domains.

1. Enable MSDP.
   CONFIGURATION mode
   `ip multicast-msdp`
2. Peer PIM systems in different administrative domains.
   CONFIGURATION mode
   `ip msdp peer connect-source`

```
    R3(conf)#ip multicast-msdp
    R3(conf)#ip msdp peer 192.168.0.1 connect-source Loopback 0
    R3(conf)#do show ip msdp summary

    Peer Addr    Local Addr    State    Source    SA    Up/Down
Description
```

To view details about a peer, use the `show ip msdp peer` command in EXEC privilege mode.

Multicast sources in remote domains are stored on the RP in the source-active cache (SA cache). The system does not create entries in the multicast routing table until there is a local receiver for the corresponding multicast group.

```
R3#show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 192.168.0.3(639) Connect Source: Lo 0
    State: Established Up/Down Time: 00:15:20
    Timers: KeepAlive 30 sec, Hold time 75 sec
    SourceActive packet count (in/out): 8/0
    SAs learned from this peer: 1
    SA Filtering:
    Input (S,G) filter: none
    Output (S,G) filter: none
```

Manage the Source-Active Cache

Each SA-originating RP caches the sources inside its domain (domain-local), and the sources which it has learned from its peers (domain-remote).

By caching sources:

- domain-local receivers experience a lower join latency
- RPs can transmit SA messages periodically to prevent SA storms
- only sources that are in the cache are advertised in the SA to prevent transmitting multiple copies of the same source information

Viewing the Source-Active Cache

To view the source-active cache, use the following command.

- View the SA cache.

  EXEC Privilege mode

  `show ip msdp sa-cache`

```
R3#show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr  SourceAddr  RPAddr       LearnedFrom  Expire UpTime
239.0.0.1  10.11.4.2   192.168.0.1  192.168.0.1  76     00:10:44
```

Limiting the Source-Active Cache

Set the upper limit of the number of active sources that the Dell EMC Networking OS caches.

The default active source limit is 500K messages. When the total number of active sources reaches the specified limit, subsequent active sources are dropped even if they pass the reverse path forwarding (RPF) and policy check.

To limit the number of sources that SA cache stores, use the following command.

- Limit the number of sources that can be stored in the SA cache.

  EXEC Privilege mode

  `show ip msdp sa-limit`

If the total number of active sources is already larger than the limit when limiting is applied, the sources that are already in Dell EMC Networking OS are not discarded. To enforce the limit in such a situation, use the `clear ip msdp sa-cache` command to clear all existing entries.

Clearing the Source-Active Cache

To clear the source-active cache, use the following command.

- Clear the SA cache of all, local, or rejected entries, or entries for a specific group.

CONFIGURATION mode

```
clear ip msdp sa-cache [group-address | local | rejected-sa]
```

Enabling the Rejected Source-Active Cache

To cache rejected sources, use the following command.

Active sources can be rejected because the RPF check failed, the SA limit is reached, the peer RP is unreachable, or the SA message has a format error.

- Cache rejected sources.

  CONFIGURATION mode

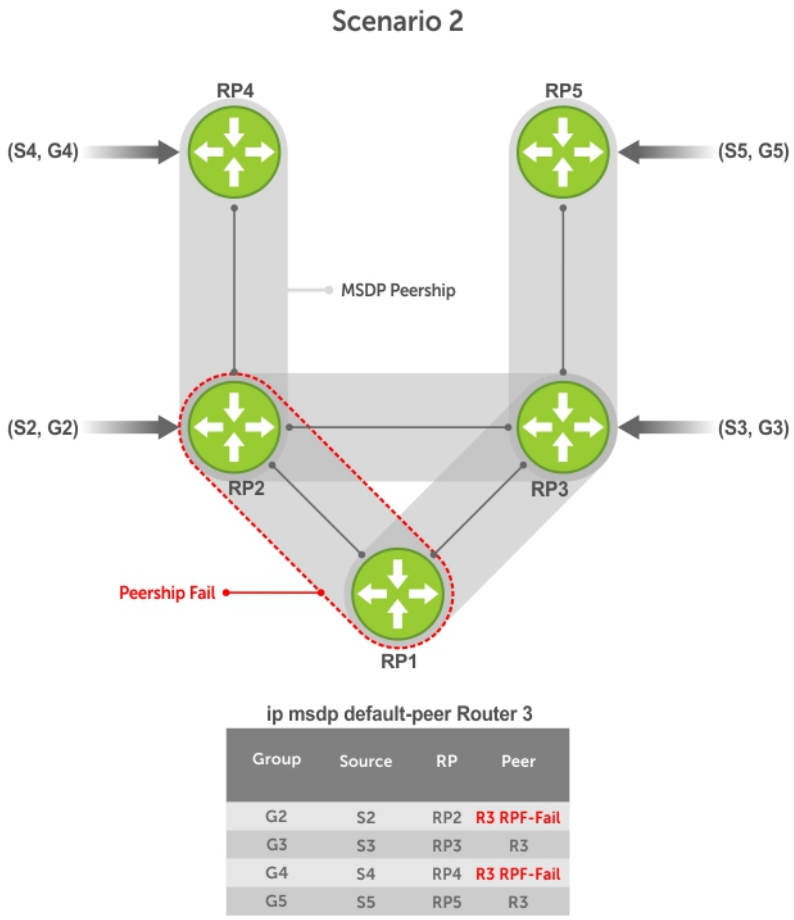
  ```
  ip msdp cache-rejected-sa
  ```

Accept Source-Active Messages that Fail the RFP Check

A default peer is a peer from which active sources are accepted even though they fail the RFP check.

Referring to the following illustrations:

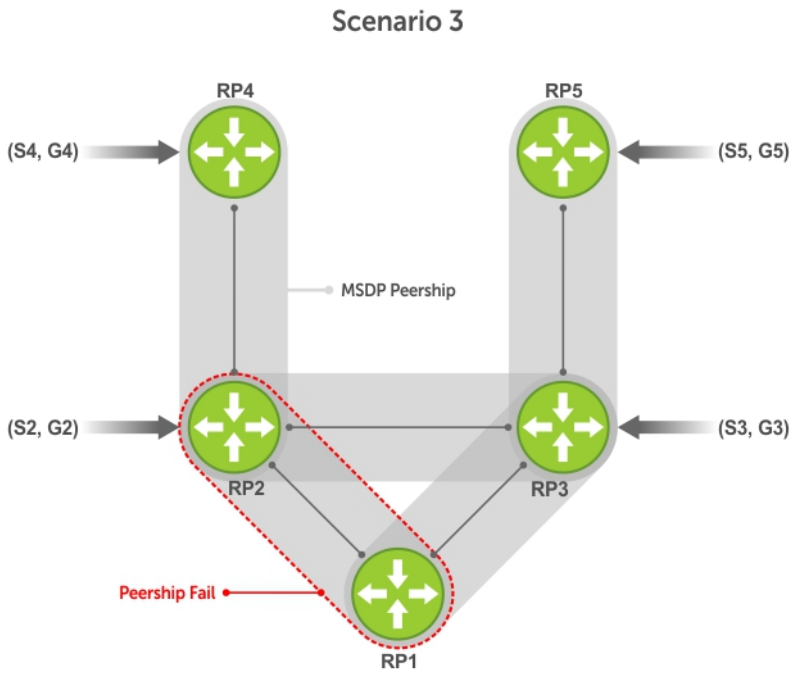
- In Scenario 1, all MSPD peers are up.
- In Scenario 2, the peership between RP1 and RP2 is down, but the link (and routing protocols) between them is still up. In this case, RP1 learns all active sources from RP3, but the sources from RP2 and RP4 are rejected because the reverse path to these routers is through Interface A.
- In Scenario 3, RP3 is configured as a default MSDP peer for RP1 and so the RPF check is disregarded for RP3.
- In Scenario 4, RP1 has a default peer plus an access list. The list permits RP4 so the RPF check is disregarded for active sources from it, but RP5 (and all others because of the implicit deny all) are subject to the RPF check and fail, so those active sources are rejected.



ip msdp default-peer Router 3

Group	Source	RP	Peer
G2	S2	RP2	R3 RPF-Fail
G3	S3	RP3	R3
G4	S4	RP4	R3 RPF-Fail
G5	S5	RP5	R3

Figure 91. MSDP Default Peer, Scenario 2



ip msdp default-peer Router 3

Group	Source	RP	Peer
G2	S2	RP2	R3
G3	S3	RP3	R3
G4	S4	RP4	R3
G5	S5	RP5	R3

Figure 92. MSDP Default Peer, Scenario 3

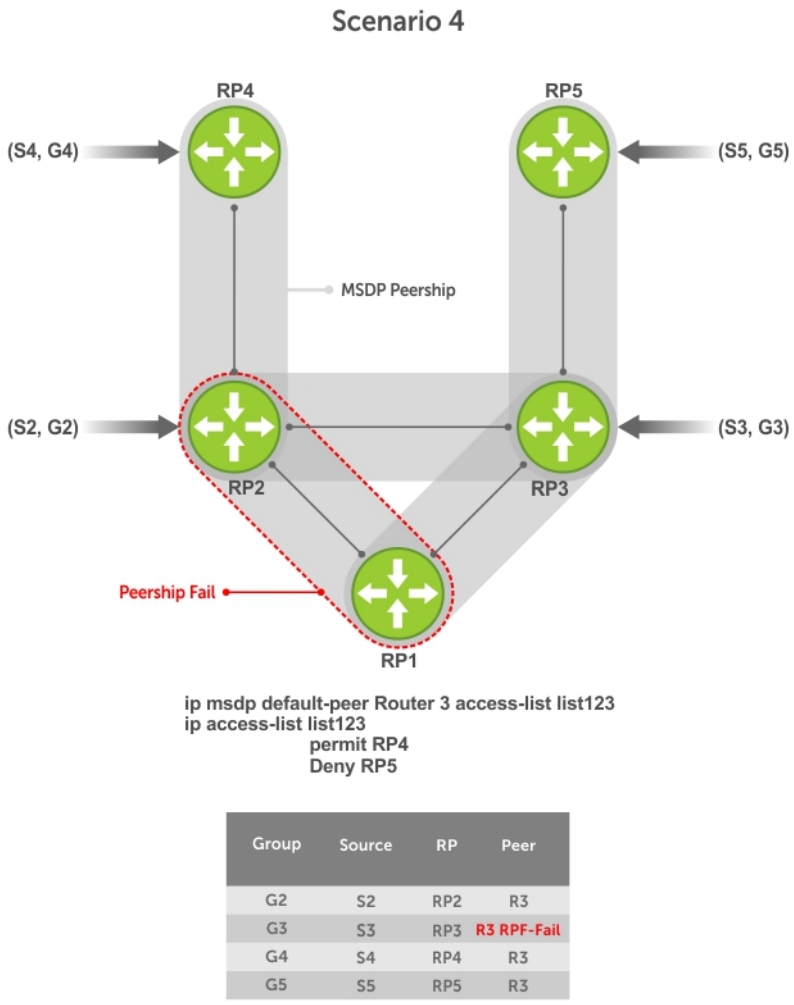


Figure 93. MSDP Default Peer, Scenario 4

Specifying Source-Active Messages

To specify messages, use the following command.

- Specify the forwarding-peer and originating-RP from which all active sources are accepted without regard for the RPF check.

  CONFIGURATION mode

  `ip msdp default-peer ip-address list`

  If you do not specify an access list, the peer accepts all sources that peer advertises. All sources from RPs that the ACL denies are subject to the normal RPF check.

```
DellEMC(conf)#ip msdp peer 10.0.50.2 connect-source Vlan 50
DellEMC(conf)#ip msdp default-peer 10.0.50.2 list fifty

DellEMC(conf)#ip access-list standard fifty
DellEMC(conf)#seq 5 permit host 200.0.0.50

DellEMC#ip msdp sa-cache
MSDP Source-Active Cache - 3 entries
GroupAddr    SourceAddr   RPAddr       LearnedFrom   Expire   UpTime
229.0.50.2   24.0.50.2    200.0.0.50   10.0.50.2     73       00:13:49
229.0.50.3   24.0.50.3    200.0.0.50   10.0.50.2     73       00:13:49
229.0.50.4   24.0.50.4    200.0.0.50   10.0.50.2     73       00:13:49
```

```
DellEMC#ip msdp sa-cache rejected-sa
MSDP Rejected SA Cache
  3 rejected SAs received, cache-size 32766
UpTime    GroupAddr    SourceAddr  RPAddr      LearnedFrom  Reason
00:33:18  229.0.50.64  24.0.50.64  200.0.1.50  10.0.50.2    Rpf-Fail
00:33:18  229.0.50.65  24.0.50.65  200.0.1.50  10.0.50.2    Rpf-Fail
00:33:18  229.0.50.66  24.0.50.66  200.0.1.50  10.0.50.2    Rpf-Fail
```

Limiting the Source-Active Messages from a Peer

To limit the source-active messages from a peer, use the following commands.

1. OPTIONAL: Store sources that are received after the limit is reached in the rejected SA cache.

   CONFIGURATION mode

   ```
   ip msdp cache-rejected-sa
   ```

2. Set the upper limit for the number of sources allowed from an MSDP peer.

   CONFIGURATION mode

   ```
   ip msdp peer peer-address sa-limit
   ```

   The default limit is **100K**.

If the total number of sources received from the peer is already larger than the limit when this configuration is applied, those sources are not discarded. To enforce the limit in such a situation, first clear the SA cache.

Preventing MSDP from Caching a Local Source

You can prevent MSDP from caching an active source based on source and/or group. Because the source is not cached, it is not advertised to remote RPs.

1. OPTIONAL: Cache sources that are denied by the redistribute list in the rejected SA cache.

   CONFIGURATION mode

   ```
   ip msdp cache-rejected-sa
   ```

2. Prevent the system from caching local SA entries based on source and group using an extended ACL.

   CONFIGURATION mode

   ```
   ip msdp redistribute list
   ```

When you apply this filter, the SA cache is not affected immediately. When sources that are denied by the ACL time out, they are not refreshed. Until they time out, they continue to reside in the cache. To apply the redistribute filter to entries already present in the SA cache, first clear the SA cache. You may optionally store denied sources in the rejected SA cache.

```
R1(conf)#do show run msdp
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 0
ip msdp redistribute list mylocalfilter
ip msdp cache-rejected-sa 1000
R1_E600(conf)#do show run acl
!
ip access-list extended mylocalfilter
  seq 5 deny ip host 239.0.0.1 host 10.11.4.2
  seq 10 deny ip any any
R1_E600(conf)#do show ip msdp sa-cache
R1_E600(conf)#do show ip msdp sa-cache rejected-sa
MSDP Rejected SA Cache
  1 rejected SAs received, cache-size 1000
UpTime    GroupAddr  SourceAddr  RPAddr       LearnedFrom  Reason
00:02:20  239.0.0.1  10.11.4.2   192.168.0.1  local        Redistribute
```

Preventing MSDP from Caching a Remote Source

To prevent MSDP from caching a remote source, use the following commands.

1. OPTIONAL: Cache sources that the SA filter denies in the rejected SA cache.

   CONFIGURATION mode

   ```
   ip msdp cache-rejected-sa
   ```

2. Prevent the system from caching remote sources learned from a specific peer based on source and group.

   CONFIGURATION mode

   ```
   ip msdp sa-filter list out peer list ext-acl
   ```

As shown in the following example, R1 is advertising source 10.11.4.2. It is already in the SA cache of R3 when an ingress SA filter is applied to R3. The entry remains in the SA cache until it expires and is *not* stored in the rejected SA cache.

```
[Router 3]
R3(conf)#do show run msdp
!
ip multicast-msdp
ip msdp peer 192.168.0.1 connect-source Loopback 0
ip msdp sa-filter in 192.168.0.1 list myremotefilter
R3(conf)#do show run acl
!
ip access-list extended myremotefilter
  seq 5 deny ip host 239.0.0.1 host 10.11.4.2
R3(conf)#do show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr  SourceAddr  RPAddr       LearnedFrom  Expire  UpTime
239.0.0.1  10.11.4.2   192.168.0.1  192.168.0.1  1       00:03:59
R3(conf)#do show ip msdp sa-cache
R3(conf)#
R3(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 0.0.0.0(639) Connect Source: Lo 0
    State: Listening Up/Down Time: 00:01:19
    Timers: KeepAlive 30 sec, Hold time 75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
```

Preventing MSDP from Advertising a Local Source

To prevent MSDP from advertising a local source, use the following command.

- Prevent an RP from advertising a source in the SA cache.

  CONFIGURATION mode

  ```
  ip msdp sa-filter list in peer list ext-acl
  ```

In the following example, R1 stops advertising source 10.11.4.2. Because it is already in the SA cache of R3, the entry remains there until it expires.

```
[Router 1]
R1(conf)#do show run msdp
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 0
ip msdp sa-filter out 192.168.0.3 list mylocalfilter
R1(conf)#do show run acl
!
ip access-list extended mylocalfilter
  seq 5 deny ip host 239.0.0.1 host 10.11.4.2
  seq 10 deny ip any any
R1(conf)#do show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
```

```
GroupAddr  SourceAddr  RPAddr       LearnedFrom  Expire  UpTime
239.0.0.1  10.11.4.2   192.168.0.1  local        70      00:27:20
R3(conf)#do show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr  SourceAddr  RPAddr       LearnedFrom  Expire  UpTime
239.0.0.1  10.11.4.2   192.168.0.1  192.168.0.1  1       00:10:29

[Router 3]
R3(conf)#do show ip msdp sa-cache
R3(conf)#
```

To display the configured SA filters for a peer, use the `show ip msdp peer` command from EXEC Privilege mode.

Logging Changes in Peership States

To log changes in peership states, use the following command.

- Log peership state changes.

  CONFIGURATION mode

  `ip msdp log-adjacency-changes`

Terminating a Peership

MSDP uses TCP as its transport protocol. In a peering relationship, the peer with the lower IP address initiates the TCP session, while the peer with the higher IP address listens on port 639.

- Terminate the TCP connection with a peer.

  CONFIGURATION mode

  `ip msdp shutdown`

After the relationship is terminated, the peering state of the terminator is SHUTDOWN, while the peering state of the peer is INACTIVE.

```
[Router 3]
R3(conf)#ip msdp shutdown 192.168.0.1
R3(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 0.0.0.0(0) Connect Source: Lo 0
    State: Shutdown Up/Down Time: 00:00:18
    Timers: KeepAlive 30 sec, Hold time 75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
   SA Filtering:
   Input (S,G) filter: myremotefilter
   Output (S,G) filter: none
[Router 1]
R1(conf)#do show ip msdp peer

Peer Addr: 192.168.0.3
    Local Addr: 0.0.0.0(0) Connect Source: Lo 0
    State: Inactive Up/Down Time: 00:00:03
    Timers: KeepAlive 30 sec, Hold time 75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
```

Clearing Peer Statistics

To clear the peer statistics, use the following command.

- Reset the TCP connection to the peer and clear all peer statistics.

  CONFIGURATION mode

```
clear ip msdp peer peer-address
```

```
R3(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 192.168.0.3(639) Connect Source: Lo 0
    State: Established Up/Down Time: 00:04:26
    Timers: KeepAlive 30 sec, Hold time 75 sec
    SourceActive packet count (in/out): 5/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
R3(conf)#do clear ip msdp peer 192.168.0.1
R3(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 0.0.0.0(0) Connect Source: Lo 0
    State: Inactive Up/Down Time: 00:00:04
    Timers: KeepAlive 30 sec, Hold time 75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
```

Debugging MSDP

To debug MSDP, use the following command.

- Display the information exchanged between peers.

  CONFIGURATION mode

  ```
  debug ip msdp
  ```

```
R1(conf)#do debug ip msdp
All MSDP debugging has been turned on
R1(conf)#03:16:08 : MSDP-0: Peer 192.168.0.3, sent Keepalive msg
03:16:09 : MSDP-0: Peer 192.168.0.3,  rcvd Keepalive msg
03:16:27 : MSDP-0: Peer 192.168.0.3,  sent Source Active msg
03:16:38 : MSDP-0: Peer 192.168.0.3,  sent Keepalive msg
03:16:39 : MSDP-0: Peer 192.168.0.3,  rcvd Keepalive msg
03:17:09 : MSDP-0: Peer 192.168.0.3,  sent Keepalive msg
03:17:10 : MSDP-0: Peer 192.168.0.3,  rcvd Keepalive msg
03:17:27 : MSDP-0: Peer 192.168.0.3,  sent Source Active msg
Input (S,G) filter: none
  Output (S,G) filter: none
```

MSDP with Anycast RP

Anycast RP uses MSDP with PIM-SM to allow more than one active group to use RP mapping.

PIM-SM allows only active groups to use RP mapping, which has several implications:

- **traffic concentration**: PIM-SM allows only one active group to RP mapping which means that all traffic for the group must, at least initially, travel over the same part of the network. You can load balance source registration between multiple RPs by strategically mapping groups to RPs, but this technique is less effective as traffic increases because preemptive load balancing requires prior knowledge of traffic distributions.
- **lack of scalable register decasulation**: With only a single RP per group, all joins are sent to that RP regardless of the topological distance between the RP, sources, and receivers, and data is transmitted to the RP until the SPT switch threshold is reached.
- **slow convergence when an active RP fails**: When you configure multiple RPs, there can be considerable convergence delay involved in switching to the backup RP.

Anycast RP relieves these limitations by allowing multiple RPs per group, which can be distributed in a topologically significant manner according to the locations of the sources and receivers.

1. All the RPs serving a given group are configured with an identical anycast address.
2. Sources then register with the topologically closest RP.
3. RPs use MSDP to peer with each other using a unique address.

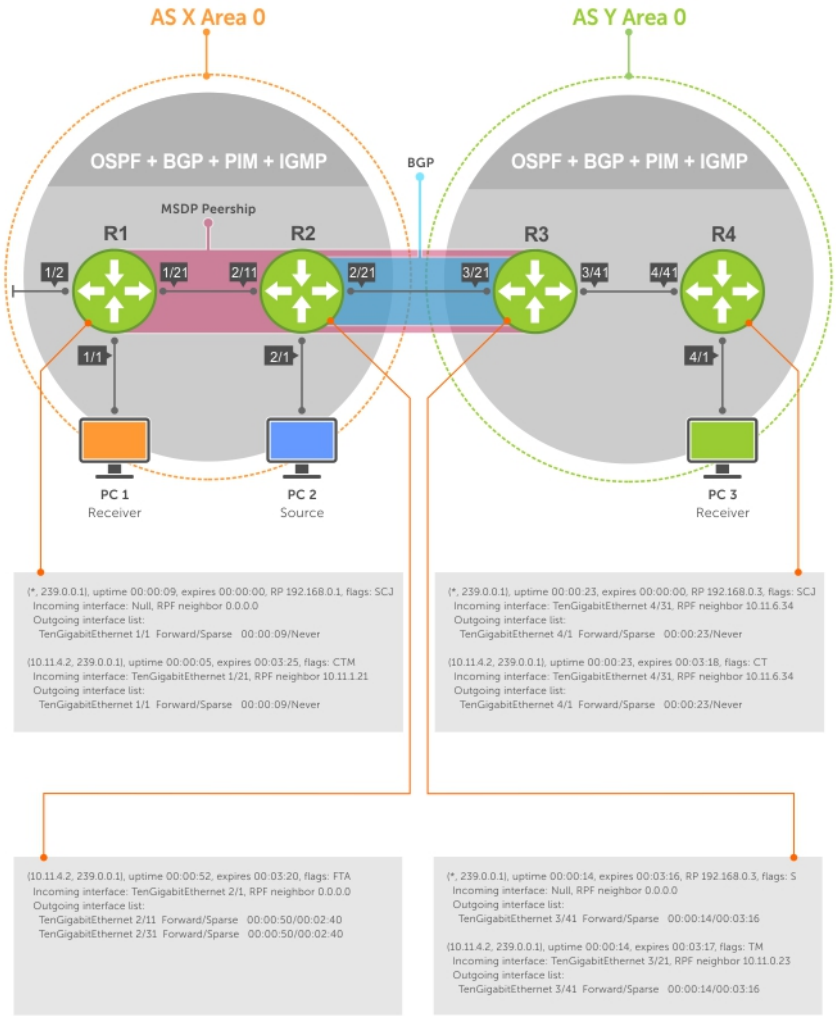


Figure 94. MSDP with Anycast RP

Configuring Anycast RP

To configure anycast RP, use the following commands.

1. In each routing domain that has multiple RPs serving a group, create a Loopback interface on each RP serving the group with the same IP address.

   CONFIGURATION mode

   `interface loopback`
2. Make this address the RP for the group.

   CONFIGURATION mode

   `ip pim rp-address`
3. In each routing domain that has multiple RPs serving a group, create another Loopback interface on each RP serving the group with a unique IP address.

   CONFIGURATION mode

   `interface loopback`
4. Peer each RP with every other RP using MSDP, specifying the unique Loopback address as the connect-source.

   CONFIGURATION mode

```
ip msdp peer
```

5. Advertise the network of each of the unique Loopback addresses throughout the network.

   ROUTER OSPF mode

   ```
   network
   ```

Reducing Source-Active Message Flooding

RPs flood source-active messages to all of their peers away from the RP.

When multiple RPs exist within a domain, the RPs forward received active source information back to the originating RP, which violates the RFP rule. You can prevent this unnecessary flooding by creating a mesh-group. A mesh in this context is a topology in which each RP in a set of RPs has a peership with all other RPs in the set. When an RP is a member of the mesh group, it forwards active source information only to its peers outside of the group.

To create a mesh group, use the following command.

- Create a mesh group.

  CONFIGURATION mode

  ```
  ip msdp mesh-group
  ```

Specifying the RP Address Used in SA Messages

The default originator-id is the address of the RP that created the message. In the case of Anycast RP, there are multiple RPs all with the same address.

To use the (unique) address of another interface as the originator-id, use the following command.

- Use the address of another interface as the originator-id instead of the RP address.

  CONFIGURATION mode

  ```
  ip msdp originator-id
  ```

The following example shows an R1 configuration for MSDP with Anycast RP.

```
ip multicast-routing
!
interface TenGigabitEthernet 1/1
  ip pim sparse-mode
  ip address 10.11.3.1/24
  no shutdown
!
interface TenGigabitEthernet 1/2
  ip address 10.11.2.1/24
  no shutdown
!
interface TenGigabitEthernet 1/21
  ip pim sparse-mode
  ip address 10.11.1.12/24
  no shutdown
!
interface Loopback 0
  ip pim sparse-mode
  ip address 192.168.0.1/32
  no shutdown
!
interface Loopback 1
  ip address 192.168.0.11/32
  no shutdown
!
router ospf 1
  network 10.11.2.0/24 area 0
  network 10.11.1.0/24 area 0
  network 10.11.3.0/24 area 0
  network 192.168.0.11/32 area 0
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 1
```

```
ip msdp peer 192.168.0.22 connect-source Loopback 1
ip msdp mesh-group AS100 192.168.0.22
ip msdp originator-id Loopback 1!

ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

The following example shows an R2 configuration for MSDP with Anycast RP.

```
ip multicast-routing
!
interface TenGigabitEthernet 2/1
  ip pim sparse-mode
  ip address 10.11.4.1/24
  no shutdown
!
interface TenGigabitEthernet 2/11
  ip pim sparse-mode
  ip address 10.11.1.21/24
  no shutdown
!
interface TenGigabitEthernet 2/31
  ip pim sparse-mode
  ip address 10.11.0.23/24
  no shutdown
!
interface Loopback 0
  ip pim sparse-mode
  ip address 192.168.0.1/32
  no shutdown
!
interface Loopback 1
  ip address 192.168.0.22/32
  no shutdown
!
router ospf 1
  network 10.11.1.0/24 area 0
  network 10.11.4.0/24 area 0
  network 192.168.0.22/32 area 0
  redistribute static
  redistribute connected
  redistribute bgp 100
!
router bgp 100
  redistribute ospf 1
  neighbor 192.168.0.3 remote-as 200
  neighbor 192.168.0.3 ebgp-multihop 255
  neighbor 192.168.0.3 no shutdown
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 1
ip msdp peer 192.168.0.11 connect-source Loopback 1
ip msdp mesh-group AS100 192.168.0.11
ip msdp originator-id Loopback 1
!
ip route 192.168.0.3/32 10.11.0.32
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

The following example shows an R3 configuration for MSDP with Anycast RP.

```
ip multicast-routing
!
interface TenGigabitEthernet 3/21
  ip pim sparse-mode
  ip address 10.11.0.32/24
  no shutdown

interface TenGigabitEthernet 3/41
  ip pim sparse-mode
  ip address 10.11.6.34/24
  no shutdown
!
```

```
interface Loopback 0
  ip pim sparse-mode
  ip address 192.168.0.3/32
  no shutdown
!
router ospf 1
  network 10.11.6.0/24 area 0
  network 192.168.0.3/32 area 0
  redistribute static
  redistribute connected
  redistribute bgp 200
!
router bgp 200
  redistribute ospf 1
  neighbor 192.168.0.22 remote-as 100
  neighbor 192.168.0.22 ebgp-multihop 255
  neighbor 192.168.0.22 update-source Loopback 0
  neighbor 192.168.0.22 no shutdown
!
ip multicast-msdp
ip msdp peer 192.168.0.11 connect-source Loopback 0
ip msdp peer 192.168.0.22 connect-source Loopback 0
ip msdp sa-filter out 192.168.0.22
!
ip route 192.168.0.1/32 10.11.0.23
ip route 192.168.0.22/32 10.11.0.23
!
ip pim rp-address 192.168.0.3 group-address 224.0.0.0/4
```

MSDP Sample Configurations

The following examples show the running-configurations described in this chapter.

For more information, see the illustrations in the Related Configuration Tasks section.

MSDP Sample Configuration: R1 Running-Config

```
ip multicast-routing
!
interface TenGigabitEthernet 1/1
  ip pim sparse-mode
  ip address 10.11.3.1/24
  no shutdown
!
interface TenGigabitEthernet 1/2
  ip address 10.11.2.1/24
  no shutdown
!
interface TenGigabitEthernet 1/21
  ip pim sparse-mode
  ip address 10.11.1.12/24
  no shutdown
!
interface Loopback 0
  ip pim sparse-mode
  ip address 192.168.0.1/32
  no shutdown
!
router ospf 1
  network 10.11.2.0/24 area 0
  network 10.11.1.0/24 area 0
  network 192.168.0.1/32 area 0
  network 10.11.3.0/24 area 0
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 0
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

MSDP Sample Configuration: R2 Running-Config

```
ip multicast-routing
!
interface TenGigabitEthernet 2/1
  ip pim sparse-mode
  ip address 10.11.4.1/24
  no shutdown
!
interface TenGigabitEthernet 2/11
  ip pim sparse-mode
  ip address 10.11.1.21/24
  no shutdown
!
interface TenGigabitEthernet 2/31
  ip pim sparse-mode
  ip address 10.11.0.23/24
  no shutdown
!
interface Loopback 0
  ip address 192.168.0.2/32
  no shutdown
!
router ospf 1
  network 10.11.1.0/24 area 0
  network 10.11.4.0/24 area 0
  network 192.168.0.2/32 area 0
  redistribute static
  redistribute connected
  redistribute bgp 100
!
router bgp 100
  redistribute ospf 1
  neighbor 192.168.0.3 remote-as 200
  neighbor 192.168.0.3 ebgp-multihop 255
  neighbor 192.168.0.3 update-source Loopback 0
  neighbor 192.168.0.3 no shutdown
!
ip route 192.168.0.3/32 10.11.0.32
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

MSDP Sample Configuration: R3 Running-Config

```
ip multicast-routing
!
interface TenGigabitEthernet 3/21
  ip pim sparse-mode
  ip address 10.11.0.32/24
  no shutdown
!
interface TenGigabitEthernet 3/41
  ip pim sparse-mode
  ip address 10.11.6.34/24
  no shutdown
!
interface ManagementEthernet 1/1
  ip address 10.11.80.3/24
  no shutdown
!
interface Loopback 0
  ip pim sparse-mode
  ip address 192.168.0.3/32
  no shutdown
!
router ospf 1
  network 10.11.6.0/24 area 0
  network 192.168.0.3/32 area 0
  redistribute static
  redistribute connected
  redistribute bgp 200
!
```

```
router bgp 200
  redistribute ospf 1
  neighbor 192.168.0.2 remote-as 100
  neighbor 192.168.0.2 ebgp-multihop 255
  neighbor 192.168.0.2 update-source Loopback 0
  neighbor 192.168.0.2 no shutdown
!
ip multicast-msdp
ip msdp peer 192.168.0.1 connect-source Loopback 0
!
ip route 192.168.0.2/32 10.11.0.23
```

MSDP Sample Configuration: R4 Running-Config

```
ip multicast-routing
!
interface TenGigabitEthernet 4/1
  ip pim sparse-mode
  ip address 10.11.5.1/24
  no shutdown
!
interface TenGigabitEthernet 4/22
  ip address 10.10.42.1/24
  no shutdown
!
interface TenGigabitEthernet 4/31
  ip pim sparse-mode
  ip address 10.11.6.43/24
  no shutdown
!
interface Loopback 0
  ip address 192.168.0.4/32
  no shutdown
!
router ospf 1
  network 10.11.5.0/24 area 0
  network 10.11.6.0/24 area 0
  network 192.168.0.4/32 area 0
!
ip pim rp-address 192.168.0.3 group-address 224.0.0.0/4
```

32

Multiple Spanning Tree Protocol (MSTP)

Multiple spanning tree protocol (MSTP) — specified in IEEE 802.1Q-2003 — is a rapid spanning tree protocol (RSTP)-based spanning tree variation that improves per-VLAN spanning tree plus (PVST+). MSTP allows multiple spanning tree instances and allows you to map many VLANs to one spanning tree instance to reduce the total number of required instances.

Protocol Overview

MSTP — specified in IEEE 802.1Q-2003 — is a rapid spanning tree protocol (RSTP)-based spanning tree variation that improves on per-VLAN spanning tree plus (PVST+). MSTP allows multiple spanning tree instances and allows you to map many VLANs to one spanning tree instance to reduce the total number of required instances.

In contrast, PVST+ allows a spanning tree instance for each VLAN. This 1:1 approach is not suitable if you have many VLANs, because each spanning tree instance costs bandwidth and processing resources.

In the following illustration, three VLANs are mapped to two multiple spanning tree instances (MSTI). VLAN 100 traffic takes a different path than VLAN 200 and 300 traffic. The behavior demonstrates how you can use MSTP to achieve load balancing.

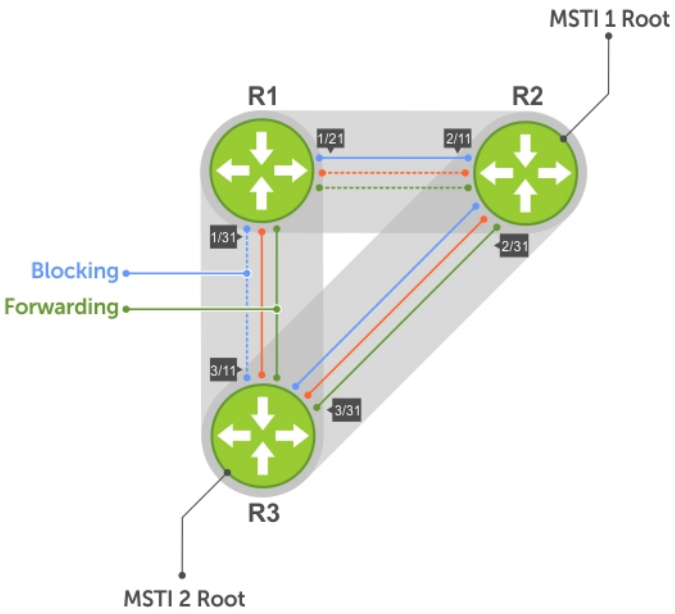


Figure 95. MSTP with Three VLANs Mapped to TWO Spanning Tree Instances

Topics:

- Spanning Tree Variations
- Configure Multiple Spanning Tree Protocol
- Enable Multiple Spanning Tree Globally
- Adding and Removing Interfaces
- Creating Multiple Spanning Tree Instances
- Influencing MSTP Root Selection
- Interoperate with Non-Dell Bridges
- Changing the Region Name or Revision

- Modifying Global Parameters
- Modifying the Interface Parameters
- Setting STP path cost as constant
- Configuring an EdgePort
- Flush MAC Addresses after a Topology Change
- MSTP Sample Configurations
- Debugging and Verifying MSTP Configurations

Spanning Tree Variations

The Dell EMC Networking OS supports four variations of spanning tree, as shown in the following table.

Table 63. Spanning Tree Variations

Dell EMC Networking Term	IEEE Specification
Spanning Tree Protocol (STP)	802 .1d
Rapid Spanning Tree Protocol (RSTP)	802 .1w
Multiple Spanning Tree Protocol (MSTP)	802 .1s
Per-VLAN Spanning Tree Plus (PVST+)	Third Party

Implementation Information

MSTP is implemented as follows in Dell EMC Networking OS:

- The Dell EMC Networking OS MSTP implementation is based on IEEE 802.1Q-2003 and interoperates only with bridges that also use this standard implementation.
- MSTP is compatible with STP and RSTP.
- Dell EMC Networking OS supports only one MSTP region.
- When you enable MSTP, all ports in Layer 2 mode participate in MSTP.
- You can configure 64 MSTIs including the default instance 0 (CIST).

Configure Multiple Spanning Tree Protocol

Configuring multiple spanning tree is a four-step process.

1. Configure interfaces for Layer 2.
2. Place the interfaces in VLANs.
3. Enable the multiple spanning tree protocol.
4. Create multiple spanning tree instances and map VLANs to them.

Related Configuration Tasks

The following are the related configuration tasks for MSTP.

- Creating Multiple Spanning Tree Instances
- Adding and Removing Interfaces
- Influencing MSTP Root Selection
- Interoperate with Non-Dell EMC Networking OS Bridges
- Changing the Region Name or Revision
- Modifying Global Parameters
- Modifying the Interface Parameters
- Configuring an EdgePort
- Flush MAC Addresses after a Topology Change
- Debugging and Verifying MSTP Configurations

- Prevent Network Disruptions with BPDU Guard
- Enabling SNMP Traps for Root Elections and Topology Changes
- Configuring Spanning Trees as Hitless

Enable Multiple Spanning Tree Globally

MSTP is not enabled by default. To enable MSTP globally, use the following commands.

When you enable MSTP, all physical, VLAN, and port-channel interfaces that are enabled and in Layer 2 mode are automatically part of the MSTI 0.

- Within an MSTI, only one path from any bridge to any other bridge is enabled.
- Bridges block a redundant path by disabling one of the link ports.

1. Enter PROTOCOL MSTP mode.

   CONFIGURATION mode

   `protocol spanning-tree mstp`
2. Enable MSTP.

   PROTOCOL MSTP mode

   `no disable`

To verify that MSTP is enabled, use the `show config` command in PROTOCOL MSTP mode.

```
DellEMC(conf)#protocol spanning-tree mstp
DellEMC(config-mstp)#show config
!
protocol spanning-tree mstp
   no disable
DellEMC#
```

Adding and Removing Interfaces

To add and remove interfaces, use the following commands.

To add an interface to the MSTP topology, configure it for Layer 2 and add it to a VLAN.

If you previously disabled MSTP on the interface using the `no spanning-tree 0` command, to enable MSTP, use the following command.

- `spanning-tree 0`

To remove an interface from the MSTP topology, use the `no spanning-tree 0` command.

Creating Multiple Spanning Tree Instances

To create multiple spanning tree instances, use the following command.

A single MSTI provides no more benefit than RSTP. To take full advantage of MSTP, create multiple MSTIs and map VLANs to them.

- Create an MSTI.

  PROTOCOL MSTP mode

  `msti`

  Specify the keyword `vlan` then the VLANs that you want to participate in the MSTI.

The following examples shows the `msti` command.

```
DellEMC(conf)#protocol spanning-tree mstp
DellEMC(conf-mstp)#msti 1 vlan 100
DellEMC(conf-mstp)#msti 2 vlan 200-300
DellEMC(conf-mstp)#show config
!
```

```
protocol spanning-tree mstp
  no disable
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200-300
```

All bridges in the MSTP region must have the same VLAN-to-instance mapping.

To view which instance a VLAN is mapped to, use the `show spanning-tree mst vlan` command from EXEC Privilege mode.

```
DellEMC(conf-mstp)#name my-mstp-region
DellEMC(conf-mstp)#exit
DellEMC(conf)#do show spanning-tree mst config
MST region name: my-mstp-region
Revision: 0
MSTI VID
1 100
2 200-300
```

To view the forwarding/discarding state of the ports participating in an MSTI, use the `show spanning-tree msti` command from EXEC Privilege mode.

```
DellEMC#show spanning-tree msti 1
MSTI 1 VLANs mapped 100

Root Identifier has priority 32768, Address 0001.e806.953e
Root Bridge hello time 2, max age 20, forward delay 15, max hops 19
Bridge Identifier has priority 32768, Address 0001.e80d.b6d6
Configured hello time 2, max age 20, forward delay 15, max hops 20
Current root has priority 32768, Address 0001.e806.953e
Number of topology changes 2, last change occured 1d2h ago on Te 1/21

Port 374 (TenGigabitEthernet 1/21) is root Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.374
Designated root has priority 32768, address 0001.e806.953e
Designated bridge has priority 32768, address 0001.e806.953e
Designated port id is 128.374, designated path cost 20000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 93671, received 46843
The port is not in the Edge port mode

Port 384 (TenGigabitEthernet 1/31) is alternate Discarding
Port path cost 20000, Port priority 128, Port Identifier 128.384
Designated root has priority 32768, address 0001.e806.953e
Designated bridge has priority 32768, address 0001.e809.c24a
Designated port id is 128.384, designated path cost 20000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 39291, received 7547
The port is not in the Edge port mode
```

Influencing MSTP Root Selection

MSTP determines the root bridge, but you can assign one bridge a lower priority to increase the probability that it becomes the root bridge.

To change the bridge priority, use the following command.

- Assign a number as the bridge priority.

  PROTOCOL MSTP mode

  `msti` *`instance`* `bridge-priority` *`priority`*

  A lower number increases the probability that the bridge becomes the root bridge.

  The range is from 0 to 61440, in increments of 4096.

  The default is **32768**.

By default, the simple configuration shown previously yields the same forwarding path for both MSTIs. The following example shows how R3 is assigned bridge priority 0 for MSTI 2, which elects a different root bridge than MSTI 2.

To view the bridge priority, use the `show config` command from PROTOCOL MSTP mode.

```
R3(conf-mstp)#msti 2 bridge-priority 0
1d2h51m: %RPM0-P:RP2 %SPANMGR-5-STP_ROOT_CHANGE: MSTP root changed for instance 2. My
Bridge ID: 0:0001.e809.c24a Old Root: 32768:0001.e806.953e New Root: 0:0001.e809.c24a

R3(conf-mstp)#show config
!
protocol spanning-tree mstp
  no disable
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300
  MSTI 2 bridge-priority 0
```

Interoperate with Non-Dell Bridges

Dell EMC Networking OS supports only one MSTP region.

A region is a combination of three unique qualities:

- **Name** is a mnemonic string you assign to the region. The default region name is **null**.
- **Revision** is a 2-byte number. The default revision number OS is **0**.
- VLAN-to-instance mapping is the placement of a VLAN in an MSTI.

For a bridge to be in the same MSTP region as another, all three of these qualities must match exactly. The default values for the name and revision number must match on all Dell EMC Networking OS devices. If there are non-Dell devices that participate in MSTP, ensure these values match on all devices.

NOTE: Some non-Dell devices may implement a non-null default region name. SFTOS, for example, uses the Bridge ID, while others may use a MAC address.

Changing the Region Name or Revision

To change the region name or revision, use the following commands.

- Change the region name.

  PROTOCOL MSTP mode

  `name` *`name`*
- Change the region revision number.

  PROTOCOL MSTP mode

  `revision` *`number`*

To view the current region name and revision, use the `show spanning-tree mst configuration` command from EXEC Privilege mode.

```
DellEMC(conf-mstp)#name my-mstp-region
DellEMC(conf-mstp)#exit
DellEMC(conf)#do show spanning-tree mst config
MST region name: my-mstp-region
Revision: 0
MSTI   VID
  1    100
  2    200-300
```

Modifying Global Parameters

The root bridge sets the values for forward-delay, hello-time, max-age, and max-hops and overwrites the values set on other MSTP bridges.

- **Forward-delay** — the amount of time an interface waits in the Listening state and the Learning state before it transitions to the Forwarding state.
- **Hello-time** — the time interval in which the bridge sends MSTP bridge protocol data units (BPDUs).

- **Max-age** — the length of time the bridge maintains configuration information before it refreshes that information by recomputing the MST topology.
- **Max-hops** — the maximum number of hops a BPDU can travel before a receiving switch discards it.

NOTE: Dell EMC Networking recommends that only experienced network administrators change MSTP parameters. Poorly planned modification of MSTP parameters can negatively affect network performance.

To change the MSTP parameters, use the following commands on the root bridge.

1. Change the forward-delay parameter.

   PROTOCOL MSTP mode

   `forward-delay` *seconds*

   The range is from 4 to 30.

   The default is **15 seconds**.
2. Change the hello-time parameter.

   PROTOCOL MSTP mode

   `hello-time` *seconds*

   NOTE: With large configurations (especially those configurations with more ports) Dell EMC Networking recommends increasing the hello-time.

   The range is from 1 to 10.

   The default is **2 seconds**.
3. Change the max-age parameter.

   PROTOCOL MSTP mode

   `max-age` *seconds*

   The range is from 6 to 40.

   The default is **20 seconds**.
4. Change the max-hops parameter.

   PROTOCOL MSTP mode

   `max-hops` *number*

   The range is from 1 to 40.

   The default is **20**.

To view the current values for MSTP parameters, use the `show running-config spanning-tree mstp` command from EXEC privilege mode.

```
DellEMC(conf-mstp)#forward-delay 16
DellEMC(conf-mstp)#exit
DellEMC(conf)#do show running-config spanning-tree mstp
!
protocol spanning-tree mstp
no disable
name my-mstp-region
MSTI 1 VLAN 100
MSTI 2 VLAN 200-300
forward-delay 16
MSTI 2 bridge-priority 4096
DellEMC(conf)#
```

Modifying the Interface Parameters

You can adjust two interface parameters to increase or decrease the probability that a port becomes a forwarding port.

- **Port cost** is a value that is based on the interface type. The greater the port cost, the less likely the port is selected to be a forwarding port.
- **Port priority** influences the likelihood that a port is selected to be a forwarding port in case that several ports have the same port cost.

The following lists the default values for port cost by interface.

Table 64. Default Values for Port Costs by Interface

Port Cost	Default Value
100-Mb/s Ethernet interfaces	200000
1-Gigabit Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
25-Gigabit Ethernet interfaces	800
40-Gigabit Ethernet interfaces	500
50-Gigabit Ethernet interfaces	400
100-Gigabit Ethernet interfaces	200
Port Channel with 100 Mb/s Ethernet interfaces	100000
Port Channel with 1-Gigabit Ethernet interfaces	10000
Port Channel with 10-Gigabit Ethernet interfaces	1000
Port Channel with 25-Gigabit Ethernet interfaces	400
Port Channel with 50-Gigabit Ethernet interfaces	200
Port Channel with 100-Gigabit Ethernet interfaces	100

To change the port cost or priority of an interface, use the following commands.

1. Change the port cost of an interface.
   INTERFACE mode
   `spanning-tree msti` *`number`* `cost` *`cost`*
   The range is from 0 to 200000.
   For the default, refer to the default values shown in the table..
2. Change the port priority of an interface.
   INTERFACE mode
   `spanning-tree msti` *`number`* `priority` *`priority`*
   The range is from 0 to 240, in increments of 16.
   The default is **128**.

To view the current values for these interface parameters, use the `show config` command from INTERFACE mode.

Setting STP path cost as constant

You can set the path cost to be constant for port-channel regardless of the operation status of the port-channel member ports.

To set the STP path cost, use the `port-channel path-cost custom` command from the PROTOCOL SPANNING-TREE mode.

The following is the example configuration:

```
DELLEMC(conf)#protocol spanning-tree pvst
DELLEMC(conf-pvst)#port-channel path-cost custom
```

This command is support in all STP modes such as STP, RSTP, MSTP, and PVST.

To change the path cost to achieve new IEEE standard behavior, use `no port-channel path-cost custom` command.

Dell EMC Networking OS behavior does not change when a new member port is added to the port-channel or an existing member port is deleted from the configuration. The path cost reflects the cumulative speed of the port-channel member ports.

Configuring an EdgePort

The EdgePort feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner.

In this mode, an interface forwards frames by default until it receives a BPDU that indicates that it should behave otherwise; it does not go through the Learning and Listening states. The `bpduguard shutdown-on-violation` option causes the interface hardware to be shut down when it receives a BPDU. When you implement only `bpduguard`, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree drops packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation. This feature is the same as PortFast mode in spanning tree.

CAUTION: Configure EdgePort only on links connecting to an end station. EdgePort can cause loops if you enable it on an interface connected to a network.

To enable EdgePort on an interface, use the following command.

- Enable EdgePort on an interface.

  INTERFACE mode

  ```
  spanning-tree mstp edge-port [bpduguard | shutdown-on-violation]
  ```

  Dell EMC Networking OS Behavior: Regarding `bpduguard shutdown-on-violation` behavior:
  - If the interface to be shut down is a port channel, all the member ports are disabled in the hardware.
  - When you add a physical port to a port channel already in the Error Disable state, the new member port is also disabled in the hardware.
  - When you remove a physical port from a port channel in the Error Disable state, the error disabled state is cleared on this physical port (the physical port is enabled in the hardware).
  - You can clear the Error Disabled state with any of the following methods:
    - Use the `shutdown` command on the interface.
    - Disable the `shutdown-on-violation` command on the interface (using the `no spanning-tree stp-id portfast [bpduguard | [shutdown-on-violation]]` command).
    - Disable spanning tree on the interface (using the `no spanning-tree` command in INTERFACE mode).
    - Disabling global spanning tree (using the `no spanning-tree` command in CONFIGURATION mode).

To verify that EdgePort is enabled, use the `show config` command from INTERFACE mode.

```
DellEMC(conf-if-te-3/11)#spanning-tree mstp edge-port
DellEMC(conf-if-te-3/11)#show config
!
interface TenGigabitEthernet 3/11
  no ip address
  switchport
  spanning-tree mstp edge-port
  spanning-tree MSTI 1 priority 144
  no shutdown
DellEMC(conf-if-te-3/11)#
```

Flush MAC Addresses after a Topology Change

Dell EMC Networking OS has an optimized MAC address flush mechanism for RSTP, MSTP, and PVST+ that flushes addresses only when necessary, which allows for faster convergence during topology changes.

However, you may activate the flushing mechanism defined by 802.1Q-2003 using the `tc-flush-standard` command, which flushes MAC addresses after every topology change notification.

To view the enable status of this feature, use the `show running-config spanning-tree mstp` command from EXEC Privilege mode.

MSTP Sample Configurations

The running-configurations support the topology shown in the following illustration.

The configurations are from Dell EMC Networking OS systems.

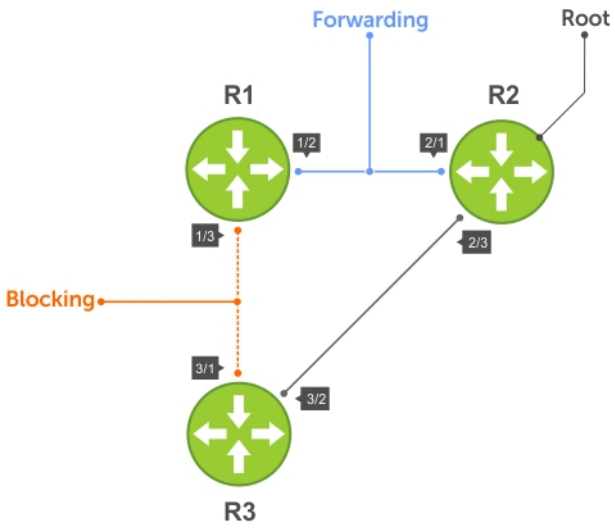


Figure 96. MSTP with Three VLANs Mapped to Two Spanning Tree Instances

Router 1 Running-Configuration

This example uses the following steps:

1. Enable MSTP globally and set the region name and revision map MSTP instances to the VLANs.
2. Assign Layer-2 interfaces to the MSTP topology.
3. Create VLANs mapped to MSTP instances tag interfaces to the VLANs.

```
(Step 1)
protocol spanning-tree mstp
  no disable
  name Tahiti
  revision 123
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300
!
(Step 2)
interface TenGigabitEthernet 1/21
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 1/31
  no ip address
  switchport
  no shutdown
!
(Step 3)
interface Vlan 100
  no ip address
  tagged TenGigabitEthernet 1/21,31
  no shutdown
!
interface Vlan 200
  no ip address
  tagged TenGigabitEthernet 1/21,31
  no shutdown
!
interface Vlan 300
  no ip address
  tagged TenGigabitEthernet 1/21,31
  no shutdown
```

Router 2 Running-Configuration

This example uses the following steps:

1. Enable MSTP globally and set the region name and revision map MSTP instances to the VLANs.
2. Assign Layer-2 interfaces to the MSTP topology.
3. Create VLANs mapped to MSTP instances tag interfaces to the VLANs.

```
(Step 1)
protocol spanning-tree mstp
  no disable
  name Tahiti
  revision 123
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300
!
(Step 2)
interface TenGigabitEthernet 2/11
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 2/31
  no ip address
  switchport
  no shutdown
!
(Step 3)
interface Vlan 100
  no ip address
  tagged TenGigabitEthernet 2/11,31
 no shutdown
!
interface Vlan 200
  no ip address
  tagged TenGigabitEthernet 2/11,31
  no shutdown
!
interface Vlan 300
  no ip address
  tagged TenGigabitEthernet 2/11,31
  no shutdown
```

Router 3 Running-Configuration

This example uses the following steps:

1. Enable MSTP globally and set the region name and revision map MSTP instances to the VLANs.
2. Assign Layer-2 interfaces to the MSTP topology.
3. Create VLANs mapped to MSTP instances tag interfaces to the VLANs.

```
(Step 1)
protocol spanning-tree mstp
  no disable
  name Tahiti
  revision 123
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300
!
(Step 2)
interface TenGigabitEthernet 3/11
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 3/21
  no ip address
  switchport
  no shutdown
```

```
!
(Step 3)
interface Vlan 100
  no ip address
  tagged TenGigabitEthernet 3/11,21
  no shutdown
!
interface Vlan 200
  no ip address
  tagged TenGigabitEthernet 3/11,21
  no shutdown
!
interface Vlan 300
  no ip address
  tagged TenGigabitEthernet 3/11,21
  no shutdown
```

SFTOS Example Running-Configuration

This example uses the following steps:

1. Enable MSTP globally and set the region name and revision map MSTP instances to the VLANs.
2. Assign Layer-2 interfaces to the MSTP topology.
3. Create VLANs mapped to MSTP instances tag interfaces to the VLANs.

```
(Step 1)
spanning-tree
spanning-tree configuration name Tahiti
spanning-tree configuration revision 123
spanning-tree MSTi instance 1
spanning-tree MSTi vlan 1 100
spanning-tree MSTi instance 2
spanning-tree MSTi vlan 2 200
spanning-tree MSTi vlan 2 300

(Step 2)
interface 1/0/31
  no shutdown
  spanning-tree port mode enable
  switchport protected 0
exit

interface 1/0/32
  no shutdown
  spanning-tree port mode enable
  switchport protected 0
exit

(Step 3)
interface vlan 100
  tagged 1/0/31
  tagged 1/0/32
exit

interface vlan 200
  tagged 1/0/31
  tagged 1/0/32
exit

interface vlan 300
  tagged 1/0/31
  tagged 1/0/32
exit
```

Debugging and Verifying MSTP Configurations

To debut and verify MSTP configuration, use the following commands.

- Display BPDUs.

  EXEC Privilege mode

  `debug spanning-tree mstp bpdu`
- Display MSTP-triggered topology change messages.

  `debug spanning-tree mstp events`

To ensure all the necessary parameters match (region name, region version, and VLAN to instance mapping), examine your individual routers.

To show various portions of the MSTP configuration, use the `show spanning-tree mst` commands.

To view the overall MSTP configuration on the router, use the `show running-configuration spanning-tree mstp` in EXEC Privilege mode.

To monitor and verify that the MSTP configuration is connected and communicating as desired, use the `debug spanning-tree mstp bpdu` command.

Key items to look for in the debug report include:

- MSTP flags indicate communication received from the same region.
  - As shown in the following, the MSTP routers are located in the same region.
  - Does the debug log indicate that packets are coming from a “Different Region”? If so, one of the key parameters is not matching.
- MSTP Region Name and Revision.
  - The configured name and revisions must be identical among all the routers.
  - Is the Region name blank? That may mean that a name was configured on one router and but was not configured or was configured differently on another router (spelling and capitalization counts).
- MSTP Instances.
  - To verify the VLAN to MSTP instance mapping, use the `show` commands.
  - Are there “extra” MSTP instances in the Sending or Received logs? This may mean that an additional MSTP instance was configured on one router but not the others.

The following example shows the `show run spanning-tree mstp` command.

```
DellEMC#show run spanning-tree mstp
!
protocol spanning-tree mstp
name Tahiti
revision 123
MSTI 1 VLAN 100
MSTI 2 VLAN 200,300
```

The following example shows viewing the debug log of a successful MSTP configuration.

```
DellEMC#debug spanning-tree mstp bpdu
MSTP debug bpdu is ON
DellEMC#
4w0d4h : MSTP: Sending BPDU on Te 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x6e
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 0
Regional Bridge Id: 32768:0001.e806.953e, CIST Port Id: 128:470
Msg Age: 0, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver3 Len: 96
Name: Tahiti, Rev: 123, Int Root Path Cost: 0
Rem Hops: 20, Bridge Id: 32768:0001.e806.953e
4w0d4h : INST 1: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
    Brg/Port Prio: 32768/128, Rem Hops: 20
INST 2: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
    Brg/Port Prio: 32768/128, Rem Hops: 20

4w0d4h : MSTP: Received BPDU on Te 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78 (Indicates MSTP routers are in the [single] region.)
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 0
Regional Bridge Id: 32768:0001.e806.953e, CIST Port Id: 128:470
Msg Age: 0, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver3 Len: 96
```

```
Name: Tahiti, Rev: 123 (MSTP region name and revision), Int Root Path Cost: 0
Rem Hops: 19, Bridge Id: 32768:0001.e8d5.cbbd
4w0d4h : INST 1 (MSTP Instance): Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root
Cost: 0
    Brg/Port Prio: 32768/128, Rem Hops: 19
INST 2 (MSTP Instance): Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
    Brg/Port Prio: 32768/128, Rem Hops: 19
Indicates MSTP
routers are in the
(single) region
MSTP Instance
MSTP Region name
```

The following example shows viewing the debug log of an unsuccessful MSTP configuration.

```
4w0d4h : MSTP: Received BPDU on Te 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78Different Region (Indicates MSTP routers are in
different regions and are not communicating with each other.)
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 0
Regional Bridge Id: 32768:0001.e806.953e, CIST Port Id: 128:470
Msg Age: 0, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver
Name: Tahiti, Rev: 123, Int Root Path Cost: 0
Rem Hops: 20, Bridge Id: 32768:0001.e8d5.cbbd
4w0d4h : INST 1: Flags: 0x70, Reg Root: 32768:0001.e8d5.cbbd, Int
    Brg/Port Prio: 32768/128, Rem Hops: 20
INST 2: Flags: 0x70, Reg Root: 32768:0001.e8d5.cbbd, Int Root Cost
    Brg/Port Prio: 32768/128, Rem Hops: 20
```

33

Multicast Features

NOTE: Multicast routing is supported on secondary IP addresses; it is not supported on IPv6.

NOTE: Multicast routing is supported across default and non-default virtual routing and forwarding (VRFs).

The Dell EMC Networking operating system (OS) supports the following multicast protocols:

- PIM Sparse-Mode (PIM-SM)
- Internet Group Management Protocol (IGMP)
- Multicast Source Discovery Protocol (MSDP)

Topics:

- Enabling IP Multicast
- Implementation Information
- Multicast Policies
- Understanding Multicast Traceroute (mtrace)
- Printing Multicast Traceroute (mtrace) Paths
- Supported Error Codes
- mtrace Scenarios

Enabling IP Multicast

Before enabling any multicast protocols, you must enable multicast routing. To enable multicast routing, use the following commands.

- Enable multicast routing for IPv4.

  CONFIGURATION mode

  ```
  ip multicast-routing
  ```

- Enable multicast routing for IPv6.

  CONFIGURATION mode

  ```
  ipv6 multicast-routing
  ```

Implementation Information

Because protocol control traffic in the Dell EMC Networking OS is redirected using the MAC address, and multicast control traffic and multicast data traffic might map to the same MAC address, the Dell EMC Networking OS might forward data traffic with certain MAC addresses to the CPU in addition to control traffic.

As the upper five bits of an IP Multicast address are dropped in the translation, 32 different multicast group IDs map to the same Ethernet address. For example, 224.0.0.5 is a known IP address for open shortest path first (OSPF) that maps to the multicast MAC address 01:00:5e:00:00:05. However, 225.0.0.5, 226.0.0.5, and so on, map to the same multicast MAC address. The Layer 2 forwarding information base (FIB) alone cannot differentiate multicast control traffic and multicast data traffic with the same address, so if you use IP address 225.0.0.5 for data traffic, both the multicast data and OSPF control traffic match the same entry and are forwarded to the CPU. Therefore, do not use well-known protocol multicast addresses for data transmission, such as the following:

Protocol	Ethernet Address
OSPF	01:00:5e:00:00:05
	01:00:5e:00:00:06

Protocol	Ethernet Address
RIP	01:00:5e:00:00:09
NTP	01:00:5e:00:01:01
VRRP	01:00:5e:00:00:12
PIM-SM	01:00:5e:00:00:0d

- The Dell EMC Networking OS implementation of MTRACE is in accordance with IETF draft *draft-fenner-traceroute-ipm*.
- Multicast is not supported on secondary IP addresses.
- If you enable multicast routing, egress Layer 3 ACL is not applied to multicast data traffic.
- Multicast traffic can be forwarded to a maximum of 15 VLANs with the same outgoing interface.

Dell EMC Networking OS does not support multicast routing in the following VLT scenarios:

- In a VLT enabled PIM router, multicast routing is not supported when there are multiple PIM spanned paths to reach source or RP. The workaround is to configure only one PIM spanned path to reach any PIM router in the aggregation or spine.
- If a source is connected to a non-spanned interface of the VLT peer nodes and the RP is reachable on a spanned interface from both the VLT nodes, there is a possibility to receive duplicate traffic on the source. To avoid this, you should configure the source to be reachable on a spanned interface.

Multicast Policies

The Dell EMC Networking OS supports multicast features for IPv4.

IPv4 Multicast Policies

The following sections describe IPv4 multicast policies.

- Limiting the Number of Multicast Routes
- Preventing a Host from Joining a Group
- Preventing a PIM Router from Forming an Adjacency
- Preventing a Source from Registering with the RP
- Preventing a PIM Router from Processing a Join

Limiting the Number of Multicast Routes

When the total number of multicast routes on a system limit is reached, the Dell EMC Networking OS does not process Internet group management protocol (IGMP) or multicast listener discovery protocol (MLD) joins to protocol-independent multicast (PIM) — though it still processes leave messages — until the number of entries decreases below 95% of the limit.

When the limit falls below 95% after hitting the maximum, the system begins relearning route entries through IGMP, MLD, and multicast source discovery protocol (MSDP).

- If you increase the limit after it is reached, subsequent join requests are accepted. In this case, increase the limit by at least 10% for IGMP and MLD to resume.
- If you decrease the limit after it is reached, the Dell EMC Networking OS does not clear the existing sessions. Entries are cleared after a timeout (you may also clear entries using the `clear ip mroute` command).

NOTE: The Dell EMC Networking OS waits at least 30 seconds between stopping and starting IGMP join processing. You may experience this delay when manipulating the limit after it is reached.

When the multicast route limit is reached, the following displays:

```
      3w1d13h: %RPM0-P:RP2 %PIM-3-PIM_TIB_LIMIT: PIM TIB limit reached. No new routes
will
be learnt until TIB level falls below low watermark.
      3w1d13h: %RPM0-P:RP2 %PIM-3-PIM_TIB_LIMIT: PIM TIB below low watermark. Route
learning
will begin.
```

To limit the number of multicast routes, use the following command.

- Limit the total number of multicast routes on the system.
  CONFIGURATION mode
  `ip multicast-limit`
  The range is from 1 to 16000.
  The default is **4000**.

NOTE: The IN-L3-McastFib CAM partition stores multicast routes and is a separate hardware limit that exists per port-pipe. Any software-configured limit may supersede this hardware space limitation. The opposite is also true, the CAM partition might not be exhausted at the time the system-wide route limit is reached using the `ip multicast-limit` command.

Preventing a Host from Joining a Group

You can prevent a host from joining a particular group by blocking specific IGMP reports using an extended access list containing the permissible source-group pairs.

NOTE: For rules in IGMP access lists, *source* is the multicast source, not the source of the IGMP packet. For IGMPv2, use the keyword `any` for *source* (as shown in the following example) because the IGMPv2 hosts do not know in advance who the source is for the group in which they are interested.

To apply the access list, use the following command.

- Apply the access list.
  INTERFACE mode
  `ip igmp access-group` *`access-list-name`*

Dell EMC Networking OS Behavior: Do not enter the `ip igmp access-group` command before creating the access-list. If you do, after entering your first deny rule, the Dell EMC Networking OS clears the multicast routing table and re-learns all groups, even those not covered by the rules in the access-list, because there is an implicit *deny all* rule at the end of all access-lists. Therefore, configuring an IGMP join request filter in this order might result in data loss. If you must enter the `ip igmp access-group` command before creating the access-list, prevent the Dell EMC Networking OS from clearing the routing table by entering a *permit any* rule with a high sequence number before you enter any other rules.

In the following example, virtual local area network (VLAN) 400 is configured with an access list to permit only IGMP reports for group 239.0.0.1. Though Receiver 2 sends a membership report for groups 239.0.0.1 and 239.0.0.2, a multicast routing table entry is created only for group 239.0.0.1. VLAN 300 has no access list limiting Receiver 1, so both IGMP reports are accepted and two corresponding entries are created in the routing table.

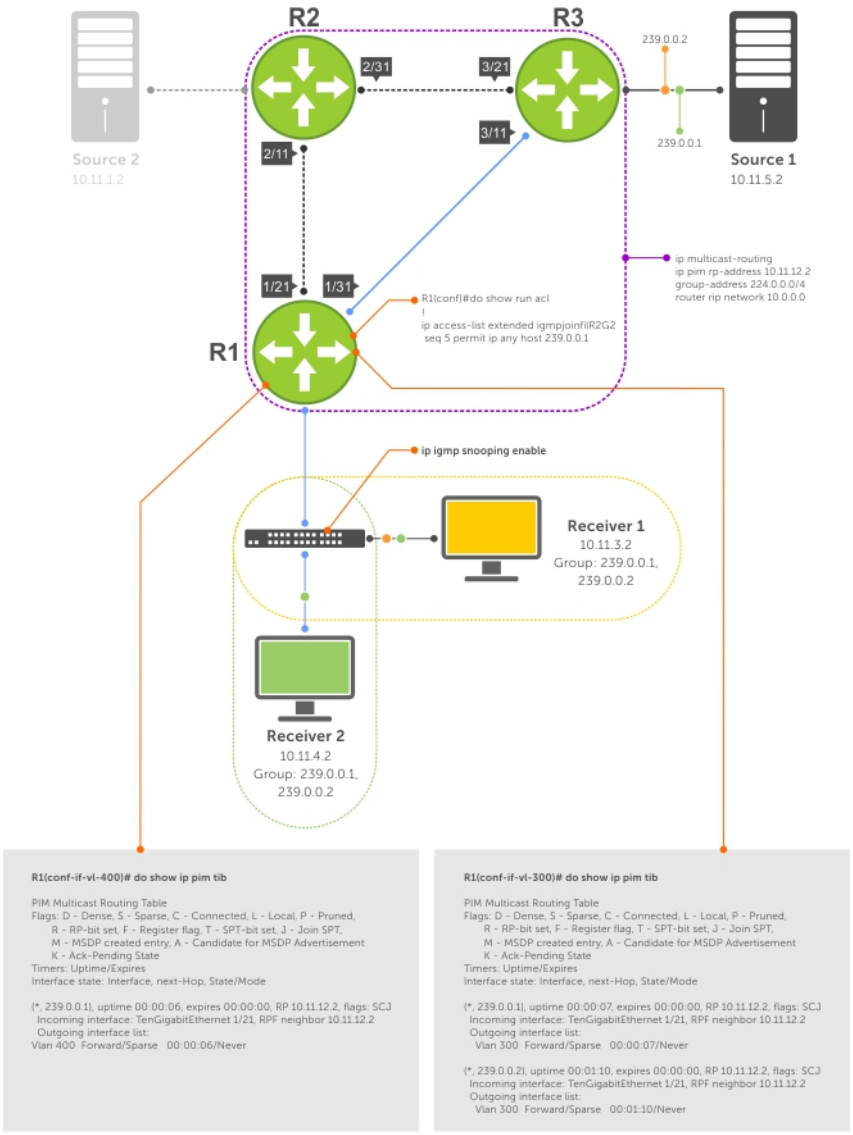


Figure 97. Preventing a Host from Joining a Group

The following table lists the location and description shown in the previous illustration.

Table 65. Preventing a Host from Joining a Group — Description

Location	Description
1/21	• Interface TenGigabitEthernet 1/21 • ip pim sparse-mode • ip address 10.11.12.1/24 • no shutdown
1/31	• Interface TenGigabitEthernet 1/31 • ip pim sparse-mode • ip address 10.11.13.1/24 • no shutdown
2/1	• Interface TenGigabitEthernet 2/1 • ip pim sparse-mode • ip address 10.11.1.1/24 • no shutdown

Table 65. Preventing a Host from Joining a Group — Description (continued)

Location	Description
2/11	• Interface TenGigabitEthernet 2/11 • ip pim sparse-mode • ip address 10.11.12.2/24 • no shutdown
2/31	• Interface TenGigabitEthernet 2/31 • ip pim sparse-mode • ip address 10.11.23.1/24 • no shutdown
3/1	• Interface TenGigabitEthernet 3/1 • ip pim sparse-mode • ip address 10.11.5.1/24 • no shutdown
3/11	• Interface TenGigabitEthernet 3/11 • ip pim sparse-mode • ip address 10.11.13.2/24 • no shutdown
3/21	• Interface TenGigabitEthernet 3/21 • ip pim sparse-mode • ip address 10.11.23.2/24 • no shutdown
Receiver 1	• Interface VLAN 300 • ip pim sparse-mode • ip address 10.11.3.1/24 • untagged TenGigabitEthernet1/1 • no shutdown
Receiver 2	• Interface VLAN 400 • ip pim sparse-mode • ip address 10.11.4.1/24 • untagged TenGigabitEthernet 1/2 • **ip igmp access-group igmpjoinfilR2G2** • no shutdown

Preventing a PIM Router from Forming an Adjacency

To prevent a router from participating in PIM (for example, to configure stub multicast routing), use the following command.

- Prevent a router from participating in PIM.

  INTERFACE mode

  ```
  ip pim neighbor-filter
  ```

Preventing a Source from Registering with the RP

To prevent the PIM source DR from sending register packets to route processor (RP) for the specified multicast source and group, use the following command. If the source DR never sends register packets to the RP, no hosts can ever discover the source and create a shortest path tree (SPT) to it.

- Prevent a source from transmitting to a particular group.

  CONFIGURATION mode

  ```
  ip pim register-filter
  ```

In the following example, Source 1 and Source 2 are both transmitting packets for groups 239.0.0.1 and 239.0.0.2. R3 has a PIM register filter that only permits packets destined for group 239.0.0.2. An entry is created for group 239.0.0.1 in the routing table,

but no outgoing interfaces are listed. R2 has no filter, so it is allowed to forward both groups. As a result, Receiver 1 receives only one transmission, while Receiver 2 receives duplicate transmissions.

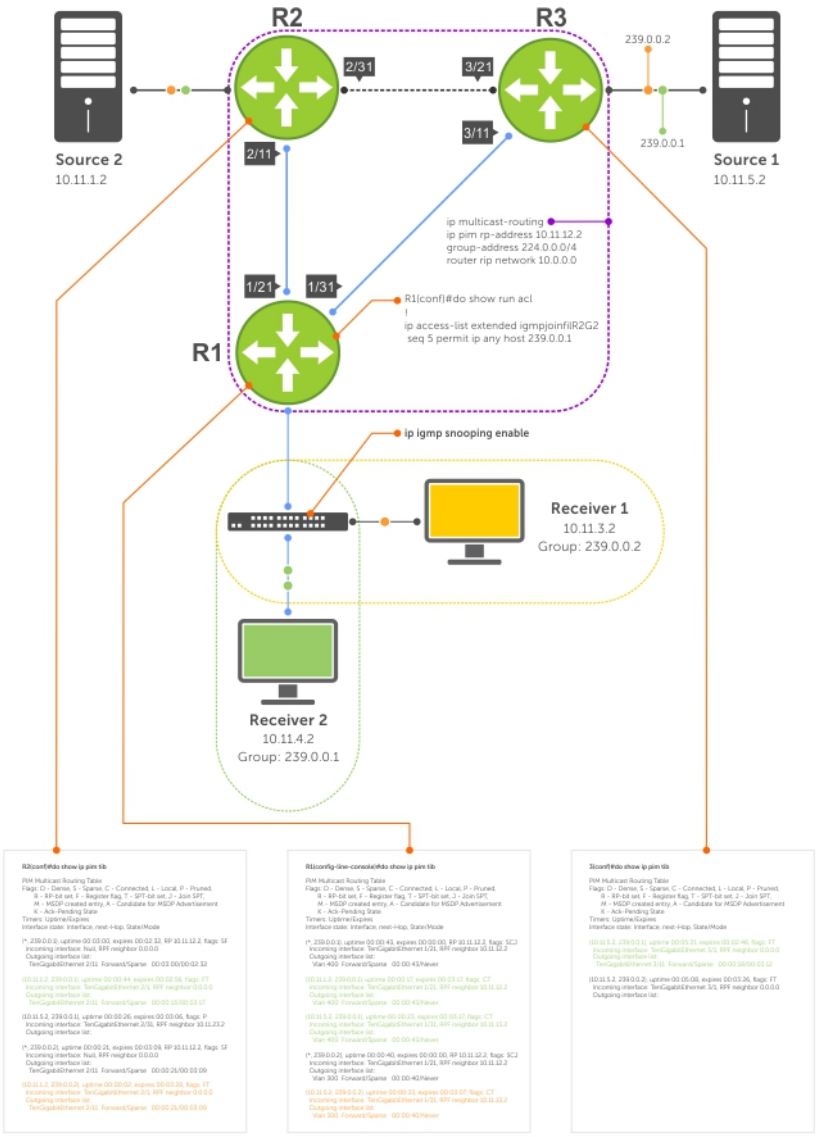


Figure 98. Preventing a Source from Transmitting to a Group

The following table lists the location and description shown in the previous illustration.

Table 66. Preventing a Source from Transmitting to a Group — Description

Location	Description
1/21	• Interface TenGigabitEthernet 1/21 • ip pim sparse-mode • ip address 10.11.12.1/24 • no shutdown
1/31	• Interface TenGigabitEthernet 1/31 • ip pim sparse-mode • ip address 10.11.13.1/24 • no shutdown
2/1	• Interface TenGigabitEthernet 2/1 • ip pim sparse-mode • ip address 10.11.1.1/24

Table 66. Preventing a Source from Transmitting to a Group — Description (continued)

Location	Description
	• no shutdown
2/11	• Interface TenGigabitEthernet 2/11 • ip pim sparse-mode • ip address 10.11.12.2/24 • no shutdown
2/31	• Interface TenGigabitEthernet 2/31 • ip pim sparse-mode • ip address 10.11.23.1/24 • no shutdown
3/1	• Interface TenGigabitEthernet 3/1 • ip pim sparse-mode • ip address 10.11.5.1/24 • no shutdown
3/11	• Interface TenGigabitEthernet 3/11 • ip pim sparse-mode • ip address 10.11.13.2/24 • no shutdown
3/21	• Interface TenGigabitEthernet 3/21 • ip pim sparse-mode • ip address 10.11.23.2/24 • no shutdown
Receiver 1	• Interface VLAN 300 • ip pim sparse-mode • ip address 10.11.3.1/24 • untagged TenGigabitEthernet 1/1 • no shutdown
Receiver 2	• Interface VLAN 400 • ip pim sparse-mode • ip address 10.11.4.1/24 • untagged TenGigabitEthernet 1/2 • no shutdown

Preventing a PIM Router from Processing a Join

To permit or deny PIM Join/Prune messages on an interface using an extended IP access list, use the following command.

NOTE: Dell EMC Networking recommends not using the `ip pim join-filter` command on an interface between a source and the RP router. Using this command in this scenario could cause problems with the PIM-SM source registration process resulting in excessive traffic being sent to the CPU of both the RP and PIM DR of the source.

Excessive traffic generates when the join process from the RP back to the source is blocked due to a new source group being permitted in the join-filter. This results in the new source becoming stuck in registering on the DR and the continuous generation of user datagram protocol (UDP)-encapsulated registration messages between the DR and RP routers which are being sent to the CPU.

- Prevent the PIM SM router from creating a state based on multicast source and/or group.

  `ip pim join-filter`

  NOTE: When you configure a join filter, it is applicable for both ingress and egress flows. There is no option to specify in or out parameters while configuring a join filter.

Understanding Multicast Traceroute (mtrace)

Multicast Traceroute (mtrace) is a multicast diagnostic facility used for tracing multicast paths.

Mtrace enables you to trace the path that a multicast packet takes from its source to the destination. When you initiate mtrace from a source to a destination, an mtrace Query packet with IGMP type 0x1F is sent to the last-hop multicast router for the given destination. The mtrace query packet is forwarded hop-by-hop untill it reaches the last-hop router.

NOTE: If the system initiating the mtrace is the last-hop router, then the Query message will not be initiated. Instead, the router sends the request message to it previous router.

The last-hop router converts this query packet to a request packet by adding a response data block. This response data block contains the last-hop router's interface address. The response data block inserted by the router also contains the following information:

- Incoming interface details
- Outgoing interface details
- Previous-hop router address
- Forwarding Code
- Query Arrival Time
- Routing Protocol details

The last-hop router calculates the path to reach the source in the reverse direction of the multicast data traffic. Based on this calculation, the last-hop router estimates the possible next-hop neighbor that is located in the direction towards the source and forwards the request packet to that neighbor.

Each router along the multicast path fills its response block in a similar manner. When the mtrace request reaches the first-hop router, it sends the response (with IGMP type 0x1E) to the response destination address specified in the mtrace query.

The response may be returned before reaching the first-hop router if a fatal error condition such as "no route" is encountered along the path.

If a multicast router along the path does not implement the mtrace feature or if there is an outage, no response is returned.

When the initiator does not get a response for a specified time interval, the system performs a hop-by-hop expanding-length search to pin point the location in the network where the problem has occurred.

NOTE: You cannot configure the wait time. It is fixed to 3 seconds.

Important Points to Remember

- Destination address of the mtrace query message can be either a unicast or a multicast address.
  NOTE: When you use mtrace to trace a specific multicast group, the query is sent with the group's address as the destination. Retries of the query use the unicast address of the receiver.
- When you issue an mtrace without specifying a group address (weak mtrace), the destination address is considered as the unicast address of the receiver.
- If the CLI session is terminated after the mtrace command is issued, then the response is ignored.
- System ignores any stray mtrace responses that it receives.
- Duplicate query messages as identified by the IP source, and Query ID (tuple) are ignored. However, duplicate request messages are not ignored in a similar manner.
- The system supports up to a maximum of eleven mtrace clients at a time.
  NOTE: The maximum number of clients are subject to performance restrictions in the new platform.
- Mtrace supports only IPv4 address family.

Printing Multicast Traceroute (mtrace) Paths

Dell EMC Networking OS supports Multicast traceroute.

MTRACE is an IGMP-based tool that prints the network path that a multicast packet takes from a source to a destination, for a particular group. Dell EMC Networking OS has mtrace client and mtrace transit functionality.

- **MTRACE Client** — an mtrace client transmits mtrace queries and print the details from received responses.

- **MTRACE Transit** — when a Dell EMC Networking system is an intermediate router between the source and destination in an MTRACE query, Dell EMC Networking OS computes the RPF neighbor for the source, fills in the request, and forwards the request to the RPF neighbor. When a Dell EMC Networking system is the last hop to the destination, Dell EMC Networking OS sends a response to the query.

To print the network path, use the following command.

- Print the network path that a multicast packet takes from a multicast source to receiver, for a particular group.

  EXEC Privilege mode

  mtrace *multicast-source-address multicast-receiver-address multicast-group-address*

```
From source (?) to destination (?)

-------------------------------------------------------------------------------------
--
|Hop|     OIF IP              |Proto|   Forwarding Code            |Source Network/
Mask|
-------------------------------------------------------------------------------------
--
  0    "destination ip(to)" -->    Destination
 -1    "Outgoing intf addr"   "Proto"   "Err/fwd code if present"   "Src Mask"
 -2    "Outgoing intf addr"   "Proto"   "Err/fwd code if present"   "Src Mask"
            .
            .
-"n"  "source ip(from)" -->  Source
-------------------------------------------------------------------------------------
---
```

  The mtrace command traverses the path of the response data block in the reverse direction of the multicast data traffic. As a result, the tabular output of the mtrace command displays the destination details in the first row, followed by the RPF router details along the path in the consequent rows, and finally the source details in the last row. The tabular output contains the following columns:

  - Hop — a hop number(counted negatively to indicate reverse-path)
  - OIF IP — outgoing interface address
  - Proto — multicast routing protocol
  - Forwarding code — error code as present in the response blocks
  - Source Network/Mask — source mask

The following is an example of tracing a multicast route.

```
R1>mtrace 103.103.103.3 1.1.1.1 226.0.0.3
Type Ctrl-C to abort.

Querying reverse path for source 103.103.103.3 to destination 1.1.1.1 via group 226.0.0.3
From source (?) to destination (?)

-----------------------------------------------------------------
|Hop|    OIF IP     |Proto| Forwarding Code |Source Network/Mask|
-----------------------------------------------------------------
  0  1.1.1.1          -->    Destination
 -1  1.1.1.1          PIM   Reached RP/Core   103.103.103.0/24
 -2  101.101.101.102  PIM       -             103.103.103.0/24
 -3  2.2.2.1          PIM       -             103.103.103.0/24
 -4  103.103.103.3    -->    Source
-----------------------------------------------------------------
```

The following table explains the output of the mtrace command:

Table 67. mtrace Command Output — Explained

Command Output	Description
`Querying reverse path for source 103.103.103.3 to destination 1.1.1.1 via group 226.0.0.3`	mtrace traverses the reverse path from the given destination to the given source for the given group

Table 67. mtrace Command Output — Explained (continued)

Command Output	Description
`From source (?) to destination (?)`	In case the provided source or destination IP can be resolved to a hostname the corresponding name will be displayed. In cases where the IP cannot be resolved, it is displayed as (?)
`0 1.1.1.1 --> Destination`	The first row in the table corresponds to the destination provided by the user.
`-1 1.1.1.1 PIM Reached RP/Core 103.103.103.0/24`	The information in each of the response blocks is displayed as follows: • o (-1) Hop count is always a negative number to indicate reverse path • o (1.1.1.1) Outgoing interface address at that node for the source and group • o (PIM) Multicast protocol used at the node to retrieve the information • o (Reached RP/Core) Forwarding code in mtrace to denote that RP node is reached • o (103.103.103.0/24) Source network and mask. In case (*G) tree is used, this field will have the value as (shared tree). In case no value is noted in the record or in case of error like No Route or Wrong Last Hop the value (default) will be displayed
-4 103.103.103.3 --> Source	The last line in the table corresponds to the source address provided by the user.

Supported Error Codes

Error codes denote problems in the network that has caused the mtrace query to fail.

These codes not only provide error information but also provide general information such as RP node reachability information.

The response data block filled in by the last-hop router contains a Forwarding code field. Forwarding code can be added at any node and is not restricted to the last hop router. This field is used to record error codes before forwarding the response to the next neighbor in the path towards the source. In a response data packet, the following error codes are supported:

Table 68. Supported Error Codes

Error Code	Error Name	Description
0x00	NO_ERROR	No error
0x01	WRONG_IF	Traceroute request arrived on the wrong interface. The router does not use this interface to forward packets to the source, group, and destination.
0x05	NO_ROUTE	The router has no route for the source or group and cannot determine a potential route.
0x06	WRONG_LAST_HOP	The router is not the proper last-hop router.
0x08	REACHED_RP	Reached Rendezvous Point or Core.
0x09	RPF_IF	Traceroute request arrived on the expected RPF interface for this source and group.
0x0A	NO_MULTICAST	Traceroute request arrived on an interface which is not enabled for multicast.

Table 68. Supported Error Codes (continued)

Error Code	Error Name	Description
0x81	NO_SPACE	There is not enough room to insert another response data block in the packet.

mtrace Scenarios

This section describes various scenarios that may result when an `mtrace` command is issued.

The following table describes various scenarios when the `mtrace` command is issued:

Table 69. Mtrace Scenarios

Scenario	Output

When you want to trace a route with the multicast tree for a source, group, and destination, you can specify all the parameters in the command. Mtrace will trace the complete path from source to destination by using the multicast tables for that group.	`R1>mtrace 103.103.103.3 1.1.1.1 226.0.0.3`
	`Type Ctrl-C to abort.`
	
	`Querying reverse path for source 103.103.103.3 to destination 1.1.1.1 via group 226.0.0.3`
	`From source (?) to destination (?)`
	
	`-----------------------------------------------------------------`
	`|Hop|  OIF IP  |Proto| Forwarding Code |Source Network/Mask|`
	`-----------------------------------------------------------------`
	`  0  1.1.1.1  -->  Destination`
	`-1  1.1.1.1  PIM  Reached RP/Core 103.103.103.0/24`
	`-2  101.101.101.102  PIM  - 103.103.103.0/24`
	`-3  2.2.2.1  PIM  - 103.103.103.0/24`
	`-4  103.103.103.3  -->  Source`
	`-----------------------------------------------------------------`

You can issue the mtrace command specifying the source multicast tree and multicast group without specifying the destination. Mtrace traces the complete path traversing through the multicast group to reach the source. The output displays the destination and the first hop (-1) as 0 to indicate any PIM enabled interface on the node.	`R1>mtrace 103.103.103.3 1.1.1.1 226.0.0.3`
	`Type Ctrl-C to abort.`
	
	`Querying reverse path for source 103.103.103.3 via group 226.0.0.3`
	`From source (?) to this node`
	
	`-----------------------------------------------------------------`
	`|Hop|  OIF IP  |Proto| Forwarding Code |Source Network/Mask|`
	`-----------------------------------------------------------------`
	`  0  0.0.0.0*  -->  Destination`
	`-1  0.0.0.0*  PIM  - 103.103.103.0/24`
	`-2  2.2.2.1  PIM  - 103.103.103.0/24`
	`-3  103.103.103.3  -->  Source`
	`-----------------------------------------------------------------`
	`* - Any PIM enabled interface on this node`

Table 69. Mtrace Scenarios (continued)

Scenario	Output
You invoke a weak mtrace request by specifying only the source without specifying the mulicast tree or multicast group information for the source. Mtrace traces a path towards the source by using the RPF neighbor at each node.	R1>mtrace 103.103.103.3 Type Ctrl-C to abort. Querying reverse path for source 103.103.103.3 via RPF From source (?) to this node ---------- -------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- -------- 0 0.0.0.0* --> Destination -1 0.0.0.0* PIM - 103.103.103.0/24 -2 2.2.2.1 PIM - 103.103.103.0/24 -3 103.103.103.3 --> Source ---------- --------- * - Any PIM enabled interface on this node R1>mtrace 103.103.103.3 1.1.1.1 Type Ctrl-C to abort. Querying reverse path for source 103.103.103.3 to destination 1.1.1.1 via RPF From source (?) to destination (?) ---------- -------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- -------- 0 1.1.1.1 --> Destination -1 1.1.1.1 PIM - 103.103.103.0/24 -2 101.101.101.102 PIM - 103.103.103.0/24 -3 2.2.2.1 PIM - 103.103.103.0/24 -4 103.103.103.3 --> Source ---------- ---------
You can issue the mtrace command by providing the source and multicast information. However, if the multicast group is a shared group (*,G), then mtrace traces the path of the shared tree until it reaches the RP. The source mask field reflects the shared tree that is being used to trace the path. The shared tree is used even in case where the source provided is not valid.	R1>mtrace 3.3.3.3 1.1.1.1 226.0.0.3 Type Ctrl-C to abort. Querying reverse path for source 3.3.3.3 to destination 1.1.1.1 via group 226.0.0.3 From source (?) to destination (?) ---------- -------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- -------- 0 1.1.1.1 --> Destination -1 1.1.1.1 PIM - shared tree -2 12.12.12.1 PIM Reached RP/Core shared tree ---------- ---------

Table 69. Mtrace Scenarios (continued)

Scenario	Output

When you issue the mtrace command with the source and multicast group information, if a multicast route is not present on a particular node, then the NO ROUTE error code is displayed on the node. In this scenario, the Source Network/Mask column for that particular node displays the the value as default.	`R1>mtrace 6.6.6.6 4.4.4.5 234.1.1.1`
	`Type Ctrl-C to abort.`
	
	`Querying reverse path for source 6.6.6.6 to destination 4.4.4.5 via group 234.1.1.1`
	`From source (?) to destination (?)`
	
	`----------------------------------------------------------------`
	`|Hop|  OIF IP  |Proto| Forwarding Code |Source Network/Mask|`
	`----------------------------------------------------------------`
	`  0  4.4.4.5  -->  Destination`
	`-1  4.4.4.4  PIM  -  6.6.6.0/24`
	`-2  20.20.20.2  PIM  -  6.6.6.0/24`
	`-3  10.10.10.1  PIM  No route  default`
	`----------------------------------------------------------------`

If you invoke a weak mtrace query (without the multicast group details) and the RPF neighbor on one of the nodes to the source is not PIM enabled, the output of the command displays a NO ROUTE error code in the Forwarding Code column. In the command output, the entry for that node in the Source Network/Mask column displays the value as default.	`R1>mtrace 6.6.6.6 4.4.4.5`
	`Type Ctrl-C to abort.`
	
	`Querying reverse path for source 6.6.6.6 to destination 4.4.4.5 via group 234.1.1.1`
	`From source (?) to destination (?)`
	
	`----------------------------------------------------------------`
	`|Hop|  OIF IP  |Proto| Forwarding Code |Source Network/Mask|`
	`----------------------------------------------------------------`
	`  0  4.4.4.5  -->  Destination`
	`-1  4.4.4.4  PIM  -  6.6.6.0/24`
	`-2  20.20.20.2  PIM  -  6.6.6.0/24`
	`-3  10.10.10.1  PIM  No route  default`
	`----------------------------------------------------------------`

If a multicast tree is not formed due to a configuration issue (for example, PIM is not enabled on one of the interfaces on the path), you can invoke a weak mtrace to identify the location in the network where the error has originated.	`R1>mtrace 6.6.6.6 4.4.4.5`
	`Type Ctrl-C to abort.`
	
	`Querying reverse path for source 6.6.6.6 to destination 4.4.4.5 via RPF`
	`From source (?) to destination (?)`
	
	`----------------------------------------------------------------`
	`|Hop|  OIF IP  |Proto| Forwarding Code |Source Network/Mask|`
	`----------------------------------------------------------------`
	`  0  4.4.4.5  -->  Destination`
	`-1  4.4.4.4  PIM  6.6.6.0/24`
	`-2  20.20.20.2  PIM  6.6.6.0/24`
	`-3  10.10.10.1  PIM  Multicast disabled  6.6.6.0/24`

Table 69. Mtrace Scenarios (continued)

Scenario	Output
	```---------- ---------```
If the destination provided in the command is not a valid receiver for the multicast group, the last hop router for the destination provides the WRONG LAST HOP error code. If the last-hop router contains a path to the source, the path is traced irrespective of the incorrect destination.	```R1>mtrace 6.6.6.6 5.5.5.5 234.1.1.1 Type Ctrl-C to abort. Querying reverse path for source 6.6.6.6 to destination 4.4.4.5 via group 234.1.1.1 From source (?) to destination (?) ---------- -------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- -------- 0 5.5.5.5 --> Destination -1 5.5.5.4 PIM Wrong Last-Hop 6.6.6.0/24 -2 20.20.20.2 PIM 6.6.6.0/24 -3 10.10.10.1 PIM 6.6.6.0/24 -4 6.6.6.6 --> Source ---------- ---------```
If a router in the network does not process mtrace and drops the packet resulting in no response, the system performs an expanding-hop search to trace the path to the router that has dropped mtrace. The output of the command displays a '*' indicating that no response is received for an mtrace request. The following message appears when the system performs a hop-by-hop search: "switching to hop-by-hop:"	```R1>mtrace 99.99.99.99 1.1.1.1 Type Ctrl-C to abort. Querying reverse path for source 99.99.99.99 to destination 1.1.1.1 via RPF From source (?) to destination (?) * * * * switching to hop-by-hop: ---------- -------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- -------- 0 1.1.1.1 --> Destination -1 1.1.1.1 PIM - 99.99.0.0/16 -2 101.101.101.102 PIM - 99.99.0.0/16 -3 2.2.2.1 PIM - 99.99.0.0/16 -4 * * * * ---------- ---------```
If there is no response for mtrace even after switching to expanded hop search, the command displays an error message.	```R1>mtrace 99.99.99.99 1.1.1.1 Type Ctrl-C to abort. Querying reverse path for source 99.99.99.99 to destination 1.1.1.1 via RPF From source (?) to destination (?) * * * * switching to hop-by-hop: ---------- -------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- --------```

Table 69. Mtrace Scenarios (continued)

Scenario	Output
	0 1.1.1.1 --> Destination -1 * * * * --- --- Timed out receiving responses Perhaps no local router has a route for source, the receiver is not a member of the multicast group or the multicast ttl is too low.
While traversing the path from source to destination, if the mtrace packet exhausts the maximum buffer size of the packet, then NO SPACE error is displayed in the output. You can initiate a new mtrace query by specifying the destination as the last IP address from the output of the previous trace query.	R1>mtrace 99.99.99.99 1.1.1.1 Type Ctrl-C to abort. Querying reverse path for source 99.99.99.99 to destination 1.1.1.1 via RPF From source (?) to destination (?) --- --- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| --- --- 0 1.1.1.1 --> Destination -1 1.1.1.1 PIM - 99.99.0.0/16 -2 101.101.101.102 PIM - 99.99.0.0/16 -3 2.2.2.1 PIM - 99.99.0.0/16 . . . -146 17.17.17.17 PIM No space in packet 99.99.0.0/16 --- ---
In a valid scenario, mtrace request packets are expected to be received on the OIF of the node. However, due to incorrect formation of the multicast tree, the packet may be received on a wrong interface. In such a scenario, a corresponding error message is displayed.	R1>mtrace 6.6.6.6 4.4.4.5 Type Ctrl-C to abort. Querying reverse path for source 6.6.6.6 to destination 4.4.4.5 via RPF From source (?) to destination (?) --- --- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| --- --- 0 4.4.4.5 --> Destination -1 4.4.4.4 PIM 6.6.6.0/24 -2 20.20.20.2 PIM 6.6.6.0/24 -3 10.10.10.1 PIM Wrong interface 6.6.6.0/24 --- --- R1>mtrace 6.6.6.6 4.4.4.5 Type Ctrl-C to abort.

Table 69. Mtrace Scenarios (continued)

Scenario	Output
	Querying reverse path for source 6.6.6.6 to destination 4.4.4.5 via RPF From source (?) to destination (?) ---------- --------- |Hop| OIF IP |Proto| Forwarding Code |Source Network/Mask| ---------- --------- 0 4.4.4.5 --> Destination -1 4.4.4.4 PIM 6.6.6.0/24 -2 20.20.20.2 PIM 6.6.6.0/24 -3 10.10.10.1 PIM RPF Interface 6.6.6.0/24 ---------- ----------

34

Object Tracking

IPv4 or IPv6 object tracking is available on Dell EMC Networking OS.

Object tracking allows the Dell EMC Networking OS client processes, such as virtual router redundancy protocol (VRRP), to monitor tracked objects (for example, interface or link status) and take appropriate action when the state of an object changes.

NOTE: In Dell EMC Networking OS release version 8.4.1.0, object tracking is supported only on VRRP.

Topics:

- Object Tracking Overview
- Object Tracking Configuration
- Displaying Tracked Objects

Object Tracking Overview

Object tracking allows you to define objects of interest, monitor their state, and report to a client when a change in an object's state occurs.

The following tracked objects are supported:

- Link status of Layer 2 interfaces
- Routing status of Layer 3 interfaces (IPv4 and IPv6)
- Reachability of IP hosts
- Reachability of IPv4 and IPv6 routes
- Metric thresholds of IPv4 and IPv6 routes
- Tracking of IP Hosts

In future releases, environmental alarms and available free memory will be supported. You can configure client applications, such as VRRP, to receive a notification when the state of a tracked object changes.

The following example shows how object tracking is performed. Router A and Router B are both connected to the internet via interfaces running OSPF. Both routers belong to a VRRP group with a virtual router at 10.0.0.1 on the local area network (LAN) side. Neither Router A nor Router B is the owner of the group. Although Router A and Router B use the same default VRRP priority (100), Router B would normally become the master for the VRRP group because it has a higher IP address.

You can create a tracked object to monitor the metric of the default route 0.0.0.0/0. After you configure the default route as a tracked object, you can configure the VRRP group to track the state of the route. In this way, the VRRP priority of the router with the better metric automatically becomes master of the VRRP group. Later, if network conditions change and the cost of the default route in each router changes, the mastership of the VRRP group is automatically reassigned to the router with the better metric.

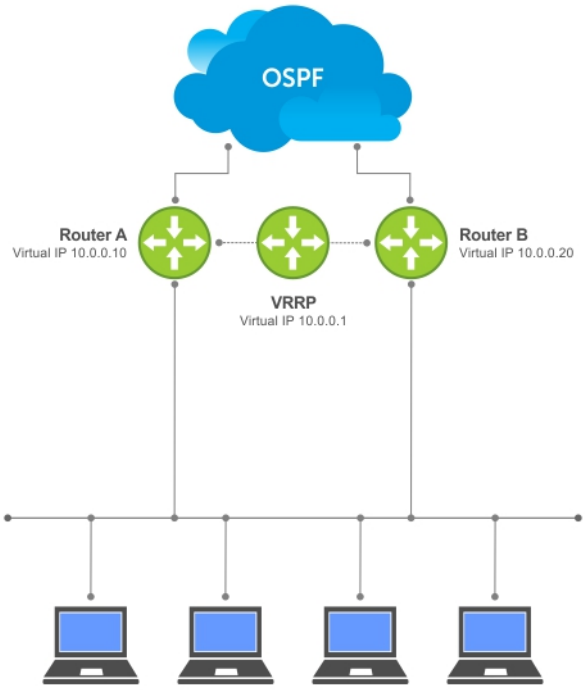


Figure 99. Object Tracking Example

When you configure a tracked object, such as an IPv4/IPv6 a route or interface, you specify an object number to identify the object. Optionally, you can also specify:

- UP and DOWN thresholds used to report changes in a route metric.
- A time delay before changes in a tracked object's state are reported to a client.

Track Layer 2 Interfaces

You can create an object to track the line-protocol state of a Layer 2 interface. In this type of object tracking, the link-level operational status (UP or DOWN) of the interface is monitored.

When the link-level status goes down, the tracked resource status is considered to be DOWN; if the link-level status goes up, the tracked resource status is considered to be UP. For logical interfaces, such as port-channels or virtual local area networks (VLANs), the link-protocol status is considered to be UP if any physical interface under the logical interface is UP.

Track Layer 3 Interfaces

You can create an object that tracks the Layer 3 state (IPv4 or IPv6 routing status) of an interface.

- The Layer 3 status of an interface is UP only if the Layer 2 status of the interface is UP and the interface has a valid IP address.
- The Layer 3 status of an interface goes DOWN when its Layer 2 status goes down or the IP address is removed from the routing table.

Track IPv4 and IPv6 Routes

You can create an object that tracks an IPv4 or IPv6 route entry in the routing table.

Specify a tracked route by its IPv4 or IPv6 address and prefix-length. Optionally specify a tracked route by a virtual routing and forwarding (VRF) instance name if the route to be tracked is part of a VRF. The next-hop address is not part of the definition of the tracked object.

A tracked route matches a route in the routing table only if the exact address and prefix length match an entry in the routing table. For example, when configured as a tracked route, 10.0.0.0/24 does not match the routing table entry 10.0.0.0/8. If no route-table entry has the exact address and prefix length, the tracked route is considered to be DOWN.

In addition to the entry of a route in the routing table, you can configure how the status of a route is tracked in either the following ways:

- By the reachability of the route's next-hop router.
- By comparing the UP or DOWN threshold for a route's metric with current entries in the route table.

Track Route Reachability

If you configure the reachability of an IP route entry as a tracked object, the UP/DOWN state of the route is determined by the entry of the next-hop address in the ARP cache.

A tracked route is considered to be reachable if there is an address resolution protocol (ARP) cache entry for the route's next-hop address. If the next-hop address in the ARP cache ages out for a route tracked for its reachability, an attempt is made to regenerate the ARP cache entry to see if the next-hop address appears before considering the route DOWN.

Track a Metric Threshold

If you configure a metric threshold to track a route, the UP/DOWN state of the tracked route is determined by the current metric for the route entered in the routing table.

To provide a common tracking interface for different clients, route metrics are scaled in the range from 0 to 255, where 0 is connected and 255 is inaccessible. The scaled metric value communicated to a client always considers a lower value to have priority over a higher value. The resulting scaled value is compared against the threshold values to determine the state of a tracked route as follows:

- If the scaled metric for a route entry is less than or equal to the UP threshold, the state of a route is UP.
- If the scaled metric for a route is greater than or equal to the DOWN threshold or the route is not entered in the routing table, the state of a route is DOWN.

The UP and DOWN thresholds are user-configurable for each tracked route. The default UP threshold is **254**; the default DOWN threshold is **255**. The notification of a change in the state of a tracked object is sent when a metric value crosses a configured threshold.

The tracking process uses a protocol-specific resolution value to convert the actual metric in the routing table to a scaled metric in the range from 0 to 255. The resolution value is user-configurable and calculates the scaled metric by dividing a route's cost by the resolution value set for the route type:

- For intermediate system to intermediate system (ISIS), you can set the resolution in the range from 1 to 1000, where the default is **10**.
- For OSPF, you can set the resolution in the range from 1 to 1592, where the default is **1**.
- The resolution value used to map static routes is not configurable. By default, Dell EMC Networking OS assigns a metric of 0 to static routes.
- The resolution value used to map router information protocol (RIP) routes is not configurable. The RIP hop-count is automatically multiplied by 16 to scale it; a RIP metric of 16 (unreachable) scales to 256, which considers the route to be DOWN. For example, to configure object tracking for a RIP route to be considered UP only if the RIP hop count is less than or equal to 4, you would configure the UP threshold to be 64 (4 x 16) and the DOWN threshold to be 65.

Set Tracking Delays

You can configure an optional UP and/or DOWN timer for each tracked object to set the time delay before a change in the state of a tracked object is communicated to clients. The configured time delay starts when the state changes from UP to DOWN or the opposite way.

If the state of an object changes back to its former UP/DOWN state before the timer expires, the timer is cancelled and the client is not notified. If the timer expires and an object's state has changed, a notification is sent to the client. For example, if the DOWN timer is running when an interface goes down and comes back up, the DOWN timer is cancelled and the client is not notified of the event.

If you do not configure a delay, a notification is sent when a change in the state of a tracked object is detected. The time delay in communicating a state change is specified in seconds.

VRRP Object Tracking

As a client, VRRP can track up to 20 objects (including route entries, and Layer 2 and Layer 3 interfaces) in addition to the 12 tracked interfaces supported for each VRRP group.

You can assign a unique priority-cost value from 1 to 254 to each tracked VRRP object or group interface. The priority cost is subtracted from the VRRP group priority if a tracked VRRP object is in a DOWN state. If a VRRP group router acts as owner-master, the run-time VRRP group priority remains fixed at 255 and changes in the state of a tracked object have no effect.

NOTE: In VRRP object tracking, the sum of the priority costs for all tracked objects and interfaces cannot equal or exceed the priority of the VRRP group.

Object Tracking Configuration

You can configure three types of object tracking for a client.

- Track Layer 2 Interfaces
- Track Layer 3 Interfaces
- Track an IPv4/IPv6 Route

For a complete listing of all commands related to object tracking, refer to the *Dell EMC Networking OS Command Line Interface Reference Guide*.

Tracking a Layer 2 Interface

You can create an object that tracks the line-protocol state of a Layer 2 interface and monitors its operational status (UP or DOWN).

You can track the status of any of the following Layer 2 interfaces:

- 1 Gigabit Ethernet: Enter `gigabitethernet slot/port` in the `track interface interface` command (see Step 1).
- 10 Gigabit Ethernet: Enter `tengigabitethernet slot/port`.
- Port channel: Enter `port-channel number`.
- VLAN: Enter `vlan vlan-id`, where valid VLAN IDs are from 1 to 4094

A line-protocol object only tracks the link-level (UP/DOWN) status of a specified interface. When the link-level status goes down, the tracked object status is DOWN; if the link-level status is up, the tracked object status is UP.

To remove object tracking on a Layer 2 interface, use the `no track object-id` command.

To configure object tracking on the status of a Layer 2 interface, use the following commands.

1. Configure object tracking on the line-protocol state of a Layer 2 interface.
   CONFIGURATION mode
   `track object-id interface interface line-protocol`
   Valid object IDs are from 1 to 500.
2. (Optional) Configure the time delay used before communicating a change in the status of a tracked interface.
   OBJECT TRACKING mode
   `delay {[up seconds] [down seconds]}`
   Valid delay times are from 0 to 180 seconds.
   The default is **0**.
3. (Optional) Identify the tracked object with a text description.
   OBJECT TRACKING mode
   `description text`
   The text string can be up to 80 characters.
4. (Optional) Display the tracking configuration and the tracked object's status.
   EXEC Privilege mode

```
show track object-id
```

```
DellEMC(conf)#track 100 interface tengigabitethernet 1/1 line-protocol
DellEMC(conf-track-100)#delay up 20
DellEMC(conf-track-100)#description San Jose data center
DellEMC(conf-track-100)#end
DellEMC#show track 100

Track 100
  Interface TenGigabitEthernet 1/1 line-protocol
  Description: San Jose data center
```

Tracking a Layer 3 Interface

You can create an object that tracks the routing status of an IPv4 or IPv6 Layer 3 interface.

You can track the routing status of any of the following Layer 3 interfaces:

- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
- For a port channel interface, enter the keywords `port-channel` then a number.
- For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

For an IPv4 interface, a routing object only tracks the UP/DOWN status of the specified IPv4 interface (the `track interface ip-routing` command).

- The status of an IPv4 interface is UP only if the Layer 2 status of the interface is UP and the interface has a valid IP address.
- The Layer 3 status of an IPv4 interface goes DOWN when its Layer 2 status goes down (for a Layer 3 VLAN, all VLAN ports must be down) or the IP address is removed from the routing table.

For an IPv6 interface, a routing object only tracks the UP/DOWN status of the specified IPv6 interface (the `track interface ipv6-routing` command).

- The status of an IPv6 interface is UP only if the Layer 2 status of the interface is UP and the interface has a valid IPv6 address.
- The Layer 3 status of an IPv6 interface goes DOWN when its Layer 2 status goes down (for a Layer 3 VLAN, all VLAN ports must be down) or the IPv6 address is removed from the routing table.

To remove object tracking on a Layer 3 IPv4/IPv6 interface, use the `no track object-id` command.

To configure object tracking on the routing status of a Layer 3 interface, use the following commands.

1. Configure object tracking on the routing status of an IPv4 or IPv6 interface.

   CONFIGURATION mode

   `track object-id interface interface {ip routing | ipv6 routing}`

   Valid object IDs are from 1 to 500.
2. (Optional) Configure the time delay used before communicating a change in the status of a tracked interface.

   OBJECT TRACKING mode

   `delay {[up seconds] [down seconds]}`

   Valid delay times are from 0 to 180 seconds.

   The default is **0**.
3. (Optional) Identify the tracked object with a text description.

   OBJECT TRACKING mode

   `description text`

   The text string can be up to 80 characters.
4. (Optional) Display the tracking configuration and the tracked object's status.

   EXEC Privilege mode

   `show track object-id`

The following is an example of configuring object tracking for an IPv4 interface:

```
DellEMC(conf)#track 101 interface tengigabitethernet 1/2 ip routing
DellEMC(conf-track-101)#delay up 20
```

```
DellEMC(conf-track-101)#description NYC metro
DellEMC(conf-track-101)#end
DellEMC#show track 101

Track 101
  Interface TenGigabitEthernet 7/2 ip routing
  Description: NYC metro
```

The following is an example of configuring object tracking for an IPv6 interface:

```
DellEMC(conf)#track 103 interface tengigabitethernet 1/11 ipv6
routing
DellEMC(conf-track-103)#description Austin access point
DellEMC(conf-track-103)#end
DellEMC#show track 103

Track 103
  Interface TenGigabitEthernet 7/11 ipv6 routing
  Description: Austin access point
```

Track an IPv4/IPv6 Route

You can create an object that tracks the reachability or metric of an IPv4 or IPv6 route.

You specify the route to be tracked by its address and prefix-length values. Optionally, for an IPv4 route, you can enter a VRF instance name if the route is part of a VPN routing and forwarding (VRF) table. The next-hop address is not part of the definition of a tracked IPv4/IPv6 route.

In order for an route's reachability or metric to be tracked, the route must appear as an entry in the routing table. A tracked route is considered to match an entry in the routing table only if the exact IPv4 or IPv6 address and prefix length match an entry in the table. For example, when configured as a tracked route, 10.0.0.0/24 does not match the routing table entry 10.0.0.0/8. Similarly, for an IPv6 address, 3333:100:200:300:400::/80 does not match routing table entry 3333:100:200:300::/64. If no route-table entry has the exact IPv4/IPv6 address and prefix length, the tracked route is considered to be DOWN.

In addition to the entry of a route in the routing table, you can configure the UP/DOWN state of a tracked route to be determined in the following ways:

- By the reachability of the route's next-hop router.

  The UP/DOWN state of the route is determined by the entry of the next-hop address in the ARP cache. A tracked route is considered to be reachable if there is an ARP cache entry for the route's next-hop address. If the next-hop address in the ARP cache ages out for a route tracked for its reachability, an attempt is made to regenerate the ARP cache entry to see if the next-hop address appears before considering the route DOWN.

- By comparing the threshold for a route's metric with current entries in the route table.

  The UP/DOWN state of the tracked route is determined by the threshold for the current value of the route metric in the routing table.

  To provide a common tracking interface for different clients, route metrics are scaled in the range from 0 to 255, where 0 is connected and 255 is inaccessible. The scaled metric value communicated to a client always considers a lower value to have priority over a higher value. The resulting scaled value is compared against the configured threshold values to determine the state of a tracked route as follows:
  - If the scaled metric for a route entry is less than or equal to the UP threshold, the state of a route is UP.
  - If the scaled metric for a route is greater than or equal to the DOWN threshold or the route is not entered in the routing table, the state of a route is DOWN.

  The UP and DOWN thresholds are user-configurable for each tracked route. The default UP threshold is **254**; the default DOWN threshold is **255**. The notification of a change in the state of a tracked object is sent when a metric value crosses a configured threshold.

  The tracking process uses a protocol-specific resolution value to convert the actual metric in the routing table to a scaled metric in the range from 0 to 255. The resolution value is user-configurable and calculates the scaled metric by dividing a route's cost by the resolution value set for the route type:
  - For ISIS, you can set the resolution in the range from 1 to 1000, where the default is **10**.
  - For OSPF, you can set the resolution in the range from 1 to 1592, where the default is **1**.
  - The resolution value used to map static routes is not configurable. By default, Dell EMC Networking OS assigns a metric of 0 to static routes.

- The resolution value used to map RIP routes is not configurable. The RIP hop-count is automatically multiplied by 16 to scale it. For example, a RIP metric of 16 (unreachable) scales to 256, which considers a route to be DOWN.

Tracking Route Reachability

Use the following commands to configure object tracking on the reachability of an IPv4 or IPv6 route.

To remove object tracking, use the `no track object-id` command.

1. Configure object tracking on the reachability of an IPv4 or IPv6 route.

   CONFIGURATION mode

   `track object-id {ip route ip-address/prefix-len | ipv6 route ipv6-address/prefix-len} reachability [vrf vrf-name]`

   Valid object IDs are from 1 to 500.

   Enter an IPv4 address in dotted decimal format; valid IPv4 prefix lengths are from / 0 to /32.

   Enter an IPv6 address in X:X:X:X::X format; valid IPv6 prefix lengths are from / 0 to /128.

   (Optional) E-Series only: For an IPv4 route, you can enter a VRF name to specify the virtual routing table to which the tracked route belongs.

2. (Optional) Configure the time delay used before communicating a change in the status of a tracked route.

   OBJECT TRACKING mode

   `delay {[up seconds] [down seconds]}`

   Valid delay times are from 0 to 180 seconds.

   The default is **0**.

3. (Optional) Identify the tracked object with a text description.

   OBJECT TRACKING mode

   `description text`

   The text string can be up to 80 characters.

4. (Optional) Display the tracking configuration and the tracked object's status.

   EXEC Privilege mode

   `show track object-id`

The following example configures object tracking on the reachability of an IPv4 route:

```
DellEMC(conf)#track 104 ip route 10.0.0.0/8 reachability
DellEMC(conf-track-104)#delay up 20 down 10
DellEMC(conf-track-104)#end
DellEMC#show track 104

Track 104
  IP route 10.0.0.0/8 reachability
  Reachability is Down (route not in route table)
    2 changes, last change 00:02:49
  Tracked by:

DellEMC#configure
DellEMC(conf)#track 4 ip route 3.1.1.0/24 reachability vrf
vrf1
```

The following example configures object tracking on the reachability of an IPv6 route:

```
DellEMC(conf)#track 105 ipv6 route 1234::/64 reachability
DellEMC(conf-track-105)#delay down 5
DellEMC(conf-track-105)#description Headquarters
DellEMC(conf-track-105)#end
DellEMC#show track 105

Track 105
  IPv6 route 1234::/64 reachability
  Description: Headquarters
```

```
  Reachability is Down (route not in route table)
    2 changes, last change 00:03:03
```

Configuring track reachability refresh interval

If there is no entry in ARP table or if the next-hop address in the ARP cache ages out for a route tracked for its reachability, an attempt is made to check if the next-hop address is reachable after a certain refresh interval to see if the next-hop address appear in the ARP cache before considering it as DOWN.

You can change the refresh interval for which the next-hop address is checked. The default refresh interval is 60 seconds. The object tracking is done after the default refresh interval everytime and whenever there is a change in reachability state, the next-hop will be considered as DOWN. The default interval is changed using the `track reachability refresh` command. You can disable the track reachability feature by setting the refresh interval to 0.

To change the refresh interval for tracking an IPv4 or IPv6 route, use the following command.

Change the reachability refresh interval for tracking of an IPv4 or IPv6 route.

CONFIGURATION mode

`track reachability refresh` *`interval`*

The refresh interval range is from 0 to 60 seconds. The default is 60 seconds.

The following example shows how to change the refresh interval for tracking the reachability of the next-hop:

```
DellEMC#configure
DellEMC(conf)#track reachability refresh 20
```

For example, consider that the next-hop address is changed and the track reachability is checked after the set refresh interval (20 seconds). If the reachability to the next-hop address is failed, the system displays the following log stating that the track object state is changed from UP to DOWN.

```
Sep 28 11:08:57 %STKUNIT1-M:CP %OTM-6-STATE: Object 2 state transition from Up to Down.
```

Tracking a Metric Threshold

Use the following commands to configure object tracking on the metric threshold of an IPv4 or IPv6 route.

To remove object tracking, use the `no track` *`object-id`* command.

1. (Optional) Reconfigure the default resolution value used by the specified protocol to scale the metric for IPv4 or IPv6 routes.

   CONFIGURATION mode

   `track resolution {ip route | ipv6 route} {isis` *`resolution-value`* `| ospf` *`resolution-value`*`}`

   The range of resolution values is:
   - ISIS routes - 1 to 1000. The default is **1**.
   - OSPF routes - 1 to 1592. The efault is **1**.
2. Configure object tracking on the metric of an IPv4 or IPv6 route.

   CONFIGURATION mode

   `track object-id {ip route` *`ip-address/prefix-len`* `| ipv6 route` *`ipv6-address/prefix-len`*`}`
   `metric threshold [vrf` *`vrf-name`*`]`

   Valid object IDs are from 1 to 500.

   Enter an IPv4 address in dotted decimal format. Valid IPv4 prefix lengths are from /0 to /32.

   Enter an IPv6 address in X:X:X:X::X format. Valid IPv6 prefix lengths are from /0 to /128.

   (Optional) E-Series only: For an IPv4 route, you can enter a VRF name.
3. (Optional) Configure the time delay used before communicating a change in the UP and/or DOWN status of a tracked route.

   OBJECT TRACKING mode

   `delay {[up` *`seconds`*`] [down` *`seconds`*`]}`

   Valid delay times are from 0 to 180 seconds.

   The default is **0**.

4. (Optional) Identify the tracked object with a text description.

   OBJECT TRACKING mode

   `description text`

   The text string can be up to 80 characters.

5. (Optional) Configure the metric threshold for the UP and/or DOWN routing status to be tracked for the specified route.

   OBJECT TRACKING mode

   `threshold metric {[up number] [down number]}`

   The default UP threshold is **254**. The routing state is UP if the scaled route metric is less than or equal to the UP threshold.

   The defult DOWN threshold is **255**. The routing state is DOWN if the scaled route metric is greater than or equal to the DOWN threshold.

6. (Optional) Display the tracking configuration.

   EXEC Privilege mode

   `show track object-id`

Example of IPv4 and IPv6 Tracking Metric Thresholds

The following example configures object tracking on the metric threshold of an IPv4 route:

```
DellEMC(conf)#track 6 ip route 2.1.1.0/24 metric threshold
DellEMC(conf-track-6)#delay down 20
DellEMC(conf-track-6)#delay up 20
DellEMC(conf-track-6)#description track ip route metric
DellEMC(conf-track-6)#threshold metric down 40
DellEMC(conf-track-6)#threshold metric up 40
DellEMC(conf-track-6)#exit
DellEMC(conf)#track 10 ip route 3.1.1.0/24 metric threshold
vrf vrf1
```

Example of IPv4 and IPv6 Tracking Metric Thresholds

The following example configures object tracking on the metric threshold of an IPv6 route:

```
DellEMC(conf)#track 8 ipv6 route 2::/64 metric threshold
DellEMC(conf-track-8)#threshold metric up 30
DellEMC(conf-track-8)#threshold metric down 40
```

Displaying Tracked Objects

To display the currently configured objects used to track Layer 2 and Layer 3 interfaces, and IPv4 and IPv6 routes, use the following `show` commands.

To display the configuration and status of currently tracked Layer 2 or Layer 3 interfaces, IPv4 or IPv6 routes, or a VRF instance, use the `show track` command. You can also display the currently configured per-protocol resolution values used to scale route metrics when tracking metric thresholds.

- Display the configuration and status of currently tracked Layer 2 or Layer 3 interfaces, IPv4 or IPv6 routes, and a VRF instance.

  `show track [object-id [brief] | interface [brief] [vrf vrf-name] | ip route [brief] [vrf vrf-name] | resolution | vrf vrf-name [brief] | brief]`

- Use the `show running-config track` command to display the tracking configuration of a specified object or all objects that are currently configured on the router.

  `show running-config track [object-id]`

Examples of Viewing Tracked Objects

```
DellEMC#show track

Track 1
  IP route 23.0.0.0/8 reachability
  Reachability is Down (route not in route table)
    2 changes, last change 00:16:08
  Tracked by:
```

```
Track 2
  IPv6 route 2040::/64 metric threshold
  Metric threshold is Up (STATIC/0/0)
    5 changes, last change 00:02:16
  Metric threshold down 255 up 254
  First-hop interface is TenGigabitEthernet 1/2
  Tracked by:
    VRRP TenGigabitEthernet 2/30 IPv6 VRID 1

Track 3
  IPv6 route 2050::/64 reachability
  Reachability is Up (STATIC)
    5 changes, last change 00:02:16
  First-hop interface is TenGigabitEthernet 1/2
  Tracked by:
    VRRP TenGigabitEthernet 2/30 IPv6 VRID 1

Track 4
  Interface TenGigabitEthernet 1/4 ip routing
  IP routing is Up
   3 changes, last change 00:03:30
  Tracked by:
```

Example of the `show track brief` Command

```
Router# show track brief

ResId  Resource                Parameter
State  LastChange
1      IP route reachability   10.16.0.0/16
```

Example of the `show track resolution` Command

```
DellEMC#show track resolution

IP Route Resolution
  ISIS     1
  OSPF     1

IPv6 Route Resolution
  ISIS       1
```

Example of the `show track vrf` Command

```
DellEMC#show track vrf red

Track 5
  IP route 192.168.0.0/24 reachability, Vrf: red
  Reachability is Up (CONNECTED)
   3 changes, last change 00:02:39
   First-hop interface is TenGigabitEthernet 1/4
```

Example of Viewing Object Tracking Configuration

```
DellEMC#show running-config track

track 1 ip route 23.0.0.0/8 reachability

track 2 ipv6 route 2040::/64 metric threshold
delay down 3
delay up 5
threshold metric up 200

track 3 ipv6 route 2050::/64 reachability

track 4 interface TenGigabitEthernet 1/4 ip routing

track 5 ip route 192.168.0.0/24 reachability vrf red
```

35

Open Shortest Path First (OSPFv2 and OSPFv3)

Open shortest path first (OSPFv2 for IPv4) and OSPF version 3 (OSPF for IPv6) are supported on Dell EMC Networking OS.

This chapter provides a general description of OSPFv2 (OSPF for IPv4) and OSPFv3 (OSPF for IPv6) as supported in the Dell EMC Networking Operating System (OS).

NOTE: The fundamental mechanisms of OSPF (flooding, DR election, area support, SPF calculations, and so on) are the same between OSPFv2 and OSPFv3. This chapter identifies and clarifies the differences between the two versions of OSPF. Except where identified, the information in this chapter applies to both protocol versions.

OSPF protocol standards are listed in the Standards Compliance chapter.

Topics:

- Protocol Overview
- OSPF with Dell EMC Networking OS
- Configuration Information
- Sample Configurations for OSPFv2
- OSPFv3 NSSA
- Configuration Task List for OSPFv3 (OSPF for IPv6)

Protocol Overview

OSPF routing is a link-state routing protocol that calls for the sending of link-state advertisements (LSAs) to all other routers within the same autonomous system (AS) areas.

Information on attached interfaces, metrics used, and other variables is included in OSPF LSAs. As OSPF routers accumulate link-state information, they use the shortest path first (SPF) algorithm to calculate the shortest path to each node.

OSPF routers initially exchange HELLO messages to set up adjacencies with neighbor routers. The HELLO process is used to establish adjacencies between routers of the AS. It is not required that every router within the AS areas establish adjacencies. If two routers on the same subnet agree to become neighbors through the HELLO process, they begin to exchange network topology information in the form of LSAs.

In OSPFv2 neighbors on broadcast and NBMA links are identified by their interface addresses, while neighbors on other types of links are identified by RID.

Autonomous System (AS) Areas

OSPF operates in a type of hierarchy.

The largest entity within the hierarchy is the autonomous system (AS), which is a collection of networks under a common administration that share a common routing strategy. OSPF is an intra-AS (interior gateway) routing protocol, although it is capable of receiving routes from and sending routes to other ASs.

You can divide an AS into a number of areas, which are groups of contiguous networks and attached hosts. Routers with multiple interfaces can participate in multiple areas. These routers, called area border routers (ABRs), maintain separate databases for each area. Areas are a logical grouping of OSPF routers identified by an integer or dotted-decimal number.

Areas allow you to further organize your routers within in the AS. One or more areas are required within the AS. Areas are valuable in that they allow sub-networks to "hide" within the AS, thus minimizing the size of the routing tables on all routers. An area within the AS may not see the details of another area's topology. AS areas are known by their area number or the router's IP address.

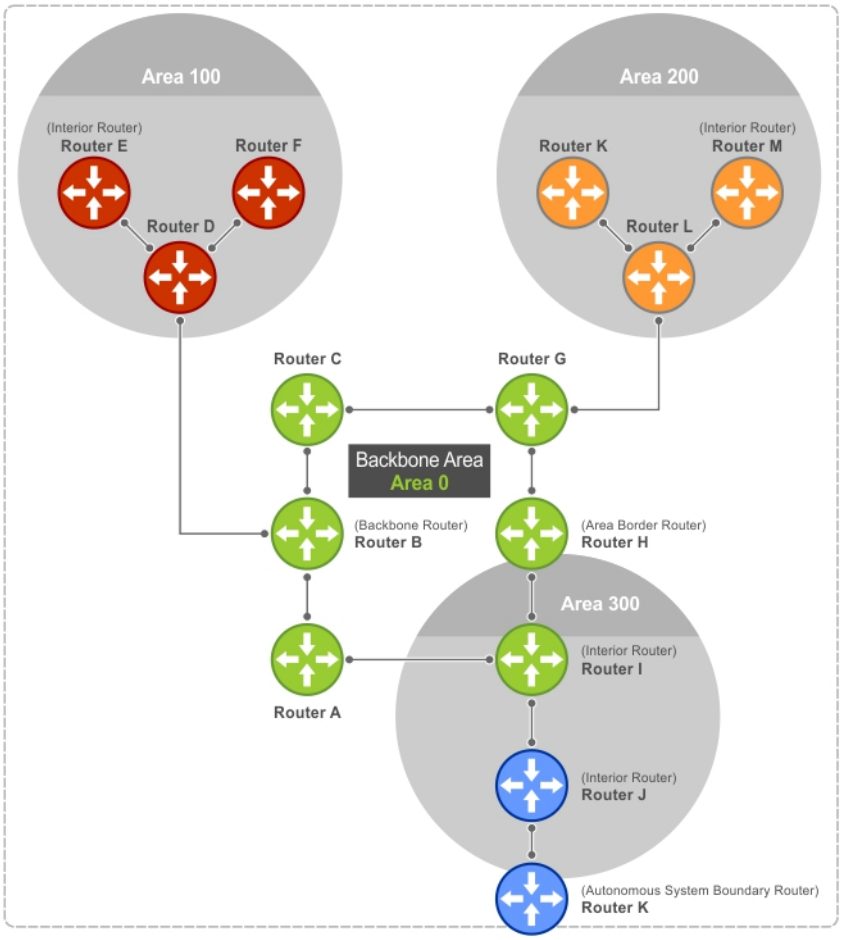


Figure 100. Autonomous System Areas

Area Types

The backbone of the network is Area 0. It is also called Area 0.0.0.0 and is the core of any AS. All other areas must connect to Area 0.

An OSPF backbone is responsible for distributing routing information between areas. It consists of all area border routers, networks not wholly contained in any area, and their attached routers.

NOTE: If you configure two non-backbone areas, then you must enable the B bit in OSPF.

The backbone is the only area with a default area number. All other areas can have their Area ID assigned in the configuration.

In the previous example, Routers A, B, C, G, H, and I are the Backbone.

- A stub area (SA) does not receive external route information, except for the default route. These areas do receive information from inter-area (IA) routes.
  NOTE: Configure all routers within an assigned stub area as stubby, and not generate LSAs that do not apply. For example, a Type 5 LSA is intended for external areas and the Stubby area routers may not generate external LSAs.
- A not-so-stubby area (NSSA) can import AS external route information and send it to the backbone. It cannot receive external AS information from the backbone or other areas.
- Totally stubby areas are referred to as no summary areas in the Dell EMC Networking OS.

Networks and Neighbors

As a link-state protocol, OSPF sends routing information to other OSPF routers concerning the state of the links between them. The state (up or down) of those links is important.

Routers that share a link become neighbors on that segment. OSPF uses the Hello protocol as a neighbor discovery and keep alive mechanism. After two routers are neighbors, they may proceed to exchange and synchronize their databases, which creates an adjacency.

Router Types

Router types are attributes of the OSPF process.

A given physical router may be a part of one or more OSPF processes. For example, a router connected to more than one area, receiving routing from a border gateway protocol (BGP) process connected to another AS acts as both an area border router and an autonomous system router.

Each router has a unique ID, written in decimal format (A.B.C.D). You do not have to associate the router ID with a valid IP address. However, to make troubleshooting easier, Dell EMC Networking recommends that the router ID and the router's IP address reflect each other.

The following example shows different router designations.

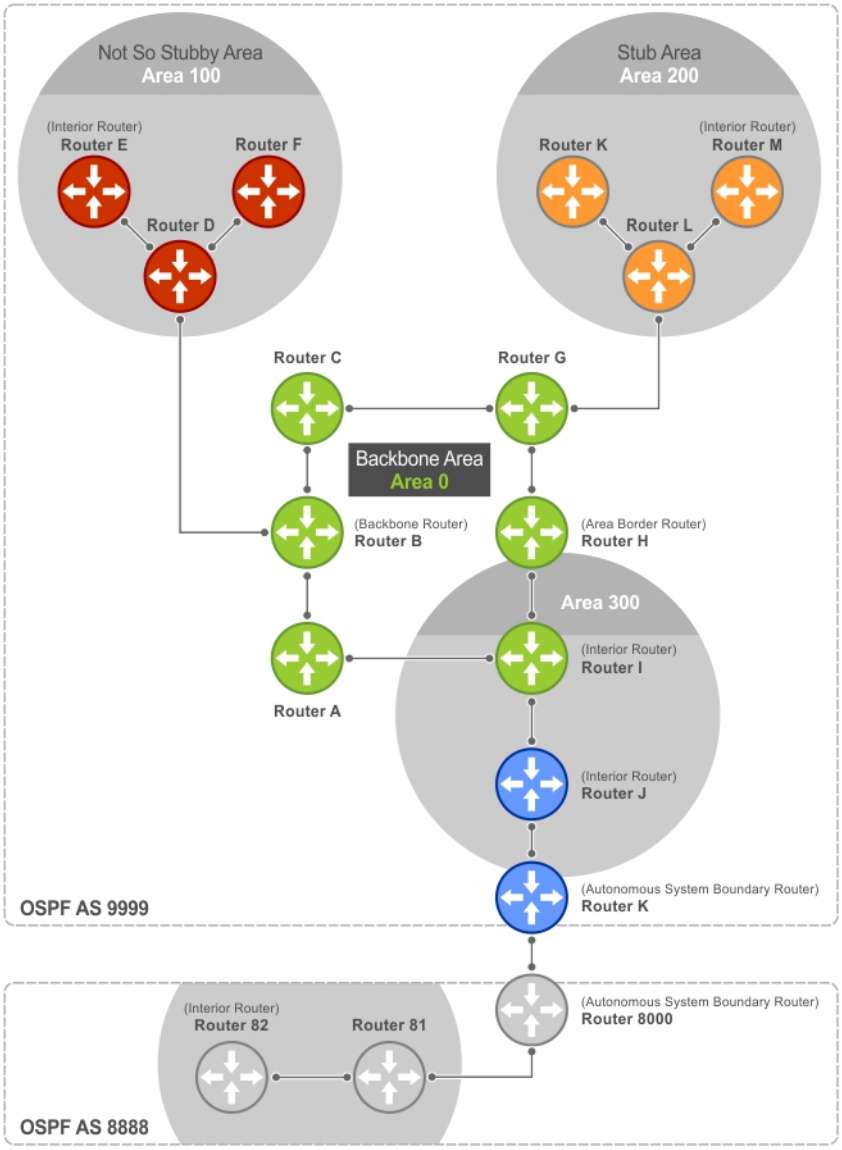


Figure 101. OSPF Routing Examples

Backbone Router (BR)

A backbone router (BR) is part of the OSPF Backbone, Area 0.

This includes all ABRs. It can also include any routers that connect only to the backbone and another ABR, but are only part of Area 0, such as Router I in the previous example.

Area Border Router (ABR)

Within an AS, an area border router (ABR) connects one or more areas to the backbone.

The ABR keeps a copy of the link-state database for every area it connects to, so it may keep multiple copies of the link state database. An ABR takes information it has learned on one of its attached areas and can summarize it before sending it out on other areas it is connected to.

An ABR can connect to many areas in an AS, and is considered a member of each area it connects to.

Autonomous System Border Router (ASBR)

The autonomous system border area router (ASBR) connects to more than one AS and exchanges information with the routers in other ASs.

Generally, the ASBR connects to a non-interior gate protocol (IGP) such as BGP or uses static routes.

Internal Router (IR)

The internal router (IR) has adjacencies with ONLY routers in the same area, as Router E, M, and I shown in the example in the Router Types.

Designated and Backup Designated Routers

OSPF elects a designated router (DR) and a backup designated router (BDR). Among other things, the DR is responsible for generating LSAs for the entire multiaccess network.

Designated routers allow a reduction in network traffic and in the size of the topological database.

- The DR maintains a complete topology table of the network and sends the updates to the other routers via multicast. All routers in an area form a slave/master relationship with the DR. Every time a router sends an update, the router sends it to the DR and BDR. The DR sends the update out to all other routers in the area.
- The BDR is the router that takes over if the DR fails.

Each router exchanges information with the DR and BDR. The DR and BDR relay the information to the other routers. On broadcast network segments, the number of OSPF packets is further reduced by the DR and BDR sending such OSPF updates to a multicast IP address that all OSPF routers on the network segment are listening on.

These router designations are not the same ad the router IDs described earlier. The DRs and BDRs are configurable in Dell EMC Networking OS. If you do not define DR or BDR, the system assigns them. OSPF looks at the priority of the routers on the segment to determine which routers are the DR and BDR. The router with the highest priority is elected the DR. If there is a tie, the router with the higher router ID takes precedence. After the DR is elected, the BDR is elected the same way. A router with a router priority set to zero cannot become the DR or BDR.

Link-State Advertisements (LSAs)

A link-state advertisement (LSA) communicates the router's local routing topology to all other local routers in the same area.

The LSA types supported by Dell EMC Networking are defined as follows:

- **Type 1: Router LSA** — The router lists links to other routers or networks in the same area. Type 1 LSAs are flooded across their own area only. The link-state ID of the Type 1 LSA is the originating router ID.
- **Type 2: Network LSA** — The DR in an area lists which routers are joined within the area. Type 2 LSAs are flooded across their own area only. The link-state ID of the Type 2 LSA is the IP interface address of the DR.

- **Type 3: Summary LSA (OSPFv2), Inter-Area-Prefix LSA (OSPFv3)** — An ABR takes information it has learned on one of its attached areas and can summarize it before sending it out on other areas it is connected to. The link-state ID of the Type 3 LSA is the destination network number.
- **Type 4: AS Border Router Summary LSA (OSPFv2), Inter-Area-Router LSA (OSPFv3)** — In some cases, Type 5 External LSAs are flooded to areas where the detailed next-hop information may not be available. An ABR floods the information for the router (for example, the ASBR where the Type 5 advertisement originated. The link-state ID for Type 4 LSAs is the router ID of the described ASBR).
- **Type 5: LSA** — These LSAs contain information imported into OSPF from other routing processes. They are flooded to all areas, except stub areas. The link-state ID of the Type 5 LSA is the external network number.
- **Type 7: External LSA** — Routers in an NSSA do not receive external LSAs from ABRs, but are allowed to send external routing information for redistribution. They use Type 7 LSAs to tell the ABRs about these external routes, which the ABR then translates to Type 5 external LSAs and floods as normal to the rest of the OSPF network.
- **Type 8: Link LSA (OSPFv3)** — This LSA carries the IPv6 address information of the local links.
- **Type 9: Link Local LSA (OSPFv2), Intra-Area-Prefix LSA (OSPFv3)** — For OSPFv2, this is a link-local "opaque" LSA as defined by RFC2370. For OSPFv3, this LSA carries the IPv6 prefixes of the router and network links.
- **Type 11 - Grace LSA (OSPFv3)** — For OSPFv3 only, this LSA is a link-local “opaque” LSA sent by a restarting OSPFv3 router during a graceful restart.

For all LSA types, there are 20-byte LSA headers. One of the fields of the LSA header is the link-state ID.

Each router link is defined as one of four types: type 1, 2, 3, or 4. The LSA includes a link ID field that identifies, by the network number and mask, the object this link connects to.

Depending on the type, the link ID has different meanings.

- 1: point-to-point connection to another router/neighboring router.
- 2: connection to a transit network IP address of the DR.
- 3: connection to a stub network IP network/subnet number.

LSA Throttling

LSA throttling provides configurable interval timers to improve OSPF convergence times.

The default OSPF static timers (5 seconds for transmission, 1 second for acceptance) ensures sufficient time for sending and resending LSAs and for system acceptance of arriving LSAs. However, some networks may require reduced intervals for LSA transmission and acceptance. Throttling timers allow for this improved convergence times.

The LSA throttling timers are configured in milliseconds, with the interval time increasing exponentially until a maximum time has been reached. If the maximum time is reached, the system, the system continues to transmit at the max-interval until twice the max-interval time has passed. At that point, the system reverts to the start-interval timer and the cycle begins again.

When you configure the LSA throttle timers, syslog messages appear, indicating the interval times, as shown below for the transmit timer (45000ms) and arrival timer (1000ms).

```
Mar 15 09:46:00: %STKUNIT0-M:CP %OSPF-4-LSA_BACKOFF: OSPF Process 10,Router lsa id
2.2.2.2 router-id 2.2.2.2 is backed off to transmit after 45000ms
```

```
Mar 15 09:46:06: %STKUNIT0-M:CP %OSPF-4-LSA_BACKOFF: OSPF Process 10,Router lsa id
3.3.3.3 rtrid 3.3.3.3 received before 1000ms time
```

NOTE: The sequence numbers are reset when previously cleared routes that are waiting for the LSA throttle timer to expire are re-enabled.

Router Priority and Cost

Router priority and cost is the method the system uses to “rate” the routers.

For example, if not assigned, the system selects the router with the highest priority as the DR. The second highest priority is the BDR.

- Priority is a numbered rating 0 to 255. The higher the number, the higher the priority.
- Cost is a numbered rating 1 to 65535. The higher the number, the greater the cost. The cost assigned reflects the cost should the router fail. When a router fails and the cost is assessed, a new priority number results.

Figure 102. Priority and Cost Examples

OSPF with Dell EMC Networking OS

The Dell EMC Networking OS supports up to 16,000 OSPF routes for OSPFv2.

Dell EMC Networking OS version 9.4(0.0) and later support only one OSPFv2 process per VRF. Dell EMC Networking OS version 9.7(0.0) and later support OSPFv3 in VRF. Also, on OSPFv3, Dell EMC Networking OS supports only one OSPFv3 process per VRF.

OSPFv2 and OSPFv3 can co-exist but you must configure them individually.

Dell EMC Networking OS supports stub areas, totally stub (no summary) and not so stubby areas (NSSAs) and supports the following LSAs, as described earlier.

- Router (type 1)
- Network (type 2)
- Network Summary (type 3)
- AS Boundary (type 4)
- LSA(type 5)
- External LSA (type 7)
- Link LSA, OSPFv3 only (type 8)
- Opaque Link-Local (type 9)
- Grace LSA, OSPFv3 only (type 11)

Graceful Restart

Graceful restart for OSPFv2 and OSPFv3 are supported on the S4820T platform in Helper and Restart modes.

When a router goes down without a graceful restart, there is a possibility for loss of access to parts of the network due to ongoing network topology changes. Additionally, LSA flooding and reconvergence can cause substantial delays. It is, therefore, desirable that the network maintains a stable topology if it is possible for data flow to continue uninterrupted.

OSPF graceful restart understands that in a modern router, the control plane and data plane functionality are separate, restarting the control plane functionality (such as the failover of the active RPM to the backup in a redundant configuration), does not necessarily have to interrupt the forwarding of data packets. This behavior is supported because the forwarding tables previously computed by an active RPM have been downloaded into the forwarding information base (FIB) on the line cards (the data plane) and are still resident. For packets that have existing FIB/CAM entries, forwarding between ingress and egress ports/VLANs, and so on, can continue uninterrupted while the control plane OSPF process comes back to full functionality and rebuilds its routing tables.

To notify its helper neighbors that the restart process is beginning, when a router is attempting to restart gracefully, it originates the following link-local Grace LSAs:

- An OSPFv2 router sends Type 9 LSAs.
- An OSPFv3 router sends Type 11 LSAs.

Type 9 and 11 LSAs include a grace period, which is the time period an OSPF router advertises to adjacent neighbor routers as the time to wait for it to return to full control plane functionality. During the grace period, neighbor OSPFv2 /v3 interfaces save the LSAs from the restarting OSPF interface. Helper neighbor routers continue to announce the restarting router as fully adjacent, as long as the network topology remains unchanged. When the restarting router completes its restart, it flushes the Type 9 and 11 LSAs, notifying its neighbors that the restart is complete. This notification happens before the grace period expires.

Dell EMC Networking routers support the following OSPF graceful restart functionality:

- Restarting role in which an enabled router performs its own graceful restart.
- Helper role in which the router's graceful restart function is to help a restarting neighbor router in its graceful restarts.
- Helper-reject role in which OSPF does not participate in the graceful restart of a neighbor.

  OSPFv2 supports *helper-only* and *restarting-only* roles. By default, both helper and restarting roles are enabled. OSPFv2 supports the helper-reject role globally on a router.

  OSPFv3 supports *helper-only* and *restarting-only* roles. The helper-only role is enabled by default. To enable the restarting role in addition to the helper-only role, configure a grace period. Reconfigure OSPFv3 graceful restart to a restarting-only role when you enable the helper-reject role on an interface. OSPFv3 supports the helper-reject role on a per-interface basis.

  Configuring helper-reject role on an OSPFv2 router or OSPFv3 interface enables the restarting-only role globally on the router or locally on the interface. In a helper-reject role, OSPF does not participate in the graceful restart of an adjacent OSPFv2/v3 router.

  If multiple OSPF interfaces provide communication between two routers, after you configure helper-reject on one interface, all other interfaces between the two routers behave as if they are in the help-reject role.

- OSPFv2 and OSPFv3 support planned-only and/or unplanned-only restarts. The default is support for both planned and unplanned restarts.

  A planned restart occurs when you enter the `redundancy force-failover rpm` command to force the primary RPM to switch to the backup RPM. During a planned restart, OSPF sends out a Grace LSA before the system switches over to the backup RPM.

  An unplanned restart occurs when an unplanned event causes the active RPM to switch to the backup RPM, such as when an active process crashes, the active RPM is removed, or a power failure happens. During an unplanned restart, OSPF sends out a Grace LSA when the backup RPM comes online.

To display the configuration values for OSPF graceful restart, enter the `show run ospf` command for OSPFv2 and the `show run ospf` and `show ipv6 ospf [vrf vrf-name] database database-summary` commands for OSPFv3.

Fast Convergence (OSPFv2, IPv4 Only)

Fast convergence allows you to define the speeds at which LSAs are originated and accepted, and reduce OSPFv2 end-to-end convergence time.

Dell EMC Networking OS allows you to accept and originate LSAs as soon as they are available to speed up route information propagation.

NOTE: The faster the convergence, the more frequent the route calculations and updates. This impacts CPU utilization and may impact adjacency stability in larger topologies.

Multi-Process OSPFv2 with VRF

Multi-process OSPF with VRF is supported on the Dell EMC Networking OS. Only one OSPFv2 process per VRF is supported.

Multi-process OSPF allows multiple OSPFv2 processes on a single router. Multiple OSPFv2 processes allow for isolating routing domains, supporting multiple route policies and priorities in different domains, and creating smaller domains for easier management. Each OSPFv2 process has a unique process ID and must have an associated router ID. There must be an equal number of interfaces and must be in Layer-3 mode for the number of processes created. For example, if you create five OSPFv2 processes on a system, there must be at least five interfaces assigned in Layer 3 mode. Each OSPFv2 process is independent. If one process loses adjacency, the other processes continue to function.

Processing SNMP and Sending SNMP Traps

Only the process in `default vrf` can process the SNMP requests and send SNMP traps.

NOTE: SNMP gets request corresponding to the OspfNbrOption field in the OspfNbrTable returns a value of 66.

RFC-2328 Compliant OSPF Flooding

In OSPF, flooding is the most resource-consuming task. The flooding algorithm described in RFC 2328 requires that OSPF flood LSAs on all interfaces, as governed by LSA's flooding scope (refer to Section 13 of the RFC.)

When multiple direct links connect two routers, the RFC 2328 flooding algorithm generates significant redundant information across all links.

By default, Dell EMC Networking OS implements an enhanced flooding procedure which dynamically and intelligently detects when to optimize flooding. Wherever possible, the OSPF task attempts to reduce flooding overhead by selectively flooding on a subset of the interfaces between two routers.

Enabling RFC-2328 Compliant OSPF Flooding

To enable OSPF flooding, use the following command.

When you enable this command, it configures Dell EMC Networking OS to flood LSAs on all interfaces.

- Enable RFC 2328 flooding.

  ROUTER OSPF mode

  `flood-2328`

To confirm RFC 2328 flooding behavior, use the `debug ip ospf packet` command. The following example shows no change in the updated packets (shown in bold).

ACKs 2 (shown in bold) is printed only for ACK packets.

The following example shows no change in the updated packets (shown in bold). ACKs 2 (shown in bold) is printed only for ACK packets.

```
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:5(LSAck) l:64 Acks 2 rid:2.2.2.2
    aid:1500 chk:0xdbee aut:0 auk: keyid:0 from:Vl 1000
       LSType:Type-5 AS External id:160.1.1.0 adv:6.1.0.0 seq:0x8000000c
       LSType:Type-5 AS External id:160.1.2.0 adv:6.1.0.0 seq:0x8000000c
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:5(LSAck) l:64 Acks 2 rid:2.2.2.2
    aid:1500 chk:0xdbee aut:0 auk: keyid:0 from:Vl 100
       LSType:Type-5 AS External id:160.1.1.0 adv:6.1.0.0 seq:0x8000000c
       LSType:Type-5 AS External id:160.1.2.0 adv:6.1.0.0 seq:0x8000000c
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:4(LSUpd) l:100 rid:6.1.0.0
    aid:0 chk:0xccbd aut:0 auk: keyid:0 from:Gi 10/21
       Number of LSA:2
      LSType:Type-5 AS External(5) Age:1 Seq:0x8000000c id:170.1.1.0 Adv:6.1.0.0
        Netmask:255.255.255.0 fwd:0.0.0.0 E2, tos:0 metric:0
      LSType:Type-5 AS External(5) Age:1 Seq:0x8000000c id:170.1.2.0 Adv:6.1.0.0
        Netmask:255.255.255.0 fwd:0.0.0.0 E2, tos:0 metric:0
```

RFC 2328 is supported by default on Dell EMC Networking OS and it is indicated in the `show ip ospf` command output.

```
DellEMC#show ip ospf
Routing Process ospf 1 with ID 2.2.2.2
Supports only single TOS (TOS0) routes
It is an Autonomous System Boundary Router
It is Flooding according to RFC 2328
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Number of area in this router is 1, normal 0 stub 0 nssa 1
--More--
```

OSPF ACK Packing

The OSPF ACK packing feature bundles multiple LS acknowledgements in a single packet, significantly reducing the number of ACK packets transmitted when the number of LSAs increases.

This feature also enhances network utilization and reduces the number of small ACK packets sent to a neighboring router. OSPF ACK packing is enabled by default and non-configurable.

Setting OSPF Adjacency with Cisco Routers

To establish an OSPF adjacency between Dell EMC Networking and Cisco routers, the hello interval and dead interval must be the same on both routers.

In Dell EMC Networking OS, the OSPF dead interval value is, by default, set to **40 seconds**, and is independent of the OSPF hello interval. Configuring a hello interval does not change the dead interval in Dell EMC Networking OS. In contrast, the OSPF dead interval on a Cisco router is, by default, four times as long as the hello interval. Changing the hello interval on the Cisco router automatically changes the dead interval.

To ensure equal intervals between the routers, use the following command.

- Manually set the dead interval of the Dell EMC Networking router to match the Cisco configuration.

  INTERFACE mode

  `ip ospf dead-interval <x>`

In the following example, the dead interval is set at 4x the hello interval (shown in bold).

```
DellEMC(conf)#int tengigabitethernet 2/2
DellEMC(conf-if-te-2/2)#ip ospf hello-interval 20
Dell(conf-if-te-2/2)#ip ospf dead-interval 80

DellEMC(conf-if-te-2/2)#
```

In the following example, the dead interval is set at 4x the hello interval (shown in bold).

```
DellEMC(conf-if-te-2/2)#ip ospf dead-interval 20
DellEMC(conf-if-te-2/2)#do show ip os int tengigabitethernet 1/3
TenGigabitEthernet 2/2 is up, line protocol is up
  Internet Address 20.0.0.1/24, Area 0
  Process ID 10, Router ID 1.1.1.2, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 1.1.1.2, Interface address 30.0.0.1
  Backup Designated Router (ID) 1.1.1.1, Interface address 30.0.0.2
  Timer intervals configured, Hello 20, Dead 80, Wait 20, Retransmit 5
  Hello due in 00:00:04
  Neighbor Count is 1, Adjacent neighbor count is 1
  Adjacent with neighbor 1.1.1.1 (Backup Designated Router)
DellEMC(conf-if-te-2/2)#
```

Configuration Information

The interfaces must be in Layer 3 mode (assigned an IP address) and enabled so that they can send and receive traffic. The OSPF process must know about these interfaces.

To make the OSPF process aware of these interfaces, they must be assigned to OSPF areas.

You must configure OSPF GLOBALLY on the system in CONFIGURATION mode.

NOTE: Loop back routes are not installed in the Route Table Manager (RTM) as non-active routes.

OSPF features and functions are assigned to each router using the `CONFIG-INTERFACE` commands for each interface.

NOTE: By default, OSPF is disabled.

Configuration Task List for OSPFv2 (OSPF for IPv4)

You can perform the following tasks to configure Open Shortest Path First version 2 (OSPF for IPv4) on the switch. Two of the tasks are mandatory; others are optional.

The following configuration tasks include two mandatory tasks and several optional tasks:

- Enabling OSPFv2 (mandatory)
- Assigning a Router ID
- Assigning an OSPFv2 Area (mandatory)
- Enable OSPFv2 on Interfaces
- Configuring Stub Areas
- Enabling Passive Interfaces
- Enabling Fast-Convergence
- Changing OSPFv2 Parameters on Interfaces
- Enabling OSPFv2 Authentication
- Creating Filter Routes
- Applying Prefix Lists
- Redistributing Routes
- Troubleshooting OSPFv2

1. Configure a physical interface. Assign an IP address, physical or Loopback, to the interface to enable Layer 3 routing.
2. Enable OSPF globally. Assign network area and neighbors.
3. Add interfaces or configure other attributes.
4. Set the time interval between when the switch receives a topology change and starts a shortest path first (SPF) calculation.

   Use `timers spf delay holdtime`

   NOTE: To set the interval time between the reception of topology changes and calculation of SPF in milli seconds, use the `timers spf delay holdtime msec` command.

   Example

```
DellEMC#
DellEMC#conf
DellEMC(conf)#router ospf 1
DellEMC(conf-router_ospf-1)#timer spf 2 5 msec
DellEMC(conf-router_ospf-1)#
DellEMC(conf-router_ospf-1)#show config
!
router ospf 1
timers spf 2 5 msec
DellEMC(conf-router_ospf-1)#
DellEMC(conf-router_ospf-1)#end
DellEMC#
```

For a complete list of the OSPF commands, refer to the *OSPF* section in the *Dell EMC Networking OS Command Line Reference Guide* document.

Enabling OSPFv2

To enable Layer 3 routing, assign an IP address to an interface (physical or Loopback). By default, OSPF, similar to all routing protocols, is disabled.

You *must* configure at least one interface for Layer 3 before enabling OSPFv2 globally.

If implementing multi-process OSPF, create an equal number of Layer 3 enabled interfaces and OSPF process IDs. For example, if you create four OSPFv2 process IDs, you must have four interfaces with Layer 3 enabled.

1. Assign an IP address to an interface.

   CONFIG-INTERFACE mode

   `ip address ip-address mask`

   The format is A.B.C.D/M.

   If you are using a Loopback interface, refer to Loopback Interfaces.
2. Enable the interface.

   CONFIG-INTERFACE mode

   `no shutdown`
3. Return to CONFIGURATION mode to enable the OSPFv2 process globally.

   CONFIGURATION mode

   `router ospf process-id [vrf {vrf name}]`

   - *vrf name*: enter the keyword `VRF` and the instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are later tied to the VRF instance.

   The range is from 0 to 65535.

   The OSPF process ID is the identifying number assigned to the OSPF process. The router ID is the IP address associated with the OSPF process.

   After the OSPF process and the VRF are tied together, the OSPF process ID cannot be used again in the system.

   If you try to enter an OSPF process ID, or if you try to enable more OSPF processes than available Layer 3 interfaces, prior to assigning an IP address to an interface and setting the `no shutdown` command, the following message displays:

```
DellEMC(conf)#router ospf 1
% Error: No router ID available.
```

Assigning a Router ID

In CONFIGURATION ROUTER OSPF mode, assign the router ID.

The router ID is not required to be the router's IP address. However, Dell EMC Networking recommends using the IP address as the router ID for easier management and troubleshooting. Optional *process-id* commands are also described.

- Assign the router ID for the OSPFv2 process.

  CONFIG-ROUTER-OSPF-id mode

  `router-id ip address`
- Disable OSPF.

  CONFIGURATION mode

  `no router ospf process-id`
- Reset the OSPFv2 process.

  EXEC Privilege mode

  `clear ip ospf process-id`
- View the current OSPFv2 status.

  EXEC mode

  `show ip ospf process-id`

```
DellEMC#show ip ospf 55555
Routing Process ospf 55555 with ID 10.10.10.10
Supports only single TOS (TOS0) routes
```

```
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
DellEMC#
```

Enabling Multi-Process OSPF (OSPFv2, IPv4 Only)

Multi-process OSPF allows multiple OSPFv2 processes on a single router.

For more information, refer to Multi-Process OSPF (OSPFv2, IPv4 Only)

When configuring a single OSPF process, follow the same steps previously described. Repeat them as often as necessary for the desired number of processes. After the process is created, all other configurations apply as usual.

1. Assign an IP address to an interface.

   CONFIG-INTERFACE mode

   `ip address ip-address mask`

   Format: A.B.C.D/M.

   If you are using a Loopback interface, refer to Loopback Interfaces.
2. Enable the interface.

   CONFIG-INTERFACE mode

   `no shutdown`
3. Return to CONFIGURATION mode to enable the OSPFv2 process globally.

   CONFIGURATION mode

   `router ospf process-id [vrf]`

   The range is from 0 to 65535.

   After the OSPF process and the VRF are tied together, the OSPF process ID cannot be used again in the system.

   If you try to enable more OSPF processes than available Layer 3 interfaces, the following message displays:

   ```
   Dell(conf)#router ospf 1
   % Error: No router ID available.
   ```

Assigning an OSPFv2 Area

After you enable OSPFv2, assign the interface to an OSPF area. Set up OSPF areas and enable OSPFv2 on an interface with the `network` command.

You must have at least one AS area: Area 0. This is the backbone area. If your OSPF network contains more than one area, configure a backbone area (Area ID 0.0.0.0). Any area besides Area 0 can have any number ID assigned to it.

The OSPFv2 process evaluates the `network` commands in the order they are configured. Assign the network address that is most explicit first to include all subnets of that address. For example, if you assign the network address 10.0.0.0 /8, you cannot assign the network address 10.1.0.0 /16 because it is already included in the first network address.

When configuring the `network` command, configure a network address and mask that is a superset of the IP subnet configured on the Layer-3 interface for OSPFv2 to use.

You can assign the area in the following step by a number or with an IP interface address.

- Enable OSPFv2 on an interface and assign a network address range to a specific OSPF area.

  CONFIG-ROUTER-OSPF-id mode

  `network ip-address mask area area-id`

  The IP Address Format is A.B.C.D/M.

  The area ID range is from 0 to 65535 or A.B.C.D/M.

Enable OSPFv2 on Interfaces

Enable and configure OSPFv2 on each interface (configure for Layer 3 protocol), and not shutdown.

You can also assign OSPFv2 to a Loopback interface as a virtual interface.

OSPF functions and features, such as MD5 Authentication, Grace Period, Authentication Wait Time, are assigned on a per interface basis.

NOTE: If using features like MD5 Authentication, ensure all the neighboring routers are also configured for MD5.

In the example below, an IP address is assigned to an interface and an OSPFv2 area is defined that includes the IP address of a Layer 3 interface.

The first bold lines assign an IP address to a Layer 3 interface, and the`no shutdown` command ensures that the interface is UP.

The second bold line assigns the IP address of an interface to an area.

Example of Enabling OSPFv2 and Assigning an Area to an Interface

```
DellEMC#(conf)#int te 4/14
Dell(conf-if-te-4/14)#ip address 10.10.10.10/24
DellEMC(conf-if-te-4/14)#no shutdown
DellEMC(conf-if-te-4/14)#ex
DellEMC(conf)#router ospf 1
DellEMC(conf-router_ospf-1)#network 1.2.3.4/24 area 0
DellEMC(conf-router_ospf-1)#network 10.10.10.10/24 area 1
DellEMC(conf-router_ospf-1)#network 20.20.20.20/24 area 2
DellEMC(conf-router_ospf-1)#
DellEMC#
```

Dell EMC Networking recommends using the interface IP addresses for the OSPFv2 router ID for easier management and troubleshooting.

To view the configuration, use the `show config` command in CONFIGURATION ROUTER OSPF mode.

OSPF, by default, sends hello packets out to all physical interfaces assigned an IP address that is a subset of a network on which OSPF is enabled.

To view currently active interfaces and the areas assigned to them, use the `show ip ospf interface` command.

Example of Viewing Active Interfaces and Assigned Areas

```
DellEMC>show ip ospf 1 interface

TenGigabitEthernet 1/17 is up, line protocol is up
  Internet Address 10.2.2.1/24, Area 0.0.0.0
  Process ID 1, Router ID 11.1.2.1, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 11.1.2.1, Interface address 10.2.2.1
  Backup Designated Router (ID) 0.0.0.0, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:04
  Neighbor Count is 0, Adjacent neighbor count is 0

TenGigabitEthernet 1/21 is up, line protocol is up
  Internet Address 10.2.3.1/24, Area 0.0.0.0
  Process ID 1, Router ID 11.1.2.1, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State BDR, Priority 1
  Designated Router (ID) 13.1.1.1, Interface address 10.2.3.2
  Backup Designated Router (ID) 11.1.2.1, Interface address 10.2.3.1
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:05
  Neighbor Count is 1, Adjacent neighbor count is 1
    Adjacent with neighbor 13.1.1.1 (Designated Router)
DellEMC>
```

Loopback interfaces also help the OSPF process. OSPF picks the highest interface address as the router-id and a Loopback interface address has a higher precedence than other interface addresses.

Example of Viewing OSPF Status on a Loopback Interface

```
DellEMC#show ip ospf 1 int

TenGigabitEthernet 1/23 is up, line protocol is up
  Internet Address 10.168.0.1/24, Area 0.0.0.1
  Process ID 1, Router ID 10.168.253.2, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State DROTHER, Priority 1
  Designated Router (ID) 10.168.253.5, Interface address 10.168.0.4
  Backup Designated Router (ID) 192.168.253.3, Interface address 10.168.0.2
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:08
  Neighbor Count is 3, Adjacent neighbor count is 2
  Adjacent with neighbor 10.168.253.5 (Designated Router)
  Adjacent with neighbor 10.168.253.3 (Backup Designated Router)

Loopback 0 is up, line protocol is up
  Internet Address 10.168.253.2/32, Area 0.0.0.1
  Process ID 1, Router ID 10.168.253.2, Network Type LOOPBACK, Cost: 1
Loopback interface is treated as a stub Host.
DellEMC#
```

Configuring Stub Areas

OSPF supports different types of LSAs to help reduce the amount of router processing within the areas.

Type 5 LSAs are not flooded into stub areas; the ABR advertises a default route into the stub area to which it is attached. Stub area routers use the default route to reach external destinations.

To ensure connectivity in your OSPFv2 network, never configure the backbone area as a stub area.

To configure a stub area, use the following commands.

1. Review all areas after they were configured to determine which areas are NOT receiving type 5 LSAs.

   EXEC Privilege mode

   `show ip ospf process-id [vrf] database database-summary`
2. Enter CONFIGURATION mode.

   EXEC Privilege mode

   `configure`
3. Enter ROUTER OSPF mode.

   CONFIGURATION mode

   `router ospf` *`process-id`* `[vrf]`

   Process ID is the ID assigned when configuring OSPFv2 globally.
4. Configure the area as a stub area.

   CONFIG-ROUTER-OSPF-id mode

   `area` *`area-id`* `stub [no-summary]`

   Use the keywords `no-summary` to prevent transmission into the area of summary ASBR LSAs.

   Area ID is the number or IP address assigned when creating the area.

To view which LSAs are transmitted, use the `show ip ospf database` *`process-id`* `database-summary` command in EXEC Privilege mode.

```
DellEMC#show ip ospf 34 database database-summary

      OSPF Router with ID (10.1.2.100) (Process ID 34)

Area      ID Router Network S-Net S-ASBR Type-7 Subtotal
2.2.2.2   1         0       0     0      0      1
3.3.3.3   1         0       0     0      0      1
DellEMC#
```

To view information on areas, use the `show ip ospf` *`process-id`* command in EXEC Privilege mode.

Enabling Passive Interfaces

A passive interface is one that does not send or receive routing information.

Enabling passive interface suppresses routing updates on an interface. Although the passive interface does not send or receive routing updates, the network on that interface is still included in OSPF updates sent via other interfaces.

To suppress the interface's participation on an OSPF interface, use the following command. This command stops the router from sending updates on that interface.

- Specify whether all or some of the interfaces are passive.

  CONFIG-ROUTEROSPF- id mode

  `passive-interface {default | interface}`

  The default is enabled passive interfaces on ALL interfaces in the OSPF process.

  Entering the physical interface type, slot, and number enables passive interface on only the identified interface.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094 (for example, passive-interface vlan 2222).

  The keyword `default` sets all interfaces on this OSPF process as passive.

  To remove the passive interface from select interfaces, use the `no passive-interface` *`interface`* command while `passive interface default` is configured.

  To enable both receiving and sending routing updates, use the `no passive-interface` *`interface`* command.

When you configure a passive interface, the `show ip ospf` *`process-id`* `interface` command adds the words `passive interface` to indicate that the hello packets are not transmitted on that interface (shown in bold).

```
DellEMC#show ip ospf 34 int

TenGigabitEthernet 1/1 is up, line protocol is down
  Internet Address 10.1.2.100/24, Area 1.1.1.1
  Process ID 34, Router ID 10.1.2.100, Network Type BROADCAST, Cost: 10
  Transmit Delay is 1 sec, State DOWN, Priority 1
  Designated Router (ID) 10.1.2.100, Interface address 0.0.0.0
  Backup Designated Router (ID) 0.0.0.0, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 13:39:46
  Neighbor Count is 0, Adjacent neighbor count is 0

TenGigabitEthernet 2/1 is up, line protocol is down
  Internet Address 10.1.3.100/24, Area 2.2.2.2
  Process ID 34, Router ID 10.1.2.100, Network Type BROADCAST, Cost: 10
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 10.1.2.100, Interface address 10.1.3.100
  Backup Designated Router (ID) 0.0.0.0, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    No Hellos (Passive interface)
  Neighbor Count is 0, Adjacent neighbor count is 0

Loopback 45 is up, line protocol is up
  Internet Address 10.1.1.23/24, Area 2.2.2.2
  Process ID 34, Router ID 10.1.2.100, Network Type LOOPBACK, Cost: 1
```

Enabling Fast-Convergence

The fast-convergence CLI sets the minimum origination and arrival LSA parameters to zero (0), allowing rapid route calculation.

When you disable fast-convergence, origination and arrival LSA parameters are set to 5 seconds and 1 second, respectively.

Setting the convergence parameter (from 1 to 4) indicates the actual convergence level. Each convergence setting adjusts the LSA parameters to zero, but the `fast-convergence` parameter setting allows for even finer tuning of the convergence speed. The higher the number, the faster the convergence.

To enable or disable fast-convergence, use the following command.

- Enable OSPF fast-convergence and specify the convergence level.

  CONFIG-ROUTEROSPF- id mode

  `fast-convergence {number}`

  The parameter range is from 1 to 4.

  The higher the number, the faster the convergence.

  When disabled, the parameter is set at 0.

  NOTE: A higher convergence level can result in occasional loss of OSPF adjacency. Generally, convergence level 1 meets most convergence requirements. Only select higher convergence levels following consultation with Dell Technical Support.

In the following examples, `Convergence Level` shows the fast-converge parameter setting and `Min LSA origination` shows the LSA parameters (shown in bold).

```
DellEMC(conf-router_ospf-1)#fast-converge 2
DellEMC(conf-router_ospf-1)#ex
DellEMC(conf)#ex
DellEMC#show ip ospf 1
Routing Process ospf 1 with ID 192.168.67.2
Supports only single TOS (TOS0) routes
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Convergence Level 2
Min LSA origination 0 secs, Min LSA arrival 0 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
DellEMC#
```

The following examples shows how to disable fast-convergence.

```
DellEMC#(conf-router_ospf-1)#no fast-converge
DellEMC#(conf-router_ospf-1)#ex
DellEMC#(conf)#ex
DellEMC##show ip ospf 1
Routing Process ospf 1 with ID 192.168.67.2
Supports only single TOS (TOS0) routes
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Convergence Level 0
Min LSA origination 5 secs, Min LSA arrival 1 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
DellEMC#
```

Changing OSPFv2 Parameters on Interfaces

In Dell EMC Networking OS, you can modify the OSPF settings on the interfaces.

Some interface parameter values must be consistent across all interfaces to avoid routing errors. For example, set the same time interval for the hello packets on all routers in the OSPF network to prevent misconfiguration of OSPF neighbors.

To change OSPFv2 parameters on the interfaces, use any or all of the following commands.

- Change the cost associated with OSPF traffic on the interface.

  CONFIG-INTERFACE mode

  `ip ospf cost`

  - *cost*: The range is from 1 to 65535 (the default depends on the interface speed).
- Change the time interval the router waits before declaring a neighbor dead.

  CONFIG-INTERFACE mode

  `ip ospf dead-interval seconds`

  - *seconds*: the range is from 1 to 65535 (the default is **40 seconds**).

  The dead interval must be four times the hello interval.

  The dead interval must be the same on all routers in the OSPF network.
- Change the time interval between hello-packet transmission.

  CONFIG-INTERFACE mode

`ip ospf hello-interval seconds`

- *seconds*: the range is from 1 to 65535 (the default is **10 seconds**).

The hello interval must be the same on all routers in the OSPF network.

- Use the MD5 algorithm to produce a message digest or key, which is sent instead of the key.
  CONFIG-INTERFACE mode
  `ip ospf message-digest-key keyid md5 key`
  - *keyid*: the range is from 1 to 255.
  - *Key*: a character string.

  NOTE: Be sure to write down or otherwise record the key. You cannot learn the key after it is configured. You must be careful when changing this key.

  NOTE: You can configure a maximum of six digest keys on an interface. Of the available six digest keys, the switches select the MD5 key that is common. The remaining MD5 keys are unused.

- Change the priority of the interface, which is used to determine the Designated Router for the OSPF broadcast network.
  CONFIG-INTERFACE mode
  `ip ospf priority number`
  - *number*: the range is from 0 to 255 (the default is **1**).
- Change the retransmission interval between LSAs.
  CONFIG-INTERFACE mode
  `ip ospf retransmit-interval seconds`
  - *seconds*: the range is from 1 to 65535 (the default is **5 seconds**).

  The retransmit interval must be the same on all routers in the OSPF network.
- Change the wait period between link state update packets sent out the interface.
  CONFIG-INTERFACE mode
  `ip ospf transmit-delay seconds`
  - *seconds*: the range is from 1 to 65535 (the default is **1 second**).

  The transmit delay must be the same on all routers in the OSPF network.

To view interface configurations, use the `show config` command in CONFIGURATION INTERFACE mode.

To view interface status in the OSPF process, use the `show ip ospf interface` command in EXEC mode.

The bold lines in the example show the change on the interface. The change is reflected in the OSPF configuration.

```
DellEMC(conf-if)#ip ospf cost 45
DellEMC(conf-if)#show config
!
interface TenGigabitEthernet 1/1
  ip address 10.1.2.100 255.255.255.0
  no shutdown
  ip ospf cost 45
DellEMC(conf-if)#end

DellEMC#show ip ospf 34 interface
  TenGigabitEthernet 1/1 is up, line protocol is up
  Internet Address 10.1.2.100/24, Area 2.2.2.2
  Process ID 34, Router ID 10.1.2.100, Network Type BROADCAST, Cost: 45
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 10.1.2.100, Interface address 10.1.2.100
  Backup Designated Router (ID) 10.1.2.100, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:06
  Neighbor Count is 0, Adjacent neighbor count is 0
DellEMC#
```

Enabling OSPFv2 Authentication

To enable or change various OSPF authentication parameters, use the following commands.

- Set a clear text authentication scheme on the interface.

  CONFIG-INTERFACE mode

  ```
  ip ospf authentication-key key
  ```

  Configure a *key* that is a text string no longer than eight characters.

  All neighboring routers must share password to exchange OSPF information.
- Set the authentication change wait time in seconds between 0 and 300 for the interface.

  CONFIG-INTERFACE mode

  ```
  ip ospf auth-change-wait-time seconds
  ```

  This setting is the amount of time OSPF has available to change its interface authentication type.

  When you configure the `auth-change-wait-time`, OSPF sends out only the old authentication scheme until the wait timer expires. After the wait timer expires, OSPF sends only the new authentication scheme. However, the new authentication scheme does not take effect immediately after the authentication change wait timer expires; OSPF accepts both the old as well as new authentication schemes for a time period that is equal to two times the configured authentication change wait timer. After this time period, OSPF accepts only the new authentication scheme.

  This transmission stops when the period ends.

  The default is **0 seconds**.

Enabling OSPFv2 Graceful Restart

Graceful restart is enabled for the global OSPF process.

The Dell EMC Networking implementation of OSPFv2 graceful restart enables you to specify:

- **grace period** — the length of time the graceful restart process can last before OSPF terminates it.
- **helper-reject neighbors** — the router ID of each restart router that does not receive assistance from the configured router.
- **mode** — the situation or situations that trigger a graceful restart.
- **role** — the role or roles the configured router can perform.

NOTE: By default, OSPFv2 graceful restart is disabled.

To enable and configure OSPFv2 graceful restart, use the following commands.

1. Enable OSPFv2 graceful-restart globally and set the grace period.

   CONFIG-ROUTEROSPF- id mode

   ```
   graceful-restart grace-period seconds
   ```

   The seconds range is from 40 and 3000.

   This setting is the time that an OSPFv2 router's neighbors advertises it as fully adjacent, regardless of the synchronization state, during a graceful restart. OSPFv2 terminates this process when the grace period ends.
2. Enter the Router ID of the OSPFv2 helper router from which the router does not accept graceful restart assistance.

   CONFIG-ROUTEROSPF- id mode

   ```
   graceful-restart helper-reject router-id
   ```

   - **Planned-only** — the OSPFv2 router supports graceful-restart for planned restarts only. A planned restart is when you manually enter a fail-over command to force the primary RPM over to the secondary RPM. During a planned restart, OSPF sends out a Grace LSA before the system switches over to the secondary RPM. OSPF also is notified that a planned restart is happening.
   - **Unplanned-only** — the OSPFv2 router supports graceful-restart for only unplanned restarts. During an unplanned restart, OSPF sends out a Grace LSA after the secondary RPM comes online.

   By default, OSPFv2 supports both planned and unplanned restarts. Selecting one or the other mode restricts OSPFv2 to the single selected mode.
3. Configure the graceful restart role or roles that this OSPFv2 router performs.

   CONFIG-ROUTEROSPF- id mode

   ```
   graceful-restart role [helper-only | restart-only]
   ```

Dell EMC Networking OS supports the following options:

- `Helper-only`: the OSPFv2 router supports graceful-restart only as a helper router.
- `Restart-only`: the OSPFv2 router supports graceful-restart only during unplanned restarts.

By default, OSPFv2 supports both restarting and helper roles. Selecting one or the other role restricts OSPFv2 to the single selected role.

To disable OSPFv2 graceful-restart after you have enabled it, use the `no graceful-restart grace-period` command in CONFIG-ROUTEROSPF- id mode. The command returns OSPF graceful-restart to its default state.

NOTE: The Helper mode is enabled by default on the device. To enable the restart mode also on the device, you must configure the grace period using the `graceful-restart grace-period` command. After you enable restart mode the router advertises the neighbor as fully adjacent during a restart.

For more information about OSPF graceful restart, refer to the *Dell EMC Networking OS Command Line Reference Guide*.

When you configure a graceful restart on an OSPFv2 router, the `show run ospf` command displays information similar to the following.

```
DellEMC#show run ospf
!
router ospf 1
  graceful-restart grace-period 300
  graceful-restart role helper-only
  graceful-restart mode unplanned-only
  graceful-restart helper-reject 10.1.1.1
  graceful-restart helper-reject 20.1.1.1
  network 10.0.2.0/24 area 0
DellEMC#
```

Creating Filter Routes

To filter routes, use prefix lists. OSPF applies prefix lists to incoming or outgoing routes.

Incoming routes must meet the conditions of the prefix lists. If they do not, OSPF does not add the route to the routing table. Configure the prefix list in CONFIGURATION PREFIX LIST mode prior to assigning it to the OSPF process.

- Create a prefix list and assign it a unique name.

  CONFIGURATION mode

  `ip prefix-list prefix-name`

  You are in PREFIX LIST mode.
- Create a prefix list with a sequence number and a deny or permit action.

  CONFIG- PREFIX LIST mode

  `seq sequence-number {deny |permit} ip-prefix [ge min-prefix-length] [le max-prefix-length]`

  The optional parameters are:
  - `ge min-prefix-length`: is the minimum prefix length to match (from 0 to 32).
  - `le max-prefix-length`: is the maximum prefix length to match (from 0 to 32).

For configuration information about prefix lists, refer to Access Control Lists (ACLs).

Applying Prefix Lists

To apply prefix lists to incoming or outgoing OSPF routes, use the following commands.

- Apply a configured prefix list to incoming OSPF routes.

  CONFIG-ROUTEROSPF-id mode

  `distribute-list prefix-list-name in [interface]`
- Assign a configured prefix list to outgoing OSPF routes.

  CONFIG-ROUTEROSPF-id

  `distribute-list prefix-list-name out [connected | isis | rip | static]`

Redistributing Routes

You can add routes from other routing instances or protocols to the OSPF process.

With the `redistribute` command, you can include RIP, static, or directly connected routes in the OSPF process.

NOTE: Do not route iBGP routes to OSPF unless there are route-maps associated with the OSPF redistribution.

To redistribute routes, use the following command.

- Specify which routes are redistributed into OSPF process.

  CONFIG-ROUTEROSPF-id mode

  `redistribute {bgp | connected | isis | rip | static} [metric` *`metric-value`* `| metric-type` *`type-value`*`] [route-map` *`map-name`*`] [tag` *`tag-value`*`]`

  Configure the following required and optional parameters:
  - `bgp, connected, isis, rip, static`: enter one of the keywords to redistribute those routes.
  - `metric` *`metric-value`*: the range is from 0 to 4294967295.
  - `metric-type` *`metric-type`*: 1 for OSPF external route type 1. 2 for OSPF external route type 2.
  - `route-map` *`map-name`*: enter a name of a configured route map.
  - `tag` *`tag-value`*: the range is from 0 to 4294967295.

To view the current OSPF configuration, use the `show running-config ospf` command in EXEC mode or the `show config` command in ROUTER OSPF mode.

```
DellEMC(conf-router_ospf)#show config
!
router ospf 34
  network 10.1.2.32 0.0.0.255 area 2.2.2.2
  network 10.1.3.24 0.0.0.255 area 3.3.3.3
  distribute-list dilling in
DellEMC(conf-router_ospf)#
```

Troubleshooting OSPFv2

Use the information in this section to troubleshoot OSPFv2 operation on the switch.

Be sure to check the following, as these questions represent typical issues that interrupt an OSPFv2 process.

NOTE: The following tasks are not a comprehensive; they provide some examples of typical troubleshooting checks.

- Have you enabled OSPF globally?
- Is the OSPF process active on the interface?
- Are adjacencies established correctly?
- Are the interfaces configured for Layer 3 correctly?
- Is the router in the correct area type?
- Have the routes been included in the OSPF database?
- Have the OSPF routes been included in the routing table (not just the OSPF database)?

Some useful troubleshooting commands are:

- `show interfaces`
- `show protocols`
- `debug IP OSPF events and/or packets`
- `show neighbors`
- `show routes`

To help troubleshoot OSPFv2, use the following commands.

- View the summary of all OSPF process IDs enables on the router.

  EXEC Privilege mode

  `show running-config ospf`
- View the summary information of the IP routes.

  EXEC Privilege mode

`show ip route summary`

- View the summary information for the OSPF database.

  EXEC Privilege mode

  `show ip ospf database`
- View the configuration of OSPF neighbors connected to the local router.

  EXEC Privilege mode

  `show ip ospf neighbor`
- View the LSAs currently in the queue.

  EXEC Privilege mode

  `show ip ospf timers rate-limit`
- View debug messages.

  EXEC Privilege mode

  `debug ip ospf` *`process-id`* `[event | packet | spf | database-timers rate-limit]`

  To view debug messages for a specific OSPF process ID, use the `debug ip ospf` *`process-id`* command.

  If you do not enter a process ID, the command applies to the first OSPF process.

  To view debug messages for a specific operation, enter one of the optional keywords:
  - `event`: view OSPF event messages.
  - `packet`: view OSPF packet information.
  - `spf`: view SPF information.
  - `database-timers rate-limit`: view the LSAs currently in the queue.

```
DellEMC#show run ospf
!
router ospf 4
  router-id 4.4.4.4
  network 4.4.4.0/28 area 1
!
ipv6 router ospf 999
  default-information originate always
  router-id 10.10.10.10
DellEMC#
```

Sample Configurations for OSPFv2

The following configurations are examples for enabling OSPFv2.

These examples are not comprehensive directions. They are intended to give you some guidance with typical configurations.

You can copy and paste from these examples to your CLI. To support your own IP addresses, interfaces, names, and so on, be sure that you make the necessary changes.

Basic OSPFv2 Router Topology

The following illustration is a sample basic OSPFv2 topology.

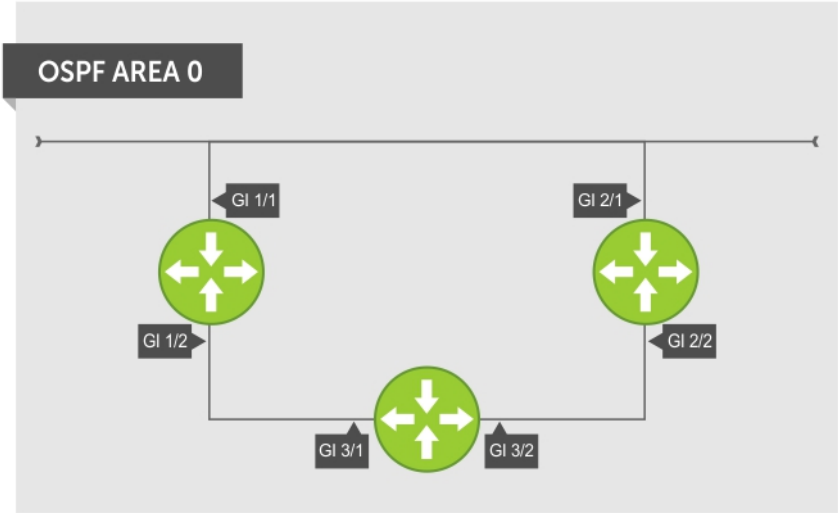


Figure 103. Basic Topology and CLI Commands for OSPFv2

OSPF Area 0 — Te 1/1 and 1/2

```
router ospf 11111
  network 10.0.11.0/24 area 0
  network 10.0.12.0/24 area 0
  network 192.168.100.0/24 area 0
!
interface TenGigabitEthernet 1/1
  ip address 10.1.11.1/24
  no shutdown
!
interface TenGigabitEthernet 1/2
  ip address 10.2.12.2/24
  no shutdown
!
interface Loopback 10
  ip address 192.168.100.100/24
  no shutdown
```

OSPF Area 0 — Te 3/1 and 3/2

```
router ospf 33333
  network 192.168.100.0/24 area 0
  network 10.0.13.0/24 area 0
  network 10.0.23.0/24 area 0
!
interface Loopback 30
  ip address 192.168.100.100/24
  no shutdown
!
interface TenGigabitEthernet 3/1
  ip address 10.1.13.3/24
  no shutdown
!
interface TenGigabitEthernet 3/2
  ip address 10.2.13.3/24
  no shutdown
```

OSPF Area 0 — Te 2/1 and 2/2

```
router ospf 22222
  network 192.168.100.0/24 area 0
  network 10.2.21.0/24 area 0
  network 10.2.22.0/24 area 0
!
interface Loopback 20
  ip address 192.168.100.20/24
  no shutdown
!
interface TenGigabitEthernet 2/1
  ip address 10.2.21.2/24
  no shutdown
!
interface TenGigabitEthernet 2/2
  ip address 10.2.22.2/24
  no shutdown
```

OSPFv3 NSSA

NSSA (Not-So-Stubby-Area) is a stub area that does not support Type-5 LSAs, but supports Type-7 LSAs to forward external links. Initially ASBR (Autonomous System Border Router) forwards the external links through Type-7 LSAs to the Area Border Router (ABR) of NSSA, which in turn converts them into Type-5 LSAs and forwards them to the rest of the OSPF domain.

NOTE: To support NSSA area, all the OSPF routers in that area should be configured with NSSA.

NSSA Options

NSSA can be configured with the following options:

1. Default-information-originate – To inject a default route using Type-7 LSAs — NSSA routers need to have access to the rest of the OSPF routers in the autonomous system. To facilitate this, the default route is injected into the NSSA area through a Type-7 LSA. This can be generated either by NSSA ASBR or NSSA ABR.
2. No-redistribute – To restrict Type-7 LSAs — When NSSA ASBR is also an ABR, redistributed external routes need not be translated from Type-7 to Type-5 LSAs. ABR will directly inject external routes through Type-5 LSAs into the OSPF domain. It does not send Type-7 LSAs into the NSSA area.
3. No-summary – To act as totally stubby area — NSSA area can be converted intoa totally stubby area to reduce the number of Type-3 LSAs. Once it is configured, NSSA ABR will inject Type-3 LSAs into the NSSA area for default routes. The remaining Type-3 LSAs are not allowed inside this area.

Configuration Task List for OSPFv3 (OSPF for IPv6)

This section describes the configuration tasks for Open Shortest Path First version 3 (OSPF for IPv6) on the switch.

The configuration options of OSPFv3 are the same as those options for OSPFv2, but you may configure OSPFv3 with differently labeled commands. Specify process IDs and areas and include interfaces and addresses in the process. Define areas as stub or totally stubby.

The interfaces must be in IPv6 Layer-3 mode (assigned an IPv6 IP address) and enabled so that they can send and receive traffic. The OSPF process must know about these interfaces. To make the OSPF process aware of these interfaces, assign them to OSPF areas.

The OSPFv3 `ipv6 ospf area` command enables OSPFv3 on the interface and places the interface in an area. With OSPFv2, two commands are required to accomplish the same tasks — the `router ospf` command to create the OSPF process, then the `network area` command to enable OSPF on an interface.

NOTE: The OSPFv2 `network area` command enables OSPF on multiple interfaces with the single command. Use the `OSPFv3 ipv6 ospf area` command on each interface that runs OSPFv3.

All IPv6 addresses on an interface are included in the OSPFv3 process that is created on the interface.

Enable OSPFv3 for IPv6 by specifying an OSPF process ID and an area in INTERFACE mode. If you have not created an OSPFv3 process, it is created automatically. All IPv6 addresses configured on the interface are included in the specified OSPF process.

NOTE: IPv6 and OSPFv3 do not support Multi-Process OSPF. You can only enable a single OSPFv3 process. To create multiple OSPF processes you need to have multiple VRFs on a switch.

Set the time interval between when the switch receives a topology change and starts a shortest path first (SPF) calculation.

`timers spf delay holdtime`

NOTE: To set the interval time between the reception of topology changes and calculation of SPF in milli seconds, use the `timers spf delay holdtime msec` command.

Example

```
DellEMC#conf
DellEMC(conf)#ipv6 router ospf 1
DellEMC(conf-ipv6-router_ospf)#timer spf 2 5 msec
DellEMC(conf-ipv6-router_ospf)#
DellEMC(conf-ipv6-router_ospf)#show config
!
ipv6 router ospf 1
timers spf 2 5 msec
DellEMC(conf-ipv6-router_ospf)#
DellEMC(conf-ipv6-router_ospf)#end
DellEMC#
```

Enabling IPv6 Unicast Routing

To enable IPv6 unicast routing, use the following command.

- Enable IPv6 unicast routing globally.

  CONFIGURATION mode

  `ipv6 unicast routing`

Applying cost for OSPFv3

Change in bandwidth directly affects the cost of OSPF routes.

- Explicitly specify the cost of sending a packet on an interface.

  INTERFACE mode

  `ipv6 ospf interface-cost`

  - *interface-cost*:The range is from 1 to 65535. Default cost is based on the bandwidth.
- Specify how the OSPF interface cost is calculated based on the reference bandwidth method. The cost of an interface is calculated as Reference Bandwidth/Interface speed.

  ROUTER OSPFv3

  `auto-cost [reference-bandwidth ref-bw]`

  To return to the default bandwidth or to assign cost based on the interface type, use the `no auto-cost [reference-bandwidth ref-bw]` command.

  - `ref-bw`: The range is from 1 to 4294967. The default is 100 megabits per second.

Assigning IPv6 Addresses on an Interface

To assign IPv6 addresses to an interface, use the following commands.

1. Assign an IPv6 address to the interface.

   CONF-INT-type slot/port mode

   `ipv6 address ipv6 address`

   IPv6 addresses are normally written as eight groups of four hexadecimal digits; separate each group by a colon (:).

   The format is A:B:C::F/128.

2. Bring up the interface.
   CONF-INT-type slot/port mode
   `no shutdown`

Assigning Area ID on an Interface

To assign the OSPFv3 process to an interface, use the following command.

The `ipv6 ospf area` command enables OSPFv3 on an interface and places the interface in the specified area. Additionally, the command creates the OSPFv3 process with ID on the router. OSPFv2 requires two commands to accomplish the same tasks — the `router ospf` command to create the OSPF process, then the `network area` command to enable OSPFv2 on an interface.

NOTE: The OSPFv2 network area command enables OSPFv2 on multiple interfaces with the single command. Use the OSPFv3 `ipv6 ospf area` command on each interface that runs OSPFv3.

- Assign the OSPFv3 process and an OSPFv3 area to this interface.
  CONF-INT-type slot/port mode
  `ipv6 ospf process-id area area-id`
  - *process-id*: the process ID number assigned.
  - *area-id*: the area ID for this interface.

Assigning OSPFv3 Process ID and Router ID Globally

To assign, disable, or reset OSPFv3 globally, use the following commands.

- Enable the OSPFv3 process globally and enter OSPFv3 mode.
  CONFIGURATION mode
  `ipv6 router ospf {process ID}`
  The range is from 0 to 65535.
- Assign the router ID for this OSPFv3 process.
  CONF-IPV6-ROUTER-OSPF mode
  `router-id {number}`
  - *number*: the IPv4 address.

  The format is A.B.C.D.

  NOTE: Enter the router-id for an OSPFv3 router as an IPv4 IP address.

- Disable OSPF.
  CONFIGURATION mode
  `no ipv6 router ospf process-id`
- Reset the OSPFv3 process.
  EXEC Privilege mode
  `clear ipv6 ospf process`

Assigning OSPFv3 Process ID and Router ID to a VRF

To assign, disable, or reset OSPFv3 on a non-default VRF, use the following commands.

- Enable the OSPFv3 process on a non-default VRF and enter OSPFv3 mode.
  CONFIGURATION mode
  `ipv6 router ospf {process ID} vrf {vrf-name}`
  The process ID range is from 0 to 65535.
- Assign the router ID for this OSPFv3 process.
  CONF-IPV6-ROUTER-OSPF mode

`router-id {number}`

- *number*: the IPv4 address.

The format is A.B.C.D.

NOTE: Enter the router-id for an OSPFv3 router as an IPv4 IP address.

- Disable OSPF.
  CONFIGURATION mode
  `no ipv6 router ospf process-id vrf {vrf-name}`
- Reset the OSPFv3 process.
  EXEC Privilege mode
  `clear ipv6 ospf [vrf vrf-name] process`

Configuring Stub Areas

To configure IPv6 stub areas, use the following command.

- Configure the area as a stub area.
  CONF-IPV6-ROUTER-OSPF mode
  `area area-id stub [no-summary]`
  - *no-summary*: use these keywords to prevent transmission in to the area of summary ASBR LSAs.
  - *Area ID*: a number or IP address assigned when creating the area. You can represent the area ID as a number from 0 to 65536 if you assign a dotted decimal format rather than an IP address.

Configuring Passive-Interface

To suppress the interface's participation on an OSPFv3 interface, use the following command.

This command stops the router from sending updates on that interface.

- Specify whether some or all some of the interfaces are passive.
  CONF-IPV6-ROUTER-OSPF mode
  `passive-interface {interface-type}`
  `Interface`: identifies the specific interface that is passive.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

To enable both receiving and sending routing updates, use the `no passive-interface interface` command.

To indicate that hello packets are not transmitted on that interface, when you configure a passive interface, the `show ipv6 ospf [vrf vrf-name] interface` command adds the words `passive interface`.

Redistributing Routes

You can add routes from other routing instances or protocols to the OSPFv3 process.

With the `redistribute` command, you can include RIP, static, or directly connected routes in the OSPF process. Route redistribution is also supported between OSPF Routing process IDs.

To add redistributing routes, use the following command.

- Specify which routes are redistributed into the OSPF process.
  CONF-IPV6-ROUTER-OSPF mode
  `redistribute {bgp | connected | static} [metric metric-value | metric-type type-value] [route-map map-name] [tag tag-value]`
  Configure the following required and optional parameters:

- `bgp | connected | static`: enter one of the keywords to redistribute those routes.
- `metric metric-value`: The range is from 0 to 4294967295.
- `metric-type metric-type`: enter `1` for OSPFv3 external route type 1 OR `2` for OSPFv3 external route type 2.
- `route-map map-name`: enter a name of a configured route map.
- `tag tag-value`: The range is from 0 to 4294967295.

Configuring a Default Route

To generate a default external route into the OSPFv3 routing domain, configure the following parameters.

To specify the information for the default route, use the following command.

- Specify the information for the default route.

  CONF-IPV6-ROUTER-OSPF mode

  `default-information originate [always [metric metric-value] [metric-type type-value]] [route-map map-name]`

  Configure the following required and optional parameters:
  - `always`: indicate that default route information is always advertised.
  - `metric metric-value`: The range is from 0 to 4294967295.
  - `metric-type metric-type`: enter `1` for OSPFv3 external route type 1 OR `2` for OSPFv3 external route type 2.
  - `route-map map-name`: enter a name of a configured route map.

Enabling OSPFv3 Graceful Restart

Follow the procedure in this section to configure graceful restart for OSPFv3. By default, OSPFv3 graceful restart is disabled and functions only in a helper role to help restarting neighbor routers in their graceful restarts when it receives a Grace LSA.

.

By default, OSPFv3 graceful restart is disabled and functions only in a helper role to help restarting neighbor routers in their graceful restarts when it receives a Grace LSA.

To enable OSPFv3 graceful restart, enter the `ipv6 router ospf process-id` command to enter OSPFv3 configuration mode. Then configure a grace period using the `graceful-restart grace-period` command. The grace period is the time that the OSPFv3 neighbors continue to advertise the restarting router as though it is fully adjacent. When you enable graceful restart (restarting role), an OSPFv3 restarting expects its OSPFv3 neighbors to help when it restarts by not advertising the broken link.

When you enable the helper-reject role on an interface using the `ipv6 ospf graceful-restart helper-reject` command, you reconfigure OSPFv3 graceful restart to function in a *restarting-only* role. OSPFv3 does not participate in the graceful restart of a neighbor.

NOTE: Enter the `ipv6 ospf graceful-restart helper-reject` command in Interface configuration mode.

NOTE: For graceful-restart configuration to work, you must configure grace-period. Use `graceful-restart grace-period` command to configure grace-period.

- Enable OSPFv3 graceful restart globally by setting the grace period (in seconds).

  CONF-IPV6-ROUTER-OSPF mode

  `graceful-restart grace-period seconds`

  The valid values are from 40 to 1800 seconds.
- Configure an OSPFv3 interface to not act on the Grace LSAs that it receives from a restarting OSPFv3 neighbor.

  INTERFACE mode

  `ipv6 ospf graceful-restart helper-reject`
- Specify the operating mode and type of events that trigger a graceful restart.

  CONF-IPV6-ROUTER-OSPF mode

  `graceful-restart mode [planned-only | unplanned-only]`
  - `Planned-only`: the OSPFv3 router supports graceful restart only for planned restarts. A planned restart is when you manually enter a `redundancy force-failover rpm` command to force the primary RPM over to the secondary

RPM. During a planned restart, OSPFv3 sends out a Grace LSA before the system switches over to the secondary RPM. OSPFv3 is notified that a planned restart is happening.

- `Unplanned-only`: the OSPFv3 router supports graceful-restart only for unplanned restarts. During an unplanned restart, OSPFv3 sends out a Grace LSA once the secondary RPM comes online.

The default is both planned and unplanned restarts trigger an OSPFv3 graceful restart. Selecting one or the other mode restricts OSPFv3 to the single selected mode.

- Disable OSPFv3 graceful-restart.

  CONF-IPV6-ROUTER-OSPF mode

  ```
  no graceful-restart grace-period
  ```

Displaying Graceful Restart

To display information on the use and configuration of OSPFv3 graceful restart, enter any of the following commands.

- Display the graceful-restart configuration for OSPFv2 and OSPFv3 (shown in the following example).

  EXEC Privilege mode

  ```
  show run ospf
  ```

- Display the Type-11 Grace LSAs sent and received on an OSPFv3 router (shown in the following example).

  EXEC Privilege mode

  ```
  show ipv6 ospf [vrf vrf-name] database grace-lsa
  ```

- Display the currently configured OSPFv3 parameters for graceful restart (shown in the following example).

  EXEC Privilege mode

  ```
  show ipv6 ospf database [vrf vrf-name] database-summary
  ```

The following example shows the `show run ospf` command.

```
DellEMC#show run ospf
!
router ospf 1
  router-id 200.1.1.1
  log-adjacency-changes
  graceful-restart grace-period 180
  network 20.1.1.0/24 area 0
  network 30.1.1.0/24 area 0
!
ipv6 router ospf 1
  log-adjacency-changes
  graceful-restart grace-period 180
```

The following example shows the `show ipv6 ospf database database-summary` command.

```
DellEMC#show ipv6 ospf database database-summary
!
OSPFv3 Router with ID (200.1.1.1) (Process ID 1)

Process 1 database summary
Type                    Count/Status
Oper Status             1
Admin Status            1
Area Bdr Rtr Status     0
AS Bdr Rtr Status       1
AS Scope LSA Count      0
AS Scope LSA Cksum sum  0
Originate New LSAS      73
Rx New LSAS             114085
Ext LSA Count           0
Rte Max Eq Cost Paths   5
GR grace-period      180
GR mode              planned and unplanned

Area 0 database summary
Type                    Count/Status
Brd Rtr Count           2
AS Bdr Rtr Count        2
```

```
LSA count                  12010
Summary LSAs               1
Rtr LSA Count              4
Net LSA Count              3
Inter Area Pfx LSA Count 12000
Inter Area Rtr LSA Count 0
Group Mem LSA Count        0
```

The following example shows the `show ipv6 ospf database grace-lsa` command.

```
DellEMC#show ipv6 ospf database grace-lsa
!
Type-11 Grace LSA (Area 0)

LS Age                : 10
Link State ID         : 6.16.192.66
Advertising Router    : 100.1.1.1
LS Seq Number         : 0x80000001
Checksum              : 0x1DF1
Length                : 36
Associated Interface : Te 5/3
Restart Interval      : 180
Restart Reason        : Switch to Redundant Processor
```

OSPFv3 Authentication Using IPsec

OSPFv3 uses IPsec to provide authentication for OSPFv3 packets. IPsec authentication ensures security in the transmission of OSPFv3 packets between IPsec-enabled routers.

IPsec is a set of protocols developed by the internet engineering task force (IETF) to support secure exchange of packets at the IP layer. IPsec supports two encryption modes: transport and tunnel.

- **Transport mode** — encrypts only the data portion (payload) of each packet, but leaves the header untouched.
- **Tunnel mode** — is more secure and encrypts both the header and payload. On the receiving side, an IPsec-compliant device decrypts each packet.

NOTE: Dell EMC Networking OS supports only Transport Encryption mode in OSPFv3 authentication with IPsec.

With IPsec-based authentication, Crypto images are used to include the IPsec secure socket application programming interface (API) required for use with OSPFv3.

To ensure integrity, data origin authentication, detection and rejection of replays, and confidentiality of the packet, RFC 4302 and RFC 4303 propose using two security protocols — authentication header (AH) and encapsulating security payload (ESP). For OSPFv3, these two IPsec protocols provide interoperable, high-quality cryptographically-based security.

- **HA** — IPsec authentication header is used in packet authentication to verify that data is not altered during transmission and ensures that users are communicating with the intended individual or organization. Insert the authentication header after the IP header with a value of 51. AH provides integrity and validation of data origin by authenticating every OSPFv3 packet. For detailed information about the IP AH protocol, refer to *RFC 4302*.
- **ESP** — encapsulating security payload encapsulates data, enabling the protection of data that follows in the datagram. ESP provides authentication and confidentiality of every packet. The ESP extension header is designed to provide a combination of security services for both IPv4 and IPv6. Insert the ESP header after the IP header and before the next layer protocol header in Transport mode. It is possible to insert the ESP header between the next layer protocol header and encapsulated IP header in Tunnel mode. However, Tunnel mode is not supported in Dell EMC Networking OS. For detailed information about the IP ESP protocol, refer to *RFC 4303*.

In OSPFv3 communication, IPsec provides security services between a pair of communicating hosts or security gateways using either AH or ESP. In an authentication policy on an interface or in an OSPF area, AH and ESP are used alone; in an encryption policy, AH and ESP may be used together. The difference between the two mechanisms is the extent of the coverage. ESP only protects IP header fields if they are encapsulated by ESP.

You decide the set of IPsec protocols that are employed for authentication and encryption and the ways in which they are employed. When you correctly implement and deploy IPsec, it does not adversely affect users or hosts. AH and ESP are designed to be cryptographic algorithm-independent.

OSPFv3 Authentication Using IPsec: Configuration Notes

OSPFv3 authentication using IPsec is implemented according to the specifications in RFC 4552.

- To use IPsec, configure an authentication (using AH) or encryption (using ESP) security policy on an interface or in an OSPFv3 area. Each security policy consists of a security policy index (SPI) and the key used to validate OSPFv3 packets. After IPsec is configured for OSPFv3, IPsec operation is invisible to the user.
  - You can only enable one security protocol (AH or ESP) at a time on an interface or for an area. Enable IPsec AH with the `ipv6 ospf authentication` command; enable IPsec ESP with the `ipv6 ospf encryption` command.
  - The security policy configured for an area is inherited by default on all interfaces in the area.
  - The security policy configured on an interface overrides any area-level configured security for the area to which the interface is assigned.
  - The configured authentication or encryption policy is applied to all OSPFv3 packets transmitted on the interface or in the area. The IPsec security associations (SAs) are the same on inbound and outbound traffic on an OSPFv3 interface.
  - There is no maximum AH or ESP header length because the headers have fields with variable lengths.
- Manual key configuration is supported in an authentication or encryption policy (dynamic key configuration using the internet key exchange [IKE] protocol is not supported).
- In an OSPFv3 authentication policy:
  - AH is used to authenticate OSPFv3 headers and certain fields in IPv6 headers and extension headers.
  - MD5 and SHA1 authentication types are supported; encrypted and unencrypted keys are supported.
- In an OSPFv3 encryption policy:
  - Both encryption and authentication are used.
  - IPsec security associations (SAs) are supported only in Transport mode (Tunnel mode is not supported).
  - ESP with null encryption is supported for authenticating only OSPFv3 protocol headers.
  - ESP with non-null encryption is supported for full confidentiality.
  - 3DES, DES, AES-CBC, and NULL encryption algorithms are supported; encrypted and unencrypted keys are supported.

NOTE: To encrypt all keys on a router, use the `service password-encryption` command in Global Configuration mode. However, this command does not provide a high level of network security. To enable key encryption in an IPsec security policy at an interface or area level, specify `7` for *`[key-encryption-type]`* when you enter the `ipv6 ospf authentication ipsec` or `ipv6 ospf encryption ipsec` command.

- To configure an IPsec security policy for authenticating or encrypting OSPFv3 packets on a physical, port-channel, or VLAN interface or OSPFv3 area, perform any of the following tasks:
  - Configuring IPsec Authentication on an Interface
  - Configuring IPsec Encryption on an Interface
  - Configuring IPsec Authentication for an OSPFv3 Area
  - Configuring IPsec Encryption for an OSPFv3 Area
  - Displaying OSPFv3 IPsec Security Policies

Configuring IPsec Authentication on an Interface

To configure, remove, or display IPsec authentication on an interface, use the following commands.

Prerequisite: Before you enable IPsec authentication on an OSPFv3 interface, first enable IPv6 unicast routing globally, configure an IPv6 address and enable OSPFv3 on the interface, and assign it to an area (refer to Configuration Task List for OSPFv3 (OSPF for IPv6)).

The SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. Configure the same authentication policy (the same SPI and key) on each OSPFv3 interface in a link.

- Enable IPsec authentication for OSPFv3 packets on an IPv6-based interface.

  INTERFACE mode

  `ipv6 ospf authentication {null | ipsec spi` *`number`* `{MD5 | SHA1} [`*`key-encryption-type`*`] `*`key`*`}`

  - `null`: causes an authentication policy configured for the area to not be inherited on the interface.
  - `ipsec spi` *`number`*: the security policy index (SPI) value. The range is from 256 to 4294967295.
  - `MD5 | SHA1`: specifies the authentication type: Message Digest 5 (`MD5`) or Secure Hash Algorithm 1 (`SHA-1`).
  - *`key-encryption-type`*: (optional) specifies if the key is encrypted. The valid values are `0` (key is not encrypted) or `7` (key is encrypted).

- `key`: specifies the text string used in authentication. All neighboring OSPFv3 routers must share key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted).

- Remove an IPsec authentication policy from an interface.

  ```
  no ipv6 ospf authentication ipsec spi number
  ```

- Remove null authentication on an interface to allow the interface to inherit the authentication policy configured for the OSPFv3 area.

  ```
  no ipv6 ospf authentication null
  ```

- Display the configuration of IPsec authentication policies on the router.

  ```
  show crypto ipsec policy
  ```

- Display the security associations set up for OSPFv3 interfaces in authentication policies.

  ```
  show crypto ipsec sa ipv6
  ```

Configuring IPsec Encryption on an Interface

To configure, remove, or display IPsec encryption on an interface, use the following commands.

Prerequisite: Before you enable IPsec encryption on an OSPFv3 interface, first enable IPv6 unicast routing globally, configure an IPv6 address and enable OSPFv3 on the interface, and assign it to an area (refer to Configuration Task List for OSPFv3 (OSPF for IPv6)).

NOTE: When you configure encryption using the `ipv6 ospf encryption ipsec` command, you enable both IPsec encryption and authentication. However, when you enable authentication on an interface using the `ipv6 ospf authentication ipsec` command, you do not enable encryption at the same time.

The SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. Configure the same authentication policy (the same SPI and key) on each OSPFv3 interface in a link.

- Enable IPsec encryption for OSPFv3 packets on an IPv6-based interface.

  INTERFACE mode

  ```
  ipv6 ospf encryption {null | ipsec spi number esp encryption-algorithm [key-encryption-type] key authentication-algorithm [key-authentication-type] key}
  ```

  - `null`: causes an encryption policy configured for the area to not be inherited on the interface.
  - `ipsec spi number`: is the security policy index (SPI) value. The range is from 256 to 4294967295.
  - `esp encryption-algorithm`: specifies the encryption algorithm used with ESP. The valid values are `3DES`, `DES`, `AES-CBC`, and `NULL`. For `AES-CBC`, only the AES-128 and AES-192 ciphers are supported.
  - *`key`*: specifies the text string used in the encryption. All neighboring OSPFv3 routers must share the same key to decrypt information. Required lengths of a non-encrypted or encrypted key are: 3DES - 48 or 96 hex digits; DES - 16 or 32 hex digits; AES-CBC - 32 or 64 hex digits for AES-128 and 48 or 96 hex digits for AES-192.
  - *`key-encryption-type`*: (optional) specifies if the key is encrypted. The valid values are `0` (key is not encrypted) or `7` (key is encrypted).
  - *`authentication-algorithm`*: specifies the encryption authentication algorithm to use. The valid values are `MD5` or `SHA1`.
  - `key`: specifies the text string used in authentication. All neighboring OSPFv3 routers must share key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted).
  - *`key-authentication-type`*: (optional) specifies if the authentication key is encrypted. The valid values are `0` or `7`.

- Remove an IPsec encryption policy from an interface.

  ```
  no ipv6 ospf encryption ipsec spi number
  ```

- Remove null encryption on an interface to allow the interface to inherit the encryption policy configured for the OSPFv3 area.

  ```
  no ipv6 ospf encryption null
  ```

- Display the configuration of IPsec encryption policies on the router.

  ```
  show crypto ipsec policy
  ```

- Display the security associations set up for OSPFv3 interfaces in encryption policies.

  ```
  show crypto ipsec sa ipv6
  ```

Configuring IPSec Authentication for an OSPFv3 Area

To configure, remove, or display IPSec authentication for an OSPFv3 area, use the following commands.

Prerequisite: Before you enable IPsec authentication on an OSPFv3 area, first enable OSPFv3 globally on the router (refer to Configuration Task List for OSPFv3 (OSPF for IPv6)).

The security policy index (SPI) value must be unique to one IPSec security policy (authentication or encryption) on the router. Configure the same authentication policy (the same SPI and key) on each interface in an OPSFv3 link.

If you have enabled IPSec encryption in an OSPFv3 area using the `area encryption` command, you cannot use the `area authentication` command in the area at the same time.

The configuration of IPSec authentication on an interface-level takes precedence over an area-level configuration. If you remove an interface configuration, an area authentication policy that has been configured is applied to the interface.

- Enable IPSec authentication for OSPFv3 packets in an area.

  CONF-IPV6-ROUTER-OSPF mode

  `area-id authentication ipsec spi number {MD5 | SHA1} [key-encryption-type] key`

  - `area area-id`: specifies the area for which OSPFv3 traffic is to be authenticated. For *area-id*, enter a number or an IPv6 prefix.
  - `spi number`: is the SPI value. The range is from 256 to 4294967295.
  - `MD5 | SHA1`: specifies the authentication type: message digest 5 (`MD5`) or Secure Hash Algorithm 1 (`SHA-1`).
  - `key-encryption-type`: (optional) specifies if the key is encrypted. The valid values are `0` (key is not encrypted) or `7` (key is encrypted).
  - `key`: specifies the text string used in authentication. All neighboring OSPFv3 routers must share key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted).
- Remove an IPSec authentication policy from an OSPFv3 area.

  `no area area-id authentication ipsec spi number`
- Display the configuration of IPSec authentication policies on the router.

  `show crypto ipsec policy`

Configuring IPsec Encryption for an OSPFv3 Area

To configure, remove, or display IPsec encryption in an OSPFv3 area, use the following commands.

Prerequisite: Before you enable IPsec encryption in an OSPFv3 area, first enable OSPFv3 globally on the router (refer to Configuration Task List for OSPFv3 (OSPF for IPv6)).

The SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. Configure the same encryption policy (the same SPI and keys) on each interface in an OPSFv3 link.

NOTE: When you configure encryption using the `area encryption` command, you enable both IPsec encryption and authentication. However, when you enable authentication on an area using the `area authentication` command, you do not enable encryption at the same time.

If you have enabled IPsec authentication in an OSPFv3 area using the `area authentication` command, you cannot use the `area encryption` command in the area at the same time.

The configuration of IPsec encryption on an interface-level takes precedence over an area-level configuration. If you remove an interface configuration, an area encryption policy that has been configured is applied to the interface.

- Enable IPsec encryption for OSPFv3 packets in an area.

  CONF-IPV6-ROUTER-OSPF mode

  `area area-id encryption ipsec spi number esp encryption-algorithm [key-encryption-type] key authentication-algorithm [key-authentication-type] key`

  - `area area-id`: specifies the area for which OSPFv3 traffic is to be encrypted. For *area-id*, enter a number or an IPv6 prefix.
  - `spi number`: is the security policy index (SPI) value. The range is from 256 to 4294967295.
  - `esp encryption-algorithm`: specifies the encryption algorithm used with ESP. The valid values are `3DES`, `DES`, `AES-CBC`, and `NULL`. For `AES-CBC`, only the AES-128 and AES-192 ciphers are supported.

- `key`: specifies the text string used in the encryption. All neighboring OSPFv3 routers must share the same key to decrypt information. The required lengths of a non-encrypted or encrypted key are: 3DES - 48 or 96 hex digits; DES - 16 or 32 hex digits; AES-CBC - 32 or 64 hex digits for AES-128 and 48 or 96 hex digits for AES-192.
- *`key-encryption-type`*: (optional) specifies if the key is encrypted. Valid values: 0 (key is not encrypted) or 7 (key is encrypted).
- *`authentication-algorithm`*: specifies the authentication algorithm to use for encryption. The valid values are `MD5` or `SHA1`.
- `key`: specifies the text string used in authentication. All neighboring OSPFv3 routers must share key to exchange information. For `MD5` authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For `SHA-1` authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted).
- *`key-authentication-type`*: (optional) specifies if the authentication key is encrypted. The valid values are `0` or `7`.

- Remove an IPsec encryption policy from an OSPFv3 area.
  `no area area-id encryption ipsec spi number`
- Display the configuration of IPsec encryption policies on the router.
  `show crypto ipsec policy`

Displaying OSPFv3 IPsec Security Policies

To display the configuration of IPsec authentication and encryption policies, use the following commands.

- Display the AH and ESP parameters configured in IPsec security policies, including the SPI number, key, and algorithms used.
  EXEC Privilege mode
  `show crypto ipsec policy [name name]`
  - `name`: displays configuration details about a specified policy.
- Display security associations set up for OSPFv3 links in IPsec authentication and encryption policies on the router.
  EXEC Privilege
  `show crypto ipsec sa ipv6 [interface interface]`
  To display information on the SAs used on a specific interface, enter `interface interface`, where interface is one of the following values:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

In the first example, the keys are not encrypted (shown in bold). In the second and third examples, the keys are encrypted (shown in bold).

The following example shows the `show crypto ipsec policy` command.

```
DellEMC#show crypto ipsec policy

Crypto IPSec client security policy data

Policy name              : OSPFv3-1-502
Policy refcount          : 1
Inbound ESP SPI          : 502 (0x1F6)
Outbound ESP SPI         : 502 (0x1F6)
Inbound ESP Auth Key     : 123456789a123456789b123456789c12
Outbound ESP Auth Key    : 123456789a123456789b123456789c12
Inbound ESP Cipher Key   : 123456789a123456789b123456789c123456789d12345678
Outbound ESP Cipher Key  : 123456789a123456789b123456789c123456789d12345678
Transform set            : esp-3des esp-md5-hmac

Crypto IPSec client security policy data

Policy name              : OSPFv3-1-500
Policy refcount          : 2
Inbound AH SPI           : 500 (0x1F4)
Outbound AH SPI          : 500 (0x1F4)
Inbound AH Key           : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97e
Outbound AH Key          : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97e
```

```
Transform set           : ah-md5-hmac

Crypto IPSec client security policy data

Policy name             : OSPFv3-0-501
Policy refcount         : 1
Inbound ESP SPI         : 501 (0x1F5)
Outbound ESP SPI        : 501 (0x1F5)
Inbound ESP Auth Key    :
bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97eb7c0c30808825fb5
Outbound ESP Auth Key   :
bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97eb7c0c30808825fb5
Inbound ESP Cipher Key  :
bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba10345a1039ba8f8a
Outbound ESP Cipher Key :
bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba10345a1039ba8f8a
Transform set           : esp-128-aes esp-sha1-hmac
```

The following example shows the `show crypto ipsec sa ipv6` command.

```
DellEMC#show crypto ipsec sa ipv6

Interface: TenGigabitEthernet 1/1
  Link Local address: fe80::201:e8ff:fe40:4d10
  IPSecv6 policy name: OSPFv3-1-500

  inbound ah sas
   spi : 500 (0x1f4)
    transform : ah-md5-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE

  outbound ah sas
   spi : 500 (0x1f4)
    transform : ah-md5-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE

  inbound esp sas

  outbound esp sas

Interface: TenGigabitEthernet 1/2
  Link Local address: fe80::201:e8ff:fe40:4d11
  IPSecv6 policy name: OSPFv3-1-600

  inbound ah sas

  outbound ah sas

  inbound esp sas
   spi : 600 (0x258)
    transform : esp-des esp-sha1-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE

 outbound esp sas
  spi : 600 (0x258)
   transform : esp-des esp-sha1-hmac
   in use settings : {Transport, }
   replay detection support : N
   STATUS : ACTIVE
```

Troubleshooting OSPFv3

The system provides several tools to troubleshoot OSPFv3 operation on the switch. This section describes typical, OSPFv3 troubleshooting scenarios.

NOTE: The following troubleshooting section is meant to be a comprehensive list, but only to provide some examples of typical troubleshooting checks.

- Have you enabled OSPF globally?
- Is the OSPF process active on the interface?
- Are the adjacencies established correctly?
- Did you configure the interfaces for Layer 3 correctly?
- Is the router in the correct area type?
- Did you include the routes in the OSPF database?
- Did you include the OSPF routes in the routing table (not just the OSPF database)?

Some useful troubleshooting commands are:

- `show ipv6 interfaces`
- `show ipv6 protocols`
- `debug ipv6 ospf events and/or packets`
- `show ipv6 neighbors`
- `show ipv6 routes`

Viewing Summary Information

To get general route, configuration, links status, and debug information, use the following commands.

- View the summary information of the IPv6 routes.
  EXEC Privilege mode
  `show ipv6 route [vrf vrf-name] summary`
- View the summary information for the OSPFv3 database.
  EXEC Privilege mode
  `show ipv6 ospf [vrf vrf-name] database`
- View the configuration of OSPFv3 neighbors.
  EXEC Privilege mode
  `show ipv6 ospf [vrf vrf-name] neighbor`
- View debug messages for all OSPFv3 interfaces.
  EXEC Privilege mode
  `debug ipv6 ospf [vrf vrf-name] [event | packet] {type slot/port}`
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a port channel interface, enter the keywords `port-channel` then a number.
  - For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

36

Policy-based Routing (PBR)

Policy-based routing (PBR) allows a switch to make routing decisions based on policies applied to an interface.

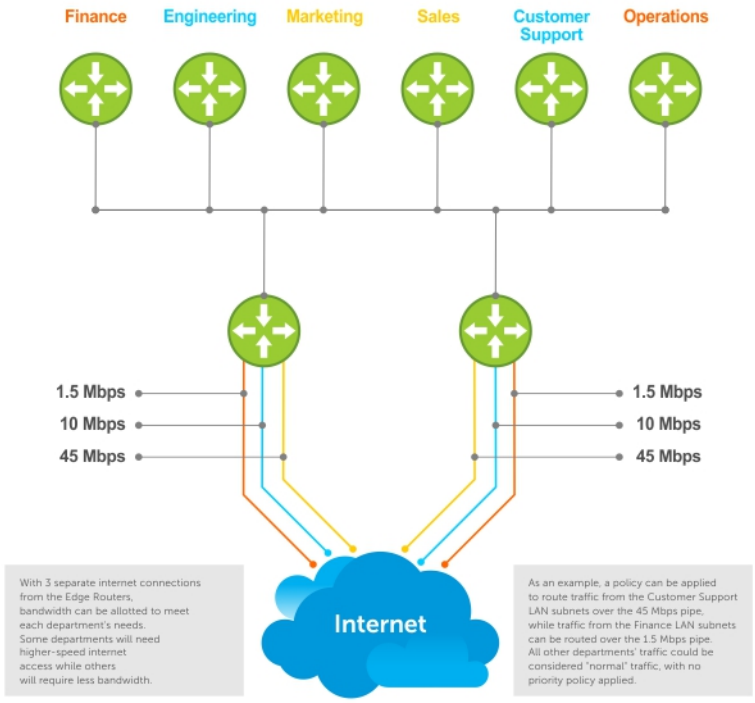
Topics:

- Overview
- Implementing PBR
- Configuration Task List for Policy-based Routing
- Create a Redirect List
- Create a Rule for a Redirect-list
- Apply a Redirect-list to an Interface using a Redirect-group
- Sample Configuration

Overview

When a router receives a packet, the router decides where to forward the packet based on the destination address in the packet, which is used to look up an entry in a routing table. However, in some cases, there may be a need to forward the packet based on other criteria: size, source, protocol type, destination, and so on. For example, a network administrator might want to forward a packet that uses transmission control protocol (TCP) across a different next-hop than packets using Internet control message protocol (ICMP). In these situations, you can a configure switch route packet according to a policy applied to interfaces.

In another scenario, when the packet comes from one source and wants to go to another destination, then route it to this next-hop or onto that specific interface. This permits routing over different links or towards different networks even while the destination is the same but depending on where the packet originates.



To enable PBR, create a redirect list. Redirect lists are defined by rules or routing policies. You can define following parameters in routing policies or rules:

- IP address of the forwarding router (next-hop IP address)
- Protocol as defined in the header
- Source IP address and mask
- Destination IP address and mask
- Source port
- Destination port
- TCP Flags

After you apply a redirect-list to an interface, all traffic passing through it is subjected to the rules defined in the redirect-list.

Traffic is forwarded based on the following:

- Next-hop addresses are verified. If the specified next hop is reachable, traffic is forwarded to the specified next-hop.
- If the specified next-hops are not reachable, the normal routing table is used to forward the traffic.
- Dell EMC Networking OS supports multiple next-hop entries in the redirect lists.
- Redirect-lists are applied at Ingress.

PBR with Redirect-to-Tunnel Option:

You can provide a tunnel ID for a redirect rule. In this case, the resolved next hop is the tunnel interface IP. The qualifiers of the rule pertain to the inner IP details. You must provide a tunnel ID for the next hop to be a tunnel interface. If you do not provide the tunnel destination IP as the next hop, the next hop is treated as an IPv4 next hop and not a tunnel next hop.

PBR with Multiple Tracking Option:

PBR with the multiple tracking option enabled extends and introduces the capabilities of object tracking to verify the next hop IP address before forwarding the traffic to the next hop. The multiple tracking options feature is most suitable for routers which have multiple devices as the next hop (primarily indirect next-hops and/or Tunnel Interfaces). These options allow you to backup Indirect next-hop with another. Choose the specific Indirect next-hop and/or Tunnel interface which is available by sending ICMP pings to verify the reach ability and/or check the Tunnel interface UP or DOWN status, and then route traffic out to that next-hop and/or Tunnel Interface

Implementing PBR

- Non-contiguous bitmasks for PBR
- Hot-Lock PBR

Non-Contiguous Bitmasks for PBR

Non-contiguous bitmasks for PBR allows more granular and flexible control over routing policies. You can include or exclude network addresses that are in the middle of a subnet can be included or excluded. Enter specific bitmasks using the dotted decimal format.

```
DellEMC#show ip redirect-list
IP redirect-list rcl0:
 Defined as:
  seq 5 permit ip 200.200.200.200 200.200.200.200 199.199.199.199 199.199.199.199
  seq 10 redirect 1.1.1.2 tcp 234.224.234.234 255.234.234.234 222.222.222.222/24
  seq 40 ack, Next-hop reachable(via Te 8/1)
 Applied interfaces:
  Te 8/2
```

Hot-Lock PBR

Ingress and egress Hot lock PBR allows you to add or delete new rules into an existing policy (already written into content address memory [CAM]) without disruption to traffic flow. Existing entries in content addressable memory (CAM) are adjusted to accommodate the new entries. Hot Lock PBR is enabled by default.

Configuration Task List for Policy-based Routing

This section explains how to configure PBR on the system.

- Create a Redirect List
- Create a Rule for a Redirect-list
- Create a Track-ID list

- Apply a Redirect-list to an Interface using a Redirect-group

PBR Exceptions (Permit)

To create an exception to a redirect list, use the`permit` command. Exceptions are used when a forwarding decision should be based on the routing table rather than a routing policy.

The Dell EMC Networking OS assigns the first available sequence number to a rule configured without a sequence number and inserts the rule into the PBR CAM region next to the existing entries. Because the order of rules is important, ensure that you configure any necessary sequence numbers.

Never apply the `permit` statement because the redirect list covers all source and destination IP addresses.

```
ip redirect-list rcl0
seq 5 redirect 2.2.2.2 ip any any
seq 10 permit ip host 3.3.3.3 any
```

To ensure the `permit` permit statement or PBR exception is effective, use a lower sequence number, as shown:

```
ip redirect-list rcl0
seq 10 permit ip host 3.3.3.3 any
seq 15 redirect 2.2.2.2 ip any any
```

Create a Redirect List

To create a redirect list, use the following commands.

Create a redirect list by entering the list name.

CONFIGURATION mode

`ip redirect-list` *`redirect-list-name`*

`redirect-list-name`: 16 characters.

To delete the redirect list, use the `no ip redirect-list` command.

The following example creates a redirect list by the name of `xyz`.

```
DellEMC(conf)#ip redirect-list ?
WORD   Redirect-list name (max 16 chars)
DellEMC(conf)#ip redirect-list xyz
```

Create a Rule for a Redirect-list

To set the rules for the redirect list, use the following command. You can enter the command multiple times and create a sequence of redirect rules. Use the `seq` *`nn`* `redirect` version of the command to organize your rules.

Configure a rule for the redirect list.

CONF-REDIRECT-LIST mode

`seq` *{number}* `redirect` *{ip-address | tunnel tunnel-id} [track <obj-id>]{ip-protocol-number | protocol-type* `[`*bit*`]}` *{source mask* | `any` | `host` *ip-address}{destination mask* | `any` | `host` *ip-address}*

- *number* is the number in sequence to initiate this rule
- *ip-address* is the Forwarding router's address
- *tunnel* is used to configure the tunnel settings
- *tunnel-id* is used to redirect the traffic
- *track <obj-id>* is used to track the object-id
- *track* is to enable the tracking
- FORMAT: A.B.C.D

- FORMAT: slot/port
- *ip-protocol-number* or *protocol-type* is the type of protocol to be redirected
- FORMAT: 0-255 for IP protocol number, or enter protocol type
- *source ip-address* or *any* or *host ip-address* is the Source's IP address
- FORMAT: A.B.C.D/NN, or ANY or HOST IP address
- *destination ip-address* or *any* or *host ip-address* is the Destination's IP address
- FORMAT: A.B.C.D/NN, or ANY or HOST IP address

To delete a rule, use the **no redirect** command.

The redirect rule supports Non-contiguous bitmasks for PBR in the Destination router IP address

The following example shows how to create a rule for a redirect list by configuring:

- IP address of the next-hop router in the forwarding route
- IP protocol number
- Source address with mask information
- Destination address with mask information

Example: Creating a Rule

```
DellEMC(conf-redirect-list)#redirect ?
A.B.C.D                 Forwarding router's address

DellEMC(conf-redirect-list)#redirect 3.3.3.3 ?
<0-255>                 An IP protocol number
icmp                    Internet Control Message Protocol
ip                      Any Internet Protocol
tcp                     Transmission Control Protocol
udp                     User Datagram Protocol
DellEMC(conf-redirect-list)#redirect 3.3.3.3 ip ?
A.B.C.D                 Source address
any                     Any source host
host                    A single source host
DellEMC(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 ?
Mask                    A.B.C.D or /nn Mask in dotted decimal or in slash format
DellEMC(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 /32 ?
A.B.C.D                 Destination address
any                     Any destination host
host                    A single destination host
DellEMC(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 /32 77.1.1.1 ?
Mask                    A.B.C.D or /nn Mask in dotted decimal or in slash format
DellEMC(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 /32 77.1.1.1 /32 ?
DellEMC(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 /32 77.1.1.1 /32
DellEMC(conf-redirect-list)#do show ip redirect-list

IP redirect-list xyz:
 Defined as:
  seq 5 redirect 3.3.3.3 ip host 222.1.1.1 host 77.1.1.1
  Applied interfaces:
  None
```

You can apply multiple rules to a single redirect-list. The rules are applied in ascending order, starting with the rule that has the lowest sequence number in a redirect-list displays the correct method for applying multiple rules to one list.

Example: Creating Multiple Rules for a Redirect-List

```
DellEMC(conf)#ip redirect-list test
DellEMC(conf-redirect-list)#seq 10 redirect 10.1.1.2 ip 20.1.1.0/24 any
DellEMC(conf-redirect-list)#seq 15 redirect 10.1.1.3 ip 20.1.1.0/25 any
DellEMC(conf-redirect-list)#seq 20 redirect 10.1.1.3 ip 20.1.1.128/24 any
DellEMC(conf-redirect-list)#show config
!
ip redirect-list test
 seq 10 redirect 10.1.1.2 ip 20.1.1.0/24 any
 seq 15 redirect 10.1.1.3 ip 20.1.1.0/25 any
 seq 20 redirect 10.1.1.3 ip 20.1.1.0/24 any
DellEMC(conf-redirect-list)#
```

NOTE: Starting with the Dell EMC Networking OS version 9.4(0.0), the use of multiple recursive routes with the same source-address and destination-address combination in a redirect policy on an router.

A recursive route is a route for which the immediate next-hop address is learned dynamically through a routing protocol and acquired through a route lookup in the routing table. You can configure multiple recursive routes in a redirect list by entering multiple `seq redirect` commands with the same source and destination address and specify a different next-hop IP address. In this way, the recursive routes are used as different forwarding routes for dynamic failover. If the primary path goes down and the recursive route is removed from the routing table, the `seq redirect` command is ignored and the next command in the list with a different route is used.

Apply a Redirect-list to an Interface using a Redirect-group

IP redirect lists are supported on physical interfaces as well as virtual local area network (VLAN) and port-channel interfaces.

NOTE: When you apply a redirect-list on a port-channel, when traffic is redirected to the next hop and the destination port-channel is shut down, the traffic is dropped. However, the traffic redirected to the destination port-channel is sometimes switched.

To apply a redirect list to an interface, use the following command. You can apply multiple redirect-lists can be applied to a redirect-group. It is also possible to create two or more redirect-groups on one interface for backup purposes.

Apply a redirect list (policy-based routing) to an interface.

INTERFACE mode

`ip redirect-group` *`redirect-list-name`* `test l2-switch`

- *`redirect-list-name`* is the name of a redirect list to apply to this interface.
- FORMAT: up to 16 characters
- You can use the l2–switch option to apply the re-direct list to Layer2 traffic.

NOTE: You can apply the l2–switch option to redirect Layer2 traffic only on a VLAN interface. This VLAN interface must be configured with an IP address for ARP resolution. The Layer2 PBR option matches the layer2 traffic flow. If you un-configure this option, then the Layer2 traffic is not matched.The Layer3 routing is not affected on the same interface on which Layer2 PBR is applied. The port from which Layer2 packets egress and the destination MAC are re-written from static ARP. Layer 2 packets with the re-written destination MAC are forwarded through the outgoing port on the same incoming VLAN interface. The `l2-switch` option ensures that the outgoing VLAN and MAC-SA are changed and TTL is not decremented.

To delete the redirect list from this interface, use the `no ip redirect-group` command.

In this example, the list `xyz` is applied to the 1/1 interface.

Example: Applying a Redirect-list to an Interface

```
DellEMC(conf-if-te-1/1)#ip redirect-group xyz
DellEMC(conf-if-te-1/1)#
```

Example: Applying a Redirect-list to an Interface

```
DellEMC(conf-if-te-1/1)#ip redirect-group test
DellEMC(conf-if-te-1/1)#ip redirect-group xyz
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
 no ip address
 ip redirect-group test
 ip redirect-group xyz
 shutdown
DellEMC(conf-if-te-1/1)#
```

```
DellEMC(conf-if-gi-1/1)#ip redirect-group test
DellEMC(conf-if-gi-1/1)#ip redirect-group xyz
DellEMC(conf-if-gi-1/1)#show config
```

```
!
interface GigabitEthernet 1/1
 no ip address
 ip redirect-group test
 ip redirect-group xyz
 shutdown
DellEMC(conf-if-gi-1/1)#
```

In addition to supporting multiple redirect-lists in a redirect-group, multiple redirect-groups are supported on a single interface. Dell EMC Networking OS has the capability to support multiple groups on an interface for backup purposes.

Show Redirect List Configuration

To view the configuration redirect list configuration, use the following commands.

1. View the redirect list configuration and the associated interfaces.

   EXEC mode

   `show ip redirect-list` *`redirect-list-name`*

2. View the redirect list entries programmed in the CAM.

   EXEC mode

   `show cam pbr`

   `show cam-usage`

   List the redirect list configuration using the `show ip redirect-list redirect-list-name` command. The non-contiguous mask displays in dotted format (x.x.x.x). The contiguous mask displays in /x format.

```
DellEMC#show ip redirect-list explicit_tunnel
IP redirect-list explicit_tunnel:
Defined as:
seq 5 redirect tunnel 1 track 1 tcp 155.55.2.0/24 222.22.2.0/24, Track 1 [up], Next-
hop reachable (via Te 1/32)
seq 10 redirect tunnel 1 track 1 tcp any any, Track 1 [up], Next-hop reachable (via
Te 1/32)
seq 15 redirect tunnel 2 udp 155.55.0.0/16 host 144.144.144.144, Track 1 [up], Next-
hop reachable (via Te 1/32)
seq 35 redirect 155.1.1.2 track 5 ip 7.7.7.0/24 8.8.8.0/24, Track 5 [up], Next-hop
reachable (via Po 5)
seq 30 redirect 155.1.1.2 track 6 icmp host 8.8.8.8 any, Track 5 [up], Next-hop
reachable (via Po 5)
seq 35 redirect 42.1.1.2 icmp host 8.8.8.8 any, Next-hop reachable (via Vl 20)
seq 40 redirect 43.1.1.2 tcp 155.55.2.0/24 222.22.2.0/24, Next-hop reachable (via Vl
30)
seq 45 redirect 31.1.1.2 track 200 ip 12.0.0.0 255.0.0.197 13.0.0.0 255.0.0.197,
Track 200 [up], Next-hop reachable (via Te 1/32)
                                                                              ,
Track 200 [up], Next-hop reachable (via Vl 20)
                                                                              ,
Track 200 [up], Next-hop reachable (via Po 5)
                                                                              ,
Track 200 [up], Next-hop reachable (via Po 7)
                                                                              ,
Track 200 [up], Next-hop reachable (via Te 2/18)
                                                                              ,
Track 200 [up], Next-hop reachable (via Te 2/19)
```

Use the `show ip redirect-list` (without the list name) to display all the redirect-lists configured on the device.

```
DellEMC#show ip redirect-list

IP redirect-list rcl0:
 Defined as:
   seq 5 permit ip 200.200.200.200 200.200.200.200 199.199.199.199 199.199.199.199
   seq 10 redirect 1.1.1.2 tcp 234.224.234.234 255.234.234.234 222.222.222.222/24 eq
40 ack, Next-hop reachable
(via Te 2/1),
```

```
 Applied interfaces:
  Te 2/2
```

NOTE: If you apply the redirect-list to an interface, the output of the `show ip redirect-list redirect-list-name` command displays reachability status for the specified next-hop.

Example: Showing CAM PBR Configuration

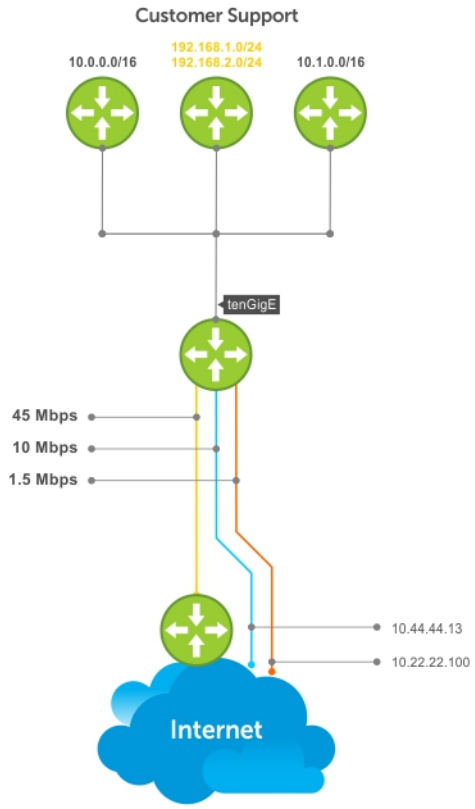
```
DellEMC#show cam pbr stack-unit 1 port-set 0

TCP Flag: Bit 5 - URG, Bit 4 - ACK, Bit 3 - PSH, Bit 2 - RST, Bit 1 - SYN, Bit 0 - FIN

Cam    Port VlanID Proto Tcp    Src    Dst    SrcIp     DstIp   Next-hop          Egress
Index Flag  Port   Port  MAC    Port
-------------------------------------------------------------------------------------
---------------------------
06080 0 N/A    IP     0x0    0 0 200.200.200.200 200.200.200.200 199.199.199.199
199.199.199.199 N/A  NA
06081 0 N/A    TCP    0x10   0   40 234.234.234.234 255.234.234.234 222.222.222.222/24
00:00:00:00:00:09 8/1
```

Sample Configuration

You can use the following example configuration to set up a PBR. These are not comprehensive directions but are intended to give you a guidance with typical configurations. You can copy and paste from these examples to your CLI. Make the necessary changes to support your own IP addresses, interfaces, names, and so on.

The Redirect-List GOLD defined in this example creates the following rules:

- description Route Gold traffic to the DS3
- seq 5 redirect 10.99.99.254 ip 192.168.1.0/24 any " Redirect to next-hop router IP 10.99.99.254 any traffic originating in 192.168.1.0/24"
- seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any " Redirect to next-hop router IP 10.99.99.254 any traffic originating in 192.168.2.0/24"
- seq 15 permit ip any



Create the Redirect-List GOLD

```
EDGE_ROUTER(conf-if-Te-2/23)#ip redirect-list GOLD
EDGE_ROUTER(conf-redirect-list)#description Route GOLD traffic to ISP_GOLD.
EDGE_ROUTER(conf-redirect-list)#direct 10.99.99.254 ip 192.168.1.0/24 any
EDGE_ROUTER(conf-redirect-list)#redirect 10.99.99.254 ip 192.168.2.0/24 any
EDGE_ROUTER(conf-redirect-list)# seq 15 permit ip any any
EDGE_ROUTER(conf-redirect-list)#show config
!
ip redirect-list GOLD
 description Route GOLD traffic to ISP_GOLD.
 seq 5 redirect 10.99.99.254 ip 192.168.1.0/24 any
 seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any
 seq 15 permit ip any any
```

Assign Redirect-List GOLD to Interface 2/11

```
EDGE_ROUTER(conf)#int Te 2/11
EDGE_ROUTER(conf-if-Te-2/11)#ip add 192.168.3.2/24
EDGE_ROUTER(conf-if-Te-2/11)#no shut
EDGE_ROUTER(conf-if-Te-2/11)#
EDGE_ROUTER(conf-if-Te-2/11)#ip redirect-group GOLD
EDGE_ROUTER(conf-if-Te-2/11)#no shut
EDGE_ROUTER(conf-if-Te-2/11)#end
EDGE_ROUTER(conf-redirect-list)#end


EDGE_ROUTER#
```

View Redirect-List GOLD

```
EDGE_ROUTER#show ip redirect-list

IP redirect-list GOLD:
 Defined as:
  seq 5 redirect 10.99.99.254 ip 192.168.1.0/24 any, Next-hop reachable (via Te 3/23)
  seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any, Next-hop reachable (via Te 3/23)
  seq 15 permit ip any any
 Applied interfaces:
  Te 2/11
EDGE_ROUTER#
```

Creating a PBR list using Explicit Track Objects for Redirect IPs

Create Track Objects to track the Redirect IPs:

```
DellEMC#configure terminal
DellEMC(conf)#track 3 ip host 42.1.1.2 reachability
DellEMC(conf-track-3)#probe icmp
DellEMC(conf-track-3)#track 4 ip host 43.1.1.2 reachability
DellEMC(conf-track-4)#probe icmp
DellEMC(conf-track-4)#end
```

Create a Redirect-list with Track Objects pertaining to Redirect-IPs:

```
DellEMC#configure terminal
DellEMC(conf)#ip redirect-list redirect_list_with_track
DellEMC(conf-redirect-list)#redirect 42.1.1.2 track 3 tcp 155.55.2.0/24 222.22.2.0/24
DellEMC(conf-redirect-list)#redirect 42.1.1.2 track 3 tcp any any
DellEMC(conf-redirect-list)#redirect 42.1.1.2 track 3 udp 155.55.0.0/16 host
144.144.144.144
DellEMC(conf-redirect-list)#redirect 42.1.1.2 track 3 udp any host 144.144.144.144
DellEMC(conf-redirect-list)#redirect 43.1.1.2 track 4 ip host 7.7.7.7 host
144.144.144.144
DellEMC(conf-redirect-list)#end
```

Verify the Status of the Track Objects (Up/Down):

```
DellEMC#show track brief

ResId  Resource                    Parameter                State  LastChange
1      Interface ip routing        Tunnel 1                 Up     00:02:16
2      Interface ipv6 routing      Tunnel 2                 Up     00:03:31
3      IP Host reachability        42.1.1.2/32              Up     00:00:59
4      IP Host reachability        43.1.1.2/32              Up     00:00:59

Apply the Redirect Rule to an Interface:

DellEMC#
DellEMC(conf)#int TenGigabitEthernet 2/28
DellEMC(conf-if-te-2/28)#ip redirect-group redirect_list_with_track
DellEMC(conf-if-te-2/28)#end
```

Verify the Applied Redirect Rules:

```
DellEMC#show ip redirect-list redirect_list_with_track

IP redirect-list redirect_list_with_track
 Defined as:
  seq 5 redirect 42.1.1.2 track 3 tcp 155.55.2.0/24 222.22.2.0/24, Track 3 [up], Next-
hop reachable (via Vl 20)
  seq 10 redirect 42.1.1.2 track 3 tcp any any, Track 3 [up], Next-hop reachable (via Vl
20)
  seq 15 redirect 42.1.1.2 track 3 udp 155.55.0.0/16 host 144.144.144.144, Track 3 [up],
Next-hop reachable (via Vl 20)
  seq 20 redirect 42.1.1.2 track 3 udp any host 144.144.144.144, Track 3 [up], Next-hop
reachable (via Vl 20)
```

```
  seq 25 redirect 43.1.1.2 track 4 ip host 7.7.7.7 host 144.144.144.144, Track 4 [up],
Next-hop reachable (via Vl 20)
 Applied interfaces:
  Te 2/28
DellEMC#
```

Creating a PBR list using Explicit Track Objects for Tunnel Interfaces

Creating steps for Tunnel Interfaces:

```
DellEMC#configure terminal
DellEMC(conf)#interface tunnel 1
DellEMC(conf-if-tu-1)#tunnel destination 40.1.1.2
DellEMC(conf-if-tu-1)#tunnel source 40.1.1.1
DellEMC(conf-if-tu-1)#tunnel mode ipip
DellEMC(conf-if-tu-1)#tunnel keepalive 60.1.1.2
DellEMC(conf-if-tu-1)#ip address 60.1.1.1/24
DellEMC(conf-if-tu-1)#ipv6 address 600:10::1/64
DellEMC(conf-if-tu-1)#no shutdown
DellEMC(conf-if-tu-1)#end
DellEMC#

DellEMC#configure terminal
DellEMC(conf)#interface tunnel 2
DellEMC(conf-if-tu-2)#tunnel destination 441:10::2
DellEMC(conf-if-tu-2)#tunnel source 441:10::1
DellEMC(conf-if-tu-2)#tunnel mode ipv6
DellEMC(conf-if-tu-2)#tunnel keepalive 601:10::2
DellEMC(conf-if-tu-2)#ipv6 address 601:10::1/64
DellEMC(conf-if-tu-2)#no shutdown
DellEMC(conf-if-tu-2)#end
DellEMC#
```

Create Track Objects to track the Tunnel Interfaces:

```
DellEMC#configure terminal
DellEMC(conf)#track 1 interface tunnel 1 ip routing
DellEMC(conf-track-1)#exit
DellEMC(conf)#track 2 interface tunnel 2 ipv6 routing
DellEMC(conf-track-2)#end
```

Verify the Status of the Track Objects (Up/Down):

```
DellEMC#show track brief

ResId  Resource                    Parameter     State  LastChange
1      Interface ip routing        Tunnel 1        Up     00:00:00
2      Interface ipv6 routing      Tunnel 2        Up     00:00:00
DellEMC#
```

Create a Redirect-list with Track Objects pertaining to Tunnel Interfaces:

```
DellEMC#configure terminal
DellEMC(conf)#ip redirect-list explicit_tunnel
DellEMC(conf-redirect-list)#redirect tunnel 1 track 1 tcp 155.55.2.0/24 222.22.2.0/24
DellEMC(conf-redirect-list)#redirect tunnel 1 track 1 tcp any any
DellEMC(conf-redirect-list)#redirect tunnel 1 track 1 udp 155.55.0.0/16 host
144.144.144.144
DellEMC(conf-redirect-list)#redirect tunnel 2 track 2 tcp 155.55.2.0/24 222.22.2.0/24
DellEMC(conf-redirect-list)#redirect tunnel 2 track 2 tcp any any
DellEMC(conf-redirect-list)#end
DellEMC#
```

Apply the Redirect Rule to an Interface:

```
DellEMC#configure terminal
DellEMC(conf)#interface TenGigabitEthernet 2/28
DellEMC(conf-if-te-2/28)#ip redirect-group  explicit_tunnel
DellEMC(conf-if-te-2/28)#exit
DellEMC(conf)#end
```

Verify the Applied Redirect Rules:

```
DellEMC#show ip redirect-list explicit_tunnel

IP redirect-list explicit_tunnel:
 Defined as:
  seq 5 redirect tunnel 1 track 1 tcp 155.55.2.0/24 222.22.2.0/24, Track 1 [up], Next-
hop reachable (via Te 1/32)
  seq 10 redirect tunnel 1 track 1 tcp any any, Track 1 [up], Next-hop reachable (via Te
1/32)
  seq 15 redirect tunnel 1 track 1 udp 155.55.0.0/16 host 144.144.144.144, Track 1 [up],
Next-hop reachable (via Te 1/32)
  seq 20 redirect tunnel 2 track 2 tcp 155.55.2.0/24 222.22.2.0/24, Track 2 [up], Next-
hop reachable (via Te 1/33)
  seq 25 redirect tunnel 2 track 2 tcp any any, Track 2 [up], Next-hop reachable (via Te
1/33)
 Applied interfaces:
  Te 2/28
DellEMC#
```

37

PIM Sparse-Mode (PIM-SM)

Protocol-independent multicast sparse-mode (PIM-SM) is a multicast protocol that forwards multicast traffic to a subnet only after a request using a PIM Join message; this behavior is the opposite of PIM-Dense mode, which forwards multicast traffic to all subnets until a request to stop.

Topics:

- Implementation Information
- Protocol Overview
- Configuring PIM-SM
- Enable PIM-SM
- Configuring S,G Expiry Timers
- Configuring a Static Rendezvous Point
- Configuring a Designated Router
- Creating Multicast Boundaries and Domains

Implementation Information

The following information is necessary for implementing PIM-SM.

- The Dell EMC Networking implementation of PIM-SM is based on IETF *Internet Draft draft-ietf-pim-sm-v2-new-05*.
- The platform supports a maximum of 95 IPv4 and IPv6 PIM interfaces and 2000 multicast entries including (*,G), and (S,G) entries. The maximum number of PIM neighbors is the same as the maximum number of PIM-SM interfaces.
- The SPT-Threshold is zero, which means that the last-hop designated router (DR) joins the shortest path tree (SPT) to the source after receiving the first multicast packet.
- Dell EMC Networking OS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.
- Dell EMC Networking OS supports PIM-SM on physical, virtual local area network (VLAN), and port-channel interfaces.

NOTE: Multicast routing is supported across default and non-default VRFs.

Protocol Overview

PIM-SM initially uses unidirectional shared trees to forward multicast traffic; that is, all multicast traffic must flow only from the rendezvous point (RP) to the receivers.

After a receiver receives traffic from the RP, PM-SM switches to SPT to forward multicast traffic. Every multicast group has an RP and a unidirectional shared tree (group-specific shared tree).

Requesting Multicast Traffic

A host requesting multicast traffic for a particular group sends an Internet group management protocol (IGMP) Join message to its gateway router.

The gateway router is then responsible for joining the shared tree to the RP (RPT) so that the host can receive the requested traffic.

1. After receiving an IGMP Join message, the receiver gateway router (last-hop DR) creates a (*,G) entry in its multicast routing table for the requested group. The interface on which the join message was received becomes the outgoing interface associated with the (*,G) entry.
2. The last-hop DR sends a PIM Join message to the RP. All routers along the way, including the RP, create an (*,G) entry in their multicast routing table, and the interface on which the message was received becomes the outgoing interface associated with the (*,G) entry. This process constructs an RPT branch to the RP.

3. If a host on the same subnet as another multicast receiver sends an IGMP report for the same multicast group, the gateway takes no action. If a router between the host and the RP receives a PIM Join message for which it already has a (*,G) entry, the interface on which the message was received is added to the outgoing interface list associated with the (*,G) entry, and the message is not (and does not need to be) forwarded towards the RP.

Refuse Multicast Traffic

A host requesting to leave a multicast group sends an IGMP Leave message to the last-hop DR. If the host is the only remaining receiver for that group on the subnet, the last-hop DR is responsible for sending a PIM Prune message up the RPT to prune its branch to the RP.

1. After receiving an IGMP Leave message, the gateway removes the interface on which it is received from the outgoing interface list of the (*,G) entry. If the (*,G) entry has no remaining outgoing interfaces, multicast traffic for that group is no longer forwarded to that subnet.
2. If the (*,G) entry has no remaining outgoing interfaces, the last-hop DR sends a PIM Prune message to towards the RP. All routers along the way remove the interface on which the message was received from the outgoing interface list of the (*,G) entry. If on any router there is at least one outgoing interface listed for that (*,G) entry, the Prune message is not forwarded.

Send Multicast Traffic

With PIM-SM, all multicast traffic must initially originate from the RP. A source must unicast traffic to the RP so that the RP can learn about the source and create an SPT to it. Then the last-hop DR may create an SPT directly to the source.

1. The source gateway router (first-hop DR) receives the multicast packets and creates an (S,G) entry in its multicast routing table. The first-hop DR encapsulates the initial multicast packets in PIM Register packets and unicasts them to the RP.
2. The RP decapsulates the PIM Register packets and forwards them if there are any receivers for that group. The RP sends a PIM Join message towards the source. All routers between the RP and the source, including the RP, create an (S,G) entry and list the interface on which the message was received as an outgoing interface, thus recreating a SPT to the source.
3. After the RP starts receiving multicast traffic via the (S,G), it unicasts a Register-Stop message to the first-hop DR so that multicast packets are no longer encapsulated in PIM Register packets and unicast. After receiving the first multicast packet from a particular source, the last-hop DR sends a PIM Join message to the source to create an SPT to it.
4. There are two paths, then, between the receiver and the source, a direct SPT and an RPT. One router receives a multicast packet on two interfaces from the same source in this case; this router prunes the shared tree by sending a PIM Prune message to the RP that tells all routers between the source and the RP to remove the outgoing interface from the (*,G) entry, and tells the RP to prune its SPT to the source with a Prune message.

Dell EMC Networking OS Behavior: When the router creates an SPT to the source, there are then two paths between the receiver and the source, the SPT and the RPT. Until the router can prune itself from the RPT, the receiver receives duplicate multicast packets which may cause disruption. Therefore, the router must prune itself from the RPT as soon as possible. Dell EMC Networking OS optimizes the shared to shortest-path tree switchover latency by copying and forwarding the first (S,G) packet received on the SPT to the PIM task immediately upon arrival. The arrival of the (S,G) packet confirms for PIM that the SPT is created, and that it can prune itself from the shared tree.

Important Point to Remember

If you use a Loopback interface with a /32 mask as the RP, you must enable PIM Sparse-mode on the interface.

Configuring PIM-SM

Configuring PIM-SM is a three-step process.

1. Enable IPv4 multicast routing (refer to the following step).
2. Select a rendezvous point.
3. Enable PIM-SM on an interface.

Enable multicast routing.

CONFIGURATION mode

```
ip multicast-routing [vrf vrf-name]
```

Related Configuration Tasks

The following are related PIM-SM configuration tasks.

- Configuring S,G Expiry Timers
- Configuring a Static Rendezvous Point
- Configuring a Designated Router
- Creating Multicast Boundaries and Domains

Enable PIM-SM

You must enable PIM-SM on each participating interface.

1. Enable IPv4 multicast routing on the system.
   CONFIGURATION mode
   `ip multicast-routing [vrf vrf-name]`
2. Enable PIM-Sparse mode.
   INTERFACE mode
   `ip pim sparse-mode`

To display which interfaces are enabled with PIM-SM, use the `show ip pim interface` command from EXEC Privilege mode.

Following is an example of `show ip pim interface` command output:

```
DellEMC#show ip pim interface
Address          Interface Ver/   Nbr    Query  DR         DR
                           Mode   Count  Intvl  Prio
165.87.34.5      Te 1/10   v2/S   0      30     1          165.87.34.5
10.1.1.2         Vl 10     v2/S   1      30     1          10.1.1.2
20.1.1.5         Vl 20     v2/S   1      30     1          20.1.1.5
165.87.31.200    Vl 30     v2/S   1      30     1          165.87.31.201
```

```
DellEMC#show ipv6 pim interface
Interface     Ver/     Nbr      Query    DR
              Mode     Count    Intvl    Prio
Te 1/1        v2/S     1        30       101
Address : fe80::4e76:25ff:fee5:32c2
DR      : this router

Te 1/2        v2/S     1        30       100
Address : fe80::4e76:25ff:fee5:32c2
DR      : this router

Te 1/3        v2/S     0        30       1
Address : fe80::4e76:25ff:fee5:32c2
DR      : this router
```

NOTE: You can influence the selection of the Rendezvous Point by enabling PIM-Sparse mode on a Loopback interface and assigning a low IP address.

To display PIM neighbors for each interface, use the `show ip pim neighbor [detail]` command EXEC Privilege mode.

Following is an example of `show ip pim neighbor` command output:

```
DellEMC#show ip pim neighbor
Neighbor     Interface Uptime/Expires     Ver  DR
Address                                       Prio/Mode
127.87.5.5   Te 1/11   01:44:59/00:01:16  v2   1 / S
127.87.3.5   Te 1/12   01:45:00/00:01:16  v2   1 / DR
```

```
127.87.50.5  Te 2/13   00:03:08/00:01:37  v2   1 / S
DellEMC#
```

```
DellEMC#show ipv6 pim neighbor
Neighbor                  Interface Uptime/Expires     Ver  DR
Address                                                    Prio/Mode
fe80::201:e8ff:fe02:140f  Te 1/11   01:44:59/00:01:16  v2   1 / S
fe80::201:e8ff:fe00:6265  Te 1/12   01:45:00/00:01:16  v2   1 / DR
DellEMC#
```

To display the PIM routing table, use the `show ip pim tib [group-address [source-address]]` command from EXEC privilege mode.

Following is an example of `show ip pim tib` command output:

```
DellEMC#show ip pim tib

PIM Multicast Routing Table
Flags: D - Dense, S - Sparse, C - Connected, L - Local, P - Pruned,
       R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT,
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 192.1.2.1), uptime 00:29:36, expires 00:03:26, RP 10.87.2.6, flags: SCJ
  Incoming interface: TenGigabitEthernet 1/12, RPF neighbor 10.87.3.5
  Outgoing interface list:
    TenGigabitEthernet 1/11
    TenGigabitEthernet 1/13

(10.87.31.5, 192.1.2.1), uptime 00:01:24, expires 00:02:26, flags: FT
  Incoming interface: TenGigabitEthernet 1/10, RPF neighbor 0.0.0.0
  Outgoing interface list:
    TenGigabitEthernet 1/11
    TenGigabitEthernet 1/12
    TenGigabitEthernet 1/13
--More--
```

```
DellEMC#show ipv6 pim tib
PIM Multicast Routing Table
Flags: D - Dense, S - Sparse, C - Connected, L - Local, P - Pruned,
       R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT,
 Timers: Uptime/Expires
 Interface state: Interface, next-Hop, State/Mode

 (*, ff0e::225:1:2:1), uptime 00:29:36, expires 00:03:26, RP 10.87.2.6, flags: SCJ
 Incoming interface: TenGigabitEthernet 1/12, RPF neighbor 35::1
 Outgoing interface list:
 TenGigabitEthernet 1/11
 TenGigabitEthernet 1/13

 (25::1, ff0e::225:1:2:1), uptime 00:01:24, expires 00:02:26, flags: FT
 Incoming interface: TenGigabitEthernet 1/10, RPF neighbor 35::1
 Outgoing interface list:
 TenGigabitEthernet 1/11
 TenGigabitEthernet 1/12
 TenGigabitEthernet 1/13
```

Configuring S,G Expiry Timers

You can configure a global expiry time (for all [S,G] entries). By default, [S,G] entries expire in 210 seconds.

When you create, delete, or update an expiry time, the changes are applied when the keep alive timer refreshes.

To configure a global expiry time, use the following command.

Enable global expiry timer for S, G entries.
CONFIGURATION mode

`ip pim sparse-mode sg-expiry-timer seconds`

The range is from 211 to 86,400 seconds.

The default is **210**.

NOTE: The global expiry time for all [S, G] entries can vary from 360 to 420 seconds.

Configuring a Static Rendezvous Point

The rendezvous point (RP) is a PIM-enabled interface on a router that acts as the root a group-specific tree; every group must have an RP.

- Identify an RP by the IP address of a PIM-enabled or Loopback interface.

  `ip pim rp-address address group-address group-address mask [override]`

Overriding Bootstrap Router Updates

PIM-SM routers must know the address of the RP for each group for which they have (*,G) entry.

This address is obtained automatically through the bootstrap router (BSR) mechanism or a static RP configuration.

Use the following command if you have configured a static RP for a group. If you do not use the `override` option with the following command, the RPs advertised in the BSR updates take precedence over any statically configured RPs.

- Use the `override` option to use statically configured RP instead of the longest prefix match of the configured group address associated with a dynamic RP.

  `ip pim rp-address address {group-address group-address mask} [override]`

To display the assigned RP for a group, use the `show ip pim rp [mapping | group-address]` command from EXEC privilege mode.

Following is an example of `show ip pim rp` command output:

```
DellEMC#show ip pim rp
Group         RP
225.0.1.40    165.87.50.5
226.1.1.1     165.87.50.5
```

To display the assigned RP for a group range (group-to-RP mapping), use the `show ip pim rp mapping` command in EXEC privilege mode.

Following is an example of `show ip pim rp mapping` command output:

```
DellEMC#show ip pim rp mapping
PIM Group-to-RP Mappings
Group(s): 224.0.0.0/4, Static
  RP: 165.87.50.5, v2
```

Configuring a Designated Router

Multiple PIM-SM routers might be connected to a single local area network (LAN) segment. One of these routers is elected to act on behalf of directly connected hosts. This router is the designated router (DR).

The DR is elected using hello messages. Each PIM router learns about its neighbors by periodically sending a hello message out of each PIM-enabled interface. Hello messages contain the IPv4 address of the interface out of which it is sent and a DR priority value. The router with the greatest priority value is the DR. If the priority value is the same for two routers, then the router with the greatest IPv4 address is the DR. By default, the DR priority value is 192, so the IP address determines the DR.

- Assign a DR priority value.

  INTERFACE mode

  `ip pim dr-priority priority-value`

- Change the interval at which a router sends hello messages.

INTERFACE mode

`ip pim query-interval` *seconds*

- Display the current value of these parameter.

  EXEC Privilege mode

  `show ip pim interface`

Following is an example of `show ip pim summary` command output:

```
DellEMC# show ip pim summary

PIM TIB version 495
Uptime 22:44:52
Entries in PIM-TIB/MFC : 2/2

Active Modes :
        PIM-SNOOPING

Interface summary:
        1 active PIM interface
        0 passive PIM interfaces
        3 active PIM neighbors

TIB summary:
       1/1 (*,G) entries in PIM-TIB/MFC
       1/1 (S,G) entries in PIM-TIB/MFC
       0/0 (S,G,Rpt) entries in PIM-TIB/MFC

       0 PIM nexthops
       0 RPs
       0 sources
       0 Register states

Message summary:
       2582/2583 Joins sent/received
       5/0 Prunes sent/received
       0/0 Candidate-RP advertisements sent/received
       0/0 BSR messages sent/received
       0/0 State-Refresh messages sent/received
       0/0 MSDP updates sent/received
       0/0 Null Register messages sent/received
       0/0 Register-stop messages sent/received

Data path event summary:
       0 no-cache messages received
       0 last-hop switchover messages received
       0/0 pim-assert messages sent/received
       0/0 register messages sent/received
DellEMC#
```

```
Dell#show ipv6 pim interface
Interface Ver/ Nbr   Query DR
          Mode Count Intvl Prio

Te 1/3   v2/S 1     30    1
  Address : fe80::201:e8ff:fe02:140f
  DR : this router

Te 1/11  v2/S 0     30    1
  Address : fe80::201:e8ff:fe02:1417
  DR : this router
Dell#
```

Creating Multicast Boundaries and Domains

A PIM domain is a contiguous set of routers that all implement PIM and are configured to operate within a common boundary defined by PIM multicast border routers (PMBRs).

PMBRs connect each PIM domain to the rest of the Internet.

Create multicast boundaries and domains by filtering inbound and outbound bootstrap router (BSR) messages per interface. The following command is applied to the subsequent inbound and outbound updates. Timeout removes existing BSR advertisements.

- Create multicast boundaries and domains by filtering inbound and outbound BSR messages per interface.

  ```
  ip pim bsr-border
  ```
- Remove candidate RP advertisements.

  ```
  clear ip pim rp-mapping
  ```

38

PIM Source-Specific Mode (PIM-SSM)

PIM source-specific mode (PIM-SSM) is a multicast protocol that forwards multicast traffic from a single source to a subnet. In the other versions of protocol independent multicast (PIM), a receiver subscribes to a group only. The receiver receives traffic not just from the source in which it is interested but from all sources sending to that group. PIM-SSM requires that receivers specify the sources in which they are interested using IGMPv3 include messages to avoid receiving unwanted traffic.

PIM-SSM is more efficient than PIM-SM because it immediately creates shortest path trees (SPT) to the source rather than first using shared trees. PIM-SM requires a shared tree rooted at the RP because IGMPv2 receivers do not know about the source sending multicast data. Multicast traffic passes from the source to the receiver through the RP, until the receiver learns the source address, at which point it switches to the SPT. PIM-SSM uses IGMPv3. Because receivers subscribe to a source and group, the RP and shared tree is unnecessary; only SPTs are used. On Dell EMC Networking systems, it is possible to use PIM-SM with IGMPv3 to achieve the same result, but PIM-SSM eliminates the unnecessary protocol overhead.

PIM-SSM also solves the multicast address allocation problem. Applications must use unique multicast addresses because if multiple applications use the same address, receivers receive unwanted traffic. However, global multicast address space is limited. Currently GLOP/EGLOP is used to statically assign Internet-routable multicast addresses, but each autonomous system number yields only 255 multicast addresses. For short-term applications, an address could be leased, but no global dynamic multicast address allocation scheme has been accepted yet. PIM-SSM eliminates the need for unique multicast addresses because routing decisions for (S1, G1) are independent from (S2, G1). As a result, subnets do not receive unwanted traffic when multiple applications use the same address.

Topics:

- Implementation Information
- Configure PIM-SSM
- Enabling PIM-SSM
- Use PIM-SSM with IGMP Version 2 Hosts
- Electing an RP using the BSR Mechanism

Implementation Information

- The Dell EMC Networking implementation of PIM-SSM is based on RFC 3569.
- The Dell EMC Networking OS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.

Important Points to Remember

- The default SSM range is 232/8 always. Applying an SSM range does not overwrite the default range. Both the default range and SSM range are effective even when the default range is not added to the SSM ACL.
- Extended ACLs cannot be used for configuring SSM range. Be sure to create the ACL first and then apply it to the SSM range.
- The default range is always supported, so range can never be smaller than the default.

Configure PIM-SSM

Configuring PIM-SSM is a two-step process.

1. Configure PIM-SSM.
2. Enable PIM-SSM for a range of addresses.

Related Configuration Tasks

- Use PIM-SSM with IGMP Version 2 Hosts

Enabling PIM-SSM

To enable PIM-SSM, follow these steps.

1. Create an ACL that uses permit rules to specify what range of addresses should use SSM.
   CONFIGURATION mode
   `ip access-list standard name`
2. Enter the `ip pim ssm-range` command and specify the ACL you created.
   CONFIGURATION mode
   `ip pim ssm-range acl-name`

To display address ranges in the PIM-SSM range, use the `show ip pim ssm-range` command from EXEC Privilege mode.

```
R1(conf)#do show run pim
!
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim ssm-range ssm
R1(conf)#do show run acl
!
ip access-list standard ssm
  seq 5 permit host 239.0.0.2
R1(conf)#do show ip pim ssm-range
Group Address  / MaskLen
239.0.0.2      / 32
```

Use PIM-SSM with IGMP Version 2 Hosts

PIM-SSM requires receivers that support IGMP version 3. You can employ PIM-SSM even when receivers support only IGMP version 1 or version 2 by translating (*,G) entries to (S,G) entries.

Translate (*,G) entries to (S,G) entries using the `ip igmp ssm-map acl` command source from CONFIGURATION mode. In a standard access list, specify the groups or the group ranges that you want to map to a source. Then, specify the multicast source.

- When an SSM map is in place and Dell EMC Networking OS cannot find any matching access lists for a group, it continues to create (*,G) entries because there is an implicit deny for unspecified groups in the ACL.
- When you remove the mapping configuration, Dell EMC Networking OS removes the corresponding (S,G) states that it created and re-establishes the original (*,G) states.
- You may enter multiple `ssm-map` commands for different access lists. You may also enter multiple `ssm-map` commands for the same access list, as long as they use different source addresses.
- When an extended ACL is associated with this command, Dell EMC Networking OS displays an error message. If you apply an extended ACL before you create it, Dell EMC Networking OS accepts the configuration, but when the ACL is later defined, Dell EMC Networking OS ignores the ACL and the stated mapping has no effect.

To display the source to which a group is mapped, use the `show ip igmp ssm-map [group]` command. If you use the `group` option, the command displays the group-to-source mapping even if the group is not currently in the IGMP group table. If you do not specify the `group` option, the display is a list of groups currently in the IGMP group table that has a group-to-source mapping.

To display the list of sources mapped to a group currently in the IGMP group table, use the `show ip igmp groups group detail` command.

Configuring PIM-SSM with IGMPv2

```
R1(conf)#do show run pim
!
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim ssm-range ssm
R1(conf)#do show run acl
!
ip access-list standard map
seq 5 permit host 239.0.0.2
!
ip access-list standard ssm
  seq 5 permit host 239.0.0.2
R1(conf)#ip igmp ssm-map map 10.11.5.2
R1(conf)#do show ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address   Interface   Mode          Uptime    Expires  Last Reporter
239.0.0.2       Vlan 300    IGMPv2-Compat 00:00:07  Never    10.11.3.2
   Member Ports: Te 1/1
239.0.0.1 Vlan 400 INCLUDE 00:00:10 Never 10.11.4.2
R1(conf)#show ip igmp ssm-map

Interface             Vlan 101
Group                 226.0.0.0
Uptime                10:40:31
Expires               Never
Router mode           IGMPv2
Last reporter         110.0.101.22
Group SSM Mapped source list
Source address        Expires
110.1.1.250           00:02:08
172.16.84.250         00:02:08
R1(conf)#do show ip igmp ssm-map 239.0.0.2
SSM Map Information
Group     : 239.0.0.2
Source(s) : 10.11.5.2
R1(conf)#show ip igmp groups detail

Interface             TenGigabitEthernet 0/10
Group                 225.1.1.1
Uptime                00:03:01
Expires               00:02:09
Router mode           INCLUDE
Last reporter         165.87.34.100
Group source list
Source address        Expires
165.87.32.21          Never
```

```
R1(conf)#do show run pim
!
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim ssm-range ssm
R1(conf)#do show run acl
!
ip access-list standard map
seq 5 permit host 239.0.0.2
!
ip access-list standard ssm
  seq 5 permit host 239.0.0.2
R1(conf)#ip igmp ssm-map map 10.11.5.2
R1(conf)#do show ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address   Interface   Mode          Uptime    Expires  Last Reporter
239.0.0.2       Vlan 300    IGMPv2-Compat 00:00:07  Never    10.11.3.2
   Member Ports: Te 1/1
239.0.0.1 Vlan 400 INCLUDE 00:00:10 Never 10.11.4.2
R1(conf)#do show ip igmp ssm-map
IGMP Connected Group Membership
```

```
Group Address  Interface   Mode           Uptime    Expires  Last Reporter
239.0.0.2      Vlan 300    IGMPv2-Compat  00:00:36  Never    10.11.3.2
   Member Ports: Te 1/1
R1(conf)#do show ip igmp ssm-map 239.0.0.2
SSM Map Information
Group      : 239.0.0.2
Source(s)  : 10.11.5.2
R1(conf)#do show ip igmp groups detail

Interface           Vlan 300
Group               239.0.0.2
Uptime              00:00:01
Expires             Never
Router mode         IGMPv2-Compat
Last reporter       10.11.3.2
Last reporter mode  IGMPv2
Last report         received Join
Group source        list
Source address      Uptime Expires
10.11.5.2 00:00:01  Never

Interface             Vlan 400
Group                 239.0.0.1
Uptime                00:00:05
Expires               Never
Router mode           INCLUDE
Last reporter         10.11.4.2
Last reporter mode    INCLUDE
Last report received  ALLOW
Group source list
Source address  Uptime    Expires
10.11.5.2       00:00:05  00:02:04
  Member Ports: Te 1/2
```

Electing an RP using the BSR Mechanism

Every PIM router within a domain must map a particular multicast group address to the same RP. The group-to-RP mapping may be statically or dynamically configured. RFC 5059 specifies a dynamic, self-configuring method called the Bootstrap Router (BSR) mechanism, by which an RP is elected from a pool of RP candidates (C-RPs).

Some routers within the domain are configured to be C-RPs. Other routers are configured to be Bootstrap Router candidates (C-BSRs); one router is elected the BSR for the domain and the BSR is responsible for forwarding BSM containing RP-set information to other routers.

The RP election process is as follows:

1. C-BSRs flood their candidacy throughout the domain in a BSM. Each message contains a BSR priority value, and the C-BSR with the highest priority value becomes the BSR.
2. Each C-RP unicasts periodic Candidate-RP-Advertisements to the BSR. Each message contains an RP priority value and the group ranges for which it is a C-RP.
3. The BSR collects the most efficient group-to-RP mappings and periodically updates it to all PIM routes in the network.
4. The BSR floods the RP-Set throughout the domain periodically in case new C-RPs are announced, or an RP failure occurs.

Constraints

1. When a multicast group range is removed from the ACL group list, the E-BSR sends the advertisements to the group with hold-time as 0 only when the C-RP timer expires. Till the timer expires, the C-RP will act as a RP for that multicast group.
2. In E-BSR, if the C-RP advertisements are not in synchronization with the standby, first few BCM C-RP advertisement might not have the complete list of RP mappings. Due to this, there is a possibility of RP mapping timeout and momentary traffic loss in the network.
3. If you configure a secondary VLT peer as an E-BSR and in case of ICL flap or failover, the VLT lag will be down resulting a BSM timeout in the PIM domain and a new BSR will be elected. Hence, it is recommended to configure the primary VLT peer as E-BSR.

NOTE: BSR configuration in the multicast topology should ensure that secondary VLT node is not selected as E-BSR. If selected as E-BSR during ICL flap or VLT failover, traffic disruption will be reported.

To enable BSR election for IPv4 or IPv6, perform the following steps:

1. Enter the following IPv4 or IPv6 command to make a PIM router a BSR candidate:
   CONFIGURATION
   ```
   ip pim bsr-candidate
   ipv6 pim bsr-candidate
   ```
2. Enter the following IPv4 or IPv6 command to make a PIM router a RP candidate:
   CONFIGURATION
   ```
   ip pim rp-candidate
   ipv6 pim rp-candidate
   ```
3. Display IPv4 or IPv6 Bootstrap Router information.
   EXEC Privilege
   ```
   show ip pim bsr-router
   ```
   Example:

```
DellEMC# show ip pim bsr-router
PIMv2 Bootstrap information
This system is the Bootstrap Router (v2)
   BSR address: 7.7.7.7 (?)
   BSR Priority: 0, Hash mask length: 30
   Next bootstrap message in 00:00:08
This system is a candidate BSR
    Candidate BSR address: 7.7.7.7, priority: 0, hash mask length: 30
DellEMC#
```

   ```
   show ipv6 pim bsr-router
   ```
   Example:

```
DellEMC#show ipv6 pim bsr-router
PIMv2 Bootstrap information
  BSR address: 200::1 (?)
  BSR Priority: 0, Hash mask length: 126
  Expires:     00:01:43

This system is a candidate BSR
  Candidate BSR address: 100::1, priority: 0, hash mask length: 126

Next Cand_RP_advertisement in 00:00:25
  RP: 100::1(Lo 0)
DellEMC#
```

Enabling RP to Server Specific Multicast Groups

When you configure an RP candidate, its advertisement is sent to the entire multicast address range and the group-to-RP mapping is advertised for the entire range of multicast address. Starting with Dell EMC Networking OS 9.11.0.0, you can configure an RP candidate for a specified range of multicast group address.

The Configured multicast group ranges are used by the BSR protocol to advertise the candidate RPs in the bootstrap messages.

You can configure the multicast group ranges as a standard ACL list of multicast prefixes. You can then associate the configured group list with the RP candidate.

NOTE: · If there is no multicast group list configured for the RP-candidate, the RP candidate will be advertised for all the multicast groups.

To enable an RP to serve specific group of multicast addresses, perform the following step:

Enter the following command to associate a multicast group to an RP candidate:

CONFIGURATION

```
ip pim [vrf vrf-name] rp-Candidate interface [priority] [acl-name]
```

The specified acl-list is associated to the rp-candidate.

NOTE: You can create the ACL list of multicast prefix using the `ip access-list standard` command.

39

Port Monitoring

Port monitoring (also referred to as mirroring) allows you to monitor ingress and/or egress traffic on specified ports. The mirrored traffic can be sent to a port to which a network analyzer is connected to inspect or troubleshoot the traffic.

Mirroring is used for monitoring Ingress or Egress or both Ingress and Egress traffic on a specific port(s). This mirrored traffic can be sent to a port where a network sniffer can connect and monitor the traffic.

Dell EMC Networking OS supports the following mirroring techniques:

- Port-Mirroring — Port Monitoring is a method of monitoring network traffic that forwards a copy of each incoming and outgoing packet from one port of a network router to another port where the packet can be studied.
- Remote Port Monitoring (RPM) — Remote Port Monitoring allows the user to monitor traffic running across a remote device in the same network. Here the mirror traffic is carried over the L2 network, so that probe devices in the network can analyze it. It is an extension to the normal Port Monitoring feature. This feature is generally referred as RPM, where mirror traffic is carried over L2 network.
- Encapsulated Remote-Port Monitoring (ERPM) — ERPM is a feature to encapsulate mirrored packet using GRE with IP delivery so that it can be sent across a routed network.

Topics:

- Important Points to Remember
- Port Monitoring
- Configuring Port Monitoring
- Enabling Flow-Based Monitoring
- Remote Port Mirroring
- Encapsulated Remote Port Monitoring
- ERPM Behavior on a typical Dell EMC Networking OS
- Port Monitoring on VLT

Important Points to Remember

- Port Monitoring is supported on both physical and logical interfaces like virtual area network (VLAN) and port-channel.
- The monitored (the source, [MD]) and monitoring ports (the destination, [MG]) must be on the same switch.
- In general, a monitoring port should have no ip address and no shutdown as the only configuration; Dell EMC Networking OS permits a limited set of commands for monitoring ports. You can display these commands using the ? command. A monitoring port also may not be a member of a VLAN.
- There may only be one destination port (MG) in a monitoring session.
- Source port (MD) can be monitored by more than one destination port (MG).
- Destination port (MG) can be a physical interface or port-channel interface.
- A Port monitoring session can have multiple source statements.
- Range command is supported in the source statement, where we can specify a range of interfaces of (Physical, Port Channel or VLAN) types.
- One Destination Port (MG) can be used in multiple sessions.
- There can be a maximum of 128 source ports in a Port Monitoring session.
- Flow based monitoring is supported for all type of source interfaces.
- Source port (MD) can be a VLAN, where the VLAN traffic received on that port pipe where its members are present is monitored
- Single MD can be monitored on max. of 4 MG ports.

Port Monitoring

Port monitoring is supported on both physical and logical interfaces, such as VLAN and port-channel interfaces. The source port (MD) with monitored traffic and the destination ports (MG) to which an analyzer can be attached must be on the same switch.

You can configure up to 128 source ports in a monitoring session. Only one destination port is supported in a monitoring session. The platform supports multiple source-destination statements in a single monitor session.

The maximum number of source ports that can be supported in a session is 128.

The maximum number of destination ports that can be supported depends on the port mirroring directions as follows:

- 4 per port pipe, if the four destination ports mirror in one direction, either rx or tx.
- 2 per port pipe, if the two destination ports mirror in bidirection.
- 3 per port pipe, if one of the destination port mirrors bidirection and the other two ports mirror in one direction (either rx or tx).

Example of Changing the Destination Port in a Monitoring Session

In the following examples, ports 1/13, 1/14, 1/15, and 1/16 all belong to the same port-pipe. They are pointing to four different destinations (1/1, 1/2, 1/3, and 1/37). Now it is not possible for another source port from the same port-pipe (for example, 1/17) to point to another new destination (for example, 1/4). If you attempt to configure another destination (to create 5 MG port), this message displays: `% Error will be thrown in case of RPM and ERPM features.`

In the following examples, ports 1/13, 1/14, 1/15, and 1/16 all belong to the same port-pipe. They are pointing to four different destinations (1/1, 1/2, 1/3, and 1/37). Now it is not possible for another source port from the same port-pipe (for example, 1/17) to point to another new destination (for example, 1/4). If you attempt to configure another destination (to create 5 MG port), this message displays: `% Error will be thrown in case of RPM and ERPM features.`

```
DellEMC(conf-mon-sess-5)#do show moni session
  SessID  Source          Destination          Dir  Mode  Source IP       Dest IP     DSCP
TTL    Drop  Rate     Gre-Protocol FcMonitor
  ------ ------         -----------          ---  ----  ---------       --------    ----
---    ----  ----     -----------  ---------
     1   Te 1/1             Te 2/1             both Port     N/A             N/A       0
0    No    N/A         N/A         yes
     2   Te 1/1             Te 2/2             both Port     N/A             N/A       0
0    No    N/A         N/A         yes
     3   Te 1/1             Te 2/3             both Port     N/A             N/A       0
0    No    N/A         N/A         yes
     4   Te 1/1             Te 2/4             both Port     N/A             N/A       0
0    No    N/A         N/A         yes
     5   Te 1/1             Te 2/5             both Port     N/A             N/A       0
0    No    N/A         N/A         yes
DellEMC(conf-mon-sess-5)#
```

```
DellEMC(conf)#mon ses 300
DellEMC(conf-mon-sess-300)#source tengig 1/17 destination tengig 1/4 direction tx
% Error: Exceeding max MG ports for this MD port pipe.
DellEMC(conf-mon-sess-300)#
DellEMC(conf-mon-sess-300)#source tengig 1/17 destination tengig 1/1 direction tx
DellEMC(conf-mon-sess-300)#do show mon session
SessID  Source      Destination    Dir  Mode       Source IP      Dest IP         DSCP
TTL    Drop  Rate    Gre-Protocol FcMonitor
------  ------      -----------    ---  ----       ---------      --------        ----
---    ----  ----    -----------  ---------
   0    Te 1/13      Te 2/1        rx   interface   0.0.0.0       0.0.0.0          0
0    No    N/A         N/A         yes
   10   Te 1/14      Te 2/2        rx   interface   0.0.0.0       0.0.0.0          0
0    No    N/A         N/A         yes
   20   Te 1/15      Te 2/3        rx   interface   0.0.0.0       0.0.0.0          0
0    No    N/A         N/A         yes
   30   Te 1/16      Te 2/7        rx   interface   0.0.0.0       0.0.0.0          0
0    No    N/A         N/A         yes
   300  Te 1/17      Te 1/1        rx   interface   0.0.0.0       0.0.0.0          0
0    No    N/A         N/A         yes
  DellEMC#
```

Example of Configuring Another Monitoring Session with a Previously Used Destination Port

```
DellEMC(conf)#monitor session 300
DellEMC(conf-mon-sess-300)#source TenGig 1/17 destination TenGig 1/4 direction tx
% Error: Exceeding max MG ports for this MD port pipe.
DellEMC(conf-mon-sess-300)#
DellEMC(conf-mon-sess-300)#source TenGig 1/17 destination TenGig 1/1 direction tx
DellEMC(conf-mon-sess-300)#do show mon session
SessID  Source      Destination    Dir  Mode        Source IP     Dest IP         DSCP
```

```
TTL    Drop   Rate     Gre-Protocol FcMonitor
------  ------       -----------    ---  ----         ---------      --------         ----
---    ----  ----     -----------  ---------
    0     Te 1/13       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    10    Te 1/14       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    20    Te 1/15       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    30    Te 1/16       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    300   Te 1/17       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
DellEMC#
```

Example of Viewing a Monitoring Session

In the example below, 0/25 and 0/26 belong to Port-pipe 1. This port-pipe has the same restriction of only four destination ports, new or used.

```
DellEMC(conf-mon-sess-300)#do show mon session
SessID  Source       Destination    Dir  Mode         Source IP      Dest IP         DSCP
TTL    Drop   Rate     Gre-Protocol FcMonitor
------  ------       -----------    ---  ----         ---------      --------         ----
---    ----  ----     -----------  ---------
    0     Te 1/13       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    10    Te 1/14       Te 1/2         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    20    Te 1/15       Te 1/3         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    30    Te 1/16       Te 1/37        rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    100   Te 1/25       Te 1/38        rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    110   Te 1/26       Te 1/39        rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
    300   Te 1/17       Te 1/1         rx   interface   0.0.0.0        0.0.0.0          0
0     No     N/A         N/A         yes
DellEMC#
```

Given these parameters, the following illustration shows the possible port monitoring configurations.

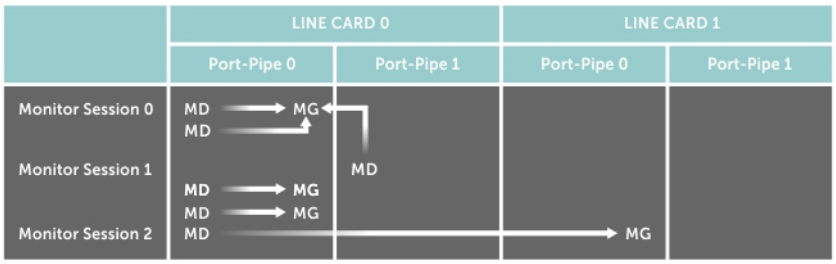


Figure 104. Port Monitoring Configurations

Dell EMC Networking OS Behavior: All monitored frames are tagged if the configured monitoring direction is egress (TX), regardless of whether the monitored port (MD) is a Layer 2 or Layer 3 port. If the MD port is a Layer 2 port, the frames are tagged with the VLAN ID of the VLAN to which the MD belongs. If the MD port is a Layer 3 port, the frames are tagged with VLAN ID 4095. If the MD port is in a Layer 3 VLAN, the frames are tagged with the respective Layer 3 VLAN ID. For example, in the configuration source TenGig 6/1 destination TeGig 6/2 direction tx, if the MD port TenGig 6/1 is an untagged member of any VLAN, all monitored frames that the MG port TeGig 6/2 receives are tagged with the VLAN ID of the MD port. Similarly, if BPDUs are transmitted, the MG port receives them tagged with the VLAN ID 4095. This behavior might result in a difference between the number of egress packets on the MD port and monitored packets on the MG port.

Dell EMC Networking OS Behavior: The platform continues to mirror outgoing traffic even after an MD participating in spanning tree protocol (STP) transitions from the forwarding to blocking.

Configuring Port Monitoring

To configure port monitoring, use the following commands.

1. Verify that the intended monitoring port has no configuration other than no shutdown, as shown in the following example.

   EXEC Privilege mode

   `show interface`

2. Create a monitoring session using the command monitor session from CONFIGURATION mode, as shown in the following example.

   CONFIGURATION mode

   `monitor session`

   ```
   monitor session type rpm/erpm
   type is an optional keyword, required only for rpm and erpm
   ```

3. Specify the source and destination port and direction of traffic, as shown in the following example.

   MONITOR SESSION mode

   `source`

To display information on currently configured port-monitoring sessions, use the `show monitor session` command from EXEC Privilege mode.

```
DellEMC(conf)#monitor session 0
DellEMC(conf-mon-sess-0)#$source te 1/1 dest te 1/2 dir rx
DellEMC(conf-mon-sess-0)#show c
!
monitor session 0
 source TenGigabitEthernet 1/1 destination TenGigabitEthernet 1/2 direction rx
DellEMC(conf-mon-sess-0)#
DellEMC(conf-mon-sess-0)#do show monitor session
  SessID  Source    Destination    Dir  Mode   Source IP    Dest IP    DSCP  TTL   Drop
Rate    Gre-Protocol FcMonitor
  ------  ------    -----------    ---  ----   ---------    --------   ----  ---   ----
----    -----------  ---------
     0   Te 1/1       Te 1/2       rx   Port   0.0.0.0      0.0.0.0     0     0    No
N/A       N/A        No

DellEMC(conf)#monitor session 0
DellEMC(conf-mon-sess-0)#source po 10 dest ten 1/2 dir rx
DellEMC(conf-mon-sess-0)#do show monitor session
  SessID  Source   Destination     Dir  Mode   Source IP    Dest IP    DSCP  TTL   Drop
Rate    Gre-Protocol FcMonitor
  ------  ------   -----------     ---  ----   ---------    --------   ----  ---   ----
----    -----------  ---------
     0   Te 1/1       Te 1/2        rx   Port  0.0.0.0      0.0.0.0     0     0    No
N/A       N/A        No
     0   Po 10        Te 1/2        rx   Port  0.0.0.0      0.0.0.0     0     0    No
N/A       N/A        No

DellEMC(conf)#monitor session 1
DellEMC(conf-mon-sess-1)#source vl 40 dest ten 1/3 dir rx
DellEMC(conf-mon-sess-1)#flow-based enable
DellEMC(conf-mon-sess-1)#exit
DellEMC(conf)#do show monitor session
  SessID  Source   Destination   Dir  Mode   Source IP    Dest IP    DSCP  TTL   Drop
Rate    Gre-Protocol FcMonitor
  ------  ------   -----------   ---  ----   ---------    --------   ----  ---   ----
----    -----------  ---------
     0   Te 1/1      Te 1/2       rx   Port  0.0.0.0      0.0.0.0     0     0    No
N/A       N/A        No
     0   Po 10       Te 1/2       rx   Port  0.0.0.0      0.0.0.0     0     0    No
N/A       N/A        No
     1   Vl 40       Te 1/3       rx   Flow  0.0.0.0      0.0.0.0     0     0    No
N/A       N/A        No
```

NOTE: Source as VLAN is achieved via Flow based mirroring. Please refer section Enabling Flow-Based Monitoring.

In the following example, the host and server are exchanging traffic which passes through the uplink interface 1/1. Port 1/1 is the monitored port and port 1/42 is the destination port, which is configured to only monitor traffic received on tengigabitethernet 1/1 (host-originated traffic).

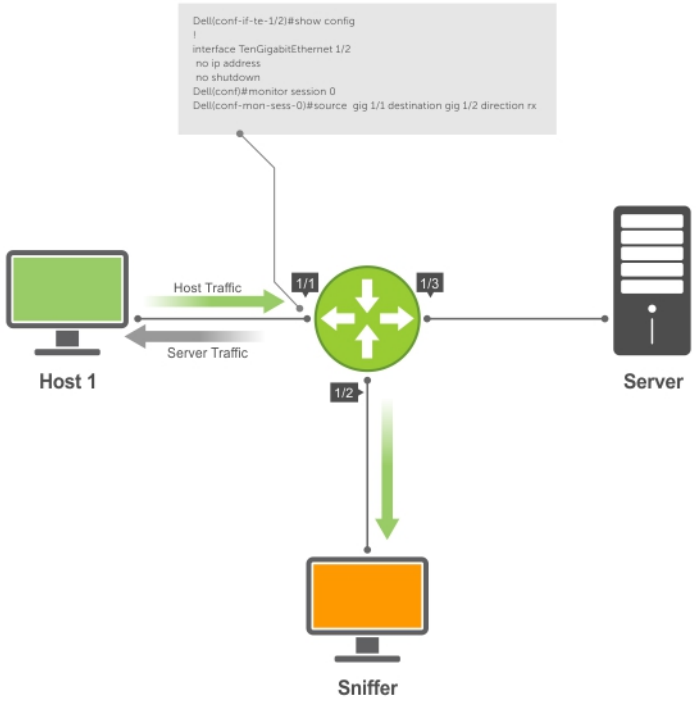


Figure 105. Port Monitoring Example

Enabling Flow-Based Monitoring

Flow-based monitoring conserves bandwidth by monitoring only specified traffic instead of all traffic on the interface. This feature is particularly useful when looking for malicious traffic. It is available for Layer 3 ingress traffic. You can specify traffic using standard or extended access-lists.

NOTE: Flow-based monitoring is supported for known unicast egress traffic.

1. Create a monitoring session.
   CONFIGURATION mode
   `monitor session session-id`
2. Enable flow-based monitoring for a monitoring session.
   MONITOR SESSION mode
   `flow-based enable`
3. Specify the source and destination port and direction of traffic.
   MONITOR SESSION mode
   `source source–port destination destination–port direction rx`
4. Define IP access-list rules that include the `monitor` keyword. For port monitoring, Dell EMC Networking OS only considers traffic matching rules with the `monitor` keyword.
   CONFIGURATION mode
   `ip access-list`
   To define access lists, see the Access Control Lists (ACLs) chapter.
5. Allocate a CAM region so that you can apply the ACL you created to the monitoring session.
   CONFIGURATION mode

```
cam-acl l2acl number ipv4acl number ipv6acl number ipv4qos number l2qos number l2pt
number ipmacacl number vman-qos number ipv4mirracl number
```

6. Apply the ACL to the monitored port.

   MONITOR SESSION mode

   ```
   ip access-group access-list-name
   ```

To view an access-list that you applied to an interface, use the `show ip accounting access-list` command from EXEC Privilege mode.

```
DellEMC(conf)#monitor session 0
DellEMC(conf-mon-sess-0)#flow-based enable
DellEMC(conf)#ip access-list ext testflow
DellEMC(config-ext-nacl)#seq 5 permit icmp any any count bytes monitor
DellEMC(config-ext-nacl)#seq 10 permit ip 102.1.1.0/24 any count bytes monitor
DellEMC(config-ext-nacl)#seq 15 deny udp any any count bytes
DellEMC(config-ext-nacl)#seq 20 deny tcp any any count bytes
DellEMC(config-ext-nacl)#exit
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#ip access-group testflow in
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  ip address 10.11.1.254/24
  ip access-group testflow in
  shutdown
DellEMC(conf-if-te-1/1)#exit
DellEMC(conf)#do show ip accounting access-list testflow
!
Extended Ingress IP access list testflow on TenGigabitEthernet 1/1
Total cam count 4
  seq 5 permit icmp any any 53 monitor 53 count bytes (0 packets 0 bytes)
  seq 10 permit ip 102.1.1.0/24 any monitor 53 count bytes (0 packets 0 bytes)
  seq 15 deny udp any any count bytes (0 packets 0 bytes)
  seq 20 deny tcp any any count bytes (0 packets 0 bytes)
DellEMC(conf)#do show monitor session 0
SessionID Source Destination Direction Mode      Type       Source IP   Dest IP   DSCP
TTL  Drop  Rate  Gre-Protocol FcMonitor
--------- ------ ----------- --------- ----      ----       ---------   --------  ----
---  ----  ----  -----------  ---------
0         Te 1/1 Te 1/2      rx        interface Flow-based  0.0.0.0    0.0.0.0    0
0    No    N/A      N/A        yes
```

Remote Port Mirroring

While local port monitoring allows you to monitor traffic from one or more source ports by directing it to a destination port on the same switch/router, remote port mirroring allows you to monitor Layer 2 and Layer 3 ingress and/or egress traffic on multiple source ports on different switches and forward the mirrored traffic to multiple destination ports on different switches.

Remote port mirroring helps network administrators monitor and analyze traffic to troubleshoot network problems in a time-saving and efficient way.

In a remote-port mirroring session, monitored traffic is tagged with a VLAN ID and switched on a user-defined, non-routable L2 VLAN. The VLAN is reserved in the network to carry only mirrored traffic, which is forwarded on all egress ports of the VLAN. Each intermediate switch that participates in the transport of mirrored traffic must be configured with the reserved L2 VLAN. Remote port monitoring supports mirroring sessions in which multiple source and destination ports are distributed across multiple switches

Remote Port Mirroring Example

Remote port mirroring uses the analyzers shown in the aggregation network in Site A.

The VLAN traffic on monitored links from the access network is tagged and assigned to a dedicated L2 VLAN. Monitored links are configured in two source sessions shown with orange and green circles. Each source session uses a separate reserved VLAN to transmit mirrored packets (mirrored source-session traffic is shown with an orange or green circle with a blue border).

The reserved VLANs transport the mirrored traffic in sessions (blue pipes) to the destination analyzers in the local network. Two destination sessions are shown: one for the reserved VLAN that transports orange-circle traffic; one for the reserved VLAN that transports green-circle traffic.

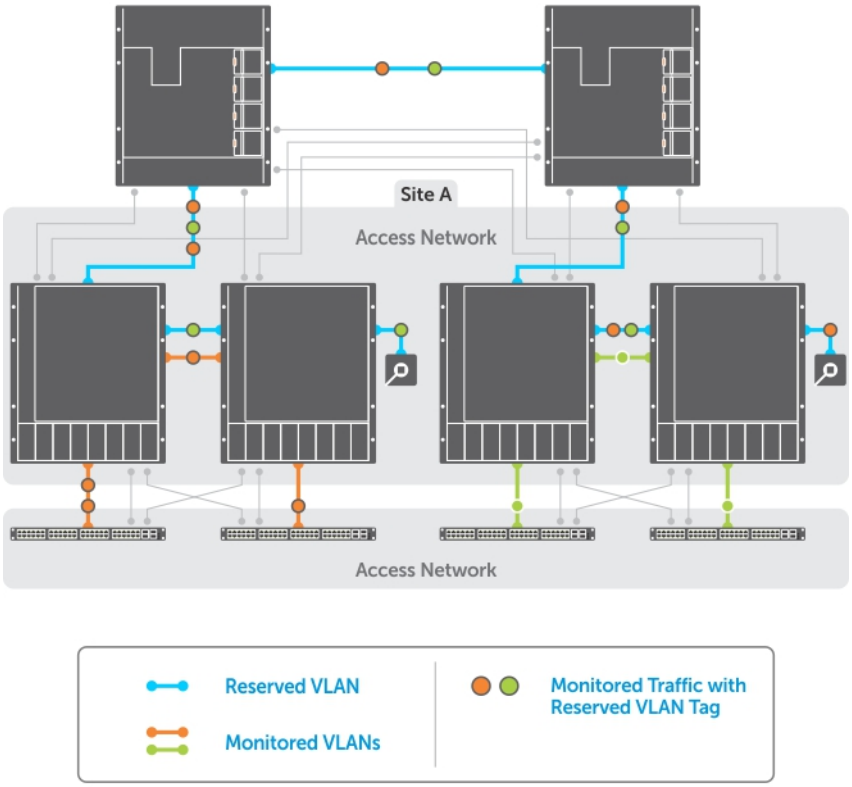


Figure 106. Remote Port Mirroring

Configuring Remote Port Mirroring

Remote port mirroring requires a source session (monitored ports on different source switches), a reserved tagged VLAN for transporting mirrored traffic (configured on source, intermediate, and destination switches), and a destination session (destination ports connected to analyzers on destination switches).

Configuration Notes

When you configure remote port mirroring, the following conditions apply:

- You can configure any switch in the network with source ports and destination ports, and allow it to function in an intermediate transport session for a reserved VLAN at the same time for multiple remote-port mirroring sessions. You can enable and disable individual mirroring sessions.
- BPDU monitoring is not required to use remote port mirroring.
- A remote port mirroring session mirrors monitored traffic by prefixing the reserved VLAN tag to monitored packets so that they are copied to the reserve VLAN.
- Mirrored traffic is transported across the network using 802.1Q-in-802.1Q tunneling. The source address, destination address and original VLAN ID of the mirrored packet are preserved with the tagged VLAN header. Untagged source packets are tagged with the reserve VLAN ID.
- You cannot configure a private VLAN or a GVRP VLAN as the reserved RPM VLAN.
- The RPM VLAN can't be a Private VLAN.
- The RPM VLAN can be used as GVRP VLAN.
- The L3 interface configuration should be blocked for RPM VLAN.
- The member port of the reserved VLAN should have MTU and IPMTU value as MAX+4 (to hold the VLAN tag parameter).

- To associate with source session, the reserved VLAN can have at max of only 4 member ports.
- To associate with destination session, the reserved VLAN can have multiple member ports.
- Reserved Vlan cannot have untagged ports

In the reserved **L2 VLAN** used for remote port mirroring:

- MAC address learning in the reserved VLAN is automatically disabled.
- The reserved VLAN for remote port mirroring can be automatically configured in intermediate switches by using GVRP.
- There is no restriction on the VLAN IDs used for the reserved remote-mirroring VLAN. Valid VLAN IDs are from 2 to 4094. The default VLAN ID is not supported.
- In mirrored traffic, packets that have the same destination MAC address as an intermediate or destination switch in the path used by the reserved VLAN to transport the mirrored traffic are dropped by the switch that receives the traffic if the switch has a L3 VLAN configured.

In a **source session** used for remote port mirroring:

- You can configure any port as a source port in a remote-port monitoring session with a maximum of three source ports per port pipe.
- Maximum number of source sessions supported on a switch: 3

  When you configure remote port mirroring, you can create a maximum of 3 source RSPAN sessions with port mirroring direction as either rx or tx. If you want to configure port mirroring for both directions (tx and rx), you can configure a maximum of 2 RSPAN sessions with one monitor session for both directions and the other session with only one direction.
- Maximum number of source ports supported in a source session: 128
- You can configure physical ports and port-channels as sources in remote port mirroring and use them in the same source session. You can use both Layer 2 (configured with the switchport command) and Layer 3 ports as source ports. You can optionally configure one or more source VLANs to specify the VLAN traffic to be mirrored on source ports.
- You can use the default VLAN and native VLANs as a source VLAN.
- You cannot configure the dedicated VLAN used to transport mirrored traffic as a source VLAN.
- Egressing remote-vlan packets are rate limited to a default value of 100 Mbps. To change the mirroring rate, configure rate-limit within the RPM session.

In a **destination session** used for remote port mirroring:

- Maximum number of destination sessions supported on a switch: 64
- Maximum number ports supported in a destination session: 64.
- You can configure any port as a destination port.
- You can configure additional destination ports in an active session.
- You can tunnel the mirrored traffic from multiple remote-port source sessions to the same destination port.
- By default, destination port sends the mirror traffic to the probe port by stripping off the rpm header. We can also configure the destination port to send the mirror traffic with the rpm header intact in the original mirror traffic..
- By default, ingress traffic on a destination port is dropped.

Restrictions

When you configure remote port mirroring, the following **restrictions** apply:

- You can configure the same source port to be used in multiple source sessions.
- You cannot configure a source port channel or source VLAN in a source session if the port channel or VLAN has a member port that is configured as a destination port in a remote-port mirroring session.
- A destination port for remote port mirroring cannot be used as a source port, including the session in which the port functions as the destination port.
- A destination port cannot be used in any spanning tree instance.
- The reserved VLAN used to transport mirrored traffic must be a L2 VLAN. L3 VLANs are not supported.
- On a source switch on which you configure source ports for remote port mirroring, you can add only one port to the dedicated RPM VLAN which is used to transport mirrored traffic. You can configure multiple ports for the dedicated RPM VLAN on intermediate and destination switches.

Displaying Remote-Port Mirroring Configurations

To display the current configuration of remote port mirroring for a specified session, enter the **show config** command in **MONITOR SESSION** configuration mode.

```
DellEMC(conf-mon-sess-2)#show config
!
monitor session 2 type rpm
 source fortyGigE 1/52 destination remote-vlan 300 direction rx
 source Port-channel 10 destination remote-vlan 300 direction rx
 no disable
```

To display the currently configured source and destination sessions for remote port mirroring on a switch, enter the **show monitor session** command in **EXEC** Privilege mode.

```
DellEMC(conf)#do show monitor session
  SessID  Source         Destination          Dir  Mode  Source IP      Dest IP
  ------  ------         -----------          ---  ----  ---------      --------
     1    remote-vlan 100   Fo 1/32/1           N/A  N/A      N/A            N/A
     1    remote-vlan 100   Po 100              N/A  N/A      N/A            N/A
     2    Fo 1/4/1         remote-vlan 300      rx   Port     N/A            N/A
     2    Po 10             remote-vlan 300     rx   Port     N/A            N/A
```

To display the current configuration of the reserved VLAN, enter the **show vlan** command.

```
DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs, R - Remote Port Mirroring VLANs, P - Primary, C
- Community, I - Isolated
       O - Openflow
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   o - OpenFlow untagged, O - OpenFlow tagged
   G - GVRP tagged, M - Vlan-stack
   i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

    NUM    Status    Description                    Q Ports
*   1      Inactive
R   100    Active                                 T Fo 1/20/1
R   300    Active                                 T Fo 1/24/1
```

Configuration procedure for Remote Port Mirroring

To configure remote port mirroring, you must configure:

1. A reserved VLAN used to transport mirrored packets on source, intermediate, and destination switches
2. A source session that consists of multiple source ports, port channels, and VLANs which are associated with the dedicated VLAN and located on different source switches
3. A destination session that consists of multiple destination ports associated with the dedicated VLAN and located on different destination switches

Configuring a RSPAN VLAN for RPM

Following are the steps for configuring a RSPAN VLAN for RPM.

You must repeat the below mentioned steps on source, intermediate, and destination switches.

1. Enter global configuration mode.

   EXEC mode

   ```
   configure terminal
   ```

2. Create a VLAN to transport mirrored traffic in RPM.

   CONFIGURATION mode

```
interface vlan vlan-id
```

3. Configure the RSPAN VLAN to be used to transport mirrored traffic in RPM.
   VLAN INTERFACE mode
   ```
   mode remote-port-mirroring
   ```
4. Configure a tagged port to carry mirrored traffic in the VLAN.
   VLAN INTERFACE mode
   ```
   tagged interface
   ```
   You can repeat this command to configure additional tagged ports for the VLAN.

Configuring a source session

Following are the steps for configuring a source session on a switch.

You can configure the below steps on other source switches to configure additional source ports for this RPM session.

1. Configure a new RPM session and specifying type as `rpm` defined a RPM session.
   CONFIGURATION mode
   ```
   monitor session session-id type rpm
   ```
   The *session-id* needs to be unique.
2. Configure the source ports or list of ports, ingress/egress traffic to be monitored.
   MONITOR SESSION mode
   ```
   source {interface | range | any} destination remote—vlan vlan-id direction {rx | tx | both}
   ```
3. Activate a RPM session. `no disable` command is mandatory in order for a RPM session to be active.
   MONITOR SESSION mode
   ```
   no disable
   ```
4. (Optional) Enable flow-based monitoring on a flow by flow basis and also for VLAN as source.
   MONITOR SESSION mode
   ```
   flow-based enable
   ```

Configuring a destination session

Following are the steps for configuring a destination session on a switch.

You can configure the below steps on other destination switches to configure additional destination ports for this RPM session.

1. Configure the destination session for RPM.
   CONFIGURATION mode
   ```
   monitor session session-id
   ```
2. Associate the Layer 2 VLAN used to transport monitored traffic with this destination session.
   MONITOR SESSION mode
   ```
   source remote-vlan vlan-id destination interface direction {rx | tx | both}
   ```
3. (Optional) Configure destination ports so that the VLAN tag is added to the monitored traffic.
   MONITOR SESSION mode
   ```
   tagged destination interface
   ```
   To configure destination ports as untagged ports, enter the `untagged destination`command.

Configuration Example of Remote Port Mirroring

This example provides a sample configuration of remote port mirroring (RPM) on a source switch, an intermediate switch, and a destination switch based on the following illustration.

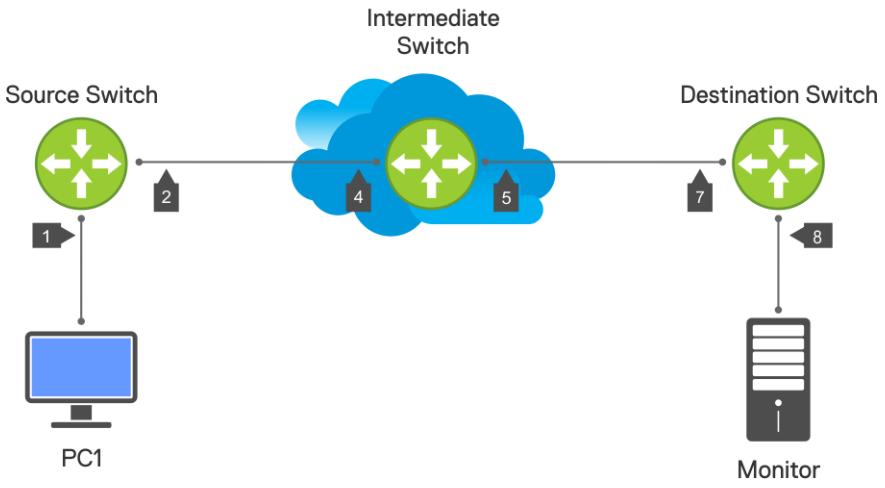


Figure 107. RPM Configuration

Following are the port numbers referred in the above illustration:

- 1 is `tengigabitethernet 1/1`
- 2 is `tengigabitethernet 1/2`
- 4 is `tengigabitethernet 1/4`
- 5 is `tengigabitethernet 1/5`
- 7 is `tengigabitethernet 1/7`
- 8 is `tengigabitethernet 1/8`

Configuring Remote Port Mirroring on a source switch

The below configuration example shows that the source is a source port and the destination is the reserved VLAN (for example, remote-vlan 10).

```
DellEMC(conf)# interface tengigabitethernet 1/2
DellEMC(conf-if-te-1/2)# switchport
DellEMC(conf-if-te-1/2)# no shutdown
DellEMC(conf-if-te-1/2)# exit
DellEMC(conf)#interface vlan 10
DellEMC(conf-if-vl-10)#mode remote-port-mirroring
DellEMC(conf-if-vl-10)#tagged tengigabitethernet 1/2
DellEMC(conf-if-vl-10)#exit
DellEMC(conf)#
DellEMC(conf)#monitor session 1 type rpm
DellEMC(conf-mon-sess-1)#source tengigabitethernet 1/1 destination remote-vlan 10 dir rx
DellEMC(conf-mon-sess-1)#no disable
DellEMC(conf-mon-sess-1)#exit
```

Configuring Remote Port Mirroring on an intermediate switch

Following is a sample configuration of RPM on an intermediate switch.

```
DellEMC(conf)#interface vlan 10
DellEMC(conf-if-vl-10)#mode remote-port-mirroring
DellEMC(conf-if-vl-10)#tagged tengigabitethernet 1/4
DellEMC(conf-if-vl-10)#tagged tengigabitethernet 1/5
DellEMC(conf-if-vl-10)#exit
```

Configuring Remote Port Mirroring on a destination switch

Following is a sample configuration of RPM on an a destination switch.

```
DellEMC(conf)#interface tengigabitethernet 1/7
DellEMC(conf-if-gi-1/9)#switchport
DellEMC(conf-if-gi-1/9)#no shutdown
DellEMC(conf-if-gi-1/9)#exit
DellEMC(conf)#interface vlan 10
DellEMC(conf-if-vl-10)#mode remote-port-mirroring
DellEMC(conf-if-vl-10)#tagged tengigabitethernet 1/7
DellEMC(conf-if-vl-10)#exit
DellEMC(conf)#monitor session 1 type rpm
DellEMC(conf-mon-sess-1)#source remote-vlan 10 destination tengigabitethernet 1/8
DellEMC(conf-mon-sess-1)#exit
```

Configuration Example of Remote Port Mirroring with flow-based enabled

This example provides a sample configuration of remote port mirroring with the flow based monitoring enabled.

Configuring Remote Port Mirroring on source switch

The below configuration example shows that the source is a source VLAN and the destination is the reserved VLAN (for example, remote-vlan 20) with the flow based monitoring enabled.

```
DellEMC(conf)#interface vlan 100
DellEMC(conf-if-vl-100)#tagged tengigabitethernet 1/1
DellEMC(conf-if-vl-100)#exit
DellEMC(conf)#interface vlan 20
DellEMC(conf-if-vl-20)#mode remote-port-mirroring
DellEMC(conf-if-vl-20)#tagged tengigabitethernet 1/2
DellEMC(conf-if-vl-20)#exit
DellEMC(conf)#monitor session 2 type rpm
DellEMC(conf-mon-sess-2)#source vlan 100 destination remote-vlan 20 dir rx
DellEMC(conf-mon-sess-2)#no disable
DellEMC(conf-mon-sess-2)#flow-based enable
DellEMC(conf-mon-sess-2)#exit
DellEMC(conf)#mac access-list standard mac_acl
DellEMC(config-std-macl)#permit 00:00:00:00:11:22 count monitor
DellEMC(config-std-macl)#exit
DellEMC(conf)#interface vlan 100
DellEMC(conf-if-vl-100)#mac access-group mac_acl1 in
DellEMC(conf-if-vl-100)#exit
```

Configuring Remote Port Mirroring on an intermediate switch

Following is a sample configuration of RPM on an intermediate switch.

```
DellEMC(conf)#interface vlan 20
DellEMC(conf-if-vl-20)#mode remote-port-mirroring
DellEMC(conf-if-vl-20)#tagged tengigabitethernet 1/4
DellEMC(conf-if-vl-20)#tagged tengigabitethernet 1/5
DellEMC(conf-if-vl-20)#exit
```

Configuring Remote Port Mirroring on a destination switch

Following is a sample configuration of RPM on a destination switch.

```
Dell(conf)#interface tengigabitethernet 1/7
Dell(conf-if-te-1/7)#switchport
Dell(conf-if-te-1/7)#no shutdown
Dell(conf-if-te-1/7)#exit
Dell(conf)#interface vlan 20
Dell(conf-if-vl-20)#mode remote-port-mirroring
Dell(conf-if-vl-20)#tagged tengigabitethernet 1/7
Dell(conf-if-vl-20)#exit
Dell(conf)#monitor session 2 type rpm
Dell(conf-mon-sess-2)#source remote-vlan 20 destination tengigabitethernet 1/8
Dell(conf-mon-sess-2)#tagged destination tengigabitethernet 1/8
Dell(conf-mon-sess-2)#exit
```

Configuration Example of RPM for port-channel

This example provides a sample configuration of remote port mirroring for the port-channel source interface.

Configuring Remote Port Mirroring on source switch

The below configuration example shows that the source is a source port-channel and the destination is the reserved VLAN (for example, remote-vlan 30).

```
DellEMC(conf)#interface tengigabitethernet 1/3
DellEMC(conf-if-te-1/3)#no shutdown
DellEMC(conf-if-te-1/3)#switchport
DellEMC(conf-if-te-1/3)#exit

DellEMC(conf)#interface vlan 30
DellEMC(conf-if-vl-30)#mode remote-port-mirroring
DellEMC(conf-if-vl-30)#tagged tengigabitethernet 1/3
DellEMC(conf-if-vl-30)#exit

DellEMC(conf)#interface port-channel 10
DellEMC(conf-if-po-10)#channel-member tengigabitethernet 1/8-9
DellEMC(conf-if-po-10)#no shutdown
DellEMC(conf-if-po-10)#exit

DellEMC(conf)#monitor session 3 type rpm
DellEMC(conf-mon-sess-3)#source port-channel 10 destination remote-vlan 30 dir both
DellEMC(conf-mon-sess-3)#no disable
DellEMC(conf-mon-sess-3)#exit
```

Configuring Remote Port Mirroring on an intermediate switch

Following is a sample configuration of RPM on an intermediate switch.

```
DellEMC(conf)#interface vlan 30
DellEMC(conf-if-vl-20)#mode remote-port-mirroring
DellEMC(conf-if-vl-20)#tagged tengigabitethernet 1/4
DellEMC(conf-if-vl-20)#tagged tengigabitethernet 1/5
DellEMC(conf-if-vl-20)#exit
```

Configuring Remote Port Mirroring on a Destination switch

Following is a sample configuration of RPM on a destination switch.

```
DellEMC(conf)#interface tengigabitethernet 1/7
DellEMC(conf-if-te-1/7)#switchport
DellEMC(conf-if-te-1/7)#no shutdown
DellEMC(conf-if-te-1/7)#exit

DellEMC(conf)#interface vlan 30
DellEMC(conf-if-vl-30)#mode remote-port-mirroring
DellEMC(conf-if-vl-30)#tagged tengigabitethernet 1/7
DellEMC(conf-if-vl-30)#exit

DellEMC(conf)#monitor session 3 type rpm
DellEMC(conf-mon-sess-3)#source remote-vlan 30 destination tengigabitethernet 1/8
DellEMC(conf-mon-sess-3)#tagged destination tengigabitethernet 1/8
DellEMC(conf-mon-sess-3)#exit
```

Encapsulated Remote Port Monitoring

Encapsulated Remote Port Monitoring (ERPM) copies traffic from source ports/port-channels or source VLANs and forwards the traffic using routable GRE-encapsulated packets to the destination IP address specified in the session.

NOTE:

When configuring ERPM, follow these guidelines

- The Dell EMC Networking OS supports ERPM source session only. Encapsulated packets terminate at the destination IP address or at the analyzer.

- You can configure up to four ERPM source sessions on switch.
- Configure the system MTU to accommodate the increased size of the ERPM mirrored packet.
- The maximum number of source ports you can define in a session is 128.
- The system encapsulates the complete ingress or egress data under GRE header, IP header, and outer MAC header and sends it out at the next hop interface as pointed by the routing table.
- Specify `flow-based enable` in case of source as VLAN or where you need monitoring on a per-flow basis.
- Specify the `monitor` keyword in the access list rules for which you want to mirror.
- The system allows you to configure up to four ERPM sessions.
- ERPM sessions do not copy locally sourced remote VLAN traffic from source trunk ports that carry RPM VLANs. ERPM sessions do not copy locally sourced ERPM GRE-encapsulated traffic from source ports.
- Flow-based mirroring is supported only for source VLAN ingress traffic.

Changes to Default Behavior

- Rate-limiting is supported for the ERSPAN traffic.
- You can configure the same port as both source and destination in an ERSPAN session.
- You can configure TTL and TOS values in the IP header of the ERSPAN traffic.

Configuration steps for ERPM

To configure an ERPM session:

Table 70. Configuration steps for ERPM

Step	Command	Purpose
1	`configure terminal`	Enter global configuration mode.
2	`monitor session <id> type erpm`	Specify a session ID and ERPM as the type of monitoring session, and enter the Monitoring-Session configuration mode. The session number needs to be unique and not already defined.
3	`source { interface | range } direction {rx | tx | both}`	Specify the source port or range of ports. Specify the ingress (rx), egress (tx), or both ingress and egress traffic to be monitored. You can enter multiple source statements in an ERPM monitoring session
4	`erpm source-ip <id> dest-ip <id> gre-protocol <value>`	Specify the source IP address, destination IP address, and GRE-protocol type value to which encapsulated mirrored traffic is sent.
5	`no flow-based enable`	ERPM to be performed on a flow-by-flow basis or if you configure a VLAN source interface. Enter the `no flow-based` command to disable flow-based ERPM.
6	`no disable`	Enter the no disable command to enable the ERPM session.

The following example shows an ERPM configuration:

```
DellEMC(conf)#monitor session 0 type erpm
DellEMC(conf-mon-sess-0)#source tengigabitethernet 1/9 direction rx
DellEMC(conf-mon-sess-0)#source port-channel 1 direction tx
DellEMC(conf-mon-sess-0)#erpm source-ip 1.1.1.1 dest-ip 7.1.1.2 gre-protocol 111
DellEMC(conf-mon-sess-0)#no disable

DellEMC(conf)#monitor session 1 type erpm
DellEMC(conf-mon-sess-1)#source vlan 11 direction rx
DellEMC(conf-mon-sess-1)#erpm source-ip 5.1.1.1 dest-ip 3.1.1.2 gre-protocol 139
DellEMC(conf-mon-sess-1)#flow-based enable
DellEMC(conf-mon-sess-1)#no disable

DellEMC# show monitor session
```

```
SessID Source  Destination Dir Mode Source IP  Dest IP  DSCP TTL  Drop  Rate  Gre-
Protocol FcMonitor Status
------ ------  ----------- --- ---- ---------  -------- ---- ---  ----  ----
---------- --------- -------
0       Te 1/9  remote-ip  rx  Port  1.1.1.1   7.1.1.2   0   255   No   100    111
   No       Enabled
0       Po 1    remote-ip  tx  Port  1.1.1.1   7.1.1.2   0   255   No   100    111
   No       Enabled
1       Vl 11   remote-ip  rx  Flow  5.1.1.1   3.1.1.2   0   255   No   100    139
   No       Enabled
```

The next example shows the configuration of an ERPM session in which VLAN 11 is monitored as the source interface and a MAC ACL filters the monitored ingress traffic.

```
DellEMC(conf)#mac access-list standard flow
DellEMC(config-std-macl)#seq 5 permit 00:00:0a:00:00:0b count monitor

DellEMC#show running-config interface vlan 11
!
interface Vlan 11
 no ip address
 tagged TenGigabitEthernet 1/1-3
 mac access-group flow in  <<<<<<<<<<<<<< Only ingress packets are supported for
mirroring
 shutdown
```

ERPM Behavior on a typical Dell EMC Networking OS

The Dell EMC Networking OS is designed to support only the Encapsulation of the data received / transmitted at the specified source port (Port A). An ERPM destination session / decapsulation of the ERPM packets at the destination Switch are not supported.

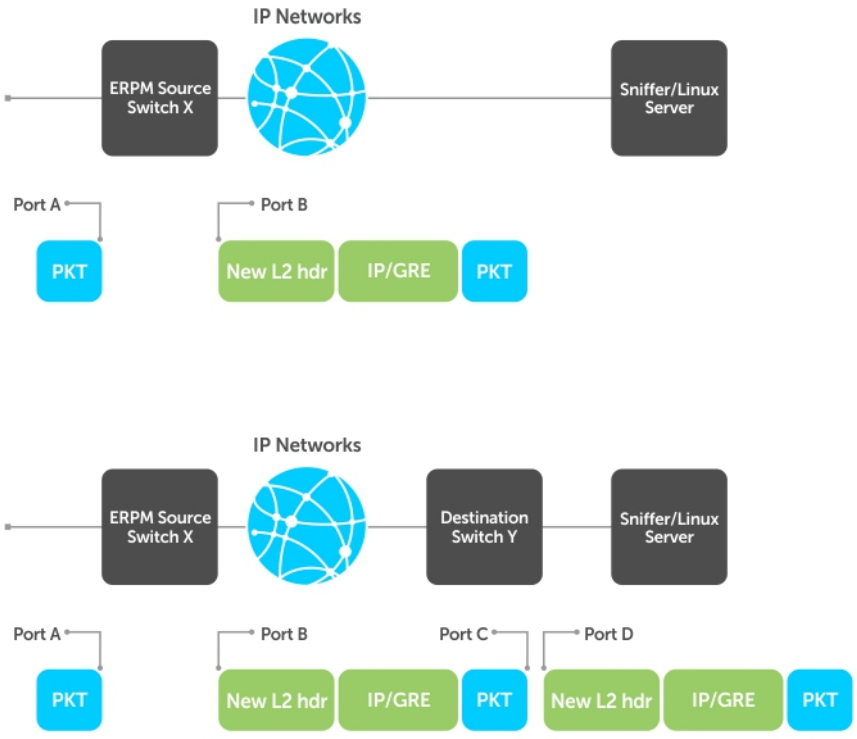


Figure 108. ERPM Behavior

As seen in the above figure, the packets received/transmitted on Port A will be encapsulated with an IP/GRE header plus a new L2 header and sent to the destination ip address (Port D's ip address) on the sniffer. The Header that gets attached to the packet is 38 bytes long.

If the sniffer does not support IP interface, a destination switch will be needed to receive the encapsulated ERPM packet and locally mirror the whole packet to the Sniffer or a Linux Server.

Decapsulation of ERPM packets at the Destination IP/ Analyzer

- In order to achieve the decapsulation of the original payload from the ERPM header. The below two methods are suggested :
  1. Using Network Analyzer
     - Install any well-known Network Packet Analyzer tool which is open source and free to download.
     - Start capture of ERPM packets on the Sniffer and save it to the trace file (for example : erpmwithheader.pcap).
     - The Header that gets attached to the packet is 38 bytes long. In case of a packet with L3 VLAN, it would be 42 bytes long. The original payload /original mirrored data starts from the 39th byte in a given ERPM packet. The first 38/42 bytes of the header needs to be ignored/ chopped off.
     - Some tools support options to edit the capture file. We can make use of such features (for example: **editcap**) and chop the ERPM header part and save it to a new trace file. This new file (i.e. the original mirrored packet) can be converted back into stream and fed to any egress interface.
  2. Using Python script
     - Either have a Linux server's ethernet port ip as the ERPM destination ip or connect the ingress interface of the server to the ERPM MirrorToPort. The analyzer should listen in the forward/egress interface. If there is only one interface, one can choose the ingress and forward interface to be same and listen in the tx direction of the interface.
     - Download/ Write a small script (for example: **erpm.py**) such that it will strip the given ERPM packet starting from the bit where GRE header ends. Basically all the bits after 0x88BE need to be removed from the packet and sent out through another interface.
     - This script **erpm.zip** is available for download at the following location: https://en.community.dell.com/techcenter/networking/m/force10_networking_scripts/20438882.aspx
     - Unzip the **erpm.zip** and copy the erpm.py file to the Linux server.
     - Run the python script using the following command:

     ```
     python erpm.py -i <ingress interface> -o <egress interface>
     ```

     erpm.py : This is the script downloaded from the script store.

<Ingress interface> : Specify the interface id which is connected to the mirroring port or this should be interface whose ip address has been specified as the destination ip address in the ERPM session.

<Egress interface> : Specify another interface on the Linux server via which the decapsulation packets can Egress. In case there is only one interface, the ingress interface itself can be specified as Egress and the analyzer can listen in the tx direction.

Port Monitoring on VLT

Devices on which VLT is configured are seen as a single device in the network. You can apply port monitoring function on the VLT devices in the network.

Port monitoring enables ingress or egress traffic traversing on a port to be sent to another port so that the traffic can be analyzed. The port to which traffic is sent for analysis is called the mirroring port. This port is connect to a port analyzer, which performs the traffic analysis function.

Depending up on the location of the port to which the port analyzer is connected, port monitoring is classified into three categories: local Port mirroring, remote port mirroring (RPM), and encapsulated remote port mirroring (ERPM).

In VLT RPM, Dell EMC Networking OS supports only 3 Mirror-to-Port (MTPs) in hardware.

By default, 2 MTPs are used for ICL lag in VLT. The remaining one MTP is used to mirror only in one direction, either rx or tx.

NOTE: For more information on port monitoring, see Port Monitoring.

The port monitoring or mirroring function when applied to VLT devices works as expected except with some restrictions. You can configure RPM or ERPM monitoring between two VLT peers. As VLT devices are seen as a single device in the network, when a fail over occurs, the source or destination port on one of the VLT peers becomes inactive causing the monitoring session to fail. As a result, Dell EMC Networking OS does not allow local Port mirroring based monitoring to be configured between VLT peers. However, you can create local Port mirroring monitoring sessions separately on individual devices that are a part of the VLT configuration.

NOTE: For more information on configuring VLT, see Configuring VLT on page 943.

VLT Non-fail over Scenario

Consider a scenario where port monitoring is configured to mirror traffic on a VLT device's port or LAG to a destination port on some other device (TOR) on the network. When there is no fail over to the VLT peer, the VLTi link (ICL LAG) also receives the mirrored traffic as the VLTi link is added as an implicit member of the RPM vlan. As a result, the mirrored traffic also reaches the peer VLT device effecting VLTi link's bandwidth usage.

To mitigate this issue, the L2 VLT egress mask drops the duplicate packets that egress out of the VLT port. If the LAG status of the peer VLT device is OPER-UP, then the other VLT peer blocks the transmission of packets received through VLTi to its port or LAG. As a result, the destination port on the device to which the packet analyzer is connected does not receive duplicate mirrored packets.

VLT Fail-over Scenario

Consider a scenario where port monitoring is configured to mirror traffic on the source port or LAG of a VLT device to a destination port on an other device on the network. A fail-over occurs when the primary VLT device fails causing the secondary VLT device to take over. At the time of failover, the mirrored packets are dropped for some time. This time period is equivalent to the gracious VLT failover recovery time.

RPM over VLT Scenarios

This section describes the restrictions that apply when you configure RPM in a VLT set up. Consider a simple VLT setup where two VLT peers are connected using VLTi and a top-of-rack switch is connected to both the VLT peers using VLT LAGs in a ring topology. In this setup, the following table describes the possible restrictions that apply when RPM is used to mirror traffic:

Table 71. RPM over VLT Scenarios

Scenario	RPM Restriction	Recommended Solution
Mirroring an Orphan Port on a VLT LAG — In this scenario, the orphan port on a VLT device is mirrored to the VLT LAG that connects a top-of-rack (TOR) switch to the VLT device. The packet analyzer is connected to the TOR switch.	The bandwidth of the VLTi link is unnecessarily used by mirrored traffic if max rate limit value is configured in the RPM mirror session.	Use ERPM session instead of RPM.
Mirroring an ICL LAG to Orphan Port — In this scenario, the ICL LAG is mirrored to any orphan port on the same VLT device. The packet analyzer is connected to the local VLT device through the orphan port.	No restrictions apply.	If the packet analyzer is directly connected to the VLT device, use local Port mirroring session instead of RPM.
Mirroring an ICL LAG to the VLT LAG — In this scenario, the ICL LAG is mirrored to the VLT LAG on the same VLT device. Packet analyzer is connected to the TOR.	No restrictions apply.	None.
Mirroring VLT LAG to Orphan Port — In this scenario, the VLT LAG is mirrored to an orphan port on the same VLT device.	No restrictions apply.	If the packet analyzer is directly connected to the VLT device, use local Port mirroring session instead of RPM.

Table 71. RPM over VLT Scenarios (continued)

Scenario	RPM Restriction	Recommended Solution
The packet analyzer is connected to the VLT device through the orphan port..		
Mirroring using Intermediate VLT device — In this scenario, the VLT device acts as the intermediate device in remote mirroring. The TOR switch contains the source-RPM configurations that enable mirroring of the VLT lag (of the TOR switch) to any orphan port in the VLT device. The packet analyzer is connected through the VLT device, but not directly to the VLT device.	No restrictions apply	None.
Mirroring Orphan Ports across VLT Devices — In this scenario, an orphan port on the primary VLT device is mirrored to another orphan port on the secondary VLT device through the ICL LAG. The port analyzer is connected to the secondary VLT device.	No restrictions apply to the RPM session. The following example shows the configuration on the primary VLT device:`source orphan port destination remote vlan direction rx/tx/both`.The following example shows the configuration on the secondary VLT device:`source remote vlan destination orphan port`.	None.
Mirroring VLT LAG across VLT Peers — In this scenario, the VLT LAG on the primary VLT peer is mirrored to an orphan port on the secondary VLT peer through the ICL LAG. The packet analyzer is connected to the secondary VLT peer.	No restrictions apply to the RPM session. The following example shows the configuration on the primary VLT device:`source VLT LAG destination remote vlan direction rx/tx/both`.The following example shows the configuration on the secondary VLT device:`source remote vlan destination orphan port`.	None
Mirroring member port of ICL LAG to Orphan Port of peer vlt device— In this scenario, a member port of the ICL LAG or a member port of the VLT LAG is mirrored to an orphan port on the peer VLT device. The packet analyzer is connected to the peer VLT device.	The bandwidth of the VLTi link is unnecessarily used by mirrored traffic if max rate limit value is configured in the RPM mirror session.	None.
Mirroring member port of ICL LAG to VLT LAG — In this scenario, a member port of the ICL LAG is mirrored to the VLT LAG on the same VLT device. The packet analyzer is connected to the TOR switch.	No restrictions apply. The bandwidth of the VLTi link is unnecessarily used by mirrored traffic if max rate limit value is configured in the RPM mirror session.	If you want to mirror traffic in the TOR locally, use local Port mirroring session instead of RPM.
Mirroring with a VLAN as source and destination — If the members of the source and destination VLANs are same in a single monitoring session.	No restrictions apply.	None.
Mirroring with an interface or LAG as source and destination --- If the source and destination interface or LAG of a monitor session are same.	No restrictions apply.	None.

40

Private VLANs (PVLAN)

The private VLAN (PVLAN) feature is supported on Dell EMC Networking OS.

For syntax details about the commands described in this chapter, refer to the Private VLANs commands chapter in the *Dell EMC Networking OS Command Line Reference Guide*.

Private VLANs extend the Dell EMC Networking OS security suite by providing Layer 2 isolation between ports within the same virtual local area network (VLAN). A PVLAN partitions a traditional VLAN into subdomains identified by a primary and secondary VLAN pair. Private VLANs block all traffic to isolated ports except traffic from promiscuous ports. Traffic received from an isolated port is forwarded only to promiscuous ports or trunk ports.

Example uses of PVLANs:

- A hotel can use an isolated VLAN in a PVLAN to provide Internet access for its guests, while stopping direct access between the guest ports.
- A service provider can provide Layer 2 security for customers and use the IP addresses more efficiently, by using a separate community VLAN per customer and at the same time using the same IP subnet address space for all community and isolated VLANs mapped to the same primary VLAN.
  - In more detail, community VLANs are especially useful in the service provider environment because multiple customers are likely to maintain servers that must be strictly separated in customer-specific groups. A set of servers owned by a customer could comprise a community VLAN, so that those servers could communicate with each other, and would be isolated from other customers. Another customer might have another set of servers in another community VLAN. Another customer might want an isolated VLAN, which has one or more ports that are also isolated from each other.

Topics:

- Private VLAN Concepts
- Using the Private VLAN Commands
- Configuration Task List
- Private VLAN Configuration Example
- Inspecting the Private VLAN Configuration

Private VLAN Concepts

Review the following PVLAN concepts before you create PVLANs on your system.

The VLAN types in a PVLAN include:

- **Community VLAN** — a type of secondary VLAN in a primary VLAN:
  - Ports in a community VLAN can communicate with each other.
  - Ports in a community VLAN can communicate with all promiscuous ports in the primary VLAN.
  - A community VLAN can only contain ports configured as host.
- **Isolated VLAN** — a type of secondary VLAN in a primary VLAN:
  - Ports in an isolated VLAN cannot talk directly to each other.
  - Ports in an isolated VLAN can only communicate with promiscuous ports in the primary VLAN.
  - An isolated VLAN can only contain ports configured as host.
- **Primary VLAN** — the base VLAN of a PVLAN:
  - A switch can have one or more primary VLANs, and it can have none.
  - A primary VLAN has one or more secondary VLANs.
  - A primary VLAN and each of its secondary VLANs decrement the available number of VLAN IDs in the switch.
  - A primary VLAN has one or more promiscuous ports.
  - A primary VLAN might have one or more trunk ports, or none.
- **Secondary VLAN** — a subdomain of the primary VLAN.
  - There are two types of secondary VLAN — community VLAN and isolated VLAN.

PVLAN port types include:

- **Community port** — a port that belongs to a community VLAN and is allowed to communicate with other ports in the same community VLAN and with promiscuous ports.
- **Host port** — in the context of a private VLAN, is a port in a secondary VLAN:
  - The port must first be assigned that role in INTERFACE mode.
  - A port assigned the host role cannot be added to a regular VLAN.
- **Isolated port** — a port that, in Layer 2, can only communicate with promiscuous ports that are in the same PVLAN.
- **Promiscuous port** — a port that is allowed to communicate with any other port type in the PVLAN:
  - A promiscuous port can be part of more than one primary VLAN.
  - A promiscuous port cannot be added to a regular VLAN.
- **Trunk port** — carries traffic between switches:
  - A trunk port in a PVLAN is always tagged.
  - In tagged mode, the trunk port carries the primary or secondary VLAN traffic. The tag on the packet helps identify the VLAN to which the packet belongs.
  - A trunk port can also belong to a regular VLAN (non-private VLAN).

Each of the port types can be any type of physical Ethernet port, including port channels (LAGs).

Using the Private VLAN Commands

To use the PVLAN feature, use the following commands.

- Enable/disable Layer 3 communication between secondary VLANs.

  INTERFACE VLAN mode

  ```
  [no] ip local-proxy-arp
  ```

  NOTE: Even after you disable `ip-local-proxy-arp` (`no ip-local-proxy-arp`) in a secondary VLAN, Layer 3 communication may happen between some secondary VLAN hosts, until the address resolution protocol (ARP) timeout happens on those secondary VLAN hosts.

- Set the mode of the selected VLAN to community, isolated, or primary.

  INTERFACE VLAN mode

  ```
  [no] private-vlan mode {community | isolated | primary}
  ```

- Map secondary VLANs to the selected primary VLAN.

  INTERFACE VLAN mode

  ```
  [no] private-vlan mapping secondary-vlan vlan-list
  ```

- Display type and status of PVLAN interfaces.

  EXEC mode or EXEC Privilege mode

  ```
  show interfaces private-vlan [interface interface]
  ```

- Display PVLANs and/or interfaces that are part of a PVLAN.

  EXEC mode or EXEC Privilege mode

  ```
  show vlan private-vlan [community | interface | isolated | primary | primary_vlan |
  interface interface]
  ```

- Display primary-secondary VLAN mapping.

  EXEC mode or EXEC Privilege mode

  ```
  show vlan private-vlan mapping
  ```

- Set the PVLAN mode of the selected port.

  INTERFACE

  ```
  switchport mode private-vlan {host | promiscuous | trunk}
  ```

NOTE: Secondary VLANs are Layer 2 VLANs, so even if they are operationally down while primary VLANs are operationally up, Layer 3 traffic is still transmitted across secondary VLANs.

NOTE: The outputs of the `show arp` and `show vlan` commands provide PVLAN data. For more information, refer to the *Dell EMC Networking OS Command Line Reference Guide*.

Configuration Task List

The following sections contain the procedures that configure a private VLAN.

- Creating PVLAN Ports
- Creating a Primary VLAN
- Creating a Community VLAN
- Creating an Isolated VLAN

Creating PVLAN ports

PVLAN ports are ports that will be assigned to the PVLAN.

1. Access INTERFACE mode for the port that you want to assign to a PVLAN.
   CONFIGURATION mode
   `interface interface`
2. Enable the port.
   INTERFACE mode
   `no shutdown`
3. Set the port in Layer 2 mode.
   INTERFACE mode
   `switchport`
4. Select the PVLAN mode.
   INTERFACE mode
   `switchport mode private-vlan {host | promiscuous | trunk}`
   - `host` (isolated or community VLAN port)
   - `promiscuous` (intra-VLAN communication port)
   - `trunk` (inter-switch PVLAN hub port)

NOTE: You cannot add interfaces that are configured as PVLAN ports to regular VLANs. You also cannot add "regular" ports (ports not configured as PVLAN ports) to PVLANs.

The following example shows the `switchport mode private-vlan` command on a port and on a port channel.

```
DellEMC#conf
DellEMC(conf)#interface TenGigabitEthernet 2/1
DellEMC(conf-if-te-2/1)#switchport mode private-vlan promiscuous

DellEMC(conf)#interface TenGigabitEthernet 2/2
DellEMC(conf-if-te-2/2)#switchport mode private-vlan host

DellEMC(conf)#interface TenGigabitEthernet 2/3
DellEMC(conf-if-te-2/3)#switchport mode private-vlan trunk

DellEMC(conf)#interface TenGigabitEthernet 2/2
DellEMC(conf-if-te-2/2)#switchport mode private-vlan host

DellEMC(conf)#interface port-channel 10
DellEMC(conf-if-po-10)#switchport mode private-vlan promiscuous
```

Creating a Primary VLAN

A primary VLAN is a port-based VLAN that is specifically enabled as a primary VLAN to contain the promiscuous ports and PVLAN trunk ports for the private VLAN.

A primary VLAN also contains a mapping to secondary VLANs, which comprise community VLANs and isolated VLANs.

1. Access INTERFACE VLAN mode for the VLAN to which you want to assign the PVLAN interfaces.
   CONFIGURATION mode

```
interface vlan vlan-id
```

2. Enable the VLAN.
   INTERFACE VLAN mode
   ```
   no shutdown
   ```
3. Set the PVLAN mode of the selected VLAN to primary.
   INTERFACE VLAN mode
   ```
   private-vlan mode primary
   ```
4. Map secondary VLANs to the selected primary VLAN.
   INTERFACE VLAN mode
   ```
   private-vlan mapping secondary-vlan vlan-list
   ```
   The list of secondary VLANs can be:
   - Specified in comma-delimited (*VLAN-ID,VLAN-ID*) or hyphenated-range format (*VLAN-ID-VLAN-ID*).
   - Specified with this command even before they have been created.
   - Amended by specifying the new secondary VLAN to be added to the list.
5. Add promiscuous ports as tagged or untagged interfaces.
   INTERFACE VLAN mode
   `tagged` *interface* or `untagged` *interface*
   Add PVLAN trunk ports to the VLAN only as tagged interfaces.
   You can enter interfaces in numeric or in range format, either comma-delimited (*slot/port,port,port*) or hyphenated (*slot/port-port*).
   You can only add promiscuous ports or PVLAN trunk ports to the PVLAN (no host or regular ports).
6. (OPTIONAL) Assign an IP address to the VLAN.
   INTERFACE VLAN mode
   ```
   ip address ip address
   ```
7. (OPTIONAL) Enable/disable Layer 3 communication between secondary VLANs.
   INTERFACE VLAN mode
   ```
   ip local-proxy-arp
   ```

NOTE: If a promiscuous or host port is untagged in a VLAN and it receives a tagged packet in the same VLAN, the packet is NOT dropped.

Creating a Community VLAN

A community VLAN is a secondary VLAN of the primary VLAN in a private VLAN.

The ports in a community VLAN can talk to each other and with the promiscuous ports in the primary VLAN.

1. Access INTERFACE VLAN mode for the VLAN that you want to make a community VLAN.
   CONFIGURATION mode
   ```
   interface vlan vlan-id
   ```
2. Enable the VLAN.
   INTERFACE VLAN mode
   ```
   no shutdown
   ```
3. Set the PVLAN mode of the selected VLAN to community.
   INTERFACE VLAN mode
   ```
   private-vlan mode community
   ```
4. Add one or more host ports to the VLAN.
   INTERFACE VLAN mode
   `tagged` *interface* or `untagged` *interface*
   You can enter the interfaces singly or in range format, either comma-delimited (*slot/port,port,port*) or hyphenated (*slot/ port-port*).

You can only add host (isolated) ports to the VLAN.

Creating an Isolated VLAN

An isolated VLAN is a secondary VLAN of a primary VLAN.

An isolated VLAN port can only talk with the promiscuous ports in that primary VLAN.

1. Access INTERFACE VLAN mode for the VLAN that you want to make an isolated VLAN.

   CONFIGURATION mode

   `interface vlan vlan-id`
2. Enable the VLAN.

   INTERFACE VLAN mode

   `no shutdown`
3. Set the PVLAN mode of the selected VLAN to isolated.

   INTERFACE VLAN mode

   `private-vlan mode isolated`
4. Add one or more host ports to the VLAN.

   INTERFACE VLAN mode

   `tagged interface` or `untagged interface`

   You can enter the interfaces singly or in range format, either comma-delimited (`slot/port,port,port`) or hyphenated (`slot/ port-port`).

   You can only add ports defined as host to the VLAN.

The following example shows the use of the PVLAN commands that are used in VLAN INTERFACE mode to configure the PVLAN member VLANs (primary, community, and isolated VLANs).

```
DellEMC#conf
DellEMC(conf)# interface vlan 10
DellEMC(conf-vlan-10)# private-vlan mode primary
DellEMC(conf-vlan-10)# private-vlan mapping secondary-vlan 100-101
DellEMC(conf-vlan-10)# untagged Te 2/1
DellEMC(conf-vlan-10)# tagged Te 2/3

DellEMC(conf)# interface vlan 101
DellEMC(conf-vlan-101)# private-vlan mode community
DellEMC(conf-vlan-101)# untagged Te 2/10

DellEMC(conf)# interface vlan 100
DellEMC(conf-vlan-100)# private-vlan mode isolated
DellEMC(conf-vlan-100)# untagged Te 2/2
```

Private VLAN Configuration Example

The following example shows a private VLAN topology.

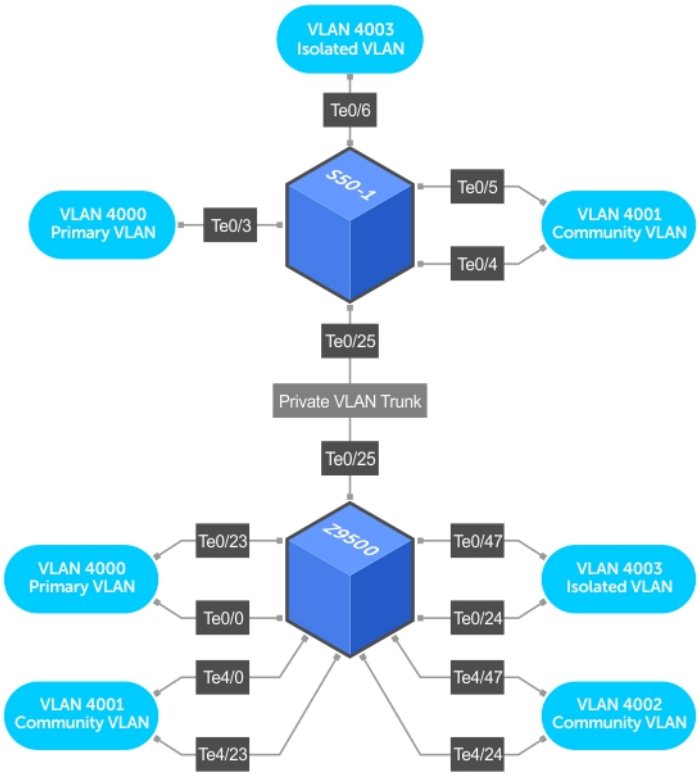


Figure 109. Sample Private VLAN Topology

The following configuration is based on the example diagram for the Z9500:

- Te 1/1 and Te 1/23 are configured as promiscuous ports, assigned to the primary VLAN, VLAN 4000.
- Te 1/25 is configured as a PVLAN trunk port, also assigned to the primary VLAN 4000.
- Te 1/24 and Te 1/47 are configured as host ports and assigned to the isolated VLAN, VLAN 4003.
- Te 4/1 and Te 23 are configured as host ports and assigned to the community VLAN, VLAN 4001.
- Te 4/24 and Te 4/47 are configured as host ports and assigned to community VLAN 4002.

The result is that:

- The ports in community VLAN 4001 can communicate directly with each other and with promiscuous ports.
- The ports in community VLAN 4002 can communicate directly with each other and with promiscuous ports.
- The ports in isolated VLAN 4003 can only communicate with the promiscuous ports in the primary VLAN 4000.
- All the ports in the secondary VLANs (both community and isolated VLANs) can only communicate with ports in the other secondary VLANs of that PVLAN over Layer 3, and only when the `ip local-proxy-arp` command is invoked in the primary VLAN.

NOTE: Even after you disable `ip-local-proxy-arp` (`no ip-local-proxy-arp`) in a secondary VLAN, Layer 3 communication may happen between some secondary VLAN hosts, until the ARP timeout happens on those secondary VLAN hosts.

In parallel, on S4810:

- Te 1/3 is a promiscuous port and Te 1/25 is a PVLAN trunk port, assigned to the primary VLAN 4000.
- Te 1/4-6 are host ports. Te 1/4 and Te 1/5 are assigned to the community VLAN 4001, while Te 1/6 is assigned to the isolated VLAN 4003.

The result is that:

- The S4810 ports would have the same intra-switch communication characteristics as described for the Z9500.
- For transmission between switches, tagged packets originating from host PVLAN ports in one secondary VLAN and destined for host PVLAN ports in the other switch travel through the promiscuous ports in the local VLAN 4000 and then through the trunk ports (1/25 in each switch).

Inspecting the Private VLAN Configuration

The standard methods of inspecting configurations also apply in PVLANs.

To inspect your PVLAN configurations, use the following commands.

- Display the specific interface configuration.

  INTERFACE mode and INTERFACE VLAN mode

  ```
  show config
  ```

- Inspect the running-config, and, with the `grep pipe` option, display a specific part of the running-config.

  ```
  show running-config | grep string
  ```

  The following example shows the PVLAN parts of the running-config from the S50V switch in the topology diagram previously shown.

- Display the type and status of the configured PVLAN interfaces.

  ```
  show interfaces private-vlan [interface interface]
  ```

  This command is specific to the PVLAN feature.

  For more information, refer to the *Security* chapter in the *Dell EMC Networking OS Command Line Reference Guide*.

- Display the configured PVLANs or interfaces that are part of a PVLAN.

  ```
  show vlan private-vlan [community | interface | isolated | primary | primary_vlan | interface interface]
  ```

  This command is specific to the PVLAN feature.

  The following examples show the results of using this command without the command options on the C300 and S50V switches in the topology diagram previously shown.

- Display the primary-secondary VLAN mapping. The following example shows the output from the S50V.

  ```
  show vlan private-vlan mapping
  ```

  This command is specific to the PVLAN feature.

The `show arp` and `show vlan` commands are revised to display PVLAN data.

The following example shows viewing a private VLAN.

```
c300-1#show vlan private-vlan

Primary Secondary Type       Active Ports
------- --------- ---------- ------ --------------
4000              Primary    Yes    Te 1/1,23,25
        4001      Community  Yes    Te 4/1,23
        4002      Community  Yes    Te 4/24,47
        4003      Isolated   Yes    Te 1/24,47
```

The following example shows using the `show vlan private-vlan mapping` command.

```
S50-1#show vlan private-vlan mapping
Private Vlan:
Primary     : 4000
Isolated    : 4003
Community   : 4001
```

NOTE: In the following example, notice the addition of the PVLAN codes – P, I, and C – in the left column.

The following example shows viewing the VLAN status.

```
S50V#show vlan
Codes: * - Default VLAN, G - GVRP VLANs, P - Primary, C - Community, I - Isolated
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
```

```
   G - GVRP tagged, M - Vlan-stack

  NUM  Status    Description               Q Ports
* 1    Inactive
  100  Inactive
P 200  Inactive  primary VLAN in PVLAN     T Te 1/19-20
I 201  Inactive  isolated VLAN in VLAN 200 T Te 1/21
```

The following example shows viewing a private VLAN configuration.

```
!
interface TenGigabitEthernet 1/3
  no ip address
  switchport
  switchport mode private-vlan promiscuous
  no shutdown
!
interface TenGigabitEthernet 1/4
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown
!
interface TenGigabitEthernet 1/5
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown
!
interface TenGigabitEthernet 1/6
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown
!
interface TenGigabitEthernet 1/25
  no ip address
  switchport
  switchport mode private-vlan trunk
  no shutdown
!
interface Vlan 4000
  private-vlan mode primary
private-vlan mapping secondary-vlan 4001-4003
no ip address
tagged TenGigabitEthernet 1/3,25
no shutdown
!
interface Vlan 4001
private-vlan mode community
```

41

Per-VLAN Spanning Tree Plus (PVST+)

Per-VLAN spanning tree plus (PVST+) is a variation of spanning tree — developed by a third party — that allows you to configure a separate spanning tree instance for each virtual local area network (VLAN).

Topics:

- Protocol Overview
- Implementation Information
- Configure Per-VLAN Spanning Tree Plus
- Enabling PVST+
- Disabling PVST+
- Influencing PVST+ Root Selection
- Modifying Global PVST+ Parameters
- Modifying Interface PVST+ Parameters
- Configuring an EdgePort
- PVST+ in Multi-Vendor Networks
- Enabling PVST+ Extend System ID
- PVST+ Sample Configurations

Protocol Overview

PVST+ is a variation of spanning tree — developed by a third party — that allows you to configure a separate spanning tree instance for each virtual local area network (VLAN).

For more information about spanning tree, refer to the Spanning Tree Protocol (STP) chapter.

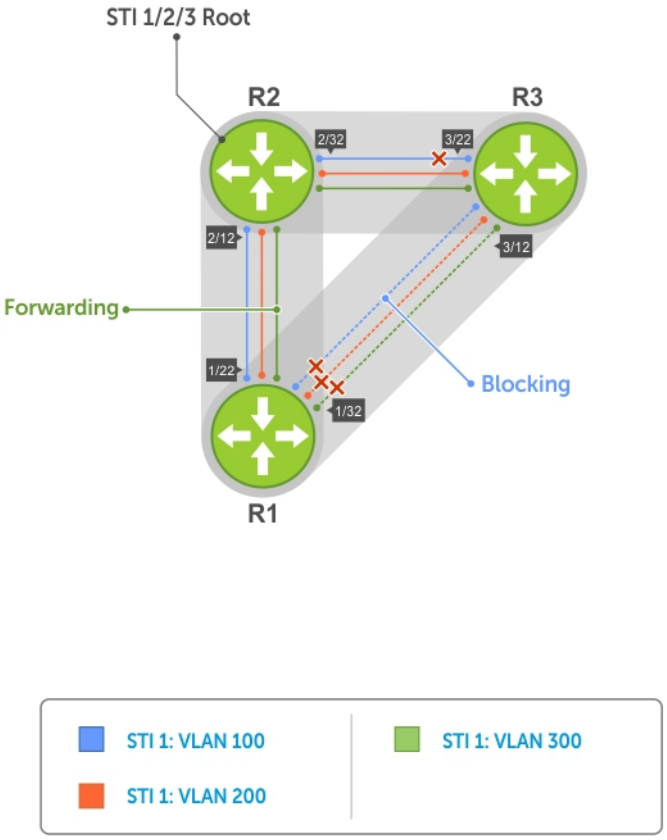


Figure 110. Per-VLAN Spanning Tree

The Dell EMC Networking OS supports three other variations of spanning tree, as shown in the following table.

Table 72. Spanning Tree Variations Dell EMC Networking OS Supports

Dell EMC Networking Term	IEEE Specification
Spanning Tree Protocol (STP)	802 .1d
Rapid Spanning Tree Protocol (RSTP)	802 .1w
Multiple Spanning Tree Protocol (MSTP)	802 .1s
Per-VLAN Spanning Tree Plus (PVST+)	Third Party

Implementation Information

- The Dell EMC Networking OS implementation of PVST+ is based on IEEE Standard 802.1w.
- The Dell EMC Networking OS implementation of PVST+ uses IEEE 802.1s costs as the default costs (as shown in the following table). Other implementations use IEEE 802.1w costs as the default costs. If you are using Dell EMC Networking systems in a multivendor network, verify that the costs are values you intended.
- You can enable PVST+ on 254 VLANs. To set up VLANs, refer to Virtual LANs (VLANs).

NOTE: Dell EMC Networking OS supports upto 64 VLANs or instances and the number of virtual ports (number of VLANs * interface groups) must not exceed 64 in a L2 and L3 scaled environment and 128 in L2 only environment.

Configure Per-VLAN Spanning Tree Plus

Configuring PVST+ is a four-step process.

1. Configure interfaces for Layer 2.
2. Place the interfaces in VLANs.
3. Enable PVST+.
4. Optionally, for load balancing, select a nondefault bridge-priority for a VLAN.

Related Configuration Tasks

- Modifying Global PVST+ Parameters
- Modifying Interface PVST+ Parameters
- Configuring an EdgePort
- Flush MAC Addresses after a Topology Change
- Prevent Network Disruptions with BPDU Guard
- Enabling SNMP Traps for Root Elections and Topology Changes
- Configuring Spanning Trees as Hitless
- PVST+ in Multi-Vendor Networks
- Enabling PVST+ Extended System ID
- PVST+ Sample Configurations

Enabling PVST+

When you enable PVST+, Dell EMC Networking OS instantiates STP on each active VLAN.

1. Enter PVST context.
   PROTOCOL PVST mode
   `protocol spanning-tree pvst`
2. Enable PVST+.
   PROTOCOL PVST mode
   `no disable`

Disabling PVST+

To disable PVST+ globally or on an interface, use the following commands.

- Disable PVST+ globally.
  PROTOCOL PVST mode
  `disable`
- Disable PVST+ on an interface, or remove a PVST+ parameter configuration.
  INTERFACE mode
  `no spanning-tree pvst`

To display your PVST+ configuration, use the `show config` command from PROTOCOL PVST mode.

```
Dell_E600(conf-pvst)#show config verbose
!
protocol spanning-tree pvst
  no disable
  vlan 100 bridge-priority 4096
```

Influencing PVST+ Root Selection

As shown in the previous per-VLAN spanning tree illustration, all VLANs use the same forwarding topology because R2 is elected the root, and all TenGigabitEthernet ports have the same cost.

The following per-VLAN spanning tree illustration changes the bridge priority of each bridge so that a different forwarding topology is generated for each VLAN. This behavior demonstrates how you can use PVST+ to achieve load balancing.

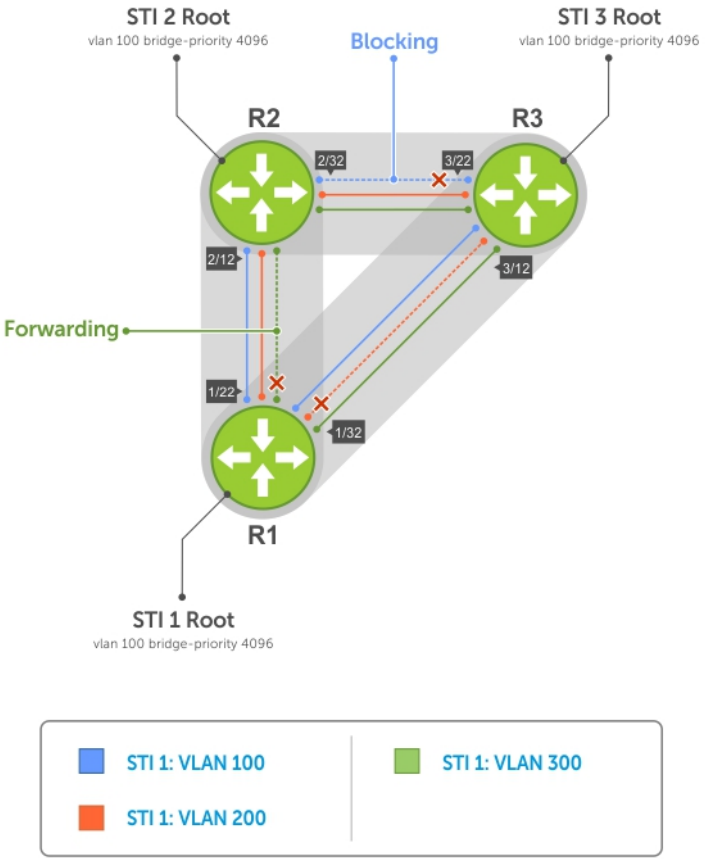


Figure 111. Load Balancing with PVST+

The bridge with the bridge value for bridge priority is elected root. Because all bridges use the default priority (until configured otherwise), the lowest MAC address is used as a tie-breaker. To increase the likelihood that a bridge is selected as the STP root, assign bridges a low non-default value for bridge priority.

To assign a bridge priority, use the following command.

- Assign a bridge priority.

  PROTOCOL PVST mode

  `vlan bridge-priority`

  The range is from 0 to 61440.

  The default is **32768**.

To display the PVST+ forwarding topology, use the `show spanning-tree pvst [vlan vlan-id]` command from EXEC Privilege mode.

```
Dell_E600(conf)#do show spanning-tree pvst vlan 100
VLAN 100
Root Identifier has priority 4096, Address 0001.e80d.b6d6
Root Bridge hello time 2, max age 20, forward delay 15
Bridge Identifier has priority 4096, Address 0001.e80d.b6d6
Configured hello time 2, max age 20, forward delay 15
We are the root of VLAN 100
```

```
Current root has priority 4096, Address 0001.e80d.b6d6
Number of topology changes 5, last change occurred 00:34:37 ago on Te 1/32

Port 375 (TenGigabitEthernet 1/22) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.375
Designated root has priority 4096, address 0001.e80d.b6:d6
Designated bridge has priority 4096, address 0001.e80d.b6:d6
Designated port id is 128.375 , designated path cost 0
Number of transitions to forwarding state 2
BPDU sent 1159, received 632
The port is not in the Edge port mode

Port 385 (TenGigabitEthernet 1/32) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.385
Designated root has priority 4096, address 0001.e80d.b6:d6
Designated bridge has priority 4096, address 0001.e80d.b6:d6
Designated port id is 128.385 , designated path cost 0
```

Modifying Global PVST+ Parameters

The root bridge sets the values for forward-delay and hello-time, and overwrites the values set on other PVST+ bridges.

- **Forward-delay** — the amount of time an interface waits in the Listening state and the Learning state before it transitions to the Forwarding state.
- **Hello-time** — the time interval in which the bridge sends bridge protocol data units (BPDUs).
- **Max-age** — the length of time the bridge maintains configuration information before it refreshes that information by recomputing the PVST+ topology.

To change PVST+ parameters on the root bridge, use the following commands.

- Change the forward-delay parameter.

  PROTOCOL PVST mode

  `vlan forward-delay`

  The range is from 4 to 30.

  The default is **15 seconds**.
- Change the hello-time parameter.

  PROTOCOL PVST mode

  `vlan hello-time`

  NOTE: With large configurations (especially those configurations with more ports), Dell EMC Networking recommends increasing the hello-time.

  The range is from 1 to 10.

  The default is **2 seconds**.
- Change the max-age parameter.

  PROTOCOL PVST mode

  `vlan max-age`

  The range is from 6 to 40.

  The default is **20 seconds**.

The values for global PVST+ parameters are given in the output of the `show spanning-tree pvst` command.

Modifying Interface PVST+ Parameters

You can adjust two interface parameters (port cost and port priority) to increase or decrease the probability that a port becomes a forwarding port.

- **Port cost** — a value that is based on the interface type. The greater the port cost, the less likely the port is selected to be a forwarding port.

- **Port priority** — influences the likelihood that a port is selected to be a forwarding port in case that several ports have the same port cost.

The following tables lists the default values for port cost by interface.

Table 73. Default Values for Port Cost

Port Cost	Default Value
100-Mb/s Ethernet interfaces	200000
1-Gigabit Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
25-Gigabit Ethernet interfaces	1700
40-Gigabit Ethernet interfaces	1400
50-Gigabit Ethernet interfaces	1200
100-Gigabit Ethernet interfaces	200
Port Channel with 100 Mb/s Ethernet interfaces	180000
Port Channel with 1-Gigabit Ethernet interfaces	18000
Port Channel with 10-Gigabit Ethernet interfaces	1800
Port Channel with 25-Gigabit Ethernet interfaces	1200
Port Channel with 50-Gigabit Ethernet interfaces	200
Port Channel with 100-Gigabit Ethernet interfaces	180

NOTE: The Dell EMC Networking OS implementation of PVST+ uses IEEE 802.1s costs as the default costs. Other implementations use IEEE 802.1w costs as the default costs. If you are using Dell EMC Networking systems in a multi-vendor network, verify that the costs are values you intended.

To change the port cost or port priority of an interface, use the following commands.

- Change the port cost of an interface.
  INTERFACE mode
  `spanning-tree pvst vlan cost.`
  The range is from 0 to 200000.
  Refer to the table for the default values.
- Change the port priority of an interface.
  INTERFACE mode
  `spanning-tree pvst vlan priority.`
  The range is from 0 to 240, in increments of 16.
  The default is **128**.

The values for interface PVST+ parameters are given in the output of the `show spanning-tree pvst` command, as previously shown.

Configuring an EdgePort

The EdgePort feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner.

In this mode an interface forwards frames by default until it receives a BPDU that indicates that it should behave otherwise; it does not go through the Learning and Listening states. The `bpduguard shutdown-on-violation` option causes the interface hardware to be shut down when it receives a BPDU. When you only implement `bpduguard`, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree drops packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation.

This feature is the same as PortFast mode in spanning tree.

CAUTION: Configure EdgePort only on links connecting to an end station. EdgePort can cause loops if you enable it on an interface connected to a network.

To enable EdgePort on an interface, use the following command.

- Enable EdgePort on an interface.

  INTERFACE mode

  ```
  spanning-tree pvst edge-port [bpduguard | shutdown-on-violation]
  ```

The EdgePort status of each interface is given in the output of the `show spanning-tree pvst` command, as previously shown.

Dell EMC Networking OS Behavior: Regarding the `bpduguard shutdown-on-violation` command behavior:

- If the interface to be shut down is a port channel, all the member ports are disabled in the hardware.
- When you add a physical port to a port channel already in an Error Disable state, the new member port is also disabled in the hardware.
- When you remove a physical port from a port channel in an Error Disable state, the Error Disabled state is cleared on this physical port (the physical port is enabled in the hardware).
- You can clear the Error Disabled state with any of the following methods:
  - Perform a `shutdown` command on the interface.
  - Disable the `shutdown-on-violation` command on the interface (the `no spanning-tree stp-id portfast [bpduguard | [shutdown-on-violation]]` command).
  - Disable spanning tree on the interface (the `no spanning-tree` command in INTERFACE mode).
  - Disabling global spanning tree (the `no spanning-tree` command in CONFIGURATION mode).

PVST+ in Multi-Vendor Networks

Some non-Dell EMC Networking systems which have hybrid ports participating in PVST+ transmit two kinds of BPDUs: an 802.1D BPDU and an untagged PVST+ BPDU.

Dell EMC Networking systems do not expect PVST+ BPDU (tagged or untagged) on an untagged port. If this situation occurs, Dell EMC Networking OS places the port in an Error-Disable state. This behavior might result in the network not converging. To prevent Dell EMC Networking OS from executing this action, use the `no spanning-tree pvst err-disable cause invalid-pvst-bpdu` command. After you configure this command, if the port receives a PVST+ BPDU, the BPDU is dropped and the port remains operational.

Enabling PVST+ Extend System ID

In the following example, ports P1 and P2 are untagged members of different VLANs. These ports are untagged because the hub is VLAN unaware. There is no data loop in this scenario; however, you can employ PVST+ to avoid potential misconfigurations.

If you enable PVST+ on the Dell EMC Networking switch in this network, P1 and P2 receive BPDUs from each other. Ordinarily, the Bridge ID in the frame matches the Root ID, a loop is detected, and the rules of convergence require that P2 move to blocking state because it has the lowest port ID.

To keep both ports in a Forwarding state, use extend system ID. Extend system ID augments the bridge ID with a VLAN ID to differentiate BPDUs on each VLAN so that PVST+ does not detect a loop and both ports can remain in a Forwarding state.

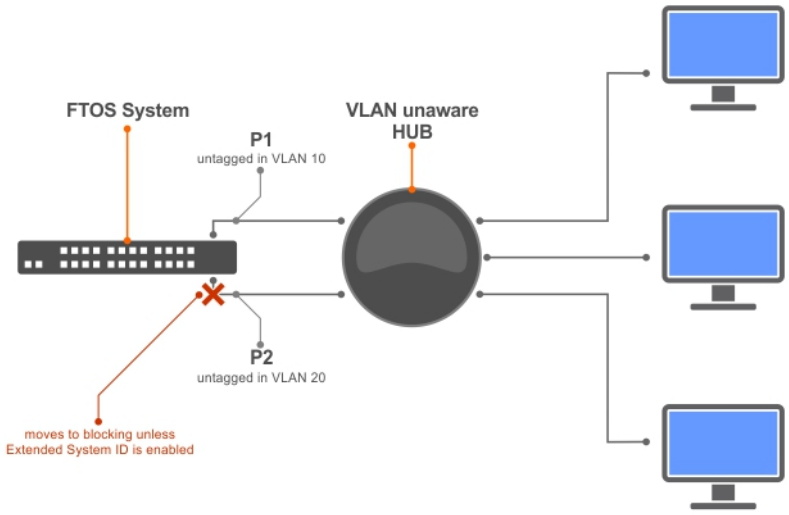


Figure 112. PVST+ with Extend System ID

- Augment the bridge ID with the VLAN ID.
  PROTOCOL PVST mode
  `extend system-id`

```
DellEMC(conf-pvst)#do show spanning-tree pvst vlan 5 brief

VLAN 5
Executing IEEE compatible Spanning Tree Protocol
Root ID Priority 32773, Address 0001.e832.73f7
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID Priority 32773 (priority 32768 sys-id-ext 5), Address 0001.e832.73f7
We are the root of Vlan 5
Configured hello time 2, max age 20, forward delay 15
```

PVST+ Sample Configurations

The following examples provide the running configurations for the topology shown in the previous illustration.

Example of PVST+ Configuration (R1)

```
interface TenGigabitEthernet 1/22
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 1/32
  no ip address
  switchport
  no shutdown
!
protocol spanning-tree pvst
  no disable
  vlan 100 bridge-priority 4096
interface Vlan 100
  no ip address
  tagged TenGigabitEthernet 1/22,32
  no shutdown
!
interface Vlan 200
  no ip address
  tagged TenGigabitEthernet 1/22,32
```

```
  no shutdown
!
interface Vlan 300
  no ip address
  tagged TenGigabitEthernet 1/22,32
  no shutdown
!
protocol spanning-tree pvst
  no disable
  vlan 100 bridge-priority 4096
```

Example of PVST+ Configuration (R2)

```
interface TenGigabitEthernet 2/12
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 2/32
  no ip address
  switchport
  no shutdown
!
interface Vlan 100
  no ip address
  tagged TenGigabitEthernet 2/12,32
  no shutdown
!
interface Vlan 200
  no ip address
  tagged TenGigabitEthernet 2/12,32
  no shutdown
!
interface Vlan 300
  no ip address
  tagged TenGigabitEthernet 2/12,32
  no shutdown
!
protocol spanning-tree pvst
  no disable
  vlan 200 bridge-priority 4096
```

Example of PVST+ Configuration (R3)

```
interface TenGigabitEthernet 3/12
  no ip address
  switchport
  no shutdown
!
interface TenGigabitEthernet 3/22
  no ip address
  switchport
  no shutdown
!
interface Vlan 100
  no ip address
  tagged TenGigabitEthernet 3/12,22
  no shutdown
!
interface Vlan 200
  no ip address
  tagged TenGigabitEthernet 3/12,22
  no shutdown
!
interface Vlan 300
  no ip address
  tagged TenGigabitEthernet 3/12,22
  no shutdown
!
protocol spanning-tree pvst
  no disable
  vlan 300 bridge-priority 4096
```

42

Quality of Service (QoS)

This chapter describes how to use and configure Quality of Service service (QoS) features on the switch.

Differentiated service is accomplished by classifying and queuing traffic, and assigning priorities to those queues.

Table 74. Dell EMC Networking Operating System (OS) Support for Port-Based, Policy-Based Features

Feature	Direction
Port-Based QoS Configurations	Ingress + Egress
Set dot1p Priorities for Incoming Traffic	Ingress
Honor dot1p Priorities on Ingress Traffic	Ingress
Configure Port-based Rate Policing	Ingress
Configure Port-based Rate Shaping	Egress
Policy-Based QoS Configurations	Ingress + Egress
Classify Traffic	Ingress
Create a Layer 3 Class Map	Ingress
Set DSCP Values for Egress Packets Based on Flow	Ingress
Create a Layer 2 Class Map	Ingress
Create a QoS Policy	Ingress + Egress
Create an Input QoS Policy	Ingress
Configure Policy-Based Rate Policing	Ingress
Set a DSCP Value for Egress Packets	Ingress
Set a dot1p Value for Egress Packets	Ingress
Create an Output QoS Policy	Egress
Configure Policy-Based Rate Shaping	Egress
Allocate Bandwidth to the Queue	Egress
Configure a Scheduler to Queue	Egress
Specify WRED Drop Precedence	Egress
Create Policy Maps	Ingress + Egress
Create Input Policy Maps	Ingress
Honor DSCP Values on Ingress Packets	Ingress
Honoring dot1p Values on Ingress Packets	Ingress

Table 74. Dell EMC Networking Operating System (OS) Support for Port-Based, Policy-Based Features (continued)

Feature	Direction
Create Output Policy Maps	Egress
Specify an Aggregate QoS Policy	Egress
Create Output Policy Maps	Egress
Enabling QoS Rate Adjustment Enabling Strict-Priority Queueing Weighted Random Early Detection	Egress
Create WRED Profiles	Egress

Figure 113. Dell EMC Networking QoS Architecture

Topics:

- Implementation Information
- Port-Based QoS Configurations
- Policy-Based QoS Configurations
- Enabling QoS Rate Adjustment
- Enabling Strict-Priority Queueing
- Weighted Random Early Detection
- Pre-Calculating Available QoS CAM Space
- Configuring Weights and ECN for WRED
- Configuring WRED and ECN Attributes
- Guidelines for Configuring ECN for Classifying and Color-Marking Packets

- Enabling Buffer Statistics Tracking

Implementation Information

The Dell EMC Networking QoS implementation complies with IEEE 802.1p *User Priority Bits for QoS Indication*.

It also implements these Internet Engineering Task Force (IETF) documents:

- RFC 2474, *Definition of the Differentiated Services Field (DS Field) in the IPv4 Headers*
- RFC 2475, *An Architecture for Differentiated Services*
- RFC 2597, *Assured Forwarding PHB Group*
- RFC 2598, *An Expedited Forwarding PHB*

You cannot configure port-based and policy-based QoS on the same interface.

Port-Based QoS Configurations

You can configure the following QoS features on an interface.

NOTE: You cannot simultaneously use egress rate shaping and ingress rate policing on the same virtual local area network (VLAN).

- Setting dot1p Priorities for Incoming Traffic
- Honoring dot1p Priorities on Ingress Traffic
- Configuring Port-Based Rate Policing
- Configuring Port-Based Rate Shaping

Setting dot1p Priorities for Incoming Traffic

Dell EMC Networking OS places traffic marked with a priority in a queue based on the following table.

If you set a dot1p priority for a port-channel, all port-channel members are configured with the same value. You cannot assign a dot1p value to an individual interface in a port-channel.

Table 75. dot1p-priority Values and Queue Numbers

dot1p	Queue Number
0	0
1	0
2	0
3	1
4	2
5	3
6	3
7	3

- Change the priority of incoming traffic on the interface.

  `dot1p-priority`

```
DellEMC#configure terminal
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#switchport
DellEMC(conf-if-te-1/1)#dot1p-priority 1
DellEMC(conf-if-te-1/1)#end
```

Honoring dot1p Priorities on Ingress Traffic

By default, Dell EMC Networking OS does not honor dot1p priorities on ingress traffic.

You can configure this feature on physical interfaces and port-channels, but you cannot configure it on individual interfaces in a port channel.

You can configure service-class dynamic dot1p from CONFIGURATION mode, which applies the configuration to all interfaces. A CONFIGURATION mode service-class dynamic dot1p entry supersedes any INTERFACE entries. For more information, refer to Mapping dot1p Values to Service Queues.

NOTE: You cannot configure `service-policy input` and `service-class dynamic dot1p` on the same interface.

- Honor dot1p priorities on ingress traffic.

  INTERFACE mode

  `service-class dynamic dot1p`

```
DellEMC#configure terminal
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#service-class dynamic dot1p
DellEMC(conf-if-te-1/1)#end
```

Priority-Tagged Frames on the Default VLAN

Priority-tagged frames are 802.1Q tagged frames with VLAN ID 0. For VLAN classification, these packets are treated as untagged. However, the dot1p value is still honored when you configure `service-class dynamic dot1p` or `trust dot1p`.

When priority-tagged frames ingress an untagged port or hybrid port, the frames are classified to the default VLAN of the port and to a queue according to their dot1p priority if you configure `service-class dynamic dotp` or `trust dot1p`. When priority-tagged frames ingress a tagged port, the frames are dropped because, for a tagged port, the default VLAN is 0.

Dell EMC Networking OS Behavior: Hybrid ports can receive untagged, tagged, and priority tagged frames. The rate metering calculation might be inaccurate for untagged ports because an internal assumption is made that all frames are treated as tagged. Internally, the ASIC adds a 4-bytes tag to received untagged frames. Though these 4-bytes are not part of the untagged frame received on the wire, they are included in the rate metering calculation resulting in metering inaccuracy.

Configuring Port-Based Rate Policing

If the interface is a member of a VLAN, you may specify the VLAN for which ingress packets are policed.

- Rate policing ingress traffic on an interface.

  INTERFACE mode

  `rate police`

The following example shows configuring rate policing.

```
DellEMC#configure terminal
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#rate police 100 40 peak 150 50
DellEMC(conf-if-te-1/1)#end
```

```
DellEMC#show interfaces tengigabitEthernet 1/2 rate police
  Rate police 300 (50) peak 800 (50)
    Traffic Monitor 0: normal 300 (50) peak 800 (50)
      Out of profile yellow 23386960 red 320605113
    Traffic Monitor 1: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 2: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 3: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 4: normal NA peak NA
      Out of profile yellow 0 red 0
```

Configuring Port-Based Rate Shaping

Rate shaping buffers, rather than drops, traffic exceeding the specified rate until the buffer is exhausted. If any stream exceeds the configured bandwidth on a continuous basis, it can consume all of the buffer space that is allocated to the port.

Dell EMC Networking OS Behavior: Rate shaping is effectively rate limiting because of its smaller buffer size. Rate shaping on tagged ports is slightly greater than the configured rate and rate shaping on untagged ports is slightly less than configured rate.

Rate shaping buffers, rather than drops, traffic exceeding the specified rate until the buffer is exhausted. If any stream exceeds the configured bandwidth on a continuous basis, it can consume all of the buffer space that is allocated to the port.

- Apply rate shaping to outgoing traffic on a port.

  INTERFACE mode

  ```
  rate shape
  ```

- Apply rate shaping to a queue.

  QoS Policy mode

  ```
  rate-shape
  ```

```
DellEMC#configure terminal
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#rate shape 500 50
DellEMC(conf-if-te-1/1)#end
```

Policy-Based QoS Configurations

Policy-based QoS configurations consist of the components shown in the following example.

Figure 114. Constructing Policy-Based QoS Configurations

Classify Traffic

Class maps differentiate traffic so that you can apply separate quality of service policies to different types of traffic.

For both class maps, Layer 2 and Layer 3, Dell EMC Networking OS matches packets against match criteria in the order that you configure them.

Creating a Layer 3 Class Map

A Layer 3 class map differentiates ingress packets based on the DSCP value or IP precedence, and characteristics defined in an IP ACL. You can also use VLAN IDs and VRF IDs to classify the traffic using layer 3 class-maps.

You may specify more than one DSCP and IP precedence value, but only one value must match to trigger a positive match for the class map.

NOTE: IPv6 and IP-any class maps cannot match on ACLs or VLANs.

Use step 1 or step 2 to start creating a Layer 3 class map.

1. Create a match-any class map.
   CONFIGURATION mode

```
class-map match-any
```

2. Create a match-all class map.
   CONFIGURATION mode
   ```
   class-map match-all
   ```
3. Specify your match criteria.
   CLASS MAP mode
   ```
   [seq sequence number] match {ip | ipv6 | ip-any}
   ```
   After you create a class-map, Dell EMC Networking OS places you in CLASS MAP mode.
   Match-any class maps allow up to five ACLs. Match-all class-maps allow only one ACL.

   NOTE: Within a `class-map`, the match rules are installed in the sequence number order.

4. Link the class-map to a queue.
   POLICY MAP mode
   ```
   service-queue
   ```

```
DellEMC(conf)#ip access-list standard acl1
DellEMC(config-std-nacl)#permit 20.0.0.0/8
DellEMC(config-std-nacl)#exit
DellEMC(conf)#ip access-list standard acl2
DellEMC(config-std-nacl)#permit 20.1.1.0/24 order 0
DellEMC(config-std-nacl)#exit
DellEMC(conf)#class-map match-all cmap1
DellEMC(conf-class-map)#match ip access-group acl1
DellEMC(conf-class-map)#exit
DellEMC(conf)#class-map match-all cmap2
DellEMC(conf-class-map)#match ip access-group acl2
DellEMC(conf-class-map)#exit
DellEMC(conf)#policy-map-input pmap
DellEMC(conf-policy-map-in)#service-queue 3 class-map cmap1
DellEMC(conf-policy-map-in)#service-queue 1 class-map cmap2
DellEMC(conf-policy-map-in)#exit
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#service-policy input pmap
```

Examples of Creating a Layer 3 IPv6 Class Map

The following example matches the IPv6 traffic with a DSCP value of 40:

```
DellEMC(conf)# class-map match-all test
DellEMC(conf-class-map)# match ipv6 dscp 40
```

The following example matches the IPv4 and IPv6 traffic with a precedence value of 3:

```
DellEMC(conf)# class-map match-any test1
DellEMC(conf-class-map)#match ip-any precedence 3
```

Creating a Layer 2 Class Map

All class maps are Layer 3 by default; however, you can create a Layer 2 class map by specifying the `layer2` option with the `class-map` command.

A Layer 2 class map differentiates traffic according to 802.1p value and/or VLAN and/or characteristics defined in a MAC ACL..

Use Step 1 or Step 2 to start creating a Layer 2 class map.

1. Create a match-any class map.
   CONFIGURATION mode
   ```
   class-map match-any
   ```
2. Create a match-all class map.
   CONFIGURATION mode
   ```
   class-map match-all
   ```

3. Specify your match criteria.

   CLASS MAP mode

   `[seq sequence number] match mac`

   After you create a class-map, Dell EMC Networking OS places you in CLASS MAP mode.

   Match-any class maps allow up to five access-lists. Match-all class-maps allow only one. You can match against only one VLAN ID.

4. Link the class-map to a queue.

   POLICY MAP mode

   `service-queue`

Determining the Order in Which ACLs are Used to Classify Traffic

When you link class-maps to queues using the `service-queue` command, Dell EMC Networking OS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities).

For example, as described in the previous example, class-map *cmap2* is matched against ingress packets before *cmap1*.

ACLs *acl1* and *acl2* have overlapping rules because the address range 20.1.1.0/24 is within 20.0.0.0/8. Therefore (without the keyword `order`), packets within the range 20.1.1.0/24 match positive against *cmap1* and are buffered in queue 7, though you intended for these packets to match positive against *cmap2* and be buffered in queue 4.

In cases such as these, where class-maps with overlapping ACL rules are applied to different queues, use the keyword `order`. Dell EMC Networking OS writes to the CAM ACL rules with lower order numbers (order numbers closer to 0) before rules with higher order numbers so that packets are matched as you intended.

- Specify the order in which you want to apply ACL rules using the keyword `order`.

  `order`

  The order can range from 0 to 254.

  By default, all ACL rules have an order of **255**.

Displaying Configured Class Maps and Match Criteria

To display all class-maps or a specific class map, use the following command.

Dell EMC Networking OS Behavior: An explicit “deny any" rule in a Layer 3 ACL used in a (match any or match all) class-map creates a "default to Queue 0" entry in the CAM, which causes unintended traffic classification. In the following example, traffic is classified in two Queues, 1 and 2. Class-map ClassAF1 is “match any,” and ClassAF2 is “match all”.

- Display all class-maps or a specific class map.

  EXEC Privilege mode

  `show qos class-map`

The following example shows incorrect traffic classifications.

```
DellEMC#show running-config policy-map-input
!
policy-map-input PolicyMapIn
  service-queue 1 class-map ClassAF1 qos-policy QosPolicyIn-1
  service-queue 2 class-map ClassAF2 qos-policy QosPolicyIn-2
DellEMC#show running-config class-map
!
class-map match-any ClassAF1
  match ip access-group AF1-FB1 set-ip-dscp 10
  match ip access-group AF1-FB2 set-ip-dscp 12
  match ip dscp 10 set-ip-dscp 14
  match ipv6 dscp 20 set-ip-dscp 14
!
class-map match-all ClassAF2
  match ip access-group AF2
  match ip dscp 18

DellEMC#show running-config ACL
!
ip access-list extended AF1-FB1
```

```
  seq 5 permit ip host 23.64.0.2 any
  seq 10 deny ip any any
!
ip access-list extended AF1-FB2
  seq 5 permit ip host 23.64.0.3 any
  seq 10 deny ip any any
!
ip access-list extended AF2
  seq 5 permit ip host 23.64.0.5 any
  seq 10 deny ip any any

DellEMC# show cam layer3-qos interface tengigabitethernet 2/4
Cam   Port Dscp Proto Tcp  Src  Dst SrcIp        DstIp     DSCP   Queue
Index                 Flag Port Port                       Marking
-------------------------------------------------------------------
20416 1    18    IP    0x0  0    0   23.64.0.5/32 0.0.0.0/0 20     2
20417 1    18    IP    0x0  0    0   0.0.0.0/0    0.0.0.0/0 -      0
20418 1    0     IP    0x0  0    0   23.64.0.2/32 0.0.0.0/0 10     1
20419 1    0     IP    0x0  0    0   0.0.0.0/0    0.0.0.0/0 -      0
20420 1    0     IP    0x0  0    0   23.64.0.3/32 0.0.0.0/0 12     1
20421 1    0     IP    0x0  0    0   0.0.0.0/0    0.0.0.0/0 -      0
20422 1    10    0     0x0  0    0   0.0.0.0/0    0.0.0.0/0 14     1
24511 1    0     0     0x0  0    0   0.0.0.0/0    0.0.0.0/0 -      0
```

In the previous example, the ClassAF1 does not classify traffic as intended. Traffic matching the first match criteria is classified to Queue 1, but all other traffic is classified to Queue 0 as a result of CAM entry 20419.

When you remove the explicit “deny any” rule from all three ACLs, the CAM reflects exactly the desired classification.

The following example shows correct traffic classifications.

```
DellEMC#show cam layer3-qos interface tengigabitethernet 2/4
Cam   Port Dscp Proto Tcp  Src  Dst SrcIp        DstIp     DSCP    Queue
Index                 Flag Port Port                       Marking
-------------------------------------------------------------------------
20416 1    18    IP    0x0  0    0   23.64.0.5/32 0.0.0.0/0 20      2
20417 1    0     IP    0x0  0    0   23.64.0.2/32 0.0.0.0/0 10      1
20418 1    0     IP    0x0  0    0   23.64.0.3/32 0.0.0.0/0 12      1
20419 1    10    0     0x0  0    0   0.0.0.0/0    0.0.0.0/0 14      1
24511 1    0     0     0x0  0    0   0.0.0.0/0    0.0.0.0/0  -      0
```

Create a QoS Policy

There are two types of QoS policies — input and output.

Input QoS policies regulate Layer 3 and Layer 2 ingress traffic. The regulation mechanisms for input QoS policies are rate policing and setting priority values.

- **Layer 3** — QoS input policies allow you to rate police and set a DSCP or dot1p value. In addition, you can configure a drop precedence for incoming packets based on their DSCP value by using a DSCP color map. For more information, see DSCP Color Maps.
- **Layer 2** — QoS input policies allow you to rate police and set a dot1p value.

Output QoS policies regulate egress traffic. The regulation mechanisms for output QoS policies are bandwidth percentage, scheduler strict, rate shaping and WRED.

NOTE: When changing a "service-queue" configuration in a QoS policy map, all QoS rules are deleted and re-added automatically to ensure that the order of the rules is maintained. As a result, the Matched Packets value shown in the `show qos statistics` command is reset.

NOTE: To avoid issues misconfiguration causes, Dell EMC Networking recommends configuring either DCBX or Egress QoS features, but not both simultaneously. If you enable both DCBX and Egress QoS at the same time, the DCBX configuration is applied and unexpected behavior occurs on the Egress QoS.

Creating an Input QoS Policy

To create an input QoS policy, use the following steps.

1. Create a Layer 3 input QoS policy.
   CONFIGURATION mode
   `qos-policy-input`
   Create a Layer 2 input QoS policy by specifying the keyword `layer2` after the `qos-policy-input` command.
2. After you create an input QoS policy, do one or more of the following:
   Configuring Policy-Based Rate Policing
   Setting a dot1p Value for Egress Packets

Configuring Policy-Based Rate Policing

To configure policy-based rate policing, use the following command.

- Configure rate police ingress traffic.
  QOS-POLICY-IN mode
  `rate-police`

Setting a dot1p Value for Egress Packets

To set a dot1p value for egress packets, use the following command.

- Set a dscp or dot1p value for egress packets.
  QOS-POLICY-IN mode
  `set mac-dot1p`

Creating an Output QoS Policy

To create an output QoS policy, use the following commands.

1. Create an output QoS policy.
   CONFIGURATION mode
   `qos-policy-output`
2. After you configure an output QoS policy, do one or more of the following:
   Scheduler Strict — Policy-based Strict-priority Queueing configuration is done through scheduler strict. It is applied to Qos-policy-output. When scheduler strict is applied to multiple Queues, high queue number takes precedence.
   Allocating Bandwidth to Queue
   Specifying WRED Drop Precedence

Configuring Policy-Based Rate Shaping

To configure policy-based rate shaping, use the following command.

- Configure rate shape egress traffic.
  QOS-POLICY-OUT mode
  `rate-shape`

Allocating Bandwidth to Queue

The switch schedules packets for egress based on Deficit Round Robin (DRR). This strategy offers a guaranteed data rate.

Allocate bandwidth to queues only in terms of percentage in 4-queue and 8-queue systems. The following table shows the default bandwidth percentage for each queue.

The following table lists the default bandwidth weights for each queue, and their equivalent percentage which is derived by dividing the bandwidth weight by the sum of all queue weights.

Table 76. Default Bandwidth Weights

Queue	Default Bandwidth Percentage for 4–Queue System	Default Bandwidth Percentage for 8–Queue System
0	6.67%	1%
1	13.33%	2%
2	26.67%	3%
3	53.33%	4%
4	-	5%
5	-	10%
6	-	25%
7	-	50%

NOTE: The system supports 4 data queues.

When you assign a percentage to one queue, note that this change also affects the amount of bandwidth that is allocated to other queues. Therefore, whenever you are allocating bandwidth to one queue, Dell EMC Networking recommends evaluating your bandwidth requirements for all other queues as well.

- Assign each queue a bandwidth percentage ranging from 1 to 100%, in increments of 1%.
  `bandwidth-percentage`

Specifying WRED Drop Precedence

You can configure the WRED drop precedence in an output QoS policy.

- Specify a WRED profile to yellow and/or green traffic.
  QOS-POLICY-OUT mode
  `wred`

For more information, refer to Applying a WRED Profile to Traffic.

Create Policy Maps

There are two types of policy maps: input and output.

Creating Input Policy Maps

There are two types of input policy-maps: Layer 3 and Layer 2.

1. Create a Layer 3 input policy map.
   CONFIGURATION mode
   `policy-map-input`
   Create a Layer 2 input policy map by specifying the keyword `layer2` with the `policy-map-input` command.
2. After you create an input policy map, do one or more of the following:
   Applying a Class-Map or Input QoS Policy to a Queue
   Applying an Input QoS Policy to an Input Policy Map
   Honoring DSCP Values on Ingress Packets
   Honoring dot1p Values on Ingress Packets
3. Apply the input policy map to an interface.

Applying a Class-Map or Input QoS Policy to a Queue

To apply a class-map or input QoS policy to a queue, use the following command.

- Assign an input QoS policy to a queue.

  POLICY-MAP-IN mode

  ```
  service-queue
  ```

Applying an Input QoS Policy to an Input Policy Map

To apply an input QoS policy to an input policy map, use the following command.

- Apply an input QoS policy to an input policy map.

  POLICY-MAP-IN mode

  ```
  policy-service-queue <number> qos-polcy
  ```

Honoring DSCP Values on Ingress Packets

Dell EMC Networking OS provides the ability to honor DSCP values on ingress packets using Trust DSCP feature.

The following table lists the standard DSCP definitions and indicates to which queues Dell EMC Networking OS maps DSCP values. When you configure trust DSCP, the matched packets and matched bytes counters are not incremented in the `show qos` statistics.

Table 77. Default DSCP to Queue Mapping

DSCP/CP hex range (XXX)xxx	DSCP Definition	Traditional IP Precedence	Internal Queue ID	DSCP/CP decimal
111XXX		Network Control	3	48–63
110XXX		Internetwork Control	3	48–63
101XXX	EF (Expedited Forwarding)	CRITIC/ECP	2	32–47
100XXX	AF4 (Assured Forwarding)	Flash Override	2	32–47
011XXX	AF3	Flash	1	16–31
010XXX	AF2	Immediate	1	16–31
001XXX	AF1	Priority	0	0–15
000XXX	BE (Best Effort)	Best Effort	0	0–15

- Enable the trust DSCP feature.

  POLICY-MAP-IN mode

  ```
  trust diffserv
  ```

Honoring dot1p Values on Ingress Packets

Dell EMC Networking OS honors dot1p values on ingress packets with the Trust dot1p feature.

The following table specifies the queue to which the classified traffic is sent based on the dot1p value.

Table 78. Default dot1p to Queue Mapping

dot1p	Queue ID
0	0
1	0
2	0

Table 78. Default dot1p to Queue Mapping (continued)

dot1p	Queue ID
3	1
4	2
5	3
6	3
7	3

The dot1p value is also honored for frames on the default VLAN. For more information, refer to Priority-Tagged Frames on the Default VLAN.

- Enable the trust dot1p feature.

  POLICY-MAP-IN mode

  ```
  trust dot1p
  ```

Mapping dot1p Values to Service Queues

All traffic is by default mapped to the same queue, Queue 0.

If you honor dot1p on ingress, you can create service classes based the queueing strategy in Honoring dot1p Values on Ingress Packets. You may apply this queuing strategy globally by entering the following command from CONFIGURATION mode.

- All dot1p traffic is mapped to Queue 0 unless you enable `service-class dynamic dot1p` on an interface or globally.
- Layer 2 or Layer 3 service policies supersede dot1p service classes.
- Create service classes.

  INTERFACE mode

  ```
  service-class dynamic dot1p
  ```

Guaranteeing Bandwidth to dot1p-Based Service Queues

To guarantee bandwidth to dot1p-based service queues, use the following command.

Apply this command in the same way as the `bandwidth-percentage` command in an output QoS policy (refer to Allocating Bandwidth to Queue). The `bandwidth-percentage` command in QOS-POLICY-OUT mode supersedes the `service-class bandwidth-percentage` command.

- Guarantee a minimum bandwidth to queues globally.

  CONFIGURATION mode

  ```
  service-class bandwidth-percentage
  ```

Applying an Input Policy Map to an Interface

To apply an input policy map to an interface, use the following command.

You can apply the same policy map to multiple interfaces, and you can modify a policy map after you apply it.

- You cannot apply a class-map and QoS policies to the same interface.
- You cannot apply an input Layer 2 QoS policy on an interface you also configure with vlan-stack access.
- If you apply a service policy that contains an ACL to more than one interface, Dell EMC Networking OS uses ACL optimization to conserve CAM space. The ACL optimization behavior detects when an ACL exists in the CAM rather than writing it to the CAM multiple times.
- Apply an input policy map to an interface.

  INTERFACE mode

  ```
  service-policy input
  ```

  Specify the keyword `layer2` if the policy map you are applying a Layer 2 policy map.

Creating Output Policy Maps

1. Create an output policy map.
   CONFIGURATION mode
   `policy-map-output`
2. After you create an output policy map, do one or more of the following:
   Applying an Output QoS Policy to a Queue
   Specifying an Aggregate QoS Policy
   Applying an Output Policy Map to an Interface
3. Apply the policy map to an interface.

Applying an Output QoS Policy to a Queue

To apply an output QoS policy to a queue, use the following command.

- Apply an output QoS policy to queues.
  INTERFACE mode
  `service-queue`

Specifying an Aggregate QoS Policy

To specify an aggregate QoS policy, use the following command.

- Specify an aggregate QoS policy.
  POLICY-MAP-OUT mode
  `policy-aggregate`

Applying an Output Policy Map to an Interface

To apply an output policy map to an interface, use the following command.

- Apply an input policy map to an interface.
  INTERFACE mode
  `service-policy output`
  You can apply the same policy map to multiple interfaces, and you can modify a policy map after you apply it.

DSCP Color Maps

This section describes how to configure color maps and how to display the color map and color map configuration.

This sections consists of the following topics:

- Creating a DSCP Color Map
- Displaying Color Maps
- Display Color Map Configuration

Creating a DSCP Color Map

You can create a DSCP color map to outline the differentiated services codepoint (DSCP) mappings to the appropriate color mapping (green, yellow, red) for the input traffic. The system uses this information to classify input traffic on an interface based on the DSCP value of each packet and assigns it an initial drop precedence of green, yellow, or red

The default setting for each DSCP value (0-63) is green (low drop precedence). The DSCP color map allows you to set the number of specific DSCP values to yellow or red. Traffic marked as yellow delivers traffic to the egress interface, which will either transmit or drop the packet based on configured queuing behavior. Traffic marked as red (high drop precedence) is dropped.

Important Points to Remember

- All DSCP values that are not specified as yellow or red are colored green (low drop precedence).
- A DSCP value cannot be in both the yellow and red lists. Setting the red or yellow list with any DSCP value that is already in the other list results in an error and no update to that DSCP list is made.
- Each color map can only have one list of DSCP values for each color; any DSCP values previously listed for that color that are not in the new DSCP list are colored green.
- If you configured a DSCP color map on an interface that does not exist or you delete a DSCP color map that is configured on an interface, that interface uses an all green color policy.

To create a DSCP color map:

1. Create the color-aware map QoS DSCP color map.

   CONFIGURATION mode

   `qos dscp-color-map` *`color-map-name`*

2. Create the color aware map profile.

   DSCP-COLOR-MAP

   `dscp {yellow | red} {`*`list-dscp-values`*`}`

3. Apply the map profile to the interface.

   CONFIG-INTERFACE mode

   `qos dscp-color-policy` *`color-map-name`*

Example: Create a DSCP Color Map

The following example creates a DSCP color map profile, color-awareness policy, and applies it to interface **1/11**.

Create the DSCP color map profile, **bat-enclave-map**, with a `yellow` drop precedence , and set the DSCP values to `9,10,11,13,15,16`

```
DellEMC(conf)# qos dscp-color-map bat-enclave-map
DellEMC(conf-dscp-color-map)# dscp yellow 9,10,11,13,15,16
DellEMC(conf-dscp-color-map)# exit
```

Assign the color map, **bat-enclave-map** to the interface.

```
DellEMC(conf)# interface tengigabitethernet 1/11
DellEMC(conf-if-te-1/11)# qos dscp-color-policy bat-enclave-map
```

Displaying DSCP Color Maps

To display DSCP color maps, use the **show qos dscp-color-map** command in EXEC mode.

Examples for Creating a DSCP Color Map

Display all DSCP color maps.

```
DellEMC# show qos dscp-color-map
Dscp-color-map mapONE
  yellow 4,7
  red 20,30
Dscp-color-map mapTWO
  yellow 16,55
```

Display a specific DSCP color map.

```
DellEMC# show qos dscp-color-map mapTWO
Dscp-color-map mapTWO
  yellow 16,55
```

Displaying a DSCP Color Policy Configuration

To display the DSCP color policy configuration for one or all interfaces, use the `show qos dscp-color-policy {summary [interface] | detail {interface}}` command in EXEC mode.

summary: Displays summary information about a color policy on one or more interfaces.

detail: Displays detailed color policy information on an interface

***interface*:** Enter the name of the interface that has the color policy configured.

Examples for Displaying a DSCP Color Policy

Display summary information about a color policy for one or more interfaces.

```
DellEMC# show qos dscp-color-policy summary
Interface     dscp-color-map
TE 1/10       mapONE
TE 1/11       mapTWO
```

Display summary information about a color policy for a specific interface.

```
DellEMC# show qos dscp-color-policy summary tengigabitethernet 1/10
Interface     dscp-color-map
TE 1/10       mapONE
```

Display detailed information about a color policy for a specific interface

```
DellEMC# show qos dscp-color-policy detail tengigabitethernet 1/10
Interface TenGigabitEthernet 1/10
Dscp-color-map mapONE
  yellow 4,7
  red 20,30
```

Enabling QoS Rate Adjustment

By default while rate limiting, policing, and shaping, Dell EMC Networking OS does not include the Preamble, SFD, or the IFG fields. These fields are overhead; only the fields from MAC destination address to the CRC are used for forwarding and are included in these rate metering calculations.

The Ethernet packet format consists of:

- Preamble: 7 bytes Preamble
- Start frame delimiter (SFD): 1 byte
- Destination MAC address: 6 bytes
- Source MAC address: 6 bytes
- Ethernet Type/Length: 2 bytes
- Payload: (variable)
- Cyclic redundancy check (CRC): 4 bytes
- Inter-frame gap (IFG): (variable)

You can optionally include overhead fields in rate metering calculations by enabling QoS rate adjustment.

QoS rate adjustment is disabled by default.

- Specify the number of bytes of packet overhead to include in rate limiting, policing, and shaping calculations.

  CONFIGURATION mode

  `qos-rate-adjust *overhead-bytes*`

  For example, to include the Preamble and SFD, type `qos-rate-adjust 8`. For variable length overhead fields, know the number of bytes you want to include.

  The default is disabled.

  The range is from 1 to 31.

Enabling Strict-Priority Queueing

In strict-priority queuing, the system de-queues all packets from the assigned queue before servicing any other queues. You can assign strict-priority to one unicast queue, using the strict-priority command.

- Policy-based per-queue rate shaping is not supported on the queue configured for strict-priority queuing. To use queue-based rate-shaping as well as strict-priority queuing at the same time on a queue, use the Scheduler Strict feature as described in Scheduler Strict .
- The `strict-priority` supersedes `bandwidth-percentage` configuration.
- A queue with strict priority can starve other queues in the same port-pipe.
- Assign strict priority to one unicast queue.

  INTERFACE mode

  `service-policy output` *`policy-map-name`*

  Enter the name for the policy map in character format (32 characters maximum).

Weighted Random Early Detection

Weighted random early detection (WRED) is a congestion avoidance mechanism that drops packets to prevent buffering resources from being consumed.

The WRED congestion avoidance mechanism drops packets to prevent buffering resources from being consumed.

Traffic is a mixture of various kinds of packets. The rate at which some types of packets arrive might be greater than others. In this case, the space on the buffer and traffic manager (BTM) (ingress or egress) can be consumed by only one or a few types of traffic, leaving no space for other types. You can apply a WRED profile to a policy-map so that specified traffic can be prevented from consuming too much of the BTM resources.

WRED uses a profile to specify minimum and maximum threshold values. The minimum threshold is the allotted buffer space for specified traffic, for example, 1000KB on egress. If the 1000KB is consumed, packets are dropped randomly at an exponential rate until the maximum threshold is reached (as shown in the following illustration); this procedure is the “early detection” part of WRED. If the maximum threshold, for example, 2000KB, is reached, all incoming packets are dropped until the buffer space consumes less than 2000KB of the specified traffic.

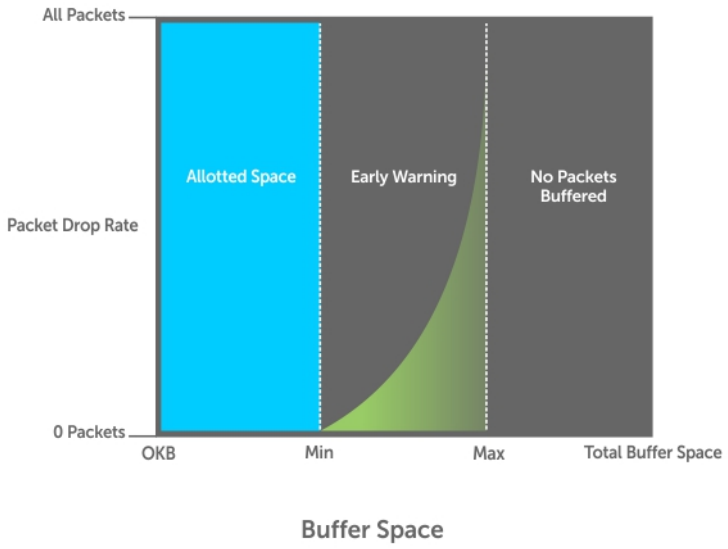


Figure 115. Packet Drop Rate for WRED

You can create a custom WRED profile or use one of the five pre-defined profiles.

Table 79. Pre-Defined WRED Profiles

Default Profile Name	Minimum Threshold	Maximum Threshold	Maximum Drop Rate
wred_drop	0	0	100
wred_teng_y	467	4671	100
wred_teng_g	467	4671	50
wred_fortyg_y	467	4671	50
wred_fortyg_g	467	4671	25

Creating WRED Profiles

To create WRED profiles, use the following commands.

1. Create a WRED profile.

   CONFIGURATION mode

   `wred-profile`
2. Specify the minimum and maximum threshold values.

   WRED mode

   `threshold`

Applying a WRED Profile to Traffic

After you create a WRED profile, you must specify to which traffic Dell EMC Networking OS should apply the profile.

Dell EMC Networking OS assigns a color (also called drop precedence) — red, yellow, or green — to each packet based on it DSCP value before queuing it.

DSCP is a 6–bit field. Dell EMC Networking uses the first three bits (LSB) of this field (DP) to determine the drop precedence.

- DP values of 110 and 100, 101 map to yellow; all other values map to green.
- If you do not configure Dell EMC Networking OS to honor DSCP values on ingress (refer to Honoring DSCP Values on Ingress Packets), all traffic defaults to green drop precedence.
- Assign a WRED profile to either yellow or green traffic.

  QOS-POLICY-OUT mode

  `wred`

Displaying Default and Configured WRED Profiles

To display the default and configured WRED profiles, use the following command.

- Display default and configured WRED profiles and their threshold values.

  EXEC mode

  `show qos wred-profile`

Example of the show qos wred-profile Command.

```
DellEMC#show qos wred-profile

Wred-profile-name min-threshold max-threshold max-drop-rate
wred_drop         0             0             100
wred_teng_y       467           4671          100
wred_teng_g       467           4671          50
wred_fortyg_y     467           4671          50
wred_fortyg_g     467           4671          25
```

Displaying WRED Drop Statistics

To display WRED drop statistics, use the following command.

- Display the number of packets Dell EMC Networking OS the WRED profile drops.

  EXEC Privilege mode

  ```
  show qos statistics wred-profile
  ```

```
DellEMC#show qos statistics wred-profile
Interface Te 1/1
Drop-statistic  Dropped Pkts

Green               51623
Yellow              51300
Out of Profile            0

DellEMC#
```

Displaying egress–queue Statistics

To display the number of transmitted and dropped packets and their rate on the egress queues of an interface, use the following command:

- Display the number of packets and number of bytes on the egress-queue profile.

  EXEC Privilege mode

  ```
  show qos statistics egress-queue
  ```

```
DellEMC#show qos statistics egress-queue tengigabitethernet 1/1

Interface Te 1/1
Unicast/Multicast Egress Queue Statistics
Queue# Q# Type      TxPkts      TxPkts/s       TxBytes    TxBytes/s  DroppedPkts  DroppedPkts/s  DroppedBytes
DroppedBytes/s
-------------------------------------------------------------------------------------------------------------------------------
-
  0    UCAST             0             0             0            0            0              0             0
0
  1    UCAST             0             0             0            0            0              0             0
0
  2    UCAST             0             0             0            0            0              0             0
0
  3    UCAST    1708252734       2340972  874625399808   1198000301   1357413570              0  694995747840
0
  4    UCAST             0             0             0            0            0              0             0
0
  5    UCAST             0             0             0            0            0              0             0
0
  6    UCAST             0             0             0            0            0              0             0
0
  7    UCAST             0             0             0            0            0              0             0
0
  8    UCAST             0             0             0            0            0              0             0
0
  9    UCAST          1132             1        143063          217            0              0             0
0
 10    MCAST             0             0             0            0            0              0             0
0
 11    MCAST             0             0             0            0            0              0             0
0
 12    MCAST             0             0             0            0            0              0             0
0
 13    MCAST             0             0             0            0            0              0             0
0
 14    MCAST             0             0             0            0            0              0             0
0
 15    MCAST             0             0             0            0            0              0             0
0
 16    MCAST             0             0             0            0            0              0             0
0
 17    MCAST             0             0             0            0            0              0             0
0
 18    MCAST             0             0             0            0            0              0             0
0
 19    MCAST             0             0             0            0            0              0             0
0
DellEMC#
```

Pre-Calculating Available QoS CAM Space

Before Dell EMC Networking OS version 7.3.1, there was no way to measure the number of CAM entries a policy-map would consume (the number of CAM entries that a rule uses is not predictable; from 1 to 16 entries might be used per rule depending upon its complexity). Therefore, it was possible to apply to an interface a policy-map that requires more entries than are available. In this case, the system writes as many entries as possible, and then generates an CAM-full error message (shown in the following example). The partial policy-map configuration might cause unintentional system behavior.

```
%EX2YD:12 %DIFFSERV-2-DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3
Cam(PolicyQos) for class 2 (TeGi 12/20) entries on portpipe 1
```

The `test cam-usage` command allows you to verify that there are enough available CAM entries before applying a policy-map to an interface so that you avoid exceeding the QoS CAM space and partial configurations. This command measures the size of the specified policy-map and compares it to the available CAM space in a partition for a specified port-pipe.

Test the policy-map size against the CAM space for a specific port-pipe or all port-pipes using these commands:

- `test cam-usage service-policy input` *`policy-map`* `{stack-unit }` *`number`* `port-set` *`number`*
- `test cam-usage service-policy input` *`policy-map`* `{stack-unit }` *`all`*

The output of this command, shown in the following example, displays:

- The estimated number of CAM entries the policy-map will consume.
- Whether or not the policy-map can be applied.
- The number of interfaces in a port-pipe to which the policy-map can be applied.

Specifically:

- **Available CAM** — the available number of CAM entries in the specified CAM partition for the specified line card or stack-unit port-pipe.
- **Estimated CAM** — the estimated number of CAM entries that the policy will consume when it is applied to an interface.
- **Status** — indicates whether the specified policy-map can be completely applied to an interface in the port-pipe.
  - **Allowed** — indicates that the policy-map can be applied because the estimated number of CAM entries is less or equal to the available number of CAM entries. The number of interfaces in the port-pipe to which the policy-map can be applied is given in parentheses.
  - **Exception** — indicates that the number of CAM entries required to write the policy-map to the CAM is greater than the number of available CAM entries, and therefore the policy-map cannot be applied to an interface in the specified port-pipe.

NOTE: The `show cam-usage` command provides much of the same information as the `test cam-usage` command, but whether a policy-map can be successfully applied to an interface cannot be determined without first measuring how many CAM entries the policy-map would consume; the `test cam-usage` command is useful because it provides this measurement.

- Verify that there are enough available CAM entries.
  `test cam-usage`

```
DellEMC# test cam-usage service-policy input pmap_l2  port-set 0  | port pipe

Port-pipe | CAM Partition | Available CAM | Estimated CAM | Status
=============================================================================
  0           L2ACL           500             200          Allowed(2)
```

Configuring Weights and ECN for WRED

The WRED congestion avoidance functionality drops packets to prevent buffering resources from being consumed. Traffic is a mixture of various kinds of packets. The rate at which some types of packets arrive might be greater than others. In this case, the space on the buffer and traffic manager (BTM) (ingress or egress) can be consumed by only one or few types of traffic, leaving no space for other types. You can apply a WRED profile to a policy-map so that the specified traffic can be prevented from consuming too much of the BTM resources.

WRED drops packets when the average queue length exceeds the configured threshold value to signify congestion. ECN is a capability that enhances WRED by marking the packets instead of causing WRED to drop them when the threshold value is

exceeded. If you configure ECN for WRED, devices employ ECN to mark the packets and reduce the rate of sending packets in a congested network.

In a best-effort network topology, data packets are transmitted in a manner in which latency or throughput is not maintained to be at an effective level. Packets are dropped when the network experiences a large traffic load. This best-effort network deployment is not suitable for applications that are time-sensitive, such as video on demand (VoD) or voice over IP (VoIP) applications. In such cases, you can use ECN in conjunction with WRED to resolve the dropping of packets under congested conditions.

Using ECN, the packets are marked for transmission at a later time after the network recovers from the heavy traffic state to an optimal load. In this manner, enhanced performance and throughput are achieved. Also, the devices can respond to congestion before a queue overflows and packets are dropped, enabling improved queue management.

When a packet reaches the device with ECN enabled for WRED, the average queue size is computed. To measure the average queue size, a weight factor is used. This weight factor is user-configurable. You can use the `wred weight number` command to configure the weight for the WRED average queue size. The mark probability value is the number of packets dropped when the average queue size reaches the maximum threshold value.

The weight factor is set to zero by default, which causes the same behavior as dropping of packets by WRED during network loads or also called instantaneous ECN marking. In a topology in which congestion of the network varies over time, you can specify a weight to enable a smooth, seamless averaging of packets to handle the sudden overload of packets based on the previous time sampling performed. You can specify the weight parameter for front-end and backplane ports separately in the range of 0 through 15.

You can enable WRED and ECN capabilities per queue for granularity. You can disable these functionality per queue, and you can also specify the minimum and maximum buffer thresholds for each color-coding of the packets. You can configure maximum drop rate percentage of yellow and green profiles. You can set up these parameters for both front-end and backplane ports.

Global Service Pools With WRED and ECN Settings

Support for global service pools is now available. You can configure global service pools that are shared buffer pools accessed by multiple queues when the minimum guaranteed buffers for the queue are consumed. S4820T platform support four global service-pools in the egress direction. Two service pools are used– one for loss-based queues and the other for lossless (priority-based flow control (PFC)) queues. You can enable WRED and ECN configuration on the global service-pools.

You can define WRED profiles and weight on each of the global service-pools for both loss-based and lossless (PFC) service-pools. The following events occur when you configure WRED and ECN on global service-pools:

- If WRED/ECN is enabled on the global service-pool with threshold values and if it is not enabled on the queues, WRED/ECN are not effective based on global service-pool WRED thresholds. The queue on which the traffic is scheduled must contain WRED/ECN settings, which are enabled for WRED, to be valid for that traffic.
- When WRED is configured on the global service-pool (regardless of whether ECN on global service-pool is configured), and one or more queues have WRED enabled and ECN disabled, WRED is effective for the minimum of the thresholds between the queue threshold and the service-pool threshold.
- When WRED is configured on the global service-pool (regardless of whether ECN on global service-pool is configured), and one or more queues are enabled with both WRED and ECN, ECN marking takes effect. The packets are ECN marked up to shared- buffer limits as determined by the shared-ratio for that global service-pool.

WRED/ECN configurations for the queues that belong to backplane ports are common to all the backplane ports and cannot be specified separately for each backplane port granularity. This behavior occurs to prevent system-level complexities in enabling this support for backplane ports. Also, WRED/ECN is not supported for multicast packets.

The following table describes the WRED and ECN operations that occur for various scenarios of WRED and ECN configuration on the queue and service pool. (X denotes not-applicable in the table, 1 indicates that the setting is enabled, 0 represents a disabled setting.)

Table 80. Scenarios of WRED and ECN Configuration

Queue Configuration		**Service-Pool Configuration**		**WRED Threshold Relationship** **Q threshold = Q-T,** **Service pool threshold = SP-T**	**Expected Functionality**
WRED	ECN	WRED	ECN		
0	0	X	X	X	WRED/ECN not applicable

Table 80. Scenarios of WRED and ECN Configuration (continued)

Queue Configuration		Service-Pool Configuration		WRED Threshold Relationship Q threshold = Q-T, Service pool threshold = SP-T	Expected Functionality
1	0	0	X	X	Queue based WRED, No ECN marking
		1	X	Q-T < SP-T	
				SP-T < Q-T	SP based WRED, No ECN marking
1	1	0	X	X	Queue-based ECN marking above queue threshold.
		1	X	Q-T < SP-T	ECN marking to shared buffer limits of the service-pool and then packets are tail dropped.
				SP-T < Q-T	Same as above but ECN marking starts above SP-T.

Configuring WRED and ECN Attributes

The functionality to configure a weight factor for the WRED and ECN functionality for backplane ports is supported on the platform.

WRED drops packets when the average queue length exceeds the configured threshold value to signify congestion. Explicit Congestion Notification (ECN) is a capability that enhances WRED by marking the packets instead of causing WRED to drop them when the threshold value is exceeded. If you configure ECN for WRED, devices employ this functionality of ECN to mark the packets and reduce the rate of sending packets in a congested, heavily-loaded network.

To configure the weight factor for WRED and ECN capabilities, global buffer pools for multiple queues, and associating a service class with ECN marking, perform the following:

1. Configure the weight factor for the computation of average-queue size. This weight value applies to front-end ports.

   QOS-POLICY-OUT mode

   ```
   DellEMC(conf-qos-policy-out)#wred—profile weight number
   ```

2. Configure a WRED profile, and specify the threshold and maximum drop rate.

   WRED mode

   ```
   DellEMC(conf-wred) #wred—profile thresh-1
   DellEMC(conf-wred) #threshold min 100 max 200 max-drop-rate 40
   ```

3. Configure another WRED profile, and specify the threshold and maximum drop rate.

   WRED mode

   ```
   DellEMC(conf-wred) #wred—profile thresh-2
   DellEMC(conf-wred) #threshold min 300 max 400 max-drop-rate 80
   ```

4. Create a global buffer pool that is a shared buffer pool accessed by multiple queues when the minimum guaranteed buffers for the queue are consumed. S4820T platform supports four global service-pools in the egress direction.

   mode

   ```
   DellEMC(conf) #service-pool wred green pool0 thresh-1 pool1 thresh-2
   DellEMC(conf) #service-pool wred yellow pool0 thresh-3 pool1 thresh-4
   DellEMC(conf) #service-pool wred weight pool0 11 pool1 4
   ```

5. Create a service class and associate the threshold weight of the shared buffer with each of the queues per port in the egress direction.

   INTERFACE mode

   ```
   DellEMC(conf-if-te-0/8)#Service-class buffer shared-threshold-weight
   ```

Guidelines for Configuring ECN for Classifying and Color-Marking Packets

Keep the following points in mind while configuring the marking and mapping of incoming packets using ECN fields in IPv4 headers:

- Currently Dell EMC Networking OS supports matching only the following TCP flags:
  - ACK
  - FIN
  - SYN
  - PSH
  - RST
  - URG

In the existing software, ECE/CWR TCP flag qualifiers are not supported.

- Because this functionality forcibly marks all the packets matching the specific match criteria as 'yellow', Dell EMC Networking OS does not support Policer based coloring and this feature concurrently.
- If single rate two color policer is configured along with this feature, then by default all packets less than PIR would be considered as "Green" But 'Green' packets matching the specific match criteria for which 'color-marking' is configured will be over-written and marked as "Yellow".
- If two rate three color policer is configured along with this feature then,
  - x < CIR – will be marked as "Green"
  - CIR < x< PIR – will be marked as "Yellow"
  - PIR < x – will be marked as "Red"

But 'Green' packets matching the specific match criteria for which 'color-marking' is configured will be over-written and marked as "Yellow".

Sample configuration to mark non-ecn packets as "yellow" with Multiple traffic class

Consider the example where there are no different traffic classes that is all the packets are egressing on the default 'queue0'.

Dell EMC Networking OS can be configured as below to mark the non-ecn packets as yellow packets.

```
!
ip access-list standard ecn_0
 seq 5 permit any ecn 0

class-map match-any ecn_0_cmap
 match ip access-group ecn_0 set-color  yellow

!
policy-map-input ecn_0_pmap
 service-queue 0 class-map ecn_0_cmap
```

Applying this policy-map "ecn_0_pmap" will mark all the packets with 'ecn == 0' as yellow packets on queue0 (default queue).

Classifying Incoming Packets Using ECN and Color-Marking

Explicit Congestion Notification (ECN) is a capability that enhances WRED by marking the packets instead of causing WRED to drop them when the threshold value is exceeded. If you configure ECN for WRED, devices employ this functionality of ECN to mark the packets and reduce the rate of sending packets in a congested, heavily-loaded network.

ECN is a mechanism using which network switches indicate congestion to end hosts for initiating appropriate action. End hosts uses two least significant bits of ToS to indicate that it is ECT. When intermediate network node encounters congestion, remarks ECT to CE for end host to take appropriate action. During congestion, ECN enabled packets are not subject to any kind of drops like WRED except tail drops. Though ECN & WRED are independent technologies, BRCM has made WRED a mandatory for ECN to work.

On ECN deployment, the non-ECN packets that are transmitted on the ECN-WRED enabled interface will be considered as Green packets and will be subject to the early WRED drops. Typically the TCP-acks, OAM, ICMP ping packets will be non-ECN in nature and it is not desirable for this packets getting WRED dropped.

In such a condition, it is necessary that the switch is capable to take differentiated actions for ECN/Non-ECN packets. After classifying packets to ECN/Non-ECN, marking ECN and Non-ECN packets to different color packets is performed.

Policy based ingress QOS involves the following three steps to achieve QOS:

1. Classification of incoming traffic.
2. Specify the differentiated actions for different traffic class.
3. Attach the policy-map to the interface.

Dell EMC Networking OS support different types of match qualifiers to classify the incoming traffic.

Match qualifiers can be directly configured in the class-map command or it can be specified through one or more ACL which in turn specifies the combination of match qualifiers.

Until Release 9.3(0.0), support is available for classifying traffic based on the 6-bit DSCP field of the IPv4 packet.

As a part of this feature, the 2-bit ECN field of the IPv4 packet will also be available to be configured as one of the match qualifier. This way the entire 8-bit ToS field of the IPv4 header shall be used to classify traffic.

The Dell EMC Networking OS Release 9.3(0.0) supports the following QOS actions in the ingress policy based QOS:

1. Rate Policing
2. Queuing
3. Marking

For the L3 Routed packets, the DSCP marking is the only marking action supported in the software. As a part of this feature, the additional marking action to set the “color” of the traffic will be provided.

Until Release 9.3(0.0), the software has the capability to qualify only on the 6-bit DSCP part of the ToS field in IPv4 Header. You can now accept and process incoming packets based on the 2-bit ECN part of the ToS field in addition to the DSCP categorization. The IPv4 ACLs (standard and Extended) are enhanced to add this qualifier. This new keyword ‘ecn’ is present for all L3 ACL types (TCP/UDP/IP/ICMP) at the level where the ‘DSCP’ qualifier is positioned in the current ACL commands.

Dell EMC Networking OS supports the capability to contain DSCP and ECN classifiers simultaneously for the same ACL entry.

You can use the ecn keyword with the ip access-list standard, ip access-list extended, seq, and permit commands for standard and extended IPv4 ACLs to match incoming packets with the specified ECN values.

Similar to ‘dscp’ qualifier in the existing L3 ACL command, the ‘ecn’ qualifier can be used along with all other supported ACL match qualifiers such as SIP/DIP/TCP/UDP/SRC PORT/DST PORT/ ICMP.

Until Release 9.3(0.0), ACL supports classification based on the below TCP flags:

- ACK
- FIN
- SYN
- PSH
- RST
- URG

You can now use the ‘ecn’ match qualifier along with the above TCP flag for classification.

The following combination of match qualifiers is acceptable to be configured for the Dell EMC Networking OS software through L3 ACL command:

- Classification based on DSCP only
- Classification based on ECN only
- Classification based on ECN and DSCP concurrently

You can now use the set-color yellow keyword with the match ip access-group command to mark the color of the traffic as ‘yellow’ would be added in the ‘match ip’ sequence of the class-map configuration.

By default, all packets are considered as ‘green’ (without the rate-policer and trust-diffserve configuration) and hence support would be provided to mark the packets as ‘yellow’ alone will be provided.

By default Dell EMC Networking OS drops all the ‘RED’ or ‘violate’ packets.

The following combination of marking actions to be specified match sequence of the class-map command:

- set a new DSCP for the packet
- set the packet color as ‘yellow’
- set the packet color as ‘yellow’ and set a new DSCP for the packet

This marking action to set the color of the packet is allowed only on the ‘match-any’ logical operator of the class-map.

This marking-action can be configured for all of the below L3 match sequence types:

- match ip access-group
- match ip dscp
- match ip precedence
- match ip vlan

Sample configuration to mark non-ecn packets as “yellow” with single traffic class

Consider the use case where the packet with DSCP value “40” need to be enqueued in queue#2 and packets with DSCP value as 50 need to be enqueued in queue#3. And all the packets with ecn value as ‘0’ must be marked as ‘yellow’.

The above requirement can be achieved using either of the two approaches.

The above requirement can be achieved using either of the two approaches.

Approach without explicit ECN match qualifiers for ECN packets:

```
!
ip access-list standard dscp_50
 seq 5 permit any dscp 50
!
ip access-list standard dscp_40
 seq 5 permit any dscp 40
!
ip access-list standard dscp_50_non_ecn
 seq 5 permit any dscp 50 ecn 0

!
ip access-list standard dscp_40_non_ecn
 seq 5 permit any dscp 40 ecn 0
!
class-map match-any class_dscp_40
 match ip access-group dscp_40_non_ecn set-color yellow
 match ip access-group dscp_40
!
class-map match-any class_dscp_50
 match ip access-group dscp_50_non_ecn set-color yellow
 match ip access-group dscp_50
!
```

```
policy-map-input pmap_dscp_40_50
 service-queue 2 class-map class_dscp_40
 service-queue 3 class-map class_dscp_50
```

Approach with explicit ECN match qualifiers for ECN packets:

```
!
ip access-list standard dscp_50_ecn
 seq 5 permit any dscp 50 ecn 1
 seq 10 permit any dscp 50 ecn 2
 seq 15 permit any dscp 50 ecn 3

!
ip access-list standard dscp_40_ecn
 seq 5 permit any dscp 40 ecn 1
 seq 10 permit any dscp 40 ecn 2
 seq 15 permit any dscp 40 ecn 3

!
ip access-list standard dscp_50_non_ecn
 seq 5 permit any dscp 50 ecn 0

!
ip access-list standard dscp_40_non_ecn
 seq 5 permit any dscp 40 ecn 0
!
class-map match-any class_dscp_40
 match ip access-group dscp_40_non_ecn set-color yellow
 match ip access-group dscp_40_ecn
!
class-map match-any class_dscp_50
 match ip access-group dscp_50_non_ecn set-color yellow
 match ip access-group dscp_50_ecn
!
policy-map-input pmap_dscp_40_50
 service-queue 2 class-map class_dscp_40
 service-queue 3 class-map class_dscp_50
```

Applying Layer 2 Match Criteria on a Layer 3 Interface

To process Layer 3 packets that contain a dot1p (IEEE 802.1p) VLAN Layer 2 header, configure VLAN tags on a Layer 3 port interface which is configured with an IP address but has no VLAN associated with it. You can also configure a VLAN sub-interface on the port interface and apply a policy map that classifies packets using the dot1p VLAN ID.

To apply an input policy map with Layer 2 match criteria to a Layer 3 port interface, use the `service-policy input policy-name layer 2` command in Interface configuration mode.

To apply a Layer 2 policy on a Layer 3 interface:

1. Configure an interface with an IP address or a VLAN sub-interface

   CONFIGURATION mode

   ```
   DellEMC(conf)# interface fo 1/49/1
   ```

   INTERFACE mode

   ```
   DellEMC(conf-if-fo-1/49/1)# ip address 90.1.1.1/16
   ```

2. Configure a Layer 2 QoS policy with Layer 2 (Dot1p or source MAC-based) match criteria.

   CONFIGURATION mode

   ```
   DellEMC(conf)# policy-map-input l2p layer2
   ```

3. Apply the Layer 2 policy on a Layer 3 interface.

   INTERFACE mode

   ```
   DellEMC(conf-if-fo-1/49/1)# service-policy input l2p layer2
   ```

Applying DSCP and VLAN Match Criteria on a Service Queue

You can configure Layer 3 class maps which contain both a Layer 3 Differentiated Services Code Point (DSCP) and IP VLAN IDs as match criteria to filter incoming packets on a service queue on the switch.

To configure a Layer 3 class map to classify traffic according to both an IP VLAN ID and DSCP value, use the `match ip vlan` *`vlan-id`* command in class-map input configuration mode. You can include the class map in a policy map, and apply the class and policy map to a service queue using the `service-queue` command. In this way, the system applies the match criteria in a class map according to queue priority (queue numbers closer to 0 have a lower priority).

To configure IP VLAN and DSCP match criteria in a Layer 3 class map, and apply the class and policy maps to a service queue:

1. Create a match-any or a match-all Layer 3 class map, depending on whether you want the packets to meet all or any of the match criteria. By default, a Layer 3 class map is created if you do not enter the `layer2` option with the class-map command. When you create a class map, you enter the class-map configuration mode.

   CONFIGURATION mode

   ```
   Dell(conf)#class-map match-all pp_classmap
   ```

2. Configure a DSCP value as a match criterion.

   CLASS-MAP mode

   ```
   Dell(conf-class-map)#match ipdscp 5
   ```

3. Configure an IP VLAN ID as a match criterion.

   CLASS-MAP mode

   ```
   Dell(conf-class-map)#match ip vlan 5
   ```

4. Create a QoS input policy.

   CONFIGURATION mode

   ```
   Dell(conf)#qos-policy-input pp_qospolicy
   ```

5. Configure the DSCP value to be set on matched packets.

   QOS-POLICY-IN mode

   ```
   Dell(conf-qos-policy-in)#set ip-dscp 5
   ```

6. Create an input policy map.

   CONFIGURATION mode

   ```
   Dell(conf)#policy-map-input pp_policmap
   ```

7. Create a service queue to associate the class map and QoS policy map.

   POLICY-MAP mode

   ```
   Dell(conf-policy-map-in)#service-queue 0 class-map pp_classmap qos-policy pp_qospolicy
   ```

Enabling Buffer Statistics Tracking

You can enable the tracking of statistical values of buffer spaces at a global level. The buffer statistics tracking utility operates in the max use count mode that enables the collection of maximum values of counters.

To configure the buffer statistics tracking utility, perform the following step:

1. Enable the buffer statistics tracking utility and enter the Buffer Statistics Snapshot configuration mode.

   CONFIGURATION mode

   ```
   DellEMC(conf)#buffer-stats-snapshot
   DellEMC(conf)#no disable
   ```

   Enable this utility to be able to configure the parameters for buffer statistics tracking. By default, buffer statistics tracking is disabled.

2. Enable the buffer statistics tracking utility and enter the Buffer Statistics Snapshot configuration mode.

   CONFIGURATION mode

   ```
   DellEMC(conf)#buffer-stats-snapshot
   DellEMC(conf)#no disable
   ```

Enable this utility to be able to configure the parameters for buffer statistics tracking. By default, buffer statistics tracking is disabled.

3. To view the buffer statistics tracking resource information depending on the type of buffer information, such as device-level details, queue-based snapshots, or priority group-level snapshot in the egress and ingress direction of traffic, use `show hardware stack-unit <id> buffer-stats-snapshot unit <id> resource x`

   EXEC/EXEC Privilege mode

```
DellEMC#show hardware stack-unit 1 buffer-stats-snapshot unit 3 resource interface
all queue mcast 3
Unit 1 unit: 3 port: 1 (interface Fo 1/144)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 5 (interface Fo 1/148)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 9 (interface Fo 1/152)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 13 (interface Fo 1/156)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 17 (interface Fo 1/160)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 21 (interface Fo 1/164)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 25 (interface Fo 1/168)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 29 (interface Fo 1/172)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 33 (interface Fo 1/176)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
MCAST       3      0

Unit 1 unit: 3 port: 37 (interface Fo 1/180)
---------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
---------------------------------------
```

4. Use `show hardware buffer-stats-snapshot resource interface` *`interface`*`{priority-group { id | all } | queue { ucast{id | all}{ mcast {id | all} | all}` to view buffer statistics tracking resource information for a specific interface.

   EXEC/EXEC Privilege mode

```
DellEMC# show hardware buffer-stats-snapshot resource interface fortyGigE 0/0 queue
all
Unit 0 unit: 0 port: 1 (interface Fo 0/0)
----------------------------------------
Q# TYPE     Q#     TOTAL BUFFERED CELLS
----------------------------------------
UCAST       0      0
UCAST       1      0
UCAST       2      0
UCAST       3      0
UCAST       4      0
UCAST       5      0
UCAST       6      0
UCAST       7      0
UCAST       8      0
UCAST       9      0
UCAST       10     0
UCAST       11     0
MCAST       0      0
MCAST       1      0
MCAST       2      0
MCAST       3      0
MCAST       4      0
MCAST       5      0
MCAST       6      0
MCAST       7      0
MCAST       8      0
```

43

Routing Information Protocol (RIP)

The Routing Information Protocol (RIP) tracks distances or hop counts to nearby routers when establishing network connections and is based on a distance-vector algorithm.

RIP is based on a distance-vector algorithm; it tracks distances or hop counts to nearby routers when establishing network connections.

RIP protocol standards are listed in the Standards Compliance chapter.

Topics:

- Protocol Overview
- Implementation Information
- Configuration Information

Protocol Overview

RIP is the oldest interior gateway protocol.

There are two versions of RIP: RIP version 1 (RIPv1) and RIP version 2 (RIPv2). These versions are documented in RFCs 1058 and 2453.

RIPv1

RIPv1 learns where nodes in a network are located by automatically constructing a routing data table.

The routing table is established after RIP sends out one or more broadcast signals to all adjacent nodes in a network. Hop counts of these signals are tracked and entered into the routing table, which defines where nodes in the network are located.

The information that is used to update the routing table is sent as either a request or response message. In RIPv1, automatic updates to the routing table are performed as either one-time requests or periodic responses (every 30 seconds). RIP transports its responses or requests by means of user datagram protocol (UDP) over port 520.

RIP must receive regular routing updates to maintain a correct routing table. Response messages containing a router's full routing table are transmitted every 30 seconds. If a router does not send an update within a certain amount of time, the hop count to that route is changed to unreachable (a route hop metric of 16 hops). Another timer sets the amount of time before the unreachable routes are removed from the routing table.

This first RIP version does not support variable length subnet mask (VLSM) or classless inter-domain routing (CIDR) and is not widely used.

RIPv2

RIPv2 adds support for subnet fields in the RIP routing updates, thus qualifying it as a classless routing protocol.

The RIPv2 message format includes entries for route tags, subnet masks, and next hop addresses. Another enhancement included in RIPv2 is multicasting for route updates on IP multicast address 224.0.0.9.

Implementation Information

Dell EMC Networking OS supports both versions of RIP and allows you to configure one version globally and the other version on interfaces or both versions on the interfaces.

The following table lists the defaults for RIP in Dell EMC Networking OS.

Table 81. RIP Defaults

Feature	Default
Interfaces running RIP	• Listen to RIPv1 and RIPv2 • Transmit RIPv1
RIP timers	• update timer = 30 seconds • invalid timer = 180 seconds • holddown timer = 180 seconds • flush timer = 240 seconds
Auto summarization	Enabled
ECMP paths supported	16

Configuration Information

By default, RIP is disabled in Dell EMC Networking OS.

To configure RIP, you must use commands in two modes: ROUTER RIP and INTERFACE. Commands executed in the ROUTER RIP mode configure RIP globally, while commands executed in the INTERFACE mode configure RIP features on that interface only.

RIP is best suited for small, homogeneous networks. You must configure all devices within the RIP network to support RIP if they are to participate in the RIP.

Configuration Task List

The following is the configuration task list for RIP.

- Enabling RIP Globally (mandatory)
- Configure RIP on Interfaces (optional)
- Controlling RIP Routing Updates (optional)
- Setting Send and Receive Version (optional)
- Generating a Default Route (optional)
- Controlling Route Metrics (optional)
- Summarize Routes (optional)
- Controlling Route Metrics
- Debugging RIP

For a complete listing of all commands related to RIP, refer to the *Dell EMC Networking OS Command Reference Interface Guide*.

Enabling RIP Globally

By default, RIP is not enabled in Dell EMC Networking OS.

To enable RIP globally, use the following commands.

1. Enter ROUTER RIP mode and enable the RIP process on Dell EMC Networking OS.

   CONFIGURATION mode

   ```
   router rip
   ```

2. Assign an IP network address as a RIP network to exchange routing information.

   ROUTER RIP mode

   ```
   network ip-address
   ```

After designating networks with which the system is to exchange RIP information, ensure that all devices on that network are configured to exchange RIP information.

The Dell EMC Networking OS default is to send RIPv1 and to receive RIPv1 and RIPv2. To change the RIP version globally, use the `version` command in ROUTER RIP mode.

To view the global RIP configuration, use the `show running-config` command in EXEC mode or the `show config` command in ROUTER RIP mode.

```
DellEMC(conf-router_rip)#show config
!
router rip
  network 10.0.0.0
DellEMC(conf-router_rip)#
```

When the RIP process has learned the RIP routes, use the `show ip rip database` command in EXEC mode to view those routes.

```
DellEMC#show ip rip database
Total number of routes in RIP database: 978
160.160.0.0/16
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/4
160.160.0.0/16    auto-summary
2.0.0.0/8
    [120/1] via 29.10.10.12, 00:01:22, Fa 1/4
2.0.0.0/8         auto-summary
4.0.0.0/8
    [120/1] via 29.10.10.12, 00:01:22, Fa 1/4
4.0.0.0/8         auto-summary
8.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/4
8.0.0.0/8         auto-summary
12.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/4
12.0.0.0/8        auto-summary
20.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/4
20.0.0.0/8        auto-summary
29.10.10.0/24 directly connected,Fa 1/4
29.0.0.0/8        auto-summary
31.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/4
31.0.0.0/8        auto-summary
192.162.2.0/24
    [120/1] via 29.10.10.12, 00:01:21, Fa 1/4
192.162.2.0/24    auto-summary
192.161.1.0/24
    [120/1] via 29.10.10.12, 00:00:27, Fa 1/4
192.161.1.0/24    auto-summary
192.162.3.0/24
    [120/1] via 29.10.10.12, 00:01:22, Fa 1/4
192.162.3.0/24    auto-summary
```

```
DellEMC#show ip rip database
Total number of routes in RIP database: 978
160.160.0.0/16
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/49
160.160.0.0/16    auto-summary
2.0.0.0/8
    [120/1] via 29.10.10.12, 00:01:22, Fa 1/49
2.0.0.0/8         auto-summary
4.0.0.0/8
    [120/1] via 29.10.10.12, 00:01:22, Fa 1/49
4.0.0.0/8         auto-summary
8.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/49
8.0.0.0/8         auto-summary
12.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/49
12.0.0.0/8        auto-summary
20.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/49
20.0.0.0/8        auto-summary
29.10.10.0/24 directly connected,Fa 1/49
29.0.0.0/8        auto-summary
31.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 1/49
```

```
31.0.0.0/8          auto-summary
192.162.2.0/24
    [120/1] via 29.10.10.12, 00:01:21, Fa 1/49
192.162.2.0/24      auto-summary
192.161.1.0/24
    [120/1] via 29.10.10.12, 00:00:27, Fa 1/49
192.161.1.0/24      auto-summary
192.162.3.0/24
    [120/1] via 29.10.10.12, 00:01:22, Fa 1/49
192.162.3.0/24      auto-summary
```

To disable RIP globally, use the `no router rip` command in CONFIGURATION mode.

Configure RIP on Interfaces

When you enable RIP globally on the system, interfaces meeting certain conditions start receiving RIP routes.

By default, interfaces that you enable and configure with an IP address in the same subnet as the RIP network address receive RIPv1 and RIPv2 routes and send RIPv1 routes.

Assign IP addresses to interfaces that are part of the same subnet as the RIP network identified in the `network` command syntax.

Controlling RIP Routing Updates

By default, RIP broadcasts routing information out all enabled interfaces, but you can configure RIP to send or to block RIP routing information, either from a specific IP address or a specific interface.

To control which devices or interfaces receive routing updates, configure a direct update to one router and configure interfaces to block RIP updates from other sources.

To control the source of RIP route information, use the following commands.

- Define a specific router to exchange RIP information between it and the Dell EMC Networking system.

  ROUTER RIP mode

  `neighbor` *`ip-address`*

  You can use this command multiple times to exchange RIP information with as many RIP networks as you want.
- Disable a specific interface from sending or receiving RIP routing information.

  ROUTER RIP mode

  `passive-interface` *`interface`*

Assigning a Prefix List to RIP Routes

Another method of controlling RIP (or any routing protocol) routing information is to filter the information through a prefix list. A prefix list is applied to incoming or outgoing routes.

Those routes must meet the conditions of the prefix list; if not, Dell EMC Networking OS drops the route. Prefix lists are globally applied on all interfaces running RIP. Configure the prefix list in PREFIX LIST mode prior to assigning it to the RIP process.

For configuration information about prefix lists, refer to Access Control Lists (ACLs).

To apply prefix lists to incoming or outgoing RIP routes, use the following commands.

- Assign a configured prefix list to all incoming RIP routes.

  ROUTER RIP mode

  `distribute-list` *`prefix-list-name`* `in`
- Assign a configured prefix list to all outgoing RIP routes.

  ROUTER RIP mode

  `distribute-list` *`prefix-list-name`* `out`

To view the current RIP configuration, use the `show running-config` command in EXEC mode or the `show config` command in ROUTER RIP mode.

Adding RIP Routes from Other Instances

In addition to filtering routes, you can add routes from other routing instances or protocols to the RIP process.

With the `redistribute` command, you can include open shortest path first (OSPF), static, or directly connected routes in the RIP process.

To add routes from other routing instances or protocols, use the following commands.

- Include directly connected or user-configured (static) routes in RIP.

  ROUTER RIP mode

  `redistribute {connected | static} [metric metric-value] [route-map map-name]`

  - *metric-value*: the range is from 0 to 16.
  - *map-name*: the name of a configured route map.
- Include IS-IS routes in RIP.

  ROUTER RIP mode

  `redistribute isis [level-1 | level-1-2 | level-2] [metric metric-value] [route-map map-name]`

  - *metric-value*: the range is from 0 to 16.
  - *map-name*: the name of a configured route map.
- Include specific OSPF routes in RIP.

  ROUTER RIP mode

  `redistribute ospf process-id [match external {1 | 2} | match internal] [metric value] [route-map map-name]`

  Configure the following parameters:
  - *process-id*: the range is from 1 to 65535.
  - `metric`: the range is from 0 to 16.
  - *map-name*: the name of a configured route map.

To view the current RIP configuration, use the `show running-config` command in EXEC mode or the `show config` command in ROUTER RIP mode.

Setting the Send and Receive Version

To change the RIP version globally or on an interface in Dell EMC Networking OS, use the following command.

To specify the RIP version, use the `version` command in ROUTER RIP mode. To set an interface to receive only one or the other version, use the `ip rip send version` or the `ip rip receive version` commands in INTERFACE mode.

You can set one RIP version globally on the system using `system`. This command sets the RIP version for RIP traffic on the interfaces participating in RIP unless the interface was specifically configured for a specific RIP version.

- Set the RIP version sent and received on the system.

  ROUTER RIP mode

  `version {1 | 2}`
- Set the RIP versions received on that interface.

  INTERFACE mode

  `ip rip receive version [1] [2]`
- Set the RIP versions sent out on that interface.

  INTERFACE mode

  `ip rip send version [1] [2]`

To see whether the `version` command is configured, use the `show config` command in ROUTER RIP mode.

The following example shows the RIP configuration after the `ROUTER RIP mode version` command is set to RIPv2. When you set the `ROUTER RIP mode version` command, the interface (`TenGigabitEthernet 1/1`) participating in the RIP process is also set to send and receive RIPv2 (shown in bold).

To view the routing protocols configuration, use the `show ip protocols` command in EXEC mode.

```
DellEMC#show ip protocols

  Routing Protocols is RIP
  Sending updates every 30 seconds, next due in 23
  Invalid after 180 seconds, hold down 180, flushed after 240
  Output delay 8 milliseconds between packets
  Automatic network summarization is in effect
  Outgoing filter for all interfaces is
  Incoming filter for all interfaces is
  Default redistribution metric is 1
  Default version control: receive version 2, send version 2
       Interface           Recv  Send
       TenGigabitEthernet  1/1  2   2
  Routing for Networks:
       10.0.0.0

  Routing Information Sources:
  Gateway      Distance          Last Update

  Distance: (default is 120)

DellEMC#
```

To configure an interface to receive or send both versions of RIP, include 1 and 2 in the command syntax. The command syntax for sending both RIPv1 and RIPv2 and receiving only RIPv2 is shown in the following example.

```
DellEMC(conf-if)#ip rip send version 1 2
DellEMC(conf-if)#ip rip receive version 2
```

The following example of the `show ip protocols` command confirms that both versions are sent out that interface. This interface no longer sends and receives the same RIP versions as Dell EMC Networking OS does globally (shown in bold).

```
DellEMC#show ip protocols
  Routing Protocols is RIP
  Sending updates every 30 seconds, next due in 11
  Invalid after 180 seconds, hold down 180, flushed after 240
  Output delay 8 milliseconds between packets
  Automatic network summarization is in effect
  Outgoing filter for all interfaces is
  Incoming filter for all interfaces is
  Default redistribution metric is 1
  Default version control: receive version 2, send version 2
         Interface          Recv   Send
         TenGigabitEthernet 1/1  2    1  2
  Routing for Networks:
         10.0.0.0

Routing Information Sources:
  Gateway      Distance        Last Update

  Distance: (default is 120)

DellEMC#
```

Generating a Default Route

Traffic is forwarded to the default route when the traffic's network is not explicitly listed in the routing table.

Default routes are not enabled in RIP unless specified. Use the `default-information originate` command in ROUTER RIP mode to generate a default route into RIP. In Dell EMC Networking OS, default routes received in RIP updates from other routes are advertised if you configure the `default-information originate` command.

- Specify the generation of a default route in RIP.

  ROUTER RIP mode

  `default-information originate [always] [metric value] [route-map route-map-name]`

  - `always`: Enter the keyword `always` to always generate a default route.

- *value* The range is from 1 to 16.
- *route-map-name*: The name of a configured route map.

To confirm that the default route configuration is completed, use the `show config` command in ROUTER RIP mode.

Summarize Routes

Routes in the RIPv2 routing table are summarized by default, thus reducing the size of the routing table and improving routing efficiency in large networks.

By default, the `autosummary` command in ROUTER RIP mode is enabled and summarizes RIP routes up to the classful network boundary.

If you must perform routing between discontiguous subnets, disable automatic summarization. With automatic route summarization disabled, subnets are advertised.

The `autosummary` command requires no other configuration commands. To disable automatic route summarization, enter `no autosummary` in ROUTER RIP mode.

NOTE: If you enable the `ip split-horizon` command on an interface, the system does not advertise the summarized address.

Controlling Route Metrics

As a distance-vector protocol, RIP uses hop counts to determine the best route, but sometimes the shortest hop count is a route over the lowest-speed link.

To manipulate RIP routes so that the routing protocol prefers a different route, manipulate the route by using the `offset` command.

Exercise caution when applying an `offset` command to routers on a broadcast network, as the router using the `offset` command is modifying RIP advertisements before sending out those advertisements.

The `distance` command also allows you to manipulate route metrics. To assign different weights to routes so that the ones with the lower weight or administrative distance assigned are preferred, use the `distance` command.

To set route matrixes, use the following commands.

- Apply a weight to all routes or a specific route and ACL.

  ROUTER RIP mode

  `distance weight [ip-address mask [access-list-name]]`

  Configure the following parameters:
  - *weight*: the range is from 1 to 255. The default is **120**.
  - *ip-address mask*: the IP address in dotted decimal format (A.B.C.D), and the mask in slash format (/x).
  - *access-list-name*: the name of a configured IP ACL.
- Apply an additional number to the incoming or outgoing route metrics.

  ROUTER RIP mode

  `offset-list access-list-name {in | out} offset [interface]`

  Configure the following parameters:
  - *prefix-list-name*: the name of an established Prefix list to determine which incoming routes are modified
  - *offset*: the range is from 0 to 16.
  - *interface*: the type, slot, and number of an interface.

To view the configuration changes, use the `show config` command in ROUTER RIP mode.

Debugging RIP

The `debug ip rip` command enables RIP debugging.

When you enable debugging, you can view information on RIP protocol changes or RIP routes.

To enable RIP debugging, use the following command.

- `debug ip rip [interface | database | events | trigger]`

  EXEC privilege mode

  Enable debugging of RIP.

The following example shows the confirmation when you enable the debug function.

```
DellEMC#debug ip rip
RIP protocol debug is ON
DellEMC#
```

To disable RIP, use the `no debug ip rip` command.

RIP Configuration Example

The examples in this section show the command sequence to configure RIPv2 on the two routers shown in the following illustration — *Core 2* and *Core 3*.

The host prompts used in the following example reflect those names. The examples are divided into the following groups of command sequences:

- Configuring RIPv2 on Core 2
- Core 2 RIP Output
- RIP Configuration on Core 3
- Core 3 RIP Output
- RIP Configuration Summary

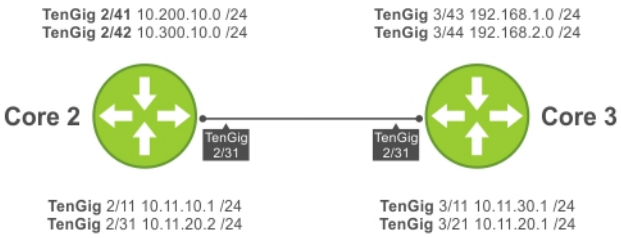


Figure 116. RIP Topology Example

RIP Configuration on Core2

The following example shows how to configure RIPv2 on a host named Core2.

```
Core2(conf-if-te-2/3)#
Core2(conf-if-te-2/3)#router rip
Core2(conf-router_rip)#ver 2
Core2(conf-router_rip)#network 10.200.10.0
Core2(conf-router_rip)#network 10.300.10.0
Core2(conf-router_rip)#network 10.11.10.0
Core2(conf-router_rip)#network 10.11.20.0
Core2(conf-router_rip)#show config
!
router rip
  network 10.0.0.0
  version 2
Core2(conf-router_rip)#
```

Core 2 RIP Output

The examples in the section show the core 2 RIP output.

- To display Core 2 RIP database, use the `show ip rip database` command.

- To display Core 2 RIP setup, use the `show ip route` command.
- To display Core 2 RIP activity, use the `show ip protocols` command.

The following example shows the `show ip rip database` command to view the learned RIP routes on Core 2.

```
Core2(conf-router_rip)#end
00:12:24: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by console
Core2#show ip rip database
Total number of routes in RIP database: 7
10.11.30.0/24
    [120/1] via 10.11.20.1, 00:00:03, TenGigabitEthernet 2/3
10.300.10.0/24     directly connected,TenGigabitEthernet 2/4
10.200.10.0/24     directly connected,TenGigabitEthernet 2/5
10.11.20.0/24      directly connected,TenGigabitEthernet 2/6
10.11.10.0/24      directly connected,TenGigabitEthernet 2/11
10.0.0.0/8         auto-summary
192.168.1.0/24
    [120/1] via 10.11.20.1, 00:00:03, TenGigabitEthernet 2/3
192.168.1.0/24     auto-summary
192.168.2.0/24
    [120/1] via 10.11.20.1, 00:00:03, TenGigabitEthernet 2/3
192.168.2.0/24     auto-summary
Core2#
```

The following example shows the `show ip route` command to show the RIP setup on Core 2.

```
Core2#show ip route

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

Destination Gateway            Dist/Metric Last Change
----------- ------- ----------- -----------
C    10.11.10.0/24   Direct, Te 2/11         0/0   00:02:26
C    10.11.20.0/24   Direct, Te 2/3          0/0   00:02:02
R    10.11.30.0/24   via 10.11.20.1, Te 2/3  120/1 00:01:20
C    10.200.10.0/24  Direct, Te 2/4          0/0   00:03:03
C    10.300.10.0/24  Direct, Te 2/5          0/0   00:02:42
R    192.168.1.0/24  via 10.11.20.1, Te 2/3  120/1 00:01:20
R    192.168.2.0/24  via 10.11.20.1, Te 2/3  120/1 00:01:20
Core2#
R    192.168.1.0/24  via 10.11.20.1, Te 2/3  120/1 00:05:22
R    192.168.2.0/24  via 10.11.20.1, Te 2/3  120/1 00:05:22

Core2#
```

The following example shows the `show ip protocols` command to show the RIP configuration activity on Core 2.

```
Core2#show ip protocols
Routing Protocol is "RIP"
  Sending updates every 30 seconds, next due in 17
  Invalid after 180 seconds, hold down 180, flushed after 240
  Output delay 8 milliseconds between packets
  Automatic network summarization is in effect
  Outgoing filter for all interfaces is
  Incoming filter for all interfaces is
  Default redistribution metric is 1
  Default version control: receive version 2, send version 2
      Interface Recv Send
      TenGigabitEthernet 2/4  2 2
      TenGigabitEthernet 2/5  2 2
      TenGigabitEthernet 2/3  2 2
      TenGigabitEthernet 2/11 2 2
Routing for Networks:
```

```
      10.300.10.0
      10.200.10.0
      10.11.20.0
      10.11.10.0

Routing Information Sources:
Gateway      Distance    Last Update
10.11.20.1   120         00:00:12

Distance: (default is 120)
Core2#
```

RIP Configuration on Core3

The following example shows how to configure RIPv2 on a host named Core3.

```
Core3(conf)#router rip
Core3(conf-router_rip)#version 2
Core3(conf-router_rip)#network 192.168.1.0
Core3(conf-router_rip)#network 192.168.2.0
Core3(conf-router_rip)#network 10.11.30.0
Core3(conf-router_rip)#network 10.11.20.0
Core3(conf-router_rip)#show config
!
router rip
  network 10.0.0.0
  network 192.168.1.0
  network 192.168.2.0
  version 2
Core3(conf-router_rip)#
```

Core 3 RIP Output

The examples in this section show the core 2 RIP output.

- To display Core 3 RIP database, use the `show ip rip database` command.
- To display Core 3 RIP setup, use the `show ip route` command.
- To display Core 3 RIP activity, use the `show ip protocols` command.

The following example shows the `show ip rip database` command to view the learned RIP routes on Core 3.

```
Core3#show ip rip database
Total number of routes in RIP database: 7
10.11.10.0/24
    [120/1] via 10.11.20.2, 00:00:13, TenGigabitEthernet 3/21
10.200.10.0/24
    [120/1] via 10.11.20.2, 00:00:13, TenGigabitEthernet 3/21
10.300.10.0/24
    [120/1] via 10.11.20.2, 00:00:13, TenGigabitEthernet 3/21
10.11.20.0/24      directly connected,TenGigabitEthernet 3/21
10.11.30.0/24      directly connected,TenGigabitEthernet 3/11
10.0.0.0/8         auto-summary
192.168.1.0/24     directly connected,TenGigabitEthernet 3/23
192.168.1.0/24     auto-summary
192.168.2.0/24     directly connected,TenGigabitEthernet 3/24
192.168.2.0/24     auto-summary
Core3#
```

The following command shows the `show ip routes` command to view the RIP setup on Core 3.

```
Core3#show ip routes

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
```

```
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set
   Destination    Gateway           Dist/Metric Last Change
  ----------- ------- -----------          -----------
R  10.11.10.0/24  via 10.11.20.2, Te 3/21   120/1      00:01:14
C  10.11.20.0/24  Direct, Te 3/21           0/0        00:01:53
C  10.11.30.0/24  Direct, Te 3/11           0/0        00:06:00
R  10.200.10.0/24 via 10.11.20.2, Te        3/21 120/1 00:01:14
R  10.300.10.0/24 via 10.11.20.2, Te        3/21 120/1 00:01:14
C  192.168.1.0/24 Direct, Te                3/23 0/0   00:06:53
C  192.168.2.0/24 Direct, Te                3/24 0/0   00:06:26
Core3#
```

The following example shows the `show ip protocols` command to show the RIP configuration activity on Core 3.

```
Core3#show ip protocols

Routing Protocol is "RIP"
  Sending updates every 30 seconds, next due in 6
  Invalid after 180 seconds, hold down 180, flushed after 240
  Output delay 8 milliseconds between packets
  Automatic network summarization is in effect
  Outgoing filter for all interfaces is
  Incoming filter for all interfaces is
  Default redistribution metric is 1
  Default version control: receive version 2, send version 2
      Interface Recv Send
      TenGigabitEthernet 3/21 2 2
      TenGigabitEthernet 3/11 2 2
      TenGigabitEthernet 3/24 2 2
      TenGigabitEthernet 3/23 2 2
Routing for Networks:
      10.11.20.0
      10.11.30.0
      192.168.2.0
      192.168.1.0

Routing Information Sources:
  Gateway     Distance  Last Update
  10.11.20.2  120       00:00:22

Distance: (default is 120)

Core3#
```

RIP Configuration Summary

The following example shows viewing the RIP configuration on Core 2.

```
!
interface TenGigabitEthernet 2/1
  ip address 10.11.10.1/24
  no shutdown
!
interface TenGigabitEthernet 2/3
  ip address 10.11.20.2/24
  no shutdown
!
interface TenGigabitEthernet 2/4
  ip address 10.200.10.1/24
  no shutdown
!
interface TenGigabitEthernet 2/5
  ip address 10.250.10.1/24
  no shutdown
router rip
version 2
10.200.10.0
```

```
10.300.10.0
10.11.10.0
10.11.20.0
```

The following example shows viewing the RIP configuration on Core 3.

```
!
interface TenGigabitEthernet 3/1
  ip address 10.11.30.1/24
  no shutdown

!
interface TenGigabitEthernet 3/2
  ip address 10.11.20.1/24
  no shutdown

!
interface TenGigabitEthernet 3/4
  ip address 192.168.1.1/24
  no shutdown

!
interface TenGigabitEthernet 3/5
  ip address 192.168.2.1/24
  no shutdown

!
router rip
version 2
network 10.11.20.0
network 10.11.30.0
network 192.168.1.0
network 192.168.2.0
```

44

Remote Monitoring (RMON)

RMON is an industry-standard implementation that monitors network traffic by sharing network monitoring information. RMON provides both 32-bit and 64-bit monitoring facility and long-term statistics collection on Dell EMC Networking Ethernet interfaces.

RMON operates with the simple network management protocol (SNMP) and monitors all nodes on a local area network (LAN) segment. RMON monitors traffic passing through the router and segment traffic not destined for the router. The monitored interfaces may be chosen by using alarms and events with standard management information bases (MIBs).

Topics:

- Implementation Information
- Fault Recovery

Implementation Information

Configure SNMP prior to setting up RMON.

For a complete SNMP implementation description, refer to Simple Network Management Protocol (SNMP).

Configuring RMON requires using the RMON CLI and includes the following tasks:

- Setting the rmon Alarm
- Configuring an RMON Event
- Configuring RMON Collection Statistics
- Configuring the RMON Collection History

RMON implements the following standard request for comments (RFCs) (for more information, refer to the Standards Compliance chapter).

- RFC-2819
- RFC-3273
- RFC-3434
- RFC-4502

Fault Recovery

RMON provides the following fault recovery functions.

- **Interface Down** — When an RMON-enabled interface goes down, monitoring continues. However, all data values are registered as 0xFFFFFFFF (32 bits) or ixFFFFFFFFFFFFFFFF (64 bits). When the interface comes back up, RMON monitoring processes resumes.

  NOTE: A network management system (NMS) should be ready to interpret a down interface and plot the interface performance graph accordingly.
- **Device Down** — When a device goes down, all sampled data is lost. But the RMON configurations are saved in the configuration file. The sampling process continues after the chassis returns to operation.
- **Platform Adaptation** — RMON supports all Dell EMC Networking chassis and all Dell EMC Networking Ethernet interfaces.

Setting the RMON Alarm

To set an alarm on any MIB object, use the `rmon alarm` or `rmon hc-alarm` command in GLOBAL CONFIGURATION mode.

- Set an alarm on any MIB object.
  CONFIGURATION mode

`[no] rmon alarm number variable interval {delta | absolute} rising-threshold [value event-number] falling-threshold value event-number [owner string]`

OR

`[no] rmon hc-alarm number variable interval {delta | absolute} rising-threshold value event-number falling-threshold value event-number [owner string]`

Configure the alarm using the following optional parameters:

- *number*: alarm number, an integer from 1 to 65,535, the value must be unique in the RMON Alarm Table.
- *variable*: the MIB object to monitor — the variable must be in SNMP OID format; for example, 1.3.6.1.2.1.1.3. The object type must be a 32-bit integer for the `rmon alarm` command and 64 bits for the `rmon hc-alarm` command.
- *interval*: time in seconds the alarm monitors the MIB variable, the value must be between 1 to 3,600.
- `delta`: tests the change between MIB variables, this option is the alarmSampleType in the RMON Alarm table.
- `absolute`: tests each MIB variable directly, this option is the alarmSampleType in the RMON Alarm table.
- `rising-threshold value`: value at which the rising-threshold alarm is triggered or reset. For the `rmon alarm` command, this setting is a 32-bits value, for the `rmon hc-alarm` command, this setting is a 64-bits value.
- *event-number*: event number to trigger when the rising threshold exceeds its limit. This value is identical to the alarmRisingEventIndex in the alarmTable of the RMON MIB. If there is no corresponding rising-threshold event, the value should be zero.
- `falling-threshold value`: value at which the falling-threshold alarm is triggered or reset. For the `rmon alarm` command, this setting is a 32-bits value, for the `rmon hc-alarm` command this setting is a 64 bits value.
- *event-number*: event number to trigger when the falling threshold exceeds its limit. This value is identical to the alarmFallingEventIndex in the alarmTable of the RMON MIB. If there is no corresponding falling-threshold event, the value should be zero.
- `owner string`: (Optional) specifies an owner for the alarm, this setting is the alarmOwner object in the alarmTable of the RMON MIB. Default is a **null-terminated string**.

To disable the alarm, use the `no` form of the command.

The following example configures RMON alarm number 10. The alarm monitors the MIB variable 1.3.6.1.2.1.2.2.1.20.1 (ifEntry.ifOutErrors) once every 20 seconds until the alarm is disabled, and checks the rise or fall of the variable. The alarm is triggered when the 1.3.6.1.2.1.2.2.1.20.1 value shows a MIB counter increase of 15 or more (such as from 100000 to 100015). The alarm then triggers event number 1, which is configured with the RMON event command. Possible events include a log entry or an SNMP trap. If the 1.3.6.1.2.1.2.2.1.20.1 value changes to 0 (falling-threshold 0), the alarm is reset and can be triggered again.

```
DellEMC(conf)#rmon alarm 10 1.3.6.1.2.1.2.2.1.20.1 20 delta rising-threshold
15 1 falling-threshold 0 1 owner nms1
```

Configuring an RMON Event

To add an event in the RMON event table, use the `rmon event` command in GLOBAL CONFIGURATION mode.

- Add an event in the RMON event table.

  CONFIGURATION mode

  `[no] rmon event number [log] [trap community] [description string] [owner string]`

  - *number*: assigned event number, which is identical to the eventIndex in the eventTable in the RMON MIB. The value must be an integer from 1 to 65,535 and be unique in the RMON Event Table.
  - *log*: (Optional) generates an RMON log entry when the event is triggered and sets the eventType in the RMON MIB to log or log-and-trap. Default is no log.
  - `trap community`: (Optional) SNMP community string used for this trap. Configures the setting of the eventType in the RMON MIB for this row as either snmp-trap snmptrap or log-and-trap. This value is identical to the eventCommunityValue in the eventTable in the RMON MIB. Default is public.
  - `description string`: (Optional) specifies a description of the event, which is identical to the event description in the eventTable of the RMON MIB. The default is a **null-terminated string**.
  - `owner string`: (Optional) owner of this event, which is identical to the eventOwner in the eventTable of the RMON MIB. Default is a **null-terminated string**.

To disable RMON on the interface, use the `no` form of this command.

In the following example, the configuration creates RMON event number 1, with the description “High ifOutErrors”, and generates a log entry when an alarm triggers the event. The user *nms1* owns the row that is created in the event table by

this command. This configuration also generates an SNMP trap when the event is triggered using the SNMP community string "eventtrap".

```
DellEMC(conf)#rmon event 1 log trap eventtrap description "High ifOutErrors" owner nms1
```

Configuring RMON Collection Statistics

To enable RMON MIB statistics collection on an interface, use the `RMON collection statistics` command in INTERFACE CONFIGURATION mode.

- Enable RMON MIB statistics collection.

  CONFIGURATION INTERFACE (config-if) mode

  `[no] rmon collection statistics {controlEntry` *`integer`*`} [owner` *`ownername`*`]`

  - `controlEntry`: specifies the RMON group of statistics using a value.
  - *`integer`*: a value from 1 to 65,535 that identifies the RMON Statistics Table. The value must be unique in the RMON Statistic Table.
  - `owner`: (Optional) specifies the name of the owner of the RMON group of statistics.
  - *`ownername`*: (Optional) records the name of the owner of the RMON group of statistics. The default is a **null-terminated string**.

To remove a specified RMON statistics collection, use the `no` form of this command.

The following command example enables the RMON statistics collection on the interface, with an ID value of 20 and an owner of *john*.

```
DellEMC(conf-if-mgmt)#rmon collection statistics controlEntry 20 owner john
```

Configuring the RMON Collection History

To enable the RMON MIB history group of statistics collection on an interface, use the `rmon collection history` command in INTERFACE CONFIGURATION mode.

- Configure the RMON MIB history group of statistics collection.

  CONFIGURATION INTERFACE (config-if) mode

  `[no] rmon collection history {controlEntry` *`integer`*`} [owner` *`ownername`*`] [buckets` *`bucket-number`*`] [interval` *`seconds`*`]`

  - `controlEntry`: specifies the RMON group of statistics using a value.
  - *`integer`*: a value from 1 to 65,535 that identifies the RMON group of statistics. The value must be a unique index in the RMON History Table.
  - `owner`: (Optional) specifies the name of the owner of the RMON group of statistics. The default is a **null-terminated string**.
  - *`ownername`*: (Optional) records the name of the owner of the RMON group of statistics.
  - `buckets`: (Optional) specifies the maximum number of buckets desired for the RMON collection history group of statistics.
  - *`bucket-number`*: (Optional) a value associated with the number of buckets specified for the RMON collection history group of statistics. The value is limited to from 1 to 1000. The default is **50** (as defined in RFC-2819).
  - `interval`: (Optional) specifies the number of seconds in each polling cycle.
  - `seconds`: (Optional) the number of seconds in each polling cycle. The value is ranged from 5 to 3,600 (Seconds). The default is **1,800** (as defined in RFC-2819).

To remove a specified RMON history group of statistics collection, use the `no` form of this command.

The following command example enables an RMON MIB collection history group of statistics with an ID number of 20 and an owner of *john*, both the sampling interval and the number of buckets use their respective defaults.

```
DellEMC(conf-if-mgmt)#rmon collection history controlEntry 20 owner john
```

45

Rapid Spanning Tree Protocol (RSTP)

The Rapid Spanning Tree Protocol (RSTP) is a Layer 2 protocol — specified by IEEE 802.1w — that is essentially the same as spanning-tree protocol (STP) but provides faster convergence and interoperability with switches configured with STP and multiple spanning tree protocol (MSTP).

Topics:

- Protocol Overview
- Configuring Rapid Spanning Tree
- Important Points to Remember
- Configuring Interfaces for Layer 2 Mode
- Enabling Rapid Spanning Tree Protocol Globally
- Adding and Removing Interfaces
- Modifying Global Parameters
- Modifying Interface Parameters
- Enabling SNMP Traps for Root Elections and Topology Changes
- Influencing RSTP Root Selection
- Configuring an EdgePort
- Configuring Fast Hellos for Link State Detection

Protocol Overview

RSTP is a Layer 2 protocol — specified by IEEE 802.1w — that is essentially the same as spanning-tree protocol (STP) but provides faster convergence and interoperability with switches configured with STP and multiple spanning tree protocol (MSTP).

The Dell EMC Networking OS supports three other variations of spanning tree, as shown in the following table.

Table 82. Spanning Tree Variations Dell EMC Networking OS Supports

Dell EMC Networking Term	IEEE Specification
Spanning Tree Protocol (STP)	802.1d
Rapid Spanning Tree Protocol (RSTP)	802.1w
Multiple Spanning Tree Protocol (MSTP)	802.1s
Per-VLAN Spanning Tree Plus (PVST+)	Third Party

Configuring Rapid Spanning Tree

Configuring RSTP is a two-step process.

1. Configure interfaces for Layer 2.
2. Enable the rapid spanning tree protocol.

Related Configuration Tasks

- Adding and Removing Interfaces
- Modifying Global Parameters
- Modifying Interface Parameters
- Configuring an EdgePort

- Prevent Network Disruptions with BPDU Guard
- Influencing RSTP Root Selection
- Configuring Spanning Trees as Hitless
- Enabling SNMP Traps for Root Elections and Topology Changes
- Configuring Fast Hellos for Link State Detection
- Flush MAC Addresses after a Topology Change

Important Points to Remember

- RSTP is disabled by default.
- Dell EMC Networking OS supports only one Rapid Spanning Tree (RST) instance.
- All interfaces in virtual local area networks (VLANs) and all enabled interfaces in Layer 2 mode are automatically added to the RST topology.
- Adding a group of ports to a range of VLANs sends multiple messages to the rapid spanning tree protocol (RSTP) task, avoid using the `range` command. When using the `range` command, Dell EMC Networking recommends limiting the range to five ports and 40 VLANs.

RSTP and VLT

Virtual link trunking (VLT) provides loop-free redundant topologies and does not require RSTP.

RSTP can cause temporary port state blocking and may cause topology changes after link or node failures. Spanning tree topology changes are distributed to the entire Layer 2 network, which can cause a network-wide flush of learned media access control (MAC) and address resolution protocol (ARP) addresses, requiring these addresses to be re-learned. However, enabling RSTP can detect potential loops caused by non-system issues such as cabling errors or incorrect configurations. RSTP is useful for potential loop detection but to minimize possible topology changes after link or node failure, configure it using the following specifications.

The following recommendations help you avoid these issues and the associated traffic loss caused by using RSTP when you enable VLT on both VLT peers:

- Configure any ports at the edge of the spanning tree's operating domain as edge ports, which are directly connected to end stations or server racks. Ports connected directly to Layer 3-only routers not running STP should have RSTP disabled or be configured as edge ports.
- Ensure that the primary VLT node is the root bridge and the secondary VLT peer node has the second-best bridge ID in the network. If the primary VLT peer node fails, the secondary VLT peer node becomes the root bridge, avoiding problems with spanning tree port state changes that occur when a VLT node fails or recovers.
- Even with this configuration, if the node has non-VLT ports using RSTP that are not configured as edge ports and are connected to other layer 2 switches, spanning tree topology changes can still be detected after VLT node recovery. To avoid this scenario, ensure that you configure any non-VLT ports as edge ports or have RSTP disabled.

Configuring Interfaces for Layer 2 Mode

To configure and enable interfaces in Layer 2 mode, use the following commands.

All interfaces on all bridges that participate in Rapid Spanning Tree must be in Layer 2 and enabled.

1. If the interface has been assigned an IP address, remove it.

   INTERFACE mode

   ```
   no ip address
   ```

2. Place the interface in Layer 2 mode.

   INTERFACE mode

   ```
   switchport
   ```

3. Enable the interface.

   INTERFACE mode

   ```
   no shutdown
   ```

To verify that an interface is in Layer 2 mode and enabled, use the `show config` command from INTERFACE mode. The bold lines indicate that the interface is in Layer 2 mode.

Enabling Rapid Spanning Tree Protocol Globally

Enable RSTP globally on all participating bridges; it is not enabled by default.

When you enable RSTP, all physical and port-channel interfaces that are enabled and in Layer 2 mode are automatically part of the RST topology.

- Only one path from any bridge to any other bridge is enabled.
- Bridges block a redundant path by disabling one of the link ports.

To enable RSTP globally for all Layer 2 interfaces, use the following commands.

1. Enter PROTOCOL SPANNING TREE RSTP mode.

   CONFIGURATION mode

   ```
   protocol spanning-tree rstp
   ```

2. Enable RSTP.

   PROTOCOL SPANNING TREE RSTP mode

   ```
   no disable
   ```

To disable RSTP globally for all Layer 2 interfaces, enter the `disable` command from PROTOCOL SPANNING TREE RSTP mode.

To verify that RSTP is enabled, use the `show config` command from PROTOCOL SPANNING TREE RSTP mode. The bold line indicates that RSTP is enabled.

```
DellEMC(conf-rstp)#show config
!
protocol spanning-tree rstp
no disable
DellEMC(conf-rstp)#
```

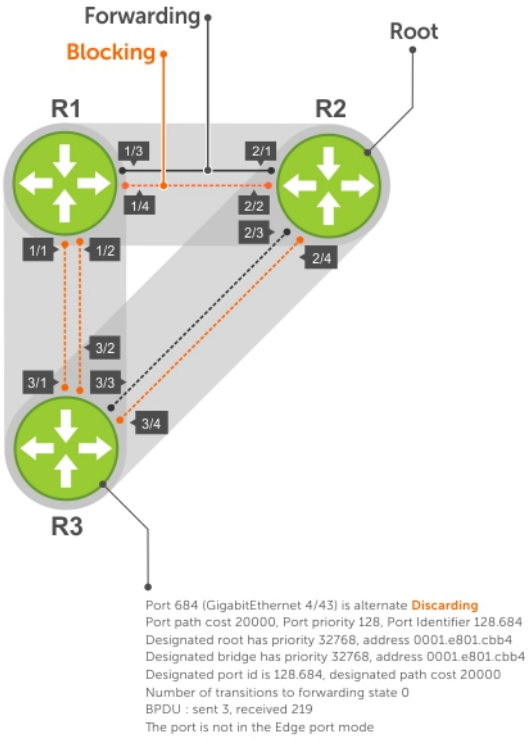


Figure 117. Rapid Spanning Tree Enabled Globally

To view the interfaces participating in RSTP, use the `show spanning-tree rstp` command from EXEC privilege mode. If a physical interface is part of a port channel, only the port channel is listed in the command output.

```
DellEMC#show spanning-tree rstp
Root Identifier has priority 32768, Address 0001.e801.cbb4
Root Bridge hello time 2, max age 20, forward delay 15, max hops 0
Bridge Identifier has priority 32768, Address 0001.e801.cbb4
Configured hello time 2, max age 20, forward delay 15, max hops 0
We are the root
Current root has priority 32768, Address 0001.e801.cbb4
Number of topology changes 4, last change occurred 00:02:17 ago on Te 1/26

Port 377 (TenGigabitEthernet 2/1) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.377
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.377, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 9
The port is not in the Edge port mode

Port 378 (TenGigabitEthernet 2/2) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.378
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.378, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 2
The port is not in the Edge port mode

Port 379 (TenGigabitEthernet 2/3) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.379
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.379, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 5
The port is not in the Edge port mode

Port 380 (TenGigabitEthernet 2/4) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.380
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.380, designated path cost 0

Number of transitions to forwarding state 1
BPDU : sent 147, received 3
The port is not in the Edge port mode
```

To confirm that a port is participating in RSTP, use the `show spanning-tree rstp brief` command from EXEC privilege mode.

```
R3#show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID Priority 32768, Address 0001.e801.cbb4
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID Priority 32768, Address 0001.e80f.1dad
Configured hello time 2, max age 20, forward delay 15
Interface                              Designated
Name    PortID  Prio Cost   Sts Cost  Bridge ID             PortID
---------- -------- ----   ------- --- ------- ------------- --------
Te 3/1  128.681 128  20000  BLK 20000 32768 0001.e80b.88bd  128.469
Te 3/2  128.682 128  20000  BLK 20000 32768 0001.e80b.88bd  128.470
Te 3/3  128.683 128  20000  FWD 20000 32768 0001.e801.cbb4  128.379
Te 3/4  128.684 128  20000  BLK 20000 32768 0001.e801.cbb4  128.380
Interface
Name    Role  PortID  Prio Cost  Sts Cost   Link-type Edge
------- --- ------ -------- ---- ------- --- -----------
Te 3/1  Altr  128.681 128 20000  BLK 20000  P2P       No
Te 3/2  Altr  128.682 128 20000  BLK 20000  P2P       No
Te 3/3  Root  128.683 128 20000  FWD 20000  P2P       No
```

```
Te 3/4   Altr   128.684 128 20000   BLK 20000   P2P         No
R3#
```

Adding and Removing Interfaces

To add and remove interfaces, use the following commands.

To add an interface to the Rapid Spanning Tree topology, configure it for Layer 2 and it is automatically added. If you previously disabled RSTP on the interface using the command `no spanning-tree 0` command, re-enable it using the `spanning-tree 0` command.

- Remove an interface from the Rapid Spanning Tree topology.

  `no spanning-tree 0`

Modifying Global Parameters

You can modify RSTP parameters.

The root bridge sets the values for forward-delay, hello-time, and max-age and overwrites the values set on other bridges participating in the Rapid Spanning Tree group.

- **Forward-delay** — the amount of time an interface waits in the Listening state and the Learning state before it transitions to the Forwarding state.
- **Hello-time** — the time interval in which the bridge sends RSTP BPDUs.
- **Max-age** — the length of time the bridge maintains configuration information before it refreshes that information by recomputing the RST topology.

NOTE: Dell EMC Networking recommends that only experienced network administrators change the Rapid Spanning Tree group parameters. Poorly planned modification of the RSTP parameters can negatively affect network performance.

The following table displays the default values for RSTP.

Table 83. RSTP Default Values

RSTP Parameter	Default Value
Forward Delay	**15 seconds**
Hello Time	**2 seconds**
Max Age	**20 seconds**
Port Cost: • 100-Mb/s Ethernet interfaces • 1-Gigabit Ethernet interfaces • 10-Gigabit Ethernet interfaces • 40-Gigabit Ethernet interfaces • Port Channel with 100 Mb/s Ethernet interfaces • Port Channel with 1-Gigabit Ethernet interfaces • Port Channel with 10-Gigabit Ethernet interfaces • Port Channel with 40-Gigabit Ethernet interfaces	Port Cost: • **200000** • **20000** • **2000** • **1400** • **180000** • **18000** • **1800** • **600**
Port Priority	**128**

To change these parameters, use the following commands.

- Change the forward-delay parameter.

  PROTOCOL SPANNING TREE RSTP mode

  `forward-delay` *`seconds`*

  The range is from 4 to 30.

  The default is **15 seconds**.
- Change the hello-time parameter.

  PROTOCOL SPANNING TREE RSTP mode

`hello-time seconds`

NOTE: With large configurations (especially those configurations with more ports) Dell EMC Networking recommends increasing the hello-time.

The range is from 1 to 10.

The default is **2 seconds**.

- Change the max-age parameter.

  PROTOCOL SPANNING TREE RSTP mode

  `max-age seconds`

  The range is from 6 to 40.

  The default is **20 seconds**.

To view the current values for global parameters, use the `show spanning-tree rstp` command from EXEC privilege mode.

Enabling SNMP Traps for Root Elections and Topology Changes

To enable SNMP traps, use the following command.

- Enable SNMP traps for RSTP, MSTP, and PVST+ collectively.

  `snmp-server enable traps xstp`

Modifying Interface Parameters

On interfaces in Layer 2 mode, you can set the port cost and port priority values.

- **Port cost** — a value that is based on the interface type. The previous table lists the default values. The greater the port cost, the less likely the port is selected to be a forwarding port.
- **Port priority** — influences the likelihood that a port is selected to be a forwarding port in case that several ports have the same port cost.

To change the port cost or priority of an interface, use the following commands.

- Change the port cost of an interface.

  INTERFACE mode

  `spanning-tree rstp cost cost`

  The range is from 0 to 65535.

  The default is listed in the previous table.

- Change the port priority of an interface.

  INTERFACE mode

  `spanning-tree rstp priority priority-value`

  The range is from 0 to 15.

  The default is **128**.

To view the current values for interface parameters, use the `show spanning-tree rstp` command from EXEC privilege mode.

Enabling SNMP Traps for Root Elections and Topology Changes

To enable SNMP traps collectively, use this command.

Enable SNMP traps for RSTP, MSTP, and PVST+ collectively.

`snmp-server enable traps xstp`

Influencing RSTP Root Selection

RSTP determines the root bridge, but you can assign one bridge a lower priority to increase the likelihood that it is selected as the root bridge.

To change the bridge priority, use the following command.

- Assign a number as the bridge priority or designate it as the primary or secondary root.

  PROTOCOL SPANNING TREE RSTP mode

  `bridge-priority` *`priority-value`*

  - *`priority-value`* The range is from 0 to 65535. The lower the number assigned, the more likely this bridge becomes the root bridge.

  The default is **32768**. Entries must be multiples of 4096.

A console message appears when a new root bridge has been assigned. The following example example shows the console message after the `bridge-priority` command is used to make R2 the root bridge (shown in bold).

```
DellEMC(conf-rstp)#bridge-priority 4096
04:27:59: %RPM0-P:RP2 %SPANMGR-5-STP_ROOT_CHANGE: RSTP root changed. My Bridge ID:
4096:0001.e80b.88bd Old Root: 32768:0001.e801.cbb4 New Root: 4096:0001.e80b.88bd
```

Configuring an EdgePort

The EdgePort feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner.

In this mode an interface forwards frames by default until it receives a BPDU that indicates that it should behave otherwise; it does not go through the Learning and Listening states. The `bpduguard shutdown-on-violation` option causes the interface hardware to be shut down when it receives a BPDU. When only `bpduguard` is implemented, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree drops packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation. This feature is the same as PortFast mode in Spanning Tree.

CAUTION: Configure EdgePort only on links connecting to an end station. If you enable EdgePort on an interface connected to a network, it can cause loops.

Dell EMC Networking OS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

- If the interface to be shut down is a port channel, all the member ports are disabled in the hardware.
- When you add a physical port to a port channel already in the Error Disable state, the new member port is also disabled in the hardware.
- When you remove a physical port from a port channel in the Error Disable state, the error disabled state is cleared on this physical port (the physical port is enabled in the hardware).
- You can clear the Error Disabled state with any of the following methods:
  - Perform an `shutdown` command on the interface.
  - Disable the `shutdown-on-violation` command on the interface (the `no spanning-tree stp-id portfast [bpduguard | [shutdown-on-violation]]` command).
  - Disable spanning tree on the interface (the `no spanning-tree` command in INTERFACE mode).
  - Disable global spanning tree (the `no spanning-tree` command in CONFIGURATION mode).

To enable EdgePort on an interface, use the following command.

- Enable EdgePort on an interface.

  INTERFACE mode

  `spanning-tree rstp edge-port [bpduguard | shutdown-on-violation]`

To verify that EdgePort is enabled on a port, use the `show spanning-tree rstp` command from EXEC privilege mode or the `show config` command from INTERFACE mode.

NOTE: Dell EMC Networking recommends using the `show config` command from INTERFACE mode.

In the following example, the bold line indicates that the interface is in EdgePort mode.

```
DellEMC(conf-if-te-2/1)#show config
!
```

```
interface TenGigabitEthernet 2/1
  no ip address
  switchport
  spanning-tree rstp edge-port
  shutdown
DellEMC(conf-if-te-2/1)#
```

Configuring Fast Hellos for Link State Detection

Use RSTP fast hellos to achieve sub-second link-down detection so that convergence is triggered faster. The standard RSTP link-state detection mechanism does not offer the same low link-state detection speed.

To achieve sub-second link-down detection so that convergence is triggered faster, use RSTP fast hellos. The standard RSTP link-state detection mechanism does not offer the same low link-state detection speed.

RSTP fast hellos decrease the hello interval to the order of milliseconds and all timers derived from the hello timer are adjusted accordingly. This feature does not inter-operate with other vendors, and is available only for RSTP.

- Configure a hello time on the order of milliseconds.

  PROTOCOL RSTP mode

  `hello-time milli-second` *`interval`*

  The range is from 50 to 950 milliseconds.

```
DellEMC(conf-rstp)#do show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
  Root ID     Priority 0, Address 0001.e811.2233
  Root Bridge hello time 50 ms, max age 20, forward delay 15
  Bridge ID   Priority 0, Address 0001.e811.2233
  We are the root
  Configured hello time 50 ms, max age 20, forward delay 15
```

NOTE: The hello time is encoded in BPDUs in increments of 1/256ths of a second. The standard minimum hello time in seconds is 1 second, which is encoded as 256. Millisecond. hello times are encoded using values less than 256; the millisecond hello time equals (x/1000)*256. When you configure millisecond hellos, the default hello interval of 2 seconds is still used for edge ports; the millisecond hello interval is not used.

46

Software-Defined Networking (SDN)

The Dell EMC Networking OS supports software-defined networking (SDN). For more information, see the *SDN Deployment Guide*.

47

Security

This chapter describes several ways to provide security to the Dell EMC Networking system.

For details about all the commands described in this chapter, refer to the *Security* chapter in the *Dell EMC Networking OS Command Reference Guide*.

Topics:

- AAA Accounting
- AAA Authentication
- Obscuring Passwords and Keys
- AAA Authorization
- RADIUS
- TACACS+
- Protection from TCP Tiny and Overlapping Fragment Attacks
- Enabling SCP and SSH
- Telnet
- VTY Line and Access-Class Configuration
- Support for Change of Authorization and Disconnect Messages packets
- Role-Based Access Control
- Two Factor Authentication (2FA)
- Configuring the System to Drop Certain ICMP Reply Messages

AAA Accounting

Accounting, authentication, and authorization (AAA) accounting is part of the AAA security model.

For details about commands related to AAA security, refer to the *Security* chapter in the *Dell EMC Networking OS Command Reference Guide*.

AAA accounting enables tracking of services that users are accessing and the amount of network resources being consumed by those services. When you enable AAA accounting, the network server reports user activity to the security server in the form of accounting records. Each accounting record comprises accounting attribute/value (AV) pairs and is stored on the access control server.

As with authentication and authorization, you must configure AAA accounting by defining a named list of accounting methods and then applying that list to various virtual terminal line (VTY) lines.

Configuration Task List for AAA Accounting

The following sections present the AAA accounting configuration tasks.

- Enabling AAA Accounting (mandatory)
- Suppressing AAA Accounting for Null Username Sessions (optional)
- Configuring Accounting of EXEC and Privilege-Level Command Usage (optional)
- Configuring AAA Accounting for Terminal Lines (optional)
- Monitoring AAA Accounting (optional)

Enabling AAA Accounting

The `aaa accounting` command allows you to create a record for any or all of the accounting functions monitored.

To enable AAA accounting, use the following command.

- Enable AAA accounting and create a record for monitoring the accounting function.

CONFIGURATION mode

```
aaa accounting {commands level | dot1x | exec | rest | suppress | system} {default | name} {start-stop | wait-start | stop-only} {radius | tacacs+}
```

The variables are:

- `system`: sends accounting information of any other AAA configuration.
- `exec`: sends accounting information when a user has logged in to EXEC mode.
- `dot1x`: sends accounting information when a dot1x user has logged in to EXEC mode.
- `command level`: sends accounting of commands executed at the specified privilege level.
- `suppress`: Do not generate accounting records for a specific type of user.
- `default` | `name`: enter the name of a list of accounting methods.
- `start-stop`: use for more accounting information, to send a start-accounting notice at the beginning of the requested event and a stop-accounting notice at the end.
- `wait-start`: ensures that the TACACS+ security server acknowledges the start notice before granting the user's process request.
- `stop-only`: use for minimal accounting; instructs the TACACS+ server to send a stop record accounting notice at the end of the requested user process.
- `tacacs+`: designate the security service as TACACS+.
- `radius`: designate the security service as RADIUS.

Suppressing AAA Accounting for Null Username Sessions

When you activate AAA accounting, the Dell EMC Networking OS software issues accounting records for all users on the system, including users whose username string is NULL because of protocol translation.

An example of this is a user who comes in on a line where the AAA authentication `login method-list none` command is applied. To prevent accounting records from being generated for sessions that do not have usernames associated with them, use the following command.

- Prevent accounting records from being generated for users whose username string is NULL.

  CONFIGURATION mode

  ```
  aaa accounting suppress null-username
  ```

Configuring TACACS+ Accounting of EXEC and Privilege-Level Command Usage

The network access server monitors the accounting functions defined in the TACACS+ attribute/value (AV) pairs.

- Configure AAA accounting to monitor accounting functions defined in TACACS+.

  CONFIGURATION mode

  ```
  aaa accounting system default start-stop tacacs+
  aaa accounting command 15 default start-stop tacacs+
  ```

  System accounting can use only the default method list.

In the following sample configuration, AAA accounting is set to track all usage of EXEC commands and commands on privilege level 15.

```
DellEMC(conf)#aaa accounting exec default start-stop tacacs+
DellEMC(conf)#aaa accounting command 15 default start-stop tacacs+
```

Configuring AAA Accounting for Terminal Lines

To enable AAA accounting with a named method list for a specific terminal line (where *com15* and *execAcct* are the method list names), use the following commands.

- Configure AAA accounting for terminal lines.

  CONFIG-LINE-VTY mode

  ```
  accounting commands 15 com15
  ```

```
accounting exec execAcct
```

```
DellEMC(config-line-vty)# accounting commands 15 com15
DellEMC(config-line-vty)# accounting exec execAcct
```

Monitoring AAA Accounting

Dell EMC Networking OS does not support periodic interim accounting because the `periodic` command can cause heavy congestion when many users are logged in to the network.

No specific `show` command exists for TACACS+ accounting.

To obtain accounting records displaying information about users currently logged in, use the following command.

- Step through all active sessions and print all the accounting records for the actively accounted functions.

  CONFIGURATION mode or EXEC Privilege mode

  ```
  show accounting
  ```

```
DellEMC#show accounting
Active accounted actions on tty2, User admin Priv 1
  Task ID 1, EXEC Accounting record, 00:00:39 Elapsed, service=shell
Active accounted actions on tty3, User admin Priv 1
  Task ID 2, EXEC Accounting record, 00:00:26 Elapsed, service=shell
DellEMC#
```

RADIUS Accounting

Dell EMC Networking OS supports Remote Authentication Dial In User Service (RADIUS) protocol to transmit the RADIUS accounting messages between a Network Access Server (NAS) and an accounting server.

NAS reports the user activity to the accounting server (RADIUS or TACACS+) with accounting records. The RADIUS accounting server stores the accounting records, which is used for network management, auditing, etc.

Dell EMC Networking OS complies with RFC2866 for RADIUS Accounting.

NAS receives the accounting request from the supplicant and sends the RADIUS request packet to the accounting server after successful authentication. The RADIUS Accounting request contains a RADIUS Acct-Status-Type as `Start` or `Stop` to update the supplicant session to the accounting server.

NOTE: In RADIUS accounting, fallback behavior among RADIUS and TACACS servers is not supported as the RADIUS accounting feature is not available in Dell EMC Networking OS version earlier than 9.14.1.5.

In VLT domain, the NAS sends the RADIUS Accounting Request packets only if the NAS is configured as a VLT primary peer.

Configure RADIUS Accounting

The NAS monitors the accounting functions defined in the RADIUS Accounting attribute/value (AV) pairs.

- Configure AAA accounting to monitor accounting functions defined in RADIUS.

  CONFIGURATION mode

  ```
  aaa accounting {dot1x | exec} default {start-stop | wait-start | stop-only} radius
  ```

In the following sample configuration, AAA accounting is set to track all usage of EXEC commands and commands on privilege level 15.

```
Dell(conf)# aaa accounting dot1x default start-stop radius
Dell(conf)# aaa accounting exec default stop-only radius
```

Sample dot1x accounting records

The following lists the sample EAP and MAB accounting records

EAP START accounting record:

```
Fri May 10 12:20:43 2019
        NAS-IP-Address = 10.16.133.49
        NAS-Port-Type = Ethernet
        NAS-Port = 1010
        Calling-Station-Id = "1E-3C-39-B3-00-00"
        User-Name = "testuser1"
        NAS-Port-Id = "GigabitEthernet 1/11"
        Service-Type = Framed-User
        Acct-Session-Id = "1e-3c-39-b3-00-00-2"
        Acct-Multi-Session-Id = "1e-3c-39-b3-00-00-00-11-33-44-77-88-6c-b3-d5-5cc"
        Acct-Status-Type = Start
        Event-Timestamp = "May 10 2019 12:20:43 CDT"
        Tmp-String-9 = "ai:"
        Acct-Unique-Session-Id = "2d6c5beef615d18fa21bbde29411f6d5"
        Timestamp = 1557508843
```

EAP STOP accounting record:

```
Fri May 10 12:22:15 2019
        NAS-IP-Address = 10.16.133.49
        NAS-Port-Type = Ethernet
        NAS-Port = 1010
        Calling-Station-Id = "1E-3C-39-B3-00-00"
        User-Name = "testuser1"
        NAS-Port-Id = "GigabitEthernet 1/11"
        Service-Type = Framed-User
        Acct-Session-Time = 92
        Acct-Session-Id = "1e-3c-39-b3-00-00-2"
        Acct-Multi-Session-Id = "1e-3c-39-b3-00-00-00-11-33-44-77-88-6c-b3-d5-5cc"
        Acct-Link-Count = 1
        Acct-Terminate-Cause = User-Request
        Acct-Status-Type = Stop
        Event-Timestamp = "May 10 2019 12:22:15 CDT"
        Tmp-String-9 = "ai:"
        Acct-Unique-Session-Id = "2d6c5beef615d18fa21bbde29411f6d5"
        Timestamp = 1557508935
```

MAB START record:

```
Fri May 10 23:30:21 2019
        User-Name = "001122334455"
        Called-Station-Id = "00-11-33-44-77-88"
        Calling-Station-Id = "00-11-22-33-44-55"
        NAS-IP-Address = 10.16.133.49
        NAS-Port-Type = Ethernet
        NAS-Port = 1010
        NAS-Port-Id = "GigabitEthernet 1/11"
        Service-Type = Call-Check
        Acct-Session-Id = "00-11-22-33-44-55-4"
        Acct-Multi-Session-Id = "00-11-22-33-44-55-00-11-33-44-77-88-5e-50-d6-5cc"
        Acct-Status-Type = Start
        Event-Timestamp = "May 10 2019 23:30:21 CDT"
        Tmp-String-9 = "ai:"
        Acct-Unique-Session-Id = "5a761462ef63b815707de5fa1c5ef348"
        Timestamp = 1557549021
```

MAB STOP record:

```
Fri May 10 23:30:42 2019
        User-Name = "001122334455"
        Called-Station-Id = "00-11-33-44-77-88"
        Calling-Station-Id = "00-11-22-33-44-55"
        NAS-IP-Address = 10.16.133.49
        NAS-Port-Type = Ethernet
        NAS-Port = 1010
        NAS-Port-Id = "GigabitEthernet 1/11"
        Service-Type = Call-Check
        Acct-Session-Time = 21
        Acct-Session-Id = "00-11-22-33-44-55-4"
```

```
        Acct-Multi-Session-Id = "00-11-22-33-44-55-00-11-33-44-77-88-5e-50-d6-5cc"
        Acct-Link-Count = 1
        Acct-Terminate-Cause = Lost-Carrier
        Acct-Status-Type = Stop
        Event-Timestamp = "May 10 2019 23:30:42 CDT"
        Tmp-String-9 = "ai:"
        Acct-Unique-Session-Id = "5a761462ef63b815707de5fa1c5ef348"
        Timestamp = 1557549042
```

RADIUS Accounting attributes

The following tables describe the various types of attributes that identify the supplicant sessions:

Table 84. RADIUS Accounting Start Record Attributes for CLI user

RADIUS Attribute code	RADIUS Attribute	Description
NAS Identification Attributes		
4	NAS-IP-Address	IPv4 address of the NAS.
95	NAS-IPv6–Address	IPv6 address of the NAS.
Session Identification Attributes		
1	User-Name	User name.
5	NAS-Port	Port on which session is connected (CLI Session-Id).
31	Calling-Station-Id	Telnet/SSH client IP address.
Accounting Attributes		
40	Acct-Status-Type	START
44	Acct-Session-Id	CLI Session-Id - To match start and stop session requests.
61	NAS-Port-Type	ASYNC - for console session. VIRTUAL - for telnet/SSH session.

Table 85. RADIUS Accounting Stop Record Attributes for CLI user

RADIUS Attribute code	RADIUS Attribute	Description
NAS Identification Attributes		
4	NAS-IP-Address	IPv4 address of the NAS.
95	NAS-IPv6–Address	IPv6 address of the NAS.
Session Identification Attributes		
1	User-Name	User name.
5	NAS-Port	Port on which session is connected (CLI Session-Id).
6	Service-Type	NAS Prompt.
31	Calling-Station-Id	Telnet/SSH client IP address.
Accounting Attributes		
40	Acct-Status-Type	STOP
44	Acct-Session-Id	CLI Session-Id - To match start and stop session requests.
46	Acct-Session Time	Time the user has received the service.
49	Acct-Terminate-Cause	Reason for session termination.
61	NAS-Port-Type	ASYNC - for Console session.

Table 85. RADIUS Accounting Stop Record Attributes for CLI user (continued)

RADIUS Attribute code	RADIUS Attribute	Description
		VIRTUAL - for telnet/SSH session.

Table 86. Use cases for CLI user to trigger RADIUS Accounting Start/Stop records

CLI event	Accounting type	Attributes
CLI user authentication success	Start	Start record attributes for CLI user.
CLI user log-off	Stop	Stop record attributes with termination cause as User Request (1).
CLI user session idle timeout	Stop	Stop record attributes with termination cause as Idle Timeout (4).
CLI user session disconnects due to Dynamic authorization	Stop	Stop record attributes with termination cause as Admin Reset (6).

Table 87. RADIUS Accounting Start Record Attributes for dot1x supplicant

RADIUS Attribute code	RADIUS Attribute	Description
NAS Identification Attributes		
4	NAS-IP-Address	IPv4 address of the NAS.
95	NAS-IPv6–Address	IPv6 address of the NAS.
Session Identification Attributes		
1	User-Name	User name/ Supplicant MAC Address (for MAB).
5	NAS-Port	Port on which session is terminated.
6	Service-Type	Framed (2) for EAP /Call check (10) for MAB.
8	Framed-IP-Address	IPv4 address of supplicant.
168	Framed-IPV6-Address	IPv6 address of supplicant.
30	Called-Station-Id	Switch MAC Address.
31	Calling-Station-Id	Supplicant MAC Address.
Accounting Attributes		
40	Acct-Status-Type	START
44	Acct-Session-Id	<Supplicant MAC> | Running number
50	Acct-Multi-Session-Id	<Supplicant MAC>|<Switch MAC>|<Timestamp>
51	Acct-Link-Count	1
61	NAS-Port-Type	Ethernet

NOTE: Framed IP address attribute is available only when the attribute support is enabled in CLI and when the IP entry is present in the DHCP binding table.

Table 88. RADIUS Accounting Stop Record Attributes for dot1x supplicant

RADIUS Attribute code	RADIUS Attribute	Description
NAS Identification Attributes		
4	NAS-IP-Address	IPv4 address of the NAS.
95	NAS-IPv6–Address	IPv6 address of the NAS.
Session Identification Attributes		

Table 88. RADIUS Accounting Stop Record Attributes for dot1x supplicant (continued)

RADIUS Attribute code	RADIUS Attribute	Description
1	User-Name	User name/ Supplicant MAC Address (for MAB).
5	NAS-Port	Port on which session is terminated.
6	Service-Type	Framed (2) for EAP /Call check (10) for MAB.
8	Framed-IP-Address	IPv4 address of supplicant.
168	Framed-IPV6-Address	IPv6 address of supplicant.
30	Called-Station-Id	Switch MAC Address.
31	Calling-Station-Id	Supplicant MAC Address.
Accounting Attributes		
40	Acct-Status-Type	STOP
44	Acct-Session-Id	<Supplicant MAC> | Running number
50	Acct-Multi-Session-Id	<Supplicant MAC>|<Switch MAC>|<Timestamp>
51	Acct-Link-Count	1
46	Acct-Session Time	Time the user has received the service.
49	Acct-Terminate-Cause	Reason for session termination.
61	NAS-Port-Type	Ethernet

NOTE: During the administrative initiated reload and system failover events, the accounting Stop records for the 802.1x authorized supplicants are not sent to RADIUS server.

Table 89. Use cases for dot1x supplicant to trigger RADIUS Accounting Start/Stop records

dot1x event	Accounting type	Attributes
Dot1x user authentication success	Start	Start record attributes for dot1x supplicant.
Dot1x user logoff	Stop	Stop record attributes with termination cause as User Request (1).
Dot1x Supplicant de-auth due to link down	Stop	Stop record attributes with termination cause as Lost carrier (2).
Dot1x Supp De-Auth due to supplicant shutdown	Stop	Stop record attributes with termination cause as Lost carrier (2).
Administrative shut of dot1x authorized interface	Stop	Stop record attributes with termination cause as Lost carrier (2).
CLI user session disconnects due to dynamic authorization (CoA port bounce/session termination/ VLAN change)	Stop	Stop record attributes with termination cause as Admin Reset (6).
Deauthorization due to VLAN change	Stop	Stop record attributes with termination cause as Admin Reset (6).
CLI configuration of the dot1x authorized port to mab-only auth-type	Stop	Stop record attributes with termination cause as port-reinitialized (21).
Configure Port control to force unauth	Stop	Stop record attributes with termination cause as port-reinitialized (21).
Interface Host mode change (single/multihost/multiauth)	Stop	Stop record attributes with termination cause as port-reinitialized (21).

Table 89. Use cases for dot1x supplicant to trigger RADIUS Accounting Start/Stop records (continued)

dot1x event	Accounting type	Attributes
Configure max supplicant per interface	Stop	Stop record attributes with termination cause as port-reinitialized (21).
Supplicant goes off without explicitly sending EAP logoff	Stop	Stop record attributes with termination cause as Idle Timeout (4).
Periodic Reauth of supplicant	Stop	Stop record attributes with termination cause as Supplicant restart (19).
Failure of dot1x authorized port assignment to untagged VLAN	Stop	Stop record attributes with termination cause as Port error (8).
dot1x user Reauth-failure	Stop	Stop record attributes with termination cause as Reauthentication Failure (20).
Disable dot1x globally/interface	Stop	Stop record attributes with termination cause as Port Administratively Disabled (22).

AAA Authentication

Dell EMC Networking OS supports a distributed client/server system implemented through authentication, authorization, and accounting (AAA) to help secure networks against unauthorized access.

In the Dell EMC Networking implementation, the Dell EMC Networking system acts as a RADIUS or TACACS+ client and sends authentication requests to a central remote authentication dial-in service (RADIUS) or Terminal access controller access control system plus (TACACS+) server that contains all user authentication and network service access information.

Dell EMC Networking uses local usernames/passwords (stored on the Dell EMC Networking system) or AAA for login authentication. With AAA, you can specify the security protocol or mechanism for different login methods and different users. In Dell EMC Networking OS, AAA uses a list of authentication methods, called method lists, to define the types of authentication and the sequence in which they are applied. You can define a method list or use the default method list. User-defined method lists take precedence over the default method list.

NOTE: If a console user logs in with RADIUS authentication, the privilege level is applied from the RADIUS server if the privilege level is configured for that user in RADIUS, whether you configure RADIUS authorization.

NOTE: RADIUS and TACACS servers support VRF-awareness functionality. You can create RADIUS and TACACS groups and then map multiple servers to a group. The group to which you map multiple servers is bound to a single VRF.

Configuration Task List for AAA Authentication

The following sections provide the configuration tasks.

- Configuring AAA Authentication Login Methods
- Enabling AAA Authentication
- Enabling AAA Authentication - RADIUS

For a complete list of all commands related to login authentication, refer to the *Security* chapter in the *Dell EMC Networking OS Command Reference Guide*.

Configure Login Authentication for Terminal Lines

You can assign up to five authentication methods to a method list. Dell EMC Networking OS evaluates the methods in the order in which you enter them in each list.

If the first method list does not respond or returns an error, Dell EMC Networking OS applies the next method list until the user either passes or fails the authentication. If the user fails a method list, Dell EMC Networking OS does not apply the next method list.

Configuring AAA Authentication Login Methods

To configure an authentication method and method list, use the following commands.

Dell EMC Networking OS Behavior: If you use a method list on the console port in which RADIUS or TACACS is the last authentication method, and the server is not reachable, Dell EMC Networking OS allows access even though the username and password credentials cannot be verified. Only the console port behaves this way, and does so to ensure that users are not locked out of the system if network-wide issue prevents access to these servers.

1. Define an authentication method-list (*method-list-name*) or specify the default.

   CONFIGURATION mode

   `aaa authentication login {method-list-name | default} method1 [... method4]`

   The default method-list is applied to all terminal lines.

   Possible methods are:

   - `enable`: use the password you defined using the `enable secret`, `enable password`, or `enable sha256-password` command in CONFIGURATION mode. In general, the `enable secret` command overrules the `enable password` command. If you configure the `enable sha256-password` command, it overrules both the `enable secret` and `enable password` commands.
   - `line`: use the password you defined using the `password` command in LINE mode.
   - `local`: use the username/password database defined in the local configuration.
   - `none`: no authentication.
   - `radius`: use the RADIUS servers configured with the radius-server host command.
   - `tacacs+`: use the TACACS+ servers configured with the `tacacs-server host` command.

2. Enter LINE mode.

   CONFIGURATION mode

   `line {aux 0 | console 0 | vty number [... end-number]}`

3. Assign a *method-list-name* or the default list to the terminal line.

   LINE mode

   `login authentication {method-list-name | default}`

To view the configuration, use the `show config` command in LINE mode or the `show running-config` in EXEC Privilege mode.

NOTE: Dell EMC Networking recommends using the `none` method only as a backup. This method does not authenticate users. The `none` and `enable` methods do not work with secure shell (SSH).

You can create multiple method lists and assign them to different terminal lines.

Enabling AAA Authentication

To enable AAA authentication, use the following command.

- Enable AAA authentication.

  CONFIGURATION mode

  `aaa authentication enable {method-list-name | default} method1 [... method4]`

  - `default`: uses the listed authentication methods that follow this argument as the default list of methods when a user logs in.
  - *method-list-name*: character string used to name the list of enable authentication methods activated when a user logs in.
  - *method1 [... method4]*: any of the following: RADIUS, TACACS, enable, line, none.

If you do not set the default list, only the local enable is checked. This setting has the same effect as issuing an `aaa authentication enable default enable` command.

Enabling AAA Authentication — RADIUS

To enable authentication from the RADIUS server, and use TACACS as a backup, use the following commands.

1. Enable RADIUS and set up TACACS as backup.

CONFIGURATION mode

```
aaa authentication enable default radius tacacs
```

2. Establish a host address and password.

   CONFIGURATION mode

   ```
   radius-server host x.x.x.x key some-password
   ```

3. Establish a host address and password.

   CONFIGURATION mode

   ```
   tacacs-server host x.x.x.x key some-password
   ```

To get `enable authentication` from the RADIUS server and use TACACS as a backup, issue the following commands.

The following example shows enabling authentication from the RADIUS server.

```
DellEMC(config)# aaa authentication enable default radius tacacs
Radius and TACACS server has to be properly setup for this.
DellEMC(config)# radius-server host x.x.x.x key <some-password>
DellEMC(config)# tacacs-server host x.x.x.x key <some-password>
```

To use local enable authentication on the console, while using remote authentication on VTY lines, run the following commands.

The following example shows enabling local authentication for console and remote authentication for the VTY lines.

```
DellEMC(config)# aaa authentication enable mymethodlist radius tacacs
DellEMC(config)# line vty 0 9
DellEMC(config-line-vty)# enable authentication mymethodlist
```

Server-Side Configuration

Using AAA authentication, the switch acts as a RADIUS or TACACS+ client to send authentication requests to a TACACS+ or RADIUS server.

- **TACACS+** — When using TACACS+, Dell EMC Networking sends an initial packet with service type SVC_ENABLE, and then sends a second packet with just the password. The TACACS server must have an entry for username $enable$.
- **RADIUS** — When using RADIUS authentication, the Dell OS sends an authentication packet with the following:

  ```
  Username: $enab15$
  Password: <password-entered-by-user>
  ```

Therefore, the RADIUS server must have an entry for this username.

Configuring Re-Authentication

Starting from Dell EMC Networking OS 9.11(0.0), the system enables re-authentication of user whenever there is a change in the authenticators.

The change in authentication happens when:

- Add or remove an authentication server (RADIUS/TACACS+)
- Modify an AAA authentication/authorization list
- Change to role-only (RBAC) mode

The re-authentication is also applicable for authenticated 802.1x devices. When there is a change in the authetication servers, the supplicants connected to all the ports are forced to re-authenticate.

1. Enable the re-authentication mode.

   CONFIGURATION mode

   ```
   aaa reauthentication enable
   ```

2. You are prompted to force the users to re-authenticate while adding or removing a RADIUS/TACACS+ server.

   CONFIGURATION mode

   ```
   aaa authentication login method-list-name
   ```

Example:

```
DellEMC(config)#aaa authentication login vty_auth_list radius
Force all logged-in users to re-authenticate (y/n)?
```

3. You are prompted to force the users to re-authenticate whenever there is a change in the RADIUS server list..

   CONFIGURATION mode

   `radius-server host` *IP Address*

   Example:

```
DellEMC(config)#radius-server host 192.100.0.12
Force all logged-in users to re-authenticate (y/n)?
```

```
DellEMC(config)#no radius-server host 192.100.0.12
Force all logged-in users to re-authenticate (y/n)?
```

Obscuring Passwords and Keys

By default, the `service password-encryption` command stores encrypted passwords. For greater security, you can also use the `service obscure-passwords` command to prevent a user from reading the passwords and keys, including RADIUS, TACACS+ keys, router authentication strings, VRRP authentication by obscuring this information. Passwords and keys are stored encrypted in the configuration file and by default are displayed in the encrypted form when the configuration is displayed. Enabling the `service obscure-passwords` command displays asterisks instead of the encrypted passwords and keys. This command prevents a user from reading these passwords and keys by obscuring this information with asterisks.

Password obscuring masks the password and keys for display only but does not change the contents of the file. The string of asterisks is the same length as the encrypted string for that line of configuration. To verify that you have successfully obscured passwords and keys, use the `show running-config` command or `show startup-config` command.

If you are using role-based access control (RBAC), only the system administrator and security administrator roles can enable the `service obscure-password` command.

To enable the obscuring of passwords and keys, use the following command.

- Turn on the obscuring of passwords and keys in the configuration.

  CONFIGURATION mode

  `service obscure-passwords`

Example of Obscuring Password and Keys

```
DellEMC(config)# service obscure-passwords
```

AAA Authorization

Dell EMC Networking OS enables AAA new-model by default.

You can set authorization to be either `local` or `remote`. Different combinations of authentication and authorization yield different results. By default, Dell EMC Networking OS sets both to **local**.

Privilege Levels Overview

Limiting access to the system is one method of protecting the system and your network. However, at times, you might need to allow others access to the router and you can limit that access to a subset of commands. In Dell EMC Networking OS, you can configure a privilege level for users who need limited access to the system.

Every command in Dell EMC Networking OS is assigned a privilege level of 0, 1, or 15. You can configure up to 16 privilege levels in Dell EMC Networking OS. Dell EMC Networking OS is pre-configured with three privilege levels and you can configure 13 more. The three pre-configured levels are:

- **Privilege level 1** — is the default level for EXEC mode. At this level, you can interact with the router, for example, view some `show` commands and Telnet and ping to test connectivity, but you cannot configure the router. This level is often called the "user" level. One of the commands available in Privilege level 1 is the `enable` command, which you can use to enter a specific privilege level.
- **Privilege level 0** — contains only the `end`, `enable`, and `disable` commands.
- **Privilege level 15** — the default level for the `enable` command, is the highest level. In this level you can access any command in Dell EMC Networking OS.

Privilege levels 2 through 14 are not configured and you can customize them for different users and access.

After you configure other privilege levels, enter those levels by adding the level parameter after the `enable` command or by configuring a user name or password that corresponds to the privilege level. For more information about configuring user names, refer to Configuring a Username and Password.

By default, commands in Dell EMC Networking OS are assigned to different privilege levels. You can access those commands only if you have access to that privilege level. For example, to reach the `protocol spanning-tree` command, log in to the router, enter the `enable` command for privilege level 15 (this privilege level is the default level for the command) and then enter CONFIGURATION mode.

You can configure passwords to control access to the box and assign different privilege levels to users. Dell EMC Networking OS supports the use of passwords when you log in to the system and when you enter the `enable` command. If you move between privilege levels, you are prompted for a password if you move to a higher privilege level.

Configuration Task List for Privilege Levels

The following list has the configuration tasks for privilege levels and passwords.

- Configuring a Username and Password (mandatory)
- Configuring the Enable Password Command (mandatory)
- Configuring Custom Privilege Levels (mandatory)
- Specifying LINE Mode Password and Privilege (optional)
- Enabling and Disabling Privilege Levels (optional)

For a complete listing of all commands related to Dell EMC Networking OS privilege levels and passwords, refer to the *Security* chapter in the *Dell EMC Networking OS Command Reference Guide*.

Configuring a Username and Password

In Dell EMC Networking OS, you can assign a specific username to limit user access to the system.

To configure a username and password, use the following command.

- Assign a user name and password.

  CONFIGURATION mode

  `username` *`name`* `[access-class` *`access-list-name`*`] [nopassword | password [`*`encryption-type`*`]` *`password`*`] [privilege` *`level`*`][secret]`

  Configure the optional and required parameters:
  - *`name`*: Enter a text string up to 63 characters long.
  - `access-class` *`access-list-name`*: Enter the name of a configured IP ACL.
  - `nopassword`: Do not require the user to enter a password.
  - *`encryption-type`*: Enter 0 for plain text or 7 for encrypted text.
  - *`password`*: Enter a string.
  - `privilege` *`level`* The range is from 0 to 15.
  - `Secret:Specify the secret for the user`

To view username, use the `show users` command in EXEC Privilege mode.

Configuring the Enable Password Command

To configure Dell EMC Networking OS, use the `enable` command to enter EXEC Privilege level 15. After entering the command, Dell EMC Networking OS requests that you enter a password.

Privilege levels are not assigned to passwords, rather passwords are assigned to a privilege level. You can always change a password for any privilege level. To change to a different privilege level, enter the `enable` command, then the privilege level. If you do not enter a privilege level, the default level **15** is assumed.

To configure a password for a specific privilege level, use the following command.

- Configure a password for a privilege level.

  CONFIGURATION mode

  `enable password [level` *`level`*`] [`*`encryption-mode`*`] `*`password`*

  Configure the optional and required parameters:
  - `level` *`level`*: Specify a level from 0 to 15. Level 15 includes all levels.
  - *`encryption-type`*: Enter 0 for clear text, 7 for DES-encrypted text, or 8 for sha256-based encrypted text.
  - *`password`*: Enter a string.

  To change only the password for the `enable` command, configure only the *`password`* parameter.

To view the configuration for the `enable secret` command, use the `show running-config` command in EXEC Privilege mode.

In custom-configured privilege levels, the `enable` command is always available. No matter what privilege level you entered Dell EMC Networking OS, you can enter the `enable 15` command to access and configure all CLIs.

Configuring Custom Privilege Levels

In addition to assigning privilege levels to the user, you can configure the privilege levels of commands so that they are visible in different privilege levels.

Within Dell EMC Networking OS, commands have certain privilege levels. With the `privilege` command, you can change the default level or you can reset their privilege level back to the default.

- Assign the launch keyword (for example, `configure`) for the keyword's command mode.
- If you assign only the first keyword to the privilege level, all commands beginning with that keyword are also assigned to the privilege level. If you enter the entire command, the software assigns the privilege level to that command only.

To assign commands and passwords to a custom privilege level, use the following commands. You must be in privilege level 15.

1. Assign a user name and password.

   CONFIGURATION mode

   `username` *`name`* `[access-class` *`access-list-name`*`] [privilege` *`level`*`] [nopassword | password [`*`encryption-type`*`]` *`password Secret`*`]`

   Configure the optional and required parameters:
   - *`name`*: Enter a text string up to 63 characters(maximum) long.
   - `access-class` *`access-list-name`*: Restrict access by access-class..
   - `privilege` *`level`*: The range is from 0 to 15.
   - `nopassword`: No password is required for the user to log in.
   - *`encryption-type`*: Enter 0 for plain text or 7 for encrypted text.
   - *`password`*: Enter a string. Specify the password for the user.
   - *`Secret`*: Specify the secret for the user.
2. Configure a password for privilege level.

   CONFIGURATION mode

   `enable password [level` *`level`*`] [`*`encryption-mode`*`]` *`password`*

   Configure the optional and required parameters:
   - `level` *`level`*: specify a level from 0 to 15. Level 15 includes all levels.
   - *`encryption-type`*: enter 0 for plain text or 7 for encrypted text.
   - *`password`*: enter a string up to 32 characters long.

   To change only the password for the `enable` command, configure only the `password` parameter.

3. Configure level and commands for a mode or reset a command's level.

   CONFIGURATION mode

   `privilege` *`mode`* {`level` *`level command`* | `reset` *`command`*}

   Configure the following required and optional parameters:

   - *`mode`*: enter a keyword for the modes (`exec`, `configure`, `interface`, `line`, `route-map`, or `router`)
   - `level` *`level`*: the range is from 0 to 15. Levels 0, 1, and 15 are pre-configured. Levels 2 to 14 are available for custom configuration.
   - *`command`*: an Dell EMC Networking OS CLI keyword (up to five keywords allowed).
   - `reset`: return the command to its default privilege mode.

To view the configuration, use the `show running-config` command in EXEC Privilege mode.

The following example shows a configuration to allow a user *john* to view only EXEC mode commands and all `snmp-server` commands. Because the `snmp-server` commands are *enable* level commands and, by default, found in CONFIGURATION mode, also assign the launch command for CONFIGURATION mode, `configure`, to the same privilege level as the `snmp-server` commands.

Line 1: The user *john* is assigned privilege level 8 and assigned a password.

Line 2: All other users are assigned a password to access privilege level 8.

Line 3: The `configure` command is assigned to privilege level 8 because it needs to reach CONFIGURATION mode where the `snmp-server` commands are located.

Line 4: The `snmp-server` commands, in CONFIGURATION mode, are assigned to privilege level 8.

```
DellEMC(conf)#username john privilege 8 password john
DellEMC(conf)#enable password level 8 notjohn
DellEMC(conf)#privilege exec level 8 configure
DellEMC(conf)#privilege config level 8 snmp-server
DellEMC(conf)#end
DellEMC#show running-config
Current Configuration ...
!
hostname Force10
!
enable password level 8 notjohn
enable password Force10
!
username admin password 0 admin
username john password 0 john privilege 8
!
```

The following example shows the Telnet session for user *john*. The `show privilege` command output confirms that *john* is in privilege level 8. In EXEC Privilege mode, *john* can access only the commands listed. In CONFIGURATION mode, *john* can access only the `snmp-server` commands.

```
apollo% telnet 172.31.1.53
Trying 172.31.1.53...
Connected to 172.31.1.53.
Escape character is '^]'.
Login: john
Password:
DellEMC#show priv
Current privilege level is 8
DellEMC#?
configure      Configuring from terminal
disable        Turn off privileged commands
enable         Turn on privileged commands
exit           Exit from the EXEC
no             Negate a command
show           Show running system information
terminal       Set terminal line parameters
traceroute     Trace route to destination
DellEMC#confi
DellEMC(conf)#?
end            Exit from Configuration mode
exit           Exit from Configuration mode
no             Reset a command
```

```
snmp-server     Modify SNMP parameters
DellEMC(conf)#
```

Specifying LINE Mode Password and Privilege

You can specify a password authentication of all users on different terminal lines.

The user's privilege level is the same as the privilege level assigned to the terminal line, unless a more specific privilege level is assigned to the user.

To specify a password for the terminal line, use the following commands.

- Configure a custom privilege level for the terminal lines.

  LINE mode

  `privilege level` *`level`*

  - `level` *`level`*: The range is from 0 to 15. Levels 0, 1, and 15 are pre-configured. Levels 2 to 14 are available for custom configuration.
- Specify either a plain text or encrypted password.

  LINE mode

  `password [`*`encryption-type`*`] `*`password`*

  Configure the following optional and required parameters:
  - *`encryption-type`*: Enter 0 for plain text or 7 for encrypted text.
  - *`password`*: Enter a text string up to 32 characters long.

To view the password configured for a terminal, use the `show config` command in LINE mode.

Enabling and Disabling Privilege Levels

To enable and disable privilege levels, use the following commands.

- Set a user's security level.

  EXEC Privilege mode

  `enable` or `enable privilege-level`

  If you do not enter a privilege level, Dell EMC Networking OS sets it to 15 by default.
- Move to a lower privilege level.

  EXEC Privilege mode

  `disable` *`level-number`*

  - *`level-number`*: The level-number you wish to set.

  If you enter `disable` without a level-number, your security level is 1.

Resetting a Password

To reset a password on the system, follow these steps.

1. Connect to the system using a console.
2. Disconnect and reconnect the power cord on the system to cycle the power.
3. During system boot, press ESC when prompted to display the Grub Menu (see Example 1).
4. During system boot, press ESC when prompted during the countdown to stop the auto-boot process (see Example 2).
5. Press `C` to access the Grub boot loader command line prompt.
6. Enter the following commands at the Grub command line prompt (`grub>`).

   ```
   set stconfigignore=true
   save_env stconfigignore
   reboot
   ```
7. The system boots up with factory default configuration. The default DellEMC> system prompt displays when the system boots.

8. Copy the startup-config into the running-config.
9. To display the content of the startup-config, remove the previous authentication configuration and set the new authentication parameters. The rest of the previous configuration is preserved.

```
Version 2.00.1201. Copyright (C) 2017 American Megatrends, Inc.
EVALUATION COPY.
Press <DEL> or <F2> to enter setup.
  Grub 1.99~rc1 (Dell Force10)
  Built by root at bsdlab on Thu_Aug_18_06:51:21_UTC_2011
  Z9000 Boot selector Label 3.0.1.1 NetBoot Label 0.0.0.0
```

```
Entering Menu
Grub 1.99~rc1 (Dell Force10)
Built by root at bsdlab on Thu_Aug_18_06:51:21_UTC_2011
Z9000 Boot selector Label 3.0.1.1 NetBoot Label 0.0.0.0
+-----------------------------+
|Force10 Boot |
|                             |
|                             |
|                             |
|                             |
|                             |
|
|
|                             |
|                             |
|                             |
|                             |
+-----------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the commands
before booting or 'c' for a command-line.
```

RADIUS

Remote authentication dial-in user service (RADIUS) is a distributed client/server protocol.

This protocol transmits authentication, authorization, and configuration information between a central RADIUS server and a RADIUS client (the Dell EMC Networking system). The system sends user information to the RADIUS server and requests authentication of the user and password. The RADIUS server returns one of the following responses:

- **Access-Accept** — the RADIUS server authenticates the user.
- **Access-Reject** — the RADIUS server does not authenticate the user.

If an error occurs in the transmission or reception of RADIUS packets, you can view the error by enabling the `debug radius` command.

Transactions between the RADIUS server and the client are encrypted (the users' passwords are not sent in plain text). RADIUS uses UDP as the transport protocol between the RADIUS server host and the client.

For more information about RADIUS, refer to RFC 2865, *Remote Authentication Dial-in User Service*.

RADIUS Authentication

Dell EMC Networking OS supports RADIUS for user authentication (text password) at login and can be specified as one of the login authentication methods in the `aaa authentication login` command.

When configuring AAA authorization, you can configure to limit the attributes of services available to a user. When you enable authorization, the network access server uses configuration information from the user profile to issue the user's session. The user's access is limited based on the configuration attributes. RADIUS exec-authorization stores a user-shell profile and that is applied during user login. You may name the relevant named-lists with either a unique name or the default name. When you enable authorization by the RADIUS server, the server returns the following information to the client:

- Idle Time
- ACL Configuration Information
- Auto-Command

- Privilege Levels

After gaining authorization for the first time, you may configure these attributes.

NOTE: RADIUS authentication/authorization is done for every login. There is no difference between first-time login and subsequent logins.

Idle Time

Every session line has its own idle-time. If the idle-time value is not changed, the default value of **30 minutes** is used.

RADIUS specifies idle-time allow for a user during a session before timeout. When a user logs in, the lower of the two idle-time values (configured or default) is used. The idle-time value is updated if both of the following happens:

- The administrator changes the idle-time of the line on which the user has logged in.
- The idle-time is lower than the RADIUS-returned idle-time.

ACL Configuration Information

The RADIUS server can specify an ACL. If an ACL is configured on the RADIUS server, and if that ACL is present, the user may be allowed access based on that ACL.

If the ACL is absent, authorization fails, and a message is logged indicating this.

RADIUS can specify an ACL for the user if both of the following are true:

- If an ACL is absent.
- If there is a very long delay for an entry, or a denied entry because of an ACL, and a message is logged.

NOTE: The ACL name must be a string. Only standard ACLs in authorization (both RADIUS and TACACS) are supported. Authorization is denied in cases using Extended ACLs.

Auto-Command

You can configure the system through the RADIUS server to automatically execute a command when you connect to a specific line.

The `auto-command` command is executed when the user is authenticated and before the prompt appears to the user.

- Automatically execute a command.

  ```
  auto-command
  ```

Privilege Levels

Through the RADIUS server, you can configure a privilege level for the user to enter into when they connect to a session.

This value is configured on the client system.

- Set a privilege level.

  ```
  privilege level
  ```

Configuration Task List for RADIUS

To authenticate users using RADIUS, you must specify at least one RADIUS server so that the system can communicate with and configure RADIUS as one of your authentication methods.

The following list includes the configuration tasks for RADIUS.

- Defining a AAA Method List to be Used for RADIUS (mandatory)
- Applying the Method List to Terminal Lines (mandatory except when using default lists)
- Specifying a RADIUS Server Host (mandatory)
- Setting Global Communication Parameters for all RADIUS Server Hosts (optional)
- Monitoring RADIUS (optional)

For a complete listing of all Dell EMC Networking OS commands related to RADIUS, refer to the *Security* chapter in the *Dell EMC Networking OS Command Reference Guide*.

NOTE: RADIUS authentication and authorization are done in a single step. Hence, authorization cannot be used independent of authentication. However, if you have configured RADIUS authorization and have not configured authentication, a message is logged stating this. During authorization, the next method in the list (if present) is used, or if another method is not present, an error is reported.

To view the configuration, use the `show config` in LINE mode or the `show running-config` command in EXEC Privilege mode.

Defining a AAA Method List to be Used for RADIUS

To configure RADIUS to authenticate or authorize users on the system, create a AAA method list.

Default method lists do not need to be explicitly applied to the line, so they are not mandatory.

To create a method list, use the following commands.

- Enter a text string (up to 16 characters long) as the name of the method list you wish to use with the RADIUS authentication method.

  CONFIGURATION mode

  `aaa authentication login` *`method-list-name`* `radius`
- Create a method list with RADIUS and TACACS+ as authorization methods.

  CONFIGURATION mode

  `aaa authorization exec {`*`method-list-name`* `| default} radius tacacs+`

  Typical order of methods: RADIUS, TACACS+, Local, None.

  If RADIUS denies authorization, the session ends (RADIUS must not be the last method specified).

Applying the Method List to Terminal Lines

To enable RADIUS AAA login authentication for a method list, apply it to a terminal line.

To configure a terminal line for RADIUS authentication and authorization, use the following commands.

- Enter LINE mode.

  CONFIGURATION mode

  `line {aux 0 | console 0 | vty` *`number`* `[`*`end-number`*`]}`
- Enable AAA login authentication for the specified RADIUS method list.

  LINE mode

  `login authentication {`*`method-list-name`* `| default}`

  This procedure is mandatory if you are not using default lists.
- To use the method list.

  CONFIGURATION mode

  `authorization exec methodlist`

Specifying a RADIUS Server Host

When configuring a RADIUS server host, you can set different communication parameters, such as the UDP port, the key password, the number of retries, and the timeout.

To specify a RADIUS server host and configure its communication parameters, use the following command.

- Enter the host name or IP address of the RADIUS server host.

  CONFIGURATION mode

  `radius-server host {`*`hostname`* `|` *`ip-address`*`} [auth-port` *`port-number`*`] [retransmit` *`retries`*`] [timeout` *`seconds`*`] [key [`*`encryption-type`*`]` *`key`*`]`

  Configure the optional communication parameters for the specific host:

  - `auth-port port-number`: the range is from 0 to 65535. Enter a UDP port number. The default is **1812**.

- `retransmit retries`: the range is from 0 to 100. Default is **3**.
- `timeout seconds`: the range is from 0 to 1000. Default is **5 seconds**.
- `key [encryption-type] key`: enter 0 for plain text or 7 for encrypted text, and a string for the key. The key can be up to 42 characters long. This key must match the key configured on the RADIUS server host.

If you do not configure these optional parameters, the global default values for all RADIUS host are applied.

To specify multiple RADIUS server hosts, configure the `radius-server host` command multiple times. If you configure multiple RADIUS server hosts, Dell EMC Networking OS attempts to connect with them in the order in which they were configured. When Dell EMC Networking OS attempts to authenticate a user, the software connects with the RADIUS server hosts one at a time, until a RADIUS server host responds with an accept or reject response.

If you want to change an optional parameter setting for a specific host, use the `radius-server host` command. To change the global communication settings to all RADIUS server hosts, refer to Setting Global Communication Parameters for all RADIUS Server Hosts.

To view the RADIUS configuration, use the `show running-config radius` command in EXEC Privilege mode.

To delete a RADIUS server host, use the `no radius-server host {hostname | ip-address}` command.

Setting Global Communication Parameters for all RADIUS Server Hosts

You can configure global communication parameters (`auth-port`, `key`, `retransmit`, and `timeout` parameters) and specific host communication parameters on the same system.

However, if you configure both global and specific host parameters, the specific host parameters override the global parameters for that RADIUS server host.

To set global communication parameters for all RADIUS server hosts, use the following commands.

- Set a time interval after which a RADIUS host server is declared dead.

  CONFIGURATION mode

  `radius-server deadtime seconds`
  - `seconds`: the range is from 0 to 2147483647. The default is **0 seconds**.
- Configure a key for all RADIUS communications between the system and RADIUS server hosts.

  CONFIGURATION mode

  `radius-server key [encryption-type] key`
  - `encryption-type`: enter 7 to encrypt the password. Enter 0 to keep the password as plain text.
  - `key`: enter a string. The key can be up to 42 characters long. You cannot use spaces in the key.
- Configure the number of times Dell EMC Networking OS retransmits RADIUS requests.

  CONFIGURATION mode

  `radius-server retransmit retries`
  - `retries`: the range is from 0 to 100. Default is **3 retries**.
- Configure the time interval the system waits for a RADIUS server host response.

  CONFIGURATION mode

  `radius-server timeout seconds`
  - `seconds`: the range is from 0 to 1000. Default is **5 seconds**.

To view the configuration of RADIUS communication parameters, use the `show running-config` command in EXEC Privilege mode.

Monitoring RADIUS

To view information on RADIUS transactions, use the following command.

- View RADIUS transactions to troubleshoot problems.

  EXEC Privilege mode

  `debug radius`

Microsoft Challenge-Handshake Authentication Protocol Support for RADIUS Authentication

Dell EMC Networking OS supports Microsoft Challenge-Handshake Authentication Protocol (MS-CHAPv2) with RADIUS authentication.

RADIUS is used to authenticate Telnet, SSH, console, REST, and OMI access to the switch based on the AAA configuration. By default, the RADIUS client in the switch uses PAP (Password Authentication Protocol) for sending the login credentials to the RADIUS server. The user-password attribute is added to the access-request message that is sent to the RADIUS server. Depending on the success or failure of authentication, the RADIUS server sends back an access-accept or access-reject message respectively.

MS-CHAPv2 is secure than PAP. MS-CHAPv2 does not send user-password in the Access-Request message. It implements mutual authentication based on the random challenges. MS-CHAP-Challenge and MS-CHAP2-Response attributes are sent in the Access-Request message from the switch to the RADIUS Server. RADIUS Server validates the attributes and sends back MS-CHAPv2-Success attribute in the Access-Accept message. If the validation fails, then RADIUS Server sends back the Access-Reject Message.

Enabling MS-CHAPv2 with the RADIUS authentication

Before enabling MS-CHAPv2 authentication on the switch, you must first Enable MS-CHAPv2 support in RADIUS Server.

To enable MS-CHAPv2 for the RADIUS authentication:

1. Enable RADIUS.

   CONFIGURATION mode

   ```
   aaa authentication login default radius local
   ```
2. Specify the protocol for authentication.

   CONFIGURATION mode

   ```
   aaa radius auth-method mschapv2
   ```
3. Establish a host address and password.

   CONFIGURATION mode

   radius-server host H key K
4. Log in to switch using console or telnet or ssh with a valid user role.

When 1-factor authentication is used, the authentication succeeds enabling you to access the switch. When two-factor authentication is used, the system prompts you to enter a one-time password as a second step of authentication. If a valid one-time password is supplied, the authentication succeeds enabling you to access the switch.

TACACS+

Dell EMC Networking OS supports terminal access controller access control system (TACACS+ client, including support for login authentication.

Configuration Task List for TACACS+

The following list includes the configuration task for TACACS+ functions.

- Choosing TACACS+ as the Authentication Method
- Monitoring TACACS+
- TACACS+ Remote Authentication
- Specifying a TACACS+ Server Host

For a complete listing of all commands related to TACACS+, refer to the *Security* chapter in the *Dell EMC Networking OS Command Reference Guide*.

Choosing TACACS+ as the Authentication Method

One of the login authentication methods available is TACACS+ and the user's name and password are sent for authentication to the TACACS hosts specified.

To use TACACS+ to authenticate users, specify at least one TACACS+ server for the system to communicate with and configure TACACS+ as one of your authentication methods.

To select TACACS+ as the login authentication method, use the following commands.

1. Configure a TACACS+ server host.

   CONFIGURATION mode

   `tacacs-server host {ip-address | host}`

   Enter the IP address or host name of the TACACS+ server.

   Use this command multiple times to configure multiple TACACS+ server hosts.

2. Enter a text string (up to 16 characters long) as the name of the method list you wish to use with the TACAS+ authentication method.

   CONFIGURATION mode

   `aaa authentication login {method-list-name | default} tacacs+ [...method3]`

   The TACACS+ method must not be the last method specified.

3. Enter LINE mode.

   CONFIGURATION mode

   `line {aux 0 | console 0 | vty number [end-number]}`

4. Assign the *method-list* to the terminal line.

   LINE mode

   `login authentication {method-list-name | default}`

To view the configuration, use the `show config` in LINE mode or the `show running-config tacacs+` command in EXEC Privilege mode.

If authentication fails using the primary method, Dell EMC Networking OS employs the second method (or third method, if necessary) automatically. The fallback to the second method would happen only if the authentication failure is due to a non-reachable server or invalid TACACS server key. The fallback would not occur if the authentication failure is due to invalid credentials. For example, if the TACACS+ server is reachable, but the server key is invalid, Dell EMC Networking OS proceeds to the next authentication method. In the following example, the TACACS+ is incorrect, but the user is still authenticated by the secondary method.

First bold line: Server key purposely changed to incorrect value.

Second bold line: User authenticated using the secondary method.

```
DellEMC(conf)#
DellEMC(conf)#do show run aaa
!
aaa authentication enable default tacacs+ enable
aaa authentication enable LOCAL enable tacacs+
aaa authentication login default tacacs+ local
aaa authentication login LOCAL local tacacs+
aaa authorization exec default tacacs+ none
aaa authorization commands 1 default tacacs+ none
aaa authorization commands 15 default tacacs+ none
aaa accounting exec default start-stop tacacs+
aaa accounting commands 1 default start-stop tacacs+
aaa accounting commands 15 default start-stop tacacs+
DellEMC(conf)#
DellEMC(conf)#do show run tacacs+
!
tacacs-server key 7 d05206c308f4d35b
tacacs-server host 10.10.10.10 timeout 1
DellEMC(conf)#tacacs-server key angeline
DellEMC(conf)#%RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user admin on
vty0 (10.11.9.209)
%RPM0-P:CP %SEC-3-AUTHENTICATION_ENABLE_SUCCESS: Enable password
authentication success on vty0 ( 10.11.9.209 )
%RPM0-P:CP %SEC-5-LOGOUT: Exec session is terminated for user admin on line
```

```
vty0 (10.11.9.209)
DellEMC(conf)#username angeline password angeline
DellEMC(conf)#%RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user angeline
on vty0 (10.11.9.209)
%RPM0-P:CP %SEC-3-AUTHENTICATION_ENABLE_SUCCESS: Enable password
authentication success on vty0 ( 10.11.9.209 )
```

Monitoring TACACS+

To view information on TACACS+ transactions, use the following command.

- View TACACS+ transactions to troubleshoot problems.

  EXEC Privilege mode

  ```
  debug tacacs+
  ```

TACACS+ Remote Authentication

The system takes the access class from the TACACS+ server. Access class is the class of service that restricts Telnet access and packet sizes.

If you have configured remote authorization, the system ignores the access class you have configured for the VTY line and gets this access class information from the TACACS+ server. The system must know the username and password of the incoming user before it can fetch the access class from the server. A user, therefore, at least sees the login prompt. If the access class denies the connection, the system closes the Telnet session immediately. The following example demonstrates how to configure the access-class from a TACACS+ server. This configuration ignores the configured access-class on the VTY line. If you have configured a deny10 ACL on the TACACS+ server, the system downloads it and applies it. If the user is found to be coming from the 10.0.0.0 subnet, the system also immediately closes the Telnet connection. Note, that no matter where the user is coming from, they see the login prompt.

When configuring a TACACS+ server host, you can set different communication parameters, such as the key password.

Example of Specifying a TACACS+ Server Host

```
DellEMC(conf)#
DellEMC(conf)#aaa authentication login tacacsmethod tacacs+
DellEMC(conf)#aaa authentication exec tacacsauthorization tacacs+
DellEMC(conf)#tacacs-server host 25.1.1.2 key Force
DellEMC(conf)#
DellEMC(conf)#line vty 0 9
DellEMC(config-line-vty)#login authentication tacacsmethod
DellEMC(config-line-vty)#end
```

Specifying a TACACS+ Server Host

To specify a TACACS+ server host and configure its communication parameters, use the following command.

- Enter the host name or IP address of the TACACS+ server host.

  CONFIGURATION mode

  ```
  tacacs-server host {hostname | ip-address} [port port-number] [timeout seconds] [key key]
  ```

  Configure the optional communication parameters for the specific host:

  - `port port-number`: the range is from 0 to 65535. Enter a TCP port number. The default is **49**.
  - `timeout seconds`: the range is from 0 to 1000. Default is **10 seconds**.
  - `key key`: enter a string for the key. The key can be up to 42 characters long. This key must match a key configured on the TACACS+ server host. This parameter must be the last parameter you configure.

  If you do not configure these optional parameters, the default global values are applied.

To specify multiple TACACS+ server hosts, configure the `tacacs-server host` command multiple times. If you configure multiple TACACS+ server hosts, Dell EMC Networking OS attempts to connect with them in the order in which they were configured.

To view the TACACS+ configuration, use the `show running-config tacacs+` command in EXEC Privilege mode.

To delete a TACACS+ server host, use the `no tacacs-server host {hostname | ip-address}` command.

```
freebsd2# telnet 2200:2200:2200:2200:2200::2202
Trying 2200:2200:2200:2200:2200::2202...
Connected to 2200:2200:2200:2200:2200::2202.
Escape character is '^]'.
Login: admin
Password:
DellEMC#
```

Command Authorization

The AAA command authorization feature configures Dell EMC Networking OS to send each configuration command to a TACACS server for authorization before it is added to the running configuration.

By default, the AAA authorization commands configure the system to check both EXEC mode and CONFIGURATION mode commands. Use the `no aaa authorization config-commands` command to enable only EXEC mode command checking.

If rejected by the AAA server, the command is not added to the running config, and a message displays:

```
04:07:48:  %RPM0-P:CP %SEC-3-SEC_AUTHORIZATION_FAIL: Authorization failure Command
authorization failed for user (denyall) on vty0 ( 10.11.9.209 )
```

Certain TACACS+ servers do not authenticate the device if you use the `aaa authorization commands level default local tacacs+` command. To resolve the issue, use the `aaa authorization commands level default tacacs+ local` command.

Protection from TCP Tiny and Overlapping Fragment Attacks

Tiny and overlapping fragment attack is a class of attack where configured ACL entries — denying TCP port-specific traffic — is bypassed and traffic is sent to its destination although denied by the ACL.

RFC 1858 and 3128 proposes a countermeasure to the problem. This countermeasure is configured into the line cards and enabled by default.

Enabling SCP and SSH

Secure shell (SSH) is a protocol for secure remote login and other secure network services over an insecure network. Dell EMC Networking OS is compatible with SSH versions 2, in both the client and server modes. SSH sessions are encrypted and use authentication. SSH is enabled by default.

For details about the command syntax, refer to the *Security* chapter in the *Dell EMC Networking OS Command Line Interface Reference Guide*.

Dell EMC Networking OS SCP, which is a remote file copy program that works with SSH.

NOTE: The Windows-based WinSCP client software is not supported for secure copying between a PC and a Dell EMC Networking OS-based system. Unix-based SCP client software is supported.

To use the SSH client, use the following command.

- Open an SSH connection and specify the hostname, username, port number,encryption cipher,HMAC algorithm and version of the SSH client.

  EXEC Privilege mode

  `ssh {hostname} [-l username | -p port-number | -v 2}| -c encryption cipher | -m HMAC algorithm`

  hostname is the IP address or host name of the remote device. Enter an IPv4 or IPv6 address in dotted decimal format (A.B.C.D).
- SSH V2 is enabled by default on all the modes.

- Display SSH connection information.

  EXEC Privilege mode

  ```
  show ip ssh
  ```

The following example uses the `ip ssh server version 2` command to enable SSH version 2 and the `show ip ssh` command to confirm the setting.

```
DellEMC(conf)#ip ssh server version 2
DellEMC(conf)#do show ip ssh
SSH server                : enabled.
SSH server version        : v2.
SSH server vrf            : default.
SSH server ciphers        : 3des-cbc,aes128-cbc,aes192-cbc,aes256-cbc,aes128-ctr,aes192-
ctr,aes256-ctr.
SSH server macs           : hmac-md5,hmac-md5-96,hmac-sha1,hmac-sha1-96,hmac-
sha2-256,hmac-sha2-256-96.
SSH server kex algorithms : diffie-hellman-group-exchange-sha1,diffie-hellman-group1-
sha1,diffie-hellman-group14-sha1.
Password Authentication   : enabled.
Hostbased Authentication  : disabled.
RSA       Authentication  : disabled.
   Vty          Encryption      HMAC            Remote IP
DellEMC(conf)#
```

To disable SSH server functions, use the `no ip ssh server enable` command.

Using SCP with SSH to Copy a Software Image

To use secure copy (SCP) to copy a software image through an SSH connection from one switch to another, use the following commands.

1. On Switch 1, set the SSH port number (port 22 by default).

   CONFIGURATION MODE

   ```
   ip ssh server port number
   ```

2. On Switch 1, enable SSH.

   CONFIGURATION MODE

   ```
   copy ssh server enable
   ```

3. On Switch 2, invoke SCP.

   CONFIGURATION MODE

   ```
   copy scp: flash:
   ```

4. On Switch 2, in response to prompts, enter the path to the desired file and enter the port number specified in Step 1.

   EXEC Privilege Mode

5. On the chassis, invoke SCP.

   CONFIGURATION mode

   ```
   copy scp: flash:
   ```

The following example shows the use of SCP and SSH to copy a software image from one switch running SSH server on UDP port 99 to the local switch.

Other SSH related command include:

- `crypto key generate` : generate keys for the SSH server.
- `debug ip ssh` : enables collecting SSH debug information.
- `ip scp topdir` : identify a location for files used in secure copy transfer.
- `ip ssh authentication-retries` : configure the maximum number of attempts that should be used to authenticate a user.
- `ip ssh connection-rate-limit` : configure the maximum number of incoming SSH connections per minute.
- `ip ssh hostbased-authentication enable` : enable host-based authentication for the SSHv2 server.
- `ip ssh password-authentication enable` : enable password authentication for the SSH server.
- `ip ssh pub-key-file` : specify the file the host-based authentication uses.
- `ip ssh rhostsfile` : specify the rhost file the host-based authorization uses.

- `ip ssh rsa-authentication enable` : enable RSA authentication for the SSHv2 server.
- `ip ssh rsa-authentication` : add keys for the RSA authentication.
- `show crypto` : display the public part of the SSH host-keys.
- `show ip ssh client-pub-keys` : display the client public keys used in host-based authentication.
- `show ip ssh rsa-authentication` : display the authorized-keys for the RSA authentication.

```
DellEMC#copy scp: flash:
Address or name of remote host []: 10.10.10.1
Port number of the server [22]: 99
Source file name []: test.cfg
User name to login remote host: admin
Password to login remote host:
```

Removing the RSA Host Keys and Zeroizing Storage

Use the `crypto key zeroize rsa` command to delete the host key pairs, both the public and private key information for RSA 1 and or RSA 2 types. Note that when FIPS mode is enabled there is no RSA 1 key pair. Any memory currently holding these keys is zeroized (written over with zeroes) and the NVRAM location where the keys are stored for persistence across reboots is also zeroized.

To remove the generated RSA host keys and zeroize the key storage location, use the `crypto key zeroize rsa` command in CONFIGURATION mode.

```
DellEMC(conf)#crypto key zeroize rsa
```

Configuring When to Re-generate an SSH Key

You can configure the time-based or volume-based rekey threshold for an SSH session. If both threshold types are configured, the session rekeys when either one of the thresholds is reached.

To configure the time or volume rekey threshold at which to re-generate the SSH key during an SSH session, use the `ip ssh rekey` [`time` *rekey-interval*] [`volume` *rekey-limit*] command. CONFIGURATION mode.

Configure the following parameters:

- *rekey-interval:* time-based rekey threshold for an SSH session. The range is from 10 to 1440 minutes. The default is **60** minutes.
- *rekey-limit*: volume-based rekey threshold for an SSH session. The range is from 1 to 4096 to megabytes. The default is **1024** megabytes.

Examples

The following example configures the time-based rekey threshold for an SSH session to 30 minutes.

```
DellEMC(conf)#ip ssh rekey time 30
```

The following example configures the volume-based rekey threshold for an SSH session to 4096 megabytes.

```
DellEMC(conf)#ip ssh rekey volume 4096
```

Configuring the SSH Server Key Exchange Algorithm

To configure the key exchange algorithm for the SSH server, use the `ip ssh server kex` *key-exchange-algorithm* command in CONFIGURATION mode.

key-exchange-algorithm : Enter a space-delimited list of key exchange algorithms that will be used by the SSH server.

The following key exchange algorithms are available:

- diffie-hellman-group-exchange-sha1
- diffie-hellman-group1-sha1

- diffie-hellman-group14-sha1

The default key exchange algorithms are the following:

- diffie-hellman-group-exchange-sha1
- diffie-hellman-group1-sha1
- diffie-hellman-group14-sha1

When FIPS is enabled, the default is diffie-hellman-group14-sha1.

Example of Configuring a Key Exchange Algorithm

The following example shows you how to configure a key exchange algorithm.

```
DellEMC(conf)# ip ssh server  kex diffie-hellman-group-exchange-sha1 diffie-hellman-
group14-sha1
```

Configuring the HMAC Algorithm for the SSH Server

To configure the HMAC algorithm for the SSH server, use the `ip ssh server mac` *`hmac-algorithm`* command in CONFIGURATION mode.

hmac-algorithm: Enter a space-delimited list of keyed-hash message authentication code (HMAC) algorithms supported by the SSH server.

The following HMAC algorithms are available:

- hmac-md5
- hmac-md5-96
- hmac-sha1
- hmac-sha1-96
- hmac-sha2-256

The default HMAC algorithms are the following:

- hmac-sha2-256
- hmac-sha1
- hmac-sha1-96
- hmac-md5
- hmac-md5-96

When FIPS is enabled, the default HMAC algorithm is hmac-sha2-256,hmac-sha1,hmac-sha1-96.

Example of Configuring a HMAC Algorithm

The following example shows you how to configure a HMAC algorithm list.

```
DellEMC(conf)# ip ssh server mac hmac-sha1-96
```

Configuring the HMAC Algorithm for the SSH Client

To configure the HMAC algorithm for the SSH client, use the `ip ssh mac` *`hmac-algorithm`* command in CONFIGURATION mode.

hmac-algorithm: Enter a space-delimited list of keyed-hash message authentication code (HMAC) algorithms supported by the SSH server.

The following HMAC algorithms are available:

- hmac-md5
- hmac-md5-96

- hmac-sha1
- hmac-sha1-96
- hmac-sha2-256

The default list of HMAC algorithm is in the following order:

- hmac-sha2-256
- hmac-sha1
- hmac-sha1-96
- hmac-md5
- hmac-md5-96

When FIPS is enabled, the default HMAC algorithm is hmac-sha2-256, hmac-sha1, hmac-sha1-96.

Example of Configuring a HMAC Algorithm

The following example shows you how to configure a HMAC algorithm list.

```
DellEMC(conf)# ip ssh mac hmac-sha1-96
```

Configuring the SSH Server Cipher List

To configure the cipher list supported by the SSH server, use the `ip ssh server cipher` *cipher-list* command in CONFIGURATION mode.

cipher-list-: Enter a space-delimited list of ciphers the SSH server will support.

The following ciphers are available.

- `3des-cbc`
- `aes128-cbc`
- `aes192-cbc`
- `aes256-cbc`
- `aes128-ctr`
- `aes192-ctr`
- `aes256-ctr`

The default cipher list is aes256-ctr, aes256-cbc, aes192-ctr, aes192-cbc, aes128-ctr, aes128-cbc, 3des-cbc.

Example of Configuring a Cipher List

The following example shows you how to configure a cipher list.

```
DellEMC(conf)#ip ssh server cipher 3des-cbc aes128-cbc aes128-ctr
```

Configuring the SSH Client Cipher List

To configure the cipher list supported by the SSH client, use the `ip ssh cipher` *cipher-list* command in CONFIGURATION mode.

cipher-list-: Enter a space-delimited list of ciphers the SSH Client supports.

The following ciphers are available.

- `3des-cbc`
- `aes128-cbc`
- `aes192-cbc`

- aes256-cbc
- aes128-ctr
- aes192-ctr
- aes256-ctr

The default cipher list is in the given order: aes256-ctr, aes256-cbc, aes192-ctr, aes192-cbc, aes128-ctr, aes128-cbc, 3des-cbc.

Example of Configuring a Cipher List

The following example shows you how to configure a cipher list.

```
DellEMC(conf)#ip ssh cipher aes128-ctr aes128-cbc 3des-cbc
```

Secure Shell Authentication

Secure Shell (SSH) is enabled by default using the SSH Password Authentication method.

Enabling SSH Authentication by Password

Authenticate an SSH client by prompting for a password when attempting to connect to the Dell EMC Networking system. This setup is the simplest method of authentication and uses SSH version 2.

To enable SSH password authentication, use the following command.

- Enable SSH password authentication.

  CONFIGURATION mode

  ```
  ip ssh password-authentication enable
  ```

To view your SSH configuration, use the `show ip ssh` command from EXEC Privilege mode.

```
DellEMC(conf)#ip ssh server enable
DellEMC(conf)#ip ssh password-authentication enable
DellEMC# show ip ssh
SSH server                : enabled.
SSH server version        : v2.
SSH server vrf            : default.
SSH server ciphers        : 3des-cbc,aes128-cbc,aes192-cbc,aes256-cbc,aes128-ctr,aes192-
ctr,aes256-ctr.
SSH server macs           : hmac-md5,hmac-md5-96,hmac-sha1,hmac-sha1-96,hmac-
sha2-256,hmac-sha2-256-96.
SSH server kex algorithms : diffie-hellman-group-exchange-sha1,diffie-hellman-group1-
sha1,diffie-hellman-group14-sha1.
Password Authentication   : enabled.
Hostbased Authentication  : disabled.
RSA       Authentication  : disabled.
   Vty          Encryption      HMAC            Remote IP
```

Using RSA Authentication of SSH

The following procedure authenticates an SSH client based on an RSA key using RSA authentication. This method uses SSH version 2.

1. On the SSH client (Unix machine), generate an RSA key, as shown in the following example.
2. Copy the public key *id_rsa.pub* to the Dell EMC Networking system.
3. Disable password authentication if enabled.

   CONFIGURATION mode

   ```
   no ip ssh password-authentication enable
   ```
4. Enable RSA authentication in SSH.

   CONFIGURATION Mode

   ```
   ip ssh rsa-authentication enable
   ```

5. Install user's public key for RSA authentication in SSH.

   EXEC Privilege Mode

   ```
   ip ssh rsa-authentication username username my-authorized-keys flash://public_key
   ```

   If you provide the username, the Dell EMC Networking OS installs the public key for that specific user. In case, no user is associated with the current logged-in session, the system displays the following error message.

   NOTE: If no user is associated with the current logged-in session, the system displays the following error message. `% Error: No username set for this term.`

```
admin@Unix_client#ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/admin/.ssh/id_rsa):
/home/admin/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/admin/.ssh/id_rsa.
Your public key has been saved in /home/admin/.ssh/id_rsa.pub.
```

Configuring Host-Based SSH Authentication

Authenticate a particular host. This method uses SSH version 2.

To configure host-based authentication, use the following commands.

1. Configure RSA Authentication. Refer to Using RSA Authentication of SSH.
2. Create *shosts* by copying the public RSA key to the file *shosts* in the directory *.ssh*, and write the IP address of the host to the file.

   ```
   cp /etc/ssh/ssh_host_rsa_key.pub /.ssh/shosts
   ```

   Refer to the first example.
3. Create a list of IP addresses and usernames that are permitted to SSH in a file called *rhosts*.

   Refer to the second example.
4. Copy the file *shosts* and *rhosts* to the Dell EMC Networking system.
5. Disable password authentication and RSA authentication, if configured

   CONFIGURATION mode or EXEC Privilege mode

   `no ip ssh password-authentication` or `no ip ssh rsa-authentication`
6. Enable host-based authentication.

   CONFIGURATION mode

   ```
   ip ssh hostbased-authentication enable
   ```
7. Bind *shosts* and *rhosts* to host-based authentication.

   CONFIGURATION mode

   `ip ssh pub-key-file flash://filename` or `ip ssh rhostsfile flash://filename`

The following example shows creating *shosts*.

```
admin@Unix_client# cd /etc/ssh

admin@Unix_client# ls
moduli     sshd_config      ssh_host_dsa_key.pub      ssh_host_key.pub
ssh_host_rsa_key.pub ssh_config ssh_host_dsa_key ssh_host_key
ssh_host_rsa_key

admin@Unix_client# cat ssh_host_rsa_key.pub

ssh-rsa AAAAB3NzaC1yc2EAAAABIwAAAIEA8K7jLZRVfjgHJzUOmXxuIbZx/
AyWhVgJDQh39k8v3e8eQvLnHBIsqIL8jVy1QHhUeb7GaDlJVEDAMz30myqQbJgXBBRTWgBpLWwL/
doyUXFufjiL9YmoVTkbKcFmxJEMkE3JyHanEi7hg34LChjk9hL1by8cYZP2kYS2lnSyQWk=

admin@Unix_client# ls

id_rsa id_rsa.pub shosts
```

```
admin@Unix_client# cat shosts

10.16.127.201, ssh-rsa AAAAB3NzaC1yc2EAAAABIwAAAIEA8K7jLZRVfjgHJzUOmXxuIbZx/AyW
hVgJDQh39k8v3e8eQvLnHBIsqIL8jVy1QHhUeb7GaDlJVEDAMz30myqQbJgXBBRTWgBpLWwL/
doyUXFufjiL9YmoVTkbKcFmxJEMkE3JyHanEi7hg34LChjk9hL1by8cYZP2kYS2lnSyQWk=
```

The following example shows creating *rhosts*.

```
admin@Unix_client# ls
id_rsa id_rsa.pub rhosts shosts
admin@Unix_client# cat rhosts
10.16.127.201 admin
```

Using Client-Based SSH Authentication

To SSH from the chassis to the SSH client, use the following command.

If the SSH port is a non-default value, use the `ip ssh server port number` command to change the default port number. You may only change the port number when SSH is disabled. Then use the `-p` option with the `ssh` command.

- SSH from the chassis to the SSH client.

  `ssh` *`ip_address`*

```
DellEMC#ssh 10.16.127.201 ?
-c                      Encryption cipher to use (for v2 clients only)
-l                      User name option
-m                      HMAC algorithm to use (for v2 clients only)
-p                      SSH server port option (default 22)
-v                      SSH protocol version
```

Troubleshooting SSH

To troubleshoot SSH, use the following information.

You may not bind *id_rsa.pub* to RSA authentication while logged in via the console. In this case, this message displays:`%Error: No username set for this term.`

Enable host-based authentication on the server (Dell EMC Networking system) and the client (Unix machine). The following message appears if you attempt to log in via SSH and host-based is disabled on the client. In this case, verify that host-based authentication is set to “Yes” in the *file ssh_config* (root permission is required to edit this file): `permission denied (host based).`

If the IP address in the RSA key does not match the IP address from which you attempt to log in, the following message appears. In this case, verify that the name and IP address of the client is contained in the *file /etc/*hosts: `RSA Authentication Error.`

Telnet

To use Telnet with SSH, first enable SSH, as previously described.

By default, the Telnet daemon is enabled. If you want to disable the Telnet daemon, use the following command, or disable Telnet in the startup config. To enable or disable the Telnet daemon, use the `[no] ip telnet server enable` command.

The Telnet server or client is VRF-aware. You can enable a Telnet server or client to listen to a specific VRF by using the *`vrf vrf-instance-name`* parameter in the `telnet` command. This capability enables a Telent server or client to look up the correct routing table and establish a connection.

Example of Using Telnet for Remote Login

```
DellEMC(conf)#ip telnet server enable
DellEMC(conf)#no ip telnet server enable
```

VTY Line and Access-Class Configuration

Various methods are available to restrict VTY access in . These depend on which authentication scheme you use — line, local, or remote.

Table 90. VTY Access

Authentication Method	VTY access-class support?	Username access-class support?	Remote authorization support?
Line	YES	NO	NO
Local	NO	YES	NO
TACACS+	YES	NO	YES (with version 5.2.1.0 and later)
RADIUS	YES	NO	YES (with version 6.1.1.0 and later)

provides several ways to configure access classes for VTY lines, including:

- VTY Line Local Authentication and Authorization
- VTY Line Remote Authentication and Authorization

VTY Line Local Authentication and Authorization

retrieves the access class from the local database.

To use this feature:

1. Create a username.
2. Enter a password.
3. Assign an access class.
4. Enter a privilege level.

You can assign line authentication on a per-VTY basis; it is a simple password authentication, using an access-class as authorization.

Configure local authentication globally and configure access classes on a per-user basis.

can assign different access classes to different users by username. Until users attempt to log in, does not know if they will be assigned a VTY line. This means that incoming users always see a login prompt even if you have excluded them from the VTY line with a deny-all access class. After users identify themselves, retrieves the access class from the local database and applies it. (then can close the connection if a user is denied access.)

NOTE: If a VTY user logs in with RADIUS authentication, the privilege level is applied from the RADIUS server only if you configure RADIUS authentication.

The following example shows how to allow or deny a Telnet connection to a user. Users see a login prompt even if they cannot log in. No access class is configured for the VTY line. It defaults from the local database.

Example of Configuring VTY Authorization Based on Access Class Retrieved from a Local Database (Per User)

```
DellEMC(conf)#user gooduser password abc privilege 10 access-class permitall
DellEMC(conf)#user baduser password abc privilege 10 access-class denyall
DellEMC(conf)#
DellEMC(conf)#aaa authentication login localmethod local
DellEMC(conf)#
DellEMC(conf)#line vty 0 9
DellEMC(config-line-vty)#login authentication localmethod
DellEMC(config-line-vty)#end
```

VTY Line Remote Authentication and Authorization

retrieves the access class from the VTY line.

The takes the access class from the VTY line and applies it to ALL users. does not need to know the identity of the incoming user and can immediately apply the access class. If the authentication method is RADIUS, TACACS+, or line, and you have configured an access class for the VTY line, immediately applies it. If the access-class is set to deny all or deny for the incoming subnet, closes the connection without displaying the login prompt. The following example shows how to deny incoming connections from subnet 10.0.0.0 without displaying a login prompt. The example uses TACACS+ as the authentication mechanism.

Example of Configuring VTY Authorization Based on Access Class Retrieved from the Line (Per Network Address)

```
DellEMC(conf)#ip access-list standard deny10
DellEMC(conf-ext-nacl)#permit 10.0.0.0/8
DellEMC(conf-ext-nacl)#deny any
DellEMC(conf)#
DellEMC(conf)#aaa authentication login tacacsmethod tacacs+
DellEMC(conf)#tacacs-server host 256.1.1.2 key Force10
DellEMC(conf)#
DellEMC(conf)#line vty 0 9
DellEMC(config-line-vty)#login authentication tacacsmethod
DellEMC(config-line-vty)#
DellEMC(config-line-vty)#access-class deny10
DellEMC(config-line-vty)#end
(same applies for radius and line authentication)
```

VTY MAC-SA Filter Support

supports MAC access lists which permit or deny users based on their source MAC address.

With this approach, you can implement a security policy based on the source MAC address.

To apply a MAC ACL on a VTY line, use the same `access-class` command as IP ACLs.

The following example shows how to deny incoming connections from subnet 10.0.0.0 without displaying a login prompt.

Example of Configuring VTY Authorization Based on MAC ACL for the Line (Per MAC Address)

```
DellEMC(conf)#mac access-list standard sourcemac
DellEMC(config-std-mac)#permit 00:00:5e:00:01:01
DellEMC(config-std-mac)#deny any
DellEMC(conf)#
DellEMC(conf)#line vty 0 9
DellEMC(config-line-vty)#access-class sourcemac
DellEMC(config-line-vty)#end
```

Support for Change of Authorization and Disconnect Messages packets

The Network Access Server (NAS) uses RADIUS to authenticate AAA or dot1x user-access to the switch. The RADIUS service does not support unsolicited messages sent from the RADIUS server to the NAS.

However, there are many instances in which it is desirable for changes to be made to session characteristics, without requiring the NAS to initiate the exchange. For example, it may be desirable for administrators to be able to terminate user sessions in progress.

Alternatively, if the user changes authorization level, this change may require that authorization attributes be added or deleted from the user sessions.

To overcome these limitations, Dell EMC Networking OS provides RADIUS extension commands in order to enable unsolicited messages to be sent to the NAS. These extension commands provide support for Disconnect Messages (DMs) and Change-of-Authorization (CoA) packets. DMs cause user sessions to be terminated immediately; whereas, CoA packets modify session authorization attributes such as VLAN IDs, user privileges, and so on.

Change of Authorization (CoA) packets

Using the CoA packets, the NAS can handle authorization of dot1x sessions by processing the following requests from the Dynamic Authorization Client (DAC): Re-authentication of the supplicant, Port disable, and Port bounce.

The CoA packets constitute one message request (CoA request) and one of the following two possible responses:

- Change of Authorization Acknowledgement (CoA-Ack) - If the authorization state change is successful, then NAS sends a CoA-Ack.
- Change of Authorization non-Acknowledgement (CoA-Nak) - If the authorization state change is not successful, then the NAS sends a CoA-Nak, which is a negative acknowledgement.

Disconnect Messages

Using the Disconnect Messages, the NAS can disconnect AAA and dot1x sessions. NAS can disconnect AAA sessions using either username or a combination of the username and session id. NAS can disconnect dot1x sessions using NAS-port, or calling-station ID, or both.

The disconnect messages constitue one message request (DM request) and one of the following two possible responses:

- Disconnect Acknowledgement (DM-Ack) - If the session is disconnected successfully, then NAS sends a DM-Ack.
- Disconnect non-Acknowledgement (DM-Nak) - If the session is not disconnected successfully, then NAS sends a DM-Nak.

Attributes

In Disconnect messsage requests and CoA-Request packets, certain attributes are used to uniquely identify the NAS as well as user sessions on the NAS.

The combination of NAS and session identification attributes included in a CoA-request or a disconnect-message request must match at least one session in order for a request to be successful; otherwise, a disconnect-Nak or CoA-Nak is sent. For disconnect-user operations using DMs, if all NAS identification attributes match, and more than one session matches all of the session identification attributes, then a CoA-request or a disconnect-message request applies to all matching sessions.

The following tables describe the various types of attributes that identify the NAS and the user sessions:

Table 91. NAS Identification Attributes

Attribute code	Attribute	Description
4	NAS-IP-Address	IPv4 address of the NAS.
95	NAS-IPv6–Address	IPv6 address of the NAS.

Table 92. Change of Authorization (CoA) Attribute

Attribute code	Attribute	Description
5	NAS-Port	Port associated with the session to be processed for EAP or MAB users or the VTY ID for AAA sessions.

Table 93. Session Identification Attributes

Attribute code	Attribute	Description
31	Calling-Station-Id (MAC Address)	The link address from which session is connected.

Table 94. Vendor-specific Attributes

Attribute code	Attribute	Description
26	Vendor-specific	NAS supports the following values for the vendor-specific attributes:

Table 94. Vendor-specific Attributes

Attribute code	Attribute	Description
		• t=26(vendor-speific);l=*length*;*vendor-identification-attribute*;Length=*value*;data="cmd=re-authenticate" • t=26(vendor-speific);l=*length*;*vendor-identification-attribute*;Length=*value*;data="cmd=disable-host-port" • t=26(vendor-speific);l=*length*;*vendor-identification-attribute*;Length=*value*;data="cmd=bounce-host-port" • t=26(vendor-speific);l=*length*;*vendor-identification-attribute*;Length=*value*;data="cmd=terminate-session" • t=26(vendor-speific);l=*length*;*vendor-identification-attribute*;Length=*value*;data="cmd=disconnect-user" The vendor identification attribute can be one of the following: • v=9(Cisco);Vendor-Type=1(cisco-av-pair) Length = *value* • v=6027 (Force10);Vendor-Type=1(Force10-av-pair) Length = *value*

Table 95. DM Attributes

Attribute code	Attribute	Description
1	User-Name(Mandatory)	Name of the user associated with one or more sessions.

Mandatory attributes

The following tables describe the mandatory attributes for various message types:

Table 96. CoA EAP/MAB Session(s) Re-authenticate

Radius Attribute code	Radius Attribute	Description	Mandatory
NAS Identification Attributes			
4	NAS-IP-Address	IPv4 address of the NAS.	No
95	NAS-IPv6–Address	IPv6 address of the NAS.	No
Session Identification Attributes			
5	NAS-Port	Port on which session is terminated	Yes, If Calling-Station-Id attribute is not provided
31	Calling-Station-Id	The link address from which session is connected.	Yes, If NAS-Port attribute is not provided
Authorization Attributes			
26	Vendor-Specific	t=26(vendor-specific);l=*length*;*vendor-identification-attribute*;Length=*value*; Data="cmd=re-authenticate"	Yes

Table 97. CoA EAP/MAB Disable Port

Radius Attribute code	Radius Attribute	Description	Mandatory
NAS Identification Attributes			
4	NAS-IP-Address	IPv4 address of the NAS.	No
95	NAS-IPv6–Address	IPv6 address of the NAS.	No
Session Identification Attributes			
5	NAS-Port	Port on which session is terminated	Yes

Table 97. CoA EAP/MAB Disable Port (continued)

Radius Attribute code	Radius Attribute	Description	Mandatory
Authorization Attributes			
26	Vendor-Specific	t=26(vendor-specific);l=*length*;*vendor-identification-attribute*;Length=*value*; Data="cmd=bounce-host-port"	Yes

Table 98. CoA EAP/MAB Bounce Port

Radius Attribute code	Radius Attribute	Description	Mandatory
NAS Identification Attributes			
4	NAS-IP-Address	IPv4 address of the NAS.	No
95	NAS-IPv6–Address	IPv6 address of the NAS.	No
Session Identification Attributes			
5	NAS-Port	Port on which session is terminated	Yes
Authorization Attributes			
26	Vendor-Specific	t=26(vendor-specific);l=*length*;*vendor-identification-attribute*;Length=*value*; Data="cmd=bounce-host-port"	Yes

Table 99. DM EAP/MAB Session(s) disconnect

Radius Attribute code	Radius Attribute	Description	Mandatory
NAS Identification Attributes			
4	NAS-IP-Address	IPv4 address of the NAS.	No
95	NAS-IPv6–Address	IPv6 address of the NAS.	No
Session Identification Attributes			
5	NAS-Port	Port on which session is terminated	Yes, if Calling-Station-Id attribute is not provided
31	Calling-Station-Id	The link address from which session is connected.	Yes, If NAS-Port attribute is not provided
Authorization Attributes			
26	Vendor-Specific	t=26(vendor-specific);l=*length*;*vendor-identification-attribute*;Length=*value*; Data="cmd=terminate-session"	Yes

Table 100. DM AAA Session(s) disconnect

Radius Attribute code	Radius Attribute	Description	Mandatory
NAS Identification Attributes			
4	NAS-IP-Address	IPv4 address of the NAS.	No
95	NAS-IPv6–Address	IPv6 address of the NAS.	No
Session Identification Attributes			
1	User-Name	Name of the user associated with one or more sessions - AAA user name	Yes

Table 100. DM AAA Session(s) disconnect (continued)

Radius Attribute code	Radius Attribute	Description	Mandatory
5	NAS-Port	Port on which session is terminated	No
Authorization Attributes			
26	Vendor-Specific	t=26(vendor-specific);l=*length*;*vendor-identification-attribute*;Length=*value*; Data="cmd=disconnect-user"	Yes

Error-cause Values

It is possible that a Dynamic Authorization Server cannot honor Disconnect Message request or CoA request packets for some reason.

The Error-Cause Attribute provides more detail on the cause of the problem. It may be included within CoA-Nak and Disconnect-Nak packets.

The following table describes various error causes for the CoA and DM requests:

Table 101. Error-cause Values

Serial Number	Error-cause	Scenarios
1	Unsupported Attributes(401)	• CoA or DM request containing one or more unsupported attributes. • DM requests containing attributes other than NAS/Session identification attributes.
2	Invalid Attribute Value(407)	• CoA or DM request containing the incorrect NAS-Port, calling-station-id, and Vendor-Specific attribute values.
3	NAS Identification Mismatch(403)	• CoA request containing NAS-IP-Address or NAS-IPV6-Address that does not match NAS.
4	Administratively Prohibited(501)	• NAS is configured to ignore the CoA or DM request. Also, dot1x is not configured on the NAS-Port.
5	Session Context Not Found(503)	• CoA or DM request containing session identification attributes that does not match any of the NAS user sessions.
6	Resource Unavailable(506)	• Internal CoA or DM message processing errors.
7	Missing Attribute(402)	• CoA or DM request without Vendor-specific attribute or invalid Vendor-specific attribute. • CoA with re-authenticate or terminate request not containing calling-station-id or NAS-Port attribute. • CoA with disable-port or bounce-port request not containing NAS-Port attribute. • DM request not containing user-name attribute.

CoA Packet Processing

This section lists various actions that the NAS performs during CoA packet processing.

The following activities are performed by NAS:

- responds with CoA-Nak, if no matching session is found for the session identification attributes in CoA; Error-Cause value is "Session Context Not Found" (503).
- responds with CoA-Nak, for any internal processing error in NAS; Error-Cause value is "Resources Unavailable" (506).
- ignores attributes that are supported as per RFC but irrelevant to the CoA operations.
- responds to a CoA-Request containing one or more incorrect attribute values with a CoA-Nak; Error-Cause value is "Invalid Attribute Value" (407).

NOTE:

The Invalid Attribute Value Error-Cause is applicable to following scenarios:

- if the CoA request contains incorrect Vendor-Specific attribute value.
- if the CoA request contains incorrect NAS-port or calling-station-id values.

- rejects the CoA-Request containing NAS-IP-Address or NAS-IPV6-Address attribute that does not match the NAS with a CoA-Nak; Error-Cause value is “NAS Identification Mismatch” (403).
- responds with a CoA-Nak, if it is configured to prohibit honoring of corresponding CoA-Request messages; Error-Cause value is “Administratively Prohibited” (501).

NOTE:

The Administratively Prohibited Error-Cause is also applicable to following scenarios:

- if the dot1x feature is not enabled in the NAS-port.
- if the NAS-port state is administratively down.

CoA or DM Discard

This section lists various actions that the NAS performs during CoA or DM discard.

The following activities are performed by NAS:

- discards the packet, if dynamic authorization feature is not enabled in NAS.
- discards the packet, if the configured shared key entry is not found for the source IP address of the packet.
- discards the packet with invalid code field. NAS supports the following radius codes.
  - Disconnect-Request (40)
  - CoA-Request (43)
- discards the duplicate packets, if NAS is currently processing the original packet. NAS identifies the duplicate packet with the following fields:
  - Source IP address
  - Source UDP port
  - Identifier
  - VRF ID
- discards the packets, if length of the packet is shorter than the length field value.
- discards the packets, if length of the packet is shorter than 20 or longer than 4096.
- discards the packets, if request authenticator does not match the calculated MD5 checksum. NAS calculates the MD5 hash using following fields from the request:
  - Code
  - Identifier
  - Length
  - 16 Zero Octets
  - Request Attributes
  - Shared secret (based on the source IP address of the packet)
- discards the packets, if the message-authenticator received in the request is invalid. The message-authenticator is calculated using the following fields:
  - Code Type
  - Identifier
  - Length
  - Request Authenticator
  - Attributes

Disconnect Message Processing

This section lists various actions that the NAS performs during DM processing.

The following activities are performed by NAS:

- responds with DM-Nak, if no matching session is found in NAS for the session identification attributes in DM; Error-Cause value is “Session Context Not Found” (503).

- responds with DM-Nak for any internal processing error in NAS; Error-Cause value is “Resources Unavailable” (506).
- ignores attributes that are supported as per RFC but are irrelevant to the DM operation.
- responds to a disconnect message containing one or more incorrect attributes values with a Disconnect-NAK; Error-Cause value is “Invalid Attribute Value” (407).
- responds to a disconnect message containing unsupported attributes with DM-Nak; Error-Cause value is “Unsupported Attributes” (401).

  NOTE: Unsupported attributes are the ones that are not mentioned in the RFC 5176 but present in the disconnect message that is received by the NAS.
- rejects the disconnect message containing NAS-IP-Address or NAS-IPV6-Address attribute that does not match NAS with DM-Nak; Error-Cause value is “NAS Identification Mismatch” (403).
- responds with a DM-Nak, if the NAS is configured to prohibit honoring of disconnect messages; Error-Cause value is “Administratively Prohibited” (501).

Configuring DAC

You can configure trusted dynamic authorization clients (DACs).

This setting enables you to configure more than one DAC. Duplicate configurations are not allowed.

1. Enter the following command to enter dynamic authorization mode:
   `radius dynamic-auth`
2. Enter the following command to configure DAC:
   `client host-name`

```
Dell(conf-dynamic-auth#)client testhost
```

Configuring the port number

You can configure the port number on which the NAS receives CoA or DM requests.

This setting enables you to specify an optional port number on which to receive CoA or DM requests. The default value is 3799.

Enter the following command to configure the port number:
`port port-number`
The range for the port number value that you can specify is from 1 to 65535.

```
Dell(conf-dynamic-auth#)port 2000
```

Configuring shared key

You can configure a global shared key for the dynamic authorization clients (DACs).

1. Enter the following command to enter dynamic authorization mode:
   `radius dynamic-auth`
2. Enter the following command to configure the global shared key value:
   `client-key encryption-type key`

```
Dell(conf-dynamic-auth#)client-key 7 password
```

Disconnecting administrative users logged in through RADIUS

Dell EMC Networking OS enables you to configure disconnect messages (DMs) to disconnect RADIUS administrative users who are logged in through an AAA interface.

Before disconnecting an administrative user using the disconnect messages, ensure that the following prerequisites are satisfied:

- Shared key is configured in NAS for DAC.
- NAS server listens on the Management IP UDP port 3799 (default) or the port configured through CLI.
- AAA session for the user is active.

NAS uses the user-name or both the user-name as well as the NAS-Port attribute to identify the AAA user session. NAS disconnects all sessions related to the user, if the user-name is provided without NAS-port.

1. Enter the following command to configure the dynamic authorization feature:
   `radius dynamic-auth`
2. Enter the following command to terminate the 802.1x user session:
   `disconnect-user`
   NAS disconnects the administrative users who are connected through an AAA interface.

```
Dell(conf#)radius dynamic-auth
Dell(conf-dynamic-auth#)disconnect-user
```

NAS takes the following actions:

- validates the DM request and the session identification attributes.
- sends a DM-Nak with an error-cause of 402 (missing attribute), if the DM request does not contain the User-Name.
- sends a DM-Ack, if it is able to successfully disconnect the admin user.
- sends a DM-Nak with an error-cause value of 506 (resource unavailable), if it is not able to disconnect the admin user.
- sends a DM-Nak with an error-cause value of 501 (administratively prohibited), if disconnect-user feature is not enabled in NAS.

Configuring CoA to bounce 802.1x enabled ports

Dell EMC Networking OS provides RADIUS extension commands that enables you to configure port bounce settings for the 802.1x enabled port.

Before configuring port bounce settings on a 802.1x enabled port, ensure that the following prerequisites are satisfied:

- Shared key is configured in NAS for DAC.
- NAS server listens on the Management IP UDP port 3799 (default) or the port configured through CLI.
- The user is logged-in through 802.1X enabled physical port and successfully authenticated with Radius Server.

When DAC initiates a port bounce operation, the NAS server causes the links on the authentication port to flap. This incident in turn triggers re-negotiation on one of the ports that is flapped.

1. Enter the following command to configure the dynamic authorization feature:
   `radius dynamic-auth`
2. Enter the following command to configure port-bounce setttings on a 802.1x enabled port:
   `coa-bounce-port`
   NAS disables the authentication port that is hosting the session and re-enables it after 10 seconds. All user sessions connected to this authentication port are affected.

```
Dell(conf#)radius dynamic-auth
Dell(conf-dynamic-auth#)coa-bounce-port
```

NAS takes the following actions whenever port-bounce is triggered:

- validates the CoA request and the session identification attributes.
- sends a CoA-Nak with an error-cause of 402 (missing attribute), if the CoA request does not contain the NAS-port attributes.
- uses the NAS-port attribute to identify the 802.1x enabled interface.
- sends a CoA-Nak with an error-cause value of 503 (session context not found), if it is unable to retrieve 802.1x enabled interface using the NAS-port attribute.
- sends a CoA-Ack if it is successfully able to flap the port.
- discards the packet, if simultaneous requests are received for the same NAS Port.

Configuring CoA to re-authenticate 802.1x sessions

Dell EMC Networking OS provides RADIUS extension commands that enables you to configure re-authentication of 802.1x user sessions. When you configure this feature, the DAC sends the CoA request to re-authenticate the 802.1x uer session when ever the authorization level of the user's profile changes.

Before configuring re-authentication of 802.1x sessions, ensure that the following prerequisites are satisfied:

- Shared key is configured in NAS for DAC.
- NAS server listens on the Management IP UDP port 3799 (default) or the port configured through CLI.

- The user is logged-in through 802.1X enabled physical port and successfully authenticated with Radius Server.

To initiate 802.1x session re-authentication, the DAC sends a standard CoA request that contains one or more session identification attributes. NAS uses the calling-station-id or the NAS-port attributes to identify a 802.1x user session. In case of the EAP or MAB users, the MAC address is the calling-station-id of the supplicant and the NAS-port is the interface identifier. If both these attributes are present in the CoA request, NAS retrieves the supplicant connected to the interface. The EAP or MAB user sessions are re-authenticated and the NAS sends a CoA-Ack to the user, in case the re-authentication is successful.

1. Enter the following command to configure the dynamic authorization feature:
   `radius dynamic-auth`
2. Enter the following command to configure the re-authentication of 802.1x sessions:
   `coa-reauthenticate`
   NAS re-initiates the user authentication state.

```
Dell(conf#)radius dynamic-auth
Dell(conf-dynamic-auth#)coa-reauthenticate
```

NAS takes the following actions whenever re-authentication is triggered:

- validates the CoA request and the session identification attributes.
- sends a CoA-Nak with an error-cause of 402 (missing attribute), if the CoA request does not contain both the calling-station-id as well as the NAS-port attribute.
- sends a CoA-Ack if the re-authentication of the 802.1x session is successful.
- sends a CoA-Nak with an error-cause value of 506 (resource unavailable), if it is unable to initiate the re-authentication process.
- sends a CoA-Nak if user authentication fails due to unresponsive supplicant or RADIUS server.
- sends a CoA-Ack, if the user is configured with static MAB profile.
- discards the packet, if simultaneous requests are received for the same calling-station-id or NAS-port or both.
- returns an error-cause value of 503 (session context not found), if it is not able to retrieve the session using the calling-station-id or NAS-port attribute or both.
- sends NAK if user is configured with forced-unauthorization.
- sends-ACK if user is configured with forced-authorization.

Terminating the 802.1x user session

Dell EMC Networking OS provides RADIUS extension commands that terminate the 802.1x user session. When this request is initiated, the NAS disconnects the 802.1x user session without disabling the physical port that authenticated the current session.

Before terminating the 802.1x user session, ensure that the following prerequisites are satisfied:

- Shared key is configured in NAS for DAC.
- NAS server listens on the Management IP UDP port 3799 (default) or the port configured through CLI.
- The user is logged-in through 802.1X enabled physical port and successfully authenticated with Radius Server.

NAS uses the calling-station-id or the NAS-port attributes to identify the 802.1x session. In case of the EAP and MAB users, the calling-station-id is the MAC address of the supplicant and the NAS-port attribute is the interface identifier. Using these atrributes, the NAS retrieves the supplicant that is connected to the interface.

1. Enter the following command to configure the dynamic authorization feature:
   `radius dynamic-auth`
2. Enter the following command to terminate the 802.1x user session:
   `terminate-session`
   NAS terminates the 802.1x user session without disabling the physical port.

```
Dell(conf#)radius dynamic-auth
Dell(conf-dynamic-auth#)terminate-session
```

NAS takes the following actions whenever session termination is triggered:

- validates the DM request and the session identification attributes.
- sends a DM-Nak with an error-cause of 402 (missing attribute), if the DM request does not contain the calling-station-id and NAS-port attributes.
- returns an error-cause value of 503 (session context not found), if it is not able to retrieve the session using the calling-station-id or NAS-port attribute or both.
- sends a DM-Ack, if it is able to terminate the session.

- sends a DM-Nak with an error-cause value of 506 (resource unavailable), if it is not able to apply changes to the existing session.
- discards the packet, if simultaneous requests are received for the same NAS-port or calling-station-id, or both.

Disabling 802.1x enabled port

Dell EMC Networking OS provides RADIUS extension commands that enables you to disable 802.1x enabled ports. This command administratively shuts down the port causing the termination of the dot1x user session. This command is useful when a port is known to cause issue in the network and needs to be disabled.

Before disabling the 802.1x enabled port, ensure that the following prerequisites are satisfied:

- Shared key is configured in NAS for DAC.
- NAS server listens on the Management IP UDP port 3799 (default) or the port configured through CLI.
- The user is logged-in through 802.1X enabled physical port and successfully authenticated with Radius Server.

To initiate shutting down of the 802.1x enabled port, the DAC sends a standard CoA request that contains one or more session identification attributes. NAS uses the NAS-port attributes to identify the 802.1x enabled physical port.

1. Enter the following command to configure the dynamic authorization feature:
   `radius dynamic-auth`
2. Enter the following command to disable the 802.1x enabled physical port:
   `coa-disable-port`
   NAS administratively shuts down the 802.1x enabled port that is hosting the session. You can re-enable this port only through a non-RADIUS mechanism or through bounce-port request.

```
Dell(conf#)radius dynamic-auth
Dell(conf-dynamic-auth#)coa-disable-port
```

NAS takes the following actions:

- validates the CoA request and the session identification attributes.
- sends a CoA-Nak with an error-cause of 402 (missing attribute), if the CoA request does not contain the NAS-port attribute.
- returns an error-cause value of 503 (session context not found), if it is not able to retrieve the port information using the NAS-port attribute.
- sends a CoA-Ack, if it is able to successfully disable the 802.1x enabled port.
- sends a CoA-Nak with an error-cause value of 506 (resource unavailable), if it is not able to disable the 802.1x enabled port.
- discards the packet, if simultaneous requests are received for the same NAS Port.

Important points to remember

Virtual link truncking (VLT) scenario

This section describes how the secondary NAS processes the PE port authorization RADIUS requests to the primary NAS.

- The NAS VLT chassis member processes the RADIUS dynamic authorization message locally if the role of chassis is primary.
- The NAS secondary VLT chassis member forwards the RADIUS dynamic authorization message authorizing dual-homed Port Extender (PE) ports to the primary VLT peer. NAS secondary VLT chassis member forwards the response to DAC after receiving it from the primary VLT peer.
- The NAS VLT secondary chassis member processes the RADIUS dynamic authorization message authorizing non-PE Control Bridge (CB) ports locally.

RPM failover scenario

This section describes how the NAS handles virtual IP failovers to the secondary RPM.

- The NAS Route Processor Module (RPM) processes the RADIUS dynamic authorization message only if the role of RPM is active.
- The NAS standby RPM processes the retransmitted CoA or DM messages without requiring a chassis reboot if primary RPM fails and standby becomes primary.

Stack failover scenario

This section describes the stack failover scenario.

- The NAS stacking module processes the RADIUS dynamic authorization messages only if the role of module is master.
- The NAS standby stacking module processes the retransmitted CoA or DM messages without requiring a chassis reboot, if the master module fails and the standby module becomes the master.

Configuring replay protection

NAS enables you to configure the replay protection window period.

NAS drops the packets if duplicate packets are received within replay protection window period. The default value is 5 minutes.

Enter the following command to configure replay protection:
`replay-prot-window minutes`
NAS considers the new replay protection window value from next window period. The range is from 1 to 10 minutes. The default is 5 minutes.

```
Dell(conf-dynamic-auth#)replay-prot-window 10
```

Rate-limiting RADIUS packets

NAS enables you to allow or reject RADIUS dynamic authorization packets based on the rate-limiting value that you specify.

NAS lets you to configure number of RADIUS dynamic authorization packets allowed per minute. The default value is 30 packets per minute. NAS discards the packets, if the number of RADIUS dynamic authorization packets in the current interval cross the configured rate-limit value.

Enter the following command to configure rate-limiting:
`rate-limit number`
NAS considers the rate limit change value from the next interval period. The range is from 10 to 60 packets per minute. The default is 30 packets per minute.

```
Dell(conf-dynamic-auth#)rate-limit 50
```

Configuring time-out value

You can configure a time-out value for the back-end task to respond to CoA or DM requests.

This setting enables the DAS to determine the amount of time to wait before a back-end response is received. The default value is 10 minutes.

Enter the following command to configure the time-out value:
`da-rsp-timeout value`

```
Dell(conf-dynamic-auth#)da-rsp-timeout 20
```

Role-Based Access Control

With Role-Based Access Control (RBAC), access and authorization is controlled based on a user's role. Users are granted permissions based on their user roles, not on their individual user ID. User roles are created for job functions and through those roles they acquire the permissions to perform their associated job function.

This chapter consists of the following sections:

- Overview
- Privilege-or-Role Mode Versus Role-only Mode
- Configuring Role-based Only AAA Authorization

- System-Defined RBAC User Roles
- Creating a New User Role
- Modifying Command Permissions for Roles
- Adding and Deleting Users from a Role
- Role Accounting
- Configuring AAA Authentication for Roles
- Configuring AAA Authorization for Roles
- Configuring an Accounting for Roles
- Applying an Accounting Method to a Role
- Displaying Active Accounting Sessions for Roles
- Configuring TACACS+ and RADIUS VSA Attributes for RBAC
- Displaying User Roles
- Displaying Accounting for User Roles
- Displaying Information About Roles Logged into the Switch
- Display Role Permissions Assigned to a Command

Overview of RBAC

With Role-Based Access Control (RBAC), access and authorization is controlled based on a user's role. Users are granted permissions based on their user roles, not on their individual user ID. User roles are created for job functions and through those roles they acquire the permissions to perform their associated job function. Each user can be assigned only a single role. Many users can have the same role.

The Dell EMC Networking OS supports the constrained RBAC model. With a constrained RBAC model, you can inherit permissions when you create a new user role, restrict or add commands a user can enter and the actions the user can perform. This allows for greater flexibility in assigning permissions for each command to each role and as a result, it is easier and much more efficient to administer user rights. If a user's role matches one of the allowed user roles for that command, then command authorization is granted.

A constrained RBAC model provides for separation of duty and as a result, provides greater security than the hierarchical RBAC model. Essentially, a constrained model puts some limitations around each role's permissions to allow you to partition of tasks. However, some inheritance is possible.

Default command permissions are based on CLI mode (such as configure, interface, router), any specific command settings, and the permissions allowed by the privilege and role commands. The role command allows you to change permissions based on the role. You can modify the permissions specific to that command and/or command option. For more information, see Modifying Command Permissions for Roles .

NOTE: When you enter a user role, you have already been authenticated and authorized. You do not need to enter an enable password because you will be automatically placed in EXEC Priv mode.

For greater security, the ability to view event, audit, and security system log is associated with user roles. For information about these topics, see Audit and Security Logs.

Privilege-or-Role Mode versus Role-only Mode

By default, the system provides access to commands determined by the user's role or by the user's privilege level. The user's role takes precedence over a user's privilege level. If the system is in "privilege or role" mode, then all existing user IDs can continue to access the switch even if they do not have a user role defined. To change to more secure mode, use role-based AAA authorization. When role-based only AAA authorization is configured, access to commands is determined only by the user's role. For more information, see **Configuring Role-based Only AAA Authorization.**

Configuring Role-based Only AAA Authorization

You can configure authorization so that access to commands is determined only by the user's role. If the user has no user role, access to the system is denied as the user is not able to login successfully. When you enable role-based only AAA authorization using the `aaa authorization role-only` command in Configuration mode, the Dell EMC Networking OS checks to ensure that you do not lock yourself out and that the user authentication is available for all terminal lines.

Pre-requisites

Before you enable role-based only AAA authorization:

1. Locally define a system administrator user role. This gives you access to login with full permissions even if network connectivity to remote authentication servers is not available.
2. Configure login authentication on the console. This ensures that all users are properly identified through authentication no matter the access point.

   If you do not configure login authentication on the console, the system displays an error when you attempt to enable role-based only AAA authorization.

3. Specify an authentication method list—RADIUS, TACACS+, or Local.

   You must specify at least local authentication. For consistency, the best practice is to define the same authentication method list across all lines, in the same order of comparison; for example VTY and console port.

   You could also use the default authentication method to apply to all the LINES; for example, console port and VTY.

   NOTE: The authentication method list must be in the same order as the authorization method list. For example, if you configure the authentication method list in the following order—TACACS+, local—Dell EMC Networking recommends configuring the authorization method list in the same order—TACACS+, local.

4. Specify the authorization method list—RADIUS, TACACS+, or Local. At a minimum, you must specify local authorization.

   For consistency, the best practice is to define the same authorization method list across all lines, in the same order of comparison; for example, VTY and console port.

   You could also use the default authorization method list to apply to all the LINES—console port, VTY.

   If you do not, the following error displays when you attempt to enable role-based only AAA authorization:

   ```
   % Error: Exec authorization must be applied to more than one line to be useful, e.g.
   console and vty lines. Could use default authorization method list as alternative.
   ```

5. Verify the configuration is applied to the console or VTY line.

   ```
   DellEMC(conf)#do show running-config line
   !
   line console 0
   login authentication test
   authorization exec test
   exec-timeout 0 0
   line vty 0
   login authentication test
   authorization exec test
   line vty 1
   login authentication test
   authorization exec test
   ```

To enable role-based only AAA authorization, enter the following command in Configuration mode:

```
DellEMC(conf)#aaa authorization role-only
```

System-Defined RBAC User Roles

By default, the Dell EMC Networking OS provides 4 system defined user roles. You can create up to 8 additional user roles.

NOTE: You cannot delete any system defined roles.

The system defined user roles are as follows:

- Network Operator (netoperator) - This user role has no privilege to modify any configuration on the switch. You can access Exec mode (monitoring) to view the current configuration and status information.
- Network Administrator (netadmin): This user role can configure, display, and debug the network operations on the switch. You can access all of the commands that are available from the network operator user role. This role does not have access to the commands that are available to the system security administrator for cryptography operations, AAA, or the commands reserved solely for the system administrator.
- Security Administrator (secadmin): This user role can control the security policy across the systems that are within a domain or network topology. The security administrator commands include FIPS mode enablement, password policies, inactivity timeouts, banner establishment, and cryptographic key operations for secure access paths.
- System Administrator (sysadmin). This role has full access to all the commands in the system, exclusive access to commands that manipulate the file system formatting, and access to the system shell. This role can also create user IDs and user roles.

The following summarizes the modes that the predefined user roles can access.

Role	Modes
netoperator	
netadmin	Exec Config Interface Router IP Route-map Protocol MAC
secadmin	Exec Config Line
sysadmin	Exec Config Interface Line Router IP Route-map Protocol MAC

User Roles

This section describes how to create a new user role and configure command permissions and contains the following topics.

- Creating a New User Role
- Modifying Command Permissions for Roles
- Adding and Deleting Users from a Role

Creating a New User Role

Instead of using the system defined user roles, you can create a new user role that best matches your organization. When you create a new user role, you can first inherit permissions from one of the system defined roles. Otherwise you would have to create a user role's command permissions from scratch. You then restrict commands or add commands to that role

NOTE: You can change user role permissions on system pre-defined user roles or user-defined user roles.

Important Points to Remember

Consider the following when creating a user role:

- Only the system administrator and user-defined roles inherited from the system administrator can create roles and user names. Only the system administrator, security administrator, and roles inherited from these can use the "role" command to modify command permissions. The security administrator and roles inherited by security administrator can only modify permissions for commands they already have access to.
- Make sure you select the correct role you want to inherit.
- If you inherit a user role, you cannot modify or delete the inheritance. If you want to change or remove the inheritance, delete the user role and create it again. If the user role is in use, you cannot delete the user role.

1. Create a new user role

   CONFIGURATION mode

   ```
   userrole name [inherit existing-role-name]
   ```

2. Verify that the new user role has inherited the security administrator permissions.

   ```
   DellEMC(conf)#do show userroles
   ```

   EXEC Privilege mode

3. After you create a user role, configure permissions for the new user role.

Example of Creating a User Role

The configuration in the following example creates a new user role, **myrole**, which inherits the security administrator (secadmin) permissions.

Create a new user role, **myrole** and inherit security administrator permissions.

```
DellEMC(conf)#userrole myrole inherit secadmin
```

Verify that the user role, **myrole**, has inherited the security administrator permissions. The output highlighted in **bold** indicates that the user role has successfully inherited the security administrator permissions.

```
DellEMC(conf)#do show userroles

************* Mon Apr 28 14:46:25 PDT 2014 **************

Authorization Mode:  role or privilege
Role         Inheritance  Modes
netoperator
netadmin                  Exec Config Interface Router IP Route-map Protocol MAC
secadmin                  Exec Config Line
sysadmin                  Exec Config Interface Line Router IP Route-map Protocol MAC.
myrole       secadmin     Exec Config Line
```

Modifying Command Permissions for Roles

You can modify (add or delete) command permissions for newly created user roles and system defined roles using the `role mode { { { addrole | deleterole } role-name } | reset } command` command in Configuration mode.

NOTE: You cannot modify system administrator command permissions.

If you add or delete command permissions using the `role` command, those changes only apply to the specific user role. They do not apply to other roles that have inheritance from that role. Authorization and accounting only apply to the roles specified in that configuration.

When you modify a command for a role, you specify the role, the mode, and whether you want to restrict access using the `deleterole` keyword or grant access using the `addrole` keyword followed by the command you are controlling access.

The following output displays the modes available for the `role` command.

```
DellEMC(conf)#role  ?
configure            Global configuration mode
exec                 Exec Mode
interface            Interface configuration mode
line                 Line Configuration mode
route-map            Route map configuration mode
router               Router configuration mode
```

Examples: Deny Network Administrator from Using the show users Command.

The following example denies the `netadmin` role from using the `show users` command and then verifies that `netadmin` cannot access the `show users` command in exec mode. Note that the `netadmin` role is not listed in the `Role access: secadmin,sysadmin`, which means the `netadmin` cannot access the `show users` command.

```
DellEMC(conf)#role exec deleterole netadmin show users

DellEMC#show role mode exec show users
Role access: secadmin,sysadmin
```

Example: Allow Security Administrator to Configure Spanning Tree

The following example allows the security administrator (secadmin) to configure the spanning tree protocol. Note *command* is protocol spanning-tree.

```
DellEMC(conf)#role configure addrole secadmin protocol spanning-tree
```

Example: Allow Security Administrator to Access Interface Mode

The following example allows the security administrator (`secadmin`) to access Interface mode.

```
DellEMC(conf)#role configure addrole secadmin ?
LINE       Initial keywords of the command to modify
DellEMC(conf)#role configure addrole secadmin interface
```

Example: Allow Security Administrator to Access Only 10-Gigabit Ethernet Interfaces

The following example allows the security administrator (`secadmin`) to only access 10-Gigabit Ethernett interfaces and then shows that the `secadmin`, highlighted in bold, can now access Interface mode. However, the `secadmin` can only access 10-Gigabit Ethernet interfaces.

```
DellEMC(conf)#role configure addrole secadmin ?
LINE            Initial keywords of the command to modify
DellEMC(conf)#role configure addrole secadmin interface tengigabitethernet

DellEMC(conf)#show role mode configure interface
Role access: netadmin, secadmin, sysadmin
```

Example: Verify that the Security Administrator Can Access Interface Mode

The following example shows that the `secadmin` role can now access Interface mode (highlighted in bold).

```
Role        Inheritance  Modes
netoperator
netadmin                 Exec Config Interface Router IP RouteMap Protocol MAC
secadmin                 Exec Config Interface Line
sysadmin                 Exec Config Interface Line Router IP RouteMap Protocol MAC
```

Example: Remove Security Administrator Access to Line Mode.

The following example removes the `secadmin` access to LINE mode and then verifies that the security administrator can no longer access LINE mode, using the `show role mode configure line` command in EXEC Privilege mode.

```
DellEMC(conf)#role configure deleterole secadmin ?
LINE          Initial keywords of the command to modify
DellEMC(conf)#role configure deleterole secadmin line

DellEMC(conf)#do show role mode ?
configure                      Global configuration mode
exec                 Exec Mode
interface            Interface configuration mode
line                 Line Configuration mode
route-map            Route map configuration mode
router               Router configuration mode

DellEMC(conf)#do show role mode configure line
Role access:sysadmin
```

Example: Grant and Remove Security Administrator Access to Configure Protocols

By default, the system defined role, `secadmin`, is not allowed to configure protocols. The following example first grants the `secadmin` role to configure protocols and then removes access to configure protocols.

```
DellEMC(conf)#role configure addrole secadmin protocol
DellEMC(conf)#role configure deleterole secadmin protocol
```

Example: Resets Only the Security Administrator role to its original setting.

The following example resets only the `secadmin` role to its original setting.

```
DellEMC(conf)#no role configure addrole secadmin protocol
```

Example: Reset System-Defined Roles and Roles that Inherit Permissions

In the following example the command protocol permissions are reset to their original setting or one or more of the system-defined roles and any roles that inherited permissions from them.

```
DellEMC(conf)#role configure reset protocol
```

Adding and Deleting Users from a Role

To create a user name that is authenticated based on a user role, use the `username` *name* password *encryption-type* `password role` *role-name* command in CONFIGURATION mode.

Example

The following example creates a user name that is authenticated based on a user role.

```
DellEMC(conf)# username john password 0 password role secadmin
```

The following example deletes a user role.

NOTE: If you already have a user ID that exists with a privilege level, you can add the user role to username that has a privilege

```
DellEMC(conf)# no username john
```

The following example adds a user, to the secadmin user role.

```
DellEMC(conf)# username john role secadmin password 0 password
```

AAA Authentication and Authorization for Roles

This section describes how to configure AAA Authentication and Authorization for Roles.

Configuration Task List for AAA Authentication and Authorization for Roles

This section contains the following AAA Authentication and Authorization for Roles configuration tasks:

- Configuring AAA Authentication for Roles
- Configuring AAA Authorization for Roles
- Configuring TACACS+ and RADIUS VSA Attributes for RBAC

Configure AAA Authentication for Roles

Authentication services verify the user ID and password combination. Users with defined roles and users with privileges are authenticated with the same mechanism. There are six methods available for authentication: **radius, tacacs+, local, enable, line,** and **none**.

When role-based only AAA authorization is enabled, the **enable, line,** and **none** methods are not available. Each of these three methods allows users to be verified with either a password that is not specific to their user ID or with no password at all. Because of the lack of security these methods are not available for role only mode. When the system is in role-only mode, users that have only privilege levels are denied access to the system because they do not have a role. For information about role only mode, see Configuring Role-based Only AAA Authorization.

NOTE: Authentication services only validate the user ID and password combination. To determine which commands are permitted for users, configure authorization. For information about how to configure authorization for roles, see Configure AAA Authorization for Roles.

To configure AAA authentication, use the **aaa authentication** command in CONFIGURATION mode.

```
aaa authentication login {method-list-name | default} method [… method4]
```

Configure AAA Authorization for Roles

Authorization services determine if the user has permission to use a command in the CLI. Users with only privilege levels can use commands in privilege-or-role mode (the default) provided their privilege level is the same or greater than the privilege level of those commands. Users with defined roles can use commands provided their role is permitted to use those commands. Role inheritance is also used to determine authorization.

Users with roles and privileges are authorized with the same mechanism. There are six methods available for authorization: `radius`, `tacacs+`, `local`, `enable`, `line`, and `none`.

When role-based only AAA authorization is enabled, the `enable`, `line`, and `none` methods are not available. Each of these three methods allows users to be authorized with either a password that is not specific to their userid or with no password at all. Because of the lack of security, these methods are not available for role-based only mode.

To configure AAA authorization, use the `aaa authorization exec` command in CONFIGURATION mode. The `aaa authorization exec` command determines which CLI mode the user will start in for their session; for example, Exec mode or Exec Privilege mode. For information about how to configure authentication for roles, see Configure AAA Authentication for Roles.

```
aaa authorization exec {method-list-name | default} method [… method4]
```

You can further restrict users' permissions, using the `aaa authorization command` command in CONFIGURATION mode.

```
aaa authorization command {method-list-name | default} method [… method4]
```

Examples of Applying a Method List

The following configuration example applies a method list: TACACS+, RADIUS and local:

```
!
radius-server host 10.16.150.203 key <clear-text>
!
tacacs-server host 10.16.150.203 key  <clear-text>
!
aaa authentication login ucraaa tacacs+ radius local
aaa authorization exec ucraaa tacacs+ radius local
aaa accounting commands role netadmin ucraaa start-stop tacacs+
!
```

The following configuration example applies a method list other than default to each VTY line.

NOTE: Note that the methods were not applied to the console so the default methods (if configured) are applied there.

```
!
line console 0
exec-timeout 0 0
line vty 0
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 1
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 2
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 3
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 4
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 5
login authentication ucraaa
```

```
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 6
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 7
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 8
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
line vty 9
login authentication ucraaa
authorization exec ucraaa
accounting commands role netadmin ucraaa
!
```

Configuring TACACS+ and RADIUS VSA Attributes for RBAC

For RBAC and privilege levels, the Dell EMC Networking OS RADIUS and TACACS+ implementation supports two vendor-specific options: privilege level and roles. The Dell EMC Networking vendor-ID is `6027` and the supported option has attribute of type string, which is titled “Force10-avpair”. The value is a string in the following format:

```
protocol : attribute sep value
```

“attribute” and “value” are an attribute-value (AV) pair defined in the Dell EMC Networking OS TACACS+ specification, and “sep” is “=”. These attributes allow the full set of features available for TACACS+ authorization and are authorized with the same attributes for RADIUS.

Example for Configuring a VSA Attribute for a Privilege Level 15

The following example configures an AV pair which allows a user to login from a network access server with a privilege level of 15, to have access to EXEC commands.

The format to create a Dell EMC Networking AV pair for privilege level is `shell:priv-lvl=<number>` where number is a value between 0 and 15.

`Force10-avpair= "shell:priv-lvl=15"`

Example for Creating a AVP Pair for System Defined or User-Defined Role

The following section shows you how to create an AV pair to allow a user to login from a network access server to have access to commands based on the user’s role. The format to create an AV pair for a user role is `Force10-avpair= "shell:role=<user-role>"` where *user-role* is a user defined or system-defined role.

In the following example, you create an AV pair for a system-defined role, sysadmin.

```
Force10-avpair= "shell:role=sysadmin"
```

In the following example, you create an AV pair for a user-defined role. You must also define a role, using the `userrole myrole inherit` command on the switch to associate it with this AV pair.

```
Force10-avpair= "shell:role=myrole"
```

The string, “myrole”, is associated with a TACACS+ user group. The user IDs are associated with the user group.

Role Accounting

This section describes how to configure role accounting and how to display active sessions for roles.

This sections consists of the following topics:

- Configuring AAA Accounting for Roles
- Applying an Accounting Method to a Role

- Displaying Active Accounting Sessions for Roles

Configuring AAA Accounting for Roles

To configure AAA accounting for roles, use the **aaa accounting** command in CONFIGURATION mode.

```
aaa accounting {system | exec | commands {level | role role-name}} {name | default}
{start-stop | wait-start | stop-only} {tacacs+}
```

Example of Configuring AAA Accounting for Roles

The following example shows you how to configure AAA accounting to monitor commands executed by the users who have a `secadmin` user role.

```
DellEMC(conf)#aaa accounting command role secadmin default start-stop tacacs+
```

Applying an Accounting Method to a Role

To apply an accounting method list to a role executed by a user with that user role, use the `accounting` command in LINE mode.

```
accounting {exec | commands {level | role role-name}} method-list
```

Example of Applying an Accounting Method to a Role

The following example applies the accounting default method to the user role secadmin (security administrator).

```
DellEMC(conf-vty-0)# accounting commands role secadmin default
```

Displaying Active Accounting Sessions for Roles

To display active accounting sessions for each user role, use the `show accounting` command in EXEC mode.

Example of Displaying Active Accounting Sessions for Roles

```
DellEMC#show accounting
Active accounted actions on tty2, User john Priv 1 Role netoperator
Task ID 1, EXEC Accounting record, 00:00:30 Elapsed,
service=shell
Active accounted actions on tty3, User admin Priv 15 Role sysadmin
Task ID 2, EXEC Accounting record, 00:00:26 Elapsed,
service=shell
```

Display Information About User Roles

This section describes how to display information about user roles and consists of the following topics:

- Displaying User Roles
- Displaying Information About Roles Logged into the Switch
- Displaying Active Accounting Sessions for Roles

Displaying User Roles

To display user roles using the `show userrole` command in EXEC Privilege mode, use the `show userroles` and `show users` commands in EXEC privilege mode.

Example of Displaying User Roles

```
DellEMC#show userroles
Role       Inheritance  Modes
```

```
netoperator              Exec
netadmin                 Exec Config Interface Line Router IP Routemap Protocol MAC
secadmin                 Exec Config
sysadmin                 Exec Config Interface Line Router IP Routemap Protocol MAC

testadmin  netadmin      Exec Config Interface Line Router IP Routemap Protocol MAC
```

Displaying Role Permissions Assigned to a Command

To display permissions assigned to a command, use the `show role` command in EXEC Privilege mode. The output displays the user role and or permission level.

Example of Role Permissions Assigned to a Command

```
DellEMC#show role mode ?
configure              Global configuration mode
exec                   Exec Mode
interface              Interface configuration mode
line                   Line Configuration mode
route-map              Route map configuration mode
router                 Router configuration mode

DellEMC#show role mode configure username
Role access: sysadmin

DellEMC##show role mode configure password-attributes
Role access: secadmin,sysadmin

DellEMC#show role mode configure interface
Role access: netadmin, sysadmin

DellEMC#show role mode configure line
Role access: netadmin,sysadmin
```

Displaying Information About Users Logged into the Switch

To display information on all users logged into the switch, using the `show users` command in EXEC Privilege mode. The output displays privilege level and/or user role. The mode is displayed at the start of the output and both the privilege and roles for all users is also displayed. If the role is not defined, the system displays "unassigned" .

Example of Displaying Information About Users Logged into the Switch

```
DellEMC#show users
Authorization Mode:  role or privilege

 Line         User    Role      Privilege Host(s) Location
 0 console 0  admin   sysadmin    15      idle
*3 vty 1      sec1    secadmin    14      idle    172.31.1.4
 4 vty 2      ml1     netadmin    12      idle    172.31.1.5
```

Two Factor Authentication (2FA)

Two factor authentication also known as 2FA, strengthens the login security by providing one time password (OTP) in addition to username and password. 2FA supports RADIUS authentications with Console, Telnet, and SSHv2.

To perform 2FA, follow these steps:

- When the Network access server (NAS) prompts for the username and password, provide the inputs.
- If the credentials are valid:
  - RADIUS server sends a request to the SMS–OTP daemon to generate an OTP for the user.
  - A challenge authentication is sent from the RADIUS server as Reply–Message attribute.
  - If the Reply–Message attribute is not sent from the RADIUS server, the default text is the `Response`.
  - 2FA is successful only on providing the correct OTP.

- If the credentials are invalid, the authentication fails.

NOTE: 2FA does not support RADIUS authentications done with REST, Web UI, and OMI.

Handling Access-Challenge Message

To provide a two-step verification in addition to the username and password, NAS prompts for additional information. An Access-Challenge request is sent from the RADIUS server to NAS.

The RADIUS server returns one of the following responses:

- **Access-Challenge**—If the user credentials are valid, the NAS server receives an Access-Challenge request from the RADIUS server.
- **Access-Accept**—NAS validates the username and password. If the credentials are valid, the RADIUS server sends an Access-Request to the short message service one time password (SMS-OTP) daemon to generate an OTP. The OTP is sent to the user's e-mail ID or mobile. If the OTP is valid, the RADIUS server authenticates the 2FA user and sends an Access-Accept response to NAS.
- **Access-Reject**—NAS validates the OTP and if the OTP is invalid, the RADIUS server does not authenticate the user and sends an Access-Reject response to NAS.

Configuring Challenge Response Authentication for SSHv2

To configure challenge response authentication for SSHv2, perform the following steps:

1. Enable challenge response authentication for SSHv2.

   CONFIGURATION mode

   ```
   ip ssh challenge-response-authentication enable
   ```

2. View the configuration.

   EXEC mode

   ```
   show ip ssh
   ```

```
DellEMC# show ip ssh
SSH server                : enabled.
SSH server version        : v2.
SSH server vrf            : default.
SSH server ciphers        : aes256-ctr,aes256-cbc,aes192-ctr,aes192-cbc,aes128-
ctr,aes128-cbc,3des-cbc.
SSH server macs           : hmac-sha2-256,hmac-sha1,hmac-sha1-96,hmac-md5,hmac-md5-96.
SSH server kex algorithms : diffie-hellman-group-exchange-sha1,diffie-hellman-group1-
sha1,diffie-hellman-group14-sha1.
Password Authentication   : enabled.
Hostbased Authentication  : disabled.
RSA       Authentication  : disabled.
Challenge Response Auth   : enabled.
   Vty          Encryption      HMAC            Remote IP
   2            aes128-cbc      hmac-md5        10.16.127.141
   4            aes128-cbc      hmac-md5        10.16.127.141
*  5            aes128-cbc      hmac-md5        10.16.127.141
DellEMC#
```

SMS-OTP Mechanism

A short message service one time password (SMS-OTP) is a free RADIUS module to implement two factor authentication. There are multiple 2FA mechanisms that can be deployed with the RADIUS. Mechanisms such as the Google authenticator do not rely on the Access-Challenge message and the SMS-OTP module rely on the Access-challenge message. The main objective of this feature is to handle the Access-Challenge messages and sends the Access-Request message with user's response.

This module requires NAS for handling the access challenge from the RADIUS server. NAS sends the input OTP in an Access-Request to the RADIUS server, and the user authentication succeeds or fails depending upon the Access-Accept or Access-Reject response received at NAS from the RADIUS server.

Configuring the System to Drop Certain ICMP Reply Messages

You can configure the Dell EMC Networking OS to drop ICMP reply messages. When you configure the `drop icmp` command, the system drops the ICMP reply messages from the front end and management interfaces. By default, the Dell EMC Networking OS responds to all the ICMP messages.

- Drop the ICMP or ICMPv6 message type.
  `drop {icmp | icmp6}`
  CONFIGURATION mode.

You can configure the Dell EMC Networking OS to suppress the following ICMPv4 and ICMP6 message types:

Table 102. Suppressed ICMP message types

ICMPv4 message types
Echo reply (0)
All sub types of destination unreachable (3)
Source quench (4)
Redirect (5)
Router advertisement (9)
Router solicitation (10)
Time exceeded (11)
IP header bad (12)
Timestamp request (13)
Timestamp reply (14)
Information request (15)
Information reply (16)
Address mask request (17)
Address mask reply (18)

NOTE: The Dell EMC Networking OS does not suppress the ICMP message type `echo request (8)`.

Table 103. Suppressed ICMPv6 message types

ICMPv6 message types
Destination unreachable (1)
Time exceeded (3)
IPv6 header bad (4)
Echo reply (129)
Who are you request (139)

Table 103. Suppressed ICMPv6 message types (continued)

ICMPv6 message types
Who are you reply (140)
Mtrace response (200)
Mtrace messages (201)

NOTE: The Dell EMC Networking OS does not suppress the following ICMPv6 message types:

- Packet too big (2)
- Echo request (128)
- Multicast listener query (130)
- Multicast listener report (131)
- Multicast listener done (132)
- Router solicitation (133)
- Router advertisement (134)
- Neighbor solicitation (135)
- Neighbor advertisement (136)
- Redirect (137)
- Router renumbering (138)
- MLD v2 listener report (143)
- Duplicate Address Request (157)
- Duplicate Address Confirmation (158)

48

Service Provider Bridging

Service provider bridging provides the ability to add a second VLAN ID tag in an Ethernet frame and is referred to as VLAN stacking in the Dell EMC Networking OS.

Topics:

- VLAN Stacking
- VLAN Stacking Packet Drop Precedence
- Dynamic Mode CoS for VLAN Stacking
- Layer 2 Protocol Tunneling
- Provider Backbone Bridging

VLAN Stacking

VLAN stacking, also called Q-in-Q, is defined in IEEE 802.1ad — Provider *Bridges*, which is an amendment to IEEE 802.1Q — Virtual *Bridged Local Area Networks*. It enables service providers to use 802.1Q architecture to offer separate VLANs to customers with no coordination between customers, and minimal coordination between customers and the provider.

Using only 802.1Q VLAN tagging all customers would have to use unique VLAN IDs to ensure that traffic is segregated, and customers and the service provider would have to coordinate to ensure that traffic mapped correctly across the provider network. Even under ideal conditions, customers and the provider would still share the 4094 available VLANs.

Instead, 802.1ad allows service providers to add their own VLAN tag to frames traversing the provider network. The provider can then differentiate customers even if they use the same VLAN ID, and providers can map multiple customers to a single VLAN to overcome the 4094 VLAN limitation. Forwarding decisions in the provider network are based on the provider VLAN tag only, so the provider can map traffic through the core independently; the customer and provider only coordinate at the provider edge.

At the access point of a VLAN-stacking network, service providers add a VLAN tag, the S-Tag, to each frame before the 802.1Q tag. From this point, the frame is double-tagged. The service provider uses the S-Tag, to forward the frame traffic across its network. At the egress edge, the provider removes the S-Tag, so that the customer receives the frame in its original condition, as shown in the following illustration.

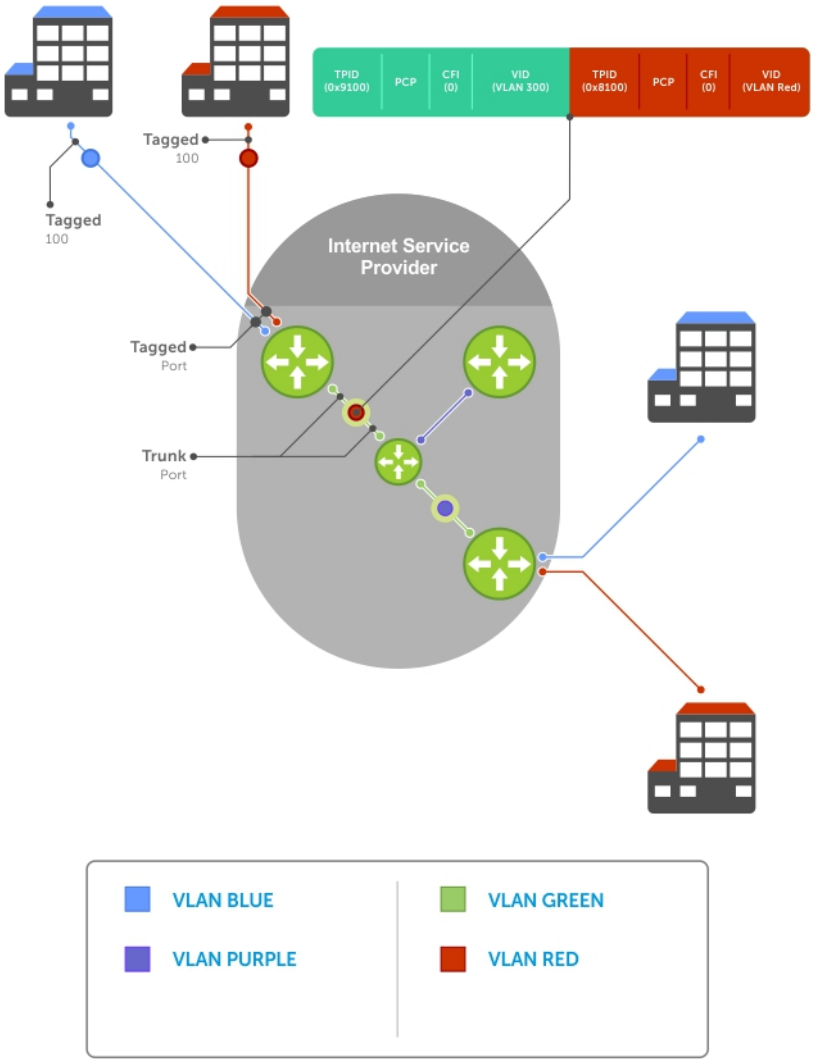


Figure 118. VLAN Stacking in a Service Provider Network

Important Points to Remember

- Interfaces that are members of the Default VLAN and are configured as VLAN-Stack access or trunk ports do not switch untagged traffic. To switch traffic, add these interfaces to a non-default VLAN-Stack-enabled VLAN.
- Dell EMC Networking cautions against using the same MAC address on different customer VLANs, on the same VLAN-Stack VLAN.
- You cannot ping across the trunk port link if one or both of the systems is an S4820T.
- This limitation becomes relevant if you enable the port as a multi-purpose port (carrying single-tagged and double-tagged traffic).

Configure VLAN Stacking

Configuring VLAN-Stacking is a three-step process.

1. Creating Access and Trunk Ports
2. Assign access and trunk ports to a VLAN (Creating Access and Trunk Ports).
3. Enabling VLAN-Stacking for a VLAN.

Related Configuration Tasks

- Configuring the Protocol Type Value for the Outer VLAN Tag
- Configuring Dell EMC Networking OS Options for Trunk Ports
- Debugging VLAN Stacking
- VLAN Stacking in Multi-Vendor Networks

Creating Access and Trunk Ports

To create access and trunk ports, use the following commands.

- **Access port** — a port on the service provider edge that directly connects to the customer. An access port may belong to only one service provider VLAN.
- **Trunk port** — a port on a service provider bridge that connects to another service provider bridge and is a member of multiple service provider VLANs.

Physical ports and port-channels can be access or trunk ports.

1. Assign the role of access port to a Layer 2 port on a provider bridge that is connected to a customer.

   INTERFACE mode

   `vlan-stack access`

2. Assign the role of trunk port to a Layer 2 port on a provider bridge that is connected to another provider bridge.

   INTERFACE mode

   `vlan-stack trunk`

3. Assign all access ports and trunk ports to service provider VLANs.

   INTERFACE VLAN mode

   `member`

To display the VLAN-Stacking configuration for a switchport, use the `show config` command from INTERFACE mode.

```
DellEMC#show run interface tengigabitEthernet 1/1
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  vlan-stack access
  no shutdown
DellEMC#show run interface tengigabitEthernet 1/2
!
interface TenGigabitEthernet 1/2
  no ip address
  switchport
  vlan-stack trunk
  no shutdown
```

Enable VLAN-Stacking for a VLAN

To enable VLAN-Stacking for a VLAN, use the following command.

- Enable VLAN-Stacking for the VLAN.

  INTERFACE VLAN mode

  `vlan-stack compatible`

To display the status and members of a VLAN, use the `show vlan` command from EXEC Privilege mode. Members of a VLAN-Stacking-enabled VLAN are marked with an M in column Q.

```
DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

NUM   Status       Q Ports
* 1   Active       U Te 3/0-5,18
```

```
   2  Inactive
   3  Inactive
   4  Inactive
   5  Inactive
   6  Active       M Po1(Te 3/14-15)
                   M Te 3/13
DellEMC#
```

Configuring the Protocol Type Value for the Outer VLAN Tag

The tag protocol identifier (TPID) field of the S-Tag is user-configurable.

To set the S-Tag TPID, use the following command.

- Select a value for the S-Tag TPID.

  CONFIGURATION mode

  `vlan-stack protocol-type`

  The default is **9100**.

To display the S-Tag TPID for a VLAN, use the `show running-config` command from EXEC privilege mode. Dell EMC Networking OS displays the S-Tag TPID only if it is a non-default value.

Configuring Dell EMC Networking OS Options for Trunk Ports

802.1ad trunk ports may also be tagged members of a VLAN so that it can carry single and double-tagged traffic.

You can enable trunk ports to carry untagged, single-tagged, and double-tagged VLAN traffic by making the trunk port a hybrid port.

To configure trunk ports, use the following commands.

1. Configure a trunk port to carry untagged, single-tagged, and double-tagged traffic by making it a hybrid port.

   INTERFACE mode

   `portmode hybrid`

   NOTE: You can add a trunk port to an 802.1Q VLAN as well as a Stacking VLAN only when the TPID 0x8100.

2. Add the port to a 802.1Q VLAN as tagged or untagged.

   INTERFACE VLAN mode

   `[tagged | untagged]`

In the following example, `TenGigabitEthernet 1/1` is a trunk port that is configured as a hybrid port and then added to VLAN 100 as untagged VLAN 101 as tagged, and VLAN 103, which is a stacking VLAN.

```
DellEMC(conf)#interface tenigabitethernet 1/1
DellEMC(conf-if-te-1/1)#portmode hybrid
DellEMC(conf-if-te-1/1)#switchport
DellEMC(conf-if-te-1/1)#vlan-stack trunk
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  no ip address
  portmode hybrid
  switchport
  vlan-stack trunk
  shutdown
DellEMC(conf-if-te-1/1)#interface vlan 100
DellEMC(conf-if-vl-100)#untagged tengigabitethernet 1/1
DellEMC(conf-if-vl-100)#interface vlan 101
DellEMC(conf-if-vl-101)#tagged tengigabitethernet 1/1
DellEMC(conf-if-vl-101)#interface vlan 103
DellEMC(conf-if-vl-103)#vlan-stack compatible
DellEMC(conf-if-vl-103-stack)#member tengigabitethernet 1/1
DellEMC(conf-if-vl-103-stack)#do show vlan

Codes: * - Default VLAN, G - GVRP VLANs
```

```
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

NUM    Status     Description  Q Ports
* 1    Inactive
  100  Inactive               U Te 1/1
  101  Inactive               T Te 1/1
  103  Inactive               M Te 1/1
```

Debugging VLAN Stacking

To debug VLAN stacking, use the following command.

- Debug the internal state and membership of a VLAN and its ports.

  `debug member`

The port notations are as follows:

- **MT** — stacked trunk
- **MU** — stacked access port
- **T** — 802.1Q trunk port
- **U** — 802.1Q access port
- **NU** — Native VLAN (untagged)

```
DellEMC# debug member vlan 603
vlan id  : 603
ports    : Te 2/4 (MT), Te 3/1(MU), Te 3/25(MT), Te 3/26(MT), Te 3/27(MU)

DellEMC#debug member port tengigabitethernet 2/4
vlan id  : 603 (MT), 100(T), 101(NU)
DellEMC#
```

VLAN Stacking in Multi-Vendor Networks

The first field in the VLAN tag is the tag protocol identifier (TPID), which is 2 bytes. In a VLAN-stacking network, after the frame is double tagged, the outer tag TPID must match the TPID of the next-hop system.

While 802.1Q requires that the inner tag TPID is 0x8100, it does not require a specific value for the outer tag TPID. Systems may use any 2-byte value; Dell EMC Networking OS uses 0x9100 (shown in the following) while non-Dell EMC Networking systems might use a different value.

If the next-hop system's TPID does not match the outer-tag TPID of the incoming frame, the system drops the frame. For example, as shown in the following, the frame originating from Building A is tagged VLAN RED, and then double-tagged VLAN PURPLE on egress at R4. The TPID on the outer tag is 0x9100. R2's TPID must also be 0x9100, and it is, so R2 forwards the frame.

Given the matching-TPID requirement, there are limitations when you employ Dell EMC Networking systems at network edges, at which, frames are either double tagged on ingress (R4) or the outer tag is removed on egress (R3).

VLAN Stacking

The default TPID for the outer VLAN tag is 0x9100. The system allows you to configure both bytes of the 2 byte TPID.

Previous versions allowed you to configure the first byte only, and thus, the systems did not differentiate between TPIDs with a common first byte. For example, 0x8100 and any other TPID beginning with 0x81 were treated as the same TPID, as shown in the following illustration. Dell EMC Networking OS Versions 8.2.1.0 and later differentiate between 0x9100 and 0x91XY, also shown in the following illustration.

You can configure the first 8 bits of the TPID using the `vlan-stack protocol-type` command.

The TPID is global. Ingress frames that do not match the system TPID are treated as untagged. This rule applies for both the outer tag TPID of a double-tagged frame and the TPID of a single-tagged frame.

For example, if you configure TPID 0x9100, the system treats 0x8100 and untagged traffic the same and maps both types to the default VLAN, as shown by the frame originating from Building C. For the same traffic types, if you configure TPID 0x8100, the

system is able to differentiate between 0x8100 and untagged traffic and maps each to the appropriate VLAN, as shown by the packet originating from Building A.

Therefore, a mismatched TPID results in the port not differentiating between tagged and untagged traffic.

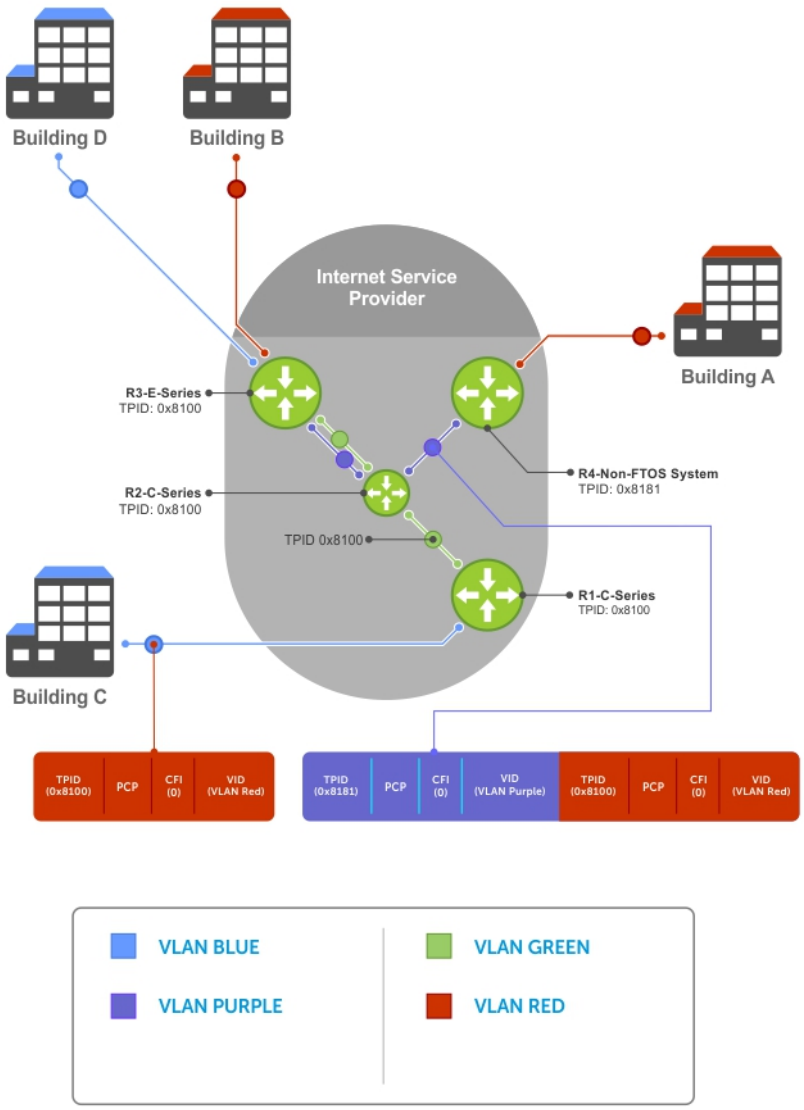


Figure 119. Single and Double-Tag TPID Match

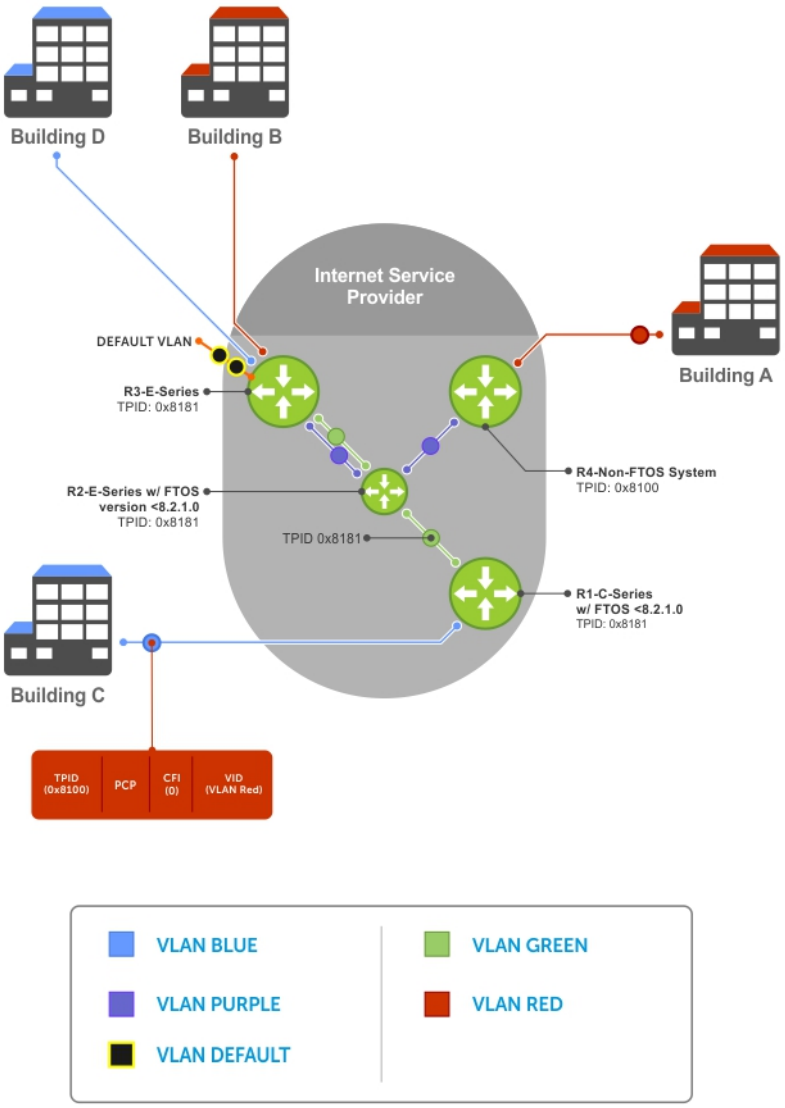


Figure 120. Single and Double-Tag First-byte TPID Match

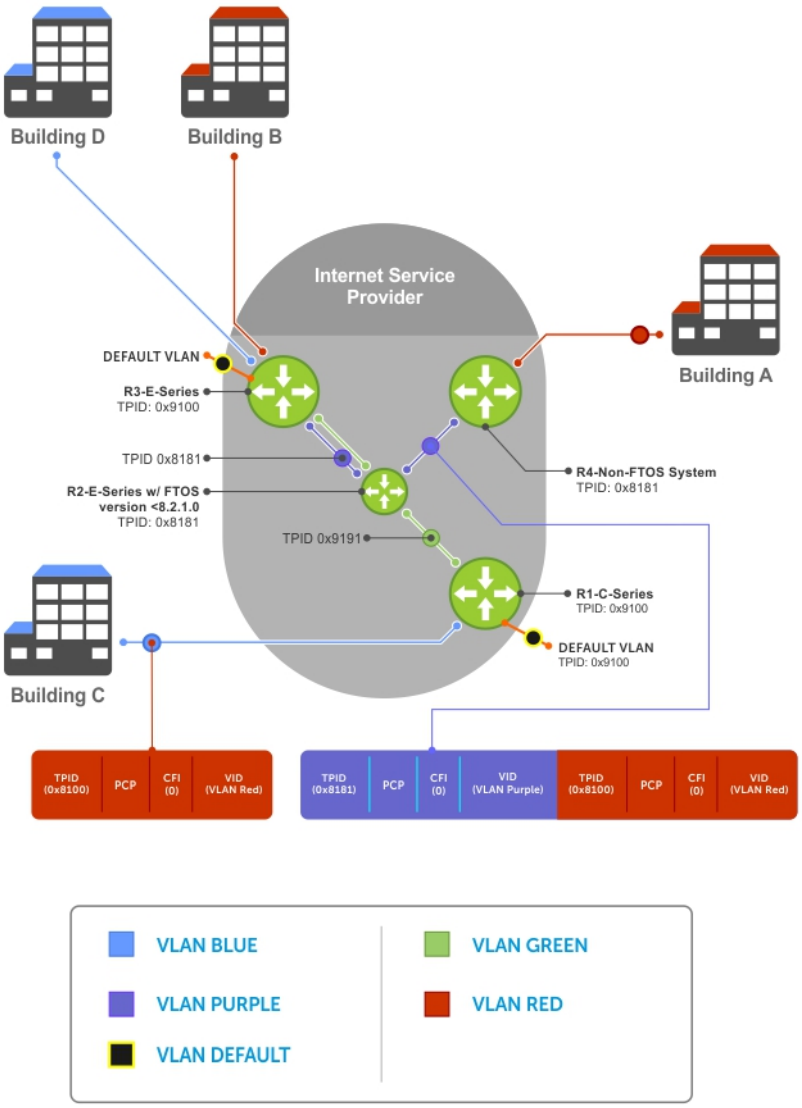


Figure 121. Single and Double-Tag TPID Mismatch

The following table details the outcome of matched and mismatched TPIDs in a VLAN-stacking network with the S-Series.

Table 104. Behaviors for Mismatched TPID

Network Position	Incoming Packet TPID	System TPID	Match Type	Pre-Version 8.2.1.0	Version 8.2.1.0+
Ingress Access Point	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	single-tag (0x8100)	0xUVWX	single-tag mismatch	switch to default VLAN	switch to default VLAN
		0x8100	single-tag match	switch to VLAN	switch to VLAN
		0x81XY	single-tag first-byte match	switch to VLAN	switch to default VLAN
Core	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	double-tag 0xUVWX	0xUVWX	double-tag match	switch to VLAN	switch to VLAN

Table 104. Behaviors for Mismatched TPID (continued)

Network Position	Incoming Packet TPID	System TPID	Match Type	Pre-Version 8.2.1.0	Version 8.2.1.0+
		0xUVYZ	double-tag first-byte match	switch to VLAN	switch to default VLAN
		0xQRST	double-tag mismatch	switch to default VLAN	switch to default VLAN
Egress Access Point	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	double-tag 0xUVWX	0xUVWX	double-tag match	switch to VLAN	switch to VLAN
		0xUVYZ	double-tag first-byte match	switch to VLAN	switch to default VLAN
		0xQRST	double-tag mismatch	switch to default VLAN	switch to default VLAN

VLAN Stacking Packet Drop Precedence

VLAN stacking packet-drop precedence is supported on the switch.

The drop eligible indicator (DEI) bit in the S-Tag indicates to a service provider bridge which packets it should prefer to drop when congested.

Enabling Drop Eligibility

Enable drop eligibility globally before you can honor or mark the DEI value.

When you enable drop eligibility, DEI mapping or marking takes place according to the defaults. In this case, the CFI is affected according to the following table.

Table 105. Drop Eligibility Behavior

Ingress	Egress	DEI Disabled	DEI Enabled
Normal Port	Normal Port	Retain CFI	Set CFI to 0.
Trunk Port	Trunk Port	Retain inner tag CFI	Retain inner tag CFI.
		Retain outer tag CFI	Set outer tag CFI to 0.
Access Port	Trunk Port	Retain inner tag CFI	Retain inner tag CFI
		Set outer tag CFI to 0	Set outer tag CFI to 0

To enable drop eligibility globally, use the following command.

- Make packets eligible for dropping based on their DEI value.
  CONFIGURATION mode
  ```
  dei enable
  ```
  By default, packets are colored green, and DEI is marked 0 on egress.

Honoring the Incoming DEI Value

To honor the incoming DEI value, you must explicitly map the DEI bit to an Dell EMC Networking OS drop precedence.

Precedence can have one of three colors.

Precedence	Description
Green	High-priority packets that are the least preferred to be dropped.
Yellow	Lower-priority packets that are treated as best-effort.
Red	Lowest-priority packets that are always dropped (regardless of congestion status).

- Honor the incoming DEI value by mapping it to an Dell EMC Networking OS drop precedence.
  INTERFACE mode
  `dei honor {0 | 1} {green | red | yellow}`
  You may enter the command once for 0 and once for 1.
  Packets with an unmapped DEI value are colored green.

To display the DEI-honoring configuration, use the `show interface dei-honor [interface slot/port]` in EXEC Privilege mode.

```
DellEMC#show interface dei-honor

Default Drop precedence: Green
Interface CFI/DEI  Drop precedence
---------------------------------------
Te 1/1    0        Green
Te 1/1    1        Yellow
Te 2/9    1        Red
Te 2/10   0        Yellow
```

Marking Egress Packets with a DEI Value

On egress, you can set the DEI value according to a different mapping than ingress.

For ingress information, refer to Honoring the Incoming DEI Value.

To mark egress packets, use the following command.

- Set the DEI value on egress according to the color currently assigned to the packet.
  INTERFACE mode
  `dei mark {green | yellow} {0 | 1}`

To display the DEI-marking configuration, use the `show interface dei-mark [interface slot/port ]` in EXEC Privilege mode.

```
DellEMC#show interface dei-mark

Default CFI/DEI Marking: 0
Interface Drop precedence CFI/DEI
-------------------------------
Te 1/1    Green           0
Te 1/1    Yellow          1
Te 2/9    Yellow          0
Te 2/10   Yellow          0
```

Dynamic Mode CoS for VLAN Stacking

One of the ways to ensure quality of service for customer VLAN-tagged frames is to use the 802.1p priority bits in the tag to indicate the level of QoS desired.

When an S-Tag is added to incoming customer frames, the 802.1p bits on the S-Tag may be configured statically for each customer or derived from the C-Tag using Dynamic Mode CoS. Dynamic Mode CoS maps the C-Tag 802.1p value to a S-Tag 802.1p value.

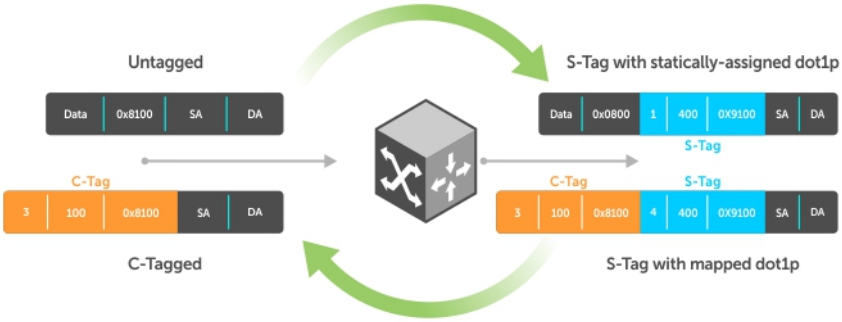


Figure 122. Statically and Dynamically Assigned dot1p for VLAN Stacking

When configuring Dynamic Mode CoS, you have two options:

- Mark the S-Tag dot1p and queue the frame according to the original C-Tag dot1p. In this case, you must have other dot1p QoS configurations; this option is classic dot1p marking.
- Mark the S-Tag dot1p and queue the frame according to the S-Tag dot1p. For example, if frames with C-Tag dot1p values 0, 6, and 7 are mapped to an S-Tag dot1p value 0, all such frames are sent to the queue associated with the S-Tag 802.1p value 0. This option requires two different CAM entries, each in a different Layer 2 ACL FP block.

NOTE: The ability to map incoming C-Tag dot1p to any S-Tag dot1p requires installing up to eight entries in the Layer 2 QoS and Layer 2 ACL table for each configured customer VLAN. The scalability of this feature is limited by the impact of the 1:8 expansion in these content addressable memory (CAM) tables.

Dell EMC Networking OS Behavior: For Option A shown in the previous illustration, when there is a conflict between the queue selected by Dynamic Mode CoS (vlan-stack dot1p-mapping) and a QoS configuration, the queue selected by Dynamic Mode CoS takes precedence. However, rate policing for the queue is determined by QoS configuration. For example, the following access-port configuration maps all traffic to Queue 0:

```
vlan-stack dot1p-mapping c-tag-dot1p 0-7 sp-tag-dot1p 1
```

However, if the following QoS configuration also exists on the interface, traffic is queued to Queue 0 but is policed at 40Mbps (`qos-policy-input` for queue 3) because class-map "a" of Queue 3 also matches the traffic. This is an expected behavior.

Examples of QoS Interface Configuration and Rate Policing

```
policy-map-input in layer2
service-queue 3 class-map a qos-policy 3
!
class-map match-any a layer2
match mac access-group a
!
mac access-list standard a
seq 5 permit any
!
qos-policy-input 3 layer2
rate-police 40
```

Likewise, in the following configuration, packets with dot1p priority 0–3 are marked as dot1p 7 in the outer tag and queued to Queue 3. Rate policing is according to `qos-policy-input` 3. All other packets will have outer dot1p 0 and hence are queued to Queue 1. They are therefore policed according to `qos-policy-input` 1.

```
policy-map-input in layer2
  service-queue 1 qos-policy 1
  service-queue 3 qos-policy 3
!
qos-policy-input 1 layer2
  rate-police 10
!
qos-policy-input 3 layer2
  rate-police 30
!
```

```
interface TenGigabitEthernet 1/21
  no ip address
  switchport
  vlan-stack access
  vlan-stack dot1p-mapping c-tag-dot1p 0-3 sp-tag-dot1p 7
  service-policy input in layer2
  no shutdown
```

Mapping C-Tag to S-Tag dot1p Values

To map C-Tag dot1p values to S-Tag dot1p values and mark the frames accordingly, use the following commands.

1. Allocate CAM space to enable queuing frames according to the C-Tag or the S-Tag.

   CONFIGURATION mode

   `cam-acl l2acl` *`number`* `ipv4acl` *`number`* `ipv6acl` *`number`* `ipv4qos` *`number`* `l2qos` *`number`* `l2pt` *`number`* `ipmacacl` *`number`* `ecfmacl` *`number`* `{vman-qos | vman-qos-dual-fp}` *`number`*

   - `vman-qos`: mark the S-Tag dot1p and queue the frame according to the original C-Tag dot1p. This method requires half as many CAM entries as `vman-qos-dual-fp`.
   - `vman-qos-dual-fp`: mark the S-Tag dot1p and queue the frame according to the S-Tag dot1p. This method requires twice as many CAM entries as `vman-qos` and FP blocks in multiples of 2.

   The default is: 0 FP blocks for `vman-qos` and `vman-qos-dual-fp`.

2. The new CAM configuration is stored in NVRAM and takes effect only after a save and reload.

   EXEC Privilege mode

   `copy running-config startup-config`

3. Reload the system.

   `reload`

4. Map C-Tag dot1p values to a S-Tag dot1p value.

   INTERFACE mode

   `vlan-stack dot1p-mapping c-tag-dot1p` *`values`* `sp-tag-dot1p` *`value`*

   Separate C-Tag values by commas. Dashed ranges are permitted.

   Dynamic Mode CoS overrides any Layer 2 QoS configuration in case of conflicts.

NOTE: Because `dot1p-mapping` marks *and* queues packets, the only remaining applicable QoS configuration is rate metering. You may use Rate Shaping or Rate Policing.

Layer 2 Protocol Tunneling

Spanning tree bridge protocol data units (BPDUs) use a reserved destination MAC address called the bridge group address, which is 01-80-C2-00-00-00.

Only spanning-tree bridges on the local area network (LAN) recognize this address and process the BPDU. When you use VLAN stacking to connect physically separate regions of a network, BPDUs attempting to traverse the intermediate network might be consumed and later dropped because the intermediate network itself might be using spanning tree (shown in the following illustration).

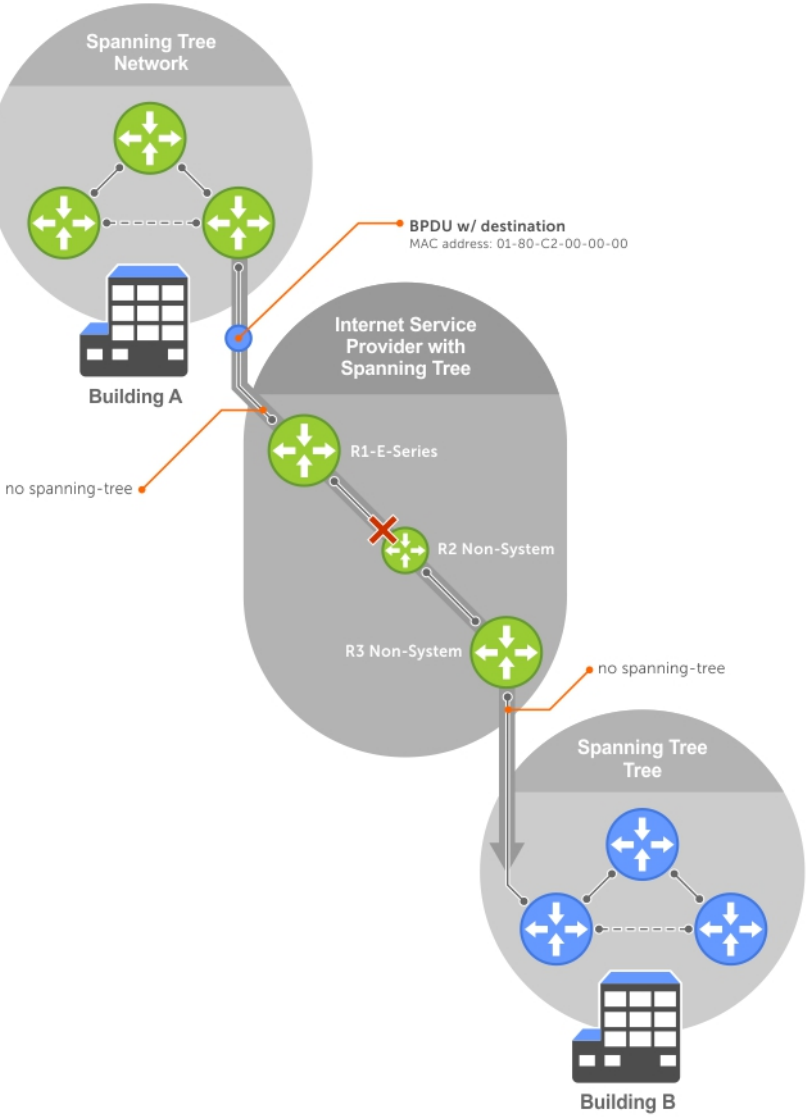


Figure 123. VLAN Stacking without L2PT

You might need to transport control traffic transparently through the intermediate network to the other region. Layer 2 protocol tunneling enables BPDUs to traverse the intermediate network by identifying frames with the Bridge Group Address, rewriting the destination MAC to a user-configured non-reserved address, and forwarding the frames. Because the frames now use a unique MAC address, BPDUs are treated as normal data frames by the switches in the intermediate network core. On egress edge of the intermediate network, the MAC address rewritten to the original MAC address and forwarded to the opposing network region (shown in the following illustration).

Dell EMC Networking OS Behavior: In Dell EMC Networking OS versions prior to 8.2.1.0, the MAC address that Dell EMC Networking systems use to overwrite the Bridge Group Address on ingress was non-configurable. The value of the L2PT MAC address was the Dell EMC Networking-unique MAC address, 01-01-e8-00-00-00. As such, with these Dell EMC Networking OS versions, Dell EMC Networking systems are required at the egress edge of the intermediate network because only Dell EMC Networking OS could recognize the significance of the destination MAC address and rewrite it to the original Bridge Group Address. In Dell EMC Networking OS version 8.2.1.0 and later, the L2PT MAC address is user-configurable, so you can specify an address that non-Dell EMC Networking systems can recognize and rewrite the address at egress edge.

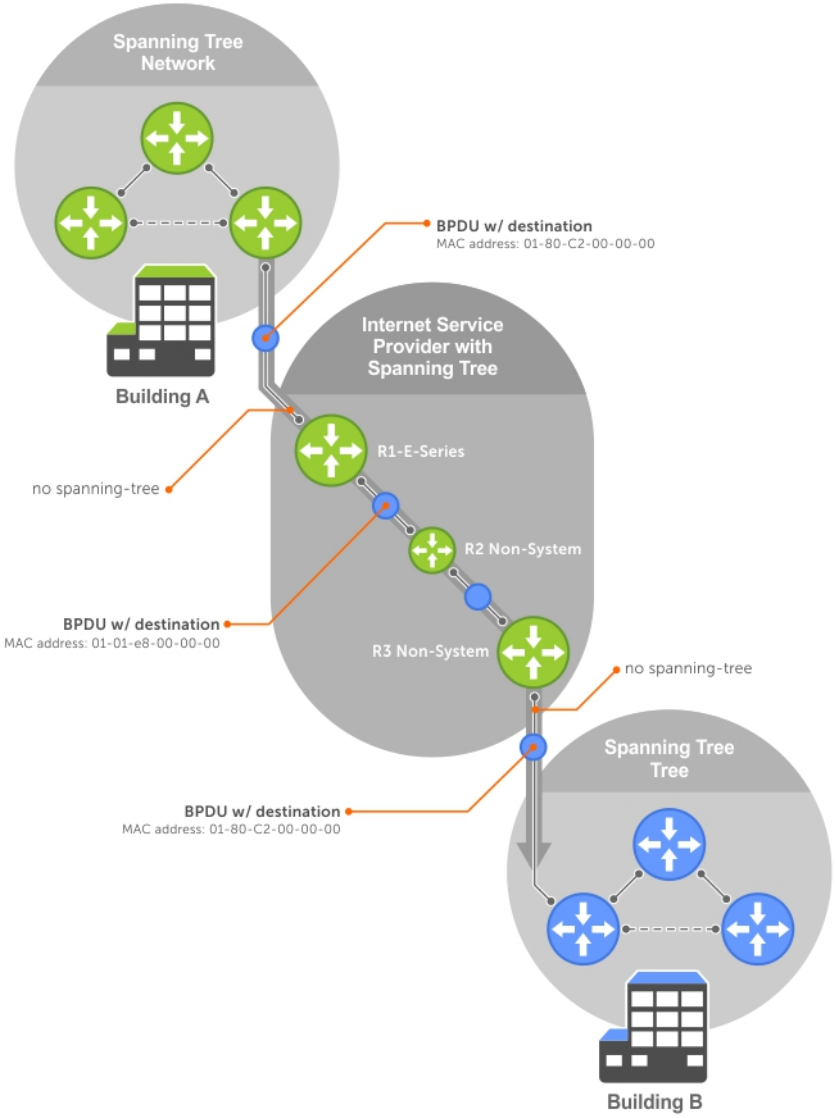


Figure 124. VLAN Stacking with L2PT

Implementation Information

- L2PT is available for STP, RSTP, MSTP, and PVST+ BPDUs.
- No protocol packets are tunneled when you enable VLAN stacking.
- L2PT requires the default CAM profile.

Enabling Layer 2 Protocol Tunneling

To enable Layer 2 protocol tunneling, use the following command.

1. Verify that the system is running the default CAM profile. Use this CAM profile for L2PT.
   EXEC Privilege mode
   ```
   show cam-profile
   ```
2. Enable protocol tunneling globally on the system.
   CONFIGURATION mode
   ```
   protocol-tunnel enable
   ```
3. Tunnel BPDUs the VLAN.

INTERFACE VLAN mode

```
protocol-tunnel stp
```

Specifying a Destination MAC Address for BPDUs

By default, Dell EMC Networking OS uses a Dell EMC Networking-unique MAC address for tunneling BPDUs. You can configure another value.

To specify a destination MAC address for BPDUs, use the following command.

- Overwrite the BPDU with a user-specified destination MAC address when BPDUs are tunneled across the provider network.

  CONFIGURATION mode

  ```
  protocol-tunnel destination-mac
  ```

  The default is 01:01:e8:00:00:00

Setting Rate-Limit BPDUs

CAM space is allocated in sections called field processor (FP) blocks.

There are a total of 13 user-configurable FP blocks. The default number of blocks for L2PT is **0**; you must allocate at least one to enable BPDU rate-limiting.

To set the rate-lime BPDUs, use the following commands.

1. Create at least one FP group for L2PT.

   CONFIGURATION mode

   ```
   cam-acl l2acl
   ```

   For details about this command, refer to CAM Allocation.
2. Save the running-config to the startup-config.

   EXEC Privilege mode

   ```
   copy running-config startup-config
   ```
3. Reload the system.

   EXEC Privilege mode

   ```
   reload
   ```
4. Set a maximum rate at which the RPM processes BPDUs for L2PT.

   VLAN STACKING mode

   ```
   protocol-tunnel rate-limit
   ```

   The default is: no rate limiting.

   The range is from 64 to 320 kbps.

Debugging Layer 2 Protocol Tunneling

To debug Layer 2 protocol tunneling, use the following command.

- Display debugging information for L2PT.

  EXEC Privilege mode

  ```
  debug protocol-tunnel
  ```

Provider Backbone Bridging

IEEE 802.1ad—Provider Bridges amends 802.1Q—Virtual Bridged Local Area Networks so that service providers can use 802.1Q architecture to offer separate VLANs to customers with no coordination between customers, and minimal coordination between customers and the provider.

802.1ad specifies that provider bridges operating spanning tree use a reserved destination MAC address called the Provider Bridge Group Address, 01-80-C2-00-00-08, to exchange BPDUs instead of the Bridge Group Address, 01-80-C2-00-00-00,

originally specified in 802.1Q. Only bridges in the service provider network use this destination MAC address so these bridges treat BPDUs originating from the customer network as normal data frames, rather than consuming them.

The same is true for GARP VLAN registration protocol (GVRP). 802.1ad specifies that provider bridges participating in GVRP use a reserved destination MAC address called the Provider Bridge GVRP Address, 01-80-C2-00-00-0D, to exchange GARP PDUs instead of the GVRP Address, 01-80-C2-00-00-21, specified in 802.1Q. Only bridges in the service provider network use this destination MAC address so these bridges treat GARP PDUs originating from the customer network as normal data frames, rather than consuming them.

Provider backbone bridging through IEEE 802.1ad eliminates the need for tunneling BPDUs with L2PT and increases the reliability of provider bridge networks as the network core need only learn the MAC addresses of core switches, as opposed to all MAC addresses received from attached customer devices.

- Use the Provider Bridge Group address as the destination MAC address in BPDUs. The xstp keyword applies this functionality to STP, RSTP, and MSTP; this functionality is not available for PVST+.

  CONFIGURATION Mode

  ```
  bpdu-destination-mac-address [xstp | gvrp] provider-bridge-group
  ```

49

sFlow

sFlow is a standard-based sampling technology embedded within switches and routers which is used to monitor network traffic. It is designed to provide traffic monitoring for high-speed networks with many switches and routers.

Topics:

- Overview
- Implementation Information
- Enabling Extended sFlow
- Enabling and Disabling sFlow on an Interface
- Enabling sFlow Max-Header Size Extended
- sFlow Show Commands
- Configuring Specify Collectors
- Changing the Polling Intervals
- Back-Off Mechanism
- sFlow on LAG ports
- Enabling Extended sFlow

Overview

The Dell EMC Networking Operating System (OS) supports sFlow version 5.

sFlow is a standard-based sampling technology embedded within switches and routers which is used to monitor network traffic. It is designed to provide traffic monitoring for high-speed networks with many switches and routers. sFlow uses two types of sampling:

- Statistical packet-based sampling of switched or routed packet flows.
- Time-based sampling of interface counters.

The sFlow monitoring system consists of an sFlow agent (embedded in the switch/router) and an sFlow collector. The sFlow agent resides anywhere within the path of the packet and combines the flow samples and interface counters into sFlow datagrams and forwards them to the sFlow collector at regular intervals. The datagrams consist of information on, but not limited to, packet header, ingress and egress interfaces, sampling parameters, and interface counters.

Application-specific integrated circuits (ASICs) typically complete packet sampling. sFlow collector analyses the sFlow datagrams received from different devices and produces a network-wide view of traffic flows.

Implementation Information

Dell EMC Networking sFlow is designed so that the hardware sampling rate is per line card port-pipe and is decided based on all the ports in that port-pipe.

If you do not enable sFlow on any port specifically, the global sampling rate is downloaded to that port and is to calculate the port-pipe's lowest sampling rate. This design supports the possibility that sFlow might be configured on that port in the future. Back-off is triggered based on the port-pipe's hardware sampling rate.

For example, if port 1 in the port-pipe has sFlow configured with a 16384 sampling rate while port 2 in the port-pipe has sFlow configured but no sampling rate set, the system applies a global sampling rate of 512 to port 2. The hardware sampling rate on the port-pipe is then set at 512 because that is the lowest configured rate on the port-pipe. When a high traffic situation occurs, a back-off is triggered and the hardware sampling rate is backed-off from 512 to 1024. Note that port 1 maintains its sampling rate of 16384; port 1 is unaffected because it maintains its configured sampling rate of 16384.:

- If the interface states are up and the sampling rate is not configured on the port, the default sampling rate is calculated based on the line speed.
- If the interface states are shut down, the sampling rate is set using the global sampling rate.

- If the global sampling rate is non-default, for example 256, and if the sampling rate is not configured on the interface, the sampling rate of the interface is the global non-default sampling rate, that is, 256.

To avoid the back-off, either increase the global sampling rate or configure all the line card ports with the desired sampling rate even if some ports have no sFlow configured.

Important Points to Remember

- The Dell EMC Networking OS implementation of the sFlow MIB supports sFlow configuration via snmpset.
- By default, sFlow collection is supported only on data ports. If you want to enable sFlow collection through management ports, use the `management egress-interface-selection` and `application sflow-collector` commands in Configuration and EIS modes respectively.
- Dell EMC Networking OS exports all sFlow packets to the collector. A small sampling rate can equate to many exported packets. A backoff mechanism is automatically applied to reduce this amount. Some sampled packets may be dropped when the exported packet rate is high and the backoff mechanism is about to or is starting to take effect. The dropEvent counter, in the sFlow packet, is always zero.
- Community list and local preference fields are not filled in extended gateway element in the sFlow datagram.
- 802.1P source priority field is not filled in extended switch element in sFlow datagram.
- Only Destination and Destination Peer AS number are packed in the *dst-as-path* field in extended gateway element.
- If the packet being sampled is redirected using policy-based routing (PBR), the sFlow datagram may contain incorrect extended gateway/router information.
- The source virtual local area network (VLAN) field in the extended switch element is not packed in case of routed packet.
- The destination VLAN field in the extended switch element is not packed in a Multicast packet.
- Up to 700 packets can be sampled and processed per second.

Enabling Extended sFlow

Extended sFlow packs additional information in the sFlow datagram depend on the type of sampled packet. The platform supports `extended-switch` information processing only.

Extended sFlow packs additional information in the sFlow datagram depending on the type of sampled packet. You can enable the following options:

- `extended-switch` — 802.1Q VLAN ID and 802.1p priority information.
- `extended-router` — Next-hop and source and destination mask length.
- `extended-gateway` — Source and destination AS number and the BGP next-hop.

NOTE: The entire AS path is not included. BGP community-list and local preference information are not included. These fields are assigned default values and are not interpreted by the collector.

- Enable extended sFlow.

  ```
  sflow [extended-switch] [extended-router] [extended-gateway] enable
  ```

  By default packing of any of the extended information in the datagram is disabled.
- Confirm that extended information packing is enabled.

  ```
  show sflow
  ```

The bold line shows that extended sFlow settings are enabled on all three types.

```
DellEMC#show sflow
sFlow services are enabled
Egress Management Interface sFlow services are disabled
Global default sampling rate: 32768
Global default counter polling interval: 20
Global default extended maximum header size: 128 bytes
Global extended information enabled: none
1 collectors configured
Collector IP addr: 100.1.1.1, Agent IP addr: 1.1.1.2, UDP port: 6343 VRF: Default
0 UDP packets exported
0 UDP packets dropped
0 sFlow samples collected

stack-unit 0 Port set 0
```

```
  Te 1/1: configured rate 16384, actual rate 16384
DellEMC#
```

If you did not enable any extended information, the show output displays the following (shown in bold).

```
DellEMC#show sflow
sFlow services are disabled
Global default sampling rate: 32768
Global default counter polling interval: 20
Global extended information enabled: none
0 collectors configured
0 UDP packets exported
0 UDP packets dropped
0 sFlow samples collected
0 sFlow samples dropped due to sub-sampling
```

Enabling and Disabling sFlow on an Interface

By default, sFlow is disabled on all interfaces.

This CLI is supported on physical ports and link aggregation group (LAG) ports.

To enable sFlow on a specific interface, use the following command.

- Enable sFlow on an interface.

  INTERFACE mode

  `[no] sflow ingress-enable`

To disable sFlow on an interface, use the `no` version of this command.

Enabling sFlow Max-Header Size Extended

To configure the maximum header size of a packet to 256 bytes, use the following commands:

- Set the maximum header size of a packet.

  CONFIGURATION mode

  INTERFACE mode

  `sflow max-header-size extended`

  By default, the maximum header size of a packet is 128 bytes. When `sflow max-header-size extended` is enabled, 256 bytes are copied. These bytes are useful for VxLAN, NvGRE, IPv4, and IPv6 tunneled packets.

  NOTE: Interface mode configuration takes priority.

- To reset the maximum header size of a packet, use the following command

  `[no] sflow max-header-size extended`

- View the maximum header size of a packet.

  `show running-config sflow`

Example of the `show sflow` command when the `sflow max-header-size extended` is configured globally

```
DellEMC(conf-if-te-1/10)#show sflow
sFlow services are enabled
Egress Management Interface sFlow services are disabled
Global default sampling rate: 32768
Global default counter polling interval: 86400
Global default extended maximum header size: 256 bytes
Global extended information enabled: none
1 collectors configured
Collector IP addr: 100.1.1.12, Agent IP addr: 100.1.1.1, UDP port: 6343 VRF: Default
0 UDP packets exported
0 UDP packets dropped
0 sFlow samples collected
```

Example of viewing the `sflow max-header-size extended` on an Interface Mode

```
DellEMC#show sflow interface tengigabitethernet 1/1
Te 1/1
sFlow type                        :Ingress
Configured sampling rate          :16384
Actual sampling rate              :16384
Counter polling interval          :20
Extended max header size     :256
Samples rcvd from h/w             :0
```

Example of the `show running-config sflow` Command

```
DellEMC#show running-config sflow
!
sflow collector 100.1.1.12 agent-addr 100.1.1.1
sflow enable
sflow max-header-size extended


DellEMC#show run int tengigabitEthernet 1/10
!
interface TenGigabitEthernet 1/10
no ip address
switchport
sflow ingress-enable
sflow max-header-size extended
no shutdown
```

sFlow Show Commands

Dell EMC Networking OS includes the following sFlow display commands.

- Displaying Show sFlow Globally
- Displaying Show sFlow on an Interface
- Displaying Show sFlow on a Line Card

Displaying Show sFlow Global

To view sFlow statistics, use the following command.

- Display sFlow configuration information and statistics.

  EXEC mode

  `show sflow`

The first bold line indicates sFlow is globally enabled.

The second bold lines indicate sFlow is enabled on Te 1/16 and Te 1/17

```
DellEMC#show sflow
sFlow services are enabled
Global default sampling rate: 32768
Global default counter polling interval: 20
1 collectors configured
Collector IP addr: 133.33.33.53, Agent IP addr: 133.33.33.116, UDP port: 6343
77 UDP packets exported
0 UDP packets dropped
165 sFlow samples collected
69 sFlow samples dropped due to sub-sampling

Stack-unit 1 Port set 0 H/W sampling rate 8192
   Te 1/16: configured rate 8192, actual rate 8192, sub-sampling rate 1
  Te 1/17: configured rate 16384, actual rate 16384, sub-sampling rate 2
```

Displaying Show sFlow on an Interface

To view sFlow information on a specific interface, use the following command.

- Display sFlow configuration information and statistics on a specific interface.

  EXEC mode

  `show sflow interface interface-name`

The following example shows the `show sflow interface` command.

```
DellEMC#show sflow interface tengigabitethernet 1/1
Te 1/1
sFlow type                      :Ingress
Configured sampling rate        :16384
Actual sampling rate            :16384
Counter polling interval        :20
Extended max header size        :128
Samples rcvd from h/w           :0
```

The following example shows the `show running-config interface` command.

```
DellEMC#show running-config interface tengigabitethernet 1/16
!
interface TenGigabitEthernet 1/16
  no ip address
  switchport
  sflow ingress-enable
  sflow sample-rate 8192
  no shutdown
```

Displaying Show sFlow on a Stack-unit

To view sFlow statistics on a specified Stack-unit, use the following command.

- Display sFlow configuration information and statistics on the specified interface.

  EXEC mode

  `show sflow stack—unit slot-number`

```
DellEMC#show sflow Stack-unit 1
 1
  Samples rcvd from h/w            :0
  Total UDP packets exported       :0
  UDP packets exported via RPM     :0
  UDP packets dropped              :36
```

Configuring Specify Collectors

The `sflow collector` command allows identification of sFlow collectors to which sFlow datagrams are forwarded.

You can specify up to two sFlow collectors. If you specify two collectors, the samples are sent to both.

- Identify sFlow collectors to which sFlow datagrams are forwarded.

  CONFIGURATION mode

  `sflow collector ip-address agent-addr ip-address [number [max-datagram-size number] ] | [max-datagram-size number ]`

  The default UDP port is **6343**.

  The default max-datagram-size is **1400**.

Changing the Polling Intervals

The `sflow polling-interval` command configures the polling interval for an interface in the maximum number of seconds between successive samples of counters sent to the collector.

This command changes the global default counter polling (20 seconds) interval. You can configure an interface to use a different polling interval.

To configure the polling intervals globally (in CONFIGURATION mode) or by interface (in INTERFACE mode), use the following command.

- Change the global default counter polling interval.

  CONFIGURATION mode or INTERFACE mode

  `sflow polling-interval` *`interval value`*

  - *`interval value`*: in seconds.

  The range is from 15 to 86400 seconds.

  The default is **20 seconds**.

Back-Off Mechanism

If the sampling rate for an interface is set to a very low value, the CPU can get overloaded with flow samples under high-traffic conditions.

In such a scenario, a binary back-off mechanism gets triggered, which doubles the sampling-rate (halves the number of samples per second) for all interfaces. The backoff mechanism continues to double the sampling-rate until the CPU condition is cleared. This is as per sFlow version 5 draft. After the back-off changes the sample-rate, you must manually change the sampling rate to the desired value.

As a result of back-off, the actual sampling-rate of an interface may differ from its configured sampling rate. You can view the actual sampling-rate of the interface and the configured sample-rate by using the `show sflow` command.

sFlow on LAG ports

When a physical port becomes a member of a LAG, it inherits the sFlow configuration from the LAG port.

Enabling Extended sFlow

Extended sFlow packs additional information in the sFlow datagram depend on the type of sampled packet. The platform supports `extended-switch` information processing only.

Extended sFlow packs additional information in the sFlow datagram depending on the type of sampled packet. You can enable the following options:

- `extended-switch` — 802.1Q VLAN ID and 802.1p priority information.
- `extended-router` — Next-hop and source and destination mask length.
- `extended-gateway` — Source and destination AS number and the BGP next-hop.

NOTE: The entire AS path is not included. BGP community-list and local preference information are not included. These fields are assigned default values and are not interpreted by the collector.

- Enable extended sFlow.

  `sflow [extended-switch] [extended-router] [extended-gateway] enable`

  By default packing of any of the extended information in the datagram is disabled.
- Confirm that extended information packing is enabled.

  `show sflow`

The bold line shows that extended sFlow settings are enabled on all three types.

```
DellEMC#show sflow
sFlow services are enabled
Egress Management Interface sFlow services are disabled
Global default sampling rate: 32768
Global default counter polling interval: 20
Global default extended maximum header size: 128 bytes
Global extended information enabled: none
1 collectors configured
Collector IP addr: 100.1.1.1, Agent IP addr: 1.1.1.2, UDP port: 6343 VRF: Default
0 UDP packets exported
0 UDP packets dropped
0 sFlow samples collected

stack-unit 0 Port set 0
  Te 1/1: configured rate 16384, actual rate 16384
DellEMC#
```

If you did not enable any extended information, the show output displays the following (shown in bold).

```
DellEMC#show sflow
sFlow services are disabled
Global default sampling rate: 32768
Global default counter polling interval: 20
Global extended information enabled: none
0 collectors configured
0 UDP packets exported
0 UDP packets dropped
0 sFlow samples collected
0 sFlow samples dropped due to sub-sampling
```

Important Points to Remember

- To export extended-gateway data, BGP must learn the IP destination address.
- If the IP destination address is not learned via BGP the Dell EMC Networking system does not export extended-gateway data.
- If the IP source address is learned via IGP, *srcAS* and *srcPeerAS* are zero.
- The *srcAS* and *srcPeerAS* might be zero even though the IP source address is learned via BGP. The c system packs the *srcAS* and *srcPeerAS* information only if the route is learned via BGP and it is reachable via the ingress interface of the packet.
- The sFlow sampling functionality is supported only for egress traffic and not for ingress traffic.

The previous points are summarized in following table.

Table 106. Extended Gateway Summary

IP SA	IP DA	srcAS and srcPeerAS	dstAS and dstPeerAS	Description
static/connected/IGP	static/connected/IGP	—	—	Extended gateway data is not exported because there is no AS information.
static/connected/IGP	BGP	0	Exported	src_as and src_peer_as are zero because there is no AS information for IGP.
BGP	static/connected/IGP	— Exported	— Exported	Prior to Dell EMC Networking OS version 7.8.1.0, extended gateway data is not exported because IP DA is not learned via BGP.

Table 106. Extended Gateway Summary (continued)

IP SA	IP DA	srcAS and srcPeerAS	dstAS and dstPeerAS	Description
				Version 7.8.1.0 allows extended gateway information in cases where the source and destination IP addresses are learned by different routing protocols, and for cases where is source is reachable over ECMP.
BGP	BGP	Exported	Exported	Extended gateway data is packed.

50

Simple Network Management Protocol (SNMP)

The Simple Network Management Protocol (SNMP) is designed to manage devices on IP networks by monitoring device operation, which might require administrator intervention.

NOTE: On Dell EMC Networking routers, standard and private SNMP management information bases (MIBs) are supported, including all *Get* and a limited number of *Set* operations (such as `set vlan` and `copy cmd`).

Topics:

- Protocol Overview
- Implementation Information
- SNMPv3 Compliance With FIPS
- Configuration Task List for SNMP
- Important Points to Remember
- Set up SNMP
- Reading Managed Object Values
- Writing Managed Object Values
- Configuring Contact and Location Information using SNMP
- Subscribing to Managed Object Value Updates using SNMP
- Enabling a Subset of SNMP Traps
- Enabling an SNMP Agent to Notify Syslog Server Failure
- Copy Configuration Files Using SNMP
- MIB Support for Power Monitoring
- MIB Support to Display the Available Memory Size on Flash
- MIB Support to Display the Software Core Files Generated by the System
- MIB Support to Display the Available Partitions on Flash
- MIB Support to Display Egress Queue Statistics
- MIB Support to ECMP Group Count
- MIB Support for entAliasMappingTable
- MIB Support for LAG
- MIB support for Port Security
- Manage VLANs using SNMP
- Managing Overload on Startup
- Enabling and Disabling a Port using SNMP
- Fetch Dynamic MAC Entries using SNMP
- Example of Deriving the Interface Index Number
- Monitoring BGP sessions via SNMP
- Monitor Port-Channels
- Troubleshooting SNMP Operation
- Transceiver Monitoring

Protocol Overview

Network management stations use SNMP to retrieve or alter management data from network elements.

A datum of management information is called a *managed object*; the value of a managed object can be static or variable. Network elements store managed objects in a database called a *management information base* (MIB).

MIBs are hierarchically structured and use object identifiers to address managed objects, but managed objects also have a textual name called an *object descriptor*.

You can download the latest MIB files from the following path:

- https://www.force10networks.com/CSPortal20/Main/SupportMain.aspx.

Implementation Information

The following describes SNMP implementation information.

- Dell EMC Networking OS supports SNMP version 1 as defined by RFC 1155, 1157, and 1212, SNMP version 2c as defined by RFC 1901, and SNMP version 3 as defined by RFC 2571.
- Dell EMC Networking OS supports up to 16 trap receivers.
- Dell EMC Networking OS implementation of the sFlow MIB supports sFlow configuration via SNMP sets.
- SNMP traps for the spanning tree protocol (STP) and multiple spanning tree protocol (MSTP) state changes are based on BRIDGE MIB (RFC 1483) for STP and IEEE 802.1 *draft ruzin-mstp-mib-02* for MSTP.

SNMPv3 Compliance With FIPS

SNMPv3 is compliant with the Federal information processing standard (FIPS) cryptography standard. The Advanced Encryption Standard (AES) Cipher Feedback (CFB) 128-bit encryption algorithm is in compliance with RFC 3826. SNMPv3 provides multiple authentication and privacy options for user configuration. A subset of these options are the FIPS-approved algorithms: HMAC-SHA1-96 for authentication and AES128-CFB for privacy. The other options are not FIPS-approved algorithms because of known security weaknesses. The AES128-CFB privacy option is supported and is compliant with RFC 3826.

The SNMPv3 feature also uses a FIPS-validated cryptographic module for all of its cryptographic operations when the system is configured with the `fips mode enable` command in Global Configuration mode. When the FIPS mode is enabled on the system, SNMPv3 operates in a FIPS-compliant manner, and only the FIPS-approved algorithm options are available for SNMPv3 user configuration. When the FIPS mode is disabled on the system, all options are available for SNMPv3 user configuration.

The following table describes the authentication and privacy options that can be configured when the FIPS mode is enabled or disabled:

Table 107. Authentication and Privacy Options

FIPS Mode	Privacy Options	Authentication Options
Disabled	des56 (DES56-CBC) aes128 (AES128-CFB)	md5 (HMAC-MD5-96) sha (HMAC-SHA1-96)
Enabled	aes128 (AES128-CFB)	sha (HMAC-SHA1-96)

To enable security for SNMP packets transferred between the server and the client, you can use the `snmp-server user` *`username`* `group` *`groupname`* `3 auth` *`authentication-type auth-password`* `priv aes128` *`priv-password`* command to specify that AES-CFB 128 encryption algorithm needs to be used.

```
DellEMC(conf)#snmp-server user snmpguy snmpmon 3 auth sha AArt61wq priv aes128 jntRR59a
```

In this example, for a specified user and a group, the AES128-CFB algorithm, the authentication password to enable the server to receive packets from the host, and the privacy password to encode the message contents are configured.

SHA authentication needs to be used with the AES-CFB128 privacy algorithm only when FIPS is enabled because SHA is then the only available authentication level. If FIPS is disabled, you can use MD5 authentication in addition to SHA authentication with the AES-CFB128 privacy algorithm

You cannot modify the FIPS mode if SNMPv3 users are already configured and present in the system. An error message is displayed if you attempt to change the FIPS mode by using the `fips mode enable` command in Global Configuration mode. You can enable or disable FIPS mode only if SNMPv3 users are not previously set up. If previously configured users exist on the system, you must delete the existing users before you change the FIPS mode.

Keep the following points in mind when you configure the AES128-CFB algorithm for SNMPv3:

1. SNMPv3 authentication provides only the `sha` option when the FIPS mode is enabled.
2. SNMPv3 privacy provides only the `aes128 privacy` option when the FIPS mode is enabled.

3. If you attempt to enable or disable FIPS mode and if any SNMPv3 users are previously configured, an error message is displayed stating you must delete all of the SNMP users before changing the FIPS mode.
4. A message is logged indicating whether FIPS mode is enabled for SNMPv3. This message is generated only when the first SNMPv3 user is configured because you can modify the FIPS mode only when users are not previously configured. This log message is provided to assist your system security auditing procedures.

Configuration Task List for SNMP

Configuring SNMP version 1 or version 2 requires a single step.

NOTE: The configurations in this chapter use a UNIX environment with net-snmp version 5.4. This environment is only one of many RFC-compliant SNMP utilities you can use to manage your Dell EMC Networking system using SNMP. Also, these configurations use SNMP version 2c.

- Creating a Community

Configuring SNMP version 3 requires configuring SNMP users in one of three methods. Refer to Setting Up User-Based Security (SNMPv3).

Related Configuration Tasks

- Managing Overload on Startup
- Reading Managed Object Values
- Writing Managed Object Values
- Subscribing to Managed Object Value Updates using SNMP
- Copying Configuration Files via SNMP
- Manage VLANs Using SNMP
- Enabling and Disabling a Port using SNMP
- Fetch Dynamic MAC Entries using SNMP
- Deriving Interface Indices
- Monitor Port-channels

Important Points to Remember

- Typically, 5-second timeout and 3-second retry values on an SNMP server are sufficient for both LAN and WAN applications. If you experience a timeout with these values, increase the timeout value to greater than 3 seconds, and increase the retry value to greater than 2 seconds on your SNMP server.
- User ACLs override group ACLs.

Set up SNMP

As previously stated, Dell EMC Networking OS supports SNMP version 1 and version 2 that are community-based security models.

The primary difference between the two versions is that version 2 supports two additional protocol operations (*informs operation* and *snmpgetbulk query*) and one additional object (*counter64 object*).

SNMP version 3 (SNMPv3) is a user-based security model that provides password authentication for user security and encryption for data security and privacy. Three sets of configurations are available for SNMP read/write operations: no password or privacy, password privileges, password and privacy privileges.

You can configure a maximum of 16 users even if they are in different groups.

Creating a Community

For SNMPv1 and SNMPv2, create a community to enable the community-based security in Dell EMC Networking OS.

The management station generates requests to either retrieve or alter the value of a management object and is called the *SNMP manager*. A network element that processes SNMP requests is called an *SNMP agent*. An *SNMP community* is a group of SNMP agents and managers that are allowed to interact. Communities are necessary to secure communication between SNMP managers and agents; SNMP agents do not respond to requests from management stations that are not part of the community.

Dell EMC Networking OS enables SNMP automatically when you create an SNMP community and displays the following message. You must specify whether members of the community may only retrieve values (read), or retrieve and alter values (read-write).

```
22:31:23: %STKUNIT0-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
```

To choose a name for the community you create, use the following command.

- Choose a name for the community.

  CONFIGURATION mode

  `snmp-server community name {ro | rw}`

To view your SNMP configuration, use the `show running-config snmp` command from EXEC Privilege mode.

```
DellEMC(conf)#snmp-server community my-snmp-community ro
22:31:23: %STKUNIT0-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
DellEMC#show running-config snmp
!
snmp-server community mycommunity ro
```

Setting Up User-Based Security (SNMPv3)

When setting up SNMPv3, you can set users up with one of the following three types of configuration for SNMP read/write operations.

Users are typically associated to an SNMP group with permissions provided, such as OID view.

- **noauth** — no password or privacy. Select this option to set up a user with no password or privacy privileges. This setting is the basic configuration. Users must have a group and profile that do not require password privileges.
- **auth** — password privileges. Select this option to set up a user with password authentication.
- **priv** — password and privacy privileges. Select this option to set up a user with password and privacy privileges.

To set up user-based security (SNMPv3), use the following commands.

- Configure the user with view privileges only (no password or privacy privileges).

  CONFIGURATION mode

  `snmp-server user name group-name 3 noauth`
- Configure an SNMP group with view privileges only (no password or privacy privileges).

  CONFIGURATION mode

  `snmp-server group group-name 3 noauth auth read name write name`
- Configure an SNMPv3 view.

  CONFIGURATION mode

  `snmp-server view view-name oid-tree {included | excluded}`

  NOTE: To give a user read and write view privileges, repeat this step for each privilege type.
- Configure the user with an authorization password (password privileges only).

  CONFIGURATION mode

  `snmp-server user name group-name 3 noauth auth md5 auth-password`
- Configure an SNMP group (password privileges only).

  CONFIGURATION mode

  `snmp-server group groupname {oid-tree} auth read name write name`

- Configure an SNMPv3 view.
  CONFIGURATION mode
  `snmp-server view view-name 3 noauth {included | excluded}`

  NOTE: To give a user read and write privileges, repeat this step for each privilege type.
- Configure an SNMP group (with password or privacy privileges).
  CONFIGURATION mode
  `snmp-server group group-name {oid-tree} priv read name write name`
- Configure the user with a secure authorization password and privacy password.
  CONFIGURATION mode
  `snmp-server user name group-name {oid-tree} auth md5 auth-password priv des56 priv password`
- Configure an SNMPv3 view.
  CONFIGURATION mode
  `snmp-server view view-name oid-tree {included | excluded}`

```
DellEMC(conf)#snmp-server host 1.1.1.1 traps {oid tree} version 3 ?
auth           Use the SNMPv3 authNoPriv Security Level
noauth         Use the SNMPv3 noAuthNoPriv Security Level
priv           Use the SNMPv3 authPriv Security Level
DellEMC(conf)#snmp-server host 1.1.1.1 traps {oid tree} version 3 noauth ?
WORD           SNMPv3 user name
```

Enable SNMPv3 traps

You must configure `notify` option for the SNMPv3 traps to work.

- Configure an SNMPv3 traps.
  CONFIGURATION mode
  `snmp-server group group-name {oid-tree} priv read name write name notify name`
  Enter the keyword `notify` then a name (a string of up to 20 characters long) as the notify view name.
- Configure an SNMPv3 view for notify.
  CONFIGURATION mode
  `snmp-server view view-name oid-tree {included | excluded}`

Reading Managed Object Values

You may only retrieve (read) managed object values if your management station is a member of the same community as the SNMP agent.

Dell EMC Networking supports RFC 4001, Textual Conventions for Internet Work Addresses that defines values representing a type of internet address. These values display for ipAddressTable objects using the `snmpwalk` command.

There are several UNIX SNMP commands that read data.

- Read the value of a single managed object.
  `snmpget -v version -c community agent-ip {identifier.instance | descriptor.instance}`
- Read the value of the managed object directly below the specified object.
  `snmpgetnext -v version -c community agent-ip {identifier.instance | descriptor.instance}`
- Read the value of many objects at once.
  `snmpwalk -v version -c community agent-ip {identifier.instance | descriptor.instance}`

In the following example, the value “4” displays in the OID before the IP address for IPv4. For an IPv6 IP address, a value of “16” displays.

```
> snmpget -v 2c -c mycommunity 10.11.131.161 sysUpTime.0
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32852616) 3 days, 19:15:26.16
> snmpget -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1.3.0
```

The following example shows reading the value of the next managed object.

```
> snmpgetnext -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1.3.0
SNMPv2-MIB::sysContact.0 = STRING:
> snmpgetnext -v 2c -c mycommunity 10.11.131.161 sysContact.0
```

The following example shows reading the value of the many managed objects at one time.

```
> snmpwalk -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1
SNMPv2-MIB::sysDescr.0 = STRING: Dell EMC Real Time Operating System Software
Dell Operating System Version: 1.0
Dell Application Software Version: E_MAIN4.9.4.0.0
Copyright (c) 1999-2014 by Dell
Build Time: Mon May 12 14:02:22 PDT 2008
SNMPv2-MIB::sysObjectID.0 = OID: SNMPv2-SMI::enterprises.6027.1.3.1
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32920954) 3 days, 19:26:49.54
SNMPv2-MIB::sysContact.0 = STRING:
```

Writing Managed Object Values

You may only alter (write) a managed object value if your management station is a member of the same community as the SNMP agent, and the object is writable.

Use the following command to write or write-over the value of a managed object.

- To write or write-over the value of a managed object.

  `snmpset -v` *`version`* `-c` *`community agent-ip {identifier.instance | descriptor.instance}syntax value`*

```
> snmpset -v 2c -c mycommunity 10.11.131.161 sysName.0 s "R5"
SNMPv2-MIB::sysName.0 = STRING: R5
```

Configuring Contact and Location Information using SNMP

You may configure system contact and location information from the Dell EMC Networking system or from the management station using SNMP.

To configure system contact and location information from the Dell EMC Networking system and from the management station using SNMP, use the following commands.

- (From a Dell EMC Networking system) Identify the system manager along with this person’s contact information (for example, an email address or phone number).

  CONFIGURATION mode

  `snmp-server contact` *`text`*

  You may use up to 55 characters.

  The default is **None**.
- (From a Dell EMC Networking system) Identify the physical location of the system (for example, San Jose, 350 Holger Way, 1st floor lab, rack A1-1).

  CONFIGURATION mode

  `snmp-server location` *`text`*

  You may use up to 55 characters.

The default is **None**.

- (From a management station) Identify the system manager along with this person's contact information (for example, an email address or phone number).

  CONFIGURATION mode

  `snmpset -v version -c community agent-ip sysContact.0 s "contact-info"`

  You may use up to 55 characters.

  The default is **None**.
- (From a management station) Identify the physical location of the system (for example, San Jose, 350 Holger Way, 1st floor lab, rack A1-1).

  CONFIGURATION mode

  `snmpset -v version -c community agent-ip sysLocation.0 s "location-info"`

  You may use up to 55 characters.

  The default is **None**.

Subscribing to Managed Object Value Updates using SNMP

By default, the Dell EMC Networking system displays some unsolicited SNMP messages (traps) upon certain events and conditions.

You can also configure the system to send the traps to a management station. Traps cannot be saved on the system.

Dell EMC Networking OS supports the following three sets of traps:

- **RFC 1157-defined traps** — coldStart, warmStart, linkDown, linkUp, authenticationFailure, and egpNeighbborLoss.
- **Dell EMC Networking enterpriseSpecific environment traps** — fan, supply, and temperature.
- **Dell EMC Networking enterpriseSpecific protocol traps** — bgp, ecfm, stp, and xstp.

To configure the system to send SNMP notifications, use the following commands.

1. Configure the Dell EMC Networking system to send notifications to an SNMP server.

   CONFIGURATION mode

   `snmp-server host ip-address [traps | informs] [version 1 | 2c |3] [community-string]`

   To send trap messages, enter the keyword `traps`.

   To send informational messages, enter the keyword `informs`.

   To send the SNMP version to use for notification messages, enter the keyword `version`.

   To identify the SNMPv1 community string, enter the name of the *community-string*.
2. Specify which traps the Dell EMC Networking system sends to the trap receiver.

   CONFIGURATION mode

   `snmp-server enable traps`

   Enable all Dell EMC Networking enterprise-specific and RFC-defined traps using the `snmp-server enable traps` command from CONFIGURATION mode.

   Enable all of the RFC-defined traps using the `snmp-server enable traps snmp` command from CONFIGURATION mode.
3. Specify the interfaces out of which Dell EMC Networking OS sends SNMP traps.

   CONFIGURATION mode

   `snmp-server trap-source`

The following example lists the RFC-defined SNMP traps and the command used to enable each. The *coldStart* and *warmStart* traps are enabled using a single command.

```
snmp authentication   SNMP_AUTH_FAIL:SNMP Authentication failed.Request with invalid
community string.
snmp coldstart        SNMP_COLD_START: Agent Initialized - SNMP COLD_START.
                      SNMP_WARM_START:Agent Initialized - SNMP WARM_START.
```

```
snmp linkdown          PORT_LINKDN:changed interface state to down:%d
snmp linkup            PORT_LINKUP:changed interface state to up:%d
```

Enabling a Subset of SNMP Traps

You can enable a subset of Dell EMC Networking enterprise-specific SNMP traps using one of the following listed command options.

To enable a subset of Dell EMC Networking enterprise-specific SNMP traps, use the following command.

- Enable a subset of SNMP traps.

  `snmp-server enable traps`

  NOTE: The `envmon` option enables all environment traps including those traps that are enabled with the `envmon supply`, `envmon temperature`, and `envmon fan` options.

  NOTE: You must configure `notify` option for the SNMPv3 traps to work.

envmon

```
CARD_SHUTDOWN: %sLine card %d down - %s
CARD_DOWN: %sLine card %d down - %s
LINECARDUP: %sLine card %d is up
CARD_MISMATCH: Mismatch: line card %d is type %s - type %s required.
RPM_STATE: RPM1 is in Active State
RPM_STATE: RPM0 is in Standby State
RPM_DOWN: RPM 0 down - hard reset
RPM_DOWN: RPM 0 down - card removed
HOT_FAILOVER: RPM Failover Completed
SFM_DISCOVERY: Found SFM 1
SFM_REMOVE: Removed SFM 1
MAJOR_SFM: Major alarm: Switch fabric down
MAJOR_SFM_CLR: Major alarm cleared: Switch fabric up
MINOR_SFM: MInor alarm: No working standby SFM
MINOR_SFM_CLR: Minor alarm cleared: Working standby SFM present
TASK SUSPENDED: SUSPENDED - svce:%d - inst:%d - task:%s
RPM0-P:CP %CHMGR-2-CARD_PARITY_ERR
ABNORMAL_TASK_TERMINATION: CRASH - task:%s %s
CPU_THRESHOLD: Cpu %s usage above threshold. Cpu5SecUsage (%d)
CPU_THRESHOLD_CLR: Cpu %s usage drops below threshold. Cpu5SecUsage (%d)
MEM_THRESHOLD: Memory %s usage above threshold. MemUsage (%d)
MEM_THRESHOLD_CLR: Memory %s usage drops below threshold. MemUsage (%d)
DETECT_STN_MOVE: Station Move threshold exceeded for Mac %s in vlan %d
CAM-UTILIZATION: Enable SNMP envmon CAM utilization traps.
```

envmon supply

```
PEM_PRBLM: Major alarm: problem with power entry module %s
PEM_OK: Major alarm cleared: power entry module %s is good
MAJOR_PS: Major alarm: insufficient power %s
MAJOR_PS_CLR: major alarm cleared: sufficient power
MINOR_PS: Minor alarm: power supply non-redundant
MINOR_PS_CLR: Minor alarm cleared: power supply redundant
```

envmon temperature

```
MINOR_TEMP: Minor alarm: chassis temperature
MINOR_TEMP_CLR: Minor alarm cleared: chassis temperature normal (%s %d
temperature is within threshold of %dC)
MAJOR_TEMP: Major alarm: chassis temperature high (%s temperature reaches or
exceeds threshold of %dC)
MAJOR_TEMP_CLR: Major alarm cleared: chassis temperature lower (%s %d
temperature is within threshold of %dC)
```

envmon fan

```
FAN_TRAY_BAD: Major alarm: fantray %d is missing or down
FAN_TRAY_OK: Major alarm cleared: fan tray %d present
FAN_BAD: Minor alarm: some fans in fan tray %d are down
FAN_OK: Minor alarm cleared: all fans in fan tray %d are good
```

vlt

```
  Enable VLT traps.

vrrp
  Enable VRRP state change traps

xstp
  %SPANMGR-5-STP_NEW_ROOT: New Spanning Tree Root, Bridge ID Priority 32768,
  Address 0001.e801.fc35.
  %SPANMGR-5-STP_TOPOLOGY_CHANGE: Bridge port TenGigabitEthernet 1/8 transitioned
  from Forwarding to Blocking state.
  %SPANMGR-5-MSTP_NEW_ROOT_BRIDGE: Elected root bridge for instance 0.
  %SPANMGR-5-MSTP_NEW_ROOT_PORT: MSTP root changed to port Te 1/8 for instance
  0. My Bridge ID: 40960:0001.e801.fc35 Old Root: 40960:0001.e801.fc35 New Root:
  32768:00d0.038a.2c01.
  %SPANMGR-5-MSTP_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e801.fc35 Mstp
  Instance Id 0 port Te 1/8 transitioned from forwarding to discarding state.

ecfm
  %ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain
  customer1 at Level 7 VLAN 1000
  %ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1
  at Level 7 VLAN 1000
  %ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain
  provider at Level 4 VLAN 3000
  %ECFM-5-ECFM_REMOTE_ALARM: Remote CCM Defect detected by MEP 3 in Domain
  customer1 at Level 7 VLAN 1000
  %ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at
  Level 7 VLAN 1000

entity
  Enable entity change traps
  Trap SNMPv2-MIB::sysUpTime.0 = Timeticks: (1487406) 4:07:54.06,
  SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::mib-2.47.2.0.1,
  SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 = INTEGER: 4
  Trap SNMPv2-MIB::sysUpTime.0 = Timeticks: (1488564) 4:08:05.64,
  SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::mib-2.47.2.0.1,
  SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 = INTEGER: 5
  Trap SNMPv2-MIB::sysUpTime.0 = Timeticks: (1489064) 4:08:10.64,
  SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::mib-2.47.2.0.1,
  SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 = INTEGER: 6
  Trap SNMPv2-MIB::sysUpTime.0 = Timeticks: (1489568)
  4:08:15.68,SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::mib-2.47.2.0.1,
  SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 = INTEGER: 7

<cr>
  SNMP Copy Config Command Completed
  %RPM0-P:CP %SNMP-4-RMON_RISING_THRESHOLD:  STACKUNIT0 rising threshold alarm from
  SNMP OID <oid>
  %RPM0-P:CP %SNMP-4-RMON_FALLING_THRESHOLD:  STACKUNIT0 falling threshold alarm from
  SNMP OID <oid>
  %RPM0-P:CP %SNMP-4-RMON_HC_RISING_THRESHOLD:  STACKUNIT0 high-capacity rising threshold
  alarm from SNMP OID <oid>
```

Enabling an SNMP Agent to Notify Syslog Server Failure

You can configure a network device to send an SNMP trap if an audit processing failure occurs due to loss of connectivity with the syslog server.

If a connectivity failure occurs on a syslog server that is configured for reliable transmission, an SNMP trap is sent and a message is displayed on the console.

The SNMP trap is sent only when a syslog connection fails and the time-interval between the last syslog notification and current time is greater than or equal to 5 minutes. This restriction also applies to the console message.

NOTE: If a syslog server failure event is generated before the SNMP agent service starts, the SNMP trap is not sent.

To enable an SNMP agent to send a trap when the syslog server is not reachable, enter the following command:

CONFIGURATION MODE

```
snmp-server enable traps snmp syslog-unreachable
```

To enable an SNMP agent to send a trap when the syslog server resumes connectivity, enter the following command:

CONFIGURATION MODE

```
snmp-server enable traps snmp syslog-reachable
```

Table 108. List of Syslog Server MIBS that have read access

MIB Object	OID	Object Values	Description
dF10SysLogTraps	1.3.6.1.4.1.6027.3.30.1.1	1 = reachable2 = unreachable	Specifies whether the syslog server is reachable or unreachable.

The following example shows the SNMP trap that is sent when connectivity to the syslog server is lost:

```
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (19738) 0:03:17.38     SNMPv2-
MIB::snmpTrapOID.0 = OID: SNMPv2-
SMI::enterprises.6027.3.30.1.1.1       SNMPv2-SMI::enterprises.6027.3.30.1.1 = STRING:
"NOT_REACHABLE: Syslog server
10.11.226.121 (port: 9140) is not reachable" SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 =
INTEGER: 2
```

Following is the sample audit log message that other syslog servers that are reachable receive:

```
Oct 21 00:46:13: dv-fedgov-s4810-6: %EVL-6-NOT_REACHABLE:Syslog server 10.11.226.121
(port: 9140) is not reachable
```

Following example shows the SNMP trap that is sent when connectivity to the syslog server is resumed:

```
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (10230) 0:01:42.30     SNMPv2-
MIB::snmpTrapOID.0 = OID: SNMPv2-
SMI::enterprises.6027.3.30.1.1.2       SNMPv2-SMI::enterprises.6027.3.30.1.1 = STRING:
"REACHABLE: Syslog server
10.11.226.121 (port: 9140) is reachable"SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 =
INTEGER: 2
```

Following is the sample audit log message that other syslog servers that are reachable receive:

```
Oct 21 05:26:04: dv-fedgov-s4810-6: %EVL-6-REACHABLE:Syslog server 10.11.226.121 (port:
9140) is reachable
```

Copy Configuration Files Using SNMP

To do the following, use SNMP from a remote client.

- copy the running-config file to the startup-config file
- copy configuration files from the Dell EMC Networking system to a server
- copy configuration files from a server to the Dell EMC Networking system

You can perform all of these tasks using IPv4 or IPv6 addresses. The examples in this section use IPv4 addresses; however, you can substitute IPv6 addresses for the IPv4 addresses in all of the examples.

The following table lists the relevant MIBs for these functions are.

Table 109. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Object Values	Description
copySrcFileType	.1.3.6.1.4.1.6027.3.5.1.1.1.1.2	1 = Dell EMC Networking OS file	Specifies the type of file to copy from. The range is:

Table 109. MIB Objects for Copying Configuration Files via SNMP (continued)

MIB Object	OID	Object Values	Description
		2 = running-config 3 = startup-config	• If copySrcFileType is running-config or startup-config, the default copySrcFileLocation is flash. • If copySrcFileType is a binary file, you must also specify copySrcFileLocation and copySrcFileName.
copySrcFileLocation	.1.3.6.1.4.1.6027.3.5.1.1.1.1.3	1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp 6 = usbflash	Specifies the location of source file. • If copySrcFileLocation is FTP or SCP, you must specify copyServerAddress, copyUserName, and copyUserPassword.
copySrcFileName	.1.3.6.1.4.1.6027.3.5.1.1.1.1.4	Path (if the file is not in the current directory) and filename.	Specifies name of the file. • If copySourceFileType is set to running-config or startup-config, copySrcFileName is not required.
copyDestFileType	.1.3.6.1.4.1.6027.3.5.1.1.1.1.5	1 = Dell EMC Networking OS file 2 = running-config 3 = startup-config	Specifies the type of file to copy to. • If copySourceFileType is running-config or startup-config, the default copyDestFileLocation is flash. • If copyDestFileType is a binary, you must specify copyDestFileLocation and copyDestFileName.
copyDestFileLocation	.1.3.6.1.4.1.6027.3.5.1.1.1.1.6	1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp	Specifies the location of destination file. • If copyDestFileLocation is FTP or SCP, you must specify copyServerAddress, copyUserName, and copyUserPassword.
copyDestFileName	.1.3.6.1.4.1.6027.3.5.1.1.1.1.7	Path (if the file is not in the default directory) and filename.	Specifies the name of destination file.
copyServerAddress	.1.3.6.1.4.1.6027.3.5.1.1.1.1.8	IP Address of the server.	The IP address of the server. • If you specify copyServerAddress, you must also specify copyUserName and copyUserPassword.
copyUserName	.1.3.6.1.4.1.6027.3.5.1.1.1.1.9	Username for the server.	Username for the FTP, TFTP, or SCP server.

Table 109. MIB Objects for Copying Configuration Files via SNMP (continued)

MIB Object	OID	Object Values	Description
			• If you specify copyUserName, you must also specify copyUserPassword.
copyUserPassword	.1.3.6.1.4.1.6027.3.5.1.1.1.1.10	Password for the server.	Password for the FTP, TFTP, or SCP server.

Copying a Configuration File

To copy a configuration file, use the following commands.

NOTE: In UNIX, enter the `snmpset` command for help using the following commands. Place the *f10-copy-config.mib* file in the directory from which you are executing the `snmpset` command or in the snmpset tool path.

1. Create an SNMP community string with read/write privileges.

   CONFIGURATION mode

   `snmp-server community community-name rw`
2. Copy the *f10-copy-config.mib* MIB from the Dell iSupport web page to the server to which you are copying the configuration file.
3. On the server, use the `snmpset` command as shown in the following example.

   `snmpset -v snmp-version -c community-name -m mib_path/f10-copy-config.mib force10system-ip-address mib-object.index {i | a | s} object-value...`

   - Every specified object must have an object value and must precede with the keyword i. Refer to the previous table.
   - *index* must be unique to all previously executed `snmpset` commands. If an index value has been used previously, a message like the following appears. In this case, increment the index value and enter the command again.

     ```
     Error in packet.
     Reason: notWritable (that object does not support modification)
     Failed object: FTOS-COPY-CONFIG-MIB::copySrcFileType.101
     ```
   - To complete the command, use as many MIB objects in the command as required by the MIB object descriptions shown in the previous table.

   NOTE: You can use the entire OID rather than the object name. Use the form: *OID.index i object-value*.

To view more information, use the following options in the `snmpset` command.

- `-c`: View the community, either public or private.
- `-m`: View the MIB files for the SNMP command.
- `-r`: Number of retries using the option
- `-t`: View the timeout.
- `-v`: View the SNMP version (either 1, 2, 2c, or 3).

The following examples show the `snmpset` command to copy a configuration. These examples assume that:

- the server OS is UNIX
- you are using SNMP version 2c
- the community name is public
- the file *f10-copy-config.mib* is in the current directory or in the snmpset tool path

Copying Configuration Files via SNMP

To copy the running-config to the startup-config from the UNIX machine, use the following command.

- Copy the running-config to the startup-config from the UNIX machine.

  `snmpset -v 2c -c public force10system-ip-address copySrcFileType.index i 2 copyDestFileType.index i 3`

The following examples show the command syntax using MIB object names and the same command using the object OIDs. In both cases, a unique index number follows the object.

The following example shows copying configuration files using MIB object names.

```
> snmpset -v 2c -r 0 -t 60 -c private -m ./f10-copy-config.mib 10.10.10.10
copySrcFileType.101
i 2 copyDestFileType.101 i 3
FTOS-COPY-CONFIG-MIB::copySrcFileType.101 = INTEGER: runningConfig(2)
FTOS-COPY-CONFIG-MIB::copyDestFileType.101 = INTEGER: startupConfig(3)
```

The following example shows copying configuration files using OIDs.

```
> snmpset -v 2c -c public -m ./f10-copy-config.mib 10.10.10.10
.1.3.6.1.4.1.6027.3.5.1.1.1.1.2.100 i 2 .1.3.6.1.4.1.6027.3.5.1.1.1.1.5.100 i 3
FTOS-COPY-CONFIG-MIB::copySrcFileType.100 = INTEGER: runningConfig(2)
FTOS-COPY-CONFIG-MIB::copyDestFileType.100 = INTEGER: startupConfig(3)
```

Copying the Startup-Config Files to the Running-Config

To copy the startup-config to the running-config from a UNIX machine, use the following command.

- Copy the startup-config to the running-config from a UNIX machine.

  `snmpset -c private -v 2c` *`force10system-ip-address`* `copySrcFileType.`*`index`* `i 3 copyDestFileType.`*`index`* `i 2`

The following example shows how to copy configuration files from a UNIX machine using the object name.

```
> snmpset -c public -v 2c -m ./f10-copy-config.mib 10.11.131.162 copySrcFileType.7 i 3
copyDestFileType.7 i 2
FTOS-COPY-CONFIG-MIB::copySrcFileType.7 = INTEGER: runningConfig(3)
FTOS-COPY-CONFIG-MIB::copyDestFileType.7 = INTEGER: startupConfig(2)
```

The following example shows how to copy configuration files from a UNIX machine using OID.

```
>snmpset -c public -v 2c 10.11.131.162 .1.3.6.1.4.1.6027.3.5.1.1.1.1.2.8 i 3
.1.3.6.1.4.1.6027.3.5.1.1.1.1.5.8 i 2
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.2.8 = INTEGER: 3
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.5.8 = INTEGER: 2
```

Copying the Startup-Config Files to the Server via FTP

To copy the startup-config to the server via FTP from the UNIX machine, use the following command.

Copy the startup-config to the server via FTP from the UNIX machine.

`snmpset -v 2c -c public -m ./f10-copy-config.mib` *`force10system-ip-address`* `copySrcFileType.`*`index`* `i 2 copyDestFileName.`*`index`* `s` *`filepath/filename`* `copyDestFileLocation.`*`index`* `i 4 copyServerAddress.`*`index`* `a` *`server-ip-address`* `copyUserName.`*`index`* `s` *`server-login-id`* `copyUserPassword.`*`index`* `s` *`server-login-password`*

- precede *`server-ip-address`* by the keyword `a`.
- precede the values for copyUsername and copyUserPassword by the keyword `s`.

```
> snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.110 i 2
copyDestFileName.110 s /home/startup-config copyDestFileLocation.110 i 4
copyServerAddress.110
a 11.11.11.11 copyUserName.110 s mylogin copyUserPassword.110 s mypass
FTOS-COPY-CONFIG-MIB::copySrcFileType.110 = INTEGER: runningConfig(2)
FTOS-COPY-CONFIG-MIB::copyDestFileName.110 = STRING: /home/startup-config
FTOS-COPY-CONFIG-MIB::copyDestFileLocation.110 = INTEGER: ftp(4)
FTOS-COPY-CONFIG-MIB::copyServerAddress.110 = IpAddress: 11.11.11.11
FTOS-COPY-CONFIG-MIB::copyUserName.110 = STRING: mylogin
FTOS-COPY-CONFIG-MIB::copyUserPassword.110 = STRING: mypass
```

Copying the Startup-Config Files to the Server via TFTP

To copy the startup-config to the server via TFTP from the UNIX machine, use the following command.

NOTE: Verify that the file exists and its permissions are set to 777. Specify the relative path to the TFTP root directory.

- Copy the startup-config to the server via TFTP from the UNIX machine.

  ```
  snmpset -v 2c -c public -m ./f10-copy-config.mib force10system-ip-address
  copySrcFileType.index i 3 copyDestFileType.index i 1 copyDestFileName.index s filepath/
  filename copyDestFileLocation.index i 3 copyServerAddress.index a server-ip-address
  ```

```
.snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10
copySrcFileType.4 i 3
copyDestFileType.4 i 1
copyDestFileLocation.4 i 3
copyDestFileName.4 s /home/myfilename
copyServerAddress.4 a 11.11.11.11
```

Copy a Binary File to the Startup-Configuration

To copy a binary file from the server to the startup-configuration on the Dell EMC Networking system via FTP, use the following command.

- Copy a binary file from the server to the startup-configuration on the Dell EMC Networking system via FTP.

  ```
  snmpset -v 2c -c public -m ./f10-copy-config.mib force10system-ip-address
  copySrcFileType.index i 1 copySrcFileLocation.index i 4 copySrcFileName.index s
  filepath/filename copyDestFileType.index i 3 copyServerAddress.index a server-ip-address
  copyUserName.index s server-login-id copyUserPassword.index s server-login-password
  ```

```
> snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.10 i 1
copySrcFileLocation.10 i 4 copyDestFileType.10 i 3 copySrcFileName.10 s /home/myfilename
copyServerAddress.10 a 172.16.1.56 copyUserName.10 s mylogin copyUserPassword.10 s mypass
```

Additional MIB Objects to View Copy Statistics

Dell EMC Networking provides more MIB objects to view copy statistics, as shown in the following table.

Table 110. Additional MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Values	Description
copyState	.1.3.6.1.4.1.6027.3.5.1.1.1.1.11	1= running 2 = successful 3 = failed	Specifies the state of the copy operation.
copyTimeStarted	.1.3.6.1.4.1.6027.3.5.1.1.1.1.12	Time value	Specifies the point in the up-time clock that the copy operation started.
copyTimeCompleted	.1.3.6.1.4.1.6027.3.5.1.1.1.1.13	Time value	Specifies the point in the up-time clock that the copy operation completed.
copyFailCause	.1.3.6.1.4.1.6027.3.5.1.1.1.1.14	1 = bad filename 2 = copy in progress 3 = disk full 4 = file exists 5 = file not found	Specifies the reason the copy request failed.

Table 110. Additional MIB Objects for Copying Configuration Files via SNMP (continued)

MIB Object	OID	Values	Description
		6 = timeout 7 = unknown	
copyEntryRowStatus	.1.3.6.1.4.1.6027.3.5.1.1.1.1.15	Row status	Specifies the state of the copy operation. Uses CreateAndGo when you are performing the copy. The state is set to *active* when the copy is completed.

Obtaining a Value for MIB Objects

To obtain a value for any of the MIB objects, use the following command.

- Get a copy-config MIB object value.

  `snmpset -v 2c -c public -m ./f10-copy-config.mib` *`force10system-ip-address`* `[`*`OID.index`* `|` *`mib-object.index`*`]`

  `index`: the index value used in the `snmpset` command used to complete the copy operation.

  NOTE: You can use the entire OID rather than the object name. Use the form: *`OID.index`*.

The following examples show the `snmpget` command to obtain a MIB object value. These examples assume that:

- the server OS is UNIX
- you are using SNMP version 2c
- the community name is public
- the file f10-copy-config.mib is in the current directory

NOTE: In UNIX, enter the `snmpset` command for help using this command.

The following examples show the command syntax using MIB object names and the same command using the object OIDs. In both cases, the same index number used in the `snmpset` command follows the object.

The following command shows how to get a MIB object value using the object name.

```
> snmpget -v 2c -c private -m ./f10-copy-config.mib 10.11.131.140 copyTimeCompleted.110
FTOS-COPY-CONFIG-MIB::copyTimeCompleted.110 = Timeticks: (1179831) 3:16:38.31
```

The following command shows how to get a MIB object value using OID.

```
> snmpget -v 2c -c private 10.11.131.140 .1.3.6.1.4.1.6027.3.5.1.1.1.1.13.110
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.13.110 = Timeticks: (1179831) 3:16:38.31
```

MIB Support for Power Monitoring

Dell EMC Networking provides MIB objects to display the information for Power Monitoring. The OIDs specific to Power Monitoring are appended to the DellITaMIbs. There are three separate DellITaMIbs available to display the current input power, average input power and average input-power start time. These statistics can also be obtained by using the CLI command:***show environment***. The following table lists the related MIB objects, OID and description for the same:

Table 111. MIB Objects to Display the Information for Power Monitoring

MIB Object	OID	Description
envMonSupplyCurrentPower	1.3.6.1.4.1.674.10895.3000.1.2.110.7.2.1.5	Displays per PSU input power (current configuration).
envMonSupplyAveragePower	1.3.6.1.4.1.674.10895.3000.1.2.110.7.2.1.6	Displays average input power.

Table 111. MIB Objects to Display the Information for Power Monitoring (continued)

MIB Object	OID	Description
envMonSupplyAvgStartTime	1.3.6.1.4.1.674.10895.3000.1.2.110.7.2.1.7	Displays average input-power start time.

SNMP Walk Example Output

```
snmpwalk -v 2c -c public 10.16.131.156 1.3.6.1.4.1.674.10895.3000.1.2.110.7.2.1.5
SNMPv2-SMI::enterprises.674.10895.3000.1.2.110.7.2.1.5.11 = INTEGER: 48
SNMPv2-SMI::enterprises.674.10895.3000.1.2.110.7.2.1.5.12 = INTEGER: 40
```

```
snmpwalk -v 2c -c public 10.16.131.156 1.3.6.1.4.1.674.10895.3000.1.2.110.7.2.1.6
SNMPv2-SMI::enterprises.674.10895.3000.1.2.110.7.2.1.6.11 = INTEGER: 31
SNMPv2-SMI::enterprises.674.10895.3000.1.2.110.7.2.1.6.12 = INTEGER: 26
```

```
snmpwalk -v 2c -c public 10.16.131.156 1.3.6.1.4.1.674.10895.3000.1.2.110.7.2.1.7
SNMPv2-SMI::enterprises.674.10895.3000.1.2.110.7.2.1.7.11 = STRING: "04/25/2017-18:32"
SNMPv2-SMI::enterprises.674.10895.3000.1.2.110.7.2.1.7.12 = STRING: "04/25/2017-18:32"
```

MIB Support to Display the Available Memory Size on Flash

Dell EMC Networking provides more MIB objects to display the available memory size on flash memory. The following table lists the MIB object that contains the available memory size on flash memory.

Table 112. MIB Objects for Displaying the Available Memory Size on Flash via SNMP

MIB Object	OID	Description
chStackUnitFlashUsageUtil	1.3.6.1.4.1.6027.3.10.1.2.9.1.6	Contains flash memory usage in percentage.

The chStackUnitUtilTable MIB table contains the chStackUnitFlashUsageUtil MIB object which contains the flash memory usage percent. The chStackUnitUtilTable is located in f10SSerChassisMib MIB.

Viewing the Available Flash Memory Size

- To view the available flash memory using SNMP, use the following command.
  `snmpget -v2c -c public 192.168.60.120 .1.3.6.1.4.1.6027.3.10.1.2.9.1.6.1`

  ```
  enterprises.6027.3.10.1.2.9.1.5.1 = Gauge32: 24
  ```

  The output above displays that 24% of the flash memory is used.

MIB Support to Display the Software Core Files Generated by the System

Dell EMC Networking provides MIB objects to display the software core files generated by the system. The chSysSwCoresTable contains the list of software core files generated by the system. The following table lists the related MIB objects.

Table 113. MIB Objects for Displaying the Software Core Files Generated by the System

MIB Object	OID	Description
chSysSwCoresTable	1.3.6.1.4.1.6027.3.10.1.2.10	This is the table that contains the list of software core files generated by the system.
chSysCoresEntry	1.3.6.1.4.1.6027.3.10.1.2.10.1	Entry number.
chSysCoresInstance	1.3.6.1.4.1.6027.3.10.1.2.10.1.1	Stores the indexed information about the available software core files.
chSysCoresFileName	1.3.6.1.4.1.6027.3.10.1.2.10.1.2	Contains the core file names and the file paths.
chSysCoresTimeCreated	1.3.6.1.4.1.6027.3.10.1.2.10.1.3	Contains the time at which core files are created.
chSysCoresStackUnitNumber	1.3.6.1.4.1.6027.3.10.1.2.10.1.4	Contains information that includes which stack unit or processor the core file was originated from.
chSysCoresProcess	1.3.6.1.4.1.6027.3.10.1.2.10.1.5	Contains information that includes the process names that generated each core file.

Viewing the Software Core Files Generated by the System

- To view the software core files generated by the system, use the following command.

  `snmpwalk -v2c -c public 192.168.60.120 .1.3.6.1.4.1.6027.3.10.1.2.10`

```
enterprises.6027.3.10.1.2.10.1.1.1.1 = 1
enterprises.6027.3.10.1.2.10.1.1.1.2 = 2
enterprises.6027.3.10.1.2.10.1.1.1.3 = 3
enterprises.6027.3.10.1.2.10.1.1.2.1 = 1
enterprises.6027.3.10.1.2.10.1.2.1.1 = "/CORE_DUMP_DIR/flashmntr.core.gz"
enterprises.6027.3.10.1.2.10.1.2.1.2 = "/CORE_DUMP_DIR/FTP_STK_MEMBER/
f10cp_l2mgr_131108080758_Stk1.acore.gz"
enterprises.6027.3.10.1.2.10.1.2.1.3 = "/CORE_DUMP_DIR/FTP_STK_MEMBER/
f10cp_vrrp_140522124357_Stk1.acore.gz"
enterprises.6027.3.10.1.2.10.1.2.2.1 =
"/CORE_DUMP_DIR/FTP_STK_MEMBER/f10cp_sysd_140617134445_Stk0.acore.gz"
enterprises.6027.3.10.1.2.10.1.3.1.1 = "Fri Mar 14 11:51:46 2014"
enterprises.6027.3.10.1.2.10.1.3.1.2 = "Fri Nov  8 08:11:16 2013"
enterprises.6027.3.10.1.2.10.1.3.1.3 = "Fri May 23 05:05:16 2014"
enterprises.6027.3.10.1.2.10.1.3.2.1 = "Tue Jun 17 14:19:26 2014"
enterprises.6027.3.10.1.2.10.1.4.1.1 = 0
enterprises.6027.3.10.1.2.10.1.4.1.2 = 1
enterprises.6027.3.10.1.2.10.1.4.1.3 = 1
enterprises.6027.3.10.1.2.10.1.4.2.1 = 0
enterprises.6027.3.10.1.2.10.1.5.1.1 = "flashmntr"
enterprises.6027.3.10.1.2.10.1.5.1.2 = "l2mgr"
enterprises.6027.3.10.1.2.10.1.5.1.3 = "vrrp" Hex: 76 72 72 70
enterprises.6027.3.10.1.2.10.1.5.2.1 = "sysd" Hex: 73 79 73 64
```

The output above displays that the software core files generated by the system.

MIB Support to Display the Available Partitions on Flash

Dell EMC Networking provides MIB objects to display the information of various partitions such as `/flash, /tmp, /usr/pkg, and /f10/ConfD`. The `dellNetFlashStorageTable` table contains the list of all partitions on disk. The following table lists the related MIB objects:

Table 114. MIB Objects to Display the Available Partitions on Flash

MIB Object	OID	Description
dellNetFlashPartitionNumber	1.3.6.1.4.1.6027.3.26.1.4.8.1.1	Index for the table.
dellNetFlashPartitionName	1.3.6.1.4.1.6027.3.26.1.4.8.1.2	Contains partition name and complete path.
dellNetFlashPartitionSize	1.3.6.1.4.1.6027.3.26.1.4.8.1.3	Contains the partition size.
dellNetFlashPartitionUsed	1.3.6.1.4.1.6027.3.26.1.4.8.1.4	Contains the amount of space used by the files on the partition.
dellNetFlashPartitionFree	1.3.6.1.4.1.6027.3.26.1.4.8.1.5	Contains the amount of free space available on the partition.
dellNetFlashPartitionMountPoint	1.3.6.1.4.1.6027.3.26.1.4.8.1.6	Symbolic or Alias name for the partition.

Viewing the Available Partitions on Flash

- To view the available partitions on flash using SNMP, use the following command:
  `snmpwalk -v 2c -c public -On 10.16.150.97 1.3.6.1.4.1.6027.3.26.1.4.8`

```
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.1 = STRING: "tmpfs"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.2 = STRING: "/dev/wd0i"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.3 = STRING: "mfs:477"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.4 = STRING: "/dev/wd0e"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.1 = INTEGER: 40960
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.2 = INTEGER: 4128782
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.3 = INTEGER: 148847
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.4 = INTEGER: 4186108
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.1 = INTEGER: 28
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.2 = INTEGER: 28
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.3 = INTEGER: 2537
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.4 = INTEGER: 76200
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.1 = INTEGER: 40932
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.2 = INTEGER: 3922316
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.3 = INTEGER: 138868
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.4 = INTEGER: 4109908
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.1 = STRING: "/tmp"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.2 = STRING: "/usr/pkg"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.3 = STRING: "/f10/ConfD/db"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.4 = STRING: "/f10/flash"
```

- If Smart Script is installed on the system, the log also shows the phone home partition.

```
snmpwalk -v 2c -c public -On 10.16.151.161 1.3.6.1.4.1.6027.3.26.1.4.8
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.1 = STRING: "/dev/ld0g"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.2 = STRING: "mfs:332"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.3 = STRING: "mfs:398"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.4 = STRING: "/dev/ld0h"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.2.5 = STRING: "tmpfs"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.1 = INTEGER: 4624894
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.2 = INTEGER: 59503
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.3 = INTEGER: 148847
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.4 = INTEGER: 2008708
.1.3.6.1.4.1.6027.3.26.1.4.8.1.3.5 = INTEGER: 51200
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.1 = INTEGER: 521636
```

```
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.2 = INTEGER: 1
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.3 = INTEGER: 2545
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.4 = INTEGER: 400528
.1.3.6.1.4.1.6027.3.26.1.4.8.1.4.5 = INTEGER: 60
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.1 = INTEGER: 3872014
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.2 = INTEGER: 56527
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.3 = INTEGER: 138860
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.4 = INTEGER: 1608180
.1.3.6.1.4.1.6027.3.26.1.4.8.1.5.5 = INTEGER: 51140
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.1 = STRING: "/usr/pkg"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.2 = STRING: "/tmpimg"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.3 = STRING: "/f10/ConfD/db"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.4 = STRING: "/f10/flash"
.1.3.6.1.4.1.6027.3.26.1.4.8.1.6.5 = STRING: "/f10/phonehome"
```

MIB Support to Display Egress Queue Statistics

Dell EMC Networking OS provides MIB objects to display the information of the packets transmitted or dropped per unicast or multicast egress queue. The following table lists the related MIB objects:

Table 115. MIB Objects to display egress queue statistics

MIB Object	OID	Description
dellNetFpEgrQTxPacketsRate	1.3.6.1.4.1.6027.3.27.1.20.1.6	Rate of Packets transmitted per Unicast/Multicast Egress queue.
dellNetFpEgrQTxBytesRate	1.3.6.1.4.1.6027.3.27.1.20.1.7	Rate of Bytes transmitted per Unicast/ Multicast Egress queue.
dellNetFpEgrQDroppedPacketsRate	1.3.6.1.4.1.6027.3.27.1.20.1.8	Rate of Packets dropped per Unicast/ Multicast Egress queue.
dellNetFpEgrQDroppedBytesRate	1.3.6.1.4.1.6027.3.27.1.20.1.9	Rate of Bytes dropped per Unicast/ Multicast Egress queue.

MIB Support to ECMP Group Count

Dell EMC Networking OS provides MIB objects to display the information of the ECMP group count information. The following table lists the related MIB objects:

Table 116. MIB Objects to display ECMP Group Count

MIB Object	OID	Description
dellNetInetCidrECMPGrpMax	1.3.6.1.4.1.6027.3.9.1.6	Total CAM for ECMP group.
dellNetInetCidrECMPGrpUsed	1.3.6.1.4.1.6027.3.9.1.7	Used CAM for ECMP group.
dellNetInetCidrECMPGrpAvl	1.3.6.1.4.1.6027.3.9.1.8	Available CAM for ECMP group.

Viewing the ECMP Group Count Information

- To view the ECMP group count information generated by the system, use the following command.

  `snmpwalk -c public -v 2c 10.16.151.191 1.3.6.1.4.1.6027.3.9`

```
SNMPv2-SMI::enterprises.6027.3.9.1.1.1.2.1.1 = Counter64: 79
SNMPv2-SMI::enterprises.6027.3.9.1.1.1.2.1.2 = Counter64: 1
SNMPv2-SMI::enterprises.6027.3.9.1.3.0 = Gauge32: 18
SNMPv2-SMI::enterprises.6027.3.9.1.4.0 = Gauge32: 1
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.10.1.1.0.24.0.0.0.0 = INTEGER: 2098693
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.10.1.1.1.32.1.4.10.1.1.1.1.4.10.1.1.1
= INTEGER: 2098693
```

```
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.10.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
INTEGER: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.20.1.1.0.24.0.0.0.0 = INTEGER:
1258296320
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.20.1.1.1.32.1.4.20.1.1.1.1.4.20.1.1.1
= INTEGER: 1258296320
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.20.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
INTEGER: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.30.1.1.0.24.0.0.0.0 = INTEGER:
1275078656
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.30.1.1.1.32.1.4.30.1.1.1.1.4.30.1.1.1
= INTEGER: 1275078656
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.30.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
INTEGER: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.70.70.70.0.24.0.0.0.0 = INTEGER:
2097157
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.70.70.70.1.32.1.4.127.0.0.1.1.4.127.0.0.1 =
INTEGER: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.70.70.70.2.32.1.4.70.70.70.2.1.4.70.70.70.2 =
INTEGER: 2097157
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.80.80.80.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
INTEGER: 2098693
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.80.80.80.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
INTEGER: 1258296320
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.80.80.80.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
INTEGER: 1275078656
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.90.90.90.0.24.0.0.0.0 = INTEGER:
2097157
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.90.90.90.1.32.1.4.127.0.0.1.1.4.127.0.0.1 =
INTEGER: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.90.90.90.2.32.1.4.90.90.90.2.1.4.90.90.90.2 =
INTEGER: 2097157
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.100.100.100.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
INTEGER: 2098693
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.100.100.100.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
INTEGER: 1258296320
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.8.1.1.4.100.100.100.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
INTEGER: 1275078656
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.10.1.1.0.24.0.0.0.0 = ""
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.10.1.1.1.32.1.4.10.1.1.1.1.4.10.1.1.1
= Hex-STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.10.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 = ""
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.20.1.1.0.24.0.0.0.0 = ""
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.20.1.1.1.32.1.4.20.1.1.1.1.4.20.1.1.1
= Hex-STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.20.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 = ""
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.30.1.1.0.24.0.0.0.0 = ""
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.30.1.1.1.32.1.4.30.1.1.1.1.4.30.1.1.1
= Hex-STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.30.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 = ""
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.70.70.70.0.24.0.0.0.0 = ""
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.70.70.70.1.32.1.4.127.0.0.1.1.4.127.0.0.1 = ""
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.70.70.70.2.32.1.4.70.70.70.2.1.4.70.70.70.2 =
Hex-STRING: 00 00 F4 FD 2C EF
SNMPv2-
```

```
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.80.80.80.0.24.1.4.10.1.1.1.1.4.10.1.1.1 = Hex-
STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.80.80.80.0.24.1.4.20.1.1.1.1.4.20.1.1.1 = Hex-
STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.80.80.80.0.24.1.4.30.1.1.1.1.4.30.1.1.1 = Hex-
STRING: 4C 76 25 F4 AB 02
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.90.90.90.0.24.0.0.0.0 = ""
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.90.90.90.1.32.1.4.127.0.0.1.1.4.127.0.0.1 = ""
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.90.90.90.2.32.1.4.90.90.90.2.1.4.90.90.90.2 =
Hex-STRING: 00 00 DA FE 04 0B
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.100.100.100.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
Hex-STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.100.100.100.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
Hex-STRING: 4C 76 25 F4 AB 02
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.9.1.1.4.100.100.100.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
Hex-STRING: 4C 76 25 F4 AB 02
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.10.1.1.0.24.0.0.0.0 = STRING: "CP"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.10.1.1.1.32.1.4.10.1.1.1.1.4.10.1.1.1
= STRING: "Fo 1/4/1"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.10.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
STRING: "CP"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.20.1.1.0.24.0.0.0.0 = STRING: "CP"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.20.1.1.1.32.1.4.20.1.1.1.1.4.20.1.1.1
= STRING: "Po 10"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.20.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
STRING: "CP"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.30.1.1.0.24.0.0.0.0 = STRING: "CP"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.30.1.1.1.32.1.4.30.1.1.1.1.4.30.1.1.1
= STRING: "Po 20"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.30.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
STRING: "CP"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.70.70.70.0.24.0.0.0.0 = STRING: "CP"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.70.70.70.1.32.1.4.127.0.0.1.1.4.127.0.0.1 =
STRING: "CP"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.70.70.70.2.32.1.4.70.70.70.2.1.4.70.70.70.2
= STRING: "Fo 1/1/1"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.80.80.80.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
STRING: "Fo 1/4/1"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.80.80.80.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
STRING: "Po 10"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.80.80.80.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
STRING: "Po 20"
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.90.90.90.0.24.0.0.0.0 = STRING: "CP"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.90.90.90.1.32.1.4.127.0.0.1.1.4.127.0.0.1 =
STRING: "CP"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.90.90.90.2.32.1.4.90.90.90.2.1.4.90.90.90.2
= STRING: "Fo 1/1/1"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.100.100.100.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
STRING: "Fo 1/4/1"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.100.100.100.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
STRING: "Po 10"
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.10.1.1.4.100.100.100.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
STRING: "Po 20"
```

```
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.10.1.1.0.24.0.0.0.0 = Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.10.1.1.1.32.1.4.10.1.1.1.1.4.10.1.1.1
= Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.10.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.20.1.1.0.24.0.0.0.0 = Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.20.1.1.1.32.1.4.20.1.1.1.1.4.20.1.1.1
= Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.20.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.30.1.1.0.24.0.0.0.0 = Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.30.1.1.1.32.1.4.30.1.1.1.1.4.30.1.1.1
= Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.30.1.1.2.32.1.4.127.0.0.1.1.4.127.0.0.1 =
Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.70.70.70.0.24.0.0.0.0 = Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.70.70.70.1.32.1.4.127.0.0.1.1.4.127.0.0.1 =
Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.70.70.70.2.32.1.4.70.70.70.2.1.4.70.70.70.2
= Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.80.80.80.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.80.80.80.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.80.80.80.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.90.90.90.0.24.0.0.0.0 = Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.90.90.90.1.32.1.4.127.0.0.1.1.4.127.0.0.1 =
Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.90.90.90.2.32.1.4.90.90.90.2.1.4.90.90.90.2
= Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.100.100.100.0.24.1.4.10.1.1.1.1.4.10.1.1.1 =
Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.100.100.100.0.24.1.4.20.1.1.1.1.4.20.1.1.1 =
Gauge32: 0
SNMPv2-
SMI::enterprises.6027.3.9.1.5.1.11.1.1.4.100.100.100.0.24.1.4.30.1.1.1.1.4.30.1.1.1 =
Gauge32: 0
SNMPv2-SMI::enterprises.6027.3.9.1.6.0 = Gauge32: 2048
SNMPv2-SMI::enterprises.6027.3.9.1.7.0 = Gauge32: 1
SNMPv2-SMI::enterprises.6027.3.9.1.8.0 = Gauge32: 2047
```

MIB Support for entAliasMappingTable

Dell EMC Networking provides a method to map the physical interface to its corresponding `ifindex` value. The `entAliasMappingTable` table contains zero or more rows, representing the logical entity mapping and physical component to external MIB identifiers. The following table lists the related MIB objects:

Table 117. MIB Objects for entAliasMappingTable

MIB Object	OID	Description
entAliasMappingTable	1.3.6.1.2.1.47.1.3.2	Contains information about `entAliasMapping` table.

Table 117. MIB Objects for entAliasMappingTable (continued)

MIB Object	OID	Description
entAliasMappingEntry	1.3.6.1.2.1.47.1.3.2.1	Contains information about a particular logical entity.
entAliasLogicalIndexOrZero	1.3.6.1.2.1.47.1.3.2.1.1	Contains a non–zero value and identifies the logical entity named by the same value of `entLogicalIndex`.
entAliasMappingIdentifier	1.3.6.1.2.1.47.1.3.2.1.2	Identifies a particular conceptual row associated with the indicated `entPhysicalIndex` and `entLogicalIndex` pair.

Viewing the entAliasMappingTable MIB

- To view the `entAliasMappingTable` generated by the system, use the following command.

  `snmpwalk -v 2c -c public -On 10.16.150.97 1.3.6.1.2.1.47.1.3.2.1`

```
.1.3.6.1.2.1.47.1.3.2.1.2.5.0 = OID: .1.3.6.1.2.1.2.2.1.1.2097157
.1.3.6.1.2.1.47.1.3.2.1.2.9.0 = OID: .1.3.6.1.2.1.2.2.1.1.2097669
.1.3.6.1.2.1.47.1.3.2.1.2.13.0 = OID: .1.3.6.1.2.1.2.2.1.1.2098181
.1.3.6.1.2.1.47.1.3.2.1.2.17.0 = OID: .1.3.6.1.2.1.2.2.1.1.2098693
.1.3.6.1.2.1.47.1.3.2.1.2.21.0 = OID: .1.3.6.1.2.1.2.2.1.1.2099205
.1.3.6.1.2.1.47.1.3.2.1.2.25.0 = OID: .1.3.6.1.2.1.2.2.1.1.2099717
.1.3.6.1.2.1.47.1.3.2.1.2.29.0 = OID: .1.3.6.1.2.1.2.2.1.1.2100228
.1.3.6.1.2.1.47.1.3.2.1.2.30.0 = OID: .1.3.6.1.2.1.2.2.1.1.2100356
.1.3.6.1.2.1.47.1.3.2.1.2.31.0 = OID: .1.3.6.1.2.1.2.2.1.1.2100484
```

MIB Support for LAG

Dell EMC Networking provides a method to retrieve the configured LACP information (Actor and Partner). Actor (local interface) is to designate the parameters and flags pertaining to the sending node, while the term Partner (remote interface) is to designate the sending node's view of its peer parameters and flags. LACP provides a standardized means for exchanging information, with partner systems, to form a link aggregation group (LAG). The following table lists the related MIB objects:

Table 118. MIB Objects for LAG

MIB Object	OID	Description
lagMIB	1.2.840.10006.300.43	Contains information about link aggregation module for managing 802.3ad.
lagMIBObjects	1.2.840.10006.300.43.1	
dot3adAgg	1.2.840.10006.300.43.1.1	
dot3adAggTable	1.2.840.10006.300.43.1.1.1	Contains information about every Aggregator that is associated with a system.
dot3adAggEntry	1.2.840.10006.300.43.1.1.1.1	Contains a list of Aggregator parameters and indexed by the `ifIndex` of the Aggregator.
dot3adAggMACAddress	1.2.840.10006.300.43.1.1.1.1.1	Contains a six octet read–only value carrying the individual MAC address assigned to the Aggregator.

Table 118. MIB Objects for LAG (continued)

MIB Object	OID	Description
dot3adAggActorSystemPriority	1.2.840.10006.300.43.1.1.1.1.2	Contains a two octet read–write value indicating the priority value associated with the Actor's system ID.
dot3adAggActorSystemID	1.2.840.10006.300.43.1.1.1.1.3	Contains a six octet read–write MAC address value used as a unique identifier for the system that contains the Aggregator.
dot3adAggAggregateOrIndividual	1.2.840.10006.300.43.1.1.1.1.4	Contains a read–only boolean value (True or False) indicating whether the Aggregator represents an Aggregate or an Individual link.
dot3adAggActorAdminKey	1.2.840.10006.300.43.1.1.1.1.5	Contains a 16–bit read–write value which is the current administrative key.
dot3adAggActorOperKey	1.2.840.10006.300.43.1.1.1.1.6	Contains a 16–bit read–write value which is the operational key.
dot3adAggPartnerSystemID	1.2.840.10006.300.43.1.1.1.1.7	Contains a six octet read–only MAC address value consisting of an unique identifier for the current Protocol partner of the Aggregator.
dot3adAggPartnerSystemPriority	1.2.840.10006.300.43.1.1.1.1.8	Contains a two octet read–only value that indicates the priority value associated with the Partner's system ID.
dot3adAggPartnerOperKey	1.2.840.10006.300.43.1.1.1.1.9	Contains the current operational value of the key for the Aggregator's current protocol partner.
dot3adAggCollectorMaxDelay	1.2.840.10006.300.43.1.1.1.1.10	Contains a 16–bit read–write attribute defining the maximum delay, in tens of microseconds, that may be imposed by the frame collector between receiving a frame from an Aggregator Parser, and either delivering the frame to its MAC Client or discarding the frame.
dot3adAggPortListTable	1.2.840.10006.300.43.1.1.2	Contains a list of all the ports associated with each Aggregator. Each LACP channel in a device occupies an entry in the table.
dot3adAggPortListEntry	1.2.840.10006.300.43.1.1.2.1	Contains a list of ports associated with a given Aggregator and indexed by the `ifIndex` of the Aggregator.
dot3adAggPortListPorts	1.2.840.10006.300.43.1.1.2.1.1	Contains a complete set of ports currently associated with the Aggregator.

Viewing the LAG MIB

- To view the LAG MIB generated by the system, use the following command.

  ```
  snmpbulkget -v 2c -c LagMIB 10.16.148.157 1.2.840.10006.300.43.1.1.1.1.1
  ```

  ```
  iso.2.840.10006.300.43.1.1.1.1.1.1258356224 = Hex-STRING: 00 01 E8 8A E8 46
  iso.2.840.10006.300.43.1.1.1.1.1.1258356736 = Hex-STRING: 00 01 E8 8A E8 46
  iso.2.840.10006.300.43.1.1.1.1.2.1258356224 = INTEGER: 32768
  iso.2.840.10006.300.43.1.1.1.1.2.1258356736 = INTEGER: 32768
  ```

```
iso.2.840.10006.300.43.1.1.1.1.3.1258356224 = Hex-STRING: 00 01 E8 8A E8 44
iso.2.840.10006.300.43.1.1.1.1.3.1258356736 = Hex-STRING: 00 01 E8 8A E8 44
iso.2.840.10006.300.43.1.1.1.1.4.1258356224 = INTEGER: 1
iso.2.840.10006.300.43.1.1.1.1.4.1258356736 = INTEGER: 1
iso.2.840.10006.300.43.1.1.1.1.5.1258356224 = INTEGER: 127
iso.2.840.10006.300.43.1.1.1.1.5.1258356736 = INTEGER: 128
```

MIB support for Port Security

Dell EMC Networking OS provides MIB objects to enable or disable port security feature on the physical and port channel interfaces.

The port security `DELL-NETWORKING-PORT-SECURITY-MIB` object contains both the global and interface level MIB objects.

Global MIB objects for port security

This section describes about the scalar MIB objects of the global MIB `dellNetPortSecGlobalObjects`.

The following table shows the scalar global MIB objects for port security.

Table 119. Global MIB Objects for Port Security

MIB Object	OID	Access or Permission	Description
dellNetGlobalPortSecurityMode	1.3.6.1.4.1.6027.3.31.1.1.1	read-write	Enables or disables port security feature globally on the device.
dellNetGlobalTotalSecureAddress	1.3.6.1.4.1.6027.3.31.1.1.2	read-only	Displays the total number of MAC addresses learnt or configured in the device.
dellNetGlobalClearSecureMacAddresses	1.3.6.1.4.1.6027.3.31.1.1.3	read-write	Deletes all the secured MAC addresses in the system based on the specific type (sticky or dynamic).
dellNetGlobalResetViolationStatus	1.3.6.1.4.1.6027.3.31.1.1.4	read-write	Resets the violation status (Err-disabled state) of all violated interfaces based on the specified type (MAC limit or station move violation).

MIB support for interface level port security

The MIB table `dellNetPortSecIfConfigTable` is used to achieve port security feature (MAC address learning limit) on an interface.

NOTE:

Port Security is not supported in VLT port channels.

The following table shows the MIB objects of the table `dellNetPortSecIfConfigTable`. The OID of the MIB table is `1.3.6.1.4.1.6027.3.31.1.2.1`.

Table 120. Interface level MIB Objects for Port Security

MIB Object	OID	Access or Permission	Description
dellNetPortSecIfPortSecurityEnable	1.3.6.1.4.1.6027.3.31.1.2.1.1.1	read-only	Specifies if the port security feature is enabled or disabled on an interface.
dellNetPortSecIfPortSecurityStatus	1.3.6.1.4.1.6027.3.31.1.2.1.1.2	read-only	Represents the port security status of an interface.

Table 120. Interface level MIB Objects for Port Security (continued)

MIB Object	OID	Access or Permission	Description
dellNetPortSecIfSecureMacLimit	1.3.6.1.4.1.6027.3.31.1.2.1.1.3	read-write	Maximum number (N) of MAC addresses to be secured on the interface
dellNetPortSecIfCurrentMacCount	1.3.6.1.4.1.6027.3.31.1.2.1.1.4	read-only	Current number of MAC addresses learnt or configured on this interface
dellNetPortSecIfStationMoveEnable	1.3.6.1.4.1.6027.3.31.1.2.1.1.5	read-write	Enable or disable station movement on the dynamically secured MAC addresses learnt on the interface.
dellNetPortSecIfSecureMacViolationAction	1.3.6.1.4.1.6027.3.31.1.2.1.1.6	read-write	Determines the action to be taken when MAC limit violation occurs in the system.
dellNetPortSecIfStmvViolationAction	1.3.6.1.4.1.6027.3.31.1.2.1.1.7	read-write	Determines the action to be taken when either dynamic or static MAC limit violation occurs in the system.
dellNetPortSecIfStickyEnable	1.3.6.1.4.1.6027.3.31.1.2.1.1.8	read-write	Enables or disables sticky port security feature on this interface.
dellNetPortSecIfClearSecureMacAddresses	1.3.6.1.4.1.6027.3.31.1.2.1.1.9	read-write	Deletes secure MAC addresses based on the specified type.
dellNetPortSecIfResetViolationStatus	1.3.6.1.4.1.6027.3.31.1.2.1.1.10	read-write	Resets the violation status of an interface based on the specified type.
dellNetPortSecIfSecureMacAgeEnable	1.3.6.1.4.1.6027.3.31.1.2.1.1.11	read-write	Enables aging of the dynamically secured MAC addresses learnt on the interface.

Enabling and viewing SNMP for port security

To enable or view `DELL-NETWORKING-PORT-SECURITY-MIB`, configure snmp-server in read-write mode using the `snmp-server community public rw` command. You can enable the port security feature on the Dell EMC Networking OS using the `snmpset` command. Also, you can view if the port security feature is enabled or disabled using the `snmpwalk` command.

To configure `dellNetPortSecIfSecureMacLimit` as 100 on an interface whose `ifIndex` is 2101252, use the following command.

```
snmpset -v 2c -c public 10.16.129.26 1.3.6.1.4.1.6027.3.31.1.2.1.1.3. 2101252 i 100
```

To remove `dellNetPortSecIfSecureMacLimit` configuration on an interface whose `ifIndex` is 2101252, use the following command.

```
snmpset -v 2c -c public 10.16.129.26 1.3.6.1.4.1.6027.3.31.1.2.1.1.3. 2101252 i 2147483647
```

To retrieve `dellNetPortSecIfSecureMacLimit` configured on an interface whose `ifIndex` is 2101252, use the following command.

```
snmpwalk -v 2c -c public 10.16.129.26 1.3.6.1.4.1.6027.3.31.1.2.1.1.3. 2101252
```

```
SNMPv2-SMI::enterprises.6027.3.31.1.2.1.1.3. 2101252 = INTEGER: 10
```

MIB objects for configuring MAC addresses

This section describes about the MIB objects `dellNetPortSecSecureStaticMacAddrTable` to configure and un-configure static MAC addresses in the system. The OID of this MIB table is `1.3.6.1.4.1.6027.3.31.1.2.2`.

The table is indexed by the following parameters:

- MAC Address (Octet string of length 6 and MAC address (in decimal) as value

- VLAN ID
- Interface Index

NOTE:

MAC addresses cannot be retrieved using `dellNetPortSecSecureStaticMacAddrTable` and `dellNetPortSecSecureMacAddrTable`. These tables are valid only if port security feature is enabled globally in the system.

Table 121. MIB Objects for configuring MAC addresses

MIB Object	OID	Access or Permission	Description
dellNetPortSecIfSecureStaticMacRowStatus	1.3.6.1.4.1.6027.3.31.1.2.2.1.4	read-write	Allows adding or deleting entries to or from the table `dellNetPortSecSecureStaticMacAddrTable`.

Enabling and viewing SNMP for static MAC addresses

You can enable and view SNMP for static MAC addresses using `snmpset` and `snmpget` command. Following example shows how to enable and view the static MAC addresses.

To configure a static MAC address (00:00:00:00:11:11) on a vlan (100) on interface whose ifIndex is (2101252), use the following command.

```
snmpset -v 2c -c public 10.16.129.26
1.3.6.1.4.1.6027.3.31.1.2.2.1.4.6.0.0.0.0.17.17.100.2101252 i 4
```

To remove the configure the above configured static MAC address, use the following command.

```
snmpset -v 2c -c public 10.16.129.26
1.3.6.1.4.1.6027.3.31.1.2.2.1.4.6.0.0.0.0.17.17.100.2101252 i 6
```

To retrieve the static MAC address configured, use the following command.

```
snmpget -v 2c -c public 10.16.129.26
1.3.6.1.4.1.6027.3.31.1.2.2.1.4.6.0.0.0.0.17.17.100.2101252
```

MIB objects for configuring MAC addresses

This section describes about the MIB table `dellNetPortSecSecureMacAddrTable` that contains the MAC database of the system.

The table is indexed by the following parameters:

- MAC Address (Octet string of length 6 and MAC address (in decimal) as value
- VLAN ID

Table 122. MIB Objects for configuring MAC addresses

MIB Object	OID	Access or Permission	Description
dellNetSecureMacIfIndex	1.3.6.1.4.1.6027.3.31.1.3.1.1.3	read-only	Shows in which interface the `dellNetSecureMacAddress` is configured or learnt.
dellNetSecureMacAddrType	1.3.6.1.4.1.6027.3.31.1.3.1.1.4	read-only	Indicates if the secure MAC address is configured as a static, dynamic, or sticky.

Viewing the Details of MAC addresses

You can retrieve the `dellNetSecureMacAddrType` details, use the `snmpwalk` command.

To retrieve the `dellNetSecureMacAddrType` on a MAC address (00:00:00:00:11:11) learnt or configured on a VLAN 10, use the following command.

```
snmpwalk -v 2c -c public 10.16.129.24 1.3.6.1.4.1.6027.3.31.1.3.1.1.4.6.0.0.0.0.17.17.10
SNMPv2-SMI::enterprises.6027.3.31.1.3.1.1.4.6.0.0.0.0.17.17.10 = INTEGER: 1
```

Manage VLANs using SNMP

The qBridgeMIB managed objects in Q-BRIDGE-MIB, defined in RFC 2674, allows you to use SNMP to manage VLANs.

Creating a VLAN

To create a VLAN, use the dot1qVlanStaticRowStatus object.

The snmpset operation shown in the following example creates VLAN 10 by specifying a value of 4 for instance 10 of the dot1qVlanStaticRowStatus object.

```
> snmpset -v2c -c mycommunity 123.45.6.78 .1.3.6.1.2.1.17.7.1.4.3.1.5.10 i 4
SNMPv2-SMI::mib-2.17.7.1.4.3.1.5.10 = INTEGER: 4
```

Assigning a VLAN Alias

Write a character string to the dot1qVlanStaticName object to assign a name to a VLAN.

```
[Unix system output]

> snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.1.1107787786 s "My
VLAN"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.1.1107787786 = STRING: "My VLAN"

[Dell system output]

DellEMC#show int vlan 10
Vlan 10 is down, line protocol is down
Vlan alias name is: My VLAN
Address is 00:01:e8:cc:cc:ce, Current address is 00:01:e8:cc:cc:ce
Interface index is 1107787786
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
```

Displaying the Ports in a VLAN

Dell EMC Networking OS identifies VLAN interfaces using an interface index number that is displayed in the output of the `show interface vlan` command.

The following example shows viewing the VLAN interface index number using SNMP.

```
DellEMC(conf)#do show interface vlan id 10
% Error: No such interface name.
R5(conf)#do show interface vlan 10
Vlan 10 is down, line protocol is down
Address is 00:01:e8:cc:cc:ce, Current address is 00:01:e8:cc:cc:ce
Interface index is 1107787786
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
```

```
LineSpeed auto
ARP type: ARPA, ARP Timeout 04:00:00
```

To display the ports in a VLAN, send an `snmpget` request for the object dot1qStaticEgressPorts using the interface index as the instance number, as shown for an S-Series.

The following example shows viewing VLAN ports using SNMP with no ports assigned.

```
> snmpget -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING:
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

The table that the Dell EMC Networking system sends in response to the `snmpget` request is a table that contains hexadecimal (hex) pairs, each pair representing a group of eight ports.

- Seven hex pairs represent a stack unit. Seven pairs accommodate the greatest number of ports available — 64 ports on the device. The last stack unit begins on the 66th bit.

The first hex pair, 00 in the previous example, represents ports 1 to 7 in Stack Unit 0. The next pair to the right represents ports 8 to 15. To resolve the hex pair into a representation of the individual ports, convert the hex pair to binary. Consider the first hex pair 00, which resolves to 0000 0000 in binary:

- Each position in the 8-character string is for one port, starting with Port 1 at the left end of the string, and ending with Port 8 at the right end. A 0 indicates that the port is not a member of the VLAN; a 1 indicates VLAN membership.

All hex pairs are 00, indicating that no ports are assigned to VLAN 10. In the following example, Port 0/2 is added to VLAN 10 as untagged; the first hex pair changes from 00 to 04.

The following example shows viewing VLAN ports using SNMP with ports assigned.

```
[Dell EMC Networking OS system output]

R5(conf)#do show vlan id 10

Codes: * - Default VLAN, G - GVRP VLANs
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

  NUM Status  Description  Q Ports
  10  Inactive             U Te 0/2

[Unix system output]

> snmpget -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00
00
```

The value 40 is in the first set of 7 hex pairs, indicating that these ports are in Stack Unit 0. The hex value 40 is 0100 0000 in binary. As described, the left-most position in the string represents Port 1. The next position from the left represents Port 2 and has a value of 1, indicating that Port 0/2 is in VLAN 10. The remaining positions are 0, so those ports are not in the VLAN.

NOTE: The table contains none of the other information the command provides, such as port speed or whether the ports are tagged or untagged.

Add Tagged and Untagged Ports to a VLAN

The value dot1qVlanStaticEgressPorts object is an array of all VLAN members.

The dot1qVlanStaticUntaggedPorts object is an array of only untagged VLAN members. All VLAN members that are not in dot1qVlanStaticUntaggedPorts are tagged.

- To add a tagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts object.
- To add an untagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts objects.

NOTE: Whether adding a tagged or untagged port, specify values for both dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts.

Example of Adding an Untagged Port to a VLAN using SNMP

In the following example, Port 0/2 is added as an untagged member of VLAN 10.

```
>snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786 x "40
00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00"
.1.3.6.1.2.1.17.7.1.4.3.1.4.1107787786 x "40 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00
00 00 00 00 00 00 00 00"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00
00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
SNMPv2-SMI::mib-2.17.7.1.4.3.1.4.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00
00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
```

Example of Adding a Tagged Port to a VLAN using SNMP

In the following example, Port 0/2 is added as a tagged member of VLAN 10.

```
>snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786 x "40
00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00"
.1.3.6.1.2.1.17.7.1.4.3.1.4.1107787786 x "00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00
00 00 00 00 00 00 00 00"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00
00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
SNMPv2-SMI::mib-2.17.7.1.4.3.1.4.1107787786 = Hex-STRING: 00 00 00 00 00 00 00 00 00 00
00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

Managing Overload on Startup

If you are running IS-IS, you can set a specific amount of time to prevent ingress traffic from being received after a reload and allow the routing protocol upgrade process to complete.

To prevent ingress traffic on a router while the IS reload is implemented, use the following command.

- Set the amount of time after an IS-IS reload is performed before ingress traffic is allowed at startup.

  `set-overload-bit on-startup isis`

The following OIDs are configurable through the `snmpset` command.

```
The node OID is 1.3.6.1.4.1.6027.3.18

F10-ISIS-MIB::f10IsisSysOloadSetOverload
F10-ISIS-MIB::f10IsisSysOloadSetOloadOnStartupUntil
F10-ISIS-MIB::f10IsisSysOloadWaitForBgp
```

```
F10-ISIS-MIB::f10IsisSysOloadV6SetOverload
F10-ISIS-MIB::f10IsisSysOloadV6SetOloadOnStartupUntil
F10-ISIS-MIB::f10IsisSysOloadV6WaitForBgp

    To enable overload bit for IPv4 set 1.3.6.1.4.1.6027.3.18.1.1 and IPv6 set
1.3.6.1.4.1.6027.3.18.1.4
    To set time to wait set 1.3.6.1.4.1.6027.3.18.1.2 and 1.3.6.1.4.1.6027.3.18.1.5
respectively
    To set time to wait till bgp session are up set 1.3.6.1.4.1.6027.3.18.1.3 and
1.3.6.1.4.1.6027.3.18.1.6
```

Enabling and Disabling a Port using SNMP

To enable and disable a port using SNMP, use the following commands.

1. Create an SNMP community on the Dell system.

   CONFIGURATION mode

   `snmp-server community`

2. From the Dell EMC Networking system, identify the interface index of the port for which you want to change the admin status.

   EXEC Privilege mode

   `show interface`

   Or, from the management system, use the `snmpwwalk` command to identify the interface index.

3. Enter the `snmpset` command to change the admin status using either the object descriptor or the OID.

   `snmpset with descriptor: snmpset -v` *`version`* `-c` *`community`* `agent-ip ifAdminStatus.`*`ifindex`* `i {1 | 2}`

   `snmpset with OID: snmpset -v` *`version`* `-c` *`community`* `agent-ip .1.3.6.1.2.1.2.2.1.7.`*`ifindex`* `i {1 | 2}`

   Choose integer `1` to change the admin status to Up, or `2` to change the admin status to Down.

Fetch Dynamic MAC Entries using SNMP

Dell EMC Networking supports the RFC 1493 dot1d table for the default VLAN and the dot1q table for all other VLANs.

NOTE: The 802.1q Q-BRIDGE MIB defines VLANs regarding 802.1d, as 802.1d itself does not define them. As a switchport must belong a VLAN (the default VLAN or a configured VLAN), all MAC address learned on a switchport are associated with a VLAN. For this reason, the Q-Bridge MIB is used for MAC address query. Moreover, specific to MAC address query, the MAC address indexes dot1dTpFdbTable only for a single forwarding database, while dot1qTpFdbTable has two indices — VLAN ID and MAC address — to allow for multiple forwarding databases and considering that the same MAC address is learned on multiple VLANs. The VLAN ID is added as the first index so that MAC addresses are read by the VLAN, sorted lexicographically. The MAC address is part of the OID instance, so in this case, lexicographic order is according to the most significant octet.

Table 123. MIB Objects for Fetching Dynamic MAC Entries in the Forwarding Database

MIB Object	OID	MIB	Description
dot1dTpFdbTable	.1.3.6.1.2.1.17.4.3	Q-BRIDGE MIB	List the learned unicast MAC addresses on the default VLAN.
dot1qTpFdbTable	.1.3.6.1.2.1.17.7.1.2. 2	Q-BRIDGE MIB	List the learned unicast MAC addresses on non-default VLANs.
dot3aCurAggFdb Table	.1.3.6.1.4.1.6027.3.2. 1.1.5	F10-LINK-AGGREGATION - MIB	List the learned MAC addresses of aggregated links (LAG).

In the following example, R1 has one dynamic MAC address, learned off of port TenGigabitEthernet 1/21, which a member of the default VLAN, VLAN 1. The SNMP walk returns the values for dot1dTpFdbAddress, dot1dTpFdbPort, and dot1dTpFdbStatus.

Each object comprises an OID concatenated with an instance number. In the case of these objects, the instance number is the decimal equivalent of the MAC address; derive the instance number by converting each hex pair to its decimal equivalent. For example, the decimal equivalent of E8 is 232, and so the instance number for MAC address 00:01:e8:06:95:ac is.0.1.232.6.149.172.

The value of dot1dTpFdbPort is the port number of the port off which the system learns the MAC address. In this case, of TenGigabitEthernet 1/21, the manager returns the integer 118.

Example of Fetching MAC Addresses Learned on the Default VLAN Using SNMP

```
----------------MAC Addresses on Force10 System------------------
DellEMC#show mac-address-table
VlanId   Mac Address      Type     Interface  State
1        00:01:e8:06:95:ac Dynamic  Te 1/21    Active
----------------Query from Management Station----------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.4.3.1
SNMPv2-SMI::mib-2.17.4.3.1.1.0.1.232.6.149.172 = Hex-STRING: 00 01 E8 06 95 AC
```

Example of Fetching MAC Addresses Learned on a Non-default VLAN Using SNMP

In the following example, TenGigabitEthernet 1/21 is moved to VLAN 1000, a non-default VLAN. To fetch the MAC addresses learned on non-default VLANs, use the object dot1qTpFdbTable. The instance number is the VLAN number concatenated with the decimal conversion of the MAC address.

```
---------------MAC Addresses on Force10 System------------
DellEMC#show mac-address-table
VlanId  Mac Address        Type     Interface  State
1000    00:01:e8:06:95:ac  Dynamic  Te 1/21    Active
---------------Query from Management Station----------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.7.1.2.2.1
```

Example of Fetching MAC Addresses Learned on a Port-Channel Using SNMP

Use dot3aCurAggFdbTable to fetch the learned MAC address of a port-channel. The instance number is the decimal conversion of the MAC address concatenated with the port-channel number.

```
--------------MAC Addresses on Force10 System-------------------
DellEMC(conf)#do show mac-address-table
VlanId  Mac Address        Type     Interface  State
1000    00:01:e8:06:95:ac  Dynamic  Po 1       Active
-------------Query from Management Station----------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.4.1.6027.3.2.1.1.5
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.1.1000.0.1.232.6.149.172.1 = INTEGER: 1000
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.2.1000.0.1.232.6.149.172.1 = Hex-STRING: 00 01
E8
06 95 AC
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.3.1000.0.1.232.6.149.172.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.4.1000.0.1.232.6.149.172.1 = INTEGER: 1
```

Example of Deriving the Interface Index Number

If you know the interface index, use the following commands to find the interface number.

```
DellEMC ~ $ snmpwalk -v 2c -c public 10.16.206.127 .1.3.6.1.2.1.2.2.1.2 | grep 2097156
IF-MIB::ifDescr.2097156 = STRING: TenGigabitEthernet 1/1

DellEMC ~ $ snmpwalk -v 2c -c public 10.16.206.127 .1.3.6.1.2.1.31.1.1.1.1 | grep 2097156
IF-MIB::ifName.2097156 = STRING: TenGigabitEthernet 1/1
```

You can use the `show interfaces` command to view the interface index.

```
DellEMC#show interface Tengigabitethernet 1/21
TenGigabitEthernet 1/21 is up, line protocol is up
Hardware is DellEMCEth, address is f8:b1:56:82:de:70
    Current address is f8:b1:56:82:de:70
```

```
Pluggable media not present
Interface index is 2097156
```

MIB Objects for Viewing the System Image on Flash Partitions

To view the system image on Flash Partition A, use the chSysSwInPartitionAImgVers object or, to view the system image on Flash Partition B, use the chSysSwInPartitionBImgVers object.

Table 124. MIB Objects for Viewing the System Image on Flash Partitions

MIB Object	OID	Description	MIB
chSysSwInPartitionAImgVers	1.3.6.1.4.1.6027.3.10.1.2.8.1.11	List the version string of the system image in Flash Partition A.	Chassis MIB
chSysSwInPartitionBImgVers	1.3.6.1.4.1.6027.3.10.1.2.8.1.12	List the version string of the system image in Flash Partition B.	Chassis MIB

Monitoring BGP sessions via SNMP

This section covers the monitoring of BGP sessions using SNMP.

BGP SNMP support for non-default VRF uses a SNMP context to distinguish multiple BGP VRF instances within a single BGP process. SNMP context is a repository of management information that can be accessed through the SNMP agent. SNMP supports multiple contexts in a device. SNMPv3 has a context name field in its PDU, which automatically allows the context name field to be mapped to a particular VRF instance without having to be mapped to a community map. SNMPv2c context has to be mapped to a community map. A new CLI command, `snmp context`, under BGP context, has been introduced to perform this function.

To map the context to a VRF instance for SNMPv2c, follow these steps:

1. Create a community and map a VRF to it. Create a context and map the context and community, to a community map.
   - sho run snmp
   - snmp-server community public ro
   - snmp-server community public ro
   - snmp-server community vrf1 ro
   - snmp-server community vrf2 ro
   - snmp-server context context1
   - snmp-server context context2
   - snmp mib community-map vrf1 context context1
   - snmp mib community-map vrf1 context context2
2. Configure `snmp context` under the VRF instances.
   - sho run bgp
   - router bgp 100
   - address-family ipv4 vrf vrf1
   - snmp context context1
   - neighbor 20.1.1.1 remote-as 200
   - neighbor 20.1.1.1 no shutdown
   - exit-address-family
   - address-family ipv4 vrf vrf2
   - snmp context context2
   - timers bgp 30 90
   - neighbor 30.1.1.1 remote-as 200
   - neighbor 30.1.1.1 no shutdown
   - exit-address-family

To map the context to a VRF instance for SNMPv3, follow these steps:

1. Create a community and map a VRF to it. Create a context and map the context and community, to a community map.
   - snmp-server community public ro

- snmp-server community VRF1 ro
- snmp-server community VRF2 ro
- snmp-server context cx1
- snmp-server context cx2
- snmp-server group admingroup 3 auth read readview write writeview
- snmp-server group admingroup 3 auth read readview context cx1
- snmp-server group admingroup 3 auth read readview context cx2
- snmp-server user admin admingroup 3 auth md5 helloworld
- snmp mib community-map VRF1 context cx1
- snmp mib community-map VRF2 context cx2
- snmp-server view readview .1 included
- snmp-server view writeview .1 included

2. Configure `snmp context` under the VRF instances.
   - sho run bgp
   - router bgp 100
   - address-family ipv4 vrf vrf1
   - snmp context context1
   - neighbor 20.1.1.1 remote-as 200
   - neighbor 20.1.1.1 no shutdown
   - exit-address-family
   - address-family ipv4 vrf vrf2
   - snmp context context2
   - timers bgp 30 90
   - neighbor 30.1.1.1 remote-as 200
   - neighbor 30.1.1.1 no shutdown
   - exit-address-family

Example of SNMP Walk Output for BGP timer configured for vrf1 (SNMPv2c)

```
snmpwalk -v 2c -c vrf1 10.16.131.125 1.3.6.1.4.1.6027.20.1.2.3
SNMPv2-SMI::enterprises.6027.20.1.2.3.1.1.1.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 950
SNMPv2-SMI::enterprises.6027.20.1.2.3.1.1.2.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 14
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.1.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 60
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.2.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 180
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.3.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 60
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.4.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 30
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.5.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 30
SNMPv2-SMI::enterprises.6027.20.1.2.3.3.1.1.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 180
SNMPv2-SMI::enterprises.6027.20.1.2.3.3.1.2.0.1.20.1.1.2.1.20.1.1.1 = Gauge32: 60
```

Example of SNMP Walk Output for BGP timer configured for vrf2 (SNMPv2c)

```
snmpwalk -v 2c -c vrf2 10.16.131.125 1.3.6.1.4.1.6027.20.1.2.3
SNMPv2-SMI::enterprises.6027.20.1.2.3.1.1.1.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 950
SNMPv2-SMI::enterprises.6027.20.1.2.3.1.1.2.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 14
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.1.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 60
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.2.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 90
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.3.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 30
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.4.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 30
SNMPv2-SMI::enterprises.6027.20.1.2.3.2.1.5.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 30
SNMPv2-SMI::enterprises.6027.20.1.2.3.3.1.1.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 90
SNMPv2-SMI::enterprises.6027.20.1.2.3.3.1.2.0.1.30.1.1.2.1.30.1.1.1 = Gauge32: 30
```

Example of SNMP Walk Output for BGP timer (SNMPv3)

```
snmpwalk -v 3 -a md5 -A helloworld -l authNoPriv -n cx1 -u admin 10.16.143.179
1.3.6.1.4.1.6027.20.1.3.6.1.4
SNMPv2-SMI::enterprises.6027.20.1.3.6.1.4.2963474636.0.1 = Gauge32: 200
SNMPv2-SMI::enterprises.6027.20.1.3.6.1.4.2963475124.0.1 = Gauge32: 100
```

Monitor Port-Channels

To check the status of a Layer 2 port-channel, use f10LinkAggMib (.1.3.6.1.4.1.6027.3.2). In the following example, Po 1 is a switchport and Po 2 is in Layer 3 mode.

Example of SNMP Trap for Monitored Port-Channels

```
[senthilnathan@lithium ~]$ snmpwalk -v 2c -c public 10.11.1.1 .1.3.6.1.4.1.6027.3.2.1.1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.2 = INTEGER: 2
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.2.1 = Hex-STRING: 00 01 E8 13 A5 C7
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.2.2 = Hex-STRING: 00 01 E8 13 A5 C8
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.3.1 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.3.2 = INTEGER: 1107755010
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.4.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.4.2 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.5.1 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.5.2 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.1 = STRING: "Gi 5/84 " << Channel member for Po1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.2 = STRING: "Gi 5/85 " << Channel member for Po2
dot3aCommonAggFdbIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.1.1107755009.1 = INTEGER: 1107755009
dot3aCommonAggFdbVlanId
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.2.1107755009.1 = INTEGER: 1
dot3aCommonAggFdbTagConfig
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.3.1107755009.1 = INTEGER: 2 (Tagged 1 or Untagged 2)
dot3aCommonAggFdbStatus
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.4.1107755009.1 = INTEGER: 1 << Status active, 2 –
status inactive
```

If we learn MAC addresses for the LAG, status is shown for those as well.

Example of Viewing Status of Learned MAC Addresses

```
dot3aCurAggVlanId
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.1.1.0.0.0.0.0.1.1 = INTEGER: 1
dot3aCurAggMacAddr
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.2.1.0.0.0.0.0.1.1 = Hex-STRING: 00 00 00 00 00
01
dot3aCurAggIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.3.1.0.0.0.0.0.1.1 = INTEGER: 1
dot3aCurAggStatus
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.4.1.0.0.0.0.0.1.1 = INTEGER: 1 << Status active, 2 –
status
inactive
```

Layer 3 LAG does not include this support. SNMP trap works for the Layer 2 / Layer 3 / default mode LAG.

Example of Viewing Changed Interface State for Monitored Ports

```
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500842) 23:36:48.42
SNMPv2-MIB::snmpTrapOID.0 = OID: IF-MIB::linkDown
IF-MIB::ifIndex.33865785 = INTEGER: 33865785
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_DN: Changed interface state
to down: Te 1/1"
2010-02-10 14:22:39 10.16.130.4 [10.16.130.4]:
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500842) 23:36:48.42
SNMPv2-MIB::snmpTrapOID.0 = OID: IF-MIB::linkDown
IF-MIB::ifIndex.1107755009 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_DN: Changed interface state to
down: Po 1"
2010-02-10 14:22:40 10.16.130.4 [10.16.130.4]:
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500932) 23:36:49.32 SNMPv2-MIB::snmpTrapOID.0 =
OID:
IF-MIB::linkUp IF-MIB::ifIndex.33865785 = INTEGER: 33865785 SNMPv2-
SMI::enterprises.6027.3.1.1.4.1.2 =
STRING: "OSTATE_UP: Changed interface state to up: Te 1/1"
2010-02-10 14:22:40 10.16.130.4 [10.16.130.4]:
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500934) 23:36:49.34 SNMPv2-MIB::snmpTrapOID.0 =
OID:
```

```
IF-MIB::linkUp IF-MIB::ifIndex.1107755009 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_UP: Changed interface state to up:
Po 1"
```

Troubleshooting SNMP Operation

When you use SNMP to retrieve management data from an SNMP agent on a Dell EMC Networking router, take into account the following behavior.

- When you query an IPv4 icmpMsgStatsInPkts object in the ICMP table by using the `snmpwalk` command, the output for echo replies may be incorrectly displayed. To correctly display this information under ICMP statistics, use the `show ip traffic` command.
- When you query an icmpStatsInErrors object in the icmpStats table by using the `snmpget` or `snmpwalk` command, the output for IPv4 addresses may be incorrectly displayed. To correctly display this information under IP and ICMP statistics, use the `show ip traffic` command.
- When you query an IPv4 icmpMsgStatsInPkts object in the ICMP table by using the `snmpwalk` command, the echo response output may not be displayed. To correctly display ICMP statistics, such as echo response, use the `show ip traffic` command.

Transceiver Monitoring

To retrieve and display the transceiver related parameters you can perform a snmpwalk transceiver table OID to retrieve transceiver details as per the MIB. This enables transceiver monitoring and identification of potential issues related to the transceivers on a switch.

- Ensure that SNMP is enabled on the device before running a query to retrieve the transceiver information.
- If Tx/Rx power, temperature, current, or voltage is not supported on optics, *snmp walk* returns a null ("") value.
- If optics are not inserted in a port, empty string is displayed.

Example of SNMP Output for Transceiver Monitoring

```
Dell $ snmpwalk -v1 -c public 10.16.150.210 1.3.6.1.4.1.6027.3.11.1.3.1.1 | grep 2106373
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.1.2113540 = STRING: "stack-unit-1
fixedmodule-5 port-1"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.2.2113540 = STRING: "Te 1/5/1"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.3.2113540 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.4.2113540 = STRING: "10GBASE-SR"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.5.2113540 = STRING: "FINISAR CORP."
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.6.2113540 = STRING: "FTLX8571D3BCL-FC"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.7.2113540 = STRING: "AL20L80"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.8.2113540 = STRING: "-2.293689"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.9.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.10.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.11.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.12.2113540 = STRING: "-1.615917"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.13.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.14.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.15.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.16.2113540 = STRING: "29.109375"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.17.2113540 = STRING: "3.286000"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.18.2113540 = STRING: "7.530000"
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.19.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.20.2113540 = ""
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.21.2113540 = ""
```

Table 125. SNMP OIDs for Transceiver Monitoring

Field (OID)	Description
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.1	Device Name
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.2	Port
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.3	Optics Present

Table 125. SNMP OIDs for Transceiver Monitoring (continued)

Field (OID)	Description
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.4	Optics Type
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.5	Vendor Name
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.6	Part Number
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.7	Serial Number
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.8	Transmit Power Lane1
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.9	Transmit Power Lane2
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.10	Transmit Power Lane3
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.11	Transmit Power Lane4
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.12	Receive Power Lane1
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.13	Receive Power Lane2
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.14	Receive Power Lane3
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.15	Receive Power Lane4
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.16	Temperature
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.17	Volltage
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.18	Transmit Bias Current Lane1
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.19	Transmit Bias Current Lane2
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.20	Transmit Bias Current Lane3
SNMPv2-SMI::enterprises.6027.3.11.1.3.1.1.21	Transmit Bias Current Lane4

51

Stacking

Using the Dell EMC Networking OS stacking feature, you can interconnect multiple switch units with stacking ports or front end user ports. The stack becomes manageable as a single switch through the stack management unit.

The system accepts Unit ID numbers from 0 to 11 and it supports stacking up to six units.

Topics:

- Stacking Overview
- Important Points to Remember
- Stacking Installation Tasks
- Stacking Configuration Tasks
- Verify a Stack Configuration
- Remove Units or Front End Ports from a Stack
- Troubleshoot a Stack

Stacking Overview

Dell EMC Networking OS elects a management (master) unit, a standby unit, and all other units are member units.

Dell EMC Networking OS presents all of the units. For example, to access Ten GigabitEthernet Port 1 on Stack Unit 1, enter `interface tengigabitethernet 1/1` from CONFIGURATION mode.

Stack Management Roles

The stack elects the management units for the stack management.

- **Stack master** — primary management unit, also called the master unit.
- **Standby** — secondary management unit.
- **Stack units** — the remaining units in the stack, also called stack members. The system supports up to six stack units.
- **Stack group** — Each set of four 10G ports or each individual 40G port correspond to a stack-group. To configure the front ports on the device to be stacking-ports, use the CLI.

The master holds the control plane and the other units maintain a local copy of the forwarding databases. From the stack master you can configure:

- System-level features that apply to all stack members.
- Interface-level features for each stack member.

The master synchronizes the following information with the standby unit:

- Stack unit topology
- Stack running configuration (which includes ACL, LACP, STP, SPAN, and so on.)
- Logs.

The master switch maintains stack operation with minimal impact in the event of:

- Switch failure
- Inter-switch stacking link failure
- Switch insertion
- Switch removal

If the master switch goes off line, the standby replaces it as the new master and the switch with the next highest priority or MAC address becomes standby.

Stack Master Election

The stack elects a master and standby unit at bootup time based on two criteria.

- **Unit priority** — User-configurable. The range is from 1 to 14. A higher value (14) means a higher priority. The default is **0**. By removing the stack-unit priority using the `no stack-unit priority` command, you can set the priority back to the default value of zero. The unit with the highest priority is elected the master management unit; the unit with the second highest priority is elected the standby unit.
- **MAC address (in case of priority tie)** — The unit with the higher MAC value becomes the master unit. The stack takes the MAC address of the master unit and retains it unless it is reloaded.

To view which switch is the stack master, enter the `show system` command.

A change in the stack master occurs when:

- You power down the stack master or bring the master switch offline.
- A failover of the master switch occurs.
- You disconnect the master switch from the stack.

When a stack reloads and all the units come up at the same time; for example, when all units boot up from flash, all units participate in the election and the master and standby are chosen based on the priority or MAC address.

When the units do not boot up at the same time, such as when some units are powered down just after reloading and powered up later to join the stack, they do not participate in the election process even though the units that boot up late may have a higher priority configured. This happens because the master and standby have already been elected, hence the unit that boots up late joins only as a member.

When an up and running standalone unit or stack is merged with another stack, based on election, the losing stack reloads and the master unit of the winning stack becomes the master of the merged stack. For more details, see sectionsAdd Units to an Existing Stack and Remove a Unit from a Stack. It is possible to reset individual units to force them to give up the management role or reload the whole stack from the CLI to ensure a fully synchronized bootup.

Example of Viewing Stack Members

```
DellEMC#show system brief
Stack MAC : 00:01:e8:8c:53:32
Reload Type : normal-reload [Next boot : normal-reload]
-- Stack Info --
Unit UnitType    Status        ReqTyp CurTyp Version          Ports
-----------------------------------------------------------------
0    Member      not present
1    Management  online        S4810  S4810  4810-8-3-12-1447 64
2    Standby     online        S4810  S4810  4810-8-3-12-1447 64
3    Member      online        S4810  S4810  4810-8-3-12-1447 64
4    Member      online        S4810  S4810  4810-8-3-12-1447 64
5    Member      not present
6    Member      not present
7    Member      not present
8    Member      not present
9    Member      not present
10   Member      not present
```

Virtual IP

You can manage the stack using a single IP, known as a virtual IP, that is retained in the stack even after a failover.

The virtual IP address is used to log in to the current master unit of the stack. Both IPv4 and IPv6 addresses are supported as virtual IPs.

Use the following command to configure a virtual IP: `Dell(conf)#virtual-ip {ip-address | ipv6-address | dhcp}`

Failover Roles

If the stack master fails (for example, is powered off), it is removed from the stack topology.

The standby unit detects the loss of peering communication and takes ownership of the stack management, switching from the standby role to the master role. The distributed forwarding tables are retained during the failover, as is the stack MAC address. The lack of a standby unit triggers an election within the remaining units for a standby role.

After the former master switch recovers, despite having a higher priority or MAC address, it does not recover its master role but instead takes the next available role.

To view failover details, use the `show redundancy` command.

MAC Addressing on Stacks

The stack has three MAC addresses: the chassis MAC, interface MAC, and null interface MAC.

All interfaces in the stack use the interface MAC address of the management unit, and the chassis MAC for the stack is the master's chassis MAC. The stack continues to use the master's chassis MAC address even after a failover. The MAC address is not refreshed until the stack is reloaded and a different unit becomes the stack master.

NOTE: If the removed management unit is brought up as a standalone unit or as part of a different stack, there is a possibility of MAC address collisions.

A standalone is added to a stack. The standalone and the master unit have the same priority, but the standalone has a lower MAC address, so the standalone reboots. In the second example, a standalone is added to a stack. The standalone has a higher priority than the stack, so the stack (excluding the new unit) reloads.

Example of Adding a Standalone with a Lower MAC Address to a Stack

```
----------STANDALONE BEFORE CONNECTION---------------
Standalone#show system brief

Stack MAC : 00:01:e8:d5:ef:81

-- Stack Info --
Unit UnitType    Status      ReqTyp CurTyp Version Ports
-----------------------------------------------------
0   Management  online     S4820T
    S4820T 7.8.1.0  52
1   Member      not present
2   Member      not present
3   Member      not present
4   Member      not present
5   Member      not present
6   Member      not present
[output omitted]
Standalone#show system | grep priority
Master priority : 0
-----------STACK BEFORE CONNECTION----------------
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

-- Stack Info --
Unit UnitType    Status       ReqTyp CurTyp Version  Ports
-----------------------------------------------------
0    Standby     online       S4820T 7.8.1.0  52
1    Management  online       S4820T 7.8.1.0  52
2    Member      not present
3    Member      not present
4    Member      not present
5    Member      not present
6    Member      not present
7 Member not present
[output omitted]
Stack#show system stack-unit 0 | grep priority
Master priority : 0
Stack#show system stack-unit 1 | grep priority
Master priority : 0
```

Example of Adding a Standalone with a Lower MAC Address and Equal Priority to a Stack

```
---------------STANDALONE AFTER CONNECTION-----------------
Standalone#%STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is
present
00:20:20: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
00:20:22: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present

Going for reboot. Reason is Stack merge
[bootup messages omitted]
---------------STACK AFTER CONNECTION--------------------
Stack# 3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
3w1d14h: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 2 (type S4820T , 52
ports)
3w1d14h: %S4820T %CHMGR-0-PS_UP: Power supply 0 in unit 2 is up
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 2 is up

Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f
Reload-Type                : normal-reload [Next boot : normal-reload]
-- Stack Info --
Unit UnitType    Status     ReqTyp CurTyp Version  Ports
-----------------------------------------------------
0   Standby    online       S4820T 7.8.1.0  52  S4820T 7.8.1.0  52
1   Management online       S4820T 7.8.1.0  52   S4820T 7.8.1.0  52
2   Member     online       S4820T 7.8.1.0  52   S4820T 7.8.1.0  52
3   Member     not present
4   Member     not present
5   Member     not present
6   Member     not present
```

Stacking LAG

When multiple links are used between stack units, Dell EMC Networking OS automatically bundles them in a stacking LAG to provide aggregated throughput and redundancy.

The stacking LAG is established automatically and transparently by Dell EMC Networking OS (without user configuration) after peering is detected and behaves as follows:

- The stacking LAG dynamically aggregates; it can lose link members or gain new links.
- Shortest path selection inside the stack: If multiple paths exist between two units in the stack, the shortest path is used.

Supported Stacking Topologies

The device supports stacking in a ring or a daisy chain topology.

Dell EMC Networking recommends the ring topology when stacking the switches to provide redundant connectivity.

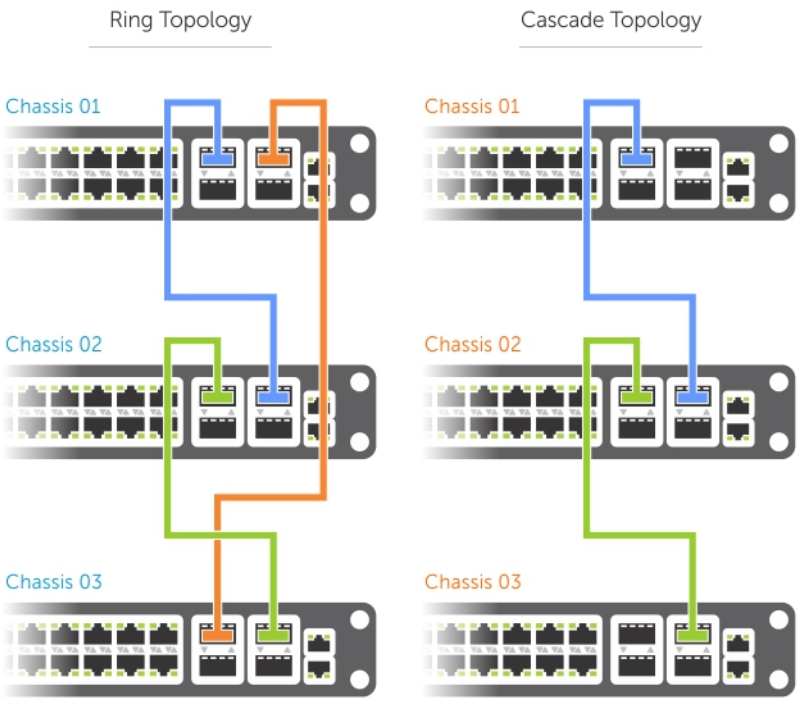


Figure 125. Supported Stacking Topologies

High Availability on Stacks

Stacks have master and standby management units analogous to Dell EMC Networking route processor modules (RPM).

The master unit synchronizes the running configuration and protocol states so that the system fails over in the event of a hardware or software fault on the master unit. In such an event, or when the master unit is removed, the standby unit becomes the stack manager and Dell EMC Networking OS elects a new standby unit. Dell EMC Networking OS resets the failed master unit: after online, it becomes a member unit; the remaining members remain online.

Example of Stack Manager Redundancy

Management Access on Stacks

You can access the stack via the console port or VTY line.

- **Console access** — You may access the stack through the console port of the master unit (stack manager) only. Similar to a standby RPM, the console port of the standby unit does not provide management capability; only a limited number of commands are available. Member units provide a limited set of commands.
- **Remote access** — You may access the master unit and standby unit in a stack through the dedicated management Ethernet interfaces with SNMP, SSH, or via Telnet.

Example of Accessing Non-Master Units on a Stack via the Console Port

```
-----------------CONSOLE ACCESS ON A STANDBY----------------------------
Dell(standby)#?
cd                              Change current directory
clear                           Reset functions
copy                            Copy from one file to another
delete                          Delete a file
dir                             List files on a filesystem
disable                         Turn off privileged commands
enable                          Turn on privileged commands
exit                            Exit from the EXEC
format                          Format a filesystem
fsck                            Filesystem check utility
```

```
pwd                                 Display current working directory
rename                              Rename a file
reset                               Reset selected card
show                                Show running system information
ssh-peer-stack-unit                 Open a SSH connection to the peer stack-unit
start                               Start shell
telnet-peer-stack-unit              Open a telnet connection to the peer stack-unit
terminal                            Set terminal line parameters
upload                              Upload file
Dell(standby)#


-----------------CONSOLE ACCESS ON A MEMBER----------------------------
Dell(stack-member-1)#?
reset-self Reset this unit alone
show Show running system information
```

You can connect two units with two or more stacking cables in case of a stacking port or cable failure. Removal of only one of the cables does not trigger a reset.

Important Points to Remember

- You can stack up to six systems.
- You cannot stack one system with other system types.
- You cannot enable stacking and virtual link trunking (VLT) simultaneously on the device. To convert a stacked unit to VLT, see Reconfiguring Stacked Switches as VLT.
- Data ports are configured as stacking ports in predefined groups of four 10G ports called stack-groups.
- When using the 40G ports, you can configure a single port as a stack port; each 40G port is a stack-group.
- All stack units must have the same version of Dell EMC Networking OS.

Stacking Installation Tasks

The following are the stacking installation tasks.

- Create a Stack
- Add Units to an Existing Stack
- Split a Stack

Create a Stack

Stacking is enabled on the device using the front end ports.

No configuration is allowed on front end ports used for stacking. Stacking can be made between 10G ports of two units or 40G ports of two units. The stack links between the two units are grouped into a single LAG.

Stack Group/Port Numbers

By default, each unit in Standalone mode is numbered stack-unit 0.

A maximum of eight 10G stack links or two 40G stack links can be made between two units in a stack. The front end ports are divided into 16 stack groups, each with 40G of bandwidth. Stack groups 0 through 11 correspond to 10G stack groups with four ports each. Stack groups 12 to 15 are one 40G port each.

The front end ports accommodate SFP, SFP+ and QSFP+.

- Ports are divided into 16 stack-groups (0 to 15) as shown in the following example. The stack groups must be of a single speed - either all 10G or all 40G.
  - stack-group 0 corresponds to ports 0-3, stack-group 1 corresponds to ports 4-7, so on through stack-group 11.
  - stack-group 12 corresponds to the 40G port 48, stack-group 13 corresponds to port 52, so on through stack group 15.

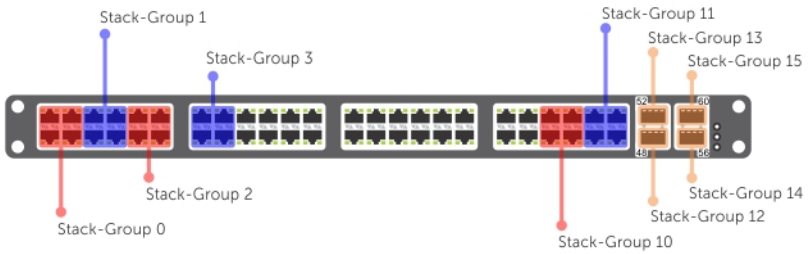


Figure 126. Stack-Group Assignments

You can connect the units while they are powered down or up. Stacking ports are bi-directional.

When a unit is added to a stack, the management unit performs a system check on the new unit to ensure the hardware type is compatible. A similar check is performed on the Dell EMC Networking OS version. If the stack is running Dell EMC Networking OS version 8.3.12.0 and the new unit is running an earlier software version, the new unit is put into a card problem state.

- If the unit is running Dell EMC Networking OS version 8.3.10.x, it is upgraded to use the same Dell EMC Networking OS version as the stack, rebooted, and joined the stack.
- If the new unit is running an Dell EMC Networking OS version prior to 8.3.10.x , the unit is put into a card problem state, Dell EMC Networking OS is not upgraded, and a syslog message is raised. The unit must be upgraded to Dell EMC Networking OS version 8.3.12.0 before you can proceed.

Syslog messages are generated by the management unit:

- before the management unit downloads its Dell EMC Networking OS version 8.3.12.0 or later to the new unit. The syslog includes the unit number, previous version, and version being downloaded.
- when the firmware synchronization is complete.
- if the system check fails, a message such as a hardware incompatibility message or incompatible uboot version is generated. If the unit is placed in a card problem state, the management unit also generates an SNMP trap.
- if the software version of the new unit predates Dell EMC Networking OS version 8.3.12.0 , the management unit puts the new unit into a card problem state and generates a syslog that identifies the unit, its Dell EMC Networking OS version, and its incompatibility for firmware synchronization.

NOTE: You must enter the `stack-unit` *`stack-unit`* `stack-group` *`stack-group`* command when adding units to a stack to ensure the units are assigned to the correct groups.

NOTE: Any scripts used to streamline the stacking configuration process must be updated to reflect the Command Mode change from EXEC to CONFIGURATION to allow the scripts to work correctly.

Enabling Front End Port Stacking

To enable the front ports on a unit for stacking, use the following commands.

NOTE: After a port has been allocated for stacking, you can only use it for stacking. If stack-group 0 is allocated for stacking, you can use ports 0, 1, 2, and 3 for stacking but not for Ethernet anymore. If only port 0 is used for stacking, ports 1, 2, and 3 are spare; they cannot be used for Ethernet.

NOTE: You can stack a maximum of eight 10G stack ports.

1. Assign a stack group for each unit.

   CONFIGURATION mode

   `stack-unit id stack-group` *`id`*

   Begin with the first port on the management unit. Next, configure both ports on each subsequent unit. Finally, return to the management unit and configure the last port.

   The range is from 0 to 15.
2. Save the stacking configuration on the ports.

   EXEC Privilege mode

   `write memory`
3. Reload the switch.

EXEC Privilege mode

`reload`

Dell EMC Networking OS automatically assigns a number to the new unit and adds it as member switch in the stack.

The new unit synchronizes its running and startup configurations with the stack.

4. After the units are reloaded, the system reboots. The units come up in a stack after the reboot completes.

To view the port assignments, use the `show system stack-unit` command.

Creating a New Stack

Prior to creating a stack, know which unit will be the management unit and which will be the standby unit.

Enable the front ports of the units for stacking. For more information, refer to Enabling Front End Port Stacking.

To create a new stack, use the following commands.

1. Power up all units in the stack.
2. Verify that each unit has the same Dell EMC Networking OS version prior to stacking them together.
   EXEC Privilege mode
   `show version`
3. Manually configure unit numbers for each unit, so that the stacking is deterministic upon boot up.
   EXEC Privilege mode
   `stack-unit` *`stack-unit-number`* `renumber` *`stack-unit-number`*.
   Renumbering causes the unit to reboot. The stack-unit default for all new units is stack-unit 0.
4. Configure the switch priority for each unit to make management unit selection deterministic.
   CONFIGURATION mode
   `stack-unit` *`stack-unit-number`* `priority` *`priority`*
5. Assign a stack group for each unit.
   CONFIGURATION mode
   `stack-unit` *`stack-unit-id`* `stack-group` *`stack-group-id`*
   Begin with the first port on the management unit. Next, configure both ports on each subsequent unit. Finally, return to the management unit and configure the last port. (refer to the following example.)
6. Connect the units using stacking cables.

   NOTE: The device does not require special stacking cables. The cables used to connect the data ports are sufficient.

7. Reload the stack one unit at a time.
   EXEC Privilege mode
   `show system brief`
   Start with the management unit, then the standby, then each of the members in order of their assigned stack number (or the position in the stack you want each unit to take).
   Allow each unit to completely boot, and verify that the stack manager detects the unit, then power the next unit.

Figure 127. Creating a New Stack

In the above example, stack unit 1 is the master management unit, stack unit 2 is the standby unit. The cables are connected to each unit.

Configure the stack groups on the units in the following order:

- Configure the first stack group on unit 1: `stack-unit 1 stack-group 13`
- Configure the stack groups on unit 2: `stack-unit 2 stack-group 14` and `stack-unit 2 stack-group 15`
- Configure the stack groups on unit 3: `stack-unit 3 stack-group 12` and `stack-unit 3 stack-group 13`
- Configure the stack groups on unit 4: `stack-unit 4 stack-group 13` and `stack-unit 4 stack-group 14`
- Configure the final stack-group on unit 1 to complete the stack: `stack-unit 1 stack-group 12`

When the stack-group configuration is complete, the system prints a syslog for reload.

```
DellEMC#configure
DellEMC(conf)#stack-unit 4 stack-group 13
DellEMC(conf)#02:39:12: %STKUNIT4-M:CP %IFMGR-6-STACK_PORTS_ADDED: Ports Fo 4/52 have
been configured as
stacking ports. Please save and reload for config to take effect
DellEMC(conf)#stack-unit 4 stack-group 14
DellEMC(conf)#02:39:15: %STKUNIT4-M:CP %IFMGR-6-STACK_PORTS_ADDED: Ports Fo 4/56 have
been configured as
stacking ports. Please save and reload for config to take effect
DellEMC(conf)#
DellEMC#02:39:18: %STKUNIT4-M:CP %SYS-5-CONFIG_I: Configured from console
```

Reload each unit in the stack. After the reload is complete, the four units come up as a stack with unit 1 as the management unit, unit 2 as the standby unit, and the remaining units as stack-members. All units in the stack can be accessed from the management unit.

NOTE: Before executing show commands, you must wait for 10 to 15 seconds for the download-agent process to start.

To view the stack unit information after the reload, use the `show system brief` command.

The following example shows how to configure two new switches for stacking using 10G ports.

```
 Dell-1(conf)#stack-unit 0 stack-group 0
Setting ports Te 1/1 Te 1/2 Te 1/3 as stack group will make their interface configs
obsolete after
a reload.
[confirm yes/no]:yes

Dell-2(conf)#stack-unit 0 stack-group 0
Setting ports Te 1/1 Te 1/2 Te 1/3 as stack group will make their interface configs
obsolete after
a reload.
[confirm yes/no]:yes

Dell-1#show system stack-ports
Topology: Ring
Interface Connection Link   Speed  Admin  Link   Trunk
                           (Gb/s)  Status Status Group
-----------------------------------------------------------------
1/1       1/1               10      up      up
1/2       1/2               10      up      up
1/3       1/3               10      up      up
```

Add Units to an Existing Stack

You can add units to an existing stack in one of three ways.

- By manually assigning a new unconfigured unit a position in an existing stack.
- By adding a configured unit to an existing stack.
- By merging two stacks.

If you are adding units to an existing stack, you can either:

- allow Dell EMC Networking OS to automatically assign the new unit a position in the stack, or
- manually determine each units position in the stack by configuring each unit to correspond with the stack before connecting it.
- If you add a unit that has a stack number that conflicts with the stack, the stack assigns the first available stack number.
- If the stack has a provision for the stack-number that will be assigned to the new unit, the provision must match the unit type, or Dell EMC Networking OS generates a type mismatch error.

After the new unit loads, it synchronizes its running and startup configurations with the stack.

Manually Assigning a New Unit to an Existing Stack

To manually assign a new unit a position in an existing stack, use the following steps.

1. On the stack, determine the next available stack-unit number, and the management priority of the management unit.
   EXEC Privilege mode
   `show system brief` or `show system stack-unit`
2. On the new unit, number it the next available stack-unit number.
   EXEC Privilege mode
   `stack-unit` *`stack-unit-number`* `renumber` *`stack-unit-new-number`*
3. (OPTIONAL) On the new unit, assign a management priority based on whether you want the new unit to be the stack manager.
   CONFIGURATION mode
   `stack-unit` *`stack-unit-number`* `priority` *`priority-number`*
4. Assign a stack group to each unit.
   CONFIGURATION mode
   `stack-unit`*`stack-unit-number`* `stack-group` *`stack-group-number`*
5. Connect the new unit to the stack using stacking cables.

The following example shows adding a stack unit with a conflicting stack number (before).

```
Dell#show system brief

Stack MAC : 00:01:e8:8a:df:e6

Reload Type : normal-reload

-- Stack Info --

Unit UnitType    Status       ReqTyp CurTyp Version Ports
-----------------------------------------------------------
0    Management  online       S4810  S4810  8-3-7-13  64
1    Member      not present
2    Member      not present
3    Standby     online       S4810  S4810  8-3-7-13  64
4    Member      not present
5    Member      not present
6    Member      not present
7    Member      not present
8    Member      not present
9    Member      not present
10   Member      not present
11   Member      not present
```

The following example shows adding a stack unit with a conflicting stack number (after).

```
Dell#show system brief
Stack MAC : 00:01:e8:8a:df:e6

Reload Type : normal-reload

-- Stack Info --

Unit UnitType    Status       ReqTyp CurTyp Version  Ports
------------------------------------------------------------
0    Management  online       S4810  S4810  8-3-7-13  64
1    Member      online       S4810  S4810  8-3-7-13  64
2    Member      not present
3    Standby     online       S4810  S4810  8-3-7-13  64
4    Member      not present
5    Member      not present
6    Member      not present
7    Member      not present
8    Member      not present
9    Member      not present
10   Member      not present
11   Member      not present
```

Adding a Configured Unit to an Existing Stack

To add a configured unit to an existing stack, use the following commands.

If a stack unit goes down and is removed from the stack, the logical provisioning configured for that stack-unit number is saved on the master and standby units. When a new unit is added to the stack, if a stack group configuration conflict occurs between the new unit and the provisioned stack unit, the configuration of the new unit takes precedence.

1. Add the configured unit to the top or bottom of the stack.
2. Power on the switch.
3. Attach cables to connect ports on the added switch to one or more existing switches in the stack.
4. Log on to the CLI and enter global configuration mode.
   - Login: username
   - Password: *****
   - DellEMC> enable
   - DellEMC# configure
5. Configure the ports on the added switch for stacking.

   CONFIGURATION mode

   `stack-unit 0 stack-group group-number`

   - `stack-unit 0`: defines the default ID unit-number in the initial configuration of a switch.
   - `stack-group group-number`: configures a port for stacking.
6. Save the stacking configuration on the ports.

   EXEC Privilege mode

   `write memory`
7. Reload the switch.

   EXEC Privilege mode

   `reload`

   Dell EMC Networking OS automatically assigns a number to the new unit and adds it as member switch in the stack. The new unit synchronizes its running and startup configurations with the stack.
8. If a standalone switch already has stack groups configured.

   Attach cables to connect the ports already configured as stack groups on the switch to one or more switches in the stack.

   Dell EMC Networking OS automatically assigns a number to the new unit and adds it as member switch in the stack. The new unit synchronizes its running and startup configurations with the stack.

Dell EMC Networking OS Behavior: When you add a switch to a stack

- If you configure the new unit with a stack number that is already assigned to a stack member, the stack avoids a numbering conflict by assigning the new switch the first available stack number.
- If the stack has been provisioned for the stack number that is assigned to the new unit, the pre-configured provisioning must match the switch type. If there is a conflict between the provisioned switch type and the new unit, a mismatch error message is displayed.

Merge Two Stacks

You may merge two stacks while they are powered and online.

To merge two stacks, connect one stack to the other using user port cables from the front end user portusing the mini-SAS cables from the stacking ports.

- Dell EMC Networking OS selects a master stack manager from the two existing managers based on the priority of the stack.
- Dell EMC Networking OS resets all the units in the losing stack; they all become stack members.
- If there is no unit numbering conflict, the stack members retain their previous unit numbers. Otherwise, the stack manager assigns new unit numbers, based on the order that they come online.
- The stack manager overwrites the startup and running config on the losing stack members with its own to synchronize the configuration on the new stack members.

Split a Stack

To split a stack, unplug the desired stacking cables.

You may do this at any time, whether the stack is powered or unpowered, and the units are online or offline. Each portion of the split stack retains the startup and running configuration of the original stack.

For a parent stack that is split into two child stacks, A and B, each with multiple units:

- If one of the new stacks receives the master and the standby management units, it is unaffected by the split.
- If one of the new stacks receives only the master unit, that unit remains the stack manager, and Dell EMC Networking OS elects a new standby management unit.
- If one of the new stacks receives only the standby unit, it becomes the master unit of the new stack, and Dell EMC Networking OS elects a new standby unit.
- If one of the new stacks receives neither the master nor the standby management unit, the stack is reset so that a new election can take place.

Stacking Configuration Tasks

Following are the stacking configuration tasks:

- Assigning Unit Numbers to Units in a Stack
- Creating a Virtual Stack Unit in a Stack
- Displaying Information About a Stack
- Influencing Management Unit Selection on a Stack
- Managing Redundancy on a Stack
- Resetting a Unit on a Stack
- Recover from Stack Link Flaps

Assigning Unit Numbers to Units in an Stack

Each unit in the stack has a stack number that is either assigned by you or Dell EMC Networking OS.

Units are numbered from 0 to 11, however, you can only stack six S4820T units. Stack numbers are stored in NVRAM and are preserved upon reload.

- Assign a stack-number to a unit.

  EXEC Privilege mode

  ```
  stack-unit old-unit-number renumber new-unit-number
  ```

Renumbering the stack manager triggers the whole stack to reload, as shown in the message below. When the stack comes back online, the master unit remains the management unit.

```
Dell#stack-unit 2 renumber 1
Renumbering master unit will reload the stack.
WARNING: Interface configuration for current unit will be lost!
Proceed to renumber [confirm yes/no]: yes
```

Creating a Virtual Stack Unit on a Stack

Use virtual stack units to configure ports on the stack before adding a new unit.

- Create a virtual stack unit.

  CONFIGURATION mode

  ```
  stack-unit stack-unit-number provision S4820TS4048–ON
  ```

Displaying Information about a Stack

To display information about the stack, use the following command.

- Display for stack-identity, status, and hardware information on every unit in a stack.

EXEC Privilege mode

`show system`

- Display most of the information in show system, but in a more convenient tabular form.

  EXEC Privilege mode

  `show system brief`
- Display the same information in show system, but only for the specified unit.

  EXEC Privilege mode

  `show system stack-unit`
- Display topology and stack link status for the entire stack.

  EXEC Privilege mode

  `show system stack-ports [status | topology]`

Display information about a switch stack using the `show system` command.

The following is an example of the `show system` command to view the stack details.

```
DellEMC#show system
Stack MAC : 00:01:e8:8a:df:e6
Reload Type : normal-reload
-- Unit 0 --
Unit Type       : Management Unit
Status          : online
Next Boot       : online
Required Type   : S4810 - 52-port GE/TE/FG (SE)
Current Type    : S4810 - 52-port GE/TE/FG (SE)
Master priority : 0
Hardware Rev    : 3.0
Num Ports       : 64
Up Time         : 57 min, 0 sec
Dell Networking OS Version    : 8-3-7-13
Jumbo Capable   : yes
POE Capable     : no
Burned In MAC : 00:01:e8:8a:df:e6
No Of MACs : 3

-- Power Supplies --
Unit Bay Status Type FanStatus
-----------------------------------------
0    0   absent      absent
0    1   up     AC   up

-- Fan Status --
Unit Bay TrayStatus Fan0 Speed Fan1 Speed
-----------------------------------------
0    0   up         up   6960  up   6960
0    1   up         up   6720  up   6720

Speed in RPM

-- Unit 1 --
Unit Type      : Member Unit
Status         : not present
Required Type  : S4810 - 52-port GE/TE/FG (SE)

-- Unit 2 --
Unit Type      : Member Unit
Status         : not present

-- Unit 3 --
Unit Type       : Standby Unit
Status          : online
Next Boot       : online
Required Type   : S4810 - 52-port GE/TE/FG (SE)
Current Type    : S4810 - 52-port GE/TE/FG (SE)
Master priority : 0
Hardware Rev    : 3.0
Num Ports       : 64
Up Time         : 57 min, 3 sec
```

```
Dell Networking OS Version    : 8-3-7-13
Jumbo Capable    : yes
POE Capable      : no
Burned In MAC    : 00:01:e8:8a:df:bf
No Of MACs       : 3

-----output truncated-----
```

The following is an example of the `show system brief` command to view the stack summary information.

```
DellEMC#show system brief

Stack MAC : 00:01:e8:8a:de:48

Reload Type : normal-reload

-- Stack Info --
Unit UnitType    Status      ReqTyp CurTyp  Version  Ports
-----------------------------------------------------------
0    Member      not present
1    Member      not present
2    Member      not present
3    Management online      S4810  S4810  8-3-12-13  64
4    Member     not present
5    Member     not present
6    Member     not present
7    Member     not present
8    Member     not present
9    Member     not present
10   Member     not present
11   Member     not present
```

The following example shows the `show system stack-ports` command.

```
DellEMC#show system stack-ports
Topology: Ring
Interface Connection Link Speed Admin  Link   Trunk
                     (Gb/s)     Status Status Group
-------------------------------------------------------------------
1/56      3/56        40        up     up
1/60      3/60        40        up     up
3/48                  40        up     down
3/52                  40        up     down
3/56      0/56        40        up     up
3/60      0/60        40        up     up
```

Influencing Management Unit Selection on a Stack

Stack priority is the system variable that Dell EMC Networking OS uses to determine which units in the stack are the master and standby management units. If multiple units tie for highest priority, the unit with the highest MAC address prevails.

If management was determined by priority only, a change in management occurs when:

- the management unit is powered down or a failover occurs.
- you disconnect the management unit from the stack.

When the management unit fails, the unit disappears from the stack topology. At that time, the standby unit detects the communication loss and switches from the standby unit role to the management unit role in the stack. From the remaining units in the stack, the system selects a new standby unit based on the unit priority using the same algorithm used when the stack was initially created. When the failed unit recovers, it takes the next available role, usually that of a stack member.

- Influence the selection of the stack management units.

  CONFIGURATION mode

  `stack-unit` *`unit-number`* `priority` *`priority-value`*

  The unit with the numerically highest priority is elected the master management unit, and the unit with the second highest priority is the standby unit.

  The range is from 1 to 14.

The default is **0**.

Managing Redundancy on a Stack

Use the following commands to manage the redundancy on a stack.

- Reset the current management unit and make the standby unit the new master unit.
  EXEC Privilege mode
  ```
  redundancy force-failover stack-unit
  ```
  A new standby is elected. When the former stack master comes back online, it becomes a member unit.
- Prevent the stack master from rebooting after a failover.
  CONFIGURATION mode
  ```
  redundancy disable-auto-reboot stack-unit
  ```
  This command does not affect a forced failover, manual reset, or a stack-link disconnect.
- Display redundancy information.
  EXEC Privilege mode
  ```
  show redundancy
  ```

Resetting a Unit on a Stack

You may reset any stack unit except for the master management unit, as shown in the following message.

```
% Error: Reset of master unit is not allowed.
```

To rest a unit on a stack, use the following commands.

- Reload a stack-unit.
  EXEC Privilege mode
  ```
  reset stack-unit unit-number
  ```
- Reload a member unit, from the unit itself.
  EXEC Privilege mode
  ```
  reset-self
  ```
- Reset a stack-unit when the unit is in a problem state.
  EXEC Privilege mode
  ```
  reset stack-unit unit-number {hard}
  ```

Verify a Stack Configuration

The light of the LED status indicator on the front panel of the stack identifies the unit's role in the stack.

- Off indicates the unit is a stack member.
- Blinking green indicates the unit is the stack standby.
- Solid green indicates the unit is the stack master (management unit).

Displaying the Status of Stacking Ports

To display the status of the stacking ports, including the topology, use the following command.

- Display the stacking ports.
  EXEC Privilege mode
  ```
  show system stack-ports
  ```

```
DellEMC#show system stack-ports
Topology: Ring
Interface Connection Link  Speed  Admin  Link   Trunk
```

```
                     (Gb/s) Status Status Group
-----------------------------------------------------
1/48   7/56          40      up     up
1/52   5/60          40      up     up
2/56   6/52          40      up     up
2/60   4/52          40      up     up
3/48   7/52          40      up     up
3/52   5/56          40      up     up
4/52   6/48          40      up     up
4/56   4/48          40      up     up
DellEMC#
```

The following example shows the parameters for the management unit in the stack.

```
DellEMC#show system stack-unit 1

-- Unit 1 --
Unit Type       : Management Unit
Status          : online
Next Boot       : online
Required Type   : S4810 - 52-port GE/TE/FG (SE)
Current Type    : S4810 - 52-port GE/TE/FG (SE)
Master priority : 0
Hardware Rev    : 3.0
Num Ports       : 64
Up Time         : 1 min, 14 sec
Dell Networking OS Version    : 4810-8-3-12-1447
Jumbo Capable   : yes
POE Capable     : no
Boot Flash      : 1.2.0.2
Memory Size     : 2147483648 bytes
Temperature     : 44C
Voltage         : ok
Serial Number   : H1DL104400018
Part Number     : Rev
Vendor Id       :
Date Code       :
Country Code    :
Piece Part ID   : N/A
PPID Revision   : N/A
Service Tag     : N/A
Expr Svc Code   : N/A
Auto Reboot     : disabled
Burned In MAC   : 00:01:e8:8c:53:32
No Of MACs      : 3

-- Power Supplies --
Unit Bay Status Type FanStatus
-----------------------------------------------
Unit Bay Status Type FanStatus
-----------------------------------------------
1    0   absent      absent
1    1   up     AC   up

-- Fan Status --
Unit Bay TrayStatus Fan0 Speed Fan1 Speed
-----------------------------------------------
1    0   up         up   7200  up   7200
1    1   up         up   7200  up   7440

Speed in RP
```

The following example shows three switches stacked together in a daisy chain topology.

```
stack-2#show system stack-ports
Topology: Daisy chain
Interface Connection Link    Speed   Admin  Link   Trunk
                     (Gb/s)  Status  Status Group
-----------------------------------------------------------
3/36      1/36       10      up      up
3/37      1/37       10      up      up
3/38      1/38       10      up      up
```

```
3/39      1/39       10      up      up
3/44      2/36       10      up      up
3/45      2/37       10      up      up
3/46      2/38       10      up      up
3/47      2/39       10      up      up
1/36      3/36       10      up      up
1/37      3/37       10      up      up
1/38      3/38       10      up      up
1/39      3/39       10      up      up
2/36      3/44       10      up      up
2/37      3/45       10      up      up
2/38      3/46       10      up      up
2/39      3/47       10      up      up
stack-2#
```

Remove Units or Front End Ports from a Stack

To remove units or front end ports from a stack, use the following instructions.

- Removing a Unit from a Stack
- Removing Front End Port Stacking

Removing a Unit from a Stack

The running-configuration and startup-configuration are synchronized on all stack units. A stack member that is disconnected from the stack maintains this configuration.

To remove a stack member from the stack, disconnect the stacking cables from the unit. You may do this at any time, whether the unit is powered or unpowered, online or offline.

NOTE: If you remove a unit in the middle of the daisy chain stack, the stack is split into multiple parts and each forms a new stack according to the stacking algorithm described throughout this chapter.

The following example shows removing a stack member (before).

```
DellEMC#show system brief

Stack MAC : 00:01:e8:8a:df:e6

Reload Type : normal-reload

-- Stack Info --
Unit UnitType   Status     ReqTyp CurTyp Version Ports
-----------------------------------------------------
0    Management online     S4810  S4810  8-3-7-13 64
1    Member     online     S4810  S4810  8-3-7-13 64
2    Member     not present
3    Standby    online     S4810  S4810  8-3-7-13 64
```

The following example shows removing a stack member (after).

```
DellEMC#show system brief

Stack MAC : 00:01:e8:8a:df:e6

Reload Type : normal-reload

-- Stack Info --
Unit UnitType   Status      ReqTyp  CurTyp Version Ports
-----------------------------------------------------
0    Management online      S4810   S4810   8-3-7-13 64
1    Member     not present S4810
2    Member     not present
3    Standby    online      S4810  S4810   8-3-7-13 64
4    Member     not present
5    Member     not present
6    Member     not present
7    Member     not present
```

```
8     Member      not present
9     Member      not present
10    Member      not present
11    Member      not present
```

NOTE: Each unit in the stack has a stack number that is either assigned by you or Dell EMC Networking OS. To manually renumber stack members, use the `stack-unit` *old-unit-number* `renumber` *new-unit-number* command. Renumbering stack members causes the entire stack to reload.

Removing Front End Port Stacking

To remove the configuration on the front end ports used for stacking, use the following commands.

1. Remove the stack group configuration that is configured.
   CONFIGURATION mode
   `no stack-unit` *id* `stack-group` *id*
2. Save the stacking configuration on the ports.
   EXEC Privilege mode
   `write memory`
3. Reload the switch.
   EXEC Privilege mode
   `reload`

After the units are reloaded, the system reboots. The units come up as standalone units after the reboot completes.

Troubleshoot a Stack

To troubleshoot a stack, use the following recovery tasks.

- Recover from Stack Link Flaps
- Recover from a Card Problem State on a Stack

Recover from Stack Link Flaps

Stack link integrity monitoring enables units to monitor their own stack ports and disable any stack port that flaps five times within 10 seconds.

Dell EMC Networking OS displays console messages for the local and remote members of a flapping link, and on the primary (master) and standby management units as KERN-2-INT messages if the flapping port belongs to either of these units.

In the following example, a stack-port on the master flaps. The remote member, Member 2, displays a console message, and the master and standby display KERN-2-INT messages.

To re-enable the downed stack-port, power cycle the offending unit.

Example of Console Messages About Flapping Link

```
----------------------MANAGMENT UNIT-------------------------------
Error: Stack Port 50 has flapped 5 times within 10 seconds.Shutting down this st
ack port now.
Error: Please check the stack cable/module and power-cycle the stack.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times w
ithin 10 seconds.Shutting down this stack port now.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Please check the stack cable/module
and power-cycle the stack.
----------------------STANDBY UNIT---------------------------------
10:55:18: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times within 10
seonds.Shutting down this stack port now.
10:55:18: %STKUNIT1-M:CP %KERN-2-INT: Error: Please check the stack cable/module
and power-cycle the stack.
----------------------MEMBER 2-------------------------------------
Error: Stack Port 51 has flapped 5 times within 10 seconds.Shutting down this stack port
```

```
now.
Error: Please check the stack cable/module and power-cycle the stack.
```

Recover from a Card Problem State on a Stack

If a unit added to a stack has a different Dell EMC Networking OS version, the unit does not come online and Dell EMC Networking OS cites a card problem error.

To recover, disconnect the new unit from the stack, change the Dell EMC Networking OS version to match the stack, and then reconnect it to the stack.

Example of Card Problem Error on a Stack - Different Dell EMC Networking OS Versions

```
stack-1#show system brief

Stack MAC : 00:01:e8:8a:fd:6e

Reload Type : normal-reload [Next boot : normal-reload]

-- Stack Info --
Unit UnitType   Status       ReqTyp CurTyp  Version   Ports
-----------------------------------------------------------
0    Standby    card problem S4810 unknown  64
1    Management online       S4810 S4810    8-3-10-223 64
2    Member     not present
3    Member     not present
4    Member     not present
5    Member     not present
6    Member     not present
7    Member     not present
8    Member     not present
9    Member     not present
10   Member     not present
11   Member     not present

-- Power Supplies --
Unit Bay Status Type FanStatus
-----------------------------------
0    0   down   DC   down
0    1   up     DC   up
1    0   absent absent
1    1   up     AC   up

-- Fan Status --
Unit Bay TrayStatus Fan0 Speed Fan1 Speed
-------------------------------------------
0    0   up         up   9360  up   9360
0    1   up         up   9600  up   9360
1    0   up         up   6720  up   6720
1    1   up         up   6960  up   6720

Speed in RPM

stack-1#
```

Recover from a Card Mismatch State on a Stack

A card mismatch occurs if the stack has a provision for the lowest available stack number which does not match the model of a newly added unit.

To recover, disconnect the new unit. Then, either:

- remove the provision from the stack, then reconnect the standalone unit, or
- renumber the standalone unit with another available stack number on the stack.

Example of Recovering from a Card Mismatch State on a Stack (S50N and S25N)

```
--------------------STANDALONE UNIT BEFORE---------------------
Standalone#show system brief
```

```
Stack MAC : 00:01:e8:d5:ef:81
-- Stack Info --
Unit UnitType   Status       ReqTyp CurTyp Version  Ports
---------------------------------------------------------
0    Management online    S50V  S50V   7.8.1.0  52
1    Member     not present  S50N
2    Member     not present  S50V
3    Member     not present  S50V
4    Member     not present
5    Member     not present
6    Member     not present
7    Member     not present
---------------------STACK BEFORE----------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
-- Stack Info --
Unit UnitType     Status      ReqTyp CurTyp Version  Ports
-----------------------------------------------------------
0    Member    not present  S25N
1    Management  online       S50N   S50N   7.8.1.0  52
2    Standby     online       S50V   S50V   7.8.1.0  52
3    Member      not present
4    Member      not present
5    Member      not present
6    Member      not present
7    Member      not present
------------------STANDALONE UNIT AFTER-------------------------
01:38:34: %STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
Going for reboot. Reason is Stack merge
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
------------------STACK AFTER to display Core 2 RIP
activity-----------------------------------
23:11:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:11:40: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
23:12:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:12:34: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52
ports)
23:12:34: %STKUNIT1-M:CP %CHMGR-3-STACKUNIT_MISMATCH: Mismatch: Stack unit 0 is type
S50V - type S25N
required

Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
-- Stack Info --
Unit UnitType   Status         ReqTyp CurTyp Version  Ports
-----------------------------------------------------------
0    Member    type mismatch  S25N  S50V    7.8.1.0  52
1    Management online         S50N   S50N   7.8.1.0  52
2    Standby    online         S50V   S50V   7.8.1.0  52
3    Member     not present
4    Member     not present
5    Member     not present
6    Member     not present
7    Member     not present
```

52

Storm Control

Storm control allows you to control unknown-unicast, muticast, and broadcast traffic on Layer 2 and Layer 3 physical interfaces.

Dell EMC Networking Operating System (OS) Behavior: Dell EMC Networking OS supports unknown-unicast, muticast, and broadcast control for Layer 2 and Layer 3 traffic.

Dell EMC Networking OS Behavior: The minimum number of packets per second (PPS) that storm control can limit on the device is two.

To view the storm control broadcast configuration `show storm-control broadcast | multicast | unknown-unicast | pfc-llfc[interface]` command.

EXEC Privilege

To view the storm control multicast configuration, use the `show storm-control broadcast | multicast | unknown-unicast | pfc-llfc [interface]` command.

EXEC Privilege

Example:

```
DellEMC#show storm-control multicast Tengigabitethernet 1/1

Multicast storm control configuration

Interface  Direction     Packets/Second
-----------------------------------------
Te 1/1     Ingress       5

DellEMC#
```

To display the storm control unknown-unicast configuration, use the `show storm-control unknown-unicast [interface]` command.

EXEC Privilege

Topics:

- Configure Storm Control

Configure Storm Control

Storm control is supported in INTERFACE mode and CONFIGURATION mode.

Configuring Storm Control from INTERFACE Mode

To configure storm control, use the following command.

From INTERFACE mode:

- You can only configure storm control for ingress traffic.
- If you configure storm control from both INTERFACE and CONFIGURATION mode, the INTERFACE mode configurations override the CONFIGURATION mode configurations.
- The storm control is calculated in packets per second.
- Configure storm control.

  INTERFACE mode
- Configure the packets per second of broadcast traffic allowed on an interface (ingress only).

  INTERFACE mode

  `storm-control broadcast packets_per_second in`

- Configure the packets per second of multicast traffic allowed on C-Series or S-Series interface (ingress only) network only.
  INTERFACE mode
  ```
  storm-control multicast packets_per_second in
  ```
- Shut down the port if it receives the PFC/LLFC packets more than the configured rate.
  INTERFACE mode
  ```
  storm-control pfc-llfc pps in shutdown
  ```
  NOTE: PFC/LLFC storm control enabled interface disables the interfaces if it receives continuous PFC/LLFC packets. It can be a result of a faulty NIC/Switch that sends spurious PFC/LLFC packets.

Configuring Storm Control from CONFIGURATION Mode

To configure storm control from CONFIGURATION mode, use the following command.

From CONFIGURATION mode you can configure storm control for ingress and egress traffic.

Do not apply per-virtual local area network (VLAN) quality of service (QoS) on an interface that has storm-control enabled (either on an interface or globally).

- Configure storm control.
  CONFIGURATION mode
- Configure the packets per second of broadcast traffic allowed in the network.
  CONFIGURATION mode
  ```
  storm-control broadcast packets_per_second in
  ```
- Configure the packets per second (pps) of multicast traffic allowed on C-Series and S-Series networks only.
  CONFIGURATION mode
  ```
  storm-control multicast packets_per_second in
  ```
- Configure the packets per second of unknown-unicast traffic allowed in or out of the network.
  CONFIGURATION mode
  ```
  storm-control unknown-unicast packets_per_second in
  ```

53

Spanning Tree Protocol (STP)

The spanning tree protocol (STP) is supported on Dell EMC Networking OS.

Topics:

- Protocol Overview
- Configure Spanning Tree
- Important Points to Remember
- Configuring Interfaces for Layer 2 Mode
- Enabling Spanning Tree Protocol Globally
- Adding an Interface to the Spanning Tree Group
- Modifying Global Parameters
- Modifying Interface STP Parameters
- Enabling PortFast
- Selecting STP Root
- STP Root Guard
- Enabling SNMP Traps for Root Elections and Topology Changes
- Configuring Spanning Trees as Hitless
- STP Loop Guard
- Displaying STP Guard Configuration

Protocol Overview

STP is a Layer 2 protocol — specified by IEEE 802.1d — that eliminates loops in a bridged topology by enabling only a single path through the network.

By eliminating loops, the protocol improves scalability in a large network and allows you to implement redundant paths, which can be activated after the failure of active paths. Layer 2 loops, which can occur in a network due to poor network design and without enabling protocols like xSTP, can cause unnecessarily high switch CPU utilization and memory consumption.

Dell EMC Networking OS supports three other variations of spanning tree, as shown in the following table.

Table 126. Dell EMC Networking OS Supported Spanning Tree Protocols

Dell EMC Networking Term	IEEE Specification
Spanning Tree Protocol (STP)	802.1d
Rapid Spanning Tree Protocol (RSTP)	802.1w
Multiple Spanning Tree Protocol (MSTP)	802.1s
Per-VLAN Spanning Tree Plus (PVST+)	Third Party

Configure Spanning Tree

Configuring spanning tree is a two-step process.

- Configuring Interfaces for Layer 2 Mode
- Enabling Spanning Tree Protocol Globally

Related Configuration Tasks

- Adding an Interface to the Spanning Tree Group
- Modifying Global Parameters
- Modifying Interface STP Parameters
- Enabling PortFast
- Prevent Network Disruptions with BPDU Guard
- STP Root Guard
- Enabling SNMP Traps for Root Elections and Topology Changes
- Configuring Spanning Trees as Hitless

Important Points to Remember

- STP is disabled by default.
- The Dell EMC Networking OS supports only one spanning tree instance (0). For multiple instances, enable the multiple spanning tree protocol (MSTP) or per-VLAN spanning tree plus (PVST+). You may only enable one flavor of spanning tree at any one time.
- All ports in virtual local area networks (VLANs) and all enabled interfaces in Layer 2 mode are automatically added to the spanning tree topology at the time you enable the protocol.
- To add interfaces to the spanning tree topology after you enable STP, enable the port and configure it for Layer 2 using the `switchport` command.
- The IEEE Standard 802.1D allows 8 bits for port ID and 8 bits for priority. The 8 bits for port ID provide port IDs for 256 ports.

Configuring Interfaces for Layer 2 Mode

All interfaces on all switches that participate in spanning tree must be in Layer 2 mode and enabled.

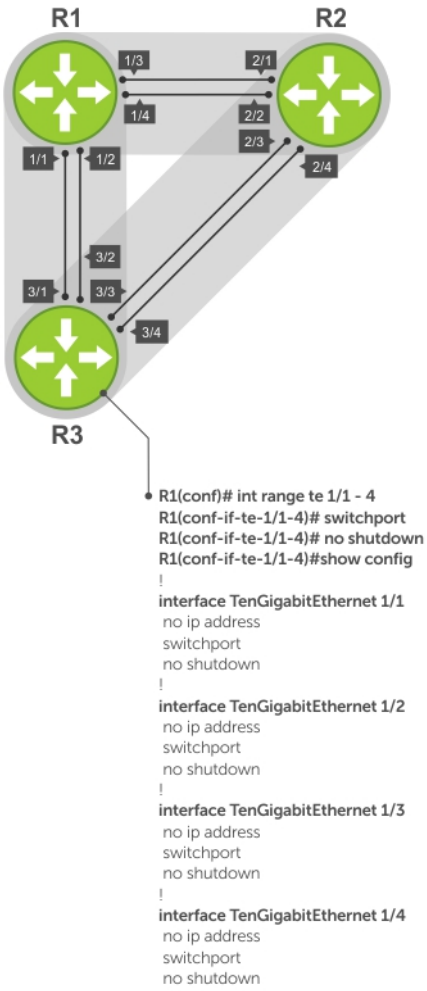


Figure 128. Example of Configuring Interfaces for Layer 2 Mode

To configure and enable the interfaces for Layer 2, use the following command.

1. If the interface has been assigned an IP address, remove it.
   INTERFACE mode
   `no ip address`
2. Place the interface in Layer 2 mode.
   INTERFACE
   `switchport`
3. Enable the interface.
   INTERFACE mode
   `no shutdown`

To verify that an interface is in Layer 2 mode and enabled, use the `show config` command from INTERFACE mode.

```
DellEMC(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
```

```
no shutdown
DellEMC(conf-if-te-1/1)#
```

Enabling Spanning Tree Protocol Globally

Enable the spanning tree protocol globally; it is not enabled by default.

When you enable STP, all physical, VLAN, and port-channel interfaces that are enabled and in Layer 2 mode are automatically part of the Spanning Tree topology.

- Only one path from any bridge to any other bridge participating in STP is enabled.
- Bridges block a redundant path by disabling one of the link ports.

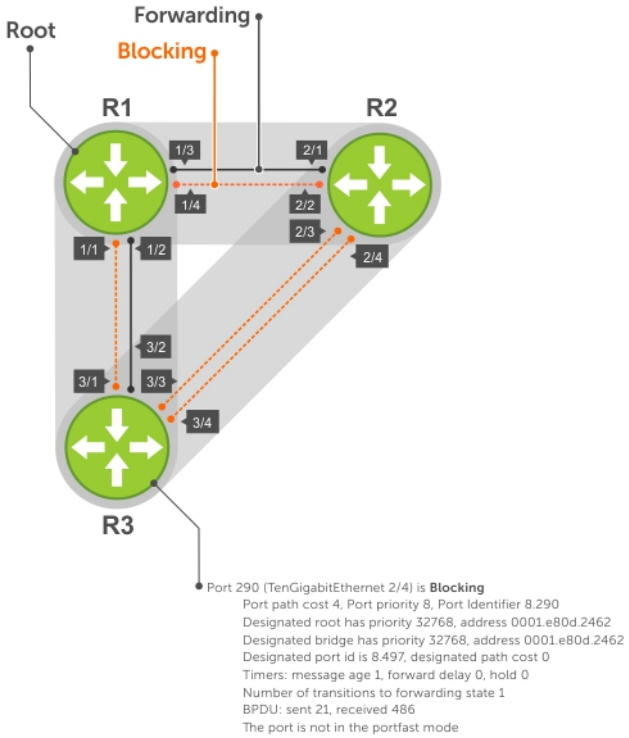


Figure 129. Spanning Tree Enabled Globally

To enable STP globally, use the following commands.

1. Enter PROTOCOL SPANNING TREE mode.
   CONFIGURATION mode
   `protocol spanning-tree 0`
2. Enable STP.
   PROTOCOL SPANNING TREE mode
   `no disable`

To disable STP globally for all Layer 2 interfaces, use the `disable` command from PROTOCOL SPANNING TREE mode.

To verify that STP is enabled, use the `show config` command from PROTOCOL SPANNING TREE mode.

```
DellEMC(conf)#protocol spanning-tree 0
DellEMC(config-span)#show config
!
protocol spanning-tree 0
  no disable
DellEMC#
```

To view the spanning tree configuration and the interfaces that are participating in STP, use the `show spanning-tree 0` command from EXEC privilege mode. If a physical interface is part of a port channel, only the port channel is listed in the command output.

```
R2#show spanning-tree 0
  Executing IEEE compatible Spanning Tree Protocol
    Bridge Identifier has priority 32768, address 0001.e826.ddb7
    Configured hello time 2, max age 20, forward delay 15
    Current root has priority 32768, address 0001.e80d.2462
    Root Port is 289 (TenGigabitEthernet 2/1), cost of root path is 4
    Topology change flag not set, detected flag not set
    Number of topology changes 3 last change occurred 0:16:11 ago
         from TenGigabitEthernet 2/3
    Timers: hold 1, topology change 35
            hello 2, max age 20, forward delay 15
    Times: hello 0, topology change 0, notification 0, aging Normal

  Port 289 (TenGigabitEthernet 2/1) is Forwarding
    Port path cost 4, Port priority 8, Port Identifier 8.289
    Designated root has priority 32768, address 0001.e80d.2462
    Designated bridge has priority 32768, address 0001.e80d.2462
    Designated port id is 8.496, designated path cost 0
    Timers: message age 1, forward delay 0, hold 0
    Number of transitions to forwarding state 1
    BPDU: sent 21, received 486
    The port is not in the portfast mode

  Port 290 (TenGigabitEthernet 2/2) is Blocking
    Port path cost 4, Port priority 8, Port Identifier 8.290
--More--
    Timers: message age 1, forward delay 0, hold 0
    Number of transitions to forwarding state 1
    BPDU: sent 21, received 486
    The port is not in the portfast mode
```

To confirm that a port is participating in Spanning Tree, use the `show spanning-tree 0 brief` command from EXEC privilege mode.

```
DellEMC#show spanning-tree 0 brief
  Executing IEEE compatible Spanning Tree Protocol
    Root ID Priority 32768, Address 0001.e80d.2462
    We are the root of the spanning tree
    Root Bridge hello time 2, max age 20, forward delay 15
    Bridge ID Priority 32768, Address 0001.e80d.2462
    Configured hello time 2, max age 20, forward delay 15
Interface                          Designated
Name    PortID Prio Cost Sts Cost  Bridge ID             PortID
-------------- ------ ---- ---- --- ----- --------------------
Te 1/1  8.496  8     4 DIS   0     32768 0001.e80d.2462  8.496
Te 1/2  8.497  8     4 DIS   0     32768 0001.e80d.2462  8.497
Te 1/3  8.513  8     4 FWD   0     32768 0001.e80d.2462  8.513
Te 1/4  8.514  8     4 FWD   0     32768 0001.e80d.2462  8.514
DellEMC#
```

Adding an Interface to the Spanning Tree Group

To add a Layer 2 interface to the spanning tree topology, use the following command.

- Enable spanning tree on a Layer 2 interface.

  INTERFACE mode

  `spanning-tree 0`

Modifying Global Parameters

You can modify the spanning tree parameters. The root bridge sets the values for forward-delay, hello-time, and max-age and overwrites the values set on other bridges participating in STP.

NOTE: Dell EMC Networking recommends that only experienced network administrators change the spanning tree parameters. Poorly planned modification of the spanning tree parameters can negatively affect network performance.

The following table displays the default values for STP.

Table 127. STP Default Values

STP Parameters	Default Value
Forward Delay	15 seconds
Hello Time	2 seconds
Max Age	20 seconds
Port Cost • 100-Mb/s Ethernet interfaces • 1-Gigabit Ethernet interfaces • 10-Gigabit Ethernet interfaces • 40-Gigabit Ethernet interfaces • Port Channel with 100 Mb/s Ethernet interfaces • Port Channel with 1-Gigabit Ethernet interfaces • Port Channel with 10-Gigabit Ethernet interfaces • Port Channel with 40-Gigabit Ethernet interfaces	• **200000** • **20000** • **2000** • **1400** • **180000** • **18000** • **1800** • **600**
Port Priority	8

- Change the `forward-delay` parameter (the wait time before the interface enters the Forwarding state).

  PROTOCOL SPANNING TREE mode

  `forward-delay` *`seconds`*

  The range is from 4 to 30.

  The default is **15 seconds**.
- Change the `hello-time` parameter (the BPDU transmission interval).

  PROTOCOL SPANNING TREE mode

  `hello-time` *`seconds`*

  NOTE: With large configurations (especially those with more ports) Dell EMC Networking recommends increasing the hello-time.

  The range is from 1 to 10.

  the default is **2 seconds**.
- Change the `max-age` parameter (the refresh interval for configuration information that is generated by recomputing the spanning tree topology).

  PROTOCOL SPANNING TREE mode

  `max-age` *`seconds`*

  The range is from 6 to 40.

  The default is **20 seconds**.

To view the current values for global parameters, use the `show spanning-tree 0` command from EXEC privilege mode. Refer to the second example in Enabling Spanning Tree Protocol Globally.

Modifying Interface STP Parameters

You can set the port cost and port priority values of interfaces in Layer 2 mode.

- **Port cost** — a value that is based on the interface type. The greater the port cost, the less likely the port is selected to be a forwarding port.
- **Port priority** — influences the likelihood that a port is selected to be a forwarding port in case that several ports have the same port cost.

The default values are listed in Modifying Global Parameters.

To change the port cost or priority of an interface, use the following commands.

- Change the port cost of an interface.

  INTERFACE mode

  `spanning-tree 0 cost cost`

  The range is from 0 to 65535.

  The default values are listed in Modifying Global Parameters.
- Change the port priority of an interface.

  INTERFACE mode

  `spanning-tree 0 priority priority-value`

  The range is from 0 to 15.

  The default is **8**.

To view the current values for interface parameters, use the `show spanning-tree 0` command from EXEC privilege mode. Refer to the second example in Enabling Spanning Tree Protocol Globally.

Enabling PortFast

The PortFast feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner.

Interfaces forward frames by default until they receive a BPDU that indicates that they should behave otherwise; they do not go through the Learning and Listening states. The `bpduguard shutdown-on-violation` option causes the interface hardware to be shut down when it receives a BPDU. When you only implement `bpduguard`, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree drops packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation.

CAUTION: Enable PortFast only on links connecting to an end station. PortFast can cause loops if it is enabled on an interface connected to a network.

To enable PortFast on an interface, use the following command.

- Enable PortFast on an interface.

  INTERFACE mode

  `spanning-tree stp-id portfast [bpduguard | [shutdown-on-violation]]`

To verify that PortFast is enabled on a port, use the `show spanning-tree` command from EXEC Privilege mode or the `show config` command from INTERFACE mode. Dell EMC Networking recommends using the `show config` command.

```
DellEMC#(conf-if-te-1/1)#show conf
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  spanning-tree 0 portfast
  no shutdown
DellEMC#(conf-if-te-1/1)#
```

Prevent Network Disruptions with BPDU Guard

Configure the Portfast (and Edgeport, in the case of RSTP, PVST+, and MSTP) feature on ports that connect to end stations. End stations do not generate BPDUs, so ports configured with Portfast/ Edgport (edgeports) do not expect to receive BDPUs.

If an edgeport does receive a BPDU, it likely means that it is connected to another part of the network, which can negatively affect the STP topology. The BPDU Guard feature blocks an edgeport after receiving a BPDU to prevent network disruptions, and Dell EMC Networking OS displays the following message.

```
3w3d0h: %RPM0-P:RP2 %SPANMGR-5-BPDU_GUARD_RX_ERROR: Received Spanning Tree BPDU on
BPDU guard port. Disable TenGigabitEthernet 3/4.
```

Enable BPDU Guard using the `bpduguard` option when enabling PortFast or EdgePort. The `bpduguard shutdown-on-violation` option causes the interface hardware to be shut down when it receives a BPDU. Otherwise, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree will only drop packets after a BPDU violation.

The following example shows a scenario in which an edgeport might unintentionally receive a BPDU. The port on the Dell EMC Networking system is configured with Portfast. If the switch is connected to the hub, the BPDUs that the switch generates might trigger an undesirable topology change. If you enable BPDU Guard, when the edge port receives the BPDU, the BPDU is dropped, the port is blocked, and a console message is generated.

NOTE: Unless you enable the `shutdown-on-violation` option, spanning-tree only drops packets after a BPDU violation; the physical interface remains up.

Dell EMC Networking OS Behavior: Regarding `bpduguard shutdown-on-violation` behavior:

- If the interface to be shut down is a port channel, all the member ports are disabled in the hardware.
- When you add a physical port to a port channel already in the Error Disable state, the new member port is also disabled in the hardware.
- When you remove a physical port from a port channel in the Error Disable state, the Error Disabled state is cleared on this physical port (the physical port is enabled in the hardware).
- You can clear the Error Disabled state with any of the following methods:
  - Perform a `shutdown` command on the interface.
  - Disable the `shutdown-on-violation` command on the interface (the `no spanning-tree` *`stp-id`* `portfast [bpduguard | [shutdown-on-violation]]` command).
  - Disable spanning tree on the interface (the `no spanning-tree` command in INTERFACE mode).
  - Disabling global spanning tree (the `no spanning-tree` in CONFIGURATION mode).

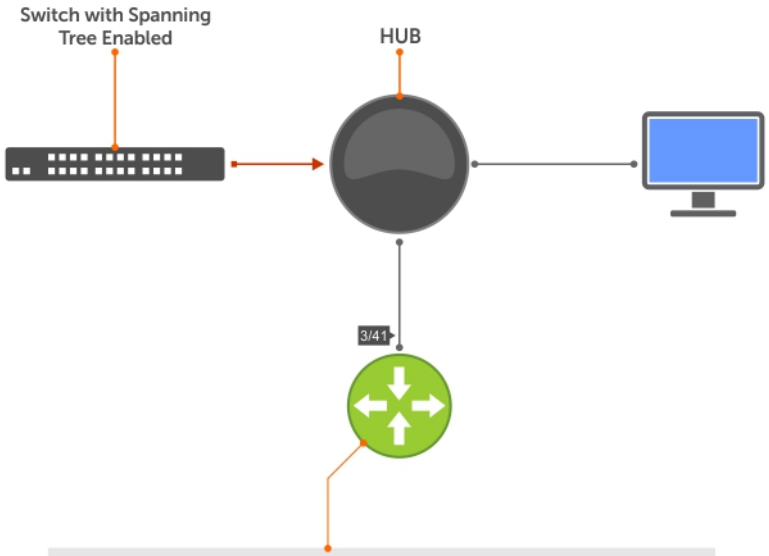


Figure 130. Enabling BPDU Guard

Dell EMC Networking OS Behavior: BPDU guard and BPDU filtering both block BPDUs, but are two separate features.

BPDU guard:

- is used on edgeports and blocks all traffic on edgeport if it receives a BPDU.
- drops the BPDU after it reaches the RP and generates a console message.

BPDU filtering:

- disables spanning tree on an interface
- drops all BPDUs at the line card without generating a console message

Example of Blocked BPDUs

```
DellEMC(conf-if-te-1/7)#do show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID Priority 32768, Address 0001.e805.fb07
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID Priority 32768, Address 0001.e85d.0e90
Configured hello time 2, max age 20, forward delay 15

Interface                              Designated
Name    PortID  Prio Cost  Sts Cost  Bridge ID            PortID
--------- -------- ---- ------- --- ------- -------------------
Te 1/6 128.263 128  20000 FWD 20000 32768 0001.e805.fb07 128.653
Te 1/7 128.264 128  20000 EDS 20000 32768 0001.e85d.0e90 128.264

Interface
Name    Role    PortID  Prio Cost  Sts Cost  Link-type Edge
--------- ------ -------- ---- ------- --- ----------------
Te 1/6 Root   128.263 128  20000 FWD 20000 P2P       No
Te 1/7 ErrDis 128.264 128  20000 EDS 20000 P2P       No
DellEMC(conf-if-te-1/7)#do show ip interface brief tengigabitEthernet 1/7
Interface           IP-Address OK Method  Status Protocol
TenGigabitEthernet 1/7 unassigned YES Manual up     up
```

Selecting STP Root

The STP determines the root bridge, but you can assign one bridge a lower priority to increase the likelihood that it becomes the root bridge. You can also specify that a bridge is the root or the secondary root.

To change the bridge priority or specify that a bridge is the root or secondary root, use the following command.

- Assign a number as the bridge priority or designate it as the root or secondary root.

  PROTOCOL SPANNING TREE mode

  `bridge-priority {priority-value | primary | secondary}`

  - *priority-value*: the range is from 0 to 65535. The lower the number assigned, the more likely this bridge becomes the root bridge.

  The primary option specifies a bridge priority of 8192.

  The secondary option specifies a bridge priority of 16384.

  The default is **32768**.

To view only the root information, use the `show spanning-tree root` command from EXEC privilege mode.

```
DellEMC#show spanning-tree 0 root
  Root ID Priority 32768, Address 0001.e80d.2462
  We are the root of the spanning tree
  Root Bridge hello time 2, max age 20, forward delay 15
DellEMC#
```

STP Root Guard

Use the STP root guard feature in a Layer 2 network to avoid bridging loops. In STP, the switch in the network with the lowest priority (as determined by STP or set with the `bridge-priority` command) is selected as the root bridge. If two switches have the same priority, the switch with the lower MAC address is selected as the root. All other switches in the network use the root bridge as the reference used to calculate the shortest forwarding path.

Because any switch in an STP network with a lower priority can become the root bridge, the forwarding topology may not be stable. The location of the root bridge can change, resulting in unpredictable network behavior. The STP root guard feature ensures that the position of the root bridge does not change.

Root Guard Scenario

For example, as shown in the following illustration (STP topology 1, upper left) Switch A is the root bridge in the network core. Switch C functions as an access switch connected to an external device. The link between Switch C and Switch B is in a Blocking state. The flow of STP BPDUs is shown in the illustration.

In STP topology 2 (shown in the upper right), STP is enabled on device D on which a software bridge application is started to connect to the network. Because the priority of the bridge in device D is lower than the root bridge in Switch A, device D is elected as root, causing the link between Switches A and B to enter a Blocking state. Network traffic then begins to flow in the directions indicated by the BPDU arrows in the topology. If the links between Switches C and A or Switches C and B cannot handle the increased traffic flow, frames may be dropped.

In STP topology 3 (shown in the lower middle), if you have enabled the root guard feature on the STP port on Switch C that connects to device D, and device D sends a superior BPDU that would trigger the election of device D as the new root bridge, the BPDU is ignored and the port on Switch C transitions from a forwarding to a root-inconsistent state (shown by the green X icon). As a result, Switch A becomes the root bridge.

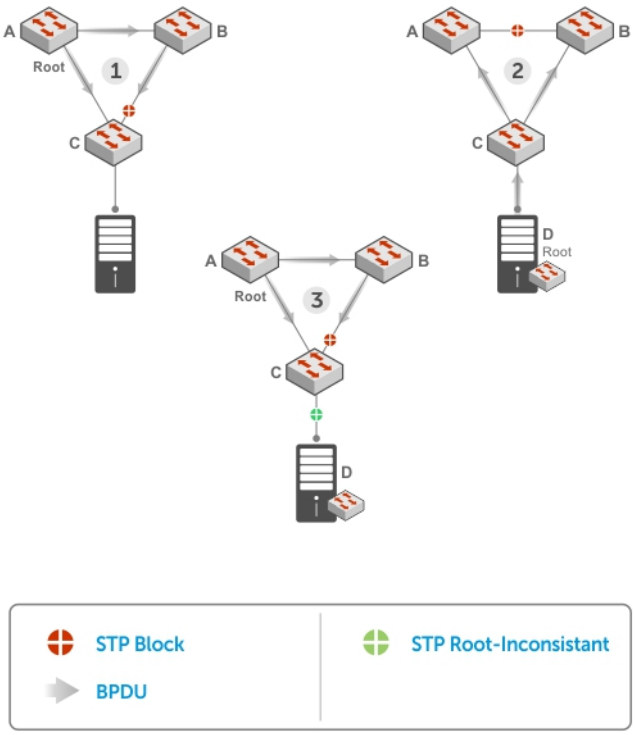


Figure 131. STP Root Guard Prevents Bridging Loops

Configuring Root Guard

Enable STP root guard on a per-port or per-port-channel basis.

Dell EMC Networking OS Behavior: The following conditions apply to a port enabled with STP root guard:

- Root guard is supported on any STP-enabled port or port-channel interface except when used as a stacking port.
- Root guard is supported on a port in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- When enabled on a port, root guard applies to all VLANs configured on the port.
- You cannot enable root guard and loop guard at the same time on an STP port. For example, if you configure root guard on a port on which loop guard is already configured, the following error message displays: • `% Error: LoopGuard is configured. Cannot configure RootGuard`.
- When used in an MSTP network, if root guard blocks a boundary port in the CIST, the port is also blocked in all other MST instances.

To enable the root guard on an STP-enabled port or port-channel interface in instance 0, use the following command.

- Enable root guard on a port or port-channel interface.

  INTERFACE mode or INTERFACE PORT-CHANNEL mode

  `spanning-tree {0 | mstp | rstp | pvst} rootguard`

  - `0`: enables root guard on an STP-enabled port assigned to instance 0.
  - `mstp`: enables root guard on an MSTP-enabled port.
  - `rstp`: enables root guard on an RSTP-enabled port.
  - `pvst`: enables root guard on a PVST-enabled port.

To disable STP root guard on a port or port-channel interface, use the `no spanning-tree 0 rootguard` command in an interface configuration mode.

To verify the STP root guard configuration on a port or port-channel interface, use the `show spanning-tree 0 guard [interface interface]` command in a global configuration mode.

Enabling SNMP Traps for Root Elections and Topology Changes

To enable SNMP traps individually or collectively, use the following commands.

- Enable SNMP traps for spanning tree state changes.

  `snmp-server enable traps stp`
- Enable SNMP traps for RSTP, MSTP, and PVST+ collectively.

  `snmp-server enable traps xstp`

Configuring Spanning Trees as Hitless

You can configure STP, RSTP, MSTP, and PVST+ to be hitless (configure all or none as hitless). When configured as hitless, critical protocol state information is synchronized between the RPMs so that RPM failover is seamless and no topology change is triggered.

To be hitless per spanning tree type or for all spanning tree types, use the following commands.

- Configure LACP to be hitless.

  CONFIGURATION mode

  `redundancy protocol lacp`
- Configure all spanning tree types to be hitless.

  CONFIGURATION mode

  `redundancy protocol xstp`

```
DellEMC(conf)#redundancy protocol xstp
DellEMC#show running-config redundancy
!
redundancy protocol xstp
DellEMC#
```

STP Loop Guard

The STP loop guard feature provides protection against Layer 2 forwarding loops (STP loops) caused by a hardware failure, such as a cable failure or an interface fault. When a cable or interface fails, a participating STP link may become unidirectional (STP requires links to be bidirectional) and an STP port does not receive BPDUs. When an STP blocking port does not receive BPDUs, it transitions to a Forwarding state. This condition can create a loop in the network.

For example, in the following example (STP topology 1, upper left), Switch A is the root switch and Switch B normally transmits BPDUs to Switch C. The link between Switch C and Switch B is in a Blocking state. However, if there is a unidirectional link failure (STP topology 1, lower left), Switch C does not receive BPDUs from Switch B. When the `max-age` timer expires, the STP port on Switch C becomes unblocked and transitions to Forwarding state. A loop is created as both Switch A and Switch C transmit traffic to Switch B.

As shown in the following illustration (STP topology 2, upper right), a loop can also be created if the forwarding port on Switch B becomes busy and does not forward BPDUs within the configured `forward-delay` time. As a result, the blocking port on Switch C transitions to a forwarding state, and both Switch A and Switch C transmit traffic to Switch B (STP topology 2, lower right).

As shown in STP topology 3 (bottom middle), after you enable loop guard on an STP port or port-channel on Switch C, if no BPDUs are received and the `max-age` timer expires, the port transitions from a blocked state to a Loop-Inconsistent state (instead of to a Forwarding state). Loop guard blocks the STP port so that no traffic is transmitted and no loop is created.

As soon as a BPDU is received on an STP port in a Loop-Inconsistent state, the port returns to a blocking state. If you disable STP loop guard on a port in a Loop-Inconsistent state, the port transitions to an STP blocking state and restarts the `max-age` timer.

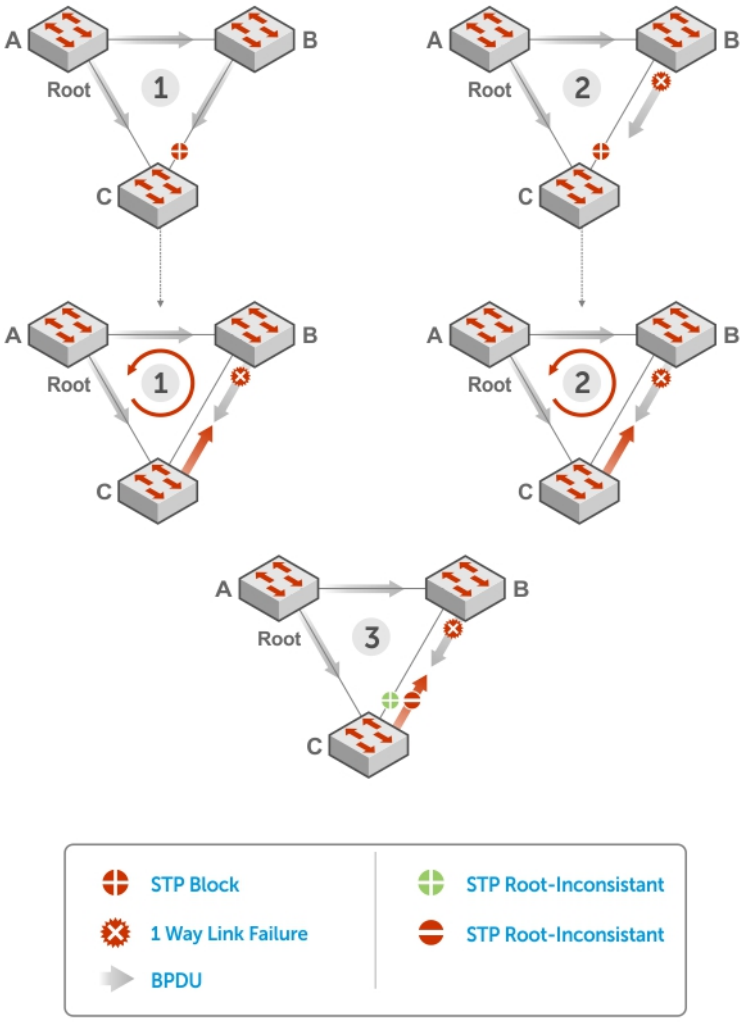


Figure 132. STP Loop Guard Prevents Forwarding Loops

Configuring Loop Guard

Enable STP loop guard on a per-port or per-port channel basis.

The following conditions apply to a port enabled with loop guard:

- Loop guard is supported on any STP-enabled port or port-channel interface.
- Loop guard is supported on a port or port-channel in any spanning tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- You cannot enable root guard and loop guard at the same time on an STP port. For example, if you configure loop guard on a port on which root guard is already configured, the following error message is displayed: `% Error: RootGuard is configured. Cannot configure LoopGuard`.
- Enabling Portfast BPDU guard and loop guard at the same time on a port results in a port that remains in a blocking state and prevents traffic from flowing through it. For example, when Portfast BPDU guard and loop guard are both configured:
  - If a BPDU is received from a remote device, BPDU guard places the port in an Err-Disabled Blocking state and no traffic is forwarded on the port.
  - If no BPDU is received from a remote device, loop guard places the port in a Loop-Inconsistent Blocking state and no traffic is forwarded on the port.

- When used in a PVST+ network, STP loop guard is performed per-port or per-port channel at a VLAN level. If no BPDUs are received on a VLAN interface, the port or port-channel transitions to a Loop-Inconsistent (Blocking) state only for this VLAN.

To enable a loop guard on an STP-enabled port or port-channel interface, use the following command.

- Enable loop guard on a port or port-channel interface.

  INTERFACE mode or INTERFACE PORT-CHANNEL mode

  `spanning-tree {0 | mstp | rstp | pvst} loopguard`

  - `0`: enables loop guard on an STP-enabled port assigned to instance 0.
  - `mstp`: enables loop guard on an MSTP-enabled port.
  - `rstp`: enables loop guard on an RSTP-enabled port.
  - `pvst`: enables loop guard on a PVST-enabled port.

To disable STP loop guard on a port or port-channel interface, use the `no spanning-tree 0 loopguard` command in an INTERFACE configuration mode.

To verify the STP loop guard configuration on a port or port-channel interface, use the `show spanning-tree 0 guard [interface interface]` command in a global configuration mode.

Displaying STP Guard Configuration

To display the STP guard configuration, use the following command.

The following example shows an STP network (instance 0) in which:

- Root guard is enabled on a port that is in a root-inconsistent state.
- Loop guard is enabled on a port that is in a listening state.
- BPDU guard is enabled on a port that is shut down (Error Disabled state) after receiving a BPDU.
- Verify the STP guard configured on port or port-channel interfaces.

  `show spanning-tree 0 guard [interface interface]`

```
DellEMC#show spanning-tree 0 guard
Interface
Name     Instance Sts          Guard type
--------- -------- ---------- ----------
Te 1/1   0        INCON(Root)  Rootguard
Te 1/2   0        LIS          Loopguard
Te 1/3   0        EDS (Shut)   Bpduguard
```

54

SupportAssist

SupportAssist sends troubleshooting data securely to Dell. SupportAssist in this Dell EMC Networking OS release does not support automated email notification at the time of hardware fault alert, automatic case creation, automatic part dispatch, or reports. SupportAssist requires Dell EMC Networking OS 9.9(0.0) and SmartScripts 9.7 or later to be installed on the Dell EMC Networking device. For more information on SmartScripts, see *Dell EMC Networking Open Automation guide*.

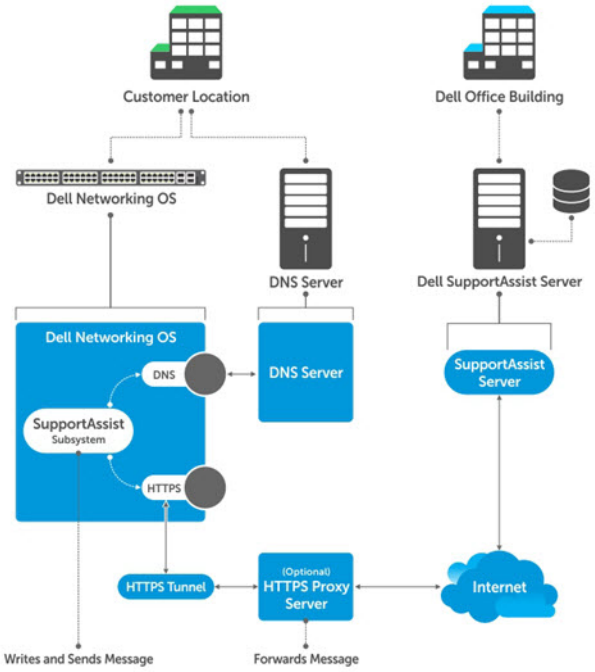


Figure 133. SupportAssist

NOTE: SupportAssist is enabled by default on the system. To disable SupportAssist, enter the `eula-consent support-assist reject` command in Global Configuration mode and save the configuration.

Topics:

- Configuring SupportAssist Using a Configuration Wizard
- Configuring SupportAssist Manually
- Configuring SupportAssist Activity
- Configuring SupportAssist Company
- Configuring SupportAssist Person
- Configuring SupportAssist Server
- Viewing SupportAssist Configuration

Configuring SupportAssist Using a Configuration Wizard

You are guided through a series of queries to configure SupportAssist. The generated commands are added to the running configuration, including the DNS resolve commands, if configured.

This command starts the configuration wizard for the SupportAssist. At any time, you can exit by entering `Ctrl-C`. If necessary, you can skip some data entry.

Enable the SupportAssist service.

CONFIGURATION mode

`support-assist activate`

```
DellEMC(conf)#support-assist activate
```

This command guides you through steps to configure SupportAssist.

Configuring SupportAssist Manually

To manually configure SupportAssist service, use the following commands.

1. Accept the end-user license agreement (EULA).

   CONFIGURATION mode

   `eula-consent {support-assist} {accept | reject}`

   NOTE: Once accepted, you do not have to accept the EULA again.

```
DellEMC(conf)# eula-consent support-assist accept
I accept the terms of the license agreement. You can reject
the license agreement by configuring this command
'eula-consent support-assist reject'.

By installing SupportAssist, you allow Dell to save your
contact information (e.g. name, phone number and/or
email address) which would be used to provide technical
support for your Dell products and services.  Dell may use
the information for providing recommendations to improve
your IT infrastructure.

Dell SupportAssist also collects and stores machine diagnostic
information, which may include but is not limited to configuration
information, user supplied contact information, names of data
volumes, IP addresses, access control lists, diagnostics &
performance information, network configuration information,
host/server configuration & performance information and related
data ("Collected Data") and transmits this information to Dell. By
downloading SupportAssist and agreeing to be bound by these terms
and the Dell end user license agreement, available at:
www.dell.com/aeula, you agree to allow Dell to provide remote
monitoring services of your IT environment and you give Dell the
right to collect the Collected Data in accordance with Dells Privacy
Policy, available at: www.dell.com/privacypolicycountryspecific,
in order to enable the performance of all of the various functions
of SupportAssist during your entitlement to receive related repair
services from Dell,.  You further agree to allow Dell to transmit
and store the Collected Data from SupportAssist in accordance with
these terms.  You agree that the provision of SupportAssist may
involve international transfers of data from you to Dell and/or
to Dells affiliates, subcontractors or business partners.  When
making such transfers, Dell shall ensure appropriate protection
is in place to safeguard the Collected Data being transferred in
connection with SupportAssist.  If you are downloading SupportAssist
on behalf of a company or other legal entity, you are further
certifying to Dell that you have appropriate authority to provide
this consent on behalf of that entity.  If you do not consent to
the collection, transmission and/or use of the Collected Data,
you may not download, install or otherwise use SupportAssist.
```

   NOTE: This step is not mandatory and you can configure SupportAssist manually without performing this step. Even before you accept or reject the EULA, the configuration data is sent to the default centrally deployed SupportAssist Server. If you reject the EULA, the configuration data is not transmitted to the SupportAssist server.

2. Move to the SupportAssist Configuration mode.

   To manually configure SupportAssist, use the following command.

   CONFIGURATION mode

```
support-assist
```

```
DellEMC(conf)#support-assist
DellEMC(conf-supportassist)#
```

3. (Optional) Configure the contact information for the company.

   SUPPORTASSIST mode

   ```
   contact-company name {company-name}[company-next-name] ... [company-next-name]
   ```

   ```
   DellEMC(conf)#support-assist
   DellEMC(conf-supportassist)#contact-company name test
   DellEMC(conf-supportassist-cmpy-test)#
   ```

4. (Optional) Configure the contact name for an individual.

   SUPPORTASSIST mode

   ```
   contact-person [first <first-name>] last <last-name>
   ```

   ```
   DellEMC(conf)#support-assist
   DellEMC(conf-supportassist)#contact-person first john last doe
   DellEMC(conf-supportassist-pers-john_doe)#
   ```

5. (Optional) Configure the name of the custom server and move to SupportAssist Server mode.

   SUPPORTASSIST mode

   ```
   server server-name
   ```

   ```
   DellEMC(conf)#support-assist
   DellEMC(conf-supportassist)#server default
   DellEMC(conf-supportassist-serv-default)#
   ```

   You can configure a maximum of two servers:

   - default server
   - custom user configured server

6. Enable all activities and servers for the SupportAssist service.

   SUPPORTASSIST mode

   ```
   enable all
   ```

   ```
   DellEMC(conf)#support-assist
   DellEMC(conf-supportassist)#enable all
   ```

7. Trigger an activity event immediately.

   EXEC Privilege mode

   ```
   support-assist activity {full-transfer | core-transfer} start now
   ```

   ```
   DellEMC#support-assist activity full-transfer start now
   ```

   ```
   DellEMC#support-assist activity core-transfer start now
   ```

Configuring SupportAssist Activity

SupportAssist Activity mode allows you to configure and view the action-manifest file for a specific activity.

To configure SupportAssist activity, use the following commands.

1. Move to the SupportAssist Activity mode for an activity. Allows you to configure customized details for a specific activity.

   SUPPORTASSIST mode

```
[no] activity {full-transfer|core-transfer|event-transfer}
```

```
DellEMC(conf-supportassist)#activity full-transfer
DellEMC(conf-supportassist-act-full-transfer)#
```

```
DellEMC(conf-supportassist)#activity core-transfer
DellEMC(conf-supportassist-act-core-transfer)#
```

```
DellEMC(conf-supportassist)#activity event-transfer
DellEMC(conf-supportassist-act-event-transfer)#
```

2. Copy an action-manifest file for an activity to the system.

   SUPPORTASSIST ACTIVITY mode

   `action-manifest get tftp | ftp | flash <file-specification> <local-file-name>`

```
DellEMC(conf-supportassist-act-full-transfer)#action-manifest get tftp://10.0.0.1/
test file
DellEMC(conf-supportassist-act-full-transfer)#
```

```
DellEMC(conf-supportassist-act-event-transfer)#action-manifest get tftp://10.0.0.1/
test file
DellEMC(conf-supportassist-act-event-transfer)#
```

3. Configure the action-manifest to use for a specific activity.

   SUPPORTASSIST ACTIVITY mode

   `[no] action-manifest install {default | <local-file-name>}`

```
DellEMC(conf-supportassist-act-full-transfer)#action-manifest install
custom_file1.json
DellEMC(conf-supportassist-act-full-transfer)#
```

```
DellEMC(conf-supportassist-act-event-transfer)#action-manifest install
custom_event_file1.json
DellEMC(conf-supportassist-act-event-transfer)#
```

4. View the list of action-manifest for a specific activity.

   SUPPORTASSIST ACTIVITY mode

   `action-manifest show {all}`

```
DellEMC(conf-supportassist-act-full-transfer)#action-manifest show all
custom_file1.json
DellEMC(conf-supportassist-act-full-transfer)#
```

```
DellEMC(conf-supportassist-act-event-transfer)#action-manifest show all
custom_event_file1.json [installed]
DellEMC(conf-supportassist-act-event-transfer)#
```

5. Remove the action-manifest file for an activity.

   SUPPORTASSIST ACTIVITY mode

   `action-manifest remove <local-file-name>`

```
DellEMC(conf-supportassist-act-full-transfer)#action-manifest remove custom_file1.json
DellEMC(conf-supportassist-act-full-transfer)#
```

```
DellEMC(conf-supportassist-act-event-transfer)#action-manifest remove
custom_event_file1.json
DellEMC(conf-supportassist-act-event-transfer)#
```

6. Enable a specific SupportAssist activity.

   By default, the full transfer includes the core files. When you disable the core transfer activity, the full transfer excludes the core files.

SUPPORTASSIST ACTIVITY mode

`[no] enable`

```
DellEMC(conf-supportassist-act-full-transfer)#enable
DellEMC(conf-supportassist-act-full-transfer)#
```

```
DellEMC(conf-supportassist-act-core-transfer)#enable
DellEMC(conf-supportassist-act-core-transfer)#
```

```
DellEMC(conf-supportassist-act-event-transfer)#enable
DellEMC(conf-supportassist-act-event-transfer)#
```

Configuring SupportAssist Company

SupportAssist Company mode allows you to configure name, address and territory information of the company. SupportAssist Company configurations are optional for the SupportAssist service.

To configure SupportAssist company, use the following commands.

1. Configure the contact information for the company.

   SUPPORTASSIST mode

   `[no] contact-company name {company-name}[company-next-name] ... [company-next-name]`

   ```
   DellEMC(conf-supportassist)#contact-company name test
   DellEMC(conf-supportassist-cmpy-test)#
   ```

2. Configure the address information for the company.

   SUPPORTASSIST COMPANY mode

   `[no] address [city company-city] [{province | region | state} name] [country company-country] [{postalcode | zipcode] company-code]`

   ```
   DellEMC(conf-supportassist-cmpy-test)#address city MyCity state MyState country
   MyCountry
   DellEMC(conf-supportassist-cmpy-test)#
   ```

3. Configure the street address information for the company.

   SUPPORTASSIST COMPANY mode

   `[no] street-address {address1}[address2]…[address8]`

   ```
   DellEMC(conf-supportassist-cmpy-test)#street-address 123 Main Street
   DellEMC(conf-supportassist-cmpy-test)#
   ```

4. Configure the territory and set the coverage for the company site.

   SUPPORTASSIST COMPANY mode

   `[no] territory company-territory`

   ```
   DellEMC(conf-supportassist-cmpy-test)#territory IN
   DellEMC(conf-supportassist-cmpy-test)#
   ```

Configuring SupportAssist Person

SupportAssist Person mode allows you to configure name, email addresses, phone, method and time zone for contacting the person. SupportAssist Person configurations are optional for the SupportAssist service.

To configure SupportAssist person, use the following commands.

1. Configure the contact name for an individual.

   SUPPORTASSIST mode

```
[no] contact-person [first <first-name>] last <last-name>
```

```
DellEMC(conf-supportassist)#contact-person first john last doe
DellEMC(conf-supportassist-pers-john_doe)#
```

2. Configure the email addresses to reach the contact person.

   SUPPORTASSIST PERSON mode

   ```
   [no] email-address primary email-address [alternate email-address]
   ```

   ```
   DellEMC(conf-supportassist-pers-john_doe)#email-address primary jdoe@mycompany.com
   DellEMC(conf-supportassist-pers-john_doe)#
   ```

3. Configure phone numbers of the contact person.

   SUPPORTASSIST PERSON mode

   ```
   [no] phone primary phone [alternate phone]
   ```

   ```
   DellEMC(conf-supportassist-pers-john_doe)#phone primary +919999999999
   DellEMC(conf-supportassist-pers-john_doe)#
   ```

4. Configure the preferred method for contacting the person.

   SUPPORTASSIST PERSON mode

   ```
   preferred-method {email | no-contact | phone]
   ```

   ```
   DellEMC(conf-supportassist-pers-john_doe)#preferred-method email
   DellEMC(conf-supportassist-pers-john_doe)#
   ```

5. Configure the time frame for contacting the person.

   SUPPORTASSIST PERSON mode

   ```
   [no] time-zone zone +-HH:MM[start-time HH:MM] [end-time HH:MM]
   ```

   ```
   DellEMC(conf-supportassist-pers-john_doe)#time-zone zone +01:24 start-time 12:00 end-
   time 23:00
   DellEMC(conf-supportassist-pers-john_doe)#
   ```

Configuring SupportAssist Server

SupportAssist Server mode allows you to configure server name and the means of reaching the server. By default, a SupportAssist server URL has been configured on the device. Configuring a URL to reach the SupportAssist remote server should be done only under the direction of Dell SupportChange.

To configure SupportAssist server, use the following commands.

1. Configure the name of the remote SupportAssist Server and move to SupportAssist Server mode.

   SUPPORTASSIST mode

   ```
   [no] server server-name
   ```

   ```
   DellEMC(conf-supportassist)#server default
   DellEMC(conf-supportassist-serv-default)#
   ```

2. Configure a proxy for reaching the SupportAssist remote server.

   SUPPORTASSIST SERVER mode

   ```
   [no] proxy-ip-address {ipv4-address | ipv6-address}port port-number [ username userid
   password [encryption-type] password ]
   ```

   ```
   DellEMC(conf-supportassist-serv-default)#proxy-ip-address 10.0.0.1 port 1024 username
   test password 0 test1
   DellEMC(conf-supportassist-serv-default)#
   ```

3. Enable communication with the SupportAssist server.

   SUPPORTASSIST SERVER mode

[no] enable

```
DellEMC(conf-supportassist-serv-default)#enable
DellEMC(conf-supportassist-serv-default)#
```

4. Configure the URL to reach the SupportAssist remote server.

   SUPPORTASSIST SERVER mode

   [no] url *uniform-resource-locator*

```
DellEMC(conf-supportassist-serv-default)#url https://192.168.1.1/index.htm
DellEMC(conf-supportassist-serv-default)#
```

Viewing SupportAssist Configuration

To view the SupportAssist configurations, use the following commands:

1. Display information on the SupportAssist feature status including any activities, status of communication, last time communication sent, and so on.

   EXEC Privilege mode

   show support-assist status

```
DellEMC#show support-assist status
SupportAssist Service: Installed
EULA: Accepted
Server: default
    Enabled: Yes
    URL: https://stor.g3.ph.dell.com
Server: Dell
    Enabled: Yes
    URL: http://1.1.1.1:1337
Service status: Enabled

Activity                 State    Last Start                Last Success
-----------------------  -------  ------------------------  ------------------------
core-transfer            Success  Feb 15 2016 09:43:41 IST    Feb 15 2016 09:43:56 IST
event-transfer           Success  Feb 15 2016 09:47:43 IST    Feb 15 2016 09:48:21
IST
full-transfer            Success  Feb 15 2016 09:36:12 IST    Feb 15 2016 09:38:27 IST
DellEMC#
```

2. Display the current configuration and changes from the default values.

   EXEC Privilege mode

   show running-config support-assist

```
DellEMC# show running-config support-assist
!
support-assist
enable all
!
activity event-transfer
  enable
  action-manifest install default
!
activity core-transfer
  enable
!
contact-company name Dell
  street-address F lane , Sector 30
  address city Brussels state HeadState country Belgium postalcode S328J3
!
contact-person first Fred last Nash
  email-address primary des@sed.com alternate sed@dol.com
  phone primary 123422 alternate 8395729
  preferred-method email
  time-zone zone +05:30 start-time 12:23 end-time 15:23
```

```
!
server Dell
  enable
  url http://1.1.1.1:1337
DellEMC#
```

3. Display the EULA for the feature.

   EXEC Privilege mode

   `show eula-consent {support-assist | *other feature*}`

```
DellEMC#show eula-consent support-assist

SupportAssist EULA has been: Accepted

Additional information about the SupportAssist EULA is as follows:

By installing SupportAssist, you allow Dell to save your contact information
(e.g. name, phone number and/or email address) which would be used to provide
technical support for your Dell products and services.  Dell may use the information
for providing recommendations to improve  your IT infrastructure.

Dell SupportAssist also collects and stores machine diagnostic information, which
may include but is not limited to configuration information, user supplied contact
information, names of data volumes, IP addresses, access control lists, diagnostics &
performance information, network configuration information, host/server configuration
& performance information and related data (Collected Data) and transmits this
information to Dell. By downloading SupportAssist and agreeing to be bound by these
terms and the Dell end user license agreement, available at: www.dell.com/aeula,
you agree to allow Dell to provide remote monitoring services of your IT environment
and you give Dell the right to collect the Collected Data in accordance with Dells
Privacy Policy, available at: www.dell.com/privacypolicycountryspecific, in order to
enable the performance of all of the various functions of SupportAssist during your
entitlement to receive related repair services from Dell,.  You further agree to
allow Dell to transmit and store the Collected Data from SupportAssist in accordance
with these terms.  You agree that the provision of SupportAssist may involve
international transfers of data from you to Dell and/or to Dells affiliates,
subcontractors or business partners.  When making such transfers, Dell shall ensure
appropriate protection is in place to safeguard the Collected Data being transferred
in connection with SupportAssist.  If you are downloading SupportAssist on behalf
of a company or other legal entity, you are further certifying to Dell that you
have appropriate authority to provide this consent on behalf of that entity.  If you
do not consent to the collection, transmission and/or use of the Collected Data,
you may not download, install or otherwise use SupportAssist.
```

55

System Time and Date

System time and date settings and the network time protocol (NTP) are supported on Dell EMC Networking OS.

You can set system times and dates and maintained through the NTP. They are also set through the Dell EMC Networking Operating System (OS) command line interfaces (CLIs) and hardware settings.

The Dell EMC Networking OS supports reaching an NTP server through different VRFs. You can configure a maximum of eight logging servers across different VRFs or the same VRF.

Topics:

- Network Time Protocol
- Dell EMC Networking OS Time and Date

Network Time Protocol

The network time protocol (NTP) synchronizes timekeeping among a set of distributed time servers and clients.

The protocol also coordinates time distribution in a large, diverse network with various interfaces. In NTP, servers maintain the time and NTP clients synchronize with a time-serving host. NTP clients choose from among several NTP servers to determine which offers the best available source of time and the most reliable transmission of information.

NTP is a fault-tolerant protocol that automatically selects the best of several available time sources to synchronize to. You can combine multiple candidates to minimize the accumulated error. Temporarily or permanently insane time sources are detected and avoided.

Dell EMC Networking recommends configuring NTP for the most accurate time. In Dell EMC Networking OS, you can configure other time sources (the hardware clock and the software clock).

NTP is designed to produce three products: clock offset, roundtrip delay, and dispersion, all of which are relative to a selected reference clock.

- **Clock offset** — represents the amount to adjust the local clock to bring it into correspondence with the reference clock.
- **Roundtrip delay** — provides the capability to launch a message to arrive at the reference clock at a specified time.
- **Dispersion** — represents the maximum error of the local clock relative to the reference clock.

Because most host time servers synchronize via another peer time server, there are two components in each of these three products, those determined by the peer relative to the primary reference source of standard time and those measured by the host relative to the peer.

In order to facilitate error control and management of the subnet itself, each of these components is maintained separately in the protocol. They provide not only precision measurements of offset and delay, but also definitive maximum error bounds, so that the user interface can determine not only the time, but the quality of the time as well.

In what may be the most common client/server model, a client sends an NTP message to one or more servers and processes the replies as received. The server interchanges addresses and ports, overwrites certain fields in the message, recalculates the checksum and returns the message immediately. Information included in the NTP message allows the client to determine the server time regarding local time and adjust the local clock accordingly. In addition, the message includes information to calculate the expected timekeeping accuracy and reliability, as well as select the best from possibly several servers.

Following conventions established by the telephone industry [BEL86], the accuracy of each server is defined by a number called the stratum, with the topmost level (primary servers) assigned as one and each level downwards (secondary servers) in the hierarchy assigned as one greater than the preceding level.

Dell EMC Networking OS synchronizes with a time-serving host to get the correct time. You can set Dell EMC Networking OS to poll specific NTP time-serving hosts for the current time. From those time-serving hosts, the system chooses one NTP host with which to synchronize and serve as a client to the NTP host. As soon as a host-client relationship is established, the networking device propagates the time information throughout its local network.

Protocol Overview

The NTP messages to one or more servers and processes the replies as received. The server interchanges addresses and ports, fills in or overwrites certain fields in the message, recalculates the checksum, and returns it immediately.

Information included in the NTP message allows each client/server peer to determine the timekeeping characteristics of its other peers, including the expected accuracies of their clocks. Using this information, each peer is able to select the best time from possibly several other clocks, update the local clock, and estimate its accuracy.

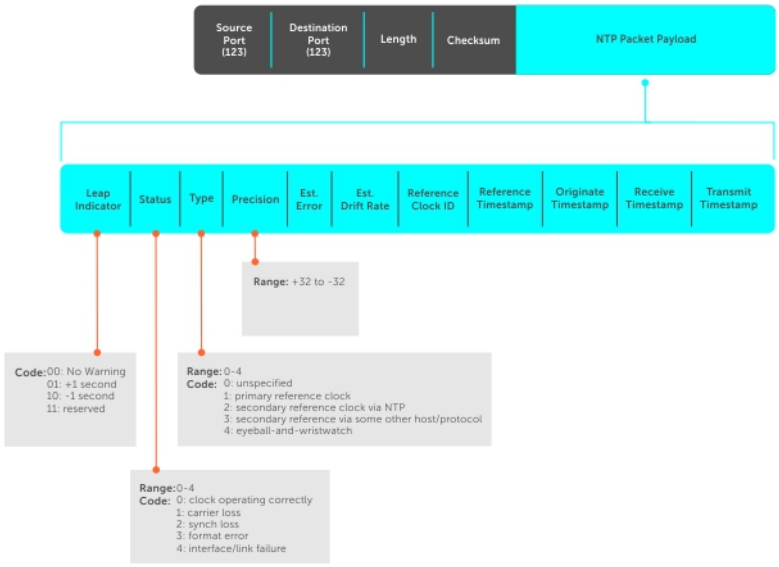


Figure 134. NTP Fields

Implementation Information

Dell EMC Networking systems can only be an NTP client.

Configure the Network Time Protocol

Configuring NTP is a one-step process.

- Enabling NTP

Related Configuration Tasks

- Configuring NTP Broadcasts
- Disabling NTP on an Interface
- Configuring a Source IP Address for NTP Packets (optional)

Enabling NTP

NTP is disabled by default.

To enable NTP, specify an NTP server to which the Dell EMC Networking system synchronizes. To specify multiple servers, enter the command multiple times. You may specify an unlimited number of servers at the expense of CPU resources.

- Specify the NTP server to which the Dell EMC Networking system synchronizes.

  CONFIGURATION mode

  ```
  ntp server ip-address
  ```

To display the system clock state with respect to NTP, use the `show ntp status` command from EXEC Privilege mode.

```
DellEMC#show ntp status
Clock is synchronized, stratum 4, reference is 10.16.151.117, vrf-id is 0
frequency is -44.862 ppm, stability is 0.050 ppm, precision is -18
reference time  deeef7ef.85eeaa10  Tue, Jul 10 2018  9:16:31.523 UTC
clock offset is -0.167449 msec, root delay is 149.194 msec
root dispersion is 54.557 msec, peer dispersion is 0.782 sec
peer mode is client
DellEMC#
```

To display the calculated NTP synchronization variables received from the server that the system uses to synchronize its clock, use the `show ntp associations` command from EXEC Privilege mode.

```
DellEMC#show ntp associations
   remote       vrf-Id     ref clock     st when poll reach   delay   offset    disp
==============================================================================
*10.16.151.117   0      45.127.112.2     3   10   16  377    0.46400  -7.4879  3.19999
* master (synced), # backup, + selected, - outlier, x falseticker
DellEMC#
```

Configuring NTP Broadcasts

With Dell EMC Networking OS, you can receive broadcasts of time information.

You can set interfaces within the system to receive NTP information through broadcast.

To configure an interface to receive NTP broadcasts, use the following commands.

- Set the interface to receive NTP packets.

  INTERFACE mode

  `ntp broadcast client`

```
2w1d11h : NTP: Maximum Slew:-0.000470, Remainder = -0.496884
```

Disabling NTP on an Interface

By default, NTP is enabled on all active interfaces. If you disable NTP on an interface, Dell EMC Networking OS drops any NTP packets sent to that interface.

To disable NTP on an interface, use the following command.

- Disable NTP on the interface.

  INTERFACE mode

  `ntp disable`

To view whether NTP is configured on the interface, use the `show config` command in INTERFACE mode. If `ntp disable` is not listed in the `show config` command output, NTP is enabled. (The `show config` command displays only non-default configuration information.)

Configuring a Source IP Address for NTP Packets

By default, the source address of NTP packets is the IP address of the interface used to reach the network.

You can configure one interface's IP address include in all NTP packets.

To configure an IP address as the source address of NTP packets, use the following command.

- Configure a source IP address for NTP packets.

  CONFIGURATION mode

  `ntp source` *`interface`*

  Enter the following keywords and slot/port or number information:

  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
- For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
- For the Management interface, enter the keyword `ManagementEthernet` then the slot/port information.
- For a port channel interface, enter the keywords `port-channel` then a number.
- For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

To view the configuration, use the `show running-config ntp` command in EXEC privilege mode (refer to the example in Configuring NTP Authentication).

Configuring NTP Authentication

NTP authentication and the corresponding trusted key provide a reliable means of exchanging NTP packets with trusted time sources.

NTP authentication begins when the first NTP packet is created following the configuration of keys. NTP authentication in Dell EMC Networking OS uses the Message Digest 5 (MD5) or Secure Hash Algorithm 1 (SHA1) algorithm and the key is embedded in the synchronization packet that is sent to an NTP time source.

Dell EMC Networking OS Behavior: Dell EMC Networking OS uses an encryption algorithm to store the authentication key that is different from previous Dell EMC Networking OS versions; Dell EMC Networking OS uses data encryption standard (DES) encryption to store the key in the startup-config when you enter the `ntp authentication-key` command. Therefore, if your system boots with a startup-configuration from an Dell EMC Networking OS version in which you have configured `ntp authentication-key`, the system cannot correctly decrypt the key and cannot authenticate the NTP packets. In this case, re-enter this command and save the running-config to the startup-config.

To configure NTP authentication, use the following commands.

1. Enable NTP authentication.

   CONFIGURATION mode

   ```
   ntp authenticate
   ```

2. Set an authentication key.

   CONFIGURATION mode

   ```
   ntp authentication-key number {md5 | sha1} key
   ```

   Configure the following parameters:

   - *number*: the range is from 1 to 65534. This *number* must be the same as the *number* in the `ntp trusted-key` command.
   - *key*: enter a text string. This text string is encrypted.

3. Define a trusted key.

   CONFIGURATION mode

   ```
   ntp trusted-key number
   ```

   Configure a number from 1 to 65534.

   The *number* must be the same as the *number* used in the `ntp authentication-key` command.

4. Configure an NTP server.

   CONFIGURATION mode

   ```
   ntp server [vrf] <vrf-name> {hostname | ipv4-address |ipv6-address} [ key keyid] [prefer] [version number][minpoll] [maxpoll]
   ```

   Configure the IP address of a server and the following optional parameters:

   - `vrf-name` : Enter the name of the VRF through which the NTP server is reachable.
   - `hostname` : Enter the keyword hostname to see the IP address or host name of the remote device.
   - `ipv4-address` : Enter an IPv4 address in dotted decimal format (A.B.C.D).
   - `ipv6-address` : Enter an IPv6 address in the format 0000:0000:0000:0000:0000:0000:0000:0000. Elision of zeros is supported.
   - `key keyid` : Configure a text string as the key exchanged between the NTP server and the client.
   - `prefer`: Enter the keyword prefer to set this NTP server as the preferred server.
   - `version number` : Enter a number as the NTP version. The range is from 1 to 4.
   - `minpoll polling-interval`: Enter the minpoll value. The range is from 4 to 16.
   - `maxpoll polling-interval`: Enter the maxpoll value. The range is from 4 to 16.

5. Configure the switch as NTP master.

   CONFIGURATION mode

   `ntp master <stratum>`

   To configure the switch as NTP Server use the `ntp master<stratum>` command. `stratum` number identifies the NTP Server's hierarchy.

The following example shows configuring an NTP server.

```
Dell EMC(conf)#show running-config ntp
!
ntp master
ntp server 10.16.127.44
ntp server 10.16.127.86
ntp server 10.16.127.144
Dell EMC (conf)#
```

```
Dell EMC#show ntp associations
   remote       vrf-Id     ref clock     st when poll reach   delay   offset    disp
==============================================================================
 LOCAL(0)         0      .LOCL.           7    7   16    7       0.000    0.000    0.002
 10.16.127.86     0      10.16.127.26     5    3   16    7       0.498  361.760    0.184
 10.16.127.144    0      10.16.127.26     5    1   16    7       0.492  359.171    0.219
 10.16.127.44     0      10.16.127.26     5    5   16    7       0.498  355.501    0.188
* master (synced), # backup, + selected, - outlier, x falseticker
Dell EMC#
```

In the above example, the LOCAL (0) determines the following:

- `LOCAL(0)` indicates that the local machine synchronizes with itself.
- `.LOCL.` indicates reference clock of the NTP master.

NOTE:

- **Leap Indicator** (`sys.leap`, `peer.leap`, `pkt.leap`) — This is a two-bit code warning of an impending leap second to be inserted in the NTP time scale. The bits are set before 23:59 on the day of insertion and reset after 00:00 on the following day. This causes the number of seconds (rollover interval) in the day of insertion to be increased or decreased by one. In the case of primary servers, the bits are set by operator intervention, while in the case of secondary servers, the bits are set by the protocol. The two bits, bit 0, and bit 1, respectively, are coded as follows:
- **Poll Interval** — integer indicating the minimum interval between transmitted messages, in seconds as a power of two. For instance, a value of six indicates a minimum interval of 64 seconds.
- **Precision** — integer indicating the precision of the various clocks, in seconds to the nearest power of two. The value must be rounded to the next larger power of two; for instance, a 50 Hz (20 ms) or 60 Hz (16.67ms) power-frequency clock is assigned the value -5 (31.25 ms), while a 1000 Hz (1 ms) crystal-controlled clock is assigned the value -9 (1.95 ms).
- **Root Delay** (`sys.rootdelay`, `peer.rootdelay`, `pkt.rootdelay`) — a signed fixed-point number indicating the total round-trip delay to the primary reference source at the root of the synchronization subnet, in seconds. This variable can take on both positive and negative values, depending on clock precision and skew.
- **Root Dispersion** (`sys.rootdispersion`, `peer.rootdispersion`, `pkt.rootdispersion`) — a signed fixed-point number indicating the maximum error relative to the primary reference source at the root of the synchronization subnet, in seconds. Only positive values greater than zero are possible.
- **Reference Clock Identifier** (`sys.refid`, `peer.refid`, `pkt.refid`) — This is a 32-bit code identifying the particular reference clock. In the case of stratum 0 (unspecified) or stratum 1 (primary reference source), this is a four-octet, left-justified, zero-padded ASCII string, for example: in the case of stratum 2 and greater (secondary reference) this is the four-octet internet address of the peer selected for synchronization.
- **Reference Timestamp** (`sys.reftime`, `peer.reftime`, `pkt.reftime`) — This is the local time, in timestamp format, when the local clock was last updated. If the local clock has never been synchronized, the value is zero.
- `Originate Timestamp`: The departure time on the server of its last NTP message. If the server becomes unreachable, the value is set to zero.
- **Receive Timestamp** — the arrival time on the client of the last NTP message from the server. If the server becomes unreachable, the value is set to zero.

- **Transmit Timestamp** — the departure time on the server of the current NTP message from the sender.
- **Filter dispersion** — the error in calculating the minimum delay from a set of sample data from a peer.

To view the NTP configuration, use the `show running-config ntp` command in EXEC privilege mode. The following example shows an encrypted authentication key (in bold). All keys are encrypted.

```
DellEMC#show running ntp
!
ntp authenticate
ntp authentication-key 345 md5 5A60910F3D211F02
ntp server 11.1.1.1 version 3
ntp trusted-key 345
DellEMC#
```

Configuring a Custom-defined Period for NTP time Synchronization

You can configure the system to send an audit log message to a syslog server if the time difference from the NTP server is greater than a threshold value (offset-threshold). However, time synchronization still occurs. To configure the offset-threshold, follow this procedure.

- Specify the threshold time interval before which the system generates an NTP audit log message if the system time deviates from the NTP server.

  CONFIGURATION mode

  `ntp offset-threshold` *`threshold-value`*

  The range for *threshold-value* is from 0 to 999.

```
DellEMC(conf)#ntp offset-threshold 9
```

Dell EMC Networking OS Time and Date

You can set the time and date using the Dell EMC Networking OS CLI.

Configuration Task List

The following is a configuration task list for configuring the time and date settings.

- Setting the Time and Date for the Switch Software Clock
- Setting the Timezone
- Setting Daylight Saving Time Once
- Setting Recurring Daylight Saving Time

Setting the Time and Date for the Switch Software Clock

You can change the order of the `month` and `day` parameters to enter the time and date as *time day month year*. You cannot delete the software clock.

The software clock runs only when the software is up. The clock restarts, based on the hardware clock, when the switch reboots.

To set the software clock, use the following command.

- Set the system software clock to the current time and date.

  EXEC Privilege mode

  `clock set` *`time month day year`*

  - *`time`*: enter the time in hours:minutes:seconds. For the hour variable, use the 24-hour format; for example, 17:15:00 is 5:15 pm.

- *month*: enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*.
- *day*: enter the number of the day. The range is from 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.
- *year*: enter a four-digit number as the year. The range is from 1993 to 2035.

```
DellEMC#clock set 16:20:00 9 september 2017
DellEMC#
```

Setting the Timezone

Universal time coordinated (UTC) is the time standard based on the International Atomic Time standard, commonly known as Greenwich Mean time.

When determining system time, include the differentiator between UTC and your local timezone. For example, San Jose, CA is the Pacific Timezone with a UTC offset of -8.

To set the clock timezone, use the following command.

- Set the clock to the appropriate timezone.

  CONFIGURATION mode

  `clock timezone timezone-name offset`

  - *timezone-name*: enter the name of the timezone. Do not use spaces.
  - *offset*: enter one of the following:
    - a number from 1 to 23 as the number of hours in addition to UTC for the timezone.
    - a minus sign (-) then a number from 1 to 23 as the number of hours.

```
DellEMC#conf
DellEMC(conf)#clock timezone Pacific -8
DellEMC(conf)#01:40:19: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Timezone
configuration changed from "UTC 0 hrs 0 mins" to "Pacific -8 hrs 0
mins"
DellEMC#
```

Set Daylight Saving Time

Dell EMC Networking OS supports setting the system to daylight saving time once or on a recurring basis every year.

Setting Daylight Saving Time Once

Set a date (and time zone) on which to convert the switch to daylight saving time on a one-time basis.

To set the clock for daylight savings time once, use the following command.

- Set the clock to the appropriate timezone and daylight saving time.

  CONFIGURATION mode

  `clock summer-time time-zone date start-month start-day start-year start-time end-month end-day end-year end-time [offset]`

  - *time-zone*: enter the three-letter name for the time zone. This name displays in the show clock output.
  - *start-month*: enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*.
  - *start-day*: enter the number of the day. The range is from 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.
  - *start-year*: enter a four-digit number as the year. The range is from 1993 to 2035.
  - *start-time*: enter the time in hours:minutes. For the hour variable, use the 24-hour format; example, 17:15 is 5:15 pm.
  - *end-month*: enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*.

- *end-day*: enter the number of the day. The range is from 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.
- *end-year*: enter a four-digit number as the year. The range is from 1993 to 2035.
- *end-time*: enter the time in hours:minutes. For the hour variable, use the 24-hour format; example, 17:15 is 5:15 pm.
- *offset*: (OPTIONAL) enter the number of minutes to add during the summer-time period. The range is from 1 to1440. The default is **60 minutes**.

```
DellEMC(conf)#clock summer-time pacific date Mar 14 2009 00:00 Nov 7 2009 00:00
DellEMC(conf)#02:02:13: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration
changed from
"none" to "Summer time starts 00:00:00 Pacific Sat Mar 14 2009;Summer time ends 00:00:00
pacific
Sat Nov 7 2009"
```

Setting Recurring Daylight Saving Time

Set a date (and time zone) on which to convert the switch to daylight saving time on a specific day every year.

If you have already set daylight saving for a one-time setting, you can set that date and time as the recurring setting with the `clock summer-time time-zone recurring` command.

To set a recurring daylight saving time, use the following command.

- Set the clock to the appropriate timezone and adjust to daylight saving time every year.

  CONFIGURATION mode

  `clock summer-time` *time-zone* `recurring` *start-week start-day start-month start-time end-week end-day end-month end-time* [*offset*]

  - *time-zone*: Enter the three-letter name for the time zone. This name displays in the show clock output.
  - *start-week*: (OPTIONAL) Enter one of the following as the week that daylight saving begins and then enter values for *start-day* through *end-time*:
    - *week-number*: Enter a number from 1 to 4 as the number of the week in the month to start daylight saving time.
    - `first`: Enter the keyword `first` to start daylight saving time in the first week of the month.
    - `last`: Enter the keyword `last` to start daylight saving time in the last week of the month.
  - *start-month*: Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*.
  - *start-day*: Enter the number of the day. The range is from 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.
  - *start-year*: Enter a four-digit number as the year. The range is from 1993 to 2035.
  - *start-time*: Enter the time in hours:minutes. For the hour variable, use the 24-hour format; example, 17:15 is 5:15 pm.
  - *end-week*: If you entered a start-week, enter the one of the following as the week that daylight saving ends:
    - *week-number*: Enter a number from 1 to 4 as the number of the week in the month to start daylight saving time.
    - `first`: Enter the keyword `first` to start daylight saving time in the first week of the month.
    - `last`: Enter the keyword `last` to start daylight saving time in the last week of the month.
  - *end-month*: Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*.
  - *end-day*: Enter the number of the day. The range is from 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.
  - *end-year*: Enter a four-digit number as the year. The range is from 1993 to 2035.
  - *end-time*: Enter the time in hours:minutes. For the hour variable, use the 24-hour format; example, 17:15 is 5:15 pm.
  - *offset*: (OPTIONAL) Enter the number of minutes to add during the summer-time period. The range is from 1 to1440. The default is **60 minutes**.

The following example shows the `clock summer-time recurring` command.

```
DellEMC(conf)#clock summer-time pacific recurring Mar 14 2009 00:00 Nov 7 2009 00:00 ?
DellEMC(conf)#02:02:13: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration
changed from
"none" to "Summer time starts 00:00:00 Pacific Sat Mar 14 2009;Summer time ends 00:00:00
pacific
Sat Nov 7 2009"
```

NOTE: If you enter <CR> after entering the `recurring` command parameter, and you have already set a one-time daylight saving time/date, the system uses that time and date as the recurring setting.

The following example shows the clock summer-time recurring parameters.

```
DellEMC(conf)#clock summer-time pacific recurring ?
<1-4>     Week number to start
first     Week number to start
last      Week number to start
<cr>
DellEMC(conf)#clock summer-time pacific recurring
DellEMC(conf)#02:10:57: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration
changed from
"Summer time starts 00:00:00 Pacific Tue Mar 14 2017 ; Summer time ends 00:00:00 pacific
Tue Nov
7 2017" to "Summer time starts 02:00:00 Pacific Tue Mar 14 2017;Summer time ends
02:00:00 pacific
Tue Nov 7 2017"
```

56

Tunneling

Tunnel interfaces create a logical tunnel for IPv4 or IPv6 traffic. Tunneling supports RFC 2003, RFC 2473, and 4213.

DSCP, hop-limits, flow label values, open shortest path first (OSPF) v2, and OSPFv3 are supported. Internet control message protocol (ICMP) error relay, PATH MTU transmission, and fragmented packets are not supported.

Topics:

- Configuring a Tunnel
- Configuring Tunnel Keepalive Settings
- Configuring a Tunnel Interface
- Configuring Tunnel Allow-Remote Decapsulation
- Configuring Tunnel source anylocal Decapsulation

Configuring a Tunnel

You can configure a tunnel in IPv6 mode, IPv6IP mode, and IPIP mode.

- If the tunnel mode is IPIP or IPv6IP, the tunnel source address and the tunnel destination address must be an IPv4 address.
- If the tunnel mode is IPv6, the tunnel source address and the tunnel destination address must be an IPv6 address.
- If the tunnel mode is IPv6 or IPIP, you can use either an IPv6 address or an IPv4 address for the logical address of the tunnel, but in IPv6IP mode, the logical address must be an IPv6 address.

The following sample configuration shows a tunnel configured in IPv6 mode (carries IPv6 and IPv4 traffic).

```
DellEMC(conf)#interface tunnel 1
DellEMC(conf-if-tu-1)#tunnel source 30.1.1.1
DellEMC(conf-if-tu-1)#tunnel destination 50.1.1.1
DellEMC(conf-if-tu-1)#tunnel mode ipip
DellEMC(conf-if-tu-1)#ip address 1.1.1.1/24
DellEMC(conf-if-tu-1)#ipv6 address 1::1/64
DellEMC(conf-if-tu-1)#no shutdown
DellEMC(conf-if-tu-1)#show config
!
interface Tunnel 1
ip address 1.1.1.1/24
ipv6 address 1::1/64
tunnel destination 50.1.1.1
tunnel source 30.1.1.1
tunnel mode ipip
no shutdown
```

The following sample configuration shows a tunnel configured in IPV6IP mode (IPv4 tunnel carries IPv6 traffic only):

```
DellEMC(conf)#interface tunnel 2
DellEMC(conf-if-tu-2)#tunnel source 60.1.1.1
DellEMC(conf-if-tu-2)#tunnel destination 90.1.1.1
DellEMC(conf-if-tu-2)#tunnel mode ipv6ip
DellEMC(conf-if-tu-2)#ipv6 address 2::1/64
DellEMC(conf-if-tu-2)#no shutdown
DellEMC(conf-if-tu-2)#show config
!
interface Tunnel 2
no ip address
ipv6 address 2::1/64
tunnel destination 90.1.1.1
tunnel source 60.1.1.1
tunnel mode ipv6ip
no shutdown
```

The following sample configuration shows a tunnel configured in IPIP mode (IPv4 tunnel carries IPv4 and IPv6 traffic):

```
DellEMC(conf)#interface tunnel 3
DellEMC(conf-if-tu-3)#tunnel source 5::5
DellEMC(conf-if-tu-3)#tunnel destination 8::9
DellEMC(conf-if-tu-3)#tunnel mode ipv6
DellEMC(conf-if-tu-3)#ip address 3.1.1.1/24
DellEMC(conf-if-tu-3)#ipv6 address 3::1/64
DellEMC(conf-if-tu-3)#no shutdown
DellEMC(conf-if-tu-3)#show config
!
interface Tunnel 3
ip address 3.1.1.1/24
ipv6 address 3::1/64
tunnel destination 8::9
tunnel source 5::5
tunnel mode ipv6
no shutdown
```

Configuring Tunnel Keepalive Settings

You can configure a tunnel keepalive target, keepalive interval, and attempts.

NOTE: By default, the tunnel keepalive is disabled.

The following sample configuration shows how to use the `tunnel keepalive` command.

```
DellEMC(conf-if-te-1/12)#show config
!
interface TenGigabitEthernet 1/12
ip address 40.1.1.1/24
ipv6 address 500:10::1/64
no shutdown
DellEMC(conf-if-te-1/12)#
DellEMC(conf)#interface tunnel 1
DellEMC(conf-if-tu-1)#ipv6 address 1abd::1/64
DellEMC(conf-if-tu-1)#ip address 1.1.1.1/24
DellEMC(conf-if-tu-1)#tunnel source 40.1.1.1
DellEMC(conf-if-tu-1)#tunnel destination 40.1.1.2
DellEMC(conf-if-tu-1)#tunnel mode ipip
DellEMC(conf-if-tu-1)#no shutdown
DellEMC(conf-if-tu-1)#tunnel keepalive 1.1.1.2 attempts 4 interval 6
DellEMC(conf-if-tu-1)#show config
!
interface Tunnel 1
ip address 1.1.1.1/24
ipv6 address 1abd::1/64
tunnel destination 40.1.1.2
tunnel source 40.1.1.1
tunnel keepalive 1.1.1.2 attempts 4 interval 6
tunnel mode ipip
no shutdown
```

Configuring a Tunnel Interface

You can configure the tunnel interface using the `ip unnumbered` and `ipv6 unnumbered` commands.

To configure the tunnel interface to operate without a unique explicit IP or IPv6 address, select the interface from which the tunnel borrows its address.

The following sample configuration shows how to use the `interface tunnel` configuration commands.

```
DellEMC(conf-if-te-1/1)#show config
 !
interface TenGigabitEthernet 1/1
ip address 20.1.1.1/24
ipv6 address 20:1::1/64
```

```
no shutdown
DellEMC(conf)#interface tunnel 1
DellEMC(conf-if-tu-1)#ip unnumbered tengigabitethernet 1/1
DellEMC(conf-if-tu-1)#ipv6 unnumbered tengigabitethernet 1/1
DellEMC(conf-if-tu-1)#tunnel source 40.1.1.1
DellEMC(conf-if-tu-1)#tunnel mode ipip decapsulate-any
DellEMC(conf-if-tu-1)#no shutdown
DellEMC(conf-if-tu-1)#show config
!
interface Tunnel 1
ip unnumbered TenGigabitEthernet 1/1
ipv6 unnumbered TenGigabitEthernet 1/1
tunnel source 40.1.1.1
tunnel mode ipip decapsulate-any
no shutdown
DellEMC(conf-if-tu-1)#
```

Configuring Tunnel Allow-Remote Decapsulation

You can configure an IPv4 or IPV6 address or prefix whose tunneled packet is accepted for decapsulation.

- If you do not configure allow-remote entries, tunneled packets from any remote peer address are accepted.
- You can configure up to eight allow-remote entries on any particular multipoint receive-only tunnel.

The following sample configuration shows how to configure a tunnel allow-remote address.

```
DellEMC(conf)#interface tunnel 1
DellEMC(conf-if-tu-1)#ipv6 address 1abd::1/64
DellEMC(conf-if-tu-1)#ip address 1.1.1.1/24
DellEMC(conf-if-tu-1)#tunnel source 40.1.1.1
DellEMC(conf-if-tu-1)#tunnel mode ipip decapsulate-any
DellEMC(conf-if-tu-1)#tunnel allow-remote 40.1.1.2
DellEMC(conf-if-tu-1)#no shutdown
DellEMC(conf-if-tu-1)#show config
!
interface Tunnel 1
ip address 1.1.1.1/24
ipv6 address 1abd::1/64
tunnel source 40.1.1.1
tunnel allow-remote 40.1.1.2
tunnel mode ipip decapsulate-any
no shutdown
```

Configuring Tunnel source anylocal Decapsulation

The `tunnel source anylocal` command allows a multipoint receive-only tunnel to decapsulate tunnel packets addressed to any IPv4 or IPv6 (depending on the tunnel mode) address configured on the switch that is operationally UP.

The `source anylocal` parameters can be used for packet decapsulation instead of the ip address or interface (`tunnel allow-remote` command), but only on multipoint receive-only mode tunnels.

The following sample configuration shows how to use the `tunnel source anylocal` command.

```
DellEMC(conf)#interface tunnel 1
DellEMC(conf-if-tu-1)#ipv6 address 1abd::1/64
DellEMC(conf-if-tu-1)#ip address 1.1.1.1/24
DellEMC(conf-if-tu-1)#tunnel source anylocal
DellEMC(conf-if-tu-1)#tunnel mode ipip decapsulate-any
DellEMC(conf-if-tu-1)#tunnel allow-remote 40.1.1.2
DellEMC(conf-if-tu-1)#no shutdown
DellEMC(conf-if-tu-1)#show config
!
interface Tunnel 1
ip address 1.1.1.1/24
ipv6 address 1abd::1/64
tunnel source anylocal
```

```
tunnel allow-remote 40.1.1.2
tunnel mode ipip decapsulate-any
no shutdown
```

57

Uplink Failure Detection (UFD)

Uplink failure detection (UFD) provides detection of the loss of upstream connectivity and, if used with network interface controller (NIC) teaming, automatic recovery from a failed link.

Topics:

- Feature Description
- How Uplink Failure Detection Works
- UFD and NIC Teaming
- Important Points to Remember
- Configuring Uplink Failure Detection
- Clearing a UFD-Disabled Interface
- Displaying Uplink Failure Detection
- Sample Configuration: Uplink Failure Detection

Feature Description

A switch provides upstream connectivity for devices, such as servers. If a switch loses its upstream connectivity, downstream devices also lose their connectivity. However, the devices do not receive a direct indication that upstream connectivity is lost because connectivity to the switch is still operational.

UFD allows a switch to associate downstream interfaces with upstream interfaces. When upstream connectivity fails, the switch disables the downstream links. Failures on the downstream links allow downstream devices to recognize the loss of upstream connectivity.

For example, as shown in the following illustration, Switches S1 and S2 both have upstream connectivity to Router R1 and downstream connectivity to the server. UFD operation is shown in Steps A through C:

- In Step A, the server configuration uses the connection to S1 as the primary path. Network traffic flows from the server to S1 and then upstream to R1.
- In Step B, the upstream link between S1 and R1 fails. The server continues to use the link to S1 for its network traffic, but the traffic is not successfully switched through S1 because the upstream link is down.
- In Step C, UFD on S1 disables the link to the server. The server then stops using the link to S1 and switches to using its link to S2 to send traffic upstream to R1.

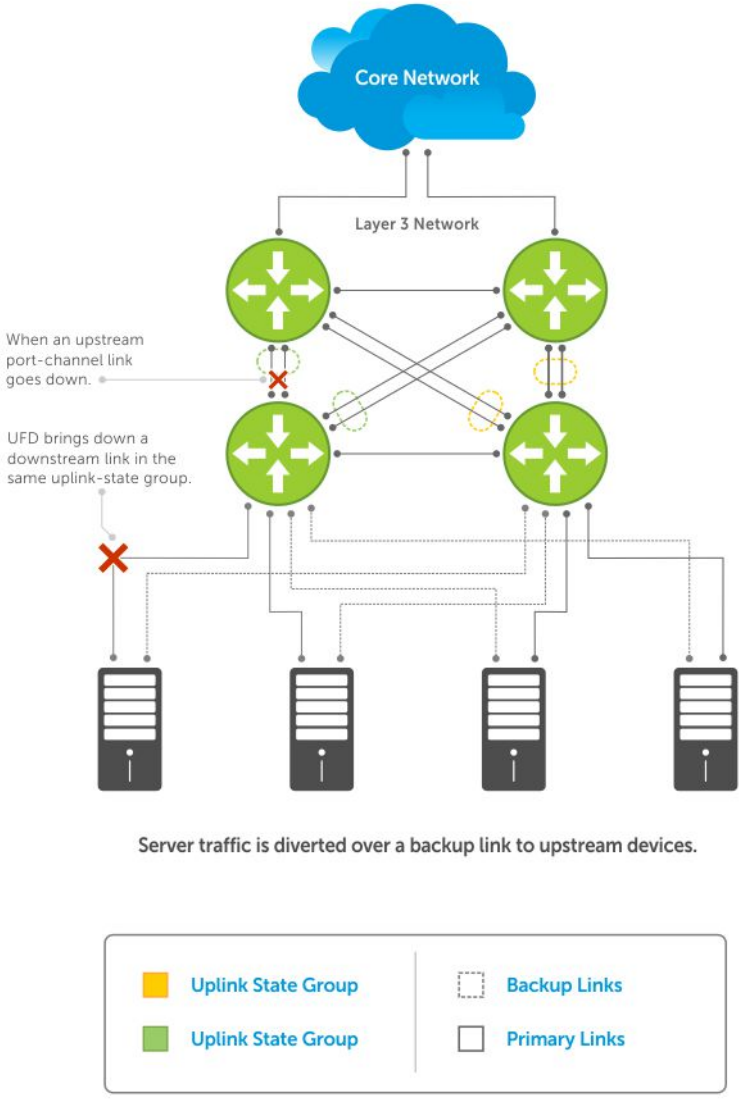


Figure 135. Uplink Failure Detection

How Uplink Failure Detection Works

UFD creates an association between upstream and downstream interfaces. The association of uplink and downlink interfaces is called an *uplink-state group*.

An interface in an uplink-state group can be a physical interface or a port-channel (LAG) aggregation of physical interfaces.

An enabled uplink-state group tracks the state of all assigned upstream interfaces. Failure on an upstream interface results in the automatic disabling of downstream interfaces in the uplink-state group. As a result, downstream devices can execute the protection or recovery procedures they have in place to establish alternate connectivity paths, as shown in the following illustration.

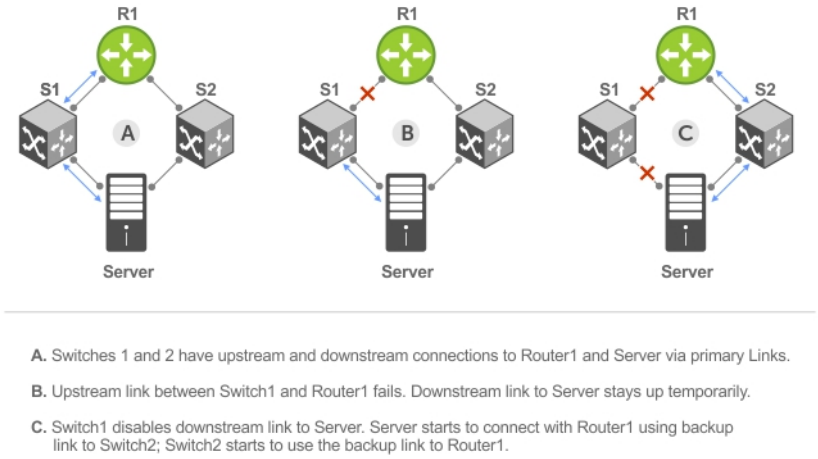


Figure 136. Uplink Failure Detection Example

If only one of the upstream interfaces in an uplink-state group goes down, a specified number of downstream ports associated with the upstream interface are put into a Link-Down state. You can configure this number and is calculated by the ratio of the upstream port bandwidth to the downstream port bandwidth in the same uplink-state group. This calculation ensures that there is no traffic drops due to insufficient bandwidth on the upstream links to the routers/switches.

By default, if all upstream interfaces in an uplink-state group go down, all downstream interfaces in the same uplink-state group are put into a Link-Down state.

Using UFD, you can configure the automatic recovery of downstream ports in an uplink-state group when the link status of an upstream port changes. The tracking of upstream link status does not have a major impact on central processing unit (CPU) usage.

UFD and NIC Teaming

To implement a rapid failover solution, you can use uplink failure detection on a switch with network adapter teaming on a server.

For more information, refer to NIC Teaming.

For example, as shown previously, the switch/ router with UFD detects the uplink failure and automatically disables the associated downstream link port to the server. To continue to transmit traffic upstream, the server with NIC teaming detects the disabled link and automatically switches over to the backup link in order.

Important Points to Remember

When you configure UFD, the following conditions apply.

- You can configure up to 16 uplink-state groups. By default, no uplink-state groups are created.
  - An uplink-state group is considered to be operationally *up* if it has at least one upstream interface in the Link-Up state.
  - An uplink-state group is considered to be operationally *down* if it has no upstream interfaces in the Link-Up state. No uplink-state tracking is performed when a group is disabled or in an Operationally Down state.
- You can assign a physical port or port-channel interfaces to an uplink-state group.
  - You can assign an interface to only one uplink-state group. Configure each interface assigned to an uplink-state group as either an upstream or downstream interface, but not both.
  - You can assign individual member ports of a port channel to the group. An uplink-state group can contain either the member ports of a port channel or the port channel itself, but not both.
  - If you assign a port channel as an upstream interface, the port channel interface enters a Link-Down state when the number of port-channel member interfaces in a Link-Up state drops below the configured `minimum number of members` parameter.

- If one of the upstream interfaces in an uplink-state group goes down, either a user-configurable set of downstream ports or all the downstream ports in the group are put in an Operationally Down state with an UFD Disabled error. The order in which downstream ports are disabled is from the lowest numbered port to the highest.
  - If one of the upstream interfaces in an uplink-state group that was down comes up, the set of UFD-disabled downstream ports (which were previously disabled due to this upstream port going down) is brought up and the UFD Disabled error is cleared.
- If you disable an uplink-state group, the downstream interfaces are not disabled regardless of the state of the upstream interfaces.
  - If an uplink-state group has no upstream interfaces assigned, you cannot disable downstream interfaces when an upstream link goes down.
- To enable the debug messages for events related to a specified uplink-state group or all groups, use the `debug uplink-state-group [group-id]` command, where the group-id is from 1 to 16.
  - To turn off debugging event messages, use the `no debug uplink-state-group [group-id]` command.
  - For an example of debug log message, refer to Clearing a UFD-Disabled Interface.

Configuring Uplink Failure Detection

To configure UFD, use the following commands.

1. Create an uplink-state group and enable the tracking of upstream links on the switch/router.

   CONFIGURATION mode

   `uplink-state-group group-id`

   - *group-id*: values are from 1 to 16.

   To delete an uplink-state group, use the `no uplink-state-group group-id` command.
2. Assign a port or port-channel to the uplink-state group as an upstream or downstream interface.

   UPLINK-STATE-GROUP mode

   `{upstream | downstream} interface`

   For interface, enter one of the following interface types:
   - 10 Gigabit Ethernet: enter `tengigabitethernet {slot/port | slot/port-range}`
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

   Where *port-range* and *port-channel-range* specify a range of ports separated by a dash (-) and/or individual ports/port channels in any order; for example:

   ```
   upstream tengigabitethernet 1/1-2,5,9,11-12
   downstream port-channel 1-3,5
   ```

   - A comma is required to separate each port and port-range entry.

   To delete an interface from the group, use the `no {upstream | downstream} interface` command.
3. Configure the number of downstream links in the uplink-state group that will be disabled (Oper Down state) if one upstream link in the group goes down.

   UPLINK-STATE-GROUP mode

   `downstream disable links {number | all}`

   - *number*: specifies the number of downstream links to be brought down. The range is from 1 to 1024.
   - `all`: brings down all downstream links in the group.

   The default is no downstream links are disabled when an upstream link goes down.

   NOTE: Downstream interfaces in an uplink-state group are put into a Link-Down state with an UFD-Disabled error message only when all upstream interfaces in the group go down.

   To revert to the default setting, use the `no downstream disable links` command.
4. (Optional) Enable auto-recovery so that UFD-disabled downstream ports in the uplink-state group come up when a disabled upstream port in the group comes back up.

   UPLINK-STATE-GROUP mode

   `downstream auto-recover`

The default is auto-recovery of UFD-disabled downstream ports is enabled.

To disable auto-recovery, use the `no downstream auto-recover` command.

5. (Optional) Enter a text description of the uplink-state group.

   UPLINK-STATE-GROUP mode

   `description text`

   The maximum length is 80 alphanumeric characters.

6. (Optional) Disable upstream-link tracking without deleting the uplink-state group.

   UPLINK-STATE-GROUP mode

   `no enable`

   The default is upstream-link tracking is automatically enabled in an uplink-state group.

   To re-enable upstream-link tracking, use the `enable` command.

Clearing a UFD-Disabled Interface

You can manually bring up a downstream interface in an uplink-state group that UFD disabled and is in a UFD-Disabled Error state.

To re-enable one or more disabled downstream interfaces and clear the UFD-Disabled Error state, use the following command.

- Re-enable a downstream interface on the switch/router that is in a UFD-Disabled Error State so that it can send and receive traffic.

  EXEC mode

  `clear ufd-disable {interface interface | uplink-state-group group-id}`

  For *interface*, enter one of the following interface types:

  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a port channel interface, enter `port-channel {1-512 | port-channel-range}`
    - Where *port-range* and *port-channel-range* specify a range of ports separated by a dash (-) and/or individual ports/port channels in any order; for example:

      ```
      tengigabitethernet 1/1-2,5,9,11-12
      port-channel 1-3,5
      ```

    - A comma is required to separate each port and port-range entry.

  `clear ufd-disable {interface interface | uplink-state-group group-id}`: re-enables all UFD-disabled downstream interfaces in the group. The range is from 1 to 16.

The following example message shows the Syslog messages that display when you clear the UFD-Disabled state from all disabled downstream interfaces in an uplink-state group by using the `clear ufd-disable uplink-state-group group-id` command. All downstream interfaces return to an operationally up state.

```
02:36:43: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 1/6
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 1/6
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 3/48
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 3/52
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 3/56
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Fo 3/48
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Fo 3/52
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Fo 3/56

02:37:29: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 1/7
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 1/7
02:37:29 : UFD: Group:3, UplinkState: DOWN

02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed uplink state group state to down: Group 3
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 3/60
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Fo 3/60

02:38:31 : UFD: Group:3, UplinkState: UP
02:38:31: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed uplink state group state to up: Group 3
```

```
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled:
Fo 3/48
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled:
Fo 3/52
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled:
Fo 3/56
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled:
Fo 3/60
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 3/48
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 3/52
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 3/56
  02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 3/60
```

Displaying Uplink Failure Detection

To display information on the UFD feature, use any of the following commands.

- Display status information on a specified uplink-state group or all groups.

  EXEC mode

  `show uplink-state-group [group-id] [detail]`

  - *group-id*: The values are from 1 to 16.
  - `detail`: displays additional status information on the upstream and downstream interfaces in each group.
- Display the current status of a port or port-channel interface assigned to an uplink-state group.

  EXEC mode

  `show interfaces interface`

  interface specifies one of the following interface types:
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
  - For a port channel interface, enter the keywords `port-channel` then a number.

  If a downstream interface in an uplink-state group is disabled (Oper Down state) by uplink-state tracking because an upstream port is down, the message error-disabled[UFD] displays in the output.
- Display the current configuration of all uplink-state groups or a specified group.

  EXEC mode or UPLINK-STATE-GROUP mode

  (For EXEC mode) `show running-config uplink-state-group [group-id]`

  (For UPLINK-STATE-GROUP mode) `show configuration`
  - *group-id*: The values are from 1 to 16.

The following example shows viewing the uplink state group status.

```
DellEMC# show uplink-state-group

Uplink State Group: 1  Status: Enabled, Up
Uplink State Group: 3  Status: Enabled, Up
Uplink State Group: 5  Status: Enabled, Down
Uplink State Group: 6  Status: Enabled, Up
Uplink State Group: 7  Status: Enabled, Up
Uplink State Group: 16 Status: Disabled, Up

DellEMC# show uplink-state-group 16
Uplink State Group: 16 Status: Disabled, Up

DellEMC#show uplink-state-group detail
(Up): Interface up (Dwn): Interface down (Dis): Interface disabled

Uplink State Group    : 1    Status: Enabled, Up
Upstream Interfaces   :
Downstream Interfaces :

Uplink State Group    : 3    Status: Enabled, Up
Upstream Interfaces   : Te 1/6(Up) Te 1/7(Up)
Downstream Interfaces : Te 3/1(Up) Te 3/3(Up) Te 3/5(Up) Te 3/6(Up)
```

```
Uplink State Group    : 5    Status: Enabled, Down
Upstream Interfaces   : Te 1/1(Dwn) Te 1/3(Dwn) Te 1/5(Dwn)
Downstream Interfaces : Te 3/2(Dis) Te 3/4(Dis) Te 3/11(Dis) Te 3/12(Dis) Te 3/13(Dis)
Te 3/14(Dis) Te 3/15(Dis)

Uplink State Group    : 6    Status: Enabled, Up
Upstream Interfaces   :
Downstream Interfaces :

Uplink State Group    : 7    Status: Enabled, Up
Upstream Interfaces   :
Downstream Interfaces :

Uplink State Group    : 16   Status: Disabled, Up
Upstream Interfaces   : Te 1/4(Dwn) Po 8(Dwn)
Downstream Interfaces : Te 1/10(Dwn)
```

The following example shows viewing the interface status with UFD information.

```
DellEMC#show interfaces tengigabitethernet 1/15
TenGigabitEthernet 1/15 is up, line protocol is down (error-disabled[UFD])
Hardware is Force10Eth, address is 00:01:e8:32:7a:47
    Current address is 00:01:e8:32:7a:47
Interface index is 280544512
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode auto
Flowcontrol rx off tx off
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:25:46
Queueing strategy: fifo
Input Statistics:
    0 packets, 0 bytes
    0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    0 Multicasts, 0 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics:
    0 packets, 0 bytes, 0 underruns
    0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    0 Multicasts, 0 Broadcasts, 0 Unicasts
    0 throttles, 0 discarded, 0 collisions
Rate info (interval 299 seconds):
    Input 00.00 Mbits/sec,    0 packets/sec, 0.00% of line-rate
    Output 00.00 Mbits/sec,   0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:01:23
```

The following example shows viewing the UFD configuration.

```
DellEMC#show running-config uplink-state-group
!
no enable
uplink state track 1
downstream TenGigabitEthernet 1/2, 4, 6, 11-19
upstream TengigabitEthernet 1/8, 12
upstream PortChannel 1
!
uplink state track 2
downstream TenGigabitEthernet 1/1, 3, 5, 7-10
upstream TengigabitEthernet 1/16, 20
```

```
DellEMC(conf-uplink-state-group-16)# show configuration
!
uplink-state-group 16
no enable
description test
downstream disable links all
downstream TenGigabitEthernet 1/21
```

```
upstream TenGigabitEthernet 1/22
upstream Port-channel 8
```

Sample Configuration: Uplink Failure Detection

The following example shows a sample configuration of UFD on a switch/router in which you configure as follows.

- Configure uplink-state group 3.
- Add downstream links Tengigabitethernet 1/1, 1/2, 1/5, 1/9, 1/11, and 1/12.
- Configure two downstream links to be disabled if an upstream link fails.
- Add upstream links Tengigabitethernet 1/3 and 1/4.
- Add a text description for the group.
- Verify the configuration with various `show` commands.

Example of Configuring UFD

```
DellEMC(conf)# uplink-state-group 3
00:08:11: %STKUNIT0-M:CP %IFMGR-5-ASTATE_UP: Changed uplink state group Admin state to
up:
Group 3
DellEMC(conf-uplink-state-group-3)# downstream tengigabitethernet 1/1-2,5,9,11-12
DellEMC(conf-uplink-state-group-3)# downstream disable links 2
DellEMC(conf-uplink-state-group-3)# upstream tengigabitethernet 1/3-4
00:10:00: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-
disabled:
Te 1/1
DellEMC#
00:10:00: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 1/1
DellEMC(conf-uplink-state-group-3)# description Testing UFD feature

DellEMC(conf-uplink-state-group-3)# show config
!
uplink-state-group 3
description Testing UFD feature
downstream disable links 2
downstream TenGigabitEthernet 1/1-2,5,9,11-12
upstream TenGigabitEthernet 1/3-4
DellEMC(conf-uplink-state-group-3)#
DellEMC(conf-uplink-state-group-3)#exit
DellEMC(conf)#exit
DellEMC#
00:13:06: %STKUNIT0-M:CP %SYS-5-CONFIG_I: Configured from console by console

DellEMC# show running-config uplink-state-group
!
uplink-state-group 3
description Testing UFD feature
downstream disable links 2
downstream TenGigabitEthernet 1/1-2,5,9,11-12
upstream TenGigabitEthernet 1/3-4

DellEMC# show uplink-state-group 3

Uplink State Group: 3 Status: Enabled, Up

DellEMC# show uplink-state-group detail

(Up): Interface up (Dwn): Interface down (Dis): Interface disabled
Uplink State Group    : 3 Status: Enabled, Up
Upstream Interfaces   : Te 1/3(Up) Te 1/4(Dwn)
Downstream Interfaces : Te 1/1(Dis) Te 1/2(Dwn) Te 1/5(Dwn) Te 1/9(Dwn) Te 1/11(Dwn)
Te 1/12(Dwn)
```

58

Upgrade Procedures

To find the upgrade procedures, go to the *Dell EMC Networking OS Release Notes* for your system type to see all the requirements needed to upgrade to the desired Dell EMC Networking OS version. To upgrade your system type, follow the procedures in the *Dell EMC Networking OS Release Notes*.

You can download the release notes of your platform at https://www.force10networks.com. Use your login ID to log in to the website.

Get Help with Upgrades

Direct any questions or concerns about the Dell EMC Networking OS upgrade procedures to the Dell Technical Support Center. You can reach Technical Support:

- On the web: https://www.dell.com/support
- By email: Dell-Force10_Technical_Support@Dell.com
- By phone: US and Canada: 866.965.5800, International: 408.965.5800.

59

Virtual LANs (VLANs)

Virtual LANs (VLANs) are a logical broadcast domain or logical grouping of interfaces in a local area network (LAN) in which all data received is kept locally and broadcast to all members of the group.

When in Layer 2 mode, VLANs move traffic at wire speed and can span multiple devices. The system supports up to 4093 port-based VLANs and one default VLAN, as specified in IEEE 802.1Q.

VLANs benefits include:

- Improved security because you can isolate groups of users into different VLANs
- Ability to create one VLAN across multiple devices

For more information about VLANs, refer to the *IEEE Standard 802.1Q Virtual Bridged Local Area Networks*. In this guide, also refer to:

- Bulk Configuration in the Interfaces chapter.
- VLAN Stacking in the Service Provider Bridging chapter.

For a complete listing of all commands related to Dell EMC Networking OS VLANs, refer to these *Dell EMC Networking OS Command Reference Guide* chapters:

- Interfaces
- 802.1X
- GARP VLAN Registration Protocol (GVRP)
- Service Provider Bridging
- Per-VLAN Spanning Tree Plus (PVST+)

The following table lists the defaults for VLANs in Dell EMC Networking OS.

Feature	Default
Spanning Tree group ID	All VLANs are part of Spanning Tree group 0.
Mode	Layer 2 (no IP address is assigned).
Default VLAN ID	VLAN 1

Topics:

- Default VLAN
- Port-Based VLANs
- VLANs and Port Tagging
- Configuration Task List
- Configuring Native VLANs
- Enabling Null VLAN as the Default VLAN

Default VLAN

When you configure interfaces for Layer 2 mode, they are automatically placed in the Default VLAN as untagged interfaces. Only untagged interfaces can belong to the Default VLAN.

The following example displays the outcome of placing an interface in Layer 2 mode. To configure an interface for Layer 2 mode, use the `switchport` command. As shown in bold, the `switchport` command places the interface in Layer 2 mode and the `show vlan` command in EXEC privilege mode indicates that the interface is now part of the Default VLAN (VLAN 1).

By default, VLAN 1 is the Default VLAN. To change that designation, use the `default vlan-id` command in CONFIGURATION mode. You cannot delete the Default VLAN.

NOTE: You cannot assign an IP address to the Default VLAN. To assign an IP address to a VLAN that is currently the Default VLAN, create another VLAN and assign it to be the Default VLAN. For more information about assigning IP addresses, refer to Assigning an IP Address to a VLAN.

- Untagged interfaces must be part of a VLAN. To remove an untagged interface from the Default VLAN, create another VLAN and place the interface into that VLAN. Alternatively, use the `no switchport` command, and Dell EMC Networking OS removes the interface from the Default VLAN.
- A tagged interface requires an additional step to remove it from Layer 2 mode. Because tagged interfaces can belong to multiple VLANs, remove the tagged interface from all VLANs using the `no tagged interface` command. Only after the interface is untagged and a member of the Default VLAN can you use the `no switchport` command to remove the interface from Layer 2 mode. For more information, refer to VLANs and Port Tagging.

Example of Configuring an Interface for Layer 2 Belonging to the Default VLAN

```
DellEMC(conf)#interface tengigabitethernet 1/2
DellEMC(conf-if)#no shut
DellEMC(conf-if)#switchport
DellEMC(conf-if)#show config
!
interface TenGigabitEthernet 1/2
  no ip address
  switchport
  no shutdown
DellEMC(conf-if)#end
DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs
  NUM  Status  Q Ports
* 1    Active  U Te 1/2
                T Te 1/1
```

Port-Based VLANs

Port-based VLANs are a broadcast domain defined by different ports or interfaces. In Dell EMC Networking OS, a port-based VLAN can contain interfaces from different line cards within the chassis. Dell EMC Networking OS supports 4094 port-based VLANs.

Port-based VLANs offer increased security for traffic, conserve bandwidth, and allow switch segmentation. Interfaces in different VLANs do not communicate with each other, adding some security to the traffic on those interfaces. Different VLANs can communicate between each other by means of IP routing. Because traffic is only broadcast or flooded to the interfaces within a VLAN, the VLAN conserves bandwidth. Finally, you can have multiple VLANs configured on one switch, thus segmenting the device.

Interfaces within a port-based VLAN must be in Layer 2 mode and can be tagged or untagged in the VLAN ID.

VLANs and Port Tagging

To add an interface to a VLAN, the interface must be in Layer 2 mode. After you place an interface in Layer 2 mode, the interface is automatically placed in the Default VLAN.

supports IEEE 802.1Q tagging at the interface level to filter traffic. When you enable tagging, a tag header is added to the frame after the destination and source MAC addresses. That information is preserved as the frame moves through the network. The following example shows the structure of a frame with a tag header. The VLAN ID is inserted in the tag header.

Figure 137. Tagged Frame Format

The tag header contains some key information that uses:

- The VLAN protocol identifier identifies the frame as tagged according to the IEEE 802.1Q specifications (2 bytes).
- Tag control information (TCI) includes the VLAN ID (2 bytes total). The VLAN ID can have 4,096 values, but two are reserved.

NOTE: The insertion of the tag header into the Ethernet frame increases the size of the frame to more than the 1,518 bytes as specified in the IEEE 802.3 standard. Some devices that are not compliant with IEEE 802.3 may not support the larger frame size.

Information contained in the tag header allows the system to prioritize traffic and to forward information to ports associated with a specific VLAN ID. Tagged interfaces can belong to multiple VLANs, while untagged interfaces can belong only to one VLAN.

Configuration Task List

This section contains the following VLAN configuration tasks.

- Creating a Port-Based VLAN (mandatory)
- Assigning Interfaces to a VLAN (optional)
- Assigning an IP Address to a VLAN (optional)
- Enabling Null VLAN as the Default VLAN

Creating a Port-Based VLAN

To configure a port-based VLAN, create the VLAN and then add physical interfaces or port channel (LAG) interfaces to the VLAN.

NOTE: The Default VLAN (VLAN 1) is part of the system startup configuration and does not require configuration.

A VLAN is active only if the VLAN contains interfaces and those interfaces are operationally up. As shown in the following example, VLAN 1 is inactive because it does not contain any interfaces. The other VLANs contain enabled interfaces and are active.

NOTE: In a VLAN, the `shutdown` command stops Layer 3 (routed) traffic only. Layer 2 traffic continues to pass through the VLAN. If the VLAN is not a routed VLAN (that is, configured with an IP address), the `shutdown` command has no affect on VLAN traffic.

When you delete a VLAN (using the `no interface vlan` *`vlan-id`* command), any interfaces assigned to that VLAN are assigned to the Default VLAN as untagged interfaces.

To create a port-based VLAN, use the following command.

- Configure a port-based VLAN (if the VLAN-ID is different from the Default VLAN ID) and enter INTERFACE VLAN mode.

  CONFIGURATION mode

  `interface vlan` *`vlan-id`*

  To activate the VLAN, after you create a VLAN, assign interfaces in Layer 2 mode to the VLAN.

To view the configured VLANs, use the `show vlan` command in EXEC Privilege mode.

```
DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM  Status   Q  Ports
* 1    Inactive U  So 9/4-11
  2    Active   U  Te 1/1,18
  3    Active   U  Te 1/2,19
  4    Active   T  Te 1/3,20
  5    Active   U  Po 1
  6    Active   U  Te 1/12
                U  So 9/0
```

Assigning Interfaces to a VLAN

You can only assign interfaces in Layer 2 mode to a VLAN using the tagged and untagged commands. To place an interface in Layer 2 mode, use the `switchport` command.

You can further designate these Layer 2 interfaces as tagged or untagged. For more information, see the Interfaces chapter and Configuring Layer 2 (Data Link) Mode. When you place an interface in Layer 2 mode by the `switchport` command, the interface is automatically designated untagged and placed in the Default VLAN.

To view which interfaces are tagged or untagged and to which VLAN they belong, use the `show vlan` command. The following example shows that six VLANs are configured, and two interfaces are assigned to VLAN 2. The Q column in the `show vlan` command example notes whether the interface is tagged (T) or untagged (U). For more information about this command, refer to the Layer 2 chapter of the *Dell EMC Networking OS Command Reference Guide*.

To tag frames leaving an interface in Layer 2 mode, assign that interface to a port-based VLAN to tag it with that VLAN ID. To tag interfaces, use the following commands.

1. Access INTERFACE VLAN mode of the VLAN to which you want to assign the interface.

   CONFIGURATION mode

   `interface vlan` *`vlan-id`*
2. Enable an interface to include the IEEE 802.1Q tag header.

   INTERFACE mode

   `tagged` *`interface`*

To view just the interfaces that are in Layer 2 mode, use the `show interfaces switchport` command in EXEC Privilege mode or EXEC mode.

The following example shows the steps to add a tagged interface (in this case, port channel 1) to VLAN 4. To view the interface's status. Interface (po 1) is tagged and in VLAN 2 and 3, use the `show vlan` command. In a port-based VLAN, use the tagged command to add the interface to another VLAN. The `show vlan` command output displays the interface's (po 1) changed status.

Except for hybrid ports, only a tagged interface can be a member of multiple VLANs. You can assign hybrid ports to two VLANs if the port is untagged in one VLAN and tagged in all others.

```
DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM Status   Q  Ports
* 1   Inactive
  2   Active   T  Po1(So 0/0-1)
               T  Te 1/1
  3   Active   T  Po1(So 0/0-1)
               T  Te 1/2

DellEMC#config
DellEMC(conf)#interface vlan 4
DellEMC(conf-if-vlan)#tagged po 1
DellEMC(conf-if-vlan)#show conf
!
interface Vlan 4
  no ip address
  tagged Port-channel 1

DellEMC(conf-if-vlan)#end

DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM Status   Q  Ports
* 1   Inactive
  2   Active   T  Po1(So 0/0-1)
               T  Te 1/1
  3   Active   T  Po1(So 0/0-1)
               T  Te 1/2
  4   Active   T  Po1(So 0/0-1)
```

When you remove a tagged interface from a VLAN (using the `no tagged interface` command), it remains tagged only if it is a tagged interface in another VLAN. If the tagged interface is removed from the only VLAN to which it belongs, the interface is placed in the Default VLAN as an untagged interface.

Moving Untagged Interfaces

To move untagged interfaces from the Default VLAN to another VLAN, use the following commands.

1. Access INTERFACE VLAN mode of the VLAN to which you want to assign the interface.

   CONFIGURATION mode

   `interface vlan` *`vlan-id`*
2. Configure an interface as untagged.

   INTERFACE mode

   `untagged` *`interface`*

   This command is available only in VLAN interfaces.

The `no untagged` *`interface`* command removes the untagged interface from a port-based VLAN and places the interface in the Default VLAN. You cannot use the `no untagged` *`interface`* command in the Default VLAN. The following example shows the steps and commands to move an untagged interface from the Default VLAN to another VLAN.

To determine interface status, use the `show vlan` command. Interface (1/2) is untagged and in the Default VLAN (vlan 1). In a port-based VLAN (vlan 4), use the `untagged` command to add the interface to that VLAN. The `show vlan` command output displays the interface's changed status (1/2). Because the Default VLAN no longer contains any interfaces, it is listed as inactive.

```
DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM  Status  Q  Ports
* 1    Active  U  Te 1/2
  2    Active  T  Po1(So 0/0-1)
               T  Te 1/3
  3    Active  T  Po1(So 0/0-1)
               T  Te 1/1
  4    Inactive
DellEMC#conf
DellEMC(conf)#interface vlan 4
DellEMC(conf-if-vlan)#untagged tengigabitethernet 1/2
DellEMC(conf-if-vlan)#show config
!
interface Vlan 4
  no ip address
  untagged TenGigabitEthernet 1/2
DellEMC(conf-if-vlan)#end

DellEMC#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM  Status   Q  Ports
* 1    Inactive
  2    Active   T  Po1(So 0/0-1)
                T  Te 1/3
  3    Active   T  Po1(So 0/0-1)
                T  Te 1/1
  4    Active   U  Te 1/2
```

The only way to remove an interface from the Default VLAN is to place the interface in Default mode by using the `no switchport` command in INTERFACE mode.

Assigning an IP Address to a VLAN

VLANs are a Layer 2 feature. For two physical interfaces on different VLANs to communicate, you must assign an IP address to the VLANs to route traffic between the two interfaces.

The `shutdown` command in INTERFACE mode does not affect Layer 2 traffic on the interface; the `shutdown` command only prevents Layer 3 traffic from traversing over the interface.

NOTE: You cannot assign an IP address to the Default VLAN (VLAN 1). To assign another VLAN ID to the Default VLAN, use the `default vlan-id vlan-id` command.

In Dell EMC Networking OS, you can place VLANs and other logical interfaces in Layer 3 mode to receive and send routed traffic. For more information, refer to Bulk Configuration.

To assign an IP address, use the following command.

- Configure an IP address and mask on the interface.

  INTERFACE mode

  ```
  ip address ip-address mask [secondary]
  ```

  - *ip-address mask* — Enter an address in dotted-decimal format (A.B.C.D) and the mask must be in slash format (/24).
  - `secondary` — This is the interface's backup IP address. You can configure up to eight secondary IP addresses.

Configuring Native VLANs

Traditionally, ports can be either untagged for membership to one VLAN or tagged for membership to multiple VLANs.

You must connect an untagged port to a VLAN-unaware station (one that does not understand VLAN tags), and you must connect a tagged port to a VLAN-aware station (one that generates and understands VLAN tags).

Native VLAN support breaks this barrier so that you can connect a port to both VLAN-aware and VLAN-unaware stations. Such ports are referred to as hybrid ports. Physical and port-channel interfaces may be hybrid ports.

Native VLAN is useful in deployments where a Layer 2 port can receive both tagged and untagged traffic on the same physical port. The classic example is connecting a voice-over-IP (VOIP) phone and a PC to the same port of the switch. The VOIP phone is configured to generate tagged packets (with VLAN = VOICE VLAN) and the attached PC generates untagged packets.

NOTE: When a hybrid port is untagged in a VLAN but it receives tagged traffic, all traffic is accepted.

NOTE: You cannot configure an existing switchport or port channel interface for Native VLAN. Interfaces must have no other Layer 2 or Layer 3 configurations when using the `portmode hybrid` command or a message similar to this displays: `% Error: Port is in Layer-2 mode Gi 5/6.`

To configure a port so that it can be a member of an untagged and tagged VLANs, use the following commands.

1. Remove any Layer 2 or Layer 3 configurations from the interface.

   INTERFACE mode
2. Configure the interface for Hybrid mode.

   INTERFACE mode

   ```
   portmode hybrid
   ```
3. Configure the interface for Switchport mode.

   INTERFACE mode

   ```
   switchport
   ```
4. Add the interface to a tagged or untagged VLAN.

   VLAN INTERFACE mode

   ```
   [tagged | untagged]
   ```

Enabling Null VLAN as the Default VLAN

In a Carrier Ethernet for Metro Service environment, service providers who perform frequent reconfigurations for customers with changing requirements occasionally enable multiple interfaces, each connected to a different customer, before the interfaces are fully configured.

This presents a vulnerability because both interfaces are initially placed in the native VLAN, VLAN 1, and for that period customers are able to access each other's networks. Dell EMC Networking OS has a Null VLAN to eliminate this vulnerability. When you enable the Null VLAN, all ports are placed into it by default, so even if you activate the physical ports of multiple customers, no traffic is allowed to traverse the links until each port is place in another VLAN.

To enable Null VLAN, use the following command.

- Disable the default VLAN, so that all ports belong to the Null VLAN until configured as a member of another VLAN.

  CONFIGURATION mode

  `default-vlan disable`

  Default: the default VLAN is enabled (`no default-vlan disable`).

60

Virtual Link Trunking (VLT)

Virtual link trunking (VLT) is a Dell EMC technology that provides two Dell EMC switches the ability to function as a single switch. VLT allows physical links between two Dell EMC switches to appear as a single virtual link to the network core or other switches such as Edge, Access, or top-of-rack (ToR). As a result, the two physical switches appear as a single switch to the connected devices.

Topics:

- Overview
- Configure Virtual Link Trunking
- RSTP Configuration
- PVST+ Configuration
- Peer Routing Configuration Example
- eVLT Configuration Example
- PIM-Sparse Mode Configuration Example
- Verifying a VLT Configuration
- Additional VLT Sample Configurations
- Troubleshooting VLT
- Reconfiguring Stacked Switches as VLT
- Specifying VLT Nodes in a PVLAN
- Configuring a VLT VLAN or LAG in a PVLAN
- Proxy ARP Capability on VLT Peer Nodes
- VLT Nodes as Rendezvous Points for Multicast Resiliency
- Configuring VLAN-Stack over VLT
- IPv6 Peer Routing in VLT Domains Overview
- Configure BFD in VLT Domain

Overview

In a traditional switched topology as shown below, spanning tree protocols (STPs) are used to block one or more links to prevent loops in the network. Although loops are prevented, bandwidth of all links is not effectively utilized by the connected devices.

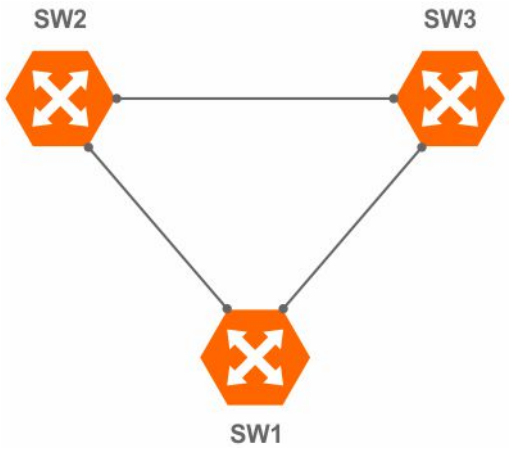


Figure 138. Traditional switched topology

VLT not only overcomes this caveat, but also provides a multipath to the connected devices. In the example shown below, the two physical VLT peers appear as a single logical device to the connected devices. As the connected devices consider the VLT

peers as a single switch, VLT eliminates STP-blocked ports. However, the two VLT devices are independent Layer2/Layer3 (L2/L3) switches for devices in the upstream network.

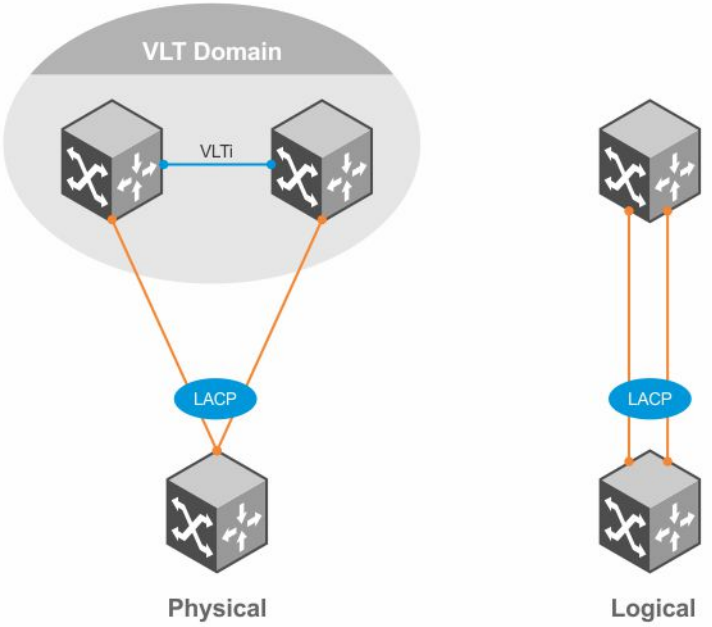


Figure 139. VLT providing multipath

VLT reduces the role of spanning tree protocols (STPs) by allowing link aggregation group (LAG) terminations on two separate distribution or core switches and supporting a loop-free topology.

To prevent the initial loop that may occur prior to VLT being established, use a spanning tree protocol. After VLT is established, you may use rapid spanning tree protocol (RSTP) to prevent loops from forming with new links that are incorrectly connected and outside the VLT domain.

VLT provides Layer 2 multipathing, creating redundancy through increased bandwidth, enabling multiple parallel paths between nodes, and load-balancing traffic where alternate paths exist.

L2/L3 control plane protocols and system management features function normally in VLT mode. Features such as VRRP and internet group management protocol (IGMP) snooping require state information coordination between the two VLT chassis. The IGMP and VLT configurations must be identical on both sides of the trunk to ensure the same behavior on both sides.

The following example shows how VLT is deployed. The switches appear as a single virtual switch from the point of view of the switch or server supporting link aggregation control protocol (LACP).

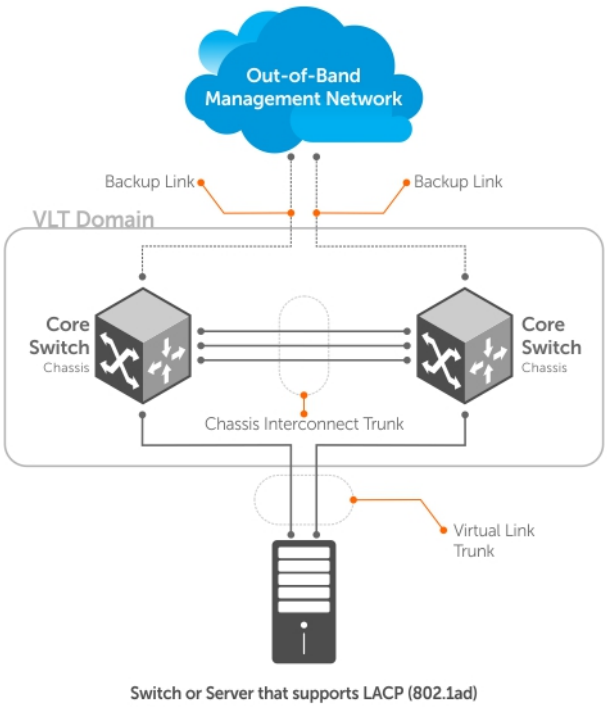


Figure 140. Example of VLT Deployment

VLT offers the following benefits:

- Allows a single device to use a LAG across two upstream devices.
- Eliminates STP-blocked ports.
- Provides a loop-free topology.
- Uses all available uplink bandwidth.
- Provides fast convergence if either the link or a device fails.
- Optimized forwarding with virtual router redundancy protocol (VRRP).
- Provides link-level resiliency.
- Assures high availability.
- Active-Active load sharing with VRRP.
- Active-Active load sharing with peer-routing for Layer-3 VLAN
- Graceful failover of LACP during reload
- Agility in VM Migration under VLT domain.

CAUTION: Dell EMC Networking does not recommend enabling Stacking and VLT simultaneously. If you enable both features at the same time, unexpected behavior may occur.

As shown in the following example, VLT presents a single logical Layer 2 domain from the perspective of attached devices that have a virtual link trunk terminating on separate chassis in the VLT domain. However, the two VLT chassis are independent Layer2/Layer3 (L2/L3) switches for devices in the upstream network. L2/L3 control plane protocols and system management features function normally in VLT mode. Features such as VRRP and internet group management protocol (IGMP) snooping require state information coordinating between the two VLT chassis. IGMP and VLT configurations must be identical on both sides of the trunk to ensure the same behavior on both sides.

The following example shows how VLT is deployed. The switches appear as a single virtual switch from the point of view of the switch or server supporting link aggregation control protocol (LACP).

VLT Terminology

The following are key VLT terms.

- **Virtual link trunk (VLT)** — The combined port channel between an attached device and the VLT peer switches.

- **VLT backup link** — The backup link monitors the connectivity between the VLT peer switches. The backup link sends configurable, periodic keep alive messages between the VLT peer switches.
- **VLT interconnect (VLTi)** — The link used to synchronize states between the VLT peer switches. Both ends must be on 10G or 40G interfaces.
- **VLT domain** — This domain includes both the VLT peer devices, VLT interconnect, and all of the port channels in the VLT connected to the attached devices. It is also associated to the configuration mode that you must use to assign VLT global parameters.
- **VLT peer device** — One of a pair of devices that are connected with the special port channel known as the VLT interconnect (VLTi).
- **Enhanced VLT (eVLT)** — Combining two VLT domains. eVLT can operate in layer 2 and layer 3 modes. eVLT is also known as mVLT.

VLT peer switches have independent management planes. A VLT interconnect between the VLT chassis maintains synchronization of L2/L3 control planes across the two VLT peer switches.

A separate backup link maintains heartbeat messages across an out-of-band (OOB) management network. The backup link ensures that node failure conditions are correctly detected and are not confused with failures of the VLT interconnect. VLT ensures that local traffic on a chassis does not traverse the VLTi and takes the shortest path to the destination via directly attached links.

The following is a summary of VLT and its functions:

- End devices (such as switches, servers, and so on) connected to a VLT domain consider the two VLT peers as a single logical switch.
- Although VLT does not require spanning tree protocols, Dell EMC Networking recommends enabling RSTP before configuring VLT to avoid possible loops from forming due to incorrect configuration.
- You can connect two VLT domains to create an eVLT topology.
- You can use eVLT as layer 2.
- Peer routing enables one VLT node to act as the default gateway for its VLT peer within a VLT domain.
- With peer routing, you need not use VRRP.
- You can use routing protocols in a VLT domain or between VLT domains (eVLT).
- VLT Proxy Gateway enables one VLT domain to act as the default gateway for its peer VLT domain in an eVLT topology.

Layer-2 Traffic in VLT Domains

In a VLT domain, the MAC address of any host connected to the VLT peers is synchronized between the VLT nodes. In the following example, VLAN 10 is spanned across three VLT domains.

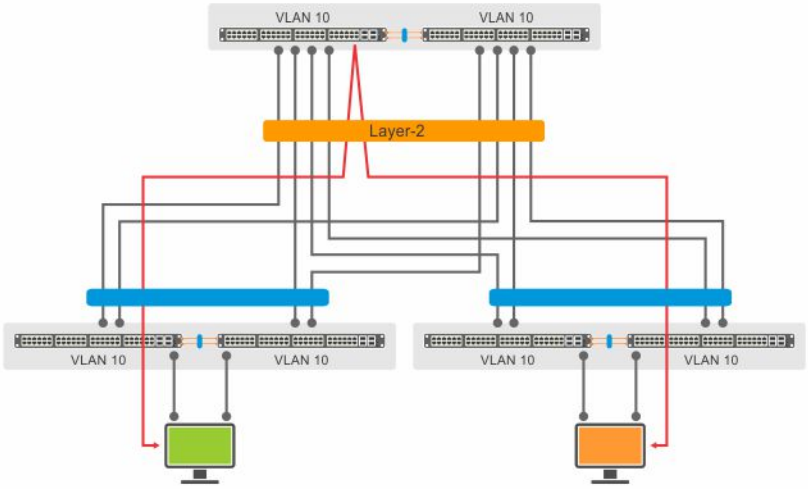


Figure 141. Layer-2 Traffic in VLT Domains

If Host 1 from a VLT domain sends a frame to Host 2 in another VLT domain, the frame can use any link shown to reach Host 2. MAC synchronization between VLT peers handles the traffic flow even if it is hashed and forwarded through the other member of the port-channel.

Viewing the MAC Synchronization Between VLT Peers

You can use the following commands to verify the MAC synchronization between VLT peers:

```
VLT-10-PEER-1#show mac-address-table count
MAC Entries for all vlans :
Dynamic Address Count :                 1007
Static Address (User-defined) Count :   1
Sticky Address Count :                  0
Total Synced Mac from Peer(N):          503
Total MAC Addresses in Use:             1008

VLT-10-PEER-1#show vlt counter mac
Total MAC VLT counters
----------------------
L2 Total MAC-Address Count:             1007



VLT-10-PEER-1#show mac-address-table
Codes: *N - VLT Peer Synced MAC
VlanId       Mac Address              Type       Interface State
10             00:00:4c:54:8b:f6 Dynamic Po 11 Active
10 00:01:e8:95:ec:97 Dynamic     Po 33 Active
10 00:01:e8:b3:ba:47 Dynamic     Po 33 Active
30 a0:00:a1:00:00:01 Dynamic     Po 11 Active
30 a0:00:a1:00:00:02 Dynamic     Po 11 Active
30 a0:00:a1:00:00:03 Dynamic     Po 11 Active
30 a0:00:a1:00:00:04 Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:05 Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:06 Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:07 Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:08 Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:09 Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:0a Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:0b Dynamic (N) Po 11 Active
30 a0:00:a1:00:00:0c Dynamic Po 11 Active


VLT-10-PEER-2#show vlt statistics mac
VLT MAC Statistics
--------------------
L2 Info Pkts sent:0, L2 Mac-sync Pkts Sent:7
L2 Info Pkts Rcvd:0, L2 Mac-sync Pkts Rcvd:9
L2 Reg Request sent:0
L2 Reg Request rcvd:0
L2 Reg Response sent:0
L2 Reg Response rcvd:0
```

Note that the MAC address synchronized by a VLT peer is differentiated with an (N) flag in the example output.

Interspersed VLANs

In Dell EMC Networking OS, the same VLAN across many racks can be extended by configuring layer-3 VLANs across the VLT nodes and the ToR switches. Spanning the VLANs in an eVLT architecture could interconnect and aggregate multiple racks with the same VLAN. With routed VLT, you can configure a VLAN as layer 3 in a VLT domain and as layer 2 VLAN in all other VLT domains. By configuring a VLAN as layer 3 in a VLT domain and as layer 2 VLAN in all other VLT domains, you can confine the ARP entries to one particular VLT domain.

At the core/aggregation layer VLT domain, you can configure common layer 3 VLANs for inter VLAN routing within the VLT domain.

VLT on Core Switches

Uplinks from servers to the access layer and from access layer to the aggregation layer are bundled in LAG groups with end-to-end Layer 2 multipathing. This set up requires “horizontal” stacking at the access layer and VLT at the aggregation layer

such that all the uplinks from servers to access and access to aggregation are in Active-Active Load Sharing mode. This example provides the highest form of resiliency, scaling, and load balancing in data center switching networks.

The following example shows stacking at the access, VLT in aggregation, and Layer 3 at the core.

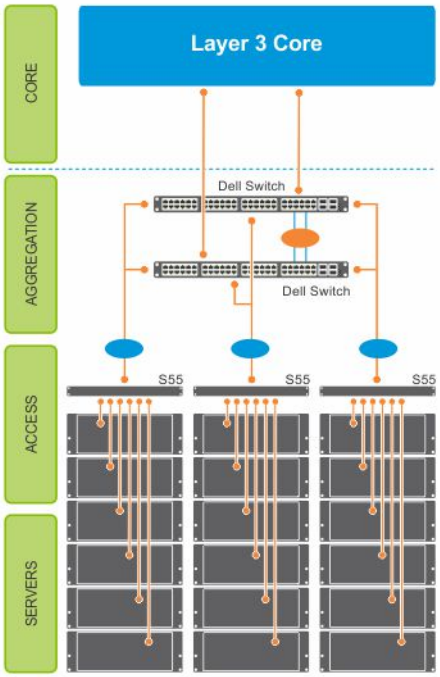


Figure 142. VLT on Core Switches

The aggregation layer is mostly in the L2/L3 switching/routing layer. For better resiliency in the aggregation, Dell EMC Networking recommends running the internal gateway protocol (IGP) on the VLTi VLAN to synchronize the L3 routing table across the two nodes on a VLT system.

Enhanced VLT

Enhanced VLT (eVLT)) refers to the ability to connect two VLT domains. An eVLT configuration creates a port channel between two VLT domains by allowing two different VLT domains, using different VLT domain ID numbers, connected by a standard link aggregation control protocol (LACP) LAG to form a loop-free Layer 2 topology in the aggregation layer.

This configuration supports a maximum of four switches, increasing the number of available ports and allowing for dual redundancy of the VLT. The following example shows how the core/aggregation port density in the Layer 2 topology is increased using eVLT. For inter-VLAN routing, you do not need a separate router.

If you enable peer routing in an eVLT topology, a VLT node acts as a proxy gateway for its peer within the VLT domain. You can also configure the two VLT domains to act as proxy gateways for each other. For more details, see the VLT Proxy Gateway chapter.

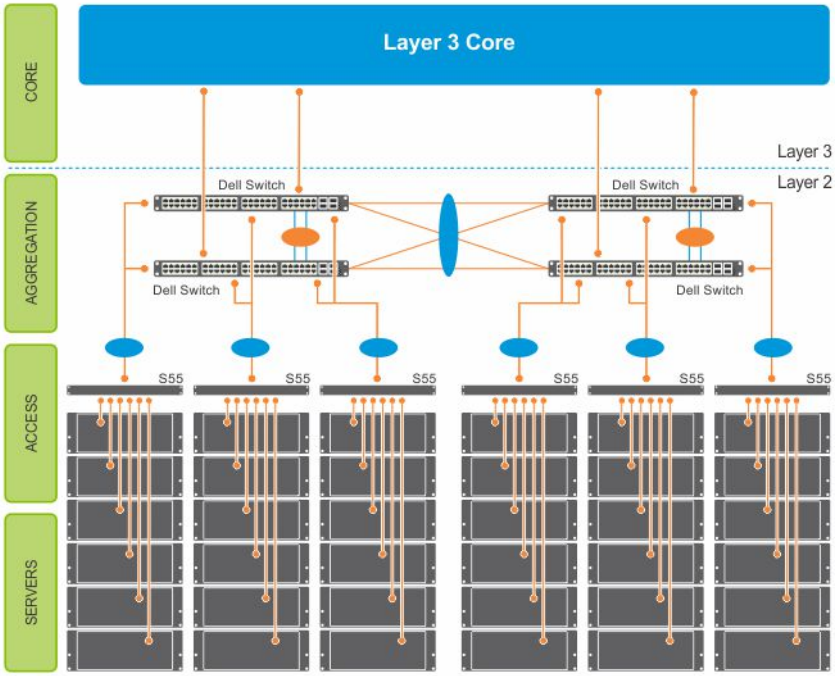


Figure 143. Enhanced VLT

Configure Virtual Link Trunking

VLT requires that you enable the feature and then configure the same VLT domain, backup link, and VLT interconnect on both peer switches.

Important Points to Remember

- You cannot enable stacking simultaneously with VLT. If you enable both at the same time, unexpected behavior can occur.
- VLT port channel interfaces must be switch ports.
- If you include RSTP on the system, configure it before VLT. Refer to Configure Rapid Spanning Tree.
- If you include PVST on the system, configure it before VLT. Refer to PVST Configuration.
- Dell EMC Networking strongly recommends that the VLTi (VLT interconnect) be a static LAG and that you disable LACP on the VLTi.
- Ensure that the spanning tree root bridge is at the Aggregation layer. Refer to RSTP and VLT for guidelines to avoid traffic loss, if you enable RSTP on the VLT device.
- If you reboot both VLT peers in BMP mode with static VLT LAGs, the DHCP server reply to the DHCP discover offer may not be forwarded by the ToR to the correct node. To avoid this issue, configure the VLT LAGs to the ToR and the ToR port channel to the VLT peers with LACP.
- If supported by the ToR, enable the lacp-ungroup feature on the ToR using the `lacp ungroup member-independent port-channel` command. If the lacp-ungroup feature is not supported on the ToR, reboot the VLT peers one at a time. After rebooting, verify that VLTi (ICL) is active before attempting DHCP connectivity.
- When a VLAN is created on VLTi peers, the VLTi port-channel is added automatically to the VLAN, whether the vlan have members or not. A VLAN creation or deletion message from the VLT peer is the trigger for adding or removing VLTi port-channels to or from a VLAN. You can manually add or remove a VLTi port-channel to a VLAN. In case a VLTi port-channel is manually removed from a VLAN, it is added back to the VLAN after reload of the VLTi peers.
- Use the `lacp ungroup member-independent` command only if the system connects to nodes using bare metal provisioning (BMP) to upgrade or boot from the network. Otherwise, when the member links flap, the links become independent switch ports and forward the traffic even before LACP is formed on that port.
- Ensure that you configure all port channels as hybrid ports and as untagged members of a VLAN where the LACP ungroup option is applicable.
- BMP uses untagged dynamic host configuration protocol (DHCP) packets to communicate with the DHCP server.

- o disable this feature on VLT and port channels, use `no lacp ungroup member-independent {vlt | port-channel}` command under the configuration mode.
- When you enable IGMP snooping on the VLT peers, ensure the value of the `delay-restore` command is not less than the query interval.
- When you enable Layer 3 routing protocols on VLT peers, make sure the delay-restore timer is set to a value that allows sufficient time for all routes to establish adjacency and exchange all the L3 routes between the VLT peers before you enable the VLT ports.
- Only use the `lacp ungroup member-independent` command if the system connects to nodes using bare metal provisioning (BMP) to upgrade or boot from the network.
- Ensure that you configure all port channels where LACP ungroup is applicable as hybrid ports and as untagged members of a VLAN. BMP uses untagged dynamic host configuration protocol (DHCP) packets to communicate with the DHCP server.
- If the DHCP server is located on the ToR and the VLTi (ICL) is down due to a failed link when a VLT node is rebooted in BMP mode, it is not able to reach the DHCP server, resulting in BMP failure.
- If the source is connected to an orphan (non-spanned, non-VLT) port in a VLT peer, the receiver is connected to a VLT (spanned) port-channel, and the VLT port-channel link between the VLT peer connected to the source and ToR is down, traffic is duplicated due to route inconsistency between peers. To avoid this scenario, Dell EMC Networking recommends configuring both the source and the receiver on a spanned VLT VLAN.
- Bulk Sync happens only for Global IPv6 Neighbors; Link-local neighbor entries are not synced.
- If all of the following conditions are true, MAC addresses may not be synced correctly:
  - VLT peers use VLT interconnect (VLTi)
  - Sticky MAC is enabled on an orphan port in the primary or secondary peer
  - MACs are currently inactive

  If this scenario occurs, use the `clear mac-address-table sticky all` command on the primary or secondary peer to correctly sync the MAC addresses.
- If you enable static ARP on only one VLT peer, entries may be overwritten during bulk sync.
- For multiple VLT LAGs configured on the same VLAN, if a host is learned on one VLT LAG and there is a station move between LAGs, the link local address redirects to the VLTi link on one of the peers. If this occurs, clear the link local address that is redirecting to the VLTi link.
- VLT Heartbeat is supported only on default VRFs.
- In a scenario where one hundred hosts are connected to a Peer1 on a non-VLT domain and traffic flows through Peer1 to Peer2; when you move these hosts from a non-VLT domain to a VLT domain and send ARP requests to Peer1, only half of these ARP requests reach Peer1, while the remaining half reach Peer2 (because of LAG hashing). The reason for this behavior is that Peer1 ignores the ARP requests that it receives on VLTi (ICL) and updates only the ARP requests that it receives on the local VLT. As a result, the remaining ARP requests still points to the Non-VLT links and traffic does not reach half of the hosts. To mitigate this issue, ensure that you configure the following settings on both the Peers (Peer1 and Peer2): `arp learn-enable` and `mac-address-table station-move refresh-arp`.
- In a topology in which two VLT peer nodes that are connected by a VLTi link and are connected to a ToR switch using a VLT LAG interface, if you configure an egress IP ACL and apply it on the VLT LAG of both peers using the `deny ip any any` command, the traffic is permitted on the VLT LAG instead of being denied. The correct behavior of dropping the traffic on the VLT LAG occurs when VLT is up on both the peer nodes. However, if VLT goes down on one of the peers, traffic traverses through VLTi and the other peer switches it to the VLT LAG. Although egress ACL is applied on the VLT nodes to deny all traffic, this egress ACL does not deny the traffic (switching traffic is not denied owing to the egress IP ACL). You cannot use egress ACLs to deny traffic properly in such a VLT scenario.
- To support Q-in-Q over VLT, ICL is implicitly made as `vlan-stack trunk port` and the TPID of the ICL is set as **8100**.
- Layer 2 Protocol Tunneling is not supported in VLT.

Configuration Notes

When you configure VLT, the following conditions apply.

- VLT domain
  - A VLT domain supports two chassis members, which appear as a single logical device to network access devices connected to VLT ports through a port channel.
  - A VLT domain consists of the two core chassis, the interconnect trunk, backup link, and the LAG members connected to attached devices.
  - Each VLT domain has a unique MAC address that you create or VLT creates automatically.
  - ARP tables are synchronized between the VLT peer nodes.

- VLT peer switches operate as separate chassis with independent control and data planes for devices attached on non-VLT ports.
- One device in the VLT domain is assigned a primary role; the other device takes the secondary role. The primary and secondary roles are required for scenarios when connectivity between the chassis is lost. VLT assigns the primary chassis role according to the lowest MAC address. You can configure the primary role manually.
- In a VLT domain, the peer switches must run the same Dell EMC Networking OS software version.
- Separately configure each VLT peer switch with the same VLT domain ID and the VLT version. If the system detects mismatches between VLT peer switches in the VLT domain ID or VLT version, the VLT Interconnect (VLTi) does not activate. To find the reason for the VLTi being down, use the `show vlt statistics` command to verify that there are mismatch errors, then use the `show vlt brief` command on each VLT peer to view the VLT version on the peer switch. If the VLT version is more than one release different from the current version in use, the VLTi does not activate.
- The chassis members in a VLT domain support connection to orphan hosts and switches that are not connected to both switches in the VLT core.

- VLT interconnect (VLTi)
  - The VLT interconnect can consist of 10G or 40G ports. A maximum of eight ports are supported. A combination of 10G and 40G ports is not supported.
  - The port channel must be in Default mode (not Switchport mode) to have VLTi recognize it.
  - The system automatically includes the required VLANs in VLTi. You do not need to manually select VLANs.
  - VLT peer switches operate as separate chassis with independent control and data planes for devices attached to non-VLT ports.
  - Port-channel link aggregation (LAG) across the ports in the VLT interconnect is required; individual ports are not supported. Dell EMC Networking strongly recommends configuring a static LAG for VLTi.
  - The VLT interconnect synchronizes L2 and L3 control-plane information across the two chassis.
  - The VLT interconnect is used for data traffic only when there is a link failure that requires using VLTi in order for data packets to reach their final destination.
  - Unknown, multicast, and broadcast traffic can be flooded across the VLT interconnect.
  - MAC addresses for VLANs configured across VLT peer chassis are synchronized over the VLT interconnect on an egress port such as a VLT LAG. MAC addresses are the same on both VLT peer nodes.
  - ARP entries configured across the VLTi are the same on both VLT peer nodes.
  - If you shut down the port channel used in the VLT interconnect on a peer switch in a VLT domain in which you did not configure a backup link, the switch's role displays in the `show vlt brief` command output as Primary instead of Standalone.
  - When you change the default VLAN ID on a VLT peer switch, the VLT interconnect may flap.
  - In a VLT domain, the following software features are supported on VLTi: link layer discovery protocol (LLDP), flow control, port monitoring, jumbo frames, and data center bridging (DCB).
  - When you enable the VLTi link, the link between the VLT peer switches is established if the following configured information is true on both peer switches:
    - the VLT system MAC address matches.
    - the VLT unit-id is not identical.

    NOTE: If you configure the VLT system MAC address or VLT unit-id on only one of the VLT peer switches, the link between the VLT peer switches is not established. Each VLT peer switch must be correctly configured to establish the link between the peers.
  - If the link between the VLT peer switches is established, changing the VLT system MAC address or the VLT unit-id causes the link between the VLT peer switches to become disabled. However, removing the VLT system MAC address or the VLT unit-id may disable the VLT ports if you happen to configure the unit ID or system MAC address on only one VLT peer at any time.
  - If the link between VLT peer switches is established, any change to the VLT system MAC address or unit-id fails if the changes made create a mismatch by causing the VLT unit-ID to be the same on both peers and/or the VLT system MAC address does not match on both peers.
  - If you replace a VLT peer node, preconfigure the switch with the VLT system MAC address, unit-id, and other VLT parameters before connecting it to the existing VLT peer switch using the VLTi connection.
  - If the size of the MTU for VLTi members is less than 1496 bytes, MAC addresses may not synchronize between VLT peers. Dell EMC Networking does not recommend using an MTU size lower than the default of 1554 bytes for VLTi members.
- VLT backup link
  - In the backup link between peer switches, heartbeat messages are exchanged between the two chassis for health checks. The default time interval between heartbeat messages over the backup link is 1 second. You can configure this interval. The range is from 1 to 5 seconds. DSCP marking on heartbeat messages is CS6.

- In order that the chassis backup link does not share the same physical path as the interconnect trunk, Dell EMC Networking recommends using the management ports on the chassis and traverse an out-of-band management network. The backup link can use user ports, but not the same ports the interconnect trunk uses.
- The chassis backup link does not carry control plane information or data traffic. Its use is restricted to health checks only.

- Virtual link trunks (VLTs) between access devices and VLT peer switches
  - To connect servers and access switches with VLT peer switches, you use a VLT port channel, as shown in Overview. Up to 48 port-channels are supported; up to 16 member links are supported in each port channel between the VLT domain and an access device.
  - The discovery protocol running between VLT peers automatically generates the ID number of the port channel that connects an access device and a VLT switch. The discovery protocol uses LACP properties to identify connectivity to a common client device and automatically generates a VLT number for port channels on VLT peers that connects to the device. The discovery protocol requires that an attached device always runs LACP over the port-channel interface.
  - VLT provides a loop-free topology for port channels with endpoints on different chassis in the VLT domain.
  - VLT uses shortest path routing so that traffic destined to hosts via directly attached links on a chassis does not traverse the chassis-interconnect link.
  - VLT allows multiple active parallel paths from access switches to VLT chassis.
  - VLT supports port-channel links with LACP between access switches and VLT peer switches. Dell EMC Networking recommends using static port channels on VLTi.
  - If VLTi connectivity with a peer is lost but the VLT backup connectivity indicates that the peer is still alive, the VLT ports on the Secondary peer are orphaned and are shut down.
    - In one possible topology, a switch uses the BMP feature to receive its IP address, configuration files, and boot image from a DHCP server that connects to the switch through the VLT domain. In the port-channel used by the switch to connect to the VLT domain, configure the port interfaces on each VLT peer as hybrid ports before adding them to the port channel (see Connecting a VLT Domain to an Attached Access Device (Switch or Server)). To configure a port in Hybrid mode so that it can carry untagged, single-tagged, and double-tagged traffic, use the `portmode hybrid` command in Interface Configuration mode as described in Configuring Native VLANs.
    - For example, if the DHCP server is on the ToR and VLTi (ICL) is down (due to either an unavailable peer or a link failure), whether you configured the VLT LAG as static or LACP, when a single VLT peer is rebooted in BMP mode, it cannot reach the DHCP server, resulting in BMP failure.
- Software features supported on VLT port-channels
  - In a VLT domain, the following software features are supported on VLT port-channels: 802.1p, ingress and egress ACLs, BGP, DHCP relay, IS-IS, OSPF, active-active PIM-SM, PIM-SSM, VRRP, Layer 3 VLANs, LLDP, flow control, port monitoring, jumbo frames, IGMP snooping, sFlow, ingress and egress ACLs, and Layer 2 control protocols RSTP and PVST only.

    NOTE: Peer VLAN spanning tree plus (PVST+) passthrough is supported in a VLT domain. PVST+ BPDUs does not result in an interface shutdown. PVST+ BPDUs for a nondefault VLAN is flooded out as any other L2 multicast packet. On a default VLAN, RTSP is part of the PVST+ topology in that specific VLAN (default VLAN).
  - In a VLT domain, ingress and egress QoS policies are supported on physical VLT ports, which can be members of VLT port channels in the domain.
    - Ingress and egress QoS policies applied on VLT ports must be the same on both VLT peers.
    - Apply the same ingress and egress QoS policies on VLTi (ICL) member ports to handle failed links.
  - For detailed information about how to use VRRP in a VLT domain, see the following *VLT and VRRP interoperability* section.
  - For information about configuring IGMP Snooping in a VLT domain, see VLT and IGMP Snooping.
  - All system management protocols are supported on VLT ports, including SNMP, RMON, AAA, ACL, DNS, FTP, SSH, Syslog, NTP, RADIUS, SCP, TACACS+, Telnet, and LLDP.
  - Enable Layer 3 VLAN connectivity VLT peers by configuring a VLAN network interface for the same VLAN on both switches.
  - Dell EMC Networking does not recommend enabling peer-routing if the CAM is full. To enable peer-routing, a minimum of two local DA spaces for wild-card functionality are required.
  - RSPAN and ERSPAN are supported on VLT.
  - FRRP is supported only on the VLTi. This feature enables configuration of an FRRP ring through VLTi. However, FRRP is not supported on any other VLT port-channel except for VLTi.
- Software features supported on VLT physical ports
  - In a VLT domain, the following software features are supported on VLT physical ports: 802.1p, LLDP, flow control, IPv6 dynamic routing, port monitoring, DHCP snooping, and jumbo frames.
- Software features not supported with VLT
  - In a VLT domain, the following software features are not supported on VLT ports: 802.1x, GVRP, and BFD.
- VLT and VRRP interoperability

- In a VLT domain, VRRP interoperates with virtual link trunks that carry traffic to and from access devices (see Overview). The VLT peers belong to the same VRRP group and are assigned master and backup roles. Each peer actively forwards L3 traffic, reducing the traffic flow over the VLT interconnect.
- VRRP elects the router with the highest priority as the master in the VRRP group. To ensure VRRP operation in a VLT domain, configure VRRP group priority on each VLT peer so that a peer is either the master or backup for all VRRP groups configured on its interfaces. For more information, see Setting VRRP Group (Virtual Router) Priority.
- To verify that a VLT peer is consistently configured for either the master or backup role in all VRRP groups, use the `show vrrp` command on each peer.
- Configure the same L3 routing (static and dynamic) on each peer so that the L3 reachability and routing tables are identical on both VLT peers. Both the VRRP master and backup peers must be able to locally forward L3 traffic in the same way.
- In a VLT domain, although both VLT peers actively participate in L3 forwarding as the VRRP master or backup router, the `show vrrp` command output displays one peer as master and the other peer as backup.

- Failure scenarios
  - On a link failover, when a VLT port channel fails, the traffic destined for that VLT port channel is redirected to the VLTi to avoid flooding.
  - When a VLT switch determines that a VLT port channel has failed (and that no other local port channels are available), the peer with the failed port channel notifies the remote peer that it no longer has an active port channel for a link. The remote peer then enables data forwarding across the interconnect trunk for packets that would otherwise have been forwarded over the failed port channel. This mechanism ensures reachability and provides loop management. If the VLT interconnect fails, the VLT software on the primary switch checks the status of the remote peer using the backup link. If the remote peer is up, the secondary switch disables all VLT ports on its device to prevent loops.
  - If all ports in the VLT interconnect fail, or if the messaging infrastructure fails to communicate across the interconnect trunk, the VLT management system uses the backup link interface to determine whether the failure is a link-level failure or whether the remote peer has failed entirely. If the remote peer is still alive (heartbeat messages are still being received), the VLT secondary switch disables its VLT port channels. If keepalive messages from the peer are not being received, the peer continues to forward traffic, assuming that it is the last device available in the network. In either case, after recovery of the peer link or reestablishment of message forwarding across the interconnect trunk, the two VLT peers resynchronize any MAC addresses learned while communication was interrupted and the VLT system continues normal data forwarding.
  - If the primary chassis fails, the secondary chassis takes on the operational role of the primary.
- The SNMP MIB reports VLT statistics.

Primary and Secondary VLT Peers

To prevent issues when connectivity between peers is lost, you can designate Primary and Secondary roles for VLT peers . You can elect or configure the Primary Peer. By default, the peer with the lowest MAC address is selected as the Primary Peer. You can configure another peer as the Primary Peer using the VLT `domain` *`domain-id`* `role priority` *`priority-value`* command.

If the VLTi link fails, the status of the remote VLT Primary Peer is checked using the backup link. If the remote VLT Primary Peer is available, the Secondary Peer disables all VLT ports to prevent loops.

If all ports in the VLTi link fail or if the communication between VLTi links fails, VLT checks the backup link to determine the cause of the failure. If the failed peer can still transmit heartbeat messages, the Secondary Peer disables all VLT member ports and any Layer 3 interfaces attached to the VLAN associated with the VLT domain. If heartbeat messages are not received, the Secondary Peer forwards traffic assumes the role of the Primary Peer. If the original Primary Peer is restored, the VLT peer reassigned as the Primary Peer retains this role and the other peer must be reassigned as a Secondary Peer. Peer role changes are reported as SNMP traps.

RSTP and VLT

VLT provides loop-free redundant topologies and does not require RSTP.

RSTP can cause temporary port state blocking and may cause topology changes after link or node failures. Spanning tree topology changes are distributed to the entire layer 2 network, which can cause a network-wide flush of learned MAC and ARP addresses, requiring these addresses to be re-learned. However, enabling RSTP can detect potential loops caused by non-system issues such as cabling errors or incorrect configurations. To minimize possible topology changes after link or node failure, RSTP is useful for potential loop detection. Configure RSTP using the following specifications.

The following recommendations help you avoid these issues and the associated traffic loss caused by using RSTP when you enable VLT on both VLT peers:

- Configure any ports at the edge of the spanning tree's operating domain as edge ports, which are directly connected to end stations or server racks. Disable RSTP on ports connected directly to Layer 3-only routers not running STP or configure them as edge ports.
- Ensure that the primary VLT node is the root bridge and the secondary VLT peer node has the second-best bridge ID in the network. If the primary VLT peer node fails, the secondary VLT peer node becomes the root bridge, avoiding problems with spanning tree port state changes that occur when a VLT node fails or recovers.
- Even with this configuration, if the node has non-VLT ports using RSTP that you did not configure as edge ports and are connected to other Layer 2 switches, spanning tree topology changes are still detected after VLT node recovery. To avoid this scenario, ensure that you configure any non-VLT ports as edge ports or disable RSTP.

VLT Bandwidth Monitoring

When bandwidth usage of the VLTi (ICL) exceeds 80%, a syslog error message (shown in the following message) and an SNMP trap are generated.

```
%STKUNIT0-M:CP %VLTMGR-6-VLT-LAG-ICL: Overall Bandwidth utilization of VLT-ICL-LAG (port-
channel 25)
crosses threshold. Bandwidth usage (80 )
```

When the bandwidth usage drops below the 80% threshold, the system generates another syslog message (shown in the following message) and an SNMP trap.

```
%STKUNIT0-M:CP %VLTMGR-6-VLT-LAG-ICL: Overall Bandwidth utilization of VLT-ICL-LAG (port-
channel 25)
reaches below threshold. Bandwidth usage (74 )VLT show remote port channel status
```

VLT and Stacking

You cannot enable stacking on the units with VLT.

If you enable stacking on a unit on which you want to enable VLT, you must first remove the unit from the existing stack. After you remove the unit, you can configure VLT on the unit.

VLT and IGMP Snooping

When configuring IGMP Snooping with VLT, ensure the configurations on both sides of the VLT trunk are identical to get the same behavior on both sides of the trunk.

When you configure IGMP snooping on a VLT node, the dynamically learned groups and multicast router ports are automatically learned on the VLT peer node.

VLT IPv6

The following features have been enhanced to support IPv6:

- **VLT Sync** — Entries learned on the VLT interface are synced on both VLT peers.
- **Non-VLT Sync** — Entries learned on non-VLT interfaces are synced on both VLT peers.
- **Tunneling** — Control information is associated with tunnel traffic so that the appropriate VLT peer can mirror the ingress port as the VLT interface rather than pointing to the VLT peer's VLTi link.
- **Statistics and Counters** — Statistical and counter information displays IPv6 information when applicable.
- **Heartbeat** — You can configure an IPv4 or IPv6 address as a backup link destination. You cannot use an IPv4 and an IPv6 address simultaneously.

VLT Port Delayed Restoration

When a VLT node boots up, if the VLT ports have been previously saved in the start-up configuration, they are not immediately enabled.

To ensure MAC and ARP entries from the VLT per node are downloaded to the newly enabled VLT node, the system allows time for the VLT ports on the new node to be enabled and begin receiving traffic.

The `delay-restore` feature waits for all saved configurations to be applied, then starts a configurable timer. After the timer expires, the VLT ports are enabled one-by-one in a controlled manner. The delay between bringing up each VLT port-channel is proportional to the number of physical members in the port-channel. The default is 90 seconds.

To change the duration of the configurable timer, use the `delay-restore` command.

If you enable IGMP snooping, IGMP queries are also sent out on the VLT ports at this time allowing any receivers to respond to the queries and update the multicast table on the new node.

This delay in bringing up the VLT ports also applies when the VLTi link recovers from a failure that caused the VLT ports on the secondary VLT peer node to be disabled.

PIM-Sparse Mode Support on VLT

The designated router functionality of the PIM Sparse-Mode multicast protocol is supported on VLT peer switches for multicast sources and receivers that are connected to VLT ports.

VLT peer switches can act as a last-hop router for IGMP receivers and as a first-hop router for multicast sources.

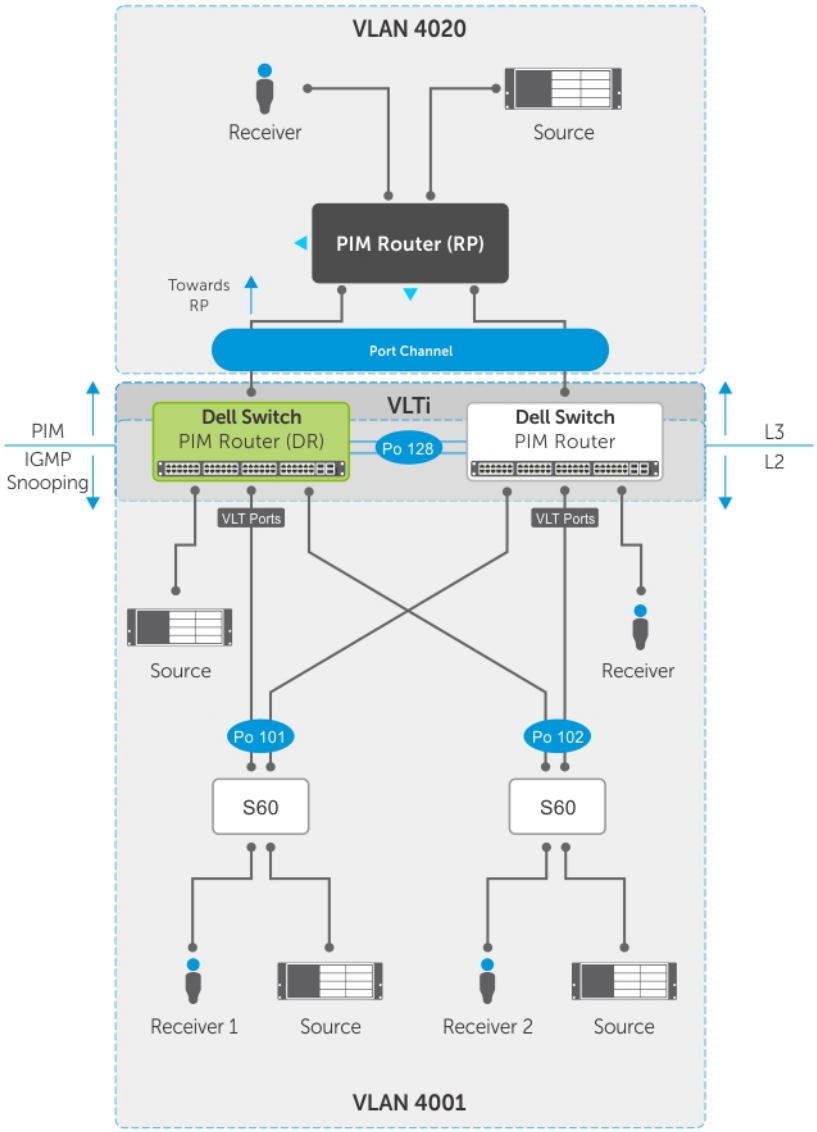


Figure 144. PIM-Sparse Mode Support on VLT

On each VLAN where the VLT peer nodes act as the first hop or last hop routers, one of the VLT peer nodes is elected as the PIM designated router. If you configured IGMP snooping along with PIM on the VLT VLANs, you must configure VLTi as the static multicast router port on both VLT peer switches. This ensures that for first hop routers, the packets from the source are redirected to the designated router (DR) if they are incorrectly hashed. In addition to being first-hop or last -hop routers, the peer node can also act as an intermediate router.

On a VLT-enabled PIM router, if any PIM neighbor is reachable through a Spanned Layer 3 (L3) VLAN interface, this must be the **only** PIM-enabled interface to reach that neighbor. A Spanned L3 VLAN is any L3 VLAN configured on both peers in a VLT domain. This does not apply to server-side L2 VLT ports because they do not connect to any PIM routers. These VLT ports can be members of multiple PIM-enabled L3 VLANs for compatibility with IGMP.

To route traffic to and from the multicast source and receiver, enable PIM on the L3 side connected to the PIM router using the `ip pim sparse-mode` command.

Each VLT peer runs its own PIM protocol independently of other VLT peers. To ensure the PIM protocol states or multicast routing information base (MRIB) on the VLT peers are synced, if the incoming interface (IIF) and outgoing interface (OIF) are Spanned, the multicast route table is synced between the VLT peers.

To verify the PIM neighbors on the VLT VLAN and on the multicast port, use the `show ip pim neighbor`, `show ip igmp snooping mrouter`, and `show running config` commands.

You can configure virtual link trunking (VLT) peer nodes as rendezvous points (RPs) in a Protocol Independent Multicast (PIM) domain.

If the VLT node elected as the designated router fails and you enable VLT Multicast Routing, multicast routes are synced to the other peer for traffic forwarding to ensure minimal traffic loss. If you did not enable VLT Multicast Routing, traffic loss occurs until the other VLT peer is selected as the DR.

VLT Routing

VLT Routing refers to the ability to run a dynamic routing protocol within a single VLT domain or between VLT domains (mVLT). In a single VLT domain, VLT routing allows routing adjacencies to be formed across the VLTi link. In eVLT, routing adjacencies are formed across the port-channel that connects the two VLT domains.

Because VLT ports are Layer 2 ports and not IP interfaces, VLT Unicast and VLT Multicast routing protocols do not operate directly on VLT ports. You must add the VLT ports as a member of one or more VLANs and assign IP addresses to these VLANs. VLT Unicast and VLT Multicast routing protocols require VLAN IP interfaces for operation. Protocols such as BGP, ISIS, OSPF, and PIM are compatible with VLT Unicast Routing and VLT Multicast Routing.

Layer 2 protocols from the ToR devices to the server are intra-rack and inter-rack. Although no spanning tree is required, interoperability with spanning trees at the aggregation layer is supported to prevent switching loops from forming due to any incorrect configuration. . Communication between devices is active-active, with no blocked links. MAC tables are synchronized between VLT nodes for bridging and you can enable IGMP snooping.

Spanned VLANs

Any VLAN configured on both VLT peer nodes is referred to as a Spanned VLAN. The VLT Interconnect (VLTi) port is automatically added as a member of the Spanned VLAN. As a result, any adjacent router connected to at least one VLT node on a Spanned VLAN subnet is directly reachable from both VLT peer nodes at the routing level.

Peer Routing

Peer routing enables one VLT node to act as a proxy gateway for the other peer in a VLT domain. When you enable routing on VLT peers, you can also enable the peer routing feature.

The following figure shows how a packet is routed when peer routing is not enabled. Due to the hashing algorithm in the port channel, a packet from a host is sent to either of the VLT port-channel members. If a packet is sent to Peer-1 which is not the destined gateway for the hosts under the ToR Switch, the packet is switched to the destined VLT peer (Peer-2) using the VLTi link. Peer-2 then routes the packet to its destination.

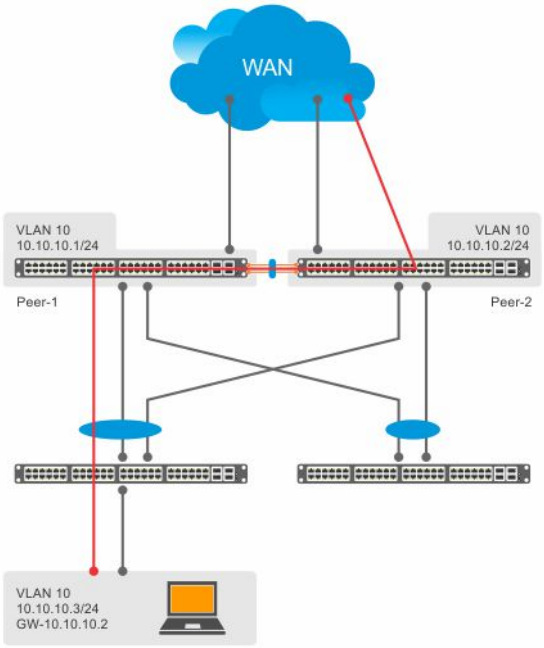


Figure 145. Packets without peer routing enabled

If you enable peer routing, a VLT node acts as a proxy gateway for its connected VLT peer as shown in the image below. Even though the gateway address of the packet is different, Peer-1 routes the packet to its destination on behalf of Peer-2 to avoid sub-optimal routing.

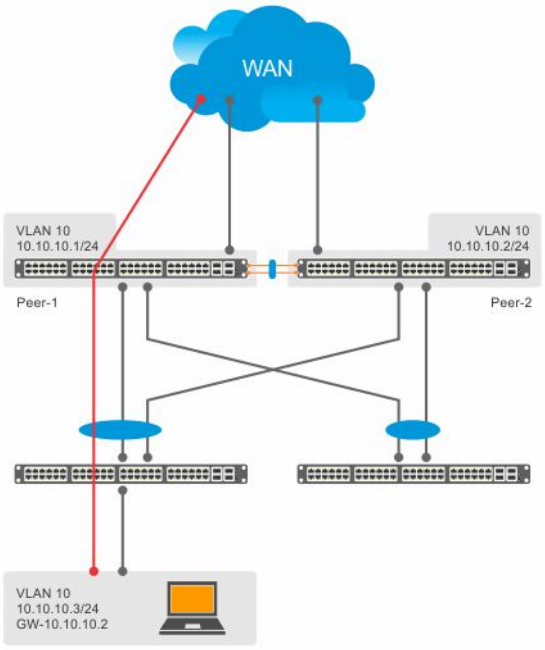


Figure 146. Packets with peer routing enabled

Benefits of Peer Routing

-
- Avoids sub-optimal routing
- Reduces latency by avoiding another hop in the traffic path.
- You can reduce the number of VLTi port channel members based on your specific design.

With peer routing, you need not configure VRRP for the participating VLANs. As both VLT nodes act as a gateway for its peer, irrespective of the gateway IP address, the traffic flows upstream without any latency. There is no limitation for the number of VLANS.

VLT Unicast Routing

VLT unicast routing is a type of VLT peer routing that locally routes unicast packets destined for the L3 endpoint of the VLT peer. This method avoids sub-optimal routing. Peer-routing syncs the MAC addresses of both VLT peers and requires two local DA entries in TCAM. If a VLT node is down, a timer that allows you to configure the amount of time needed for peer recovery provides resiliency. You can enable VLT unicast across multiple configurations using VLT links. You can enable ECMP on VLT nodes using VLT unicast.

VLT unicast routing is supported on both IPv4 and IPv6. To enable VLT unicast routing, both VLT peers must be in L3 mode. Static route and routing protocols such as RIP, OSPF, ISIS, and BGP are supported. However, point-to-point configuration is not supported. To enable VLT unicast, the VLAN configuration must be symmetrical on both peers. You cannot configure the same VLAN as Layer 2 on one node and as Layer 3 on the other node. Configuration mismatches are logged in the syslog and display in the `show vlt mismatch` command output.

If you enable VLT unicast routing, the following actions occur:

- L3 routing is enabled on any new IP address / IPv6 address configured for a VLAN interface that is up.
- L3 routing is enabled on any VLAN with an admin state of up.

NOTE: If the CAM is full, do not enable peer-routing.

NOTE: The peer routing and peer-routing-timeout is applicable for both IPv6/ IPv4.

Configuring VLT Unicast

To enable and configure VLT unicast, follow these steps.

1. Enable VLT on a switch, then configure a VLT domain and enter VLT-domain configuration mode.
   CONFIGURATION mode
   `vlt domain domain-id`
2. Enable peer-routing.
   VLT DOMAIN mode
   `peer-routing`
3. Configure the peer-routing timeout.
   VLT DOMAIN mode
   `peer-routing–timeout value`
   value: Specify a value (in seconds) from 1 to 65535. The default value is infinity (without configuring the timeout).

VLT Multicast Routing

VLT multicast routing is a type of VLT peer routing that provides resiliency to multicast routed traffic during the multicast routing protocol convergence period after a VLT link or VLT peer fails using the least intrusive method (PIM) and does not alter current protocol behavior.

Unlike VLT unicast routing, a normal multicast routing protocol does not exchange multicast routes between VLT peers. When you enable VLT multicast routing, the multicast routing table is synced between the VLT peers. Only multicast routes configured with a Spanned VLAN IP as their IIF are synced between VLT peers. For multicast routes with a Spanned VLAN IIF, only OIFs configured with a Spanned VLAN IP interface are synced between VLT peers.

The advantages of syncing the multicast routes between VLT peers are:

- **VLT resiliency** — After a VLT link or peer failure, if the traffic hashes to the VLT peer, the traffic continues to be routed using multicast until the PIM protocol detects the failure and adjusts the multicast distribution tree.
- **Optimal routing** — The VLT peer that receives the incoming traffic can directly route traffic to all downstream routers connected on VLT ports.
- **Optimal VLTi forwarding** — Only one copy of the incoming multicast traffic is sent on the VLTi for routing or forwarding to any orphan ports, rather than forwarding all the routed copies.

Important Points to Remember

- You can only use one spanned VLAN from a PIM-enabled VLT node to an external neighboring PIM router.
- If you connect multiple spanned VLANs to a PIM neighbor, or if both spanned and non-spanned VLANs can access the PIM neighbor, ECMP can cause the PIM protocol running on each VLT peer node to choose a different VLAN or IP route to reach the PIM neighbor. This can result in issues with multicast route syncing between peers.
- Both VLT peers require symmetric Layer 2 and Layer 3 configurations on both VLT peers for any spanned VLAN.
- For optimal performance, configure the VLT VLAN routing metrics to prefer VLT VLAN interfaces over non-VLT VLAN interfaces.
- When using factory default settings on a new switch deployed as a VLT node, packet loss may occur due to the requirement that all ports must be open.
- ECMP is not compatible on VLT nodes using VLT multicast. You must use a single VLAN.

Configuring VLT Multicast

To enable and configure VLT multicast, follow these steps.

1. Enable VLT on a switch, then configure a VLT domain and enter VLT-domain configuration mode.
   CONFIGURATION mode
   `vlt domain domain-id`
2. Enable peer-routing.
   VLT DOMAIN mode
   `peer-routing`

3. Configure the multicast peer-routing timeout.

   VLT DOMAIN mode

   `multicast peer-routing—timeout value`

   value: Specify a value (in seconds) from 1 to 1200.

   NOTE: Reduce the multicast peer-routing-timeout value to 10 seconds to clear the (S,G) entry in mroute in primary VLT peer. Also, the MLD leave packet must be sent after the unicast route convergence.

4. Configure a PIM-SM compatible VLT node as a designated router (DR). For more information, refer to Configuring a Designated Router.
5. Configure a PIM-enabled external neighboring router as a rendezvous point (RP). For more information, refer to Configuring a Static Rendezvous Point.
6. Configure the VLT VLAN routing metrics to prefer VLT VLAN interfaces over non-VLT VLAN interfaces. For more information, refer to Classify Traffic.
7. Configure symmetrical Layer 2 and Layer 3 configurations on both VLT peers for any spanned VLAN.

Non-VLT ARP Sync

ARP entries (including ND entries) learned on other ports are synced with the VLT peer to support station move scenarios.

NOTE: ARP entries learned on non-VLT, non-spanned VLANs are not synced with VLT peers.

RSTP Configuration

RSTP is supported in a VLT domain.

Before you configure VLT on peer switches, configure RSTP in the network. RSTP is required for initial loop prevention during the VLT startup phase. You may also use RSTP for loop prevention in the network outside of the VLT port channel. For information about how to configure RSTP, Rapid Spanning Tree Protocol (RSTP).

Run RSTP on both VLT peer switches. The primary VLT peer controls the RSTP states, such as forwarding and blocking, on both the primary and secondary peers. Dell EMC Networking recommends configuring the primary VLT peer as the RSTP primary root device and configuring the secondary VLT peer as the RSTP secondary root device.

BPDUs use the MAC address of the primary VLT peer as the RSTP bridge ID in the designated bridge ID field. The primary VLT peer sends these BPDUs on VLT interfaces connected to access devices. The MAC address for a VLT domain is automatically selected on the peer switches when you create the domain (refer to Enabling VLT and Creating a VLT Domain).

Configure both ends of the VLT interconnect trunk with identical RSTP configurations. When you enable VLT, the `show spanning-tree rstp brief` command output displays VLT information (refer to Verifying a VLT Configuration).

Preventing Forwarding Loops in a VLT Domain

During the bootup of VLT peer switches, a forwarding loop may occur until the VLT configurations are applied on each switch and the primary/secondary roles are determined.

To prevent the interfaces in the VLT interconnect trunk and RSTP-enabled VLT ports from entering a Forwarding state and creating a traffic loop in a VLT domain, take the following steps.

1. Configure RSTP in the core network and on each peer switch as described in Rapid Spanning Tree Protocol (RSTP).

   Disabling RSTP on one VLT peer may result in a VLT domain failure.
2. Enable RSTP on each peer switch.

   PROTOCOL SPANNING TREE RSTP mode

   `no disable`
3. Configure each peer switch with a unique bridge priority.

   PROTOCOL SPANNING TREE RSTP mode

   `bridge-priority`

Sample RSTP configuration

The following is a sample of an RSTP configuration:

Using the example shown in the Overview section as a sample VLT topology, the primary VLT switch sends BPDUs to an access device (switch or server) with its own RSTP bridge ID. BPDUs generated by an RSTP-enabled access device are only processed by the primary VLT switch. The secondary VLT switch tunnels the BPDUs that it receives to the primary VLT switch over the VLT interconnect. Only the primary VLT switch determines the RSTP roles and states on VLT ports and ensures that the VLT interconnect link is never blocked.

In the case of a primary VLT switch failure, the secondary switch starts sending BPDUs with its own bridge ID and inherits all the port states from the last synchronization with the primary switch. An access device never detects the change in primary/ secondary roles and does not see it as a topology change.

The following examples show the RSTP configuration that you must perform on each peer switch to prevent forwarding loops.

Configure RSTP on VLT peers to prevent forwarding loops—VLT peer 1 (primary)

```
Dell_VLTpeer1(conf)#protocol spanning-tree rstp
Dell_VLTpeer1(conf-rstp)#no disable
Dell_VLTpeer1(conf-rstp)#bridge-priority 4096
```

Configure RSTP on VLT peers to prevent forwarding loops—VLT peer 2 (secondary)

```
Dell_VLTpeer2(conf)#protocol spanning-tree rstp
Dell_VLTpeer2(conf-rstp)#no disable
Dell_VLTpeer2(conf-rstp)#bridge-priority 8192
```

NOTE: When you remove the VLT configuration, RSTP is recommended as a backup solution to avoid spanning—tree loops. In this case, you can use the following RSTP default values:

- hello–time: 2 seconds
- forward–delay: 15 seconds
- max–age: 20

There is no additional performance gain on VLT, if you configure the RSTP values other than the default values.

Configuring VLT

To configure VLT, use the following procedure.

Prerequisites:

Before you begin, make sure that both VLT peer switches are running the same Dell EMC Networking OS version and are configured for RSTP as described in the RSTP Configuration section. For VRRP operation, ensure that you configure VRRP groups and L3 routing on each VLT peer as described in *VLT and VRRP interoperability* in the Configuration Notes section. To configure VLT and create a VLT domain in which two switches are physically connected and treated as a single port channel by access devices, configure the following settings on each VLT peer device.

1. Configure the VLT interconnect for the VLT domain. The primary and secondary switch roles in the VLT domain are automatically assigned after you configure both sides of the VLTi.
   NOTE: If you use a third-party ToR unit, to avoid potential problems if you reboot the VLT peers, Dell EMC recommends using static LAGs on the VLTi between VLT peers.
2. Enable VLT and create a VLT domain ID. VLT automatically selects a system MAC address.
3. Configure a backup link for the VLT domain.
4. (Optional) Manually reconfigure the default VLT settings, such as the MAC address and VLT primary/ secondary roles.
5. Connect the peer switches in a VLT domain to an attached access device (switch or server).

Configuring a VLT Interconnect

To configure a VLT interconnect, follow these steps.

1. Configure the port channel for the VLT interconnect on a VLT switch and enter interface configuration mode.

   CONFIGURATION mode

   ```
   interface port-channel id-number
   ```

   Enter the same port-channel number configured with the `peer-link port-channel` command as described in Enabling VLT and Creating a VLT Domain.

   NOTE: To be included in the VLTi, the port channel must be in Default mode (`no switchport` or VLAN assigned).

2. Remove any IP address from the interface if already present.

   INTERFACE PORT-CHANNEL mode

   ```
   no ip address
   ```

3. Add one or more port interfaces to the port channel.

   INTERFACE PORT-CHANNEL mode

   ```
   channel-member interface
   ```

   interface: specify one of the following interface types:

   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

4. Ensure that the port channel is active.

   INTERFACE PORT-CHANNEL mode

   ```
   no shutdown
   ```

5. Repeat Steps 1 to 4 on the VLT peer switch to configure the VLT interconnect.

Enabling VLT and Creating a VLT Domain

To enable VLT and create a VLT domain, use the following steps.

1. Enable VLT on a switch, then configure a VLT domain and enter VLT-domain configuration mode.

   CONFIGURATION mode

   ```
   vlt domain domain-id
   ```

   The domain ID range is from 1 to 1000.

   Configure the same domain ID on the peer switch to allow for common peering. VLT uses the domain ID to automatically create a VLT MAC address for the domain. If you do not configure the system explicitly, the system mac-address of the primary will be the VLT MAC address for the domain.

   To disable VLT, use the `no vlt domain` command.

   NOTE: Do not use MAC addresses such as “reserved” or “multicast.”

2. Configure the IP address of the management interface on the remote VLT peer to be used as the endpoint of the VLT backup link for sending out-of-band hello messages.

   VLT DOMAIN CONFIGURATION mode

   ```
   back-up destination {ipv4-address | ipv6-address} [interval seconds] {vrf [management
   vrf-name | vrf-name]}
   ```

   NOTE: If management VRF or any specific VRF is enabled on the system, then use the `back-up destination` command with `vrf [management vrf-name | vrf-name]`

   You can optionally specify the time interval used to send hello messages. The range is from 1 to 5 seconds.

3. Configure the port channel to be used as the VLT interconnect between VLT peers in the domain.

   VLT DOMAIN CONFIGURATION mode

   ```
   peer-link port-channel id-number
   ```

4. Enable peer routing.

VLT DOMAIN CONFIGURATION mode

`peer-routing`

If you enable peer routing, a VLT node acts as the proxy gateway for its peer.

5. (Optional) After you configure a VLT domain on each peer switch and connect (cable) the two VLT peers on each side of the VLT interconnect, the system elects a primary and secondary VLT peer device (see Primary and Secondary VLT Peers). To configure the primary and secondary roles before the election process, use the `primary-priority` command. Enter a lower value on the primary peer and a higher value on the secondary peer.

   VLT DOMAIN CONFIGURATION mode

   `primary-priority` *`value`*

   The priority values are from 1 to 65535. The default is **32768**.

   If the primary peer fails, the secondary peer (with the higher priority) takes the primary role. If the primary peer (with the lower priority) later comes back online, it is assigned the secondary role (there is no preemption).

6. (Optional) Prevent a possible loop during the bootup of a VLT peer switch or a device that accesses the VLT domain.

   CONFIGURATION mode

   `lacp ungroup member-independent {vlt | port-channel` *`port-channel-id`*`}`

   LACP on VLT ports (on a VLT switch or access device), which are members of the virtual link trunk, is not brought up until the VLT domain is recognized on the access device.

7. Repeat Steps 1 to 5 on the VLT peer switch to configure the IP address of this switch as the endpoint of the VLT backup link and to configure the same port channel for the VLT interconnect.

Configuring a VLT Backup Link

To configure a VLT backup link, use the following command.

1. Specify the management interface to be used for the backup link through an out-of-band management network.

   CONFIGURATION mode

   `interface managementethernet` *`slot/port`*

   Enter the slot (0-1) and the port (0).

2. Configure an IPv4 address (A.B.C.D) or IPv6 address (X:X:X:X::X) and mask (/x) on the interface.

   MANAGEMENT INTERFACE mode

   `{ip address` *`ipv4-address/ mask`* `| ipv6 address` *`ipv6-address/ mask`*`}`

   This is the IP address to be configured on the VLT peer with the back-up destination command.

3. Ensure that the interface is active.

   MANAGEMENT INTERFACE mode

   `no shutdown`

4. Configure a VLT backup link using the IPv4 or IPv6 address of the VLT peer's management interface.

   MANAGEMENT INTERFACE mode

   `back-up destination {ip address` *`ipv4-address/ mask`* `| ipv6 address` *`ipv6-address/ mask`*`}`

5. Repeat Steps 1 to 4 on the VLT peer switch.

To set an amount of time, in seconds, to delay the system from restoring the VLT port, use the `delay-restore` command at any time. For more information, refer to VLT Port Delayed Restoration.

Configuring a VLT Port Delay Period

To configure a VLT port delay period, use the following commands.

1. Enter VLT-domain configuration mode for a specified VLT domain.

   CONFIGURATION mode

   `vlt domain` *`domain-id`*

   The range of domain IDs from 1 to 1000.

2. Enter an amount of time, in seconds, to delay the restoration of the VLT ports after the system is rebooted.

   CONFIGURATION mode

`delay-restore` *`delay-restore-time`*

The range is from 1 to 1200.

The default is **90 seconds**.

Reconfiguring the Default VLT Settings (Optional)

To reconfigure the default VLT settings, use the following commands.

1. Enter VLT-domain configuration mode for a specified VLT domain.

   CONFIGURATION mode

   `vlt domain` *`domain-id`*

   The range of domain IDs is from 1 to 1000.
2. After you configure a VLT domain on each peer switch and connect (cable) the two VLT peers on each side of the VLT interconnect, the system elects a primary and secondary VLT peer device. To configure the primary and secondary roles before the election process, use the primary-priority command. Enter a lower value on the primary peer and a higher value on the secondary peer.

   If the primary peer fails, the secondary peer (with the higher priority) takes the primary role. If the primary peer (with the lower priority) later comes back online, it is assigned the secondary role (there is no preemption).
3. (Optional) When you create a VLT domain on a switch, Dell EMC Networking OS automatically creates a VLT-system MAC address used for internal system operations.

   VLT DOMAIN CONFIGURATION mode

   `system-mac mac-address` *`mac-address`*

   To explicitly configure the default MAC address for the domain by entering a new MAC address, use the `system-mac` command. The format is aaaa.bbbb.cccc.

   Also, reconfigure the same MAC address on the VLT peer switch.

   Use this command to minimize the time required for the VLT system to synchronize the default MAC address of the VLT domain on both peer switches when one peer switch reboots.
4. (Optional) When you create a VLT domain on a switch, Dell EMC Networking OS automatically assigns a unique unit ID (0 or 1) to each peer switch.

   VLT DOMAIN CONFIGURATION mode

   `unit-id {0 | 1}`

   To explicitly configure the default values on each peer switch, use the `unit-id` command.

   Configure a different unit ID (0 or 1) on each peer switch.

   Unit IDs are used for internal system operations.

   Use this command to minimize the time required for the VLT system to determine the unit ID assigned to each peer switch when one peer switch reboots.

Connecting a VLT Domain to an Attached Access Device (Switch or Server)

To connect a VLT domain to an attached access device, use the following commands.

On a VLT peer switch: To connect to an attached device, configure the same port channel ID number on each peer switch in the VLT domain.

1. Configure the same port channel to be used to connect to an attached device and enter interface configuration mode.

   CONFIGURATION mode

   `interface port-channel` *`id-number`*
2. Remove an IP address from the interface.

   INTERFACE PORT-CHANNEL mode

   `no ip address`
3. Place the interface in Layer 2 mode.

   INTERFACE PORT-CHANNEL mode

   `switchport`
4. Add one or more port interfaces to the port channel.

INTERFACE PORT-CHANNEL mode

`channel-member interface`

interface: specify one of the following interface types:

- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

5. Ensure that the port channel is active.

   INTERFACE PORT-CHANNEL mode

   `no shutdown`

6. Associate the port channel to the corresponding port channel in the VLT peer for the VLT connection to an attached device.

   INTERFACE PORT-CHANNEL mode

   `vlt-peer-lag port-channel id-number`

7. Repeat Steps 1 to 6 on the VLT peer switch to configure the same port channel as part of the VLT domain.
8. **On an attached switch or server**: To connect to the VLT domain and add port channels to it, configure a port channel. For an example of how to verify the port-channel configuration, refer to VLT Sample Configuration.

To configure the VLAN where a VLT peer forwards received packets over the VLTi from an adjacent VLT peer that is down, use the `peer-down-vlan` parameter. When a VLT peer with BMP reboots, untagged DHCP discover packets are sent to the peer over the VLTi. Using this configuration ensures the DHCP discover packets are forwarded to the VLAN that has the DHCP server.

Configuring a VLT VLAN Peer-Down (Optional)

To configure a VLT VLAN peer-down, use the following commands.

1. Enter VLT-domain configuration mode for a specified VLT domain.

   CONFIGURATION mode

   `vlt domain domain-id`

   The range of domain IDs is from 1 to 1000.

2. Enter the port-channel number that acts as the interconnect trunk.

   VLT DOMAIN CONFIGURATION mode

   `peer-link port-channel id-number`

3. Enter the VLAN ID number of the VLAN where the VLT forwards packets received on the VLTi from an adjacent peer that is down.

   VLT DOMAIN CONFIGURATION mode

   `peer-down-vlan vlan interface number`

Configuring Enhanced VLT (Optional)

To configure enhanced VLT (eVLT) between two VLT domains on your network, use the following procedure.

For a sample configuration, refer to eVLT Configuration Example. To set up the VLT domain, use the following commands.

1. Configure the port channel to be used for the VLT interconnect on a VLT switch and enter interface configuration mode.

   CONFIGURATION mode

   `interface port-channel id-number`

   Enter the same port-channel number configured with the `peer-link port-channel` command in the Enabling VLT and Creating a VLT Domain.

2. Add one or more port interfaces to the port channel.

   INTERFACE PORT-CHANNEL mode

   `channel-member interface`

   interface: specify one of the following interface types:

   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.

3. Enter VLT-domain configuration mode for a specified VLT domain.

CONFIGURATION mode

`vlt domain domain-id`

The range of domain IDs is from 1 to 1000.

4. Enter the port-channel number that acts as the interconnect trunk.

   VLT DOMAIN CONFIGURATION mode

   `peer-link port-channel id-number`

5. Configure the IP address of the management interface on the remote VLT peer to be used as the endpoint of the VLT backup link for sending out-of-band hello messages.

   VLT DOMAIN CONFIGURATION mode

   `back-up destination ip-address [interval seconds]`

   You can optionally specify the time interval used to send hello messages. The range is from 1 to 5 seconds.

6. When you create a VLT domain on a switch, Dell EMC Networking OS automatically creates a VLT-system MAC address used for internal system operations. To explicitly configure the default MAC address for the domain by entering a new MAC address, use the following command.

   VLT DOMAIN CONFIGURATION mode

   `system-mac mac-address mac-address`

   The format is aaaa.bbbb.cccc.

   Also reconfigure the same MAC address on the VLT peer switch.

   Use this command to minimize the time required for the VLT system to synchronize the default MAC address of the VLT domain on both peer switches when one peer switch reboots.

7. When you create a VLT domain on a switch, Dell EMC Networking OS automatically assigns a unique unit ID (0 or 1) to each peer switch. To explicitly configure the default values on each peer switch, use the following command.

   VLT DOMAIN CONFIGURATION mode

   `unit-id {0 | 1}`

   The unit IDs are used for internal system operations.

   Configure a different unit ID (0 or 1) on each peer switch.

   Use this command to minimize the time required for the VLT system to determine the unit ID assigned to each peer switch when one peer switch reboots.

8. **Configure enhanced VLT.** Configure the port channel to be used for the VLT interconnect on a VLT switch and enter interface configuration mode.

   CONFIGURATION mode

   `interface port-channel id-number`

   Enter the same port-channel number configured with the `peer-link port-channel` command in the Enabling VLT and Creating a VLT Domain.

9. Place the interface in Layer 2 mode.

   INTERFACE PORT-CHANNEL mode

   `switchport`

10. Associate the port channel to the corresponding port channel in the VLT peer for the VLT connection to an attached device.

    INTERFACE PORT-CHANNEL mode

    `vlt-peer-lag port-channel id-number`

11. Ensure that the port channel is active.

    INTERFACE PORT-CHANNEL mode

    `no shutdown`

12. **Add links to the eVLT port.** Configure a range of interfaces to bulk configure.

    CONFIGURATION mode

    `interface range {port-channel id}`

13. Enable LACP on the LAN port.

    INTERFACE mode

    `port-channel-protocol lacp`

14. Configure the LACP port channel mode.

INTERFACE mode

```
port-channel number mode [active]
```

15. Ensure that the interface is active.
    MANAGEMENT INTERFACE mode
    ```
    no shutdown
    ```
16. Enable peer routing.
    VLT DOMAIN CONFIGURATION mode
    ```
    peer-routing
    ```
    If you enable peer routing, a VLT node acts as the proxy gateway for its peer.
17. Repeat steps 1 through 16 for the VLT peer node in Domain 1.
18. Repeat steps 1 through 16 for the first VLT node in Domain 2.
19. Repeat steps 1 through 16 for the VLT peer node in Domain 2.

To verify the configuration of a VLT domain, use any of the `show` commands described in Verifying a VLT Configuration.

VLT Sample Configuration

To review a sample VLT configuration setup, study these steps.

1. Configure the VLT domain with the same ID in VLT peer 1 and VLT peer 2.
   VLT DOMAIN mode
   ```
   vlt domain domain id
   ```
2. Configure the VLTi between VLT peer 1 and VLT peer 2.
3. You can configure LACP/static LAG between the peer units (not shown).
   CONFIGURATION mode
   ```
   interface port-channel port-channel id
   ```
   NOTE: To benefit from the protocol negotiations, Dell EMC Networking recommends configuring VLTs used as facing hosts/switches with LACP. Ensure both peers use the same port channel ID.
4. Configure the peer-link port-channel in the VLT domains of each peer unit.
   INTERFACE PORTCHANNEL mode
   ```
   channel-member
   ```
5. Configure the backup link between the VLT peer units (shown in the following example).
6. Configure the peer 2 management ip/ interface ip for which connectivity is present in VLT peer 1.
   EXEC Privilege mode
   ```
   show running-config vlt
   ```
7. Configure the peer 1 management ip/ interface ip for which connectivity is present in VLT peer 1.
   EXEC mode or EXEC Privilege mode
   ```
   show interfaces interface
   ```
8. Configure the VLT links between VLT peer 1 and VLT peer 2 to the top of rack unit (shown in the following example).
9. Configure the static LAG/LACP between ports connected from VLT peer 1 and VLT peer 2 to the top of rack unit.
   EXEC Privilege mode
   ```
   show running-config entity
   ```
10. Configure the VLT peer link port channel id in VLT peer 1 and VLT peer 2.
    EXEC mode or EXEC Privilege mode
    ```
    show interfaces interface
    ```
11. In the top of rack unit, configure LACP in the physical ports.
    EXEC Privilege mode
    ```
    show running-config entity
    ```
12. Verify that VLT is running.
    EXEC mode

`show vlt brief` or `show vlt detail`

13. Verify that the VLT LAG is running in both VLT peer units.

    EXEC mode or EXEC Privilege mode

    `show interfaces` *interface*

In the following sample VLT configuration steps, VLT peer 1 is Dell-2, VLT peer 2 is Dell-4, and the ToR is S60-1.

NOTE: If you use a third-party ToR unit, Dell EMC Networking recommends using static LAGs with VLT peers to avoid potential problems if you reboot the VLT peers.

Configure the VLT domain with the same ID in VLT peer 1 and VLT peer 2.

```
Dell-2(conf)#vlt domain 5
Dell-2(conf-vlt-domain)#

Dell-4(conf)#vlt domain 5
Dell-4(conf-vlt-domain)#
```

Configure the VLTi between VLT peer 1 and VLT peer 2.

1. You can configure the LACP/static LAG between the peer units (not shown).
2. Configure the peer-link port-channel in the VLT domains of each peer unit.

```
Dell-2(conf)#interface port-channel 1
Dell-2(conf-if-po-1)#channel-member TenGigabitEthernet 1/4-7
Dell-4(conf)#interface port-channel 1
Dell-4(conf-if-po-1)#channel-member TenGigabitEthernet 1/4-7
```

Configure the backup link between the VLT peer units.

1. Configure the peer 2 management ip/ interface ip for which connectivity is present in VLT peer 1.
2. Configure the peer 1 management ip/ interface ip for which connectivity is present in VLT peer 2.

```
Dell-2#show running-config vlt
!
vlt domain 5
  peer-link port-channel 1
  back-up destination 10.11.206.58


Dell-2# show interfaces managementethernet 0/0
Internet address is 10.11.206.43/16

Dell-4#show running-config vlt
!
vlt domain 5
  peer-link port-channel 1
  back-up destination 10.11.206.43
Dell-4#
Dell-4#show running-config interface managementethernet 0/0
ip address 10.11.206.58/16
  no shutdown
```

Configure the VLT links between VLT peer 1 and VLT peer 2 to the Top of Rack unit. In the following example, port Te 1/4 in VLT peer 1 is connected to Te 1/8 of ToR and port Te 1/18 in VLT peer 2 is connected to Te 1/30 of ToR.

1. Configure the static LAG/LACP between the ports connected from VLT peer 1 and VLT peer 2 to the Top of Rack unit.
2. Configure the VLT peer link port channel id in VLT peer 1 and VLT peer 2.
3. In the Top of Rack unit, configure LACP in the physical ports (shown for VLT peer 1 only. Repeat steps for VLT peer 2. The bold `vlt-peer-lag port-channel 2` indicates that port-channel 2 is the port-channel id configured in VLT peer 2).

```
Dell-2#show running-config interface tengigabitethernet 1/4
!
interface TenGigabitEthernet 1/4
  no ip address
!
  port-channel-protocol LACP
    port-channel 2 mode active
  no shutdown
```

```
configuring VLT peer lag in VLT
Dell-2#show running-config interface port-channel 2
!
interface Port-channel 2
  no ip address
  switchport
  vlt-peer-lag port-channel 2
  no shutdown

Dell-2#show interfaces port-channel 2 brief
Codes: L - LACP Port-channel

   LAG  Mode  Status  Uptime    Ports
L  2    L2L3  up      03:33:14  Te 1/4 (Up)
```

In the ToR unit, configure LACP on the physical ports.

```
s60-1#show running-config interface tengigabitethernet 1/8
!
interface TenGigabitEthernet 1/8
  no ip address
!
  port-channel-protocol LACP
    port-channel 100 mode active
  no shutdown

s60-1#show running-config interface tengigabitethernet 1/30
!
interface TenGigabitEthernet 1/30
no ip address
!
  port-channel-protocol LACP
    port-channel 100 mode active
  no shutdown


s60-1#show running-config interface port-channel 100
!
interface Port-channel 100
  no ip address
  switchport
  no shutdown

s60-1#show interfaces port-channel 100 brief
Codes: L - LACP Port-channel

   LAG  Mode  Status  Uptime    Ports
L  100  L2    up      03:33:48  Te 1/8 (Up)
                               Te 1/30 (Up)
```

Verify VLT is up. Verify that the VLTi (ICL) link, backup link connectivity (heartbeat status), and VLT peer link (peer chassis) are all up.

```
DellEMC#show vlt br
VLT Domain Brief
------------------
 Domain ID                          : 1
 Role                                    : Secondary
 Role Priority                      : 32768
 ICL Link Status                    : Up
 HeartBeat Status                   : Up
 VLT Peer Status                    : Up
 Version                            : 6(3)
 Local System MAC address           : 00:01:e8:8a:e9:91
 Remote System MAC address          : 00:01:e8:8a:e9:76
 Remote system version              : 6(3)
 Delay-Restore timer                : 90 seconds

 Delay-Restore Abort Threshold      : 60 seconds
 Peer-Routing                       : Disabled
 Peer-Routing-Timeout timer         : 0 seconds
```

```
 Multicast peer-routing timeout         : 150 seconds
 DellEMC#
```

Verify that the VLT LAG is up in VLT peer unit.

```
Dell-2#show interfaces port-channel 2 brief
Codes: L - LACP Port-channel

  LAG  Mode  Status  Uptime    Ports
L 2    L2L3  up      03:43:24  Te 1/4 (Up)

Dell-4#show interfaces port-channel 2 brief
Codes: L - LACP Port-channel

  LAG  Mode  Status  Uptime    Ports
L 2    L2L3  up      03:33:31  Te 1/18 (Up)
```

PVST+ Configuration

PVST+ is supported in a VLT domain.

Before you configure VLT on peer switches, configure PVST+ in the network. PVST+ is required for initial loop prevention during the VLT startup phase. You may also use PVST+ for loop prevention in the network outside of the VLT port channel.

Run PVST+ on both VLT peer switches. A PVST+ instance is created for every VLAN configured in the system. PVST+ instances running in the Primary Peer control the VLT-LAGs on both Primary and Secondary peers. Only the Primary VLT switch determines the PVST+ roles and states on VLT ports and ensures that the VLT interconnect link is never blocked. The PVST+ instance in Primary peer sends the role/state of VLT-LAGs for all VLANs to the Secondary peer. The Secondary peer uses this information to program the hardware. The PVST+ instance running in Secondary peer does not control the VLT-LAGs.

Dell EMC Networking recommends configuring the primary VLT peer as the primary root device for all the configured PVST+ Instances and configuring the secondary VLT peer as the secondary root device for all the configured PVST+ Instances.

Sample PVST+ Configuration

The following examples show the PVST+ configuration that you must perform on each peer switch to prevent forwarding loops.

Configure PVST+ on VLT Peers to Prevent Forwarding Loops (VLT Peer 1)

```
Dell_VLTpeer1(conf)#protocol spanning-tree pvst
Dell_VLTpeer1(conf-pvst)#no disable
Dell_VLTpeer1(conf-pvst)#vlan 1000 bridge-priority 0
```

Configure PVST+ on VLT Peers to Prevent Forwarding Loops (VLT Peer 2)

```
Dell_VLTpeer2(conf)#protocol spanning-tree pvst
Dell_VLTpeer2(conf-pvst)#no disable
Dell_VLTpeer2(conf-pvst)#vlan 1000 bridge-priority 4096
```

Configure both ends of the VLT interconnect trunk with identical PVST+ configurations. When you enable VLT, the `show spanning-tree pvst brief` command output displays VLT information.

```
DellEMC#show spanning-tree pvst vlan 1000 brief
VLAN 1000
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 0, Address 90b1.1cf4.9b79
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 0, Address 90b1.1cf4.9b79
We are the root of Vlan 1000
Configured hello time 2, max age 20, forward delay 15

Interface                                                 Designated
 Name       PortID   Prio Cost    Sts          Cost        Bridge ID          PortID
---------- -------- ---- ------  ----------- ------- -------------------- --------
Po 1       128.2    128  188     FWD(vltI)    0       0     90b1.1cf4.9b79  128.2
Po 2       128.3    128  2000    FWD(vlt)     0       0     90b1.1cf4.9b79  128.3
```

```
Te 1/10     128.230  128  2000     FWD          0        0      90b1.1cf4.9b79  128.230
Te 1/13     128.233  128  2000     FWD          0        0      90b1.1cf4.9b79  128.233

Interface
 Name       Role   PortID   Prio Cost    Sts          Cost    Link-type Edge
---------- ------ -------- ---- ------- ----------- ------- --------- ----
Po 1        Desg   128.2    128  188     FWD          0       (vltl)P2P No
Po 2        Desg   128.3    128  2000    FWD          0       (vlt) P2P No
Te 1/10     Desg   128.230  128  2000    FWD          0        P2P      Yes
Te 1/13     Desg   128.233  128  2000    FWD          0        P2P      No
DellEMC#
```

Peer Routing Configuration Example

This section provides a detailed explanation of how to configure peer routing in a VLT domain.

In the following example, devices are configured as follows:

- Access switch A1 is connected to two VLT peers (Dell-1 and Dell-2).
- The two VLT peers are connected to an upstream switch R1.
- OSPF is configured in Dell-1, Dell-2, and R1 switches.
- Dell-1 is configured as the root bridge.
- Dell-1 is configured as the VLT primary.
- As the Router ID of Dell-1 is the highest in the topology (highest loopback address of 172.17.1.1), Dell-1 is the OSPF Designated Router.
- As the Router ID of Dell-2 is the second highest in the topology (172.16.1.1), Dell-2 is the OSPF Backup Designated Router.

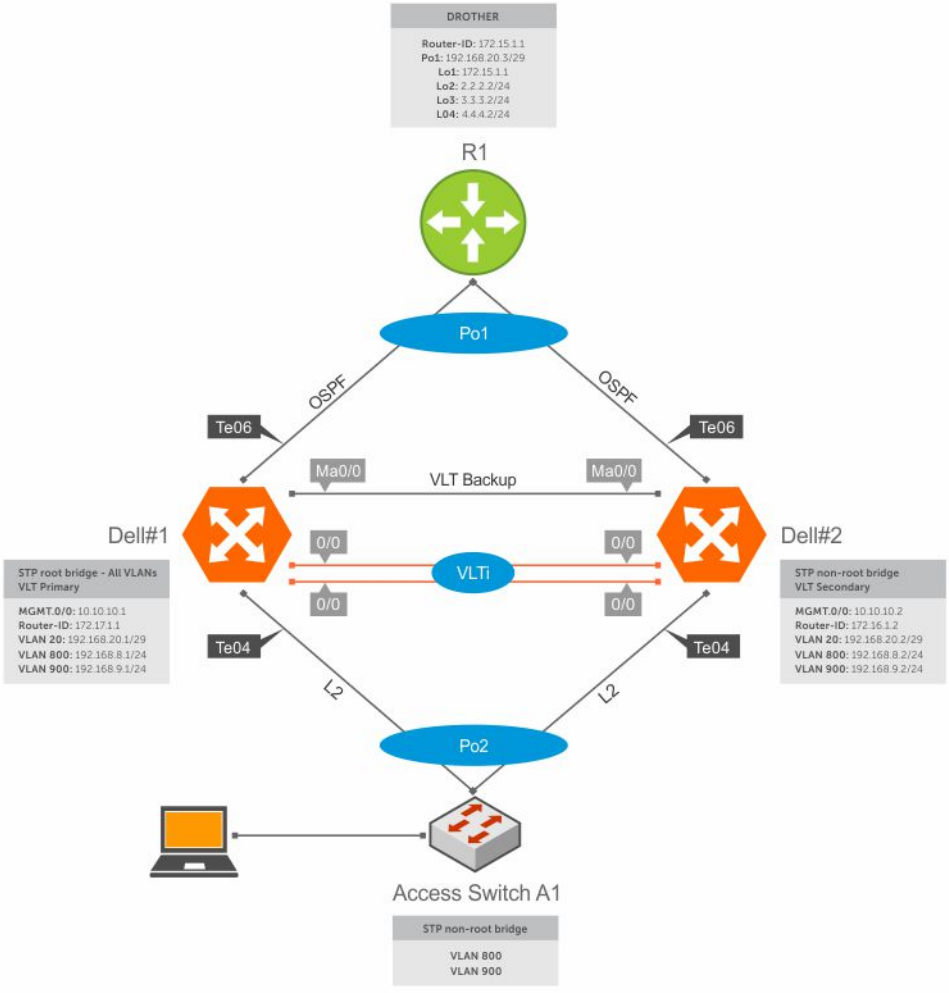


Figure 147. Peer Routing Configuration Example

Dell-1 Switch Configuration

In the following output, RSTP is enabled with a bridge priority of 0. This ensures that Dell-1 becomes the root bridge.

```
DellEMC#1#show run | find protocol
protocol spanning-tree pvst
no disable
vlan 1,20,800,900 bridge-priority 0

The following output shows the existing VLANs.

DellEMC#1#show vlan | find NUM
     NUM            Status                Description             Q  Ports
*    1               Active                                           U Po10 (Te 0/0-1)
                                                                      U Te 0/4,47
     20              Active                OSPF PEERING VLAN       U  Po1 (Te 0/6)
                                                                      V Po10 (Te 0/0-1)
     800             Active                Client-VLAN                V Po10 (Te 0/0-1)
     900             Active                Client-VLAN-2              V Po10 (Te 0/0-1)
```

The following is the configuration in interfaces:

```
DellEMC#1#sh run int ma0/0
interface ManagementEthernet 0/0
description Used_for_VLT_Keepalive
ip address 10.10.10.1/24
no shutdown
```

(The management interfaces are part of a default VRF and are isolated from the switch's data plane.) In Dell-1, te 0/0 and te 0/1 are used for VLTi.

```
DellEMC#1#sh run int te0/0
interface TenGigabitEthernet 0/0
description VLTi LINK
no ip address
no shutdown
(VLTi Physical link)
!
DellEMC#1#sh run int te0/1
interface TenGigabitEthernet 0/1
description VLTi LINK
no ip address
no shutdown
(VLTi Physical link)
```

The following example shows that te 0/0 and te 0/1 are included in port channel 10. Also note that configuration on the VLTi links does not contain the switchport command.

```
DellEMC#1#sh run int po10
interface Port-channel 10
description VLTi Port-Channel
no ip address
channel-member TenGigabitEthernet 0/0-1
no shutdown
```

Te 0/4 connects to the access switch A1.

```
DellEMC#1#sh run int te0/4
interface TenGigabitEthernet 0/4
description To_Access_Switch_A1_fa0/13
no ip address
port-channel-protocol LACP
port-channel 2 mode active
no shutdown
```

Te 0/6 connects to the uplink switch R1.

```
DellEMC#1#sh run int te0/6
interface TenGigabitEthernet 0/6
```

```
description To_CR1_fa0/13
no ip address
port-channel-protocol LACP
port-channel 1 mode active
no shutdown
```

Port channel 1 connects the uplink switch R1.

```
DellEMC#1#sh run int po1
interface Port-channel 1
description port-channel_to_R1
no ip address
switchport
vlt-peer-lag port-channel 1
no shutdown
```

Port channel 2 connects the access switch A1.

```
DellEMC#1#sh run int po2
interface Port-channel 2
description port-channel_to_access_switch_A1
no ip address
portmode hybrid
switchport
vlt-peer-lag port-channel 2
no shutdown
```

Vlan 20 is used in Dell-1, Dell-2, and R1 to form OSPF adjacency. When OSPF is converged, the routing tables in all devices are synchronized.

```
DellEMC#1#sh run int vlan 20
interface Vlan 20
description OSPF PEERING VLAN
ip address 192.168.20.1/29
untagged Port-channel 1
no shutdown
!
DellEMC#1#sh run int vlan 800
interface Vlan 800
description Client-VLAN
ip address 192.168.8.1/24
tagged Port-channel 2
no shutdown
```

The following output shows Dell-1 is configured with VLT domain 1. The `peer-link port-channel` command makes port channel 10 as the VLTi link. The peer-routing command enables peer routing between VLT peers in VLT domain 1. The IP address configured with the backup-destination command is the management IP address of the VLT peer (Dell-2).

```
DellEMC#1#sh run | find vlt
vlt domain 1
peer-link port-channel 10
back-up destination 10.10.10.2
primary-priority 4096
system-mac MAC address 90:b1:1c:f4:01:01
unit-id 0
peer routing
```

Verify if VLT on Dell-1 is functional

```
DellEMC#1#sh vlt brief
VLT Domain Brief
------------------
Domain ID:                              1
Role:                                   Primary
Role Priority:                          4096

ICL Link Status:                        Up
HeartBeat Status:                       Up
VLT Peer Status:                        Up
Local Unit Id:                          0
```

```
Version:                                    6(3)
Local System MAC address:                   90:b1:1c:f4:2c:bb
Remote System MAC address:                  90:b1:1c:f4:29:f1
Configured System MAC address:              90:b1:1c:f4:01:01
Remote system version:                      6(3)
Delay-Restore timer:                        90 seconds
Peer routing :                              Enabled
Peer routing-Timeout timer:                 0 seconds
Multicast peer routing timeout:             150 seconds
```

Verify that the heartbeat mechanism is operational

```
DellEMC#1#sh vlt backup-link
VLT Backup Link
-----------------
Destination:                                10.10.10.2
Peer HeartBeat status:                      Up
Destination VRF:                            default
HeartBeat Timer Interval:                   1
HeartBeat Timeout:                          3
UDP Port:                                   34998
HeartBeat Messages Sent:                    4
HeartBeat Messages Received:                5
```

Use the `show vlt detail` command to verify that VLT is functional and that the correct VLANs are allowed.

```
DellEMC#1#sh vlt detail

Local LAG Id    Peer LAG Id    Local Status   Peer Status   Active VLANs
------------    -----------    ------------   -----------   -------------
1               1              UP             UP            20
2               2              UP             UP            1, 800, 900
```

The following output displays the OSPF configuration in Dell-1

```
DellEMC#1#sh run | find router
router ospf 1
router-id 172.17.1.1
network 192.168.9.0/24 area 0
network 192.168.8.0/24 area 0
network 172.17.1.0/24 area 0
network 192.168.20.0/29 area 0
passive-interface default
no passive-interface vlan 20
```

While the `passive-interface default` command prevents all interfaces from establishing an OSPF neighborship, the `no passive-interface vlan 20` command enables the interface for VLAN 20, the OSPF peering VLAN, to establish OSPF adjacencies.

The following output displays that Dell-1 forms neighborship with Dell-2 and R1.

```
DellEMC#1#show ip ospf neighbor
Neighbor ID         Pri        State                  Dead Time      Address
Interface     Area
172.16.1.2          1          FULL/BDR               00:00:31       192.168.20.2
Vl 20         0
172.15.1.1          1          FULL/DROTHER           00:00:39       192.168.20.3
Vl 20         0
```

The following output displays the routes learned using OSPF.

```
DellEMC#1#show ip route ospf
  Destination    Gateway                 Dist/Metric       Last Change
  -----------    -------                 -----------       -----------
O 2.2.2.2/24     via 192.168.20.3, Vl 20 110/2             02:13:50
O 3.3.3.2/24     via 192.168.20.3, Vl 20 110/2             02:13:50
O 4.4.4.2/24     via 192.168.20.3, Vl 20 110/2             02:13:50
O 172.15.1.1/32  via 192.168.20.3, Vl 20 110/2             02:13:50
O 172.16.1.2/32  via 192.168.20.2, Vl 20 110/1             02:13:50
```

The following output displays the MAC address of all interfaces in the system. All interfaces, physical and virtual, have the same MAC address. This is the address used for peer routing.

```
DellEMC#1#show interfaces | grep Hardware
Hardware is DellEMCEth, address is 90:b1:1c:f4:2c:bd
Hardware is DellEMCEth, address is 90:b1:1c:f4:2c:bd
Hardware is DellEMCEth, address is 90:b1:1c:f4:2c:bd
Hardware is DellEMCEth, address is 90:b1:1c:f4:2c:bd
! Output truncated for brevity
```

Verify if peer routing has populated the CAM table with the correct information using the `show cam mac` command.

```
DellEMC#1#sh cam mac stack-unit 0 port-set 0
!
VlanId     Mac Address          Region          Interface
20         90:b1:1c:f4:29:f3     STATIC           Po 10
20         00:0d:bc:6e:93:00     DYNAMIC      Po 1
20         ff:ff:ff:ff:ff:ff     STATIC           00001
900         90:b1:1c:f4:29:f3     STATIC           Po 10
900         ff:ff:ff:ff:ff:ff     STATIC           00001
800         90:b1:1c:f4:29:f3     STATIC           Po 10
800         ff:ff:ff:ff:ff:ff     STATIC           00001
0             ff:ff:ff:ff:ff:ff      STATIC           00001
0            90:b1:1c:f4:2c:bd     LOCAL_DA     00001
0            90:b1:1c:f4:29:f3     LOCAL_DA     00001A
```

The above output shows that the 90:b1:1c:f4:2c:bd MAC address belongs to Dell-1. The 90:b1:1c:f4:29:f3 MAC address belongs to Dell-2. Also note that these MAC addresses are marked with LOCAL_DA. This means, these are the local destination MAC addresses used by hosts when routing is required. Packets sent to this MAC address are directly forwarded to their destinations without being sent to the peer switch.

Dell-2 Switch Configuration

In the following output, RSTP is enabled with a bridge priority of 32768, which is the second lowest in this topology. This ensures that Dell-2 becomes the root bridge if Dell-1 fails.

```
DellEMC#2#sh run | find protocol
protocol spanning-tree pvst
no disable
vlan 1,20,800,900 bridge-priority 32768
```

The following output shows the existing VLANs.

```
DellEMC#1#show vlan | find NUM
     NUM      Status         Description                Q Ports
*    1        Active                                    U Po10 (Te 0/0-1)
     20       Active         OSPF PEERING VLAN          U Po1 (Te 0/6)
                                                        V Po10 (Te 0/0-1)
     800      Active         Client-VLAN                V Po10 (Te 0/0-1)
                                 U Te 0/4
     900      Active         Client-VLAN-2              V Po10 (Te 0/0-1)
```

Configuration on Interfaces

```
DellEMC#2#sh run int ma0/0
interface ManagementEthernet 0/0
description Used_for_VLT_Keepalive
ip address 10.10.10.2/24
no shutdown
```

In Dell-2, te 0/0 and te 0/1 are used for VLTi.

In Dell-2, tw 1/1 and tw 1/2 are used for VLTi.

```
DellEMC#2#sh run int te0/0
interface TenGigabitEthernet 0/0
description VLTi LINK
no ip address
```

```
no shutdown
!
DellEMC#1#sh run int te0/1
interface TenGigabitEthernet 0/1
description VLTi LINK
no ip address
no shutdown
```

The following example shows that te 0/0 and te 0/1 are included in port channel 10. Also note that configuration on the VLTi links does not contain the switchport command.

```
Dell-2#sh run int po10
interface Port-channel 10
description VLTi Port-Channel
no ip address
channel-member TenGigabitEthernet 0/0-1
no shutdown
```

Te 0/4 connects to the access switch A1.

```
Dell-2#sh run int te0/4
interface TenGigabitEthernet 0/4
description To_Access_Switch_A1_fa0/13
no ip address
port-channel-protocol LACP
port-channel 2 mode active
no shutdown
```

Te 0/6 connects to the uplink switch R1.

```
Dell-2#sh run int te0/6
interface TenGigabitEthernet 0/6
description To_CR1_fa0/13
no ip address
port-channel-protocol LACP
port-channel 1 mode active
no shutdown
```

Port channel 1 connects the uplink switch R1.

```
Dell-2#sh run int po1
interface Port-channel 1
description port-channel_to_R1
no ip address
switchport
vlt-peer-lag port-channel 1
no shutdown
```

Port channel 2 connects the access switch A1.

```
Dell-2#sh run int po2
interface Port-channel 2
description port-channel_to_access_switch_A1
no ip address
portmode hybrid
switchport
vlt-peer-lag port-channel 2
no shutdown
```

Vlan 20 is used in Dell-1, Dell-2, and R1 to form OSPF adjacency. When OSPF is converged, the routing tables in all devices are synchronized.

```
Dell-2#sh run int vlan 20
interface Vlan 20
description OSPF PEERING VLAN
ip address 192.168.20.2/29
untagged Port-channel 1
no shutdown
!
Dell-2#sh run int vlan 800
```

```
interface Vlan 800
description Client-VLAN
ip address 192.168.8.2/24
tagged Port-channel 2
no shutdown
```

The following output shows Dell-2 is configured with VLT domain 1. The peer-link port-channel command makes port channel 10 as the VLTi link. The `peer-routing` command enables peer routing between VLT peers in VLT domain 1. The IP address configured with the backup-destination command is the management IP address of the VLT peer (Dell-1). A priority value of 55000 makes Dell-2 as the secondary VLT peer.

```
Dell-2#sh run | find vlt
vlt domain 1
peer-link port-channel 10
back-up destination 10.10.10.1
primary-priority 55000
system-mac MAC address 90:b1:1c:f4:01:01
unit-id 0
peer routing
```

Verify if VLT on Dell-1 is functional

```
Dell-2#sh vlt brief
VLT Domain Brief
------------------
Domain ID:                          1
Role:                               Secondary
Role Priority:                      55000

ICL Link Status:                    Up
HeartBeat Status:                   Up
VLT Peer Status:                    Up
Local Unit Id:                      1
Version:                            6(3)
Local System MAC address:           90:b1:1c:f4:29:f1
Remote System MAC address:          90:b1:1c:f4:2c:bb
Configured System MAC address:      90:b1:1c:f4:01:01
Remote system version:              6(3)
Delay-Restore timer:                90 seconds
Peer routing :                      Enabled
Peer routing-Timeout timer:         0 seconds
Multicast peer routing timeout:     150 seconds
```

Verify if the heartbeat mechanism is operational on Dell-2

```
Dell-2#sh vlt backup-link
VLT Backup Link
-----------------
Destination:                         10.10.10.1
Peer HeartBeat status:               Up
Destination VRF:                     default
HeartBeat Timer Interval:            1
HeartBeat Timeout:                   3
UDP Port:                            34998
HeartBeat Messages Sent:             8
HeartBeat Messages Received:         8
```

Use the show vlt detail command to verify that VLT is functional and that the correct VLANs are allowed.

```
Dell-2#sh vlt detail

Local LAG Id  Peer LAG Id    Local Status  Peer Status  Active VLANs
------------  -----------    ------------  -----------  -------------
1             1              UP            UP           20
2             2              UP            UP           1, 800, 900
```

The following output displays the OSPF configuration in Dell-2

```
Dell-2#sh run | find router
router ospf 1
```

```
router-id 172.17.1.2
network 192.168.8.0/24 area 0
network 192.168.9.0/24 area 0
network 172.16.1.0/24 area 0
network 192.168.20.0/29 area 0
passive-interface default
no passive-interface vlan 20
```

While the passive-interface default command prevents all interfaces from establishing an OSPF neighborship, the no passive-interface vlan 20 command allows the interface for VLAN 20, the OSPF peering VLAN, to establish OSPF adjacencies.

The following output displays that Dell-1 forms neighborship with Dell-2 and R1.

```
Dell-2#show ip ospf neighbor
Neighbor ID     Pri        State              Dead Time      Address         Interface
 Area
172.17.1.1      1          FULL/DR            00:00:31       192.168.20.1    Vl 20
  0
172.15.1.1      1          FULL/DROTHER       00:00:33       192.168.20.3    Vl 20
  0
```

The following output displays the routes learned using OSPF. Dell-2 also learns the routes to the loopback addresses on R1 through OSPF.

```
Dell-2#show ip route ospf
  Destination     Gateway                 Dist/Metric       Last Change
  -----------     -------                 -----------       -----------
O 2.2.2.2/24      via 192.168.20.3, Vl 20 110/2             02:15:25
O 3.3.3.2/24      via 192.168.20.3, Vl 20 110/2             02:15:25
O 4.4.4.2/24      via 192.168.20.3, Vl 20 110/2             02:15:25
O 172.15.1.1/32   via 192.168.20.3, Vl 20 110/2             02:15:25
O 172.16.1.2/32   via 192.168.20.1, Vl 20 110/1             02:15:25
```

The following output displays the MAC address of all interfaces in the system. All interfaces, physical and virtual, have the same MAC address. This is the address used for peer routing.

```
DellEMC#2#show interfaces | grep Hardware
Hardware is DellEMCEth, address is 90:b1:1c:f4:29:f3
Hardware is DellEMCEth, address is 90:b1:1c:f4:29:f3
Hardware is DellEMCEth, address is 90:b1:1c:f4:29:f3
Hardware is DellEMCEth, address is 90:b1:1c:f4:29:f3
(Output truncated for brevity)
```

Verify if peer routing has populated the CAM table with the correct information using the `show cam mac` command.

```
DellEMC#2#sh cam mac stack-unit 0 port-set 0
VlanId Mac Address       Region    Interface
20   90:b1:1c:f4:29:f3 STATIC   Po 10
20     00:0d:bc:6e:93:00 DYNAMIC   Po 1
20     ff:ff:ff:ff:ff:ff STATIC    00001
900    90:b1:1c:f4:29:f3 STATIC    Po 10
900    ff:ff:ff:ff:ff:ff STATIC    00001
800    90:b1:1c:f4:29:f3 STATIC    Po 10
800    ff:ff:ff:ff:ff:ff STATIC    00001
0      ff:ff:ff:ff:ff:ff STATIC    00001
0    90:b1:1c:f4:2c:bd LOCAL_DA  00001
0    90:b1:1c:f4:29:f3 LOCAL_DA  00001
```

The MAC addresses of the VLAN interfaces of both VLT peers are registered in the L2 CAM as LOCAL_DA addresses.

R1 Configuration

Configuration on Interfaces

```
R1#show run | find Loopback2
interface Loopback2
ip address 2.2.2.2 255.255.255.0
!
```

```
interface Loopback3
ip address 3.3.3.2 255.255.255.0
!
interface Loopback4
ip address 4.4.4.2 255.255.255.0

R1#show run int port-channel 1
interface Port-channel1
switchport
ip address 192.168.20.3 255.255.255.248

R1#show run | find router
router ospf 1
router-id 172.15.1.1
passive-interface default
no passive-interface Port-channel1
network 2.2.2.0 0.0.0.255 area 0
network 3.3.3.0 0.0.0.255 area 0
network 4.4.4.0 0.0.0.255 area 0
(The above subnets correspond to loopback interfaces lo2, lo3 and lo4. These three
loopback interfaces are advertised to the VLT pair, DellEMC#1 and DellEMC#2)
network 172.15.1.0 0.0.0.255 area 0
network 192.168.20.0 0.0.0.7 area 0
```

```
CR1#show ip ospf neighbor
(R1 is a DROTHER)
Neighbor ID Pri State    Dead Time Address      Interface
172.16.1.2  1   FULL/BDR 00:00:31  192.168.20.2 Port-channel1
172.17.1.1  1   FULL/DR  00:00:38  192.168.20.1 Port-channel1
```

```
CR1#show ip route
(Output Truncated)
  2.0.0.0/24 is subnetted, 1 subnets
C    2.2.2.0 is directly connected, Loopback2
  3.0.0.0/24 is subnetted, 1 subnets
C    3.3.3.0 is directly connected, Loopback3
O  192.168.8.0/24 [110/2] via 192.168.20.2, 02:02:34, Port-channel1
    [110/2] via 192.168.20.1, 02:02:34, Port-channel1
(OSPF-learned route back to client subnet - VLAN 800)
  4.0.0.0/24 is subnetted, 1 subnets
C    4.4.4.0 is directly connected, Loopback4
O    192.168.9.0/24 [110/2] via 192.168.20.2, 02:02:34, Port-channel1
    [110/2] via 192.168.20.1, 02:02:34, Port-channel1
(OSPF-learned route back to client subnet #2 - VLAN 900)
    172.17.0.0/24 is subnetted, 1 subnets
O      172.17.1.1 [110/1] via 192.168.20.1, 02:02:34, Port-channel1
    172.16.0.0/24 is subnetted, 1 subnets
O      172.16.1.2 [110/1] via 192.168.20.2, 02:02:34, Port-channel1
    192.168.20.0/29 is subnetted, 1 subnets
C      192.168.20.0 is directly connected, Port-channel1
(OSPF peering VLAN)
    10.0.0.0/24 is subnetted, 1 subnets
C      10.10.10.0 is directly connected, FastEthernet0/2
```

Access Switch A1 Configurations and Verification

A1 access switch is configured to not be the STP root bridge

```
A1#sh run | be spanning
spanning-tree mode pvst
spanning-tree vlan 1,800,900 priority 61440
```

```
interface Port-channeL2
description Port-Channel_to_Dell_VLT_Te0/4
switchport trunk encapsulation dot1q
switchport mode trunk
spanning-tree portfast trunk
```

This default route is configured for testing purposes, as described in the next section. The access switch (A1) is used to generate ICMP test PINGs to a loopback interface on CR1. This default route points to DellEMC#2's VLAN 800 SVI interface. It's in place to ensure that routed test traffic has DellEMC#2's MAC address as the destination address in the Ethernet frame's header

When A1 sends a packet to R1, the VLT peers act as the default gateway for each other. If the packet reaches Dell-1, irrespective of the default gateway used, Dell-1 routes the packet to R1. Similarly, if the packet reaches Dell-2, irrespective of the default gateway used, Dell-2 routes the packet to R1.

eVLT Configuration Example

The following example demonstrates the steps to configure enhanced VLT (eVLT) in a network.

In this example, you are configuring two domains. Domain 1 consists of Peer 1 and Peer 2; Domain 2 consists of Peer 3 and Peer 4, as shown in the following example.

In Domain 1, configure Peer 1 fist, then configure Peer 2. When that is complete, perform the same steps for the peer nodes in Domain 2. The interface used in this example is TenGigabitEthernet.

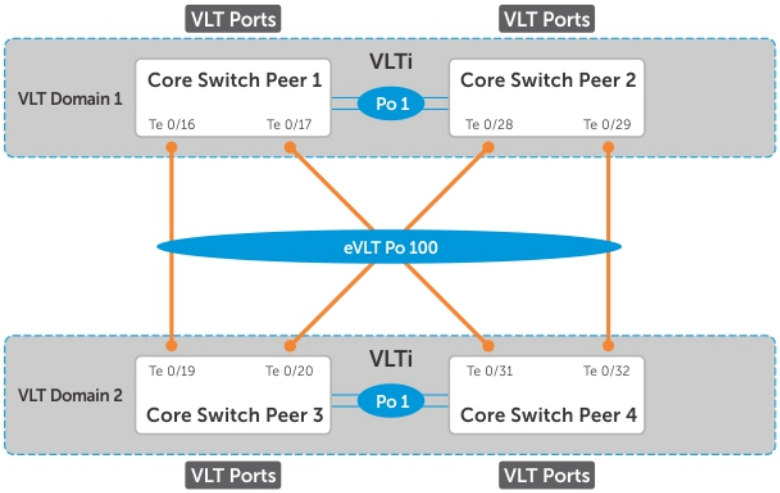


Figure 148. eVLT Configuration Example

eVLT Configuration Step Examples

In Domain 1, configure the VLT domain and VLTi on Peer 1.

```
Domain_1_Peer1#configure
Domain_1_Peer1(conf)#interface port-channel 1
Domain_1_Peer1(conf-if-po-1)# channel-member TenGigabitEthernet 1/8-9

Domain_1_Peer1(conf)#vlt domain 1000
Domain_1_Peer1(conf-vlt-domain)# peer-link port-channel 1
Domain_1_Peer1(conf-vlt-domain)# back-up destination 10.16.130.11
Domain_1_Peer1(conf-vlt-domain)# system-mac mac-address 00:0a:00:0a:00:0a
Domain_1_Peer1(conf-vlt-domain)# peer-routing
Domain_1_Peer1(conf-vlt-domain)# unit-id 0
```

Configure eVLT on Peer 1.

```
Domain_1_Peer1(conf)#interface port-channel 100
Domain_1_Peer1(conf-if-po-100)# switchport
Domain_1_Peer1(conf-if-po-100)# vlt-peer-lag port-channel 100
Domain_1_Peer1(conf-if-po-100)# no shutdown
```

Add links to the eVLT port-channel on Peer 1.

```
Domain_1_Peer1(conf)#interface range tengigabitethernet 1/16 - 17
Domain_1_Peer1(conf-if-range-te-1/16-17)# port-channel-protocol LACP
Domain_1_Peer1(conf-if-range-te-1/16-17)# port-channel 100 mode active
Domain_1_Peer1(conf-if-range-te-1/16-17)# no shutdown
```

Next, configure the VLT domain and VLTi on Peer 2.

```
Domain_1_Peer2#configure
Domain_1_Peer2(conf)#interface port-channel 1
Domain_1_Peer2(conf-if-po-1)# channel-member TenGigabitEthernet 1/8-9

Domain_1_Peer2(conf) #vlt domain 1000
Domain_1_Peer2(conf-vlt-domain)# peer-link port-channel 1
Domain_1_Peer2(conf-vlt-domain)# back-up destination 10.16.130.12
Domain_1_Peer2(conf-vlt-domain)# system-mac mac-address 00:0a:00:0a:00:0a
Domain_1_Peer2(conf-vlt-domain)# peer-routing
Domain_1_Peer2(conf-vlt-domain)# unit-id 1
```

Configure eVLT on Peer 2.

```
Domain_1_Peer2(conf)#interface port-channel 100
Domain_1_Peer2(conf-if-po-100)# switchport
Domain_1_Peer2(conf-if-po-100)# vlt-peer-lag port-channel 100
Domain_1_Peer2(conf-if-po-100)# no shutdown
```

Add links to the eVLT port-channel on Peer 2.

```
Domain_1_Peer2(conf)#interface range tengigabitethernet 1/28 - 29
Domain_1_Peer2(conf-if-range-te-1/28-29)# port-channel-protocol LACP
Domain_1_Peer2(conf-if-range-te-1/28-29)# port-channel 100 mode active
Domain_1_Peer2(conf-if-range-te-1/28-29)# no shutdown
```

In Domain 2, configure the VLT domain and VLTi on Peer 3.

```
Domain_2_Peer3#configure
Domain_2_Peer3(conf)#interface port-channel 1
Domain_2_Peer3(conf-if-po-1)# channel-member TenGigabitEthernet 1/8-9
Domain_1_Peer3#no shutdown
Domain_2_Peer3(conf)#vlt domain 200
Domain_2_Peer3(conf-vlt-domain)# peer-link port-channel 1
Domain_2_Peer3(conf-vlt-domain)# back-up destination 10.18.130.11
Domain_2_Peer3(conf-vlt-domain)# system-mac mac-address 00:0b:00:0b:00:0b
Domain_2_Peer3(conf-vlt-domain)# peer-routing
Domain_2_Peer3(conf-vlt-domain)# unit-id 0
```

Configure eVLT on Peer 3.

```
Domain_2_Peer3(conf)#interface port-channel 100
Domain_2_Peer3(conf-if-po-100)# switchport
Domain_2_Peer3(conf-if-po-100)# vlt-peer-lag port-channel 100
Domain_2_Peer3(conf-if-po-100)# no shutdown
```

Add links to the eVLT port-channel on Peer 3.

```
Domain_2_Peer3(conf)#interface range tengigabitethernet 1/19 - 20
Domain_2_Peer3(conf-if-range-te-1/19-20)# port-channel-protocol LACP
Domain_2_Peer3(conf-if-range-te-1/19-20)# port-channel 100 mode active
Domain_2_Peer3(conf-if-range-te-1/19-20)# no shutdown
```

Next, configure the VLT domain and VLTi on Peer 4.

```
Domain_2_Peer4#configure
Domain_2_Peer4(conf)#interface port-channel 1
Domain_2_Peer4(conf-if-po-1)# channel-member TenGigabitEthernet 1/8-9
Domain_1_Peer4#no shutdown

Domain_2_Peer4(conf)#vlt domain 200
```

```
Domain_2_Peer4(conf-vlt-domain)# peer-link port-channel 1
Domain_2_Peer4(conf-vlt-domain)# back-up destination 10.18.130.12
Domain_2_Peer4(conf-vlt-domain)# system-mac mac-address 00:0b:00:0b:00:0b
Domain_2_Peer4(conf-vlt-domain)# peer-routing
Domain_2_Peer4(conf-vlt-domain)# unit-id 1
```

Configure eVLT on Peer 4.

```
Domain_2_Peer4(conf)#interface port-channel 100
Domain_2_Peer4(conf-if-po-100)# switchport
Domain_2_Peer4(conf-if-po-100)# vlt-peer-lag port-channel 100
Domain_2_Peer4(conf-if-po-100)# no shutdown
```

Add links to the eVLT port-channel on Peer 4.

```
Domain_2_Peer4(conf)#interface range tengigabitethernet 1/31 - 32
Domain_2_Peer4(conf-if-range-te-1/31-32)# port-channel-protocol LACP
Domain_2_Peer4(conf-if-range-te-1/31-32)# port-channel 100 mode active
Domain_2_Peer4(conf-if-range-te-1/31-32)# no shutdown
```

PIM-Sparse Mode Configuration Example

The following sample configuration shows how to configure the PIM Sparse mode designated router functionality on the VLT domain with two VLT port-channels that are members of VLAN 4001.

For more information, refer to PIM-Sparse Mode Support on VLT.

Examples of Configuring PIM-Sparse Mode

The following example shows how to enable PIM multicast routing on the VLT node globally.

```
VLT_Peer1(conf)#ip multicast-routing
```

The following example shows how to enable PIM on the VLT port VLANs.

```
VLT_Peer1(conf)#interface vlan 4001
VLT_Peer1(conf-if-vl-4001)#ip address 140.0.0.1/24
VLT_Peer1(conf-if-vl-4001)#ip pim sparse-mode
VLT_Peer1(conf-if-vl-4001)#tagged port-channel 101
VLT_Peer1(conf-if-vl-4001)#tagged port-channel 102
VLT_Peer1(conf-if-vl-4001)#no shutdown
VLT_Peer1(conf-if-vl-4001)#exit
```

The following example shows how to configure the VLTi port as a static multicast router port for the VLAN.

```
VLT_Peer1(conf)#interface vlan 4001
VLT_Peer1(conf-if-vl-4001)#ip igmp snooping mrouter interface port-channel 128
VLT_Peer1(conf-if-vl-4001)#exit
VLT_Peer1(conf)#end
```

The following example shows how to repeat these steps on VLT Peer Node 2.

```
VLT_Peer2(conf)#ip multicast-routing

VLT_Peer2(conf)#interface vlan 4001
VLT_Peer2(conf-if-vl-4001)#ip address 140.0.0.2/24
VLT_Peer2(conf-if-vl-4001)#ip pim sparse-mode
VLT_Peer2(conf-if-vl-4001)#tagged port-channel 101
VLT_Peer2(conf-if-vl-4001)#tagged port-channel 102
VLT_Peer2(conf-if-vl-4001)#no shutdown

VLT_Peer2(conf-if-vl-4001)#ip igmp snooping mrouter interface port-channel 128
VLT_Peer2(conf-if-vl-4001)#exit
VLT_Peer2(conf)#end
```

Verifying a VLT Configuration

To monitor the operation or verify the configuration of a VLT domain, use any of the following `show` commands on the primary and secondary VLT switches.

- Display information on backup link operation.

  EXEC mode

  `show vlt backup-link`
- Display general status information about VLT domains currently configured on the switch.

  EXEC mode

  `show vlt brief`
- Display detailed information about the VLT-domain configuration, including local and peer port-channel IDs, local VLT switch status, and number of active VLANs on each port channel.

  EXEC mode

  `show vlt detail`
- Display the VLT peer status, role of the local VLT switch, VLT system MAC address and system priority, and the MAC address and priority of the locally-attached VLT device.

  EXEC mode

  `show vlt role`
- Display the current configuration of all VLT domains or a specified group on the switch.

  EXEC mode

  `show running-config vlt`
- Display statistics on VLT operation.

  EXEC mode

  `show vlt statistics`
- Display the RSTP configuration on a VLT peer switch, including the status of port channels used in the VLT interconnect trunk and to connect to access devices.

  EXEC mode

  `show spanning-tree rstp`
- Display the current status of a port or port-channel interface used in the VLT domain.

  EXEC mode

  `show interfaces` *`interface`*
  - *`interface`*: specify one of the following interface types:
    - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
    - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
    - For a port channel interface, enter the keywords `port-channel` then a number.

The following example shows the `show vlt backup-link` command.

```
Dell_VLTpeer1# show vlt backup-link

VLT Backup Link
-----------------
Destination:                10.11.200.18
Peer HeartBeat status:      Up
HeartBeat Timer Interval:   1
HeartBeat Timeout:          3
UDP Port:                   34998
HeartBeat Messages Sent:    1026
HeartBeat Messages Received: 1025

Dell_VLTpeer2# show vlt backup-link

VLT Backup Link
-----------------
Destination:                10.11.200.20
Peer HeartBeat status:      Up
HeartBeat Timer Interval:   1
HeartBeat Timeout:          3
```

```
UDP Port:                      34998
HeartBeat Messages Sent:       1030
HeartBeat Messages Received: 1014
```

The following example shows the `show vlt brief` command.

```
DellEMC#show vlt brief
VLT Domain Brief
------------------
 Domain ID                              : 1
 Role                                   : Secondary
 Role Priority                          : 32768
 ICL Link Status                        : Up
 HeartBeat Status                       : Up
 VLT Peer Status                        : Up
 Version                                : 6(3)
 Local System MAC address               : 00:01:e8:8a:e9:91
 Remote System MAC address              : 00:01:e8:8a:e9:76
 Remote system version                  : 6(3)
 Delay-Restore timer                    : 90 seconds

 Delay-Restore Abort Threshold          : 60 seconds
 Peer-Routing                           : Disabled
 Peer-Routing-Timeout timer             : 0 seconds
 Multicast peer-routing timeout         : 150 seconds
 DellEMC#
```

The following example shows the `show vlt detail` command.

```
Dell_VLTpeer1# show vlt detail

Local LAG Id Peer LAG Id Local Status Peer Status Active VLANs
------------ ----------- ------------ ----------- -------------
100          100         UP            UP         10, 20, 30
127          2           UP            UP        20, 30

Dell_VLTpeer2# show vlt detail

Local LAG Id Peer LAG Id Local Status Peer Status Active VLANs
------------ ----------- ------------ ----------- -------------
2            127         UP           UP          20, 30
100          100         UP           UP          10, 20, 30
```

The following example shows the `show vlt role` command.

```
Dell_VLTpeer1# show vlt role

VLT Role
----------
VLT Role:                    Primary
System MAC address:          00:01:e8:8a:df:bc
System Role Priority:        32768
Local System MAC address:    00:01:e8:8a:df:bc
Local System Role Priority: 32768

Dell_VLTpeer2# show vlt role

VLT Role
----------
VLT Role:                    Secondary
System MAC address:          00:01:e8:8a:df:bc
System Role Priority:        32768
Local System MAC address:    00:01:e8:8a:df:e6
Local System Role Priority: 32768
```

The following example shows the `show running-config vlt` command.

```
Dell_VLTpeer1# show running-config vlt
!
vlt domain 30
  peer-link port-channel 60
```

```
  back-up destination 10.11.200.18

Dell_VLTpeer2# show running-config vlt
!
vlt domain 30
  peer-link port-channel 60
  back-up destination 10.11.200.20
```

The following example shows the `show vlt statistics` command.

```
Dell_VLTpeer1# show vlt statistics

VLT Statistics
----------------
HeartBeat Messages Sent:     987
HeartBeat Messages Received: 986
ICL Hello's Sent:            148
ICL Hello's Received:        98

Dell_VLTpeer2# show vlt statistics

VLT Statistics
----------------
HeartBeat Messages Sent:     994
HeartBeat Messages Received: 978
ICL Hello's Sent:            89
ICL Hello's Received:        89
```

The following example shows the `show spanning-tree rstp` command.

The bold section displays the RSTP state of port channels in the VLT domain. Port channel 100 is used in the VLT interconnect trunk (VLTi) to connect to VLT peer2. Port channels 110, 111, and 120 are used to connect to access switches or servers (vlt).

```
Dell_VLTpeer1# show spanning-tree rstp brief

Executing IEEE compatible Spanning Tree Protocol
Root ID Priority 0, Address 0001.e88a.dff8
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID Priority 4096, Address 0001.e88a.d656
Configured hello time 2, max age 20, forward delay 15

Interface                                    Designated
Name   PortID  Prio Cost  Sts Cost   Bridge ID PortID
---------- -------- ---- ------- --------- ------- ------------------
Po 1   128.2   128 200000 DIS       800   4096   0001.e88a.d656 128.2
Po 3   128.4   128 200000 DIS       800   4096   0001.e88a.d656 128.4
Po 4   128.5   128 200000 DIS       800   4096   0001.e88a.d656 128.5
Po 100 128.101 128 800    FWD(VLTi) 800   0      0001.e88a.dff8 128.101
Po 110 128.111 128 00     FWD(vlt)  800   4096   0001.e88a.d656 128.111
Po 111 128.112 128 200000 DIS(vlt)  800   4096   0001.e88a.d656 128.112
Po 120 128.121 128 2000   FWD(vlt)  800   4096   0001.e88a.d656 128.121

Dell_VLTpeer2# show spanning-tree rstp brief

Executing IEEE compatible Spanning Tree Protocol
Root ID Priority 0, Address 0001.e88a.dff8
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID Priority 0, Address 0001.e88a.dff8
We are the root
Configured hello time 2, max age 20, forward delay 15

Interface                                    Designated
Name   PortID  Prio Cost  Sts      Cost    Bridge ID PortID
---------- -------- ---- ------- -------- - ------- -------------
Po 1   128.2   128 200000 DIS      0   0   0001.e88a.dff8 128.2
Po 3   128.4   128 200000 DIS      0   0   0001.e88a.dff8 128.4
Po 4   128.5   128 200000 DIS      0   0   0001.e88a.dff8 128.5
Po 100 128.101 128 800    FWD(VLTi)0   0   0001.e88a.dff8 128.101
Po 110 128.111 128 00     FWD(vlt) 0   0   0001.e88a.dff8 128.111
Po 111 128.112 128 200000 DIS(vlt) 0   0   0001.e88a.dff8 128.112
Po 120 128.121 128 2000   FWD(vlt) 0   0   0001.e88a.dff8 128.121
```

Additional VLT Sample Configurations

To configure VLT, configure a backup link and interconnect trunk, create a VLT domain, configure a backup link and interconnect trunk, and connect the peer switches in a VLT domain to an attached access device (switch or server).

Review the following examples of VLT configurations.

Configuring Virtual Link Trunking (VLT Peer 1)

Enable VLT and create a VLT domain with a backup-link and interconnect trunk (VLTi).

```
Dell_VLTpeer1(conf)#vlt domain 999
Dell_VLTpeer1(conf-vlt-domain)#peer-link port-channel 100
Dell_VLTpeer1(conf-vlt-domain)#back-up destination 10.11.206.35
Dell_VLTpeer1(conf-vlt-domain)#exit
```

Configure the backup link.

```
Dell_VLTpeer1(conf)#interface ManagementEthernet 0/0
Dell_VLTpeer1(conf-if-ma-0/0)#ip address 10.11.206.23/
Dell_VLTpeer1(conf-if-ma-0/0)#no shutdown
Dell_VLTpeer1(conf-if-ma-0/0)#exit
```

Configure the VLT interconnect (VLTi).

```
Dell_VLTpeer1(conf)#interface port-channel 100
Dell_VLTpeer1(conf-if-po-100)#no ip address
Dell_VLTpeer1(conf-if-po-100)#channel-member fortyGigE 1/48,52
Dell_VLTpeer1(conf-if-po-100)#no shutdown
Dell_VLTpeer1(conf-if-po-100)#exit
```

Configure the port channel to an attached device.

```
Dell_VLTpeer1(conf)#interface port-channel 110
Dell_VLTpeer1(conf-if-po-110)#no ip address
Dell_VLTpeer1(conf-if-po-110)#switchport
Dell_VLTpeer1(conf-if-po-110)#channel-member fortyGigE 1/56
Dell_VLTpeer1(conf-if-po-110)#no shutdown
Dell_VLTpeer1(conf-if-po-110)#vlt-peer-lag port-channel 110
Dell_VLTpeer1(conf-if-po-110)#end
```

Verify that the port channels used in the VLT domain are assigned to the same VLAN.

```
Dell_VLTpeer1# show vlan id 10
Codes: * - Default VLAN, G - GVRP VLANs, P - Primary, C - Community, I - Isolated
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack, H - Hyperpull tagged

   NUM Status Description Q Ports
   10  Active             U Po110(Fo 1/56)
                          T Po100(Fo 1/48,52)
```

Configuring Virtual Link Trunking (VLT Peer 2)

Enable VLT and create a VLT domain with a backup-link VLT interconnect (VLTi).

```
Dell_VLTpeer2(conf)#vlt domain 999
Dell_VLTpeer2(conf-vlt-domain)#peer-link port-channel 100
Dell_VLTpeer2(conf-vlt-domain)#back-up destination 10.11.206.23
Dell_VLTpeer2(conf-vlt-domain)#exit
```

Configure the backup link.

```
Dell_VLTpeer2(conf)#interface ManagementEthernet 0/0
Dell_VLTpeer2(conf-if-ma-0/0)#ip address 10.11.206.35/
Dell_VLTpeer2(conf-if-ma-0/0)#no shutdown
Dell_VLTpeer2(conf-if-ma-0/0)#exit
```

Configure the VLT interconnect (VLTi).

```
 Dell_VLTpeer2(conf)#interface port-channel 100
Dell_VLTpeer2(conf-if-po-100)#no ip address
Dell_VLTpeer2(conf-if-po-100)#channel-member fortyGigE 1/56,60
Dell_VLTpeer2(conf-if-po-100)#no shutdown
Dell_VLTpeer2(conf-if-po-100)#exit
```

Configure the port channel to an attached device.

```
Dell_VLTpeer2(conf)#interface port-channel 110
Dell_VLTpeer2(conf-if-po-110)#no ip address
Dell_VLTpeer2(conf-if-po-110)#switchport
Dell_VLTpeer2(conf-if-po-110)#channel-member fortyGigE 1/48
Dell_VLTpeer2(conf-if-po-110)#no shutdown
Dell_VLTpeer2(conf-if-po-110)#vlt-peer-lag port-channel 110
Dell_VLTpeer2(conf-if-po-110)#end
```

Verify that the port channels used in the VLT domain are assigned to the same VLAN.

```
Dell_VLTpeer2# show vlan id 10
Codes: * - Default VLAN, G - GVRP VLANs, P - Primary, C - Community, I - Isolated
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack, H - Hyperpull tagged

   NUM Status Description Q Ports
   10  Active             U Po110(Fo 1/48)
                          T Po100(Fo 1/56,60)
```

Verifying a Port-Channel Connection to a VLT Domain (From an Attached Access Switch)

On an access device, verify the port-channel connection to a VLT domain.

```
Dell_TORswitch(conf)# show running-config interface port-channel 11
!
interface Port-channel 11
no ip address
switchport
channel-member fortyGigE 1/48,52
no shutdown
```

Troubleshooting VLT

To help troubleshoot different VLT issues that may occur, use the following information.

NOTE: For information on VLT Failure mode timing and its impact, contact your Dell EMC Networking representative.

Table 128. Troubleshooting VLT

Description	Behavior at Peer Up	Behavior During Run Time	Action to Take
Bandwidth monitoring	A syslog error message and an SNMP trap is generated when the VLTi bandwidth usage goes above the 80% threshold and when it drops below 80%.	A syslog error message and an SNMP trap is generated when the VLTi bandwidth usage goes above its threshold.	Depending on the traffic that is received, the traffic can be offloaded inVLTi.
Domain ID mismatch	The VLT peer does not boot up. The VLTi is forced to a down state. A syslog error message and an SNMP trap are generated.	The VLT peer does not boot up. The VLTi is forced to a down state. A syslog error message and an SNMP trap are generated.	Verify the domain ID matches on both VLT peers.

Table 128. Troubleshooting VLT (continued)

Description	Behavior at Peer Up	Behavior During Run Time	Action to Take
Dell EMC Networking OS Version mismatch	A syslog error message is generated.	A syslog error message is generated.	Follow the correct upgrade procedure for the unit with the mismatched Dell EMC Networking OS version.
Remote VLT port channel status	N/A	N/A	Use the `show vlt detail` and `show vlt brief` commands to view the VLT port channel status information.
Spanning tree mismatch at global level	All VLT port channels go down on both VLT peers. A syslog error message is generated.	No traffic is passed on the port channels. A one-time informational syslog message is generated.	During run time, a loop may occur as long as the mismatch lasts. To resolve, enable RSTP on both VLT peers.
Spanning tree mismatch at port level	A syslog error message is generated.	A one-time informational syslog message is generated.	Correct the spanning tree configuration on the ports.
System MAC mismatch	A syslog error message and an SNMP trap are generated.	A syslog error message and an SNMP trap are generated.	Verify that the unit ID of VLT peers is not the same on both units and that the MAC address is the same on both units.
Unit ID mismatch	The VLT peer does not boot up. The VLTi is forced to a down state. A syslog error message is generated.	The VLT peer does not boot up. The VLTi is forced to a down state. A syslog error message is generated.	Verify the unit ID is correct on both VLT peers. Unit ID numbers must be sequential on peer units; for example, if Peer 1 is unit ID “0”, Peer 2 unit ID must be “1’.
Version ID mismatch	A syslog error message and an SNMP trap are generated.	A syslog error message and an SNMP trap are generated.	Verify the Dell EMC Networking OS software versions on the VLT peers is compatible. For more information, refer to the *Release Notes* for this release.
VLT LAG ID is not configured on one VLT peer	A syslog error message is generated. The peer with the VLT configured remains active.	A syslog error message is generated. The peer with the VLT configured remains active.	Verify the VLT LAG ID is configured correctly on both VLT peers.
VLT LAG ID mismatch	The VLT port channel is brought down. A syslog error message is generated.	The VLT port channel is brought down. A syslog error message is generated.	Perform a mismatch check after the VLT peer is established.

Reconfiguring Stacked Switches as VLT

To convert switches that have been stacked to VLT peers, use the following procedure.

1. Remove the current configuration from the switches. You will need to split the configuration up for each switch.
2. Copy the files to the flash memory of the appropriate switch.
3. Copy the files on the flash drive to the startup-config.
4. Reset the stacking ports to user ports for both switches.
5. Reload the stack and confirm the new configurations have been applied.

6. On the Secondary switch (stack-unit1), enter the command `stack-unit1 renumber 0`.
7. Confirm the reload query.
8. After reloading, confirm that VLT is enabled.
9. Confirm that the management ports are interconnected or connected to a switch that can transfer Heartbeat information.

Specifying VLT Nodes in a PVLAN

You can configure VLT peer nodes in a private VLAN (PVLAN). VLT enables redundancy without the implementation of Spanning Tree Protocol (STP), and provides a loop-free network with optimal bandwidth utilization.

Because the VLT LAG interfaces are terminated on two different nodes, PVLAN configuration of VLT VLANs and VLT LAGs are symmetrical and identical on both the VLT peers. PVLANs provide Layer 2 isolation between ports within the same VLAN. A PVLAN partitions a traditional VLAN into sub-domains identified by a primary and secondary VLAN pair. With VLT being a Layer 2 redundancy mechanism, support for configuration of VLT nodes in a PVLAN enables Layer 2 security functionalities. To achieve maximum VLT resiliency, you should configure the PVLAN IDs and mappings to be identical on both the VLT peer nodes.

The association of PVLAN with the VLT LAG must also be identical. After the VLT LAG is configured to be a member of either the primary or secondary PVLAN (which is associated with the primary), ICL becomes an automatic member of that PVLAN on both switches. This association helps the PVLAN data flow received on one VLT peer for a VLT LAG to be transmitted on that VLT LAG from the peer.

You can associate either a VLT VLAN or a VLT LAG to a PVLAN. First configure the VLT interconnect (VLTi) or a VLT LAG by using the `peer-link port-channel` *`id-number`* command or the VLT VLAN by using the `peer-link port-channel` *`id-number`* `peer-down-vlan vlan` *`interface number`* command and the `switchport` command. After you specify the VLTi link and VLT LAGs, you can associate the same port channel or LAG bundle that is a part of a VLT to a PVLAN by using the `interface` *`interface`* and `switchport mode private-vlan` commands.

When a VLTi port in trunk mode is a member of symmetric VLT PVLANs, the PVLAN packets are forwarded only if the PVLAN settings of both the VLT nodes are identical. You can configure the VLTi in trunk mode to be a member of non-VLT PVLANs if the VLTi is configured on both the peers. MAC address synchronization is performed for VLT PVLANs across peers in a VLT domain.

Keep the following points in mind when you configure VLT nodes in a PVLAN:

- Configure the VLTi link to be in trunk mode. Do not configure the VLTi link to be in access or promiscuous mode.
- You can configure a VLT LAG or port channel to be in trunk, access, or promiscuous port modes when you include the VLT LAG in a PVLAN. The VLT LAG settings must be the same on both the peers. If you configure a VLT LAG as a trunk port, you can associate that LAG to be a member of a normal VLAN or a PVLAN. If you configure a VLT LAG to be a promiscuous port, you can configure that LAG to be a member of PVLAN only. If you configure a VLT LAG to be in access port mode, you can add that LAG to be a member of the secondary VLAN only.
- ARP entries are synchronized even when a mismatch occurs in the PVLAN mode of a VLT LAG.

Any VLAN that contains at least one VLT port as a member is treated as a VLT VLAN. You can configure a VLT VLAN to be a primary, secondary, or a normal VLAN. However, the VLT VLAN configuration must be symmetrical across peers. If the VLT LAG is tagged to any one of the primary or secondary VLANs of a PVLAN, then both the primary and secondary VLANs are considered as VLT VLANs.

If you add an ICL or VLTi link as a member of a primary VLAN, the ICL becomes a part of the primary VLAN and its associated secondary VLANs, similar to the behavior for normal trunk ports. VLAN parity is not validated if you associate an ICL to a PVLAN. Similarly, if you dissociate an ICL from a PVLAN, although the PVLAN parity exists, ICL is removed from that PVLAN.

Association of VLTi as a Member of a PVLAN

If a VLAN is configured as a non-VLT VLAN on both the peers, the VLTi link is made a member of that VLAN if the VLTi link is configured as a PVLAN or normal VLAN on both the peers. If a PVLAN is configured as a VLT VLAN on one peer and a non-VLT VLAN on another peer, the VLTi is added as a member of that VLAN by verifying the PVLAN parity on both the peers. In such a case, if a PVLAN is present as a VLT PVLAN on at least one of the peers, then symmetric configuration of the PVLAN is validated to cause the VLTi to be a member of that VLAN. Whenever a change in the VLAN mode on one of the peers occurs, the information is synchronized with the other peer and VLTi is either added or removed from the VLAN based on the validation of the VLAN parity.

For VLT VLANs, the association between primary VLAN and secondary VLANs is examined on both the peers. Only if the association is identical on both the peers, VLTi is configured as a member of those VLANs. This behavior is because of security

functionalities in a PVLAN. For example, if a VLAN is a primary VLT VLAN on one peer and not a primary VLT VLAN on the other peer, VLTi is not made a part of that VLAN.

MAC Synchronization for VLT Nodes in a PVLAN

For the MAC addresses that are learned on non-VLT ports, MAC address synchronization is performed with the other peer if the VLTi (ICL) link is part of the same VLAN as the non-VLT port. For MAC addresses that are learned on VLT ports, the VLT LAG mode of operation and the primary to secondary association of the VLT nodes is determined on both the VLT peers. MAC synchronization is performed for the VLT LAGs only if the VLT LAG and primary-secondary VLT peer mapping are symmetrical.

The PVLAN mode of VLT LAGs on one peer is validated against the PVLAN mode of VLT LAGs on the other peer. MAC addresses that are learned on that VLT LAG are synchronized between the peers only if the PVLAN mode on both the peers is identical. For example, if the MAC address is learned on a VLT LAG and the VLAN is a primary VLT VLAN on one peer and not a primary VLT VLAN on the other peer, MAC synchronization does not occur.

Whenever a change occurs in the VLAN mode of one of the peers, this modification is synchronized with the other peers. Depending on the validation mechanism that is initiated for MAC synchronization of VLT peers, MAC addresses learned on a particular VLAN are either synchronized with the other peers, or MAC addresses synchronized from the other peers on the same VLAN are deleted. This method of processing occurs when the PVLAN mode of VLT LAGs is modified.

Because the VLTi link is only a member of symmetric VLT PVLANs, MAC synchronization takes place directly based on the membership of the VLTi link in a VLAN and the VLT LAG mode.

PVLAN Operations When One VLT Peer is Down

When a VLT port moves to the Admin or Operationally Down state on only one of the VLT nodes, the VLT Lag is still considered to be up. All the PVLAN MAC entries that correspond to the operationally down VLT LAG are maintained as synchronized entries in the device. These MAC entries are removed when the peer VLT LAG also becomes inactive or a change in PVLAN configuration occurs.

PVLAN Operations When a VLT Peer is Restarted

When the VLT peer node is rebooted, the VLAN membership of the VLTi link is preserved and when the peer node comes back online, a verification is performed with the newly received PVLAN configuration from the peer. If any differences are identified, the VLTi link is either added or removed from the VLAN. When the peer node restarts and returns online, all the PVLAN configurations are exchanged across the peers. Based on the information received from the peer, a bulk synchronization of MAC addresses that belong to spanned PVLANs is performed.

During the booting phase or when the ICL link attempts to come up, a system logging message is recorded if VLT PVLAN mismatches, PVLAN mode mismatches, PVLAN association mismatches, or PVLAN port mode mismatches occur. Also, you can view these discrepancies if any occur by using the `show vlt mismatch` command.

Interoperation of VLT Nodes in a PVLAN with ARP Requests

When an ARP request is received, and the following conditions are applicable, the IP stack performs certain operations.

- The VLAN on which the ARP request is received is a secondary VLAN (community or isolated VLAN).
- Layer 3 communication between secondary VLANs in a private VLAN is enabled by using the `ip local-proxy-arp` command in INTERFACE VLAN configuration mode.
- The ARP request is not received on the ICL

Under such conditions, the IP stack performs the following operations:

- The ARP reply is sent with the MAC address of the primary VLAN.
- The ARP request packet originates on the primary VLAN for the intended destination IP address.

The ARP request received on ICLs are not proxied, even if they are received with a secondary VLAN tag. This behavior change occurs because the node from which the ARP request was forwarded would have replied with its MAC address, and the current node discards the ARP request.

Scenarios for VLAN Membership and MAC Synchronization With VLT Nodes in PVLAN

The following table illustrates the association of the VLTi link and PVLANs, and the MAC synchronization of VLT nodes in a PVLAN (for various modes of operations of the VLT peers):

Table 129. VLAN Membership and MAC Synchronization With VLT Nodes in PVLAN

VLT LAG Mode		PVLAN Mode of VLT VLAN		ICL VLAN Membership	Mac Synchronization
Peer1	**Peer2**	**Peer1**	**Peer2**		
Trunk	Trunk	Primary	Primary	Yes	Yes
Trunk	Trunk	Primary	Normal	No	No
Trunk	Trunk	Normal	Normal	Yes	Yes
Promiscuous	Trunk	Primary	Primary	Yes	No
Trunk	Access	Primary	Secondary	No	No
Promiscuous	Promiscuous	Primary	Primary	Yes	Yes
Promiscuous	Access	Primary	Secondary	No	No
Promiscuous	Promiscuous	Primary	Primary	Yes	Yes
		- Secondary (Community)	- Secondary (Isolated)	No	No
Access	Access	Secondary (Community)	Secondary (Isolated)	No	No
		• Primary X	• Primary X	Yes	Yes
Promiscuous	Promiscuous	Primary	Primary	Yes	Yes
		- Secondary (Community)	- Secondary (Community)	Yes	Yes
		- Secondary (Isolated)	- Secondary (Isolated)	Yes	Yes
Promiscuous	Trunk	Primary	Normal	No	No
Promiscuous	Trunk	Primary	Primary	Yes	No
Access	Access	Secondary (Community)	Secondary (Community)	Yes	Yes
		- Primary VLAN X	- Primary VLAN X	Yes	Yes
Access	Access	Secondary (Isolated)	Secondary (Isolated)	Yes	Yes
		- Primary VLAN X	- Primary VLAN X	Yes	Yes
Access	Access	Secondary (Isolated)	Secondary (Isolated)	No	No

Table 129. VLAN Membership and MAC Synchronization With VLT Nodes in PVLAN (continued)

VLT LAG Mode		PVLAN Mode of VLT VLAN		ICL VLAN Membership	Mac Synchronization
Peer1	**Peer2**	**Peer1**	**Peer2**		
		- Primary VLAN X	- Primary VLAN Y	No	No
Access	Access	Secondary (Community)	Secondary (Community)	No	No
		- Primary VLAN Y	- Primary VLAN X	No	No
Promiscuous	Access	Primary	Secondary	No	No
Trunk	Access	Primary/Normal	Secondary	No	No

Configuring a VLT VLAN or LAG in a PVLAN

You can configure the VLT peers or nodes in a private VLAN (PVLAN). Because the VLT LAG interfaces are terminated on two different nodes, PVLAN configuration of VLT VLANs and VLT LAGs are symmetrical and identical on both the VLT peers. PVLANs provide Layer 2 isolation between ports within the same VLAN. A PVLAN partitions a traditional VLAN into subdomains identified by a primary and secondary VLAN pair. With VLT being a Layer 2 redundancy feature, support for configuration of VLT nodes in a PVLAN enables Layer 2 security functionalities to be achieved. This section describe how to configure a VLT VLAN or a VLT LAG (VLTi link) and assign that VLT interface to a PVLAN.

Creating a VLT LAG or a VLT VLAN

1. Configure the port channel for the VLT interconnect on a VLT switch and enter interface configuration mode
   CONFIGURATION mode
   `interface port-channel id-number`.
   Enter the same port-channel number configured with the `peer-link port-channel` command as described in Enabling VLT and Creating a VLT Domain.
   NOTE: To be included in the VLTi, the port channel must be in Default mode (`no switchport` or VLAN assigned).
2. Remove an IP address from the interface.
   INTERFACE PORT-CHANNEL mode
   `no ip address`
3. Add one or more port interfaces to the port channel.
   INTERFACE PORT-CHANNEL mode
   `channel-member interface`
   interface: specify one of the following interface types:
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
4. Ensure that the port channel is active.
   INTERFACE PORT-CHANNEL mode
   `no shutdown`
5. To configure the VLT interconnect, repeat Steps 1–4 on the VLT peer switch.
6. Enter VLT-domain configuration mode for a specified VLT domain.
   CONFIGURATION mode
   `vlt domain domain-id`
   The range of domain IDs is from 1 to 1000.
7. Enter the port-channel number that acts as the interconnect trunk.

VLT DOMAIN CONFIGURATION mode

`peer-link port-channel id-number`

8. (Optional) To configure a VLT LAG, enter the VLAN ID number of the VLAN where the VLT forwards packets received on the VLTi from an adjacent peer that is down.

   VLT DOMAIN CONFIGURATION mode

   `peer-link port-channel id-number peer-down-vlan vlan interface number`

Associating the VLT LAG or VLT VLAN in a PVLAN

1. Access INTERFACE mode for the port that you want to assign to a PVLAN.

   CONFIGURATION mode

   `interface interface`

2. Enable the port.

   INTERFACE mode

   `no shutdown`

3. Set the port in Layer 2 mode.

   INTERFACE mode

   `switchport`

4. Select the PVLAN mode.

   INTERFACE mode

   `switchport mode private-vlan {host | promiscuous | trunk}`

   - `host` (isolated or community VLAN port)
   - `promiscuous` (intra-VLAN communication port)
   - `trunk` (inter-switch PVLAN hub port)

5. Access INTERFACE VLAN mode for the VLAN to which you want to assign the PVLAN interfaces.

   CONFIGURATION mode

   `interface vlan vlan-id`

6. Enable the VLAN.

   INTERFACE VLAN mode

   `no shutdown`

7. To obtain maximum VLT resiliency, configure the PVLAN IDs and mappings to be identical on both the VLT peer nodes. Set the PVLAN mode of the selected VLAN to primary.

   INTERFACE VLAN mode

   `private-vlan mode primary`

8. Map secondary VLANs to the selected primary VLAN.

   INTERFACE VLAN mode

   `private-vlan mapping secondary-vlan vlan-list`

   The list of secondary VLANs can be:

   - Specified in comma-delimited (*VLAN-ID,VLAN-ID*) or hyphenated-range format (*VLAN-ID-VLAN-ID*).
   - Specified with this command even before they have been created.
   - Amended by specifying the new secondary VLAN to be added to the list.

Proxy ARP Capability on VLT Peer Nodes

The proxy ARP functionality is supported on VLT peer nodes.

A proxy ARP-enabled device answers the ARP requests that are destined for the other router in a VLT domain. The local host forwards the traffic to the proxy ARP-enabled device, which in turn transmits the packets to the destination.

By default, proxy ARP is enabled. To disable proxy ARP, use the `no proxy-arp` command in Interface mode. To re-enable proxy ARP, use the `ip proxy-arp` command in Interface mode. To view if proxy ARP is enabled on the interface, use the

`show config` command in INTERFACE mode. If it is not listed in the `show config` command output, it is enabled. Only nondefault information displays in the `show config` command output.

An ARP proxy operation is performed on the VLT peer node IP address when the peer VLT node is down. The ARP proxy stops working either when the peer routing timer expires or when the peer VLT node goes up. Layer 3 VLT provides a higher resiliency at the Layer 3 forwarding level. VLT peer routing allows you to replace VRRP with routed VLT to route the traffic from Layer 2 access nodes. With proxy ARP, hosts can resolve the MAC address of the VLT node even when VLT node is down.

If the ICL link is down when a VLT node receives an ARP request for the IP address of the VLT peer, owing to LAG-level hashing algorithm in the top-of-rack (ToR) switch, the incorrect VLT node responds to the ARP request with the peer MAC address. Proxy ARP is not performed when the ICL link is up and the ARP request the wrong VLT peer. In this case, ARP requests are tunneled to the VLT peer.

Proxy ARP supported on both VLT interfaces and non-VLT interfaces. Proxy ARP is supported on symmetric VLANs only. Proxy ARP is enabled by default. To support proxy ARP, the routing table must be symmetrically configured. For example, consider a sample topology in which you configure VLAN 100 on two VLT nodes, node 1 and node 2. You did not configure the ICL link between the two VLT nodes. Assume that the VLAN 100 IP address in node 1 is 10.1.1.1/24 and VLAN 100 IP address in node 2 is 20.1.1.2/24. In this case, if the ARP request for 20.1.1.1 reaches node 1, node 1 does not perform the ARP request for 20.1.1.2. Proxy ARP is supported only for the IP address that belongs to the received interface IP network. Proxy ARP is not supported if the ARP-requested IP address is different from the received interface IP subnet. For example, if you configure VLAN 100 and 200 on the VLT peers, and if you configured the VLAN 100 IP address as 10.1.1.0/24 and you configured the VLAN 200 IP address as 20.1.1.0/24, the proxy ARP is not performed if the VLT node receives an ARP request for 20.1.1.0/24 on VLAN 100.

Working of Proxy ARP for VLT Peer Nodes

Proxy ARP is enabled only when you enable peer routing on both the VLT peers. If you disable peer routing on one of the VLT peers, proxy ARP is not performed when the ICL link goes down. Proxy ARP is performed only when the VLT peer's MAC address is installed in the database. Proxy ARP is stopped when the VLT peer's MAC address is removed from the ARP database because of peer routing timer expiry. The source hardware address in the ARP response contains the VLT peer MAC address. Proxy ARP is supported for both unicast and broadcast ARP requests. Control packets, other than ARP requests destined for the VLT peers that reach the undesired and incorrect VLT node, are dropped if the ICL link is down. Further processing is not done on these control packets. The VLT node does not perform any action if it receives gratuitous ARP requests for the VLT peer IP address. Proxy ARP is also supported on secondary VLANs. When the ICL link or peer is down, and the ARP request for a private VLAN IP address reaches the wrong peer, the wrong peer responds to the ARP request with the peer MAC address.

The IP address of the VLT node VLAN interface is synchronized with the VLT peer over ICL when the VLT peers are up. Whenever you add or delete an IP address, this updated information is synchronized with the VLT peer. IP address synchronization occurs regardless of the VLAN administrative state. IP address addition and deletion serve as the trigger events for synchronization. When a VLAN state is down, the VLT peer might perform a proxy ARP operation for the IP addresses of that VLAN interface.

VLT nodes start performing Proxy ARP when the ICL link goes down. When the VLT peer comes up, proxy ARP stops for the peer VLT IP addresses. When the peer node is rebooted, the IP address synchronized with the peer is not flushed. Peer down events cause the proxy ARP to commence.

When a VLT node detects peer up, it does not perform proxy ARP for the peer IP addresses. IP address synchronization occurs again between the VLT peers.

Proxy ARP is enabled only if you enable peer routing on both the VLT peers. If you disable peer routing by using the `no peer-routing`command in VLT DOMAIN node, a notification is sent to the VLT peer to disable the proxy ARP. If you disable peer routing when ICL link is down, a notification is not sent to the VLT peer and in such a case, the VLT peer does not disable the proxy ARP operation.

When you remove the VLT domain on one of the VLT nodes, the peer routing configuration removal is notified to the peer. In this case, the VLT peer node disables the proxy ARP. When you remove the ICL link on one of the VLT nodes using the `no peer-link` command, the ICL down event is triggered on the other VLT node, which in turn starts the proxy ARP application. The VLT node, where the ICL link is deleted, flushes the peer IP addresses and does not perform proxy ARP for the additional LAG hashed ARP requests.

VLT Nodes as Rendezvous Points for Multicast Resiliency

You can configure VLT peer nodes as rendezvous points (RPs) in a Protocol Independent Multicast (PIM) domain.

PIM uses a VLT node as the RP to distribute multicast traffic to a multicast group. Messages to join the multicast group (Join messages) and data are sent towards the RP, so that receivers can discover who the senders are and begin receiving traffic destined for the multicast group.

To enable an explicit multicast routing table synchronization method for VLT nodes, you can configure VLT nodes as RPs. Multicast routing needs to identify the incoming interface for each route. The PIM running on both VLT peers enables both the peers to obtain traffic from the same incoming interface.

You can configure a VLT node to be an RP using the `ip pim rp-address` command in Global Configuration mode. When you configure a VLT node as an RP, the (*, G) routes that are synchronized from the VLT peers are ignored and not downloaded to the device. For the (S, G) routes that are synchronized from the VLT peer, after the RP starts receiving multicast traffic via these routes, these (S, G) routes are considered valid and are downloaded to the device. Only (S, G) routes are used to forward the multicast traffic from the source to the receiver.

You can configure VLT nodes, which function as RP, as Multicast source discovery protocol (MSDP) peers in different domains. However, you cannot configure the VLT peers as MSDP peers in the same VLT domain. In such instances, the VLT peer does not support the RP functionality.

If the same source or RP can be accessed over both a VLT and a non-VLT VLAN, configure better metrics for the VLT VLANs. Otherwise, it is possible that one VLT node chooses a non-VLT VLAN (if the path through the VLT VLAN was not available when the route was learned) and another VLT node selects a VLT VLAN. Such a scenario can cause duplication of packets. ECMP is not supported when you configure VLT nodes as RPs.

Backup RP is not supported if the VLT peer that functions as the RP is statically configured. With static RP configuration, if the RP reboots, it can handle new clients only after it comes back online. Until the RP returns to the active state, the VLT peer forwards the packets for the already logged-in clients. To enable the VLT peer node to retain the synchronized multicast routes or synchronized multicast outgoing interface (OIF) maps after a peer node failure, use the timeout value that you configured using the `multicast peer-routing timeout` *`value`* command. You can configure an optimal time for a VLT node to retain synced multicast routes or synced multicast outgoing interface (OIF), after a VLT peer node failure, using the `multicast peer-routing-timeout` command in VLT DOMAIN mode. Using the bootstrap router (BSR) mechanism, you can configure both the VLT nodes in a VLT domain as the candidate RP for the same group range. When an RP fails, the VLT peer automatically takes over the role of the RP. This phenomenon enables resiliency by the PIM BSR protocol.

Configuring VLAN-Stack over VLT

To configure VLAN-stack over VLT, follow these steps.

1. Configure the VLT LAG as VLAN-Stack access or Trunk mode on both the peers.
   INTERFACE PORT-CHANNEL mode
   `vlan-stack {access | trunk}`
2. Configure VLAN as VLAN-stack compatible on both the peers.
   INTERFACE VLAN mode
   `vlan-stack compatible`
3. Add the VLT LAG as a member to the VLAN-stack on both the peers.
   INTERFACE VLAN mode
   `member port-channel` *`port–channel ID`*
4. Verify the VLAN-stack configurations.
   EXEC Privilege
   `show running-config`

Sample configuration of VLAN-stack over VLT (Peer 1)

Configure the VLT domain

```
DellEMC(conf)#vlt domain 1
```

```
DellEMC(conf-vlt-domain)#peer-link port-channel 1
DellEMC(conf-vlt-domain)#back-up destination 10.16.151.116
DellEMC(conf-vlt-domain)#primary-priority 100
DellEMC(conf-vlt-domain)#system-mac mac-address 00:00:00:11:11:11
DellEMC(conf-vlt-domain)#unit-id 0
DellEMC(conf-vlt-domain)#

DellEMC#show running-config vlt
!
vlt domain 1
 peer-link port-channel 1
 back-up destination 10.16.151.116
 primary-priority 100
 system-mac mac-address 00:00:00:11:11:11
 unit-id 0
DellEMC#
```

Configure the VLT LAG as VLAN-Stack Access or Trunk Port

```
DellEMC(conf)#interface port-channel 10
DellEMC(conf-if-po-10)#switchport
DellEMC(conf-if-po-10)#vlt-peer-lag port-channel 10
DellEMC(conf-if-po-10)#vlan-stack access
DellEMC(conf-if-po-10)#no shutdown

DellEMC#show running-config interface port-channel 10
!
interface Port-channel 10
 no ip address
 switchport
 vlan-stack access
  vlt-peer-lag port-channel 10
 no shutdown
DellEMC#

DellEMC(conf)#interface port-channel 20
DellEMC(conf-if-po-20)#switchport
DellEMC(conf-if-po-20)#vlt-peer-lag port-channel 20
DellEMC(conf-if-po-20)#vlan-stack trunk
DellEMC(conf-if-po-20)#no shutdown

DellEMC#show running-config interface port-channel 20
!
interface Port-channel 20
 no ip address
 switchport
 vlan-stack trunk
 vlt-peer-lag port-channel 20
 no shutdown
DellEMC#
```

Configure the VLAN as a VLAN-Stack VLAN and add the VLT LAG as Members to the VLAN

```
DellEMC(conf)#interface vlan 50
DellEMC(conf-if-vl-50)#vlan-stack compatible
DellEMC(conf-if-vl-50-stack)#member port-channel 10
DellEMC(conf-if-vl-50-stack)#member port-channel 20

DellEMC#show running-config interface vlan 50
!
interface Vlan 50
 vlan-stack compatible
 member Port-channel 10,20
 shutdown
DellEMC#
```

Verify that the Port Channels used in the VLT Domain are Assigned to the VLAN-Stack VLAN

```
DellEMC#show vlan id 50
```

```
Codes: * - Default VLAN, G - GVRP VLANs, R - Remote Port Mirroring VLANs, P - Primary, C
- Community, I - Isolated
       O - Openflow
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   o - OpenFlow untagged, O - OpenFlow tagged
   G - GVRP tagged, M - Vlan-stack
   i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

    NUM    Status    Description                    Q Ports
    50     Active                                   M Po10(Te 1/8)
                                                    M Po20(Te 1/12)
                                                    V Po1(Te 1/30-32)
DellEMC#
```

Sample Configuration of VLAN-Stack Over VLT (Peer 2)

Configure the VLT domain

```
DellEMC(conf)#vlt domain 1
DellEMC(conf-vlt-domain)#peer-link port-channel 1
DellEMC(conf-vlt-domain)#back-up destination 10.16.151.115
DellEMC(conf-vlt-domain)#system-mac mac-address 00:00:00:11:11:11
DellEMC(conf-vlt-domain)#unit-id 1
DellEMC(conf-vlt-domain)#

DellEMC#show running-config vlt
vlt domain 1
 peer-link port-channel 1
 back-up destination 10.16.151.115
 system-mac mac-address 00:00:00:11:11:11
 unit-id 1
DellEMC#
```

Configure the VLT LAG as VLAN-Stack Access or Trunk Port

```
DellEMC(conf)#interface port-channel 10
DellEMC(conf-if-po-10)#switchport
DellEMC(conf-if-po-10)#vlt-peer-lag port-channel 10
DellEMC(conf-if-po-10)#vlan-stack access
DellEMC(conf-if-po-10)#no shutdown

DellEMC#show running-config interface port-channel 10
!
interface Port-channel 10
 no ip address
 switchport
 vlan-stack access
 vlt-peer-lag port-channel 10
 no shutdown
DellEMC#

DellEMC(conf)#interface port-channel 20
DellEMC(conf-if-po-20)#switchport
DellEMC(conf-if-po-20)#vlt-peer-lag port-channel 20
DellEMC(conf-if-po-20)#vlan-stack trunk
DellEMC(conf-if-po-20)#no shutdown

DellEMC#show running-config interface port-channel 20
!
interface Port-channel 20
 no ip address
 switchport
 vlan-stack trunk
 vlt-peer-lag port-channel 20
 no shutdown
DellEMC#
```

Configure the VLAN as a VLAN-Stack VLAN and add the VLT LAG as members to the VLAN

```
DellEMC(conf)#interface vlan 50
DellEMC(conf-if-vl-50)#vlan-stack compatible
```

```
DellEMC(conf-if-vl-50-stack)#member port-channel 10
DellEMC(conf-if-vl-50-stack)#member port-channel 20
DellEMC(conf-if-vl-50-stack)#

DellEMC#show running-config interface vlan 50
!
interface Vlan 50
 vlan-stack compatible
  member Port-channel 10,20
  shutdown
DellEMC#
```

Verify that the Port Channels used in the VLT Domain are Assigned to the VLAN-Stack VLAN

```
DellEMC#show vlan id 50

Codes: * - Default VLAN, G - GVRP VLANs, R - Remote Port Mirroring VLANs, P - Primary, C
- Community, I - Isolated
       O - Openflow
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   o - OpenFlow untagged, O - OpenFlow tagged
   G - GVRP tagged, M - Vlan-stack
   i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

    NUM    Status    Description                    Q Ports
    50     Active                                   M Po10(Te 1/8)
                                                    M Po20(Te 1/20)
                                                    V Po1(Te 1/30-32)
DellEMC#
```

IPv6 Peer Routing in VLT Domains Overview

VLT enables the physical links between two devices that are called VLT nodes or peers, and within a VLT domain, to be considered as a single logical link to external devices that are connected using LAG bundles to both the VLT peers. This capability enables redundancy without the implementation of Spanning tree protocol (STP), thereby providing a loop-free network with optimal bandwidth utilization.

IPv6 peer routing is supported on all the platforms that are compatible with IPv6 routing and support VLT. This functionality performs the following operations:

- Forwarding control traffic to the correct VLT node when the control traffic reaches the wrong VLT node due to hashing at the VLT LAG level on the ToR.
- Routing the data traffic which is destined to peer VLT node.
- Synchronizing neighbor entries learned on VLT VLAN interfaces between the primary and secondary node.
- Synchronizing the IP address of VLT VLAN interfaces between the VLT primary node and secondary node.
- Performing routing on behalf of peer VLT nodes for a configured time period when a peer VLT node goes down.

When you configure Layer 3 VLT peer routing using the `peer-routing` command in VLT DOMAIN mode, it applies for both IPv4 and IPv6 traffic in VLT domains. Layer 3 VLT provides a higher resiliency at the Layer 3 forwarding level. Routed VLT allows you to replace VRRP with routed VLT to route the traffic from the Layer 2 access nodes. With neighbor discovery (ND) synchronization, both the VLT nodes perform Layer 3 forwarding on behalf of each other.

The neighbor entries are typically learned by a node using neighbor solicitation (NS) and ND messages. These NS or neighbor advertisement (NA) messages can be either destined to the VLT node or to any nodes on the same network as the VLT interface. These learned neighbor entries are propagated to another VLT node so that the peer does not need to relearn the entries.

IPv6 Peer Routing

When you enable peer routing on VLT nodes, the MAC address of the peer VLT node is stored in the ternary content addressable memory (TCAM) space table of a station. If the data traffic destined to a VLT node, node1, reaches the other

VLT node, node2, owing to LAG-level hashing in the ToR switch, it is routed instead of forwarding the packet to node1. This processing occurs because of the match or hit for the entry in the TCAM of the VLT node2.

Synchronization of IPv6 ND Entries in a VLT Domain

Because the VLT nodes appear as a single unit, the ND entries learned via the VLT interface are expected to be the same on both VLT nodes. VLT V6 VLAN and neighbor discovery protocol monitor (NDPM) entries synchronization between VLT nodes is performed.

The VLT V6 VLAN information must synchronize with peer VLT node. Therefore, both the VLT nodes are aware of the VLT VLAN information associated with the peers. The CLI configuration and dynamic state changes of VLT V6 VLANs are notified to peer VLT node. The ND entries are generally learned by a node from Neighbor advertisements (NA).

ND entries synchronization scenarios:

- When you enable and configure VLT on both VLT node1 and node2, any dynamically learned ND entry in VLT node1 be synchronizes instantaneously to VLT node2 and vice-versa. The link-local address also synchronizes if learned on the VLT VLAN interface.
- During failure cases, when a VLT node goes down and comes back up all the ND entries learned via VLT interface must synchronize to the peer VLT node.

Synchronization of IPv6 ND Entries in a Non-VLT Domain

Layer 3 VLT provides a higher resiliency at the Layer 3 forwarding level. Routed VLT allows you to replace VRRP with routed VLT to route the traffic from Layer 2 access nodes. With ND synchronization, both the VLT nodes perform Layer 3 forwarding on behalf of each other. Synchronization of NDPM entries learned on non-VLT interfaces between the non-VLT nodes.

NDPM entries learned on non-VLT interfaces synchronize with the peer VLT nodes in case the ND entries are learned on spanned VLANs so that each node can complete Layer 3 forwarding on behalf of each other. Whenever you configure a VLAN on a VLT node, this information is communicated to the peer VLT node regardless of whether the VLAN configured is a VLT or a non-VLT interface. If the VLAN operational state (OSTATE) is up, dynamically learned ND entry in VLT node1 synchronizes to VLT node2.

Tunneling IPv6 ND in a VLT Domain

Tunneling an NA packet from one VLT node to its peer is required because an NA may reach the wrong VLT node instead of arriving at the destined VLT node. This may occur because of LAG hashing at the ToR switch. The tunneled NA carries some control information along with it so that the appropriate VLT node can mimic the ingress port as the VLT interface rather than pointing to VLT node's interconnecting link (ICL link).

The overall tunneling process involves the VLT nodes that are connected from the ToR through a LAG. The following illustration is a basic VLT setup, which describes the communication between VLT nodes to tunnel the NA from one VLT node to its peer.

NA messages can be sent in two scenarios:

- NA messages are almost always sent in response to an NS message from a node. In this case, the solicited NA has the destination address field set to the unicast MAC address of the initial NS sender. This solicited NA must be tunneled when they reach the wrong peer.
- Sometimes NA messages are sent by a node when its link-layer address changes. This NA message is sent as an unsolicited NA to advertise its new address and the destination address field is set to the link-local scope of all-nodes multicast address. This unsolicited NA packet does not have to be tunneled.

Consider a sample scenario in which two VLT nodes, Unit1 and Unit2, are connected in a VLT domain using an ICL or VLTi link. To the south of the VLT domain, Unit1 and Unit2 are connected to a ToR switch named Node B. Also, Unit1 is connected to another node, Node A, and Unit2 is linked to a node, Node C. When an NS traverses from Unit2 to Node B(ToR) and a corresponding NA reaches Unit1 because of LAG hashing, this NA is tunneled to Unit 2 along with some control information. The control information present in the tunneled NA packet is processed in such a way so that the ingress port is marked as the link from Node B to Unit 2 rather than pointing to ICL link through which tunneled NA arrived.

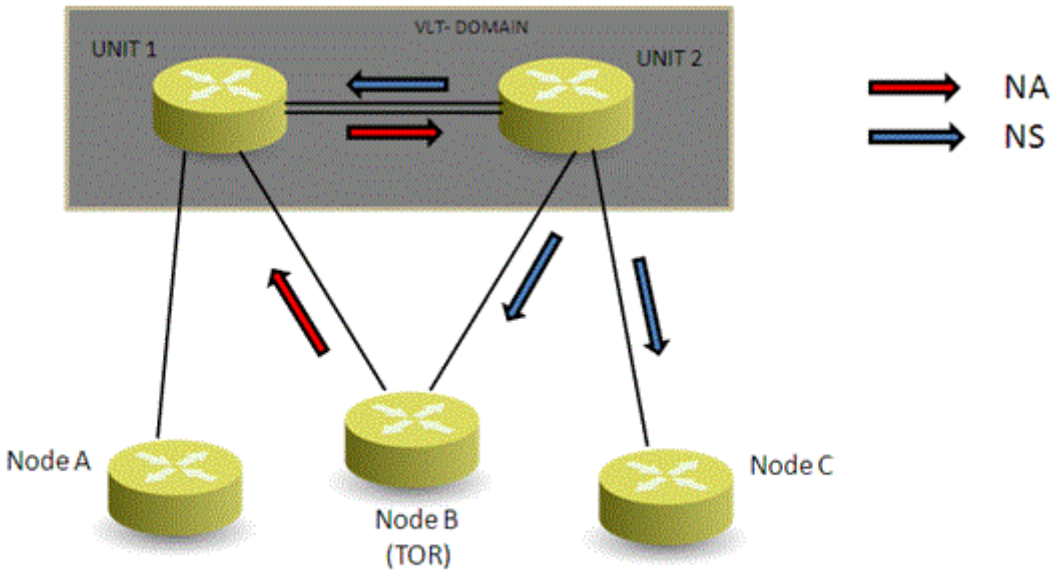


Figure 149. Sample Configuration of IPv6 Peer Routing in a VLT Domain

Sample Configuration of IPv6 Peer Routing in a VLT Domain

Consider a sample scenario as shown in the following figure in which two VLT nodes, Unit1 and Unit2, are connected in a VLT domain using an ICL or VLTi link. To the south of the VLT domain, Unit1 and Unit2 are connected to a ToR switch named Node B. Also, Unit1 is connected to another node, Node A, and Unit2 is linked to a node, Node C. The network between the ToR and the VLT nodes is Layer 2. Servers or hosts that are connected to the ToR (Node B) generate Layer 3 control/data traffic from the South or lower-end of the vertically-aligned network.

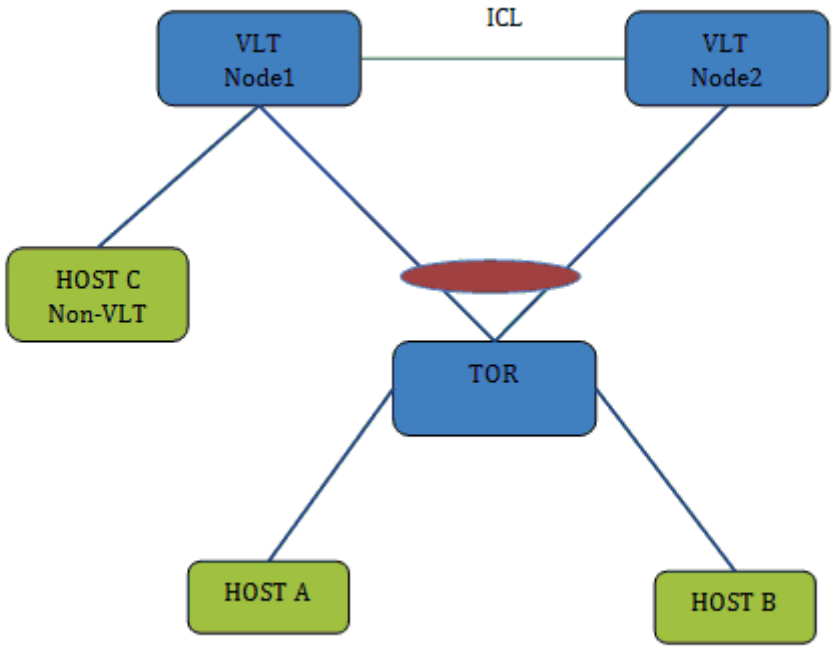


Figure 150. Sample Configuration of IPv6 Peer Routing in a VLT Domain

Neighbor Solicitation from VLT Hosts

Consider a case in which NS for VLT node1 IP reaches VLT node1 on the VLT interface and NS for VLT node1 IP reaches VLT node2 due to LAG level hashing in the ToR. When VLT node1 receives NS from VLT VLAN interface, it unicasts the NA packet on the VLT interface. When NS reaches VLT node2, it is flooded on all interfaces including ICL. When VLT node 1 receives NS on ICL, it floods the NA packet on the VLAN. If NS is unicast and if it reaches the wrong VLT peer, it is lifted to the CPU using ACL entry. Then wrong peer adds a tunnel header and forwards the packet over ICL.

Neighbor Advertisement from VLT Hosts

Consider an example in which NA for VLT node1 reaches VLT node1 on the VLT interface and NA for VLT node1 reaches VLT node2 due to LAG level hashing in ToR. When VLT node1 receives NA on VLT interface, it learns the Host MAC address on VLT interface. This learned neighbor entry is synchronized to VLT node2 as it is learned on VLT interface of Node2. If VLT node2 receives a NA packet on VLT interface which is destined to VLT node1, node 2 lifts the NA packet to CPU using an ACL entry then it adds a tunnel header to the received NA and forwards the packet to VLT node1 over ICL. When VLT node1 receives NA over ICL with tunnel header it learns the Host MAC address on VLT port channel interface. This learned neighbor entry is synchronized to VLT node2 as it is learned on VLT interface of Node2.

If NA is intended for a VLT peer and DIP is LLA of the peer, it is lifted to the CPU and tunneled to the peer. VLT nodes drop the NA packet if the NA is received over ICL without a tunneling header.

Neighbor Solicitation from Non-VLT Hosts

Consider a sample scenario in which NS for VLT node1 IP reaches VLT node1 on a non-VLT interface and NS for VLT node1 IP reaches VLT node2 on a non-VLT interface. When VLT node1 receives NS from a non-VLT interface, it unicasts the NA packet on the received interface. When NS reaches VLT node2, it floods on all interfaces including ICL. When VLT node 1 receives NS on the ICL, it floods the NA packet on the VLAN. If NS is unicast and if it reaches the wrong VLT peer, it is lifted to the CPU using the ACL entry. Then the wrong peer adds a tunnel header and forwards the packet over the ICL.

Neighbor Advertisement from Non-VLT Hosts

Consider a situation in which NA for VLT node1 reaches VLT node1 on a non-VLT interface and NA for VLT node1 reaches VLT node2 on a non-VLT interface. When VLT node1 receives NA on a VLT interface, it learns the Host MAC address on the received interface. This learned neighbor entry is synchronized to VLT node2 as it is learned on ICL. If VLT node2 receives a NA packet on non-VLT interface which is destined to VLT node1, node 2 lifts the NA packet to CPU using an ACL entry then it adds a tunnel header to the received NA and forwards the packet to VLT node1 over ICL. When VLT node1 received NA over ICL with tunnel header it learns the Host MAC address on the ICL. Host entries learned on ICL will not be synchronized to the VLT peer.

If NA is intended for VLT peer and DIP is LLA of peer, it is lifted to CPU and tunneled to the peer. VLT nodes will drop NA packet, If NA is received over ICL without tunneling header.

Traffic Destined to VLT Nodes

Hosts can send traffic to one of the VLT nodes using a global IP or Link-Local address. When the host communicates with the VLT node using LLA and traffic reaches the wrong peer due to LAG level hashing in the ToR, the wrong peer routes the packet to correct the VLT node though the destination IP is LLA. Consider a case in which traffic destined for VLT node1 reaches VLT node1 on the VLT interface and traffic destined for VLT node1 reaches VLT node2 due to LAG level hashing in the ToR.

When VLT node1 receives traffic on VLT interface, it consumes the packets and process them based on the packet type. If VLT node2 receives a packet on a VLT interface which is destined to VLT node1, it routes the packet to VLT node1 instead of switching the packet because the match that occurs for the neighbor entry in the TCAM table.

If the destination IP address is peers' link-local advertisement (LLA), the wrong VLT peer switches the traffic over ICL. This is achieved using switching egress object for peers LLA.

VLT host to North Bound traffic flow

One of the VLT peer is configured as the default gateway router on VLT hosts. If the VLT node receives Layer 3 traffic intended for the other VLT peer, it routes the traffic to next hop instead of forwarding the traffic to the VLT peer. If the neighbor entry is not present, the VLT node resolves the next hop. There may be traffic loss during the neighbor resolution period.

North-Bound to VLT host traffic flow

When a VLT node receives traffic from the north intended for the VLT host, it completes neighbor entry lookup and routes traffic to the VLT interface. If the VLT interface is not operationally up, the VLT node routes the traffic over ICL. If the neighbor entry is not present, the VLT node resolves the destination. There may be traffic loss during the neighbor resolution period.

VLT host to Non-VLT host traffic flow

When VLT node receives traffic intended to non-VLT host, it routes the traffic over non-VLT interface. If the traffic intended to non-VLT host reaches wrong VLT peer due to LAG hashing in ToR, the wrong VLT node will resolve the destination over ICL and routes the traffic over ICL. When Correct VLT node receives this routed traffic over ICL it will switch traffic to non-VLT interface.

Non-VLT host to VLT host traffic flow

When VLT node receives traffic from non-VLT host intended to VLT host, it routes the traffic to VLT interface. If VLT interface is not operationally up VLT node will route the traffic over ICL.

Non-VLT host to North Bound traffic flow

When VLT node receives traffic from non-VLT host intended to north bound with DMAC as self MAC it routes traffic to next hop. When VLT node receives traffic from non-VLT host intended to north bound with DMAC as peer MAC it will not forward the packet to VLT peer instead it will route the traffic to north bound next hop.

North Bound to Non-VLT host traffic flow

When VLT node receives traffic from north bound intended to the non-VLT host, it does neighbor entry lookup and routes traffic to VLT interface. If traffic reaches wrong VLT peer, it routes the traffic over ICL.

Non-VLT host to Non-VLT host traffic flow

When VLT node receives traffic from non-VLT host intended to the non-VLT host, it does neighbor entry lookup and routes traffic over ICL interface. If traffic reaches wrong VLT peer, it routes the traffic over ICL.

Router Solicitation

When VLT node receives router Solicitation on VLT interface/non-VLT interface it consumes the packets and will send RA back on the received interface.

VLT node will drop the RS message if it is received over ICL interface.

Router Advertisement

When VLT node receives router advertisement on VLT interface/non-VLT interface it consumes the packets.

VLT node will drop the RA message if it is received over ICL interface.

Upgrading from Releases That Do Not Support IPv6 Peer Routing

During an upgrade to Release 9.4(0.0) from earlier releases, VLT peers might contain different versions of FTOS. You must upgrade both the VLT peers to Release 9.4(0.0) to leverage the benefits of IPv6 peer routing.

Station Movement

When a host moves from VLT interface to non-VLT interface or vice versa Neighbor entry is updated and synchronized to VLT peer. When a host moves from non-VLT interface of VLT node1 to non-VLT interface of VLT node2 neighbor entry is updated and synchronized to VLT peer.

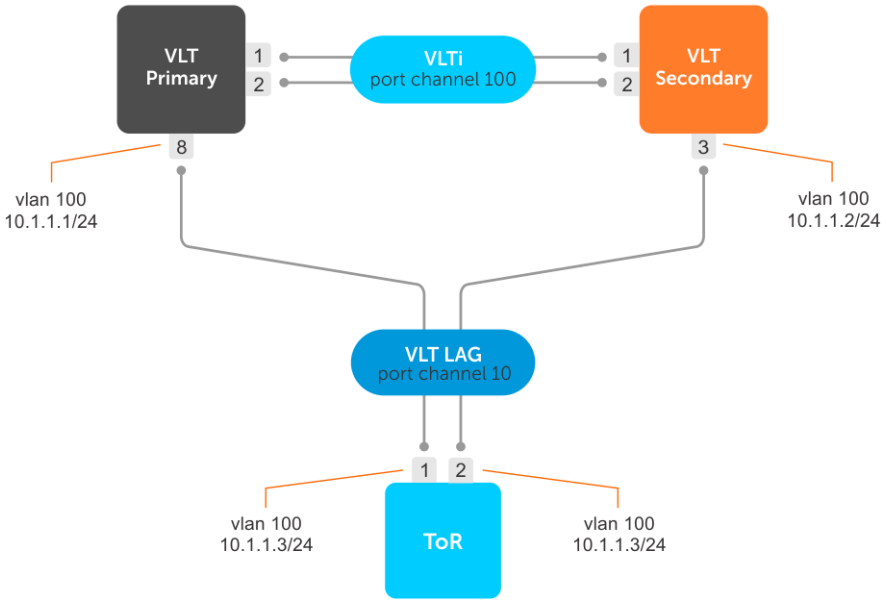
Configure BFD in VLT Domain

Dell EMC Networking OS supports Bidirectional Forwarding Detection (BFD) to detect communication failures on an interface that is a part of a VLT link aggregation group (LAG).

In VLT domain, BFD provides high availability path when there are communication failures in any one of the VLT LAG links. The VLT nodes and top of rack (ToR) use the VLT LAG links to carry the BFD packets. When any one of the VLT LAG links connected to the ToR is down, the BFD packets reach the VLT primary or secondary using the VLTi (ICL) link.

Sample BFD configuration in VLT domain

The following shows the sample configuration of BFD implementation in VLT environment:



ToR

1. Enable BFD globally.

```
TOR(conf)# bfd enable
```

2. Configure a VLT peer LAG.

```
TOR(conf)#interface tengigabitethernet 1/1
TOR(conf-if-te-1/1)#no ip address
TOR(conf-if-te-1/1)#port-channel-protocol lacp
TOR(conf-if-te-1/1)#port-channel 10 mode active
TOR(conf-if-te-1/1)#no shutdown
```

```
TOR(conf)#interface tengigabitethernet 1/2
TOR(conf-if-te-1/2)#no ip address
TOR(conf-if-te-1/2)#port-channel-protocol lacp
TOR(conf-if-te-1/2)#port-channel 10 mode active
TOR(conf-if-te-1/2)#no shutdown
```

3. Configure the port channel for the VLT interconnect on a ToR.

```
TOR(conf)# interface port-channel 10
TOR(conf-if-po-111)# no ip address
TOR(conf-if-po-111)# switchport
TOR(conf-if-po-111)# no shutdown
```

4. Configure a VLAN.

```
TOR(conf)#interface vlan 100
TOR(conf-if-vl-100)#ip address 100.1.1.3/24
TOR(conf-if-vl-100)#tagged port-channel 10
TOR(conf-if-vl-100)#arp timeout 1
TOR(conf-if-vl-100)#no shutdown
TOR(conf-if-vl-100)#exit
```

5. Enable BFD over OSPF.

```
TOR(conf)# router ospf 1
TOR(conf-router_ospf)# network 100.1.1.0/24 area 0
TOR(conf-router_ospf)# bfd all-neighbors
```

VLT Primary

1. Enable BFD globally.

```
VLT_Primary(conf)# bfd enable
```

2. Configure port channel which is used as VLTi link.

```
VLT_Primary(conf)# interface port-channel 100
VLT_Primary(conf-if-po-100)# no ip address
VLT_Primary(conf-if-po-100)# channel-member tengigabitethernet 1/1, 1/2
VLT_Primary(conf-if-po-100)# no shutdown
```

3. Enable VLT and configure a VLT domain.

```
VLT_Primary(conf)# vlt domain 100
VLT_Primary(conf-vlt-domain)# peer-link port-channel 100
VLT_Primary(conf-vlt-domain)# back-up destination 10.16.206.199
VLT_Primary(conf-vlt-domain)# peer-routing
```

4. Configure a VLT peer LAG.

```
VLT_Primary(conf)#interface tengigabitethernet 1/8
VLT_Primary(conf-if-te-1/8)#no ip address
VLT_Primary(conf-if-te-1/8)#port-channel-protocol lacp
VLT_Primary(conf-if-te-1/8)#port-channel 10 mode active
VLT_Primary(conf-if-te-1/8)#no shutdown
```

```
VLT_Primary(conf)#interface port-channel 10
VLT_Primary(conf-if-po-10)#no ip address
VLT_Primary(conf-if-po-10)#switchport
VLT_Primary(conf-if-po-10)#vlt-peer-lag port-channel 10
VLT_Primary(conf-if-po-10)#no shutdown
```

5. Configure a VLAN.

```
VLT_Primary(conf)#interface vlan 100
VLT_Primary(conf-if-vl-100)#ip address 100.1.1.1/24
VLT_Primary(conf-if-vl-100)#tagged port-channel 10
VLT_Primary(conf-if-vl-100)#no shutdown
VLT_Primary(conf-if-vl-100)#exit
```

6. Enable BFD over OSPF.

```
VLT_Primary(conf)# router ospf 1
VLT_Primary(conf-router_ospf)# network 100.1.1.0/24 area 0
VLT_Primary(conf-router_ospf)# bfd all-neighbors
```

VLT Secondary

1. Enable BFD globally.

```
VLT_Secondary(conf)# bfd enable
```

2. Configure port channel which is used as VLTi link.

```
VLT_Secondary(conf)# interface port-channel 100
VLT_Secondary(conf-if-po-100)# no ip address
VLT_Secondary(conf-if-po-100)# channel-member tengigabitethernet 1/1, 1/2
VLT_Secondary(conf-if-po-100)# no shutdown
```

3. Enable VLT and configure a VLT domain.

```
VLT_Secondary(conf)# vlt domain 100
VLT_Secondary(conf-vlt-domain)# peer-link port-channel 100
VLT_Secondary(conf-vlt-domain)# back-up destination 10.16.206.80
VLT_Secondary(conf-vlt-domain)# peer-routing
```

4. Configure a VLT peer LAG.

```
VLT_Primary(conf)#interface tengigabitethernet 1/3
VLT_Primary(conf-if-te-1/3)#no ip address
VLT_Primary(conf-if-te-1/3)#port-channel-protocol lacp
VLT_Primary(conf-if-te-1/3)#port-channel 10 mode active
VLT_Primary(conf-if-te-1/3)#no shutdown
```

```
VLT_Primary(conf)#interface port-channel 10
VLT_Primary(conf-if-po-10)#no ip address
VLT_Primary(conf-if-po-10)#switchport
VLT_Primary(conf-if-po-10)#vlt-peer-lag port-channel 10
VLT_Primary(conf-if-po-10)#no shutdown
```

5. Configure a VLAN

```
VLT_Secondary(conf)#interface vlan 100
VLT_Secondary(conf-if-vl-100)#ip address 100.1.1.2/24
VLT_Secondary(conf-if-vl-100)#tagged port-channel 10
VLT_Secondary(conf-if-vl-100)#no shutdown
VLT_Secondary(conf-if-vl-100)#exit
```

6. Enable BFD over OSPF.

```
VLT_Secondary(conf)# router ospf 1
VLT_Secondary(conf-router_ospf)# network 100.1.1.0/24 area 0
VLT_Secondary(conf-router_ospf)# bfd all-neighbors
```

Verify the BFD configuration in each node using the following show commands:

- To verify the BFD neighbors in the ToR, use `show bfd neighbors` command.

```
TOR#show bfd neighbors
LocalAddr          RemoteAddr        Interface  State Rx-int Tx-int Mult Clients
* 100.1.1.3        100.1.1.1         Vl 100     Up    200    200    3    0
* 100.1.1.3        100.1.1.2         Vl 100     Up    200    200    3    0
```

- To verify the VLTi (ICL) link is up in the VLT primary peer, use `show vlt brief` command.

```
VLT_Primary#show vlt brief
VLT Domain Brief
------------------
Domain ID:                    100
Role:                         Primary
Role Priority:                32768
ICL Link Status:              Up
HeartBeat Status:             Up
VLT Peer Status:              Up
Version:                      6(9)
Local System MAC address:     f4:8e:38:6a:97:3f
Remote System MAC address:    00:e6:e2:f5:5c:15
Remote system version:        6(9)
Delay-Restore timer:          90 seconds
```

```
Delay-Restore Abort Threshold:  60 seconds
Peer-Routing :                  Enabled
Peer-Routing-Timeout timer:     0 seconds
Multicast peer-routing timeout: 150 seconds
```

- To verify the VLTi (ICL) link is up in the VLT secondary peer, use `show vlt brief` command.

```
VLT_Secondary#show vlt brief
VLT Domain Brief
------------------
Domain ID:                      100
Role:                           Secondary
Role Priority:                  32768
ICL Link Status:                Up
HeartBeat Status:               Up
VLT Peer Status:                Up
Version:                        6(9)
Local System MAC address:       00:e6:e2:f5:5c:15
Remote System MAC address:      f4:8e:38:6a:97:3f
Remote system version:          6(9)
Delay-Restore timer:            90 seconds
Delay-Restore Abort Threshold:  60 seconds
Peer-Routing :                  Enabled
Peer-Routing-Timeout timer:     0 seconds
Multicast peer-routing timeout: 150 seconds
```

61

VLT Proxy Gateway

The virtual link trucking (VLT) proxy gateway feature allows a VLT domain to locally terminate and route L3 packets that are destined to a Layer 3 (L3) end point in another VLT domain. Enable the VLT proxy gateway using the link layer discover protocol (LLDP) method or the static configuration. For more information, see the *Command Line Reference Guide*.

Topics:

- Proxy Gateway in VLT Domains
- Configuring a Static VLT Proxy Gateway
- Configuring an LLDP VLT Proxy Gateway
- VLT Proxy Gateway Sample Topology

Proxy Gateway in VLT Domains

Using a proxy gateway, the VLT peers in a domain can route the L3 packets destined for VLT peers in another domain as long as they have L3 reachability for the IP destinations.

A proxy gateway in a VLT domain provides the following benefits:

- Avoids sub-optimal routing of packets by a VLT domain when packets are destined to the endpoint in another VLT domain.
- Provides resiliency if a VLT peer goes down by performing proxy routing for the peer's destination MAC address in another VLT domain.

A typical scenario is virtual movement of servers across data centers. Virtual movement enables live migration of running virtual machines (VMs) from one host to another without downtime. For example, consider a square VLT connecting two data centers. If a VM, `VM1` on Server Rack 1 has C as its default gateway and VM1 performs a virtual movement to Server Rack 2 with no change in default gateway. In this case, L3 packets destined for C can be routed either by C1 or D1 locally. To do this, install the local system mac address of C and D in both C1 and D1 so the packets for C and D could have a hit at C1 /D1 and be routed locally.

The following figure shows:

- Server racks, named Rack 1 and Rack 2, are part of data centers named DC1 and DC2, respectively.
- Rack 1 is connected to devices A and B in Layer 2.
- Rack 2 is connected to devices A and B in Layer 2.
- A VLT link aggregation group (LAG) is present between A and B.
- A and B are connected to core routers, C and D.
- VLT routing is present between C and D.
- C1 and D1 are Layer 3 core routers in DC2, in which VLT routing is enabled.
- The core routers C and D in the local VLT domain is connected to the core routers C1 and D1 in the remote VLT Domain using VLT links in eVLT fashion.

For more information about eVLT, refer to the Virtual Link Trunking (VLT) chapter. The core or Layer 3 routers C and D in local VLT Domain and C1 and D1 in the remote VLT Domain are then part of a Layer 3 cloud.

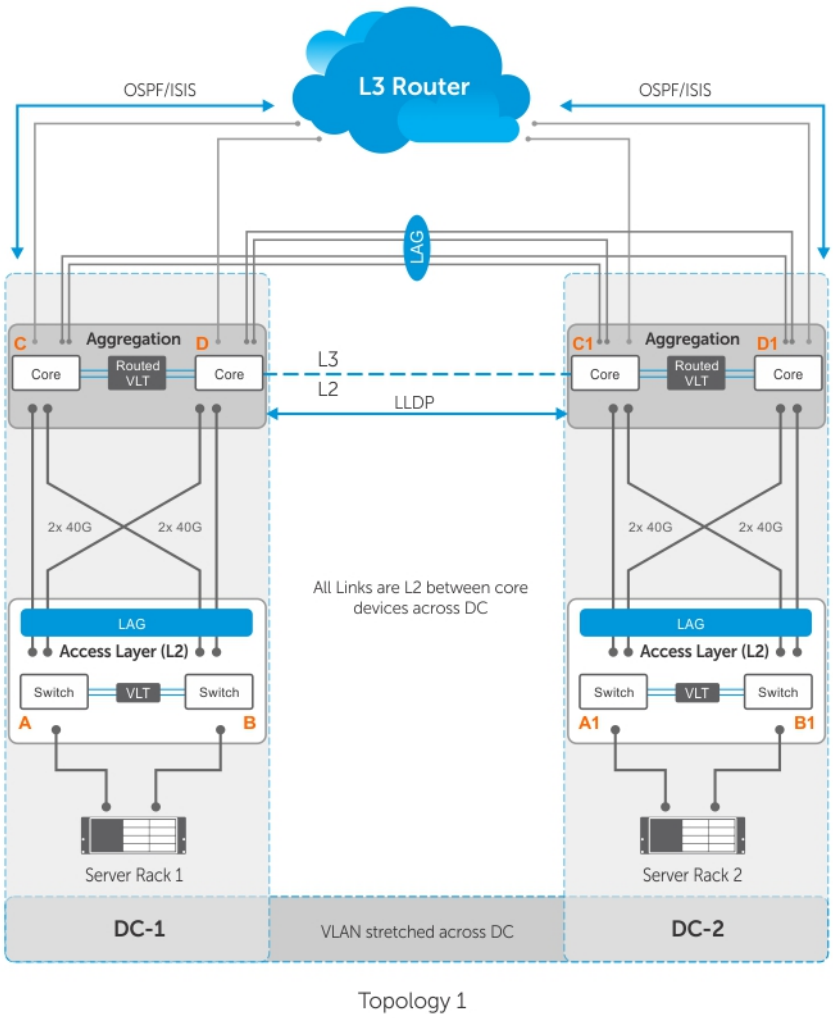


Figure 151. Sample Configuration for a VLT Proxy Gateway

Guidelines for Enabling the VLT Proxy Gateway

Keep the following points in mind when you enable a VLT proxy gateway:

- Proxy gateway is supported only for VLT; for example, across a VLT domain.
- You must enable the VLT `peer-routing` command for the VLT proxy gateway to function.
- Asymmetric virtual local area network (VLAN) configuration, such as the same VLAN configured with Layer 2 (L2) mode on one VLT domain and L3 mode on another VLT domain is not supported. You must always configure the same mode for the VLANs across the VLT domain.
- You must maintain VLAN symmetry within a VLT domain.
- The connection between DCs must be a L3 VLT in eVLT format . For more information, refer to the eVLT Configuration Example
- The trace route across the DCs can show extra hops.
- To ensure no traffic drops, you must maintain route symmetry across the VLT domains. When the routing table across DCs is not symmetrical, there is a possibility of a routing miss by a DC that does not have the route for L3 traffic. Because routing protocols are enabled and both DCs are in the same subnet, there is no dynamic route asymmetry. But if you configure a static route on one DC and not on the other, there is asymmetry.
- If the port-channel specified in the `proxy-gateway` command is not a VLT LAG, the configuration is rejected by the CLI.
- You cannot change the VLT LAG to a legacy LAG when it is part of proxy-gateway.
- You cannot change the link layer discovery protocol (LLDP) port channel interface to a legacy LAG when you enable a proxy gateway.
- Dell EMC Networking recommends the `vlt-peer-mac transmit` command only for square VLTs without diagonal links.
- The virtual router redundancy (VRRP) protocol and IPv6 routing is not supported.

- Private VLANs (PVLANs) are not supported.
- When a Virtual Machine (VM) moves from one VLT domain to the another VLT domain, the VM host sends the gratuitous ARP (GARP) , which in-turn triggers a mac movement from the previous VLT domain to the newer VLT domain.
- After a station move, if the host sends a TTL1 packet destined to its gateway; for example, a previous VLT node, the packet can be dropped.
- After a station move, if the host first PINGs its gateway; for example, a previous VLT node it results in a 40 to 60% success rate considering it takes a longer path.
- When you remove and add back a MAC address, L3 frames can be received out-of-order at the L3 cloud. This happens when proxy gateway routing and sub-optimal routing intersperse with each other.

Enable VLT Proxy Gateway

To enable the VLT proxy gateway, the system mac addresses of C and D in the local VLT domain must be installed in C1 and D1 in the remote VLT domain and vice versa. You can install the mac address in two methods - the `proxy-gateway lldp` method or the `proxy-gateway static` configuration. Proxy-gateway LLDP is a dynamic method of installing the local mac addresses in the remote VLT domain, which is achieved using a new organizational type, length, value (TLV) in LLDP packets.

LLDP Organizational TLV for Proxy Gateway

- You can configure the VLT proxy gateway in a VLT domain using the `proxy-gateway LLDP` command in proxy-gateway Configuration mode. Specify the port-channel interface of the square VLT link on which LLDP packets are sent using the`peer-domain-link port-channel` command.

  Configuring the `proxy gateway lldp` and the `peer-domain-link` port channel, LLDP sets TLV flags on the interfaces for receiving and transmitting private TLV packets. After defining these organizational TLV settings, LLDP encodes the local system mac-addresses as organizational TLVs for transmitting to the peer. If you specify the `no proxy gateway LLDP interface` command, LLDP stops transmitting and receiving proxy gateway TLV packets on the specified interfaces. However, other TLVs are not affected. From the interfaces on which you enabled the proxy gateway LLDP, LLDP decodes the TLV packets from the remote LLDP by using the new organizational TLV.

  The following requirements must be satisfied for LLDP proxy gateway to function correctly:
  - Data centers must be directly connected.
  - LLDP has a limited TLV size. As a result, information that is carried by the new TLV is limited to one or two MAC addresses.
  - You must have all related systems properly configured and set up.

- LLDP defines an organizationally specific TLV (type 127) with a unique identifier (0x0001E8) and a defined subtype (0x01) for sending or receiving information.
- LLDP uses the existing infrastructure and adds a new TLV for sending and receiving on the configured ports.
- There are only a few MAC addresses for each unit transmitted. All currently active MAC addresses are carried on the newly defined TLV.
- Dell EMC Networking devices not configured with VLT proxy gateway process standard TLVs and ignore TLVs configured with VLT proxy gateway.

The LLDP organizational TLV passes local destination MAC address information to peer VLT domain devices so they can act as a proxy gateway. To enable proxy gateway LLDP, two configurations are required:

- You must configure the global proxy gateway LLDP to enable the proxy-gateway LLDP TLV.
- You must configure the interface proxy gateway LLDP to enable or disable a proxy-gateway LLDP TLV on specific interfaces.
- The interface is typically a VLT port-channel that connects to a remote VLT domain.
- The new proxy gateway TLV is carried on the physical links under the port channel only.
- You must have at least one link connection to each unit of the VLT domain.

Following are the prerequisites for Proxy Gateway LLDP configuration:

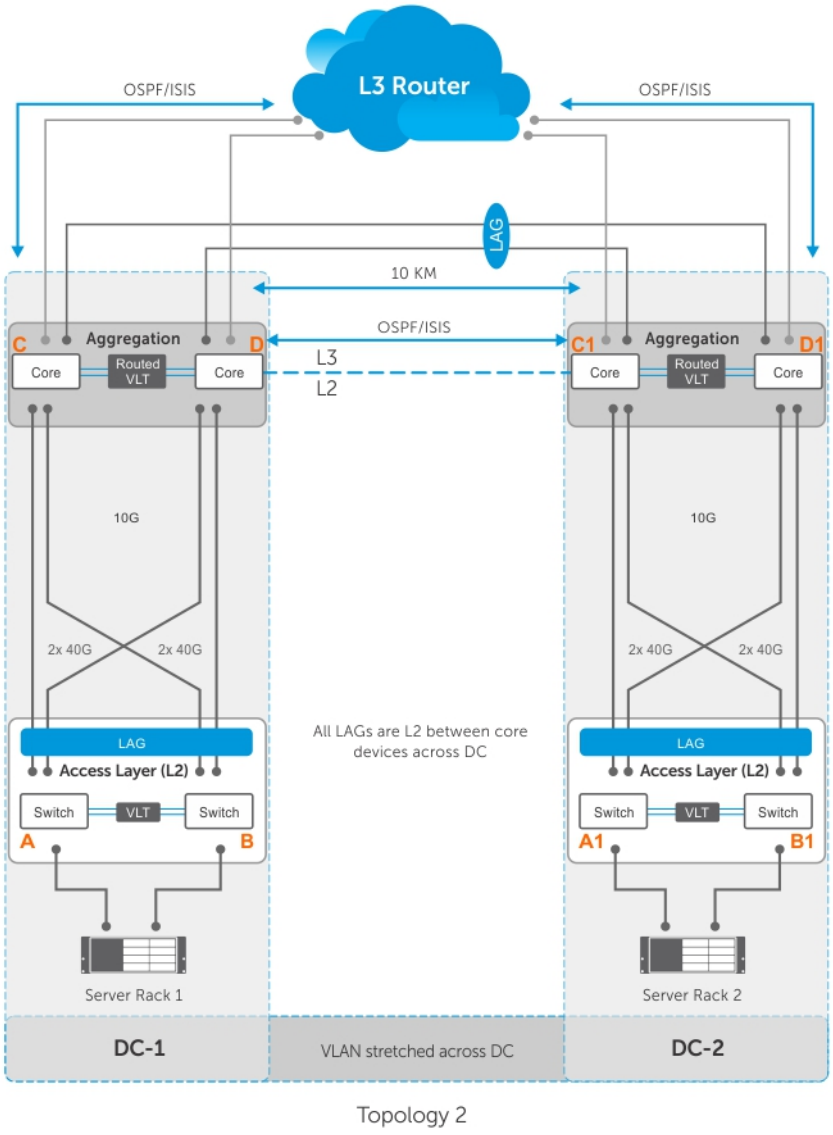
- You must globally enable LLDP.

- You cannot have interface–level LLDP disable commands on the interfaces configured for proxy gateway and you must enable both transmission and reception.
- You must connect both units of the remote VLT domain by the port channel member.
- If you connect more than one port to a unit of the remote VLT domain, the connection must be completed by the time you enable the proxy gateway LLDP.
- You cannot have other conflicting configurations (for example, you cannot have a static proxy gateway configuration).

Proxy Gateway LLDP configuration might not operate properly if one of the following conditions is true:

- Any proxy gateway configuration or LLDP configuration is not working.
- LLDP packets fail to reach the remote VLT domain devices (for example, because the system is down, rebooting, or the port's physical link connection is down).

LLDP VLT Proxy Gateway in a Square VLT Topology



Topology 2

Figure 152. Sample Configuration for a VLT Proxy Gateway

- The preceding figure shows a sample square VLT Proxy gateway topology. There are no diagonal links in the square VLT connection between the C and D in VLT domain 1 and C1 and D1 in the VLT domain 2. This causes sub-optimal routing.

For VLT Proxy Gateway to work in this scenario you must configure the `VLT-peer-mac transmit` command under VLT Domain Proxy Gateway LLDP mode, in both C and D (VLT domain 1) and C1 and D1 (VLT domain 2). This behavior is applicable only in the LLDP configuration and not required in the static configuration.

Sample Configuration

```
DellEMC(conf-vlt-domain)#proxy-gateway lldp
DellEMC(conf-vlt-domain-pxy-gw-lldp)#vlt-peer-mac transmit
```

- Assume the inter-chassis link (ICL) between C1 and D1 is shutdown and if D1 is the secondary VLT, one half of the inter DC link goes down. After VM motion, if a packet reaches D1 with the destination MAC address of D, it may be dropped. This behavior is applicable only in an LLDP configuration; in a static configuration, the packet is forwarded.
- Any L3 packet, when it gets an L3 hit and is routed, it has a time to live (TTL) decrement as expected.
- You can disable the VLT Proxy Gateway for a particular VLAN using an "Exclude-VLAN" configuration. The configuration has to be done in both the VLT domains [C and D in VLT domain 1 and C1 and D1 in VLT domain 2].

Sample Configuration LLDP Method

```
DellEMC(conf-vlt-domain)#proxy-gateway ll
DellEMC(conf-vlt-domain-pxy-gw-lldp)#peer-domain-link port-channel 1 exclude-vlan 10
```

Sample Configuration Static Method

```
DellEMC(conf-vlt-domain)#proxy-gateway static
DellEMC(conf-vlt-domain-pxy-gw-static)#remote-mac-address <xx:xx:xx:xx:xx:xx> exclude-
vlan 10
```

- Packet duplication may happen with “Exclude-VLAN” configuration – Assume you used the `exclude-vlan` option (called VLAN 10) in C and D and in C1 and D1; If packets for VLAN 10 with C’s MAC address (C is in VLT domain 1) gets an L3 hit at C1 in VLT domain 2, they are switched to both D1 (via ICL) and C via inter DC link. This may lead to packet duplication. Therefore, if C’s MAC address is learned at C1, the packet does not flood to D1 and only switches to C and avoids packet duplication.
- With the existing hardware capabilities, you can only disable VLT Proxy Gateway only for 500 VLANs, using exclude-VLAN configuration.

Configuring a Static VLT Proxy Gateway

You can configure a proxy gateway in VLT domains. A proxy gateway allows you to locally route the packets that are destined to an L3 endpoint of the other VLT domain.

Apply the following configurations in the Core L3 Routers C and D in local VLT domain and C1 and D1 in the remote VLT domain:

1. Configure `proxy-gateway static` in VLT Domain Configuration mode.
2. Configure `remote-mac-address <mac-address>` in VLT Domain Proxy Gateway LLDP mode. Configure the system mac-addresses of both C and D in C1 and also in D1 in the remote VLT domain and vice versa.

Sample Static Configuration on C switch or C1 switch

```
Switch_C#conf
Switch_C(conf)#vlt domain 1
Switch_C(conf-vlt-domain1)#proxy-gateway static
Switch_C(conf-vlt-domain1-pxy-gw-static)#remote-mac-address <xx:xx:xx:xx:xx:xx>....
```

Configuring an LLDP VLT Proxy Gateway

You can configure a proxy gateway in a VLT domain to locally route packets destined to a L3 endpoint in another VLT domain.

Apply the following configurations in the Core L3 Routers C and D in the local VLT domain and C1 and D1 in the remote VLT domain:

1. Configure `proxy-gateway lldp` in VLT Domain Configuration mode.
2. Configure `peer-domain-link port-channel <vlt portchannel ID>` in VLT Domain Proxy Gateway LLDP mode. The VLT port channel is the one that connects the remote VLT domain.

Sample Dynamic Proxy Configuration on C switch or C1 switch

```
Switch_C#conf
Switch_C(conf)#vlt domain 1
Switch_C(conf-vlt-domain1)#proxy-gateway lldp
Switch_C(conf-vlt-domain1-pxy-gw-lldp)#peer-domain-link port-channel 1....
```

VLT Proxy Gateway Sample Topology

VLT proxy gateway enables one VLT domain to act as proxy gateway for another VLT domain when a host or virtual machine is moved from one VLT domain to the other VLT domain.

The following image depicts a sample topology in which VLT proxy gateway is configured.

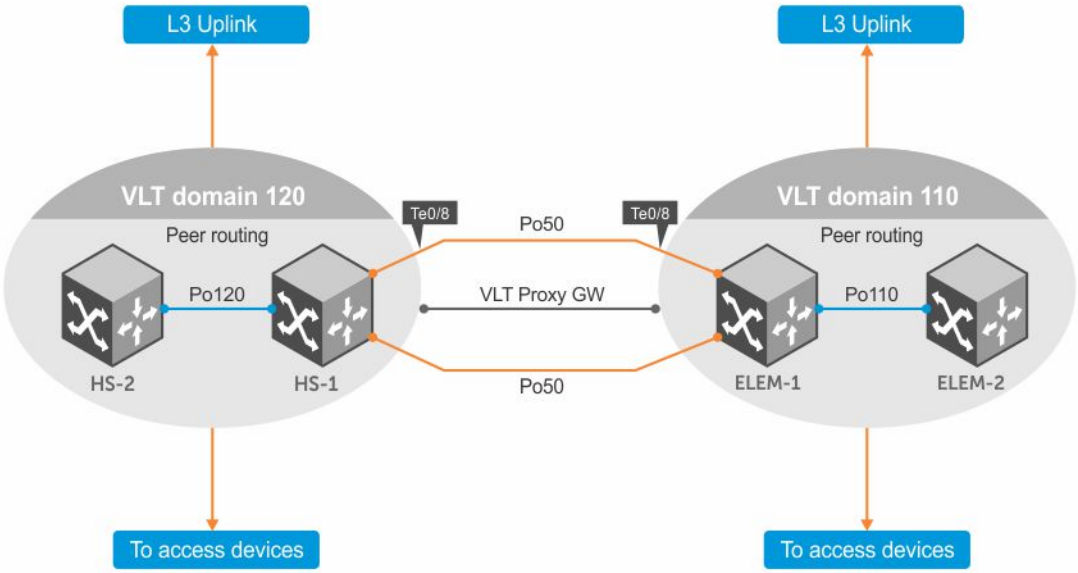


Figure 153. VLT Proxy Gateway Sample Topology

VLT Domain Configuration

Dell-1 and Dell-2 constitute VLT domain 120. Dell-3 and Dell-4 constitute VLT domain 110. These two VLT domains are connected using a VLT LAG P0 50. To know how to configure the interfaces in VLT domains, see the Configuring VLT section.

Dell-1 VLT Configuration

```
vlt domain 120
peer-link port-channel 120
back-up destination 10.1.1.3
primary-priority 4096
system-mac mac-address 02:01:e8:d8:93:e3
unit-id 0
peer-routing
! proxy-gateway static
remote-mac-address 00:01:e8:8b:ff:4f
remote-mac-address 00:01:e8:d8:93:04
```

The MAC addresses, configured using the remote-mac-address command, belong to Dell-3 and Dell-4.

```
interface TenGigabitEthernet 0/8
description "To DELL-3 10Gb"
no ip address
```

```
interface TenGigabitEthernet 0/9
description "To DELL-3 10Gb"
no ip address

!
port-channel-protocol LACP
port-channel 50 mode active
no shutdown

interface Port-channel 50
description "mVLT port channel to DELL-3"
no ip address
switchport
no spanning-tree
vlt-peer-lag port-channel 50
no shutdown
```

Note that on the inter-domain link, the `switchport` command is enabled. On a VLTi link between VLT peers in a VLT domain, the `switchport` command is not used.

VLAN 100 is used as the OSPF peering VLAN between Dell-1 and Dell-2.

```
interface Vlan 100
description OSPF Peering VLAN to Dell-2
ip address 10.10.100.1/30 ip ospf network point-to-point
no shutdown
```

VLAN 101 is used as the OSPF peering VLAN between the two VLT domains.

```
interface Vlan 101
description ospf peering vlan across VLTPG_Po50
ip address 10.10.101.1/30
tagged Port-channel 50 ip ospf network point-to-point
no shutdown
```

The following sequence of commands represent the OSPF configuration on Dell-1.

```
router ospf 1
router-id 4.4.4.4
network 10.10.100.0/30 area 0
network 10.10.101.0/30 area 0
```

The following output shows that Dell-2 and VLT domain 110 form OSPF neighborship with Dell-1.

```
Dell-1#show ip ospf neighbor
!
Neighbor ID Pri State    Dead Time Address     Interface Area
2.2.2.2     1   FULL/ -  00:00:39  10.10.100.2 Vl 100    0
3.3.3.3     1   FULL/ -  00:00:32  10.10.101.2 Vl 101    0
```

Dell-2 VLT Configuration

```
vlt domain 120
peer-link port-channel 120
back-up destination 10.1.1.2
primary-priority 24576
system-mac mac-address 02:01:e8:d8:93:e3
unit-id 1
peer-routing

proxy-gateway static
remote-mac-address 00:01:e8:8b:ff:4f
remote-mac-address 00:01:e8:d8:93:04
```

The MAC addresses, configured using the `remote-mac-address` command, belong to Dell-3 and Dell-4.

```
interface Vlan 100
description OSPF peering VLAN to Dell-1
ip address 10.10.100.2/30
ip ospf network point-to-point
no shutdown
```

The following is the OSPF configuration on Dell-2.

```
router ospf 1
router-id 2.2.2.2
network 10.10.100.0/30 area 0
```

The following output shows that Dell-1 forms OSPF neighborship with Dell-2.

```
Dell-2#sh ip ospf nei

Neighbor ID Pri State    Dead Time Address      Interface Area
4.4.4.4     1   FULL/ - 00:00:33  10.10.100.1 Vl 100    0
```

Dell-3 VLT Configuration

```
vlt domain 110
peer-link port-channel 110
back-up destination 10.1.1.1
primary-priority 4096
system-mac mac-address 02:01:e8:d8:93:02
unit-id 0
peer-routing
! proxy-gateway static
remote-mac-address 00:01:e8:d8:93:07
remote-mac-address 00:01:e8:d8:93:e5
```

These MAC addresses are the system L2 interface addresses for each switch at the remote site, Dell-1 and Dell-2.

```
interface TenGigabitEthernet 0/8
description "To Dell-1 10Gb"
no ip address

interface TenGigabitEthernet 0/9
description "To Dell-1 10Gb"
no ip address

!
port-channel-protocol LACP
port-channel 50 mode active
no shutdown
!
interface Port-channel 50
description "mVLT port channel to Dell-1"
no ip address
switchport no spanning-tree
(STP is disabled between sites)
vlt-peer-lag port-channel 50
no shutdown


interface Vlan 101
description ospf peering vlan across VLTPG_Po50
ip address 10.10.101.2/30
tagged Port-channel 50
ip ospf network point-to-point
no shutdown
!
interface Vlan 102
description ospf peering vlan to DELL-4
ip address 10.10.102.1/30
```

```
ip ospf network point-to-point
no shutdown
```

The following is the OSPF configuration on Dell-3.

```
router ospf 1
router-id 3.3.3.3
network 10.10.101.0/30 area 0
network 10.10.102.0/30 area 0
```

The following output shows that Dell-4 and VLT domain 120 form OSPF neighborship with Dell-3.

```
Dell-3#sh ip ospf nei
!
Neighbor ID Pri State   Dead Time Address     Interface Area
4.4.4.4     1   FULL/ - 00:00:33  10.10.101.1 Vl 101    0
1.1.1.1     1   FULL/ - 00:00:34  10.10.102.2 Vl 102    0
```

Dell-4 VLT Configuration

```
vlt domain 110
peer-link port-channel 110
back-up destination 10.1.1.0
primary-priority 24576
system-mac mac-address 02:01:e8:d8:93:02
unit-id 1
peer-routing
! proxy-gateway static
remote-mac-address 00:01:e8:d8:93:07
remote-mac-address 00:01:e8:d8:93:e5
```

These MAC addresses are the system L2 interface addresses for each switch at the remote site, Dell-1 and Dell-2.

```
interface Vlan 102
description ospf peering vlan to DELL-3
ip address 10.10.102.2/30
ip ospf network point-to-point
no shutdown
```

The following is the OSPF configuration on Dell-4.

```
router ospf 1
router-id 1.1.1.1
network 10.10.102.0/30 area 0
```

The following output shows that Dell-4 forms OSPF neighborship with Dell-3.

```
ELEM-2#sh ip ospf nei
!
Neighbor ID Pri State   Dead Time Address     Interface Area
3.3.3.3     1   FULL/ - 00:00:32  10.10.102.1 Vl        102
```

With the above configuration, the two VLT domains act as VLT proxy gateways for any host that is moved from one VLT domain to another.

62

Virtual Routing and Forwarding (VRF)

Virtual Routing and Forwarding (VRF) allows a physical router to partition itself into multiple Virtual Routers (VRs). The control and data plane are isolated in each VR so that traffic does NOT flow across VRs.Virtual Routing and Forwarding (VRF) allows multiple instances of a routing table to co-exist within the same router at the same time.

Topics:

- VRF Overview
- VRF Configuration Notes
- VRF Configuration
- Sample VRF Configuration
- Route Leaking VRFs
- Dynamic Route Leaking

VRF Overview

VRF improves functionality by allowing network paths to be segmented without using multiple devices. Using VRF also increases network security and can eliminate the need for encryption and authentication due to traffic segmentation.

Internet service providers (ISPs) often take advantage of VRF to create separate virtual private networks (VPNs) for customers; VRF is also referred to as VPN routing and forwarding.

VRF acts like a logical router; while a physical router may include many routing tables, a VRF instance uses only a single routing table. VRF uses a forwarding table that designates the next hop for each data packet, a list of devices that may be called upon to forward the packet, and a set of rules and routing protocols that govern how the packet is forwarded. These VRF forwarding tables prevent traffic from being forwarded outside a specific VRF path and also keep out traffic that should remain outside the VRF path.

VRF uses interfaces to distinguish routes for different VRF instances. Interfaces in a VRF can be either physical (Ethernet port or port channel) or logical (VLANs). You can configure identical or overlapping IP subnets on different interfaces if each interface belongs to a different VRF instance.

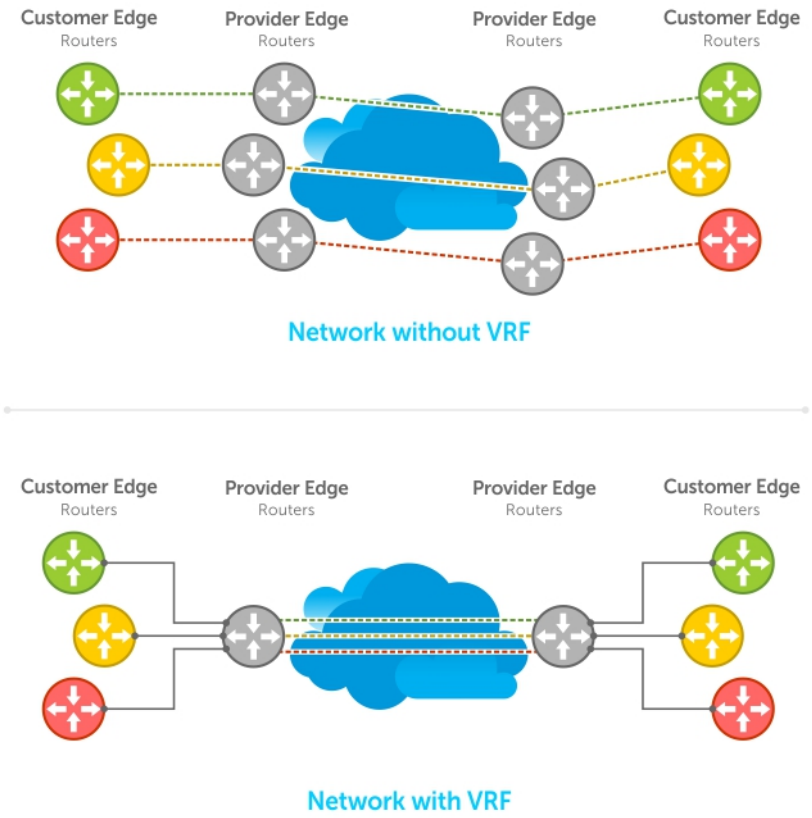


Figure 154. VRF Network Example

VRF Configuration Notes

Although there is no restriction on the number of VLANs that can be assigned to a VRF instance, the total number of routes supported in VRF is limited by the size of the IPv4 CAM.

VRF is implemented in a network device by using Forwarding Information Bases (FIBs).

A network device may have the ability to configure different virtual routers, where entries in the FIB that belong to one VRF cannot be accessed by another VRF on the same device. Only Layer 3 interfaces can belong to a VRF. VRF is supported on following types of interface:

- Physical Ethernet interfaces
- Port-channel interfaces (static & dynamic using LACP)
- VLAN interfaces
- Loopback interfaces

VRF supports route redistribution between routing protocols (including static routes) only when the routes are within the same VRF.

uses both the VRF name and VRF ID to manage VRF instances. The VRF name and VRF ID number are assigned using the ip vrf command. The VRF ID is displayed in show ip vrf command output.

The VRF ID is not exchanged between routers. VRF IDs are local to a router.

While using /32 route leak, do not use VLAN interface as Next-hop. This would result in packets being soft-forwarded by CPU, which might lead to latency and packet drop.

If the next-hop IP in a static route VRF statement is VRRP IP of another VRF, this static route does not get installed on the VRRP master.

VRF supports some routing protocols only on the default VRF (default-vrf) instance. Table 1 displays the software features supported in VRF and whether they are supported on all VRF instances or only the default VRF.

NOTE: To configure a router ID in a non-default VRF, configure at least one IP address in both the default as well as the non-default VRF.

Table 130. Software Features Supported on VRF

Feature/Capability	Support Status for Default VRF	Support Status for Non-default VRF
802.1x protocol on the VLAN port	Yes	Yes
OSPF, RIP, ISIS, BGP on physical and logical interfaces	Yes	Yes **NOTE:** RIP is not supported on non-default VRF.
Dynamic Port-channel (LACP) on VLAN port or a Layer 3 port	Yes	Yes
Static Port-channel as VLAN port or a Layer 3 port	Yes	Yes
Encapsulated Remote Port Monitoring	Yes	No
BFD on physical and logical interfaces	Yes	No
Multicast protocols (PIM-SM, MSDP)	Yes	Yes **NOTE:** MSDP is not supported in non-default VRF.
PIM-DM	No	No
Layer3 (IPv4/IPv6) ACLs on physical interfaces and LAGs	Yes	Yes **NOTE:** IPv6 ACLs supported only in default VRF
PBR, L3 QOS on physical interfaces and LAGs	Yes	No
Layer 3 (IPv4/IPv6) ACLs on VLANs	Yes	Yes **NOTE:** IPv6 ACLs supported only in default VRF.
PBR, L3 QoS on VLANs	Yes **NOTE:** QoS not supported on VLANs.	No
IPv4 ARP	Yes	Yes
sFlow	Yes	No
VRRP on physical and logical interfaces	Yes	Yes
VRRPV3	Yes	Yes
Secondary IP Addresses	Yes	Yes
Basic	Yes	Yes
OSPFv3	Yes	Yes
IS-IS	Yes	Yes
BGP	Yes	Yes
ACL	Yes	No
Multicast	No	No
NDP	Yes	Yes
RAD	Yes	Yes

DHCP

DHCP requests are not forwarded across VRF instances. The DHCP client and server must be on the same VRF instance.

VRF Configuration

The VRF configuration tasks are:

1. Enabling VRF in Configuration Mode
2. Creating a Non-Default VRF
3. Assign an Interface to a VRF

You can also:

- View VRF Instance Information
- Connect an OSPF Process to a VRF Instance
- Configure VRRP on a VRF

Loading VRF CAM

- Load CAM memory for the VRF feature.
  CONFIGURATION
  `feature vrf`

After you load VRF CAM, CLI parameters that allow you to configure non-default VRFs are made available on the system.

Creating a Non-Default VRF Instance

VRF is enabled by default on the switch and supports up to 512 VRF instances: 1 to 511 and the default VRF (0).

- Create a non-default VRF instance by specifying a name and VRF ID number, and enter VRF configuration mode.
  CONFIGURATION
  `ip vrf` *`vrf-name vrf-id`*
  The VRF ID range is from 1 to 511. 0 is the default VRF ID.

Assigning an Interface to a VRF

You must enter the `ip vrf forwarding` command before you configure the IP address or any other setting on an interface.

NOTE: You can configure an IP address or subnet on a physical or VLAN interface that overlaps the same IP address or subnet configured on another interface only if the interfaces are assigned to different VRFs. If two interfaces are assigned to the same VRF, you cannot configure overlapping IP subnets or the same IP address on them.

- Assign an interface to a VRF instance.
  INTERFACE
  `ip vrf forwarding` *`vrf-name`*

Assigning a Front-end Port to a Management VRF

Starting in 9.7(0.0) release, you can assign a front-end port to a management VRF and make the port to act as a host interface.

NOTE: You cannot assign loop-back and port-channel interfaces to a management port.

To assign a front-end port to a management VRF, perform the following steps:

1. Enter the front-end interface that you want to assign to a management interface.
   CONFIGURATION
   `interface tengigabitethernet 1/1`

2. Assign the interface to management VRF.
   INTERFACE CONFIGURATION
   `ip vrf forwarding management`
   Before assigning a front-end port to a management VRF, ensure that no IP address is configured on the interface.
3. Assign an IPv4 address to the interface.
   INTERFACE CONFIGURATION
   `ip address 10.1.1.1/24`
   Before assigning a front-end port to a management VRF, ensure that no IP address is configured on the interface.
4. Assign an IPv6 address to the interface.
   INTERFACE CONFIGURATION
   `ipv6 address 1::1`
   You can also auto configure an IPv6 address using the `ipv6 address autoconfig` command.

View VRF Instance Information

To display information about VRF configuration, enter the `show ip vrf` command. To display information on all VRF instances (including the default VRF 0), do not enter a value for *vrf-name*.

- Display the interfaces assigned to a VRF instance.
  EXEC
  `show ip vrf [vrf-name]`

Assigning an OSPF Process to a VRF Instance

OSPF routes are supported on all VRF instances. See the *Open Shortest Path First* (OSPFv2) chapter for complete OSPF configuration information.

Assign an OSPF process to a VRF instance . Return to CONFIGURATION mode to enable the OSPF process. The OSPF Process ID is the identifying number assigned to the OSPF process, and the Router ID is the IP address associated with the OSPF process.

Once the OSPF process and the VRF are tied together, the OSPF Process ID cannot be used again in the system.

- Enable the OSPFv2 process globally for a VRF instance. Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.
  CONFIGURATION
  `router ospf process-id vrf vrf name`
  The `process-id` range is from 0-65535.

Configuring VRRP on a VRF Instance

You can configure the VRRP feature on interfaces that belong to a VRF instance.

In a virtualized network that consists of multiple VRFs, various overlay networks can exist on a shared physical infrastructure. Nodes (hosts and servers) that are part of the VRFs can be configured with IP static routes for reaching specific destinations through a given gateway in a VRF. VRRP provides high availability and protection for next-hop static routes by eliminating a single point of failure in the default static routed network. .

Table 131. Configuring VRRP on a VRF

Task	Command Syntax	Command Mode
Create VRF	`ip vrf vrf1`	CONFIGURATION
Assign the VRF to an interface	`ip vrf forwarding vrf1`	INTERFACE–CONFIG

Table 131. Configuring VRRP on a VRF (continued)

Task	Command Syntax	Command Mode
Assign an IP address to the interface	`ip address 10.1.1.1 /24` `no shutdown`	

Configure the VRRP group and virtual IP address	`vrrp-group 10`
	`virtual-address 10.1.1.100`
	`show config`
	`----------`
	`!`
	`interface TenGigabitEthernet 1/13`
	`ip vrf forwarding vrf1`
	`ip address 10.1.1.1/24`
	`!`
	`vrrp-group 10`
	`virtual-address 10.1.1.100`
	`no shutdown`	

View VRRP command output for the VRF vrf1	`show vrrp vrf vrf1`
	`----------`
	`TenGigabitEthernet 1/13, IPv4 VRID: 10, Version: 2, Net: 10.1.1.1`
	`VRF: 1 vrf1`
	`State: Master, Priority: 100,`
	`Master: 10.1.1.1 (local)`
	`Hold Down: 0 sec, Preempt: TRUE,`
	`AdvInt: 1 sec`
	`Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 43, Gratuitous ARP sent: 0`
	`Virtual MAC address:`
	`00:00:5e:00:01:0a`
	`Virtual IP address:`
	`10.1.1.100`
	`Authentication: (none)`	

Configuring Management VRF

You can assign a management interface to a management VRF.

NOTE: The loopback interface cannot be added into the management VRF.

1. Create a management VRF.
   CONFIGURATION
   `ip vrf management`
2. Assign a management port to a management VRF.
   VRF MODE
   `interface management`

When Management VRF is configured, the following interface range or interface group commands are disabled:

- `ipv6 nd dad` — Duplicated Address Detection
- `ipv6 nd dns-server` — Configure DNS distribution option in RA packets originated by the router
- `ipv6 nd hop-limit` — Set hop limit advertised in RA and used in IPv6 data packets originated by the router
- `ipv6 nd managed-config-flag` — Hosts should use DHCP for address config
- `ipv6 nd max-ra-interval` — Set IPv6 Max Router Advertisement Interval
- `ipv6 nd mtu` — Configure MTU advertisements in RA packets
- `ipv6 nd other-config-flag` — Hosts should use DHCP for non-address config
- `ipv6 nd prefix` — Configure IPv6 Routing Prefix Advertisement
- `ipv6 nd ra-guard` — Configure IPv6 ra-guard
- `ipv6 nd ra-lifetime` — Set IPv6 Router Advertisement Lifetime

- `ipv6 nd reachable-time` — Set advertised reachability time
- `ipv6 nd retrans-timer` — Set NS retransmit interval used and advertised in RA
- `ipv6 nd suppress-ra` — Suppress IPv6 Router Advertisements
- `ipv6 ad <ipv6-address>` — IPv6 Address Detection
- `ipv6 ad autoconfig` — IPv6 stateless auto-configuration
- `ipv6 address <ipv6-address>` — Configure IPv6 address on an interface

NOTE: The command line help still displays relevant details corresponding to each of these commands. However, these interface range or interface group commands are not supported when Management VRF is configured.

Configuring a Static Route

- Configure a static route that points to a management interface.

  CONFIGURATION

  `management route ip-address mask managementethernet` or `management route ipv6-address prefix-length managementethernet`

  You can also have the management route to point to a front-end port in case of the management VRF. For example: `management route 2::/64 tengigabitethernet 1/1`.
- Configure a static entry in the IPv6 neighbor discovery.

  CONFIGURATION

  `ipv6 neighbor vrf management 1::1 tengigabitethernet 1/1 xx:xx:xx:xx:xx:xx`

Sample VRF Configuration

The following configuration illustrates a typical VRF set-up.

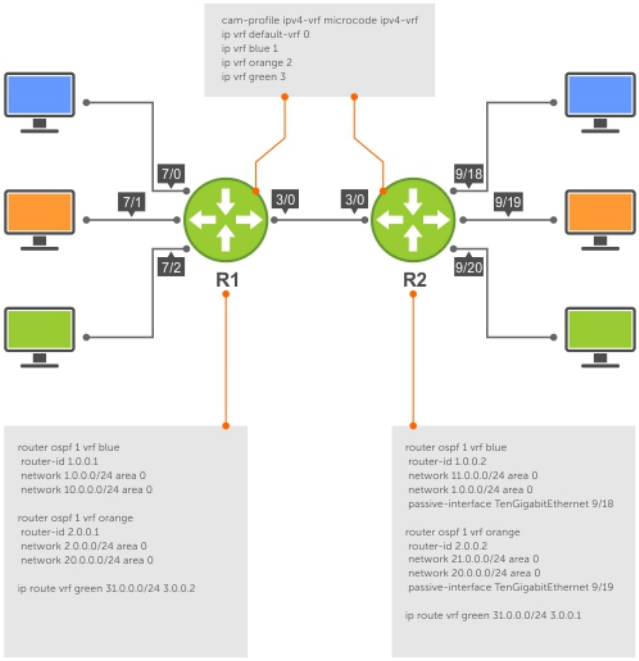


Figure 155. Setup OSPF and Static Routes

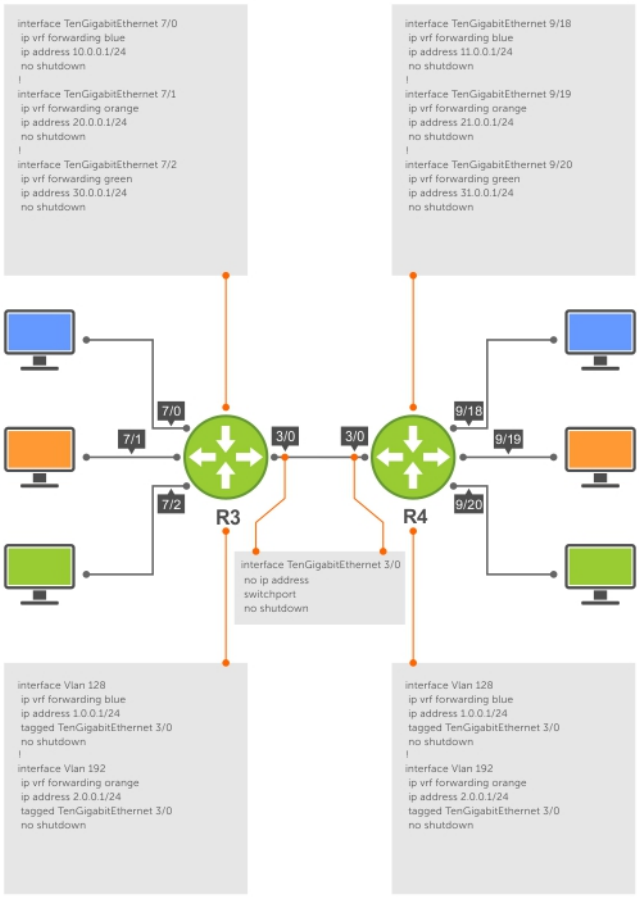


Figure 156. Setup VRF Interfaces

The following example relates to the configuration shown in the above illustrations.

Router 1

```
ip vrf blue 1
!
ip vrf orange 2
!
ip vrf green 3
!
interface TenGigabitEthernet 3/1
    no ip address
    switchport
    no shutdown
!
interface TenGigabitEthernet 1/1
    ip vrf forwarding blue
    ip address 10.0.0.1/24
    no shutdown
!
interface TenGigabitEthernet 1/2
    ip vrf forwarding orange
    ip address 20.0.0.1/24
    no shutdown
!
interface TenGigabitEthernet 1/3
    ip vrf forwarding green
    ip address 30.0.0.1/24
    no shutdown
!
interface Vlan 128
```

```
    ip vrf forwarding blue
    ip address 1.0.0.1/24
    tagged TenGigabitEthernet 3/1
    no shutdown
!
interface Vlan 192
    ip vrf forwarding orange
    ip address 2.0.0.1/24
    tagged TenGigabitEthernet 3/1
    no shutdown
!
interface Vlan 256
    ip vrf forwarding green
    ip address 3.0.0.1/24
    tagged TenGigabitEthernet 3/1
    no shutdown
!
router ospf 1 vrf blue
    router-id 1.0.0.1
    network 1.0.0.0/24 area 0
    network 10.0.0.0/24 area 0
!
router ospf 2 vrf orange
    router-id 2.0.0.1
    network 2.0.0.0/24 area 0
    network 20.0.0.0/24 area 0
!
ip route vrf green 31.0.0.0/24 3.0.0.2!
```

Router 2

```
ip vrf blue 1
!
ip vrf orange 2
!
ip vrf green 3
!
interface TenGigabitEthernet 3/1
    no ip address
    switchport
    no shutdown
!
interface TenGigabitEthernet 2/1
    ip vrf forwarding blue
    ip address 11.0.0.1/24
    no shutdown
!
interface TenGigabitEthernet 2/2
    ip vrf forwarding orange
    ip address 21.0.0.1/24
    no shutdown
!
interface TenGigabitEthernet 2/3
    ip vrf forwarding green
    ip address 31.0.0.1/24
    no shutdown
!
interface Vlan 128
    ip vrf forwarding blue
    ip address 1.0.0.2/24
    tagged TenGigabitEthernet 3/1
no shutdown
interface Vlan 192
    ip vrf forwarding orange
    ip address 2.0.0.2/24
    tagged TenGigabitEthernet 3/1
    no shutdown
!
interface Vlan 256
    ip vrf forwarding green
    ip address 3.0.0.2/24
    tagged TenGigabitEthernet 3/1
    no shutdown
```

```
!
router ospf 1 vrf blue
    router-id 1.0.0.2
    network 11.0.0.0/24 area 0
    network 1.0.0.0/24 area 0
    passive-interface TenGigabitEthernet 2/1
!
router ospf 2 vrf orange
    router-id 2.0.0.2
    network 21.0.0.0/24 area 0
    network 2.0.0.0/24 area 0
    passive-interface TenGigabitEthernet 2/2
!
ip route vrf green30.0.0.0/24 3.0.0.1
!
```

The following shows the output of the show commands on Router 1.

Router 1

```
DellEMC#show ip vrf
VRF-Name                        VRF-ID Interfaces

default-vrf                     0      Te 3/0-3,
                                       Te 1/3-47,
                                       Te 2/0-47,
                                       Ma 0/0,
                                       Ma 1/0,
                                       Nu 0,
                                       Vl 1
blue                            1      Te 1/1,
                                       Vl 128
orange                          2      Te 1/2,
                                       Vl 192
green                           3      Te 1/3,
                                       Vl 256
DellEMC#show ip ospf 1 neighbor
Neighbor ID     Pri     State         Dead Time Address Interface Area
1.0.0.2          1     FULL/DR     00:00:32 1.0.0.2         Vl 128         0

DellEMC#sh ip ospf 2 neighbor
Neighbor ID  Pri     State         Dead Time Address Interface Area
2.0.0.2          1     FULL/DR     00:00:37 2.0.0.2         Vl 192         0
DellEMC#show ip route vrf blue

Codes: C - connected, S - static, R - RIP,
            B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
            O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
            N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
            E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
            L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
            > - non-active route, + - summary route

Gateway of last resort is not set

        Destination         Gateway                 Dist/Metric         Last Change
        -----------        -------                 -----------         -----------
C         1.0.0.0/24         Direct, Vl 128     0/0                    00:20:48
C         10.0.0.0/24         Direct, Te 1/1     0/0                    00:10:06
O         11.0.0.0/24         via 1.0.0.2,         Vl 128 110/2     00:11:13

DellEMC#show ip route vrf orange

Codes: C - connected, S - static, R - RIP,
            B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
            O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
            N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
            E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
            L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
            > - non-active route, + - summary route

Gateway of last resort is not set
```

```
        Destination         Gateway                         Dist/Metric         Last
Change
        -----------         -------                         -----------
-----------
C         2.0.0.0/24         Direct, Vl 192             0/0                      00:20:55
C         20.0.0.0/24         Direct, Te 1/2             0/0                      00:10:05
O         21.0.0.0/24         via 2.0.0.2,    Vl 192         110/2
00:10:41

DellEMC#show ip route vrf green

Codes: C - connected, S - static, R - RIP,
           B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
           O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
           N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
           E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
           L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
           > - non-active route, + - summary route

Gateway of last resort is not set

        Destination         Gateway                         Dist/Metric         Last
Change
        -----------         -------                         -----------
-----------
C         3.0.0.0/24         Direct, Vl 256             0/0                      00:20:52
C         30.0.0.0/24         Direct, Te 1/3             0/0                      00:09:45
S         31.0.0.0/24         via 3.0.0.2, Vl 256     1/0                      00:09:06
================================================================================
```

The following shows the output of the show commands on Router 2.

Router 2

```
DellEMC#show ip vrf
VRF-Name                                           VRF-ID     Interfaces
default-vrf                                        0            Te 3/0-3,
                                                                  Te 2/0-17,21-47,
                                                                  Ma 0/0,
                                                                  Ma 1/0,
                                                                  Nu 0,
                                                                  Vl 1
blue                                                1            Te 2/1,
                                                                  Vl 128
orange                                                2            Te 2/2,
                                                                  Vl 192
green                                               3            Te 2/3,
                                                                  Vl 256

DellEMC#show ip ospf 1 neighbor
Neighbor ID         Pri             State Dead Time     Address     Interface     Area
1.0.0.1 1         FULL/BDR     00:00:36                1.0.0.1     Vl 128             0
!
DellEMC#sh ip ospf 2 neighbor
Neighbor ID         Pri             State Dead Time     Address     Interface     Area
2.0.0.1 1         FULL/BDR     00:00:33                2.0.0.1     Vl 192             0
!
DellEMC#show ip route vrf blue

Codes: C - connected, S - static, R - RIP,
           B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
           O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
           N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
           E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
           L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
           > - non-active route, + - summary route

Gateway of last resort is not set
    Destination         Gateway                         Dist/Metric         Last Change
    -----------         -------                         -----------         -----------
C     1.0.0.0/24         Direct,    Vl 128                0/0
00:27:21
O     10.0.0.0/24         via 1.0.0.1,    Vl 128         110/2                00:14:24
```

```
C     11.0.0.0/24         Direct, Te 2/1             0/0                    00:19:46

DellEMC#show ip route vrf orange

Codes: C - connected, S - static, R - RIP,
            B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
            O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
            N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
            E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
            L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
            > - non-active route, + - summary route

Gateway of last resort is not set

    Destination         Gateway                        Dist/Metric         Last Change
    -----------         -------                        -----------         -----------
C     2.0.0.0/24         Direct, Vl 192             0/0                    00:26:44
O     20.0.0.0/24         via 2.0.0.1, Vl 192     110/2                 00:14:22
C     21.0.0.0/24         Direct, Te 2/2             0/0                    00:20:38

DellEMC#show ip route vrf green

Codes: C - connected, S - static, R - RIP,
            B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
            O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
            N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
            E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
            L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
            > - non-active route, + - summary route

Gateway of last resort is not set

    Destination         Gateway                        Dist/Metric         Last Change
    -----------         -------                        -----------         -----------
C     3.0.0.0/24         Direct, Vl 256             0/0                    00:26:27
S     30.0.0.0/24         via 3.0.0.1, Vl 256     1/0                   00:17:03
C     31.0.0.0/24         Direct, Te 2/3             0/0                    00:20:19
DellEMC#
```

Route Leaking VRFs

Static routes can be used to redistribute routes between non-default to default/non-default VRF and vice-versa.

You can configure route leaking between two VRFs using the following command: `ip route vrf x.x.x.x s.s.s.s nh.nh.nh.nh vrf default.`

This command indicates that packets that are destined to `x.x.x.x/s.s.s.s` are reachable through `nh.nh.nh.nh` in the default VRF table. Meaning, the routes to `x.x.x.x/s.s.s.s` are leaked from the default VRF routing table into the non-default VRF routing table.

NOTE: The Dell EMC Networking OS supports route leaking only for transit traffic. If the system receives a packet on one VRF which is destined to another VRF, the packet is routed to that destination. If the system receives a packet on one VRF which is destined to the same device (such as a ping), they system drops the packet.

The following example illustrates how route leaking between two VRFs can be performed:

```
interface TenGigabitEthernet 1/9
  ip vrf forwarding VRF1
  ip address 120.0.0.1/24
interface TenGigabitEthernet 1/10
  ip vrf forwarding VRF2
  ip address 140.0.0.1/24
ip route vrf VRF1 20.0.0.0/16 140.0.0.2 vrf VRF2
ip route vrf VRF2 40.0.0.0/16 120.0.0.2 vrf VRF1
```

Dynamic Route Leaking

Route Leaking is a powerful feature that enables communication between isolated (virtual) routing domains by segregating and sharing a set of services such as VOIP, Video, and so on that are available on one routing domain with other virtual domains. Inter-VRF Route Leaking enables a VRF to leak or export routes that are present in its RTM to one or more VRFs.

Dynamic Route Leaking enables a source VRF to share both its connected routes as well as dynamically learnt routes from various protocols, such as ISIS, OSPF, BGP, and so on, with other default or non-default VRFs.

You can also leak global routes to be made available to VRFs. As the global RTM usually contains a large pool of routes, when the destination VRF imports global routes, these routes will be duplicated into the VRF's RTM. As a result, it is mandatory to use route-maps to filter out leaked routes while sharing global routes with VRFs.

Configuring Route Leaking without Filtering Criteria

You can use the `ip route-export tag` command to export all the IPv4 routes corresponding to a source VRF. For leaking IPv6 routes, use the `ipv6 route-export tag` command. This action exposes source VRF's routes (IPv4 or IPv6 depending on the command that you use) to various other VRFs. The destinations or target VRFs then import these IPv4 or IPv6 routes using the `ip route-import tag` or the `ipv6 route-import tag` command respectively.

NOTE: In Dell EMC Networking OS, you can configure at most one route-export per VRF as only one set of routes can be exposed for leaking. However, you can configure multiple route-import targets because a VRF can accept routes from multiple VRFs.

After the target VRF learns routes that are leaked by the source VRF, the source VRF in turn can leak the export target corresponding to the destination VRFs that have imported its routes. The source VRF learns the export target corresponding to the destinations VRF using the `ip route-import tag` or `ipv6 route-import tag` command. This mechanism enables reverse communication between destination VRF and the source VRF.

If the target VRF contains the same prefix (either sourced or Leaked route from some other VRF), then the Leak for that particular prefix will fail and an error-log will be thrown. Manual intervention is required to clear the unneeded prefixes. The source route will take priority over the leaked route and the leaked route is deleted.

Consider a scenario where you have created four VRF tables VRF-red, VRF-blue, VRF-Green, and VRF-shared. The VRF-shared table belongs to a particular service that should be made available only to VRF-Red and VRF-Blue but not VRF-Green.
For this purpose, routes corresponding VRF-Shared routes are leaked to only VRF-Red and VRF-Blue. And for reply, routes corresponding to VRF-Red and VRF-Blue are leaked to VRF-Shared.

For leaking the routes from VRF-Shared to VRF-Red and VRF-Blue, you can configure route-export tag on VRF-shared (source VRF, who is exporting the routes); the same route-export tag value should be configured on VRF-Red and VRF-blue as route-import tag (target VRF, that is importing the routes). For a reply communication, VRF-red and VRF-blue are configured with two different route-export tags, one for each, and those two values are configured as route-import tags on VRF-shared.

To configure route leaking, perform the following steps:

1. Configure VRF-shared using the following command:

   ```
   ip vrf vrf-shared
   interface interface-type slot/port
   ip vrf forwarding vrf-shared
   ip address ip—address mask
   ```

   A non-default VRF named VRF-Shared is created and the interface 1/4 is assigned to this VRF.
2. Configure the export target in the source VRF:.

   ```
   ip route-export 1:1
   ```
3. Configure VRF-red.

   ```
   ip vrf vrf-red
   interface-type slot/port
   ip vrf forwarding VRF-red
   ip address ip—address mask
   ```

   A non-default VRF named VRF-red is created and the interface is assigned to this VRF.
4. Configure the import target in VRF-red.

```
ip route-import 1:1
```

5. Configure the export target in VRF-red.

```
ip route-export 2:2
```

6. Configure VRF-blue.

```
ip vrf vrf-blue
interface-type slot/port
ip vrf forwarding VRF-blue
ip address ip—address mask
```

A non-default VRF named VRF-blue is created and the interface 1/12 is assigned to it.

7. Configure the import target in VRF-blue.

```
ip route-import 1:1
```

8. Configure the export target in VRF-blue.

```
ip route-import 3:3
```

9. Configure VRF-green.

```
ip vrf vrf-green
interface-type slot/port
ip vrf forwarding VRF-green
ip address ip—address mask
```

A non-default VRF named VRF-green is created and the interface is assigned to it.

10. Configure the import target in the source VRF VRF-Shared for reverse communication with VRF-red and VRF-blue.

```
ip vrf vrf-shared
ip route—import 2:2
ip route-import 3:3
```

The show run output for the above configuration is as follows:

```
ip vrf VRF-Red
    ip route-export  2:2
    ip route-import  1:1
!
ip vrf VRF-Blue
    ip route-export  3:3
    ip route-import  1:1
!
ip vrf VRF-Green
!
ip vrf VRF-shared
    ip route-export  1:1
    ip route-import  2:2
    ip route-import  3:3
```

Show routing tables of all the VRFs (without any route-export and route-import tags being configured)

```
DellEMC# show ip route vrf VRF-Red
O     11.1.1.1/32     via 111.1.1.1  110/0    00:00:10
C     111.1.1.0/24    Direct, Te 1/11 0/0       22:39:59

DellEMC# show ip route vrf VRF-Blue
O     22.2.2.2/32     via 122.2.2.2 110/0    00:00:11



C     122.2.2.0/24    Direct, Te 1/12 0/0       22:39:61

DellEMC# show ip route vrf VRF-Green
O     33.3.3.3/32     via 133.3.3.3  110/0    00:00:11



C     133.3.3.0/24    Direct, Te 1/13 0/0       22:39:61
```

```
DellEMC# show ip route vrf VRF-Shared
O      44.4.4.4/32         via 144.4.4.4 110/0     00:00:11



C      144.4.4.0/24        Direct, Te 1/4 0/0       00:32:36
```

Show routing tables of VRFs(after route-export and route-import tags are configured).

```
DellEMC# show ip route vrf VRF-Red

O      11.1.1.1/32       via 111.1.1.1   110/0    00:00:10
C      111.1.1.0/24      Direct, Te 1/11 0/0      22:39:59
O      44.4.4.4/32       via VRF-shared:144.4.4.4 0/0      00:32:36
C      144.4.4.0/24      Direct, VRF-shared:Te 1/4 0/0    00:32:36

DellEMC# show ip route vrf VRF-Blue
O      22.2.2.2/32      via 122.2.2.2       110/0    00:00:11



C      122.2.2.0/24     Direct, Te 1/12     0/0       22:39:61
O      44.4.4.4/32      via vrf-shared:144.4.4.4 0/0      00:32:36
C      144.4.4.0/24     Direct, vrf-shared:Te 1/4  0/0   00:32:36

DellEMC# show ip route vrf VRF-Green

O      33.3.3.3/32         via 133.3.3.3  110/0    00:00:11



C      133.3.3.0/24        Direct, Te 1/13 0/0     22:39:61

DellEMC# show ip route vrf VRF-Shared
O     11.1.1.1/32           via VRF-Red:111.1.1.1 110/0    00:00:10
C     111.1.1.0/24          Direct, VRF-Red:Te 1/11 0/0    22:39:59
O      22.2.2.2/32          via VRF-Blue:122.2.2.2  110/0  00:00:11
C      122.2.2.0/24         Direct, VRF-Blue:Te 1/22 0/0   22:39:61
O      44.4.4.4/32          via 144.4.4.4    110/0    00:00:11



C     144.4.4.0/24          Direct, Te 1/4    0/0    00:32:36
```

Important Points to Remember

- If the target VRF conatins the same prefix as either the sourced or Leaked route from some other VRF, then route Leaking for that particular prefix fails and the following error-log is thrown.

  ```
  SYSLOG ("Duplicate prefix found %s in the target VRF %d", address, import_vrf_id)
  with
                      The type/level is EVT_LOGWARNING.
  ```

- The source routes always take precedence over leaked routes. The leaked routes are deleted as soon as routes are locally learnt by the VRF using other means.
- For recovery, you must take appropriate action either by deleting the unwanted prefixes or issuing clear command or both.
- In the target VRF, you cannot leak routes that are imported through the route leaking feature.
- The leaked route points to the next-hop of the source routes. You cannot do any modifications to the next-hop of the leaked route in the destination VRF.
- IPv6 link local routes will never be leaked from one VRF to another.

Configuring Route Leaking with Filtering

When you initalize route leaking from one VRF to another, all the routes are exposed to the target VRF. If the size of the source VRF's RTM is considerablly large, an import operation results in the duplication of the target VRF's RTM with the source RTM entries. To mitigate this issue, you can use route-maps to filter the routes that are exported and imported into the route targets based on certain matching criteria. These match criteria include, prefix matches and portocol matches.

You can use the `match source-protocol` or `match ip-address` commands to specify matching criteria for importing or exporting routes between VRFs.

NOTE: You must use the match source-protocol or match ip-address commands in conjunction with the route-map command to be able to define the match criteria for route leaking.

Consider a scenario where you have created two VRF tables VRF-red and VRF-blue. VRF-red exports routes with the export_ospfbgp_protocol route-map to VRF-blue. VRF-blue imports these routes into its RTM.

For leaking these routes from VRF-red to VRF-blue, you can use the ip route-export route-map command on VRF-red (source VRF, that is exporting the routes); you must also specify a match criteria for these routes using the match source-protocol command. When you leak these routes into VRF-blue, only the routes (OSPF and BGP) that satisfy the matching criteria defined in route-map export_ospfbgp_protocol are exposed to VRF-blue.

While importing these routes into VRF-blue, you can further specify match conditions at the import end to define the filtering criteria based on which the routes are imported into VRF-blue. You can define a route-map import_ospf_protocol and then specify the match criteria as OSPF using the match source-protocol ospf command.

You can then use the ip route-import route-map command to import routes matching the filtering criteria defined in the import_ospf_protocol route-map. For a reply communication, VRF-blue is configured with a route-export tag. This value is then configured as route-import tag on the VRF-Red.

To configure route leaking using filtering criteria, perform the following steps:

1. Configure VRF-red:

   ```
   ip vrf vrf-red
   interface-type slot/port
   ip vrf forwarding VRF-red
   ip address ip—address mask
   ```

   A non-default VRF named VRF-red is created and the interface is assigned to this VRF.
2. Define a route-map export_ospfbgp_protocol.

   ```
   DellEMC(config)route-map export_ospfbgp_protocol permit 10
   ```
3. Define the matching criteria for the exported routes.

   ```
   DellEMC(config-route-map)match source-protocol ospf
   ```

   DellEMC(config-route-map)match source-protocol bgp

   This action specifies that the route-map contains OSPF and BGP as the matching criteria for exporting routes from vrf-red.
4. Configure the export target in the source VRF with route-map export_ospfbgp_protocol.

   ```
   ip route-export 1:1 export_ospfbgp_protocol
   ```
5. Configure VRF-blue.

   ```
   ip vrf vrf-blue
   interface-type slot/port
   ip vrf forwarding VRF-blue
   ip address ip—address mask
   ```

   A non-default VRF named VRF-blue is created and the interface 1/22 is assigned to it.
6. Define the route-map import_ospf_protocol.

   ```
   DellEMC(config)route-map import_ospf_protocol permit 10
   ```
7. Define the matching criteria for importing routes into VRF-blue.

   ```
   DellEMC(config-route-map)match source-protocol ospf
   ```

   This action specifies that the route-map contains OSPF as the matching criteria for importing routes into vrf-blue.
8. Configure the import target in VRF-blue with route-map import_ospf_protociol.

   ```
   ip route-import 1:1 import_ospf_protocol
   ```

   When you import routes into VRF-blue using the route-map import_ospf_protocol, only OSPF routes are imported into VRF-blue. Even though VRF-red has leaked both OSPF as well as BGP routes to be shared with other VRFs, this command imports only OSPF routes into VRF-blue.
9. Configure the import target in the source VRF for reverse communnication with the destination VRF.

   ```
   ip route-import 2:2
   ```

The show run output for the above configuration is as follows:

```
ip vrf vrf-Red
ip route-export     1:1 export_ospfbgp_protocol
ip route-import     2:2
! this action exports only the OSPF and BGP routes to other VRFs
!
ip vrf vrf-Blue
  ip route-export   2:2
  ip route-import   1:1 import_ospf_protocol
!this action accepts only OSPF routes from VRF-red even though both OSPF as well as BGP
routes are shared
```

The show VRF commands displays the following output:

```
DellEMC# show ip route vrf VRF-Blue
C     122.2.2.0/24 Direct, Te 1/22  0/0      22:39:61
O     22.2.2.2/32  via 122.2.2.2    110/0    00:00:11



O     44.4.4.4/32  via vrf-red:144.4.4.4   0/0    00:32:36
                                  << only OSPF and BGP leaked from VRF-red
```

Important Points to Remember

- Only Active routes are eligible for leaking. For example, if VRF-A has two routes from BGP and OSPF, in which the BGP route is not active. In this scenario, the OSPF route takes precedence over BGP. Even though the Target VRF-B has specified filtering options to match BGP, the BGP route is not leaked as that route is not active in the Source VRF.
- The export-target and import-target support only the match protocol and match prefix-list options. Other options that are configured in the route-maps are ignored.
- You can expose a unique set of routes from the Source VRF for Leaking to other VRFs. For example, in VRF-red there is no option for exporting one set of routes (for example, OSPF) to VRF- blue and another set of routes (for example, BGP routes) to some other VRF. Similarly, when two VRFs leak or export routes, there is no option to discretely filter leaked routes from each source VRF. Meaning, you cannot import one set of routes from VRF-red and another set of routes from VRF-blue.

63

Virtual Router Redundancy Protocol (VRRP)

Virtual router redundancy protocol (VRRP) is designed to eliminate a single point of failure in a statically routed network.

Topics:

- VRRP Overview
- VRRP Benefits
- VRRP Implementation
- VRRP Configuration
- Sample Configurations

VRRP Overview

VRRP is designed to eliminate a single point of failure in a statically routed network.

VRRP specifies a MASTER router that owns the next hop IP and MAC address for end stations on a local area network (LAN). The MASTER router is chosen from the virtual routers by an election process and forwards packets sent to the next hop IP address. If the MASTER router fails, VRRP begins the election process to choose a new MASTER router and that new MASTER continues routing traffic.

VRRP uses the virtual router identifier (VRID) to identify each virtual router configured. The IP address of the MASTER router is used as the next hop address for all end stations on the LAN. The other routers the IP addresses represent are BACKUP routers.

VRRP packets are transmitted with the virtual router MAC address as the source MAC address. The MAC address is in the following format: 00-00-5E-00-01-{VRID}. The first three octets are unchangeable. The next two octets (00-01) indicate the address block assigned to the VRRP protocol, and are unchangeable. The final octet changes depending on the VRRP virtual router identifier and allows for up to 255 VRRP routers on a network.

The following example shows a typical network configuration using VRRP. Instead of configuring the hosts on the network 10.10.10.0 with the IP address of either Router A or Router B as their default router; their default router is the IP address configured on the virtual router. When any host on the LAN segment wants to access the Internet, it sends packets to the IP address of the virtual router.

In the following example, Router A is configured as the MASTER router. It is configured with the IP address of the virtual router and sends any packets addressed to the virtual router through interface 1/1 to the Internet. As the BACKUP router, Router B is also configured with the IP address of the virtual router. If, for any reason, Router A becomes unavailable, VRRP elects a new MASTER Router. Router B assumes the duties of Router A and becomes the MASTER router. At that time, Router B responds to the packets sent to the virtual IP address.

All workstations continue to use the IP address of the virtual router to address packets destined to the Internet. Router B receives and forwards them on interface 1/1. Until Router A resumes operation, VRRP allows Router B to provide uninterrupted service to the users on the LAN segment accessing the Internet.

For more detailed information about VRRP, refer to *RFC 2338, Virtual Router Redundancy Protocol*.

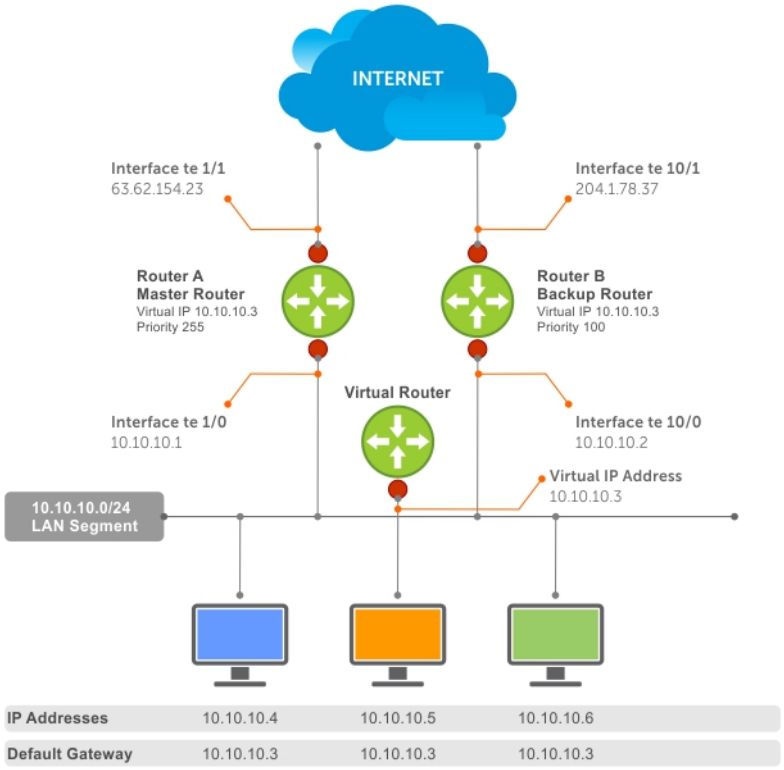


Figure 157. Basic VRRP Configuration

VRRP Benefits

With VRRP configured on a network, end-station connectivity to the network is not subject to a single point-of-failure. End-station connections to the network are redundant and are not dependent on internal gateway protocol (IGP) protocols to converge or update routing tables.

In conjunction with Virtual Link Trunking (VLT), you can configure optimized forwarding with virtual router redundancy protocol (VRRP).

NOTE: In a VLT environment, VRRP configuration acts as active-active and if route is not present in any of the VRRP nodes, the packet to the destination is dropped on that VRRP node.

VRRP Implementation

Within a single VRRP group, up to 12 virtual IP addresses are supported.

Virtual IP addresses can belong to the primary or secondary IP address' subnet configured on the interface. You can ping all the virtual IP addresses configured on the Master VRRP router from anywhere in the local subnet.

The S-Series supports varying number of maximum VRRP groups per interface. The supports a total of 2000 VRRP groups on a switch and 512 VRRP groups per interface.

The following recommendations shown may vary depending on various factors like address resolution protocol (ARP) broadcasts, IP broadcasts, or spanning tree protocol (STP) before changing the advertisement interval. When the number of packets processed by RP2/CP/FP processor increases or decreases based on the dynamics of the network, the advertisement intervals may increase or decrease accordingly.

CAUTION: Increasing the advertisement interval increases the VRRP Master dead interval, resulting in an increased failover time for Master/Backup election. Take caution when increasing the advertisement interval, as the increased dead interval may cause packets to be dropped during that switch-over time.

NOTE: In a VLT environment, VRRP configuration acts as active-active and if route is not present in any of the VRRP nodes, the packet to the destination is dropped on that VRRP node.

Table 132. Recommended VRRP Advertise Intervals

Total VRRP Groups	Recommended Advertise Interval	Groups/Interface
Less than 250	1 second	12
Between 250 and 450	2–3 seconds	24
Between 450 and 600	3–4 seconds	36
Between 600 and 800	4 seconds	48
Between 800 and 1000	5 seconds	84
Between 1000 and 1200	7 seconds	100
Between 1200 and 1500	8 seconds	120

VRRP Configuration

By default, VRRP is not configured.

Configuration Task List

The following list specifies the configuration tasks for VRRP.

- Creating a Virtual Router (mandatory)
- Configuring the VRRP Version for an IPv4 Group (optional)
- Assign Virtual IP Addresses mandatory)
- Setting VRRP Group (Virtual Router) Priority (optional)
- Configuring VRRP Authentication (optional)
- Disabling Preempt (optional)
- Changing the Advertisement Interval (optional)
- Track an Interface or Object
- Setting VRRP Initialization Delay

For a complete listing of all commands related to VRRP, refer to *Dell EMC Networking OS Command Line Reference Guide*.

Creating a Virtual Router

To enable VRRP, create a virtual router. In Dell EMC Networking Operating System (OS), the virtual router identifier (VRID) identifies a VRRP group.

To enable or delete a virtual router, use the following commands.

- Create a virtual router for that interface with a VRID.

  INTERFACE mode

  ```
  vrrp-group vrid
  ```

  The VRID range is from 1 to 255.

  NOTE: The interface must already have a primary IP address defined and be enabled, as shown in the second example.

- Delete a VRRP group.

  INTERFACE mode

  ```
  no vrrp-group vrid
  ```

The following examples how to configure VRRP.

```
DellEMC(conf)#interface tengigabitethernet 1/1
DellEMC(conf-if-te-1/1)#vrrp-group 111
DellEMC(conf-if-te-1/1-vrid-111)#
```

The following examples how to verify the VRRP configuration.

```
DellEMC(conf-if-te-1/1)#show conf
!
interface TenGigabitEthernet 1/1
  ip address 10.10.10.1/24
!
  vrrp-group 111
  no shutdown
```

Configuring the VRRP Version for an IPv4 Group

For IPv4, you can configure a VRRP group to use one of the following VRRP versions:

- VRRPv2 as defined in RFC 3768, *Virtual Router Redundancy Protocol (VRRP)*
- VRRPv3 as defined in RFC 5798, *Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6*

You can also migrate a IPv4 group from VRRPv2 to VRRP3.

To configure the VRRP version for IPv4, use the **version** command in INTERFACE mode.

Example: Configuring VRRP to Use Version 3

The following example configures the IPv4 VRRP 100 group to use VRRP protocol version 3.

```
DellEMC(conf-if-te-1/1)# vrrp-group 100
DellEMC(conf-if-te-1/1-vrid-100)#version ?
2       VRRPv2
3       VRRPv3
both    Interoperable, send VRRPv3 receive both
DellEMC(conf-if-te-1/1-vrid-100)#version 3
```

You can use the **version both** command in INTERFACE mode to migrate from VRRPv2 to VRRPv3. When you set the VRRP version to **both**, the switch sends only VRRPv3 advertisements but can receive VRRPv2 or VRRPv3 packets.

To migrate an IPv4 VRRP group from VRRPv2 to VRRPv3:

1. Set the switches with the lowest priority to “both”.
2. Set the switch with the highest priority to version to 3.
3. Set all the switches from **both** to version 3.

NOTE: Do not run VRRP version 2 and version 3 in the same group for an extended period of time

Example: Migrating an IPv4 VRRP Group from VRRPv2 to VRRPv3

NOTE: Carefully following this procedure, otherwise you might introduce dual master switches issues.

To migrate an IPv4 VRRP Group from VRRPv2 to VRRPv3:

1. Set the backup switches to VRRP version to both.

```
Dell_backup_switch1(conf-if-te-1/1-vrid-100)#version both
Dell_backup_switch2(conf-if-te-1/2-vrid-100)#version both
```

2. Set the master switch to VRRP protocol version 3.

```
Dell_master_switch(conf-if-te-1/1-vrid-100)#version 3
```

3. Set the backup switches to version 3.

```
Dell_backup_switch1(conf-if-te-1/1-vrid-100)#version 3
Dell_backup_switch2(conf-if-te-1/2-vrid-100)#version 3
```

Assign Virtual IP addresses

Virtual routers contain virtual IP addresses configured for that VRRP group (VRID). A VRRP group does not transmit VRRP packets until you assign the Virtual IP address to the VRRP group.

To activate a VRRP group on an interface (so that VRRP group starts transmitting VRRP packets), configure at least one virtual IP address in a VRRP group. The virtual IP address is the IP address of the virtual router and does not require the IP address mask.

You can configure up to 12 virtual IP addresses on a single VRRP group (VRID).

The following rules apply to virtual IP addresses:

- The virtual IP addresses must be in the same subnet as the primary or secondary IP addresses configured on the interface. Though a single VRRP group can contain virtual IP addresses belonging to multiple IP subnets configured on the interface, Dell EMC Networking recommends configuring virtual IP addresses belonging to the same IP subnet for any one VRRP group.
  - For example, an interface (on which you enable VRRP) contains a primary IP address of 50.1.1.1/24 and a secondary IP address of 60.1.1.1/24. The VRRP group (VRID 1) must contain virtual addresses belonging to either subnet 50.1.1.0/24 or subnet 60.1.1.0/24, but not from both subnets (though Dell EMC Networking OS allows the same).
- If the virtual IP address and the interface's primary/secondary IP address are the same, the priority on that VRRP group MUST be set to 255. The interface then becomes the OWNER router of the VRRP group and the interface's physical MAC address is changed to that of the owner VRRP group's MAC address.
- If you configure multiple VRRP groups on an interface, only one of the VRRP Groups can contain the interface primary or secondary IP address.
- On a stack system, if a force failover is performed on a master stack unit, the VRRP virtual addresses are disabled. To re-enable VRRP, execute the `mac-address-table station-move refresh-arp` command.

Configuring a Virtual IP Address

To configure a virtual IP address, use the following commands.

1. Configure a VRRP group.

   INTERFACE mode

   `vrrp-group` *`vrrp-id`*

   The VRID range is from 1 to 255.

2. Configure virtual IP addresses for this VRID.

   INTERFACE -VRID mode

   `virtual-address` *`ip-address1`* `[`*`...ip-address12`*`]`

   The range is up to 12 addresses.

The following example shows how to configure a virtual IP address.

```
DellEMC(conf-if-te-1/1-vrid-111)#virtual-address 10.10.10.1
DellEMC(conf-if-te-1/1-vrid-111)#virtual-address 10.10.10.2
DellEMC(conf-if-te-1/1-vrid-111)#virtual-address 10.10.10.3
```

The following example shows how to verify a virtual IP address configuration.

NOTE: In the following example, the primary IP address and the virtual IP addresses are on the same subnet.

```
DellEMC(conf-if-te-1/1)#show conf
!
interface TenGigabitEthernet 1/1
   ip address 10.10.10.1/24
!
vrrp-group 111
```

```
  priority 255
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
!
vrrp-group 222
  no shutdown
```

The following example shows the same VRRP group (VRID 111) configured on multiple interfaces on different subnets.

```
DellEMC#show vrrp
------------------
TenGigabitEthernet 1/1, VRID: 111, Version: 2 Net: 10.10.10.1
VRF: 0 default
State: Master, Priority: 255, Master: 10.10.10.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 1768, Gratuitous ARP sent: 5
Virtual MAC address:
  00:00:5e:00:01:6f
Virtual IP address:
  10.10.10.1 10.10.10.2 10.10.10.3 10.10.10.10
Authentication: (none)
------------------
TenGigabitEthernet 1/2, VRID: 111, Version: 2 Net: 10.10.2.1
VRF: 0 default
State: Master, Priority: 100, Master: 10.10.2.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 27, Gratuitous ARP sent: 2
Virtual MAC address:
  00:00:5e:00:01:6f
Virtual IP address:
  10.10.2.2 10.10.2.3
Authentication:
```

When the VRRP process completes its initialization, the State field contains either Master or Backup.

Setting VRRP Group (Virtual Router) Priority

Setting a virtual router priority to 255 ensures that router is the “owner” virtual router for the VRRP group. VRRP elects the MASTER router by choosing the router with the highest priority.

The default priority for a virtual router is **100**. The higher the number, the higher the priority. If the MASTER router fails, VRRP begins the election process to choose a new MASTER router based on the next-highest priority.

If two routers in a VRRP group come up at the same time and have the same priority value, the interface’s physical IP addresses are used as tie-breakers to decide which is MASTER. The router with the higher IP address becomes MASTER.

To configure the VRRP group’s priority, use the following command.

- Configure the priority for the VRRP group.

  INTERFACE -VRID mode

  `priority` *priority*

  The range is from 1 to 255.

  The default is **100**.

```
DellEMC(conf-if-te-1/2)#vrrp-group 111
DellEMC(conf-if-te-1/2-vrid-111)#priority 125
```

To verify the VRRP group priority, use the `show vrrp` command.

```
Dellshow vrrp
------------------
TenGigabitEthernet 1/1, VRID: 111, Net: 10.10.10.1
VRF: 0 default
State: Master, Priority: 255, Master: 10.10.10.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 2343, Gratuitous ARP sent: 5
Virtual MAC address:
```

```
  00:00:5e:00:01:6f
Virtual IP address:
  10.10.10.1 10.10.10.2 10.10.10.3 10.10.10.10
Authentication: (none)
------------------
TenGigabitEthernet 1/2, VRID: 111, Net: 10.10.2.1
VRF: 0 default
State: Master, Priority: 125, Master: 10.10.2.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 601, Gratuitous ARP sent: 2
Virtual MAC address:
  00:00:5e:00:01:6f
Virtual IP address:
  10.10.2.2 10.10.2.3
Authentication: (none)
```

Configuring VRRP Authentication

Simple authentication of VRRP packets ensures that only trusted routers participate in VRRP processes.

When you enable authentication, Dell EMC Networking OS includes the password in its VRRP transmission. The receiving router uses that password to verify the transmission.

NOTE: You must configure all virtual routers in the VRRP group the same: you must enable authentication with the same password or authentication is disabled.

NOTE: Authentication for VRRPv3 is not supported.

To configure simple authentication, use the following command.

- Configure a simple text password.

  INTERFACE-VRID mode

  `authentication-type simple [encryption-type] password`

  Parameters:
  - *encryption-type*: `0` indicates unencrypted; `7` indicates encrypted.
  - *password*: plain text.

The bold section shows the encryption type (encrypted) and the password.

```
DellEMC(conf-if-te-1/1-vrid-111)#authentication-type ?
DellEMC(conf-if-te-1/1-vrid-111)#authentication-type simple 7 force10
```

The following example shows verifying the VRRP authentication configuration using the `show conf` command. The bold section shows the encrypted password.

```
DellEMC(conf-if-te-1/1-vrid-111)#show conf
!
  vrrp-group 111
    authentication-type simple 7 387a7f2df5969da4
    priority 255
    virtual-address 10.10.10.1
    virtual-address 10.10.10.2
    virtual-address 10.10.10.3
    virtual-address 10.10.10.10
```

Disabling Preempt

The `preempt` command is enabled by default. The command forces the system to change the MASTER router if another router with a higher priority comes online.

Prevent the BACKUP router with the higher priority from becoming the MASTER router by disabling preempt.

NOTE: You must configure all virtual routers in the VRRP group the same: you must configure all with preempt enabled or configure all with preempt disabled.

Because preempt is enabled by default, disable the preempt function with the following command.

- Prevent any BACKUP router with a higher priority from becoming the MASTER router.

  INTERFACE-VRID mode

  `no preempt`

Re-enable preempt by entering the `preempt` command. When you enable preempt, it does not display in the `show` commands, because it is a default setting.

The following example shows how to disable preempt using the `no preempt` command.

```
DellEMC(conf-if-te-1/1)#vrrp-group 111
DellEMC(conf-if-te-1/1-vrid-111)#no preempt
DellEMC(conf-if-te-1/1-vrid-111)#
```

The following example shows how to verify preempt is disabled using the `show conf` command.

```
DellEMC(conf-if-te-1/1-vrid-111)#show conf
!
  vrrp-group 111
    authentication-type simple 7 387a7f2df5969da4
    no preempt
    priority 255
    virtual-address 10.10.10.1
    virtual-address 10.10.10.2
    virtual-address 10.10.10.3
    virtual-address 10.10.10.10
```

Changing the Advertisement Interval

By default, the MASTER router transmits a VRRP advertisement to all members of the VRRP group every one second, indicating it is operational and is the MASTER router.

If the VRRP group misses three consecutive advertisements, the election process begins and the BACKUP virtual router with the highest priority transitions to MASTER.

NOTE: To avoid throttling VRRP advertisement packets, Dell EMC Networking OS recommends increasing the VRRP advertisement interval to a value higher than the default value of one second. If you do change the time interval between VRRP advertisements on one router, change it on all participating routers.

If are using VRRP version 2, you must configure the timer values in multiple of whole seconds. For example a timer value of 3 seconds or 300 centisecs are valid and equivalent. However, a time value of 50 centisecs is invalid because it not a multiple of 1 second. If you are using VRRP version 3, you must configure the timer values in multiples of 25 centisecs.

If you are configured for VRRP version 2, the timer values must be in multiples of whole seconds. For example, timer value of 3 seconds or 300 centisecs are valid and equivalent. However, a timer value of 50 centisecs is invalid because it not is not multiple of 1 second.

If are using VRRP version 3, you must configure the timer values in multiples of 25 centisecs.

To change the advertisement interval in seconds or centisecs, use the following command. A centisecs is 1/100 of a second.

- Change the advertisement interval setting.

  INTERFACE-VRID mode

  `advertise-interval` *`seconds`*

  The range is from 1 to 255 seconds.

  The default is **1 second**.
- For VRRPv3, change the advertisement centisecs interval setting.

  INTERFACE-VRID mode

  `advertise-interval centisecs` *`centisecs`*

  The range is from 25 to 4075 centisecs in units of 25 centisecs.

  The default is 100 centisecs.

The following example shows how to change the advertise interval using the `advertise-interval` command.

```
DellEMC(conf-if-te-1/1)#vrrp-group 111
DellEMC(conf-if-te-1/1-vrid-111)#advertise-interval 10
DellEMC(conf-if-te-1/1-vrid-111)#
```

The following example shows how to verify the advertise interval change using the `show conf` command.

```
DellEMC(conf-if-te-1/1-vrid-111)#show conf
!
  vrrp-group 111
    advertise-interval 10
    authentication-type simple 7 387a7f2df5969da4
    no preempt
    priority 255
    virtual-address 10.10.10.1
    virtual-address 10.10.10.2
    virtual-address 10.10.10.3
    virtual-address 10.10.10.10
```

Track an Interface or Object

You can set Dell EMC Networking OS to monitor the state of any interface according to the virtual group.

Each VRRP group can track up to 12 interfaces and up to 20 additional objects, which may affect the priority of the VRRP group. If the tracked interface goes down, the VRRP group's priority decreases by a default value of **10** (also known as *cost*). If the tracked interface's state goes up, the VRRP group's priority increases by 10.

The lowered priority of the VRRP group may trigger an election. As the Master/Backup VRRP routers are selected based on the VRRP group's priority, tracking features ensure that the best VRRP router is the Master for that group. The sum of all the costs of all the tracked interfaces must be less than the configured priority on the VRRP group. If the VRRP group is configured as Owner router (priority 255), tracking for that group is disabled, irrespective of the state of the tracked interfaces. The priority of the owner group always remains at 255.

For a virtual group, you can track the line-protocol state or the routing status of any of the following interfaces with the `interface` *`interface`* parameter:

- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` then the slot/port information.
- For a port channel interface, enter the keywords `port-channel` then a number.
- For a VLAN interface, enter the keyword `vlan` then a number from 1 to 4094.

For a virtual group, you can also track the status of a configured object (the `track` *`object-id`* command) by entering its object number.

NOTE: You can configure a tracked object for a VRRP group (using the `track` *`object-id`* command in INTERFACE-VRID mode) before you actually create the tracked object (using a `track` *`object-id`* command in CONFIGURATION mode). However, no changes in the VRRP group's priority occur until the tracked object is defined and determined to be down.

In addition, if you configure a VRRP group on an interface that belongs to a VRF instance and later configure object tracking on an interface for the VRRP group, the tracked interface must belong to the VRF instance.

Tracking an Interface

To track an interface, use the following commands.

NOTE: The sum of all the costs for all tracked interfaces must be less than the configured priority of the VRRP group.

- Monitor an interface and, optionally, set a value to be subtracted from the interface's VRRP group priority.

  INTERFACE-VRID mode

  `track` *`interface`* `[priority-cost` *`cost`*`]`

  The cost range is from 1 to 254.

  The default is **10**.

- (Optional) Display the configuration and the UP or DOWN state of tracked objects, including the client (VRRP group) that is tracking an object's state.

  EXEC mode or EXEC Privilege mode

  `show track`
- (Optional) Display the configuration and the UP or DOWN state of tracked interfaces and objects in VRRP groups, including the time since the last change in an object's state.

  EXEC mode or EXEC Privilege mode

  `show vrrp`
- (Optional) Display the configuration of tracked objects in VRRP groups on a specified interface.

  EXEC mode or EXEC Privilege mode

  `show running-config interface` *`interface`*

The following example shows how to configure tracking using the `track` command.

```
DellEMC(conf-if-te-1/1)#vrrp-group 111
DellEMC(conf-if-te-1/1-vrid-111)#track Tengigabitethernet 1/2
```

The following example shows how to verify tracking using the `show conf` command.

```
DellEMC(conf-if-te-1/1-vrid-111)#show conf
!
  vrrp-group 111
    advertise-interval 10
    authentication-type simple 7 387a7f2df5969da4
    no preempt
    priority 255
    track TenGigabitEthernet 1/2
    virtual-address 10.10.10.1
    virtual-address 10.10.10.2
    virtual-address 10.10.10.3
    virtual-address 10.10.10.10
```

The following example shows verifying the tracking status.

```
DellEMC#show track

Track 2
  IPv6 route 2040::/64 metric threshold
  Metric threshold is Up (STATIC/0/0)
    5 changes, last change 00:02:16
  Metric threshold down 255 up 254
  First-hop interface is TenGigabitEthernet 1/3
  Tracked by:
    VRRP TenGigabitEthernet 1/8 IPv6 VRID 1

Track 3
  IPv6 route 2050::/64 reachability
  Reachability is Up (STATIC)
    5 changes, last change 00:02:16
  First-hop interface is TenGigabitEthernet 1/3
  Tracked by:
    VRRP TenGigabitEthernet 1/8 IPv6 VRID 1
```

The following example shows verifying the VRRP status.

```
DellEMC#show vrrp
------------------

TenGigabitEthernet 1/8, IPv6 VRID: 1, Version: 3, Net: fe80::201:e8ff:fe01:95cc
VRF: 0 default
State: Master, Priority: 100, Master: fe80::201:e8ff:fe01:95cc (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 310
Virtual MAC address:
  00:00:5e:00:02:01
Virtual IP address:
```

```
  2007::1 fe80::1
Tracking states for 2 resource Ids:
 2 - Up IPv6 route, 2040::/64, priority-cost 20, 00:02:11
 3 - Up IPv6 route, 2050::/64, priority-cost 30, 00:02:11
```

The following example shows verifying the VRRP configuration on an interface.

```
DellEMC#show running-config interface tengigabitethernet 1/8

interface TenGigabitEthernet 1/8
  no ip address
  ipv6 address 2007::30/64

vrrp-ipv6-group 1
 track 2 priority-cost 20
 track 3 priority-cost 30
  virtual-address 2007::1
  virtual-address fe80::1
no shutdown
```

Setting VRRP Initialization Delay

When configured, VRRP is enabled immediately upon system reload or boot. You can delay VRRP initialization to allow the IGP and EGP protocols to be enabled prior to selecting the VRRP Master. This delay ensures that VRRP initializes with no errors or conflicts. You can configure the delay for up to 15 minutes, after which VRRP enables normally.

NOTE: When you reload a node that contains VRRP configuration and is enabled for VLT, Dell EMC Networking recommends that you configure the reload timer by using the `vrrp delay reload` command to ensure that VRRP is functional. Otherwise, when you reload a VLT node configured for VRRP, the local destination address is not seen on the reloaded node causing suboptimal routing.

Set the delay timer on individual interfaces. The delay timer is supported on all physical interfaces, VLANs, and LAGs.

When you configure both CLIs, the later timer rules VRRP enabling. For example, if you set `vrrp delay reload 600` and `vrrp delay minimum 300`, the following behavior occurs:

- When the system reloads, VRRP waits 600 seconds (10 minutes) to bring up VRRP on all interfaces that are up and configured for VRRP.
- When an interface comes up and becomes operational, the system waits 300 seconds (5 minutes) to bring up VRRP on that interface.

To set the delay time for VRRP initialization, use the following commands.

- Set the delay time for VRRP initialization on an individual interface.

  INTERFACE mode

  `vrrp delay minimum` *seconds*

  This time is the gap between an interface coming up and being operational, and VRRP enabling.

  The seconds range is from 0 to 900.

  The default is **0**.
- Set the delay time for VRRP initialization on all the interfaces in the system configured for VRRP.

  INTERFACE mode

  `vrrp delay reload` *seconds*

  This time is the gap between system boot up completion and VRRP enabling.

  The seconds range is from 0 to 900.

  The default is **0**.

VRRP for IPv6 Configuration

This section shows VRRP IPv6 topology with CLI configurations.

Consider an example VRRP for IPv6 configuration in which the IPv6 VRRP group consists of two routers.

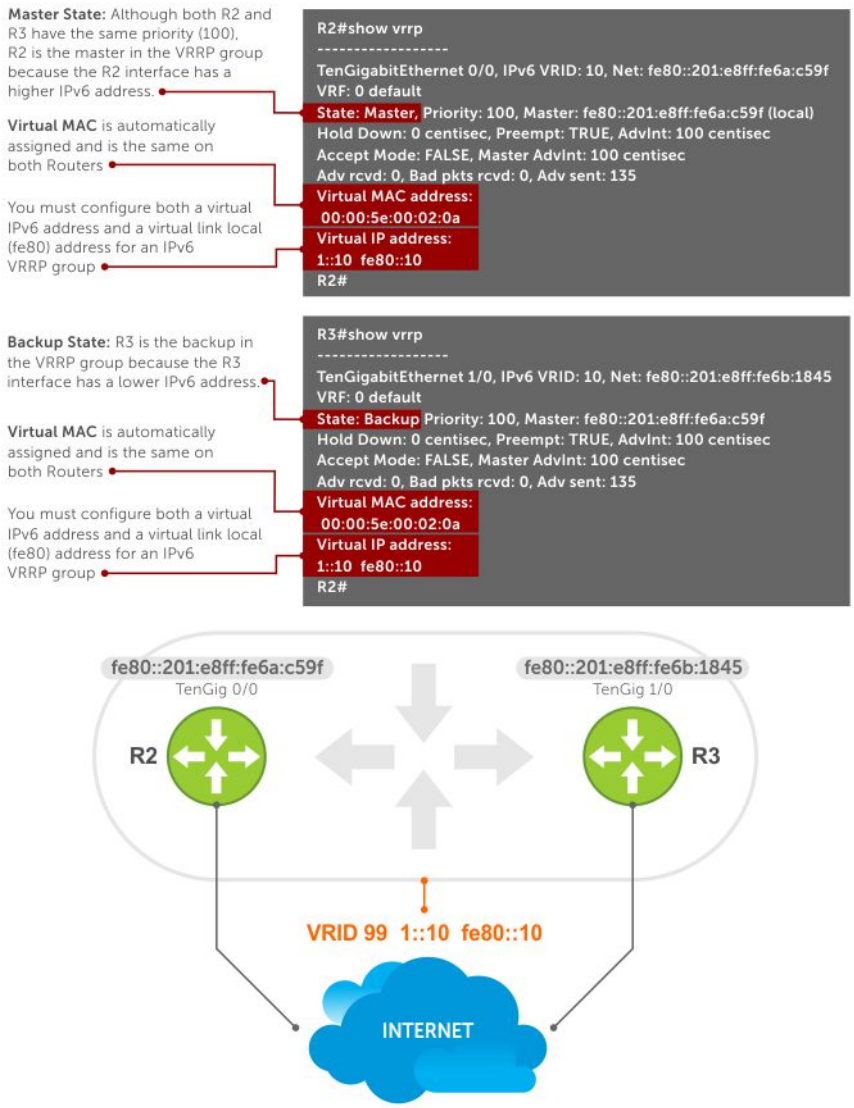


Figure 158. VRRP for IPv6 Topology

NOTE: This example does not contain comprehensive directions and is intended to provide guidance for only a typical VRRP configuration. You can copy and paste from the example to your CLI. Be sure you make the necessary changes to support your own IP addresses, interfaces, names, and so on.

NOTE: In a VRRP or VRRPv3 group, if two routers come up with the same priority and another router already has MASTER status, the router with master status continues to be master even if one of two routers has a higher IP or IPv6 address.

```
Router 2
R2(conf)#interface tengigabitethernet 1/1
R2(conf-if-te-1/1)#no ip address
R2(conf-if-te-1/1)#ipv6 address 1::1/64
R2(conf-if-te-1/1)#vrrp-group 10
```

NOTE: You must configure a virtual link local (fe80) address for each VRRPv3 group created for an interface. The VRRPv3 group becomes active as soon as you configure the link local address. Afterwards, you can configure the group's virtual IPv6 address.

```
R2(conf-if-te-1/1-vrid-10)#virtual-address fe80::10
```

NOTE: The virtual IPv6 address you configure should be the same as the IPv6 subnet to which the interface belongs.

```
R2(conf-if-te-1/1-vrid-10)#virtual-address 1::10
R2(conf-if-te-1/1-vrid-10)#no shutdown
R2(conf-if-te-1/1)#show config
interface TenGigabitEthernet 1/1
 ipv6 address 1::1/64
 vrrp-group 10
  priority 100
  virtual-address fe80::10
  virtual-address 1::10
 no shutdown
R2(conf-if-te-1/1)#end

R2#show vrrp
------------------
TenGigabitEthernet 1/1, IPv6 VRID: 10, Version: 3, Net:fe80::201:e8ff:fe6a:c59f
VRF: 0 default
State: Master, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 135
Virtual MAC address:
00:00:5e:00:02:0a
Virtual IP address:
1::10 fe80::10
```

NOTE: Although R2 and R3 have the same default priority (100), R2 is elected master in the VRRPv3 group because the Tengigabitethernet 1/1 interface has a higher IPv6 address than the Tengigabitethernet 1/2 interface on R3.

```
Router 3
R3(conf)#interface tengigabitethernet 1/2
R3(conf-if-te-1/2)#no ipv6 address
R3(conf-if-te-1/2)#ipv6 address 1::2/64
R3(conf-if-te-1/2)#vrrp-group 10
R2(conf-if-te-1/2-vrid-10)#virtual-address fe80::10
R2(conf-if-te-1/2-vrid-10)#virtual-address 1::10
R3(conf-if-te-1/2-vrid-10)#no shutdown
R3(conf-if-te-1/2)#show config
interface TenGigabitEthernet 1/2
 ipv6 address 1::2/64
 vrrp-group 10
  priority 100
  virtual-address fe80::10
  virtual-address 1::10
 no shutdown
R3(conf-if-te-1/2)#end

R3#show vrrp
------------------
TenGigabitEthernet 1/2, IPv6 VRID: 10, Version: 3, Net: fe80::201:e8ff:fe6b:1845
VRF: 0 default
State: Backup, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 11, Bad pkts rcvd: 0, Adv sent: 0
Virtual MAC address:
00:00:5e:00:02:0a
Virtual IP address:
1::10 fe80::10
```

```
DellEMC#show vrrp tengigabitethernet 0/0
TenGigabitEthernet 0/0, IPv6 VRID: 255, Version: 3, Net: fe80::201:e8ff:fe8a:fd76
VRF: 0 default
State: Backup, Priority: 90, Master: fe80::201:e8ff:fe8a:e9ed
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 214, Bad pkts rcvd: 0, Adv sent: 0
Virtual MAC address:
```

```
00:00:5e:00:02:ff
Virtual IP address:
10:1:1::255 fe80::255


DellEMC#show vrrp tengigabitethernet 2/8
TenGigabitEthernet 2/8, IPv6 VRID: 255, Version: 3, Net: fe80::201:e8ff:fe8a:e9ed
VRF: 0 default
State: Master, Priority: 110, Master: fe80::201:e8ff:fe8a:e9ed (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 120
Virtual MAC address:
00:00:5e:00:02:ff
Virtual IP address:
10:1:1::255 fe80::255
DellEMC#


DellEMC#show vrrp vrf vrf1 vlan 400
Vlan 400, IPv6 VRID: 255, Version: 3, Net: fe80::201:e8ff:fe8a:e9ed
VRF: 1 vrf1
State: Master, Priority: 200, Master: fe80::201:e8ff:fe8a:e9ed (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 339
Virtual MAC address:
00:00:5e:00:02:ff
Virtual IP address:
10:1:1::255 fe80::255

Vlan 400, IPv6 VRID: 255, Version: 3, Net: fe80::201:e8ff:fe8a:fd76
VRF: 1 vrf1
State: Backup, Priority: 90, Master: fe80::201:e8ff:fe8a:e9ed
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 399, Bad pkts rcvd: 0, Adv sent: 0
Virtual MAC address:
00:00:5e:00:02:ff
Virtual IP address:
10:1:1::255 fe80::255


DellEMC#show vrrp vrf vrf2 port-channel 1
Port-channel 1, IPv6 VRID: 255, Version: 3, Net: fe80::201:e8ff:fe8a:e9ed
VRF: 2 vrf2
State: Master, Priority: 100, Master: fe80::201:e8ff:fe8a:e9ed (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 443
Virtual MAC address:
00:00:5e:00:02:ff
Virtual IP address:
10:1:1::255 fe80::255


DellEMC#show vrrp vrf vrf2 port-channel 1
Port-channel 1, IPv6 VRID: 255, Version: 3, Net: fe80::201:e8ff:fe8a:fd76
VRF: 2 vrf2
State: Backup, Priority: 90, Master: fe80::201:e8ff:fe8a:e9ed
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 548, Bad pkts rcvd: 0, Adv sent: 0
Virtual MAC address:
00:00:5e:00:02:ff
Virtual IP address:
10:1:1::255 fe80::255
```

Sample Configurations

Before you set up VRRP, review the following sample configurations.

VRRP for an IPv4 Configuration

The following configuration shows how to enable IPv4 VRRP. This example does not contain comprehensive directions and is intended to provide guidance for only a typical VRRP configuration. You can copy and paste from the example to your CLI. To support your own IP addresses, interfaces, names, and so on, be sure that you make the necessary changes. The VRRP topology was created using the CLI configuration shown in the following example.

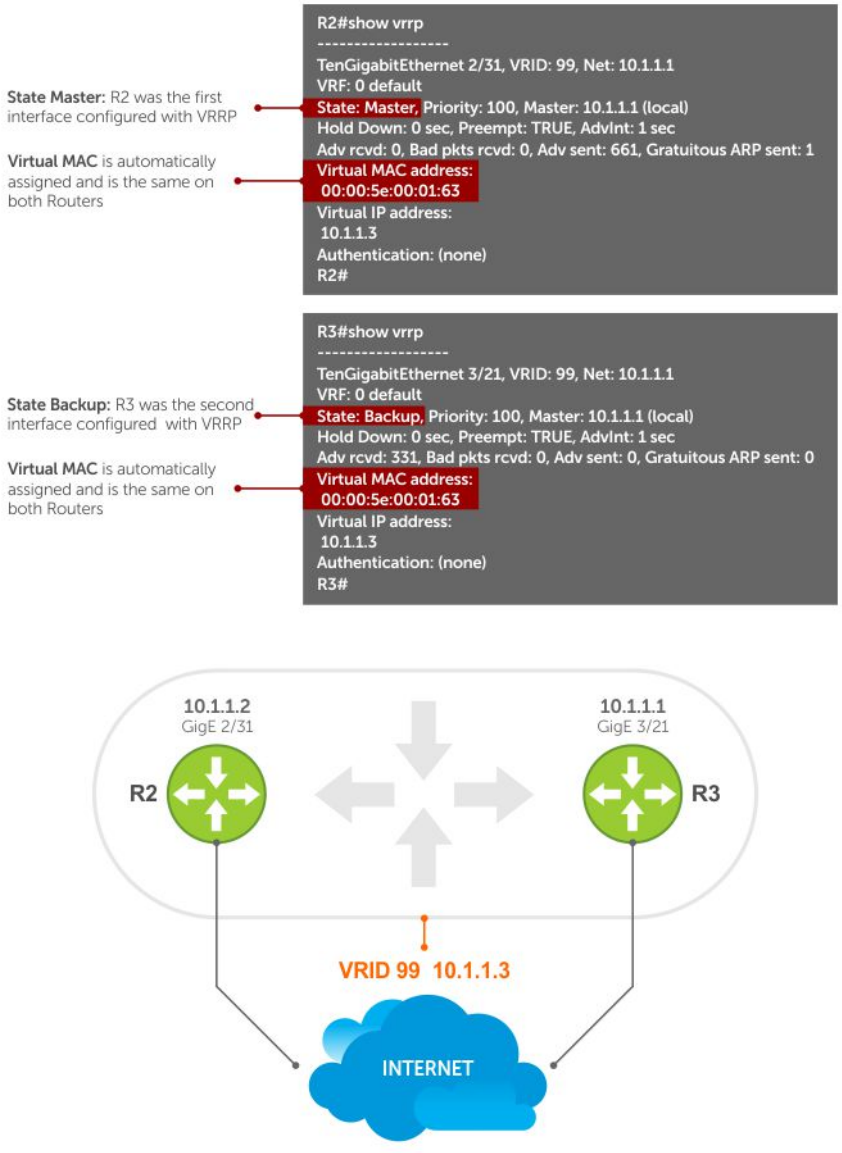


Figure 159. VRRP for IPv4 Topology

Examples of Configuring VRRP for IPv4 and IPv6

The following example shows configuring VRRP for IPv4 Router 2.

```
R2(conf)#interface tengigabitethernet 2/31
R2(conf-if-te-2/31)#ip address 10.1.1.1/24
R2(conf-if-te-2/31)#vrrp-group 99
R2(conf-if-te-2/31-vrid-99)#priority 200
```

```
R2(conf-if-te-2/31-vrid-99)#virtual 10.1.1.3
R2(conf-if-te-2/31-vrid-99)#no shut
R2(conf-if-te-2/31)#show conf
!
interface TenGigabitEthernet 2/31
  ip address 10.1.1.1/24
!
  vrrp-group 99
    priority 200
    virtual-address 10.1.1.3
  no shutdown
R2(conf-if-te-2/31)#end

R2#show vrrp
------------------
TenGigabitEthernet 2/31, VRID: 99, Net: 10.1.1.1
VRF: 0 default
State: Master, Priority: 200, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 817, Gratuitous ARP sent: 1
Virtual MAC address:
  00:00:5e:00:01:63
Virtual IP address:
  10.1.1.3
Authentication: (none)
R2#
Router 3
R3(conf)#interface tengigabitethernet 3/21
R3(conf-if-te-3/21)#ip address 10.1.1.2/24
R3(conf-if-te-3/21)#vrrp-group 99
R3(conf-if-te-3/21-vrid-99)#virtual 10.1.1.3
R3(conf-if-te-3/21-vrid-99)#no shut
R3(conf-if-te-3/21)#show conf
!
interface TenGigabitEthernet 3/21
  ip address 10.1.1.1/24
!
  vrrp-group 99
   virtual-address 10.1.1.3
  no shutdown
R3(conf-if-te-3/21)#end
R3#show vrrp
------------------
TenGigabitEthernet 3/21, VRID: 99, Net: 10.1.1.2
VRF: 0 default
State: Backup, Priority: 100, Master: 10.1.1.1
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 698, Bad pkts rcvd: 0, Adv sent: 0, Gratuitous ARP sent: 0
Virtual MAC address:
  00:00:5e:00:01:63
Virtual IP address:
  10.1.1.3
Authentication: (none)
```

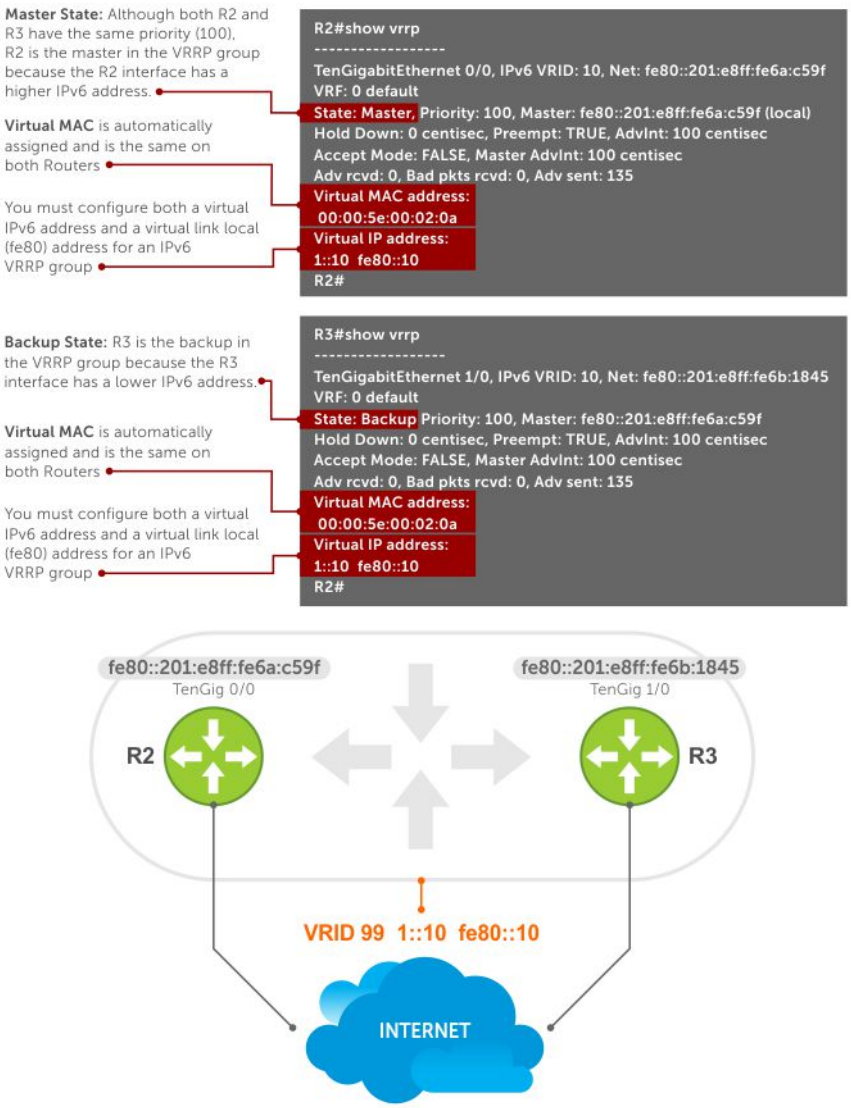


Figure 160. VRRP for an IPv6 Configuration

NOTE: In a VRRP or VRRPv3 group, if two routers come up with the same priority and another router already has MASTER status, the router with master status continues to be MASTER even if one of two routers has a higher IP or IPv6 address.

The following example shows configuring VRRP for IPv6 Router 2 and Router 3.

Configure a virtual link local (fe80) address for each VRRPv3 group created for an interface. The VRRPv3 group becomes active as soon as you configure the link local address. Afterward, you can configure the group's virtual IPv6 address.

The virtual IPv6 address you configure must be the same as the IPv6 subnet to which the interface belongs.

Although R2 and R3 have the same default, priority (100), R2 is elected master in the VRRPv3 group because the `TenGigabitethernet 1/1` interface has a higher IPv6 address than the `TenGigabitethernet 1/2` interface on R3.

Router 2

```
R2(conf)#interface tengigabitethernet 1/1
R2(conf-if-te-1/1)#no ip address
R2(conf-if-te-1/1)#ipv6 address 1::1/64
R2(conf-if-te-1/1)#vrrp-group 10
R2(conf-if-te-1/1-vrid-10)#virtual-address fe80::10
R2(conf-if-te-1/1-vrid-10)#virtual-address 1::10
R2(conf-if-te-1/1-vrid-10)#no shutdown
R2(conf-if-te-1/1)#show config
interface TenGigabitEthernet 1/1
  ipv6 address 1::1/64
```

```
  vrrp-group 10
    priority 100
    virtual-address fe80::10
    virtual-address 1::10
  no shutdown
R2(conf-if-te-1/1)#end

R2#show vrrp
------------------
TenGigabitEthernet 1/1, IPv6 VRID: 10, Version: 3, Net:fe80::201:e8ff:fe6a:c59f
VRF: 0 default
State: Master, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 135
Virtual MAC address:
00:00:5e:00:02:0a
Virtual IP address:
1::10 fe80::10
```

Router 3

```
R3(conf)#interface tengigabitethernet 1/2
R3(conf-if-te-1/2)#no ipv6 address
R3(conf-if-te-1/2)#ipv6 address 1::2/64
R3(conf-if-te-1/2)#vrrp-group 10
R2(conf-if-te-1/2-vrid-10)#virtual-address fe80::10
R2(conf-if-te-1/2-vrid-10)#virtual-address 1::10
R3(conf-if-te-1/2-vrid-10)#no shutdown
R3(conf-if-te-1/2)#show config
interface TenGigabitEthernet 1/2
  ipv6 address 1::2/64
  vrrp-group 10
    priority 100
    virtual-address fe80::10
    virtual-address 1::10
  no shutdown
R3(conf-if-te-1/2)#end

R3#show vrrp
------------------
TenGigabitEthernet 1/2, IPv6 VRID: 10, Version: 3, Net:
VRF: 0 default
fe80::201:e8ff:fe6b:1845
VRF: 0 default-vrf
State: Backup, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 11, Bad pkts rcvd: 0, Adv sent: 0
Virtual MAC address:
00:00:5e:00:02:0a
```

VRRP in a VRF Configuration

The following example shows how to enable VRRP operation in a VRF virtualized network for the following scenarios.

- Multiple VRFs on physical interfaces running VRRP.
- Multiple VRFs on VLAN interfaces running VRRP.

To view a VRRP in a VRF configuration, use the `show` commands.

VRRP in a VRF: Non-VLAN Scenario

The following example shows how to enable VRRP in a non-VLAN.

The following example shows a typical use case in which you create three virtualized overlay networks by configuring three VRFs in two switches. The default gateway to reach the Internet in each VRF is a static route with the next hop being the virtual IP address configured in VRRP. In this scenario, a single VLAN is associated with each VRF.

Both Switch-1 and Switch-2 have three VRF instances defined: VRF-1, VRF-2, and VRF-3. Each VRF has a separate physical interface to a LAN switch and an upstream VPN interface to connect to the Internet. Both Switch-1 and Switch-2 use VRRP groups on each VRF instance in order that there is one MASTER and one backup router for each VRF. In VRF-1 and VRF-2, Switch-2 serves as owner-master of the VRRP group and Switch-1 serves as the backup. On VRF-3, Switch-1 is the owner-master and Switch-2 is the backup.

In VRF-1 and VRF-2 on Switch-2, the virtual IP and node IP address, subnet, and VRRP group are the same. On Switch-1, the virtual IP address, subnet, and VRRP group are the same in VRF-1 and VRF-2, but the IP address of the node interface is unique. There is no requirement for the virtual IP and node IP addresses to be the same in VRF-1 and VRF-2; similarly, there is no requirement for the IP addresses to be different. In VRF-3, the node IP addresses and subnet are unique.

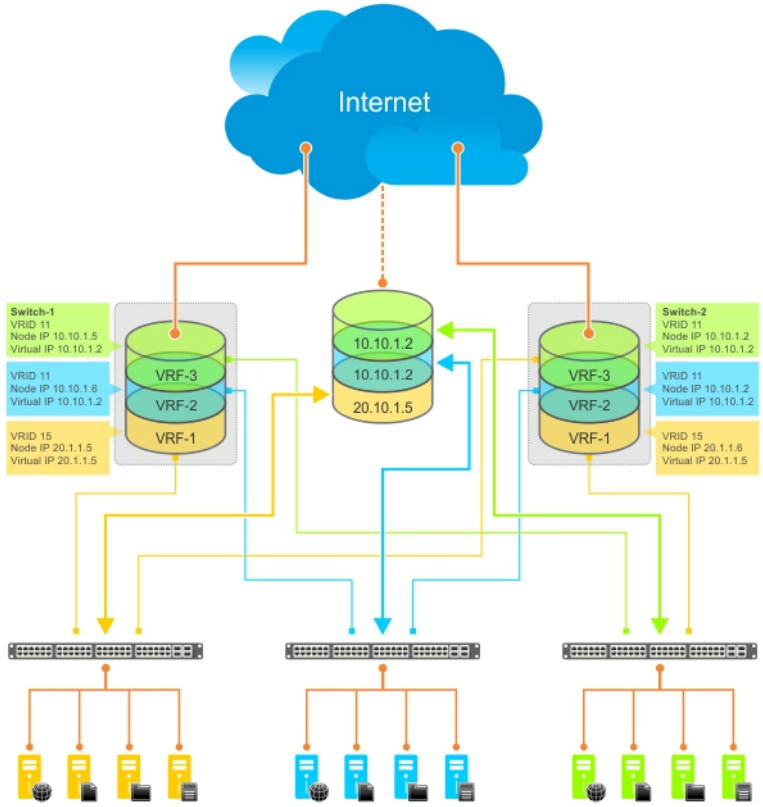


Figure 161. VRRP in a VRF: Non-VLAN Example

Example of Configuring VRRP in a VRF on Switch-1 (Non-VLAN)

Switch-1

```
S1(conf)#ip vrf default-vrf 0
!
S1(conf)#ip vrf VRF-1 1
!
S1(conf)#ip vrf VRF-2 2
!
S1(conf)#ip vrf VRF-3 3
!
S1(conf)#interface TenGigabitEthernet 1/1
S1(conf-if-te-1/1)#ip vrf forwarding VRF-1
S1(conf-if-te-1/1)#ip address 10.10.1.5/24
S1(conf-if-te-1/1)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S1(conf-if-te-1/1-vrid-101)#priority 100
S1(conf-if-te-1/1-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-te-1/1)#no shutdown
!
S1(conf)#interface TenGigabitEthernet 1/2
S1(conf-if-te-1/2)#ip vrf forwarding VRF-2
S1(conf-if-te-1/2)#ip address 10.10.1.6/24
S1(conf-if-te-1/2)#vrrp-group 11
```

```
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S1(conf-if-te-1/2-vrid-101)#priority 100
S1(conf-if-te-1/2-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-te-1/2)#no shutdown
!
S1(conf)#interface TenGigabitEthernet 1/3
S1(conf-if-te-1/3)#ip vrf forwarding VRF-3
S1(conf-if-te-1/3)#ip address 20.1.1.5/24
S1(conf-if-te-1/3)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S1(conf-if-te-1/3-vrid-105)#priority 255
S1(conf-if-te-1/3-vrid-105)#virtual-address 20.1.1.5
S1(conf-if-te-1/3)#no shutdown

DellEMC#show vrrp tengigabitethernet 2/8
------------------
TenGigabitEthernet 2/8, IPv4 VRID: 1, Version: 2, Net: 10.1.1.1
VRF: 0 default
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 119, Gratuitous ARP sent: 1
Virtual MAC address:
00:00:5e:00:01:01
Virtual IP address:
10.1.1.100
Authentication: (none)
```

Example of Configuring VRRP in a VRF on Switch-2 (Non-VLAN Configuration)

Switch-2

```
S2(conf)#ip vrf default-vrf 0
!
S2(conf)#ip vrf VRF-1 1
!
S2(conf)#ip vrf VRF-2 2
!
S2(conf)#ip vrf VRF-3 3
!
S2(conf)#interface TenGigabitEthernet 1/1
S2(conf-if-te-1/1)#ip vrf forwarding VRF-1
S2(conf-if-te-1/1)#ip address 10.10.1.2/24
S2(conf-if-te-1/1)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S2(conf-if-te-1/1-vrid-101)#priority 255
S2(conf-if-te-1/1-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-te-1/1)#no shutdown
!
S2(conf)#interface TenGigabitEthernet 1/2
S2(conf-if-te-1/2)#ip vrf forwarding VRF-2
S2(conf-if-te-1/2)#ip address 10.10.1.2/24
S2(conf-if-te-1/2)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S2(conf-if-te-1/2-vrid-101)#priority 255
S2(conf-if-te-1/2-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-te-1/2)#no shutdown
!
S2(conf)#interface TenGigabitEthernet 1/3
S2(conf-if-te-1/3)#ip vrf forwarding VRF-3
S2(conf-if-te-1/3)#ip address 20.1.1.6/24
S2(conf-if-te-1/3)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S2(conf-if-te-1/3-vrid-105)#priority 100
S2(conf-if-te-1/3-vrid-105)#virtual-address 20.1.1.5
S2(conf-if-te-1/3)#no shutdown
```

VLAN Scenario

In another scenario, to connect to the LAN, VRF-1, VRF-2, and VRF-3 use a single physical interface with multiple tagged VLANs (instead of separate physical interfaces).

In this case, you configure three VLANs: VLAN-100, VLAN-200, and VLAN-300. Each VLAN is a member of one VRF. A physical interface (`tengigabitethernet 1/1`) attaches to the LAN and is configured as a tagged interface in VLAN-100, VLAN-200, and VLAN-300. The rest of this example is similar to the non-VLAN scenario.

This VLAN scenario often occurs in a service-provider network in which you configure VLAN tags for traffic from multiple customers on customer-premises equipment (CPE), and separate VRF instances associated with each VLAN are configured on the provider edge (PE) router in the point-of-presence (POP).

VRRP in VRF: Switch-1 VLAN Configuration

Switch-1

```
S1(conf)#ip vrf VRF-1 1
!
S1(conf)#ip vrf VRF-2 2
!
S1(conf)#ip vrf VRF-3 3
!
S1(conf)#interface TenGigabitEthernet 1/1
S1(conf-if-te-1/1)#no ip address
S1(conf-if-te-1/1)#switchport
S1(conf-if-te-1/1)#no shutdown
!
S1(conf-if-te-1/1)#interface vlan 100
S1(conf-if-vl-100)#ip vrf forwarding VRF-1
S1(conf-if-vl-100)#ip address 10.10.1.5/24
S1(conf-if-vl-100)#tagged TenGigabitethernet 1/1
S1(conf-if-vl-100)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S1(conf-if-vl-100-vrid-101)#priority 100
S1(conf-if-vl-100-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-vl-100)#no shutdown
!
S1(conf-if-te-1/1)#interface vlan 200
S1(conf-if-vl-200)#ip vrf forwarding VRF-2
S1(conf-if-vl-200)#ip address 10.10.1.6/24
S1(conf-if-vl-200)#tagged TenGigabitethernet 1/1
S1(conf-if-vl-200)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S1(conf-if-vl-200-vrid-101)#priority 100
S1(conf-if-vl-200-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-vl-200)#no shutdown
!
S1(conf-if-te-1/1)#interface vlan 300
S1(conf-if-vl-300)#ip vrf forwarding VRF-3
S1(conf-if-vl-300)#ip address 20.1.1.5/24
S1(conf-if-vl-300)#tagged TenGigabitethernet 1/1
S1(conf-if-vl-300)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S1(conf-if-vl-300-vrid-101)#priority 255
S1(conf-if-vl-300-vrid-101)#virtual-address 20.1.1.5
S1(conf-if-vl-300)#no shutdown

DellEMC#show vrrp vrf vrf1 vlan 400
------------------
Vlan 400, IPv4 VRID: 1, Version: 2, Net: 10.1.1.1
VRF: 1 vrf1
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 278, Gratuitous ARP sent: 1
Virtual MAC address:
00:00:5e:00:01:01
Virtual IP address:
10.1.1.100
Authentication: (none)

DellEMC#show vrrp vrf vrf2 port-channel 1
------------------
```

```
Port-channel 1, IPv4 VRID: 1, Version: 2, Net: 10.1.1.1
VRF: 2 vrf2
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 419, Gratuitous ARP sent: 1
Virtual MAC address:
00:00:5e:00:01:01
Virtual IP address:
10.1.1.100
Authentication: (none)
```

VRRP in VRF: Switch-2 VLAN Configuration

Switch-2

```
S2(conf)#ip vrf VRF-1 1
!
S2(conf)#ip vrf VRF-2 2
!
S2(conf)#ip vrf VRF-3 3
!
S2(conf)#interface TenGigabitEthernet 1/1
S2(conf-if-te-1/1)#no ip address
S2(conf-if-te-1/1)#switchport
S2(conf-if-te-1/1)#no shutdown
!
S2(conf-if-te-1/1)#interface vlan 100
S2(conf-if-vl-100)#ip vrf forwarding VRF-1
S2(conf-if-vl-100)#ip address 10.10.1.2/24
S2(conf-if-vl-100)#tagged TenGigabitethernet 1/1
S2(conf-if-vl-100)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S2(conf-if-vl-100-vrid-101)#priority 255
S2(conf-if-vl-100-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-vl-100)#no shutdown
!
S2(conf-if-te-1/1)#interface vlan 200
S2(conf-if-vl-200)#ip vrf forwarding VRF-2
S2(conf-if-vl-200)#ip address 10.10.1.2/24
S2(conf-if-vl-200)#tagged TenGigabitethernet 1/1
S2(conf-if-vl-200)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S2(conf-if-vl-200-vrid-101)#priority 255
S2(conf-if-vl-200-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-vl-200)#no shutdown
!
S2(conf-if-te-1/1)#interface vlan 300
S2(conf-if-vl-300)#ip vrf forwarding VRF-3
S2(conf-if-vl-300)#ip address 20.1.1.6/24
S2(conf-if-vl-300)#tagged TenGigabitethernet 1/1
S2(conf-if-vl-300)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S2(conf-if-vl-300-vrid-101)#priority 100
S2(conf-if-vl-300-vrid-101)#virtual-address 20.1.1.5
S2(conf-if-vl-300)#no shutdown

DellEMC#show vrrp vrf vrf1 vlan 400
------------------
Vlan 400, IPv4 VRID: 1, Version: 2, Net: 10.1.1.1
VRF: 1 vrf1
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 278, Gratuitous ARP sent: 1
Virtual MAC address:
00:00:5e:00:01:01
Virtual IP address:
10.1.1.100
Authentication: (none)

Vlan 400, IPv4 VRID: 10, Version: 2, Net: 20.1.1.2
VRF: 1 vrf1
State: Backup, Priority: 90, Master: 20.1.1.1
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 377, Bad pkts rcvd: 0, Adv sent: 0, Gratuitous ARP sent: 0
```

```
Virtual MAC address:
00:00:5e:00:01:0a
Virtual IP address:
20.1.1.100
Authentication: (none)

DellEMC#show vrrp vrf vrf2 port-channel 1
------------------
Port-channel 1, IPv4 VRID: 1, Version: 2, Net: 10.1.1.1
VRF: 2 vrf2
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 419, Gratuitous ARP sent: 1
Virtual MAC address:
00:00:5e:00:01:01
Virtual IP address:
10.1.1.100
Authentication: (none)
```

64

Debugging and Diagnostics

This chapter describes debugging and diagnostics for the device.

Topics:

- Offline Diagnostics
- Trace Logs
- Auto Save on Crash or Rollover
- Last Restart Reason
- Hardware Watchdog Timer
- Using the Show Hardware Commands
- Enabling Environmental Monitoring
- Buffer Tuning
- Troubleshooting Packet Loss
- Enabling Application Core Dumps
- Mini Core Dumps
- Enabling TCP Dumps

Offline Diagnostics

The offline diagnostics test suite is useful for isolating faults and debugging hardware.

The diagnostics tests are grouped into three levels:

- **Level 0** — Level 0 diagnostics check for the presence of various components and perform essential path verifications. In addition, Level 0 diagnostics verify the identification registers of the components on the board.
- **Level 1** — A smaller set of diagnostic tests. Level 1 diagnostics perform status/self-test for all the components on the board and test their registers for appropriate values. In addition, Level 1 diagnostics perform extensive tests on memory devices (for example, SDRAM, flash, NVRAM, EEPROM) wherever possible.
- **Level 2** — The full set of diagnostic tests. Level 2 diagnostics are used primarily for on-board loopback tests and more extensive component diagnostics. Various components on the board are put into Loopback mode and test packets are transmitted through those components. Level 2 diagnostics also perform snake tests using virtual local area network (VLAN) configurations.

Important Points to Remember

- You can only perform offline diagnostics on an offline standalone unit or offline member unit of a stack of three or more. You cannot perform diagnostics on the management or standby unit in a stack of two or more; if you do, a message similar to this displays: `Running Diagnostics on master/standby unit is not allowed on stack`.
- Diagnostics only test connectivity, not the entire data path.
- Diagnostic results are stored on the flash of the unit on which you performed the diagnostics.
- When offline diagnostics are complete, the unit or stack member reboots automatically.

Running Offline Diagnostics

To run offline diagnostics, use the following commands.

For more information, refer to the examples following the steps.

1. Place the unit in the offline state.

   EXEC Privilege mode

   `offline stack-unit` *`stack-unit-number`*

You cannot enter this command on a MASTER or Standby stack unit.

NOTE: The system reboots when the offline diagnostics complete. This is an automatic process. The following warning message appears when you implement the `offline stack-unit` command: `Warning - Diagnostic execution will cause stack-unit to reboot after completion of diags. Proceed with Offline-Diags [confirm yes/no]:y`

After the system goes offline, you must reload or run the **online stack-unit *stack-unit-number*** command for the normal operation.

2. Confirm the offline status.

   EXEC Privilege mode

   `show system brief`

3. Start diagnostics on the unit.

   `diag stack-unit *stack-unit-number*`

   When the tests are complete, the system displays the following message and automatically reboots the unit.

```
DellEMC#00:09:42 : Diagnostic test results are stored on file: flash:/TestReport-
SU-0.txt
Diags completed... Rebooting the system now!!!
Mar 12 10:40:35: %S6000:0 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on stack unit 1
```

```
DellEMC#00:09:42 : Diagnostic test results are stored on file: flash:/TestReport-
SU-0.txt
Diags completed... Rebooting the system now!!!

Mar 12 10:40:35: %S4820:0 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on stack unit 0
```

   Diagnostic results are printed to a file in the flash using the filename format *TestReport-SU-<stack-unit-id>.txt*.

   Log messages differ somewhat when diagnostics are done on a standalone unit and on a stack member.

4. View the results of the diagnostic tests.

   EXEC Privilege mode

   `show file flash://TestReport-SU-*stack-unit-id*.txt`

The following example shows the `offline stack-unit *stack-unit-number*` command.

```
DellEMC#offline stack-unit 0
Warning - offline of unit will bring down all the protocols and
the unit will be operationally down, except for running Diagnostics.
Please make sure that stacking is not configured for Diagnostics execution.
Also reboot/online command is necessary for normal operation after the offline command
is issued.
Proceed with Offline [confirm yes/no]:yes
DellEMC#Mar 12 10:32:18: %STKUNIT0-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down -
stack unit offline
```

The following example shows the `show system brief` command.

```
DellEMC#show system brief

Stack MAC : 00:01:e8:8b:5d:8c
Reload-Type                :   normal-reload [Next boot : normal-reload]

--  Stack Info  --
Unit  UnitType     Status         ReqTyp          CurTyp          Version     Ports
-------------------------------------------------------------------------------
  0   Management   online         S4820T          S4820T          9-4(0-89)   64
  1   Member       not present
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
  8   Member       not present
```

```
  9   Member       not present
 10   Member       not present
 11   Member       not present

--  Power Supplies  --
Unit   Bay   Status       Type    FanSpeed(rpm)
--------------------------------------------------------------------------
  0     0    down         UNKNOWN       0
  0     1    up           AC          14000

--  Fan  Status  --
Unit Bay   TrayStatus  Fan0    Speed   Fan1    Speed
------------------------------------------------------------------------------
 0    0     up          up      13466   up      13466
 0    1     up          up      13653   up      13466

Speed in RPM
```

The following example shows the `diag` command (standalone unit).

```
DellEMC#diag stack-unit 0 level0
Warning - diagnostic execution will cause multiple link flaps on the peer side -
advisable to shut directly connected ports
Proceed with Diags [confirm yes/no]: yes
DellEMC#Dec 15 04:14:07: %S4820:0 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on stack
unit 0
00:12:10 : System may take additional time for Driver Init.
00:12:10 : Approximate time to complete the Diags ... 6 Mins
```

The following example shows the `diag` command (stack member).

```
[output from master unit]

DellEMC#diag stack-unit 2
Warning - the stack unit will be pulled out of the stack for diagnostic execution
Proceed with Diags [confirm yes/no]: yes
Warning - diagnostic execution will cause multiple link flaps on the peer side -
advisable
to shut directly connected ports
Proceed with Diags [confirm yes/no]: yes
DellEMC#00:03:13: %S25P:2 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on stack unit 2
00:03:13 : Approximate time to complete these Diags ... 6 Min
00:03:13 : Diagnostic test results will be stored on stack unit 2 file: flash:/
TestReport-SU-2.txt

DellEMC#00:03:35: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card
removed
00:08:50: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
00:09:00: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 2 (type S25P, 28
ports)
00:09:00: %S25P:2 %CHMGR-0-PS_UP: Power supply 0 in unit 2 is up
00:09:00: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 2 is up
[output from the console of the unit in which diagnostics are performed]

DellEMC(stack-member-2)#
Diagnostic test results are stored on file: flash:/TestReport-SU-2.txt
Diags completed... Rebooting the system now!!!
```

The following example shows the `show file flash:\\` command (standalone member).

```
DellEMC#show file flash://TestReport-SU-0.txt
diagS4810UsbAOpen[735]: ERROR: Usb device open failed

*****************************S-Series Diagnostics**********************

    Board                          : S4820T Dell Inc.
    Stack Unit Board Serial Number  : SWDG129100003
    CPU Version                    : MPC8536E, Version: 1.1
    CPLD REVISION                  : 0x9
    CPLD BOARD TYPE                : 0x9
    Diag image based on build      : 9-4(0-89)
```

```
    Stack Unit Board temperature    : 46 Degree C
    Stack Unit Number               : 0

************************Stack Unit EEPROM INFO**************************

**************MFG INFO*******************

Data in Chassis Eeprom Mfg Info is listed as...
Vendor Id             : 07
Country Code          : 06
Date Code             : 10222012
Serial Number         : SWDG129100003
Part Number           : 2P7Y5
Product Revision      : A
Product Order Number : 02P7Y52BE0005A00123SBC902

*************************** S4820T LEVEL 0 DIAGNOSTICS*************************

diagS4810DumpPowerGoodStatus[653]: ERROR: Psu : 0 Output voltage is NOT in regulation
range
Test 1.000 - Psu Power Good Test ..................................     FAIL
Test 1.001 - Psu Power Good Test ..................................  PASS
Test 1 - Psu Power Good Test  .....................................     FAIL
diagS4820DumpPsuStatus[1753]: ERROR: Psu0: Reporting fault in Current, Voltage and Fan
condition
diagS4820DumpPsuStatus[1757]: ERROR: Psu0: Output voltage is NOT in regulation range
Test 2.000 - Fan Psu Status test ..................................     FAIL
Test 2.001 - Fan Psu Status test ..................................  PASS
Test 2 - Fan Psu Status Test  .....................................     FAIL
Test 3.000 - Fan board presence Test ..............................  PASS
Test 3.001 - Fan board presence Test ..............................  PASS
Test 3 - Fan Board Presence Test  .................................  PASS
Test 4.000 - Board Fan Status Test ................................  PASS
Test 4.001 - Board Fan Status Test ................................  PASS
Test 4.002 - Board Fan Status Test ................................  PASS
Test 4.003 - Board Fan Status Test ................................  PASS
Test 4 - Board Fan Status Test  ...................................  PASS
Test 5 - Fan Tray Air Flow Status Test  ...........................  PASS
Test 6.000 - TMP75A-CPU ...........................................  PASS
Test 6.001 - TMP75B-MAC ...........................................  PASS
Test 6.002 - TMP75C-LEFT-PHY ......................................  PASS
Test 6.003 - TMP75D-RIGHT-PHY .....................................  PASS
Test 6 - Temp Sensor Access Test ..................................  PASS
Test 7 - RTC Presence Test ........................................  PASS
Test 8 - SDRAM Presence Test ......................................  PASS
```

Trace Logs

In addition to the syslog buffer, Dell EMC Networking OS buffers trace messages which are continuously written by various Dell EMC Networking OS software tasks to report hardware and software events and status information.

Each trace message provides the date, time, and name of the Dell EMC Networking OS process. All messages are stored in a ring buffer. You can save the messages to a file either manually or automatically after failover.

Auto Save on Crash or Rollover

Exception information for MASTER or standby units is stored in the *flash:/TRACE_LOG_DIR* directory. This directory contains files that save trace information when there has been a task crash or timeout.

- On a MASTER unit, you can reach the *TRACE_LOG_DIR* files by FTP or by using the `show file` command from the *flash://TRACE_LOG_DIR* directory.
- On a Standby unit, you can reach the *TRACE_LOG_DIR* files only by using the `show file` command from the *flash://TRACE_LOG_DIR* directory.

NOTE: Non-management member units do not support this functionality.

Last Restart Reason

If the system restarts for some reason (automatically or manually), the `show system` command output includes the reason for the restart.

The following table shows the reasons displayed in the output and their corresponding causes.

Table 133. Line Card Restart Causes and Reasons

Causes	Displayed Reasons
Remote power cycle of the chassis	push button reset
reload	soft reset
reboot after a crash	soft reset

Hardware Watchdog Timer

The `hardware watchdog` command automatically reboots an Dell EMC Networking OS switch/router with a single RPM that is unresponsive.

This is a last resort mechanism intended to prevent a manual power cycle.

Using the Show Hardware Commands

The `show hardware` command tree consists of commands used with the system. These commands display information from a hardware sub-component and from hardware-based feature tables.

NOTE: Use the `show hardware` commands only under the guidance of the Dell Technical Assistance Center.

The following lists the `show hardware` commands available as of the latest Dell EMC Networking OS version.

- View internal interface status of the stack-unit CPU port which connects to the external management interface.
  EXEC Privilege mode
  `show hardware stack-unit {0-11} cpu management statistics`
- View driver-level statistics for the data-plane port on the CPU for the specified stack-unit.
  EXEC Privilege mode
  `show hardware stack-unit {0-11} cpu data-plane statistics`
  This view provides insight into the packet types entering the CPU to see whether CPU-bound traffic is internal (IPC traffic) or network control traffic, which the CPU must process.
- View the modular packet buffers details per stack unit and the mode of allocation.
  EXEC Privilege mode
  `show hardware stack-unit {0-11} buffer total-buffer`
- View the modular packet buffers details per unit and the mode of allocation.
  EXEC Privilege mode
  `show hardware stack-unit {0-11} buffer unit {0-1} total-buffer`
- View the forwarding plane statistics containing the packet buffer usage per port per stack unit.
  EXEC Privilege mode
  `show hardware stack-unit {0-11} buffer unit {0-1} port {1-64 | all} buffer-info`
- View the forwarding plane statistics containing the packet buffer statistics per COS per port.
  EXEC Privilege mode
  `show hardware stack-unit {0-11} buffer unit {0-1} port {1-64} queue {0-14 | all} buffer-info`
- View input and output statistics on the party bus, which carries inter-process communication traffic between CPUs.
  EXEC Privilege mode

`show hardware stack-unit {0-11} cpu party-bus statistics`

- View the ingress and egress internal packet-drop counters, MAC counters drop, and FP packet drops for the stack unit on per port basis.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} drops unit {0-1} port {1-64}`

  This view helps identifying the stack unit/port pipe/port that may experience internal drops.
- View the input and output statistics for a stack-port interface.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} stack-port {portnumber}`
- View the counters in the field processors of the stack unit.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} counters`
- View the details of the FP Devices and Hi gig ports on the stack-unit.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} details`
- Execute a specified `bShell` command from the CLI without going into the bShell.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} execute-shell-cmd {command}`
- View the Multicast IPMC replication table from the bShell.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} ipmc-replication`
- View the internal statistics for each port-pipe (unit) on per port basis.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} port-stats [detail]`
- View the stack-unit internal registers for each port-pipe.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} register`
- View the tables from the bShell through the CLI without going into the bShell.

  EXEC Privilege mode

  `show hardware stack-unit {0-11} unit {0-1} table-dump {table name}`

Enabling Environmental Monitoring

The device components use environmental monitoring hardware to detect transmit power readings, receive power readings, and temperature updates.

To receive periodic power updates, you must enable the following command.

- Enable environmental monitoring.

  `enable optic-info-update interval`

```
DellEMC#show interfaces fortyGigE 1/52 transceiver

QSFP 52 Serial ID Base Fields
QSFP 52 Id                      = 0x0d
QSFP 52 Ext Id                  = 0x00
QSFP 52 Connector               = 0x0c
QSFP 52 Transceiver Code        = 0x04 0x00 0x00 0x00 0x00 0x00 0x00 0x00
QSFP 52 Encoding                = 0x05
QSFP 52 Length(SFM)     Km      = 0x00
QSFP 52 Length(OM3)     2m      = 0x32
QSFP 52 Length(OM2)     1m      = 0x00
QSFP 52 Length(OM1)     1m      = 0x00
QSFP 52 Length(Copper)  1m      = 0x00
QSFP 52 Vendor Rev              = 01
QSFP 52 Laser Wavelength        = 850.00 nm
QSFP 52 CheckCodeBase           = 0x26
```

```
QSFP 52 Serial ID Extended Fields
QSFP 52 BR max                    = 0
QSFP 52 BR min                    = 0
QSFP 52 Vendor SN                 = QC050955
QSFP 52 Datecode                  = 120205
QSFP 52 CheckCodeExt              = 0x2b

QSFP 52 Diagnostic Information
================================
QSFP 52 Rx Power measurement type         = Average
================================
QSFP 52 Temp High Alarm threshold         = 75.000C
QSFP 52 Voltage High Alarm threshold      = 3.630V
QSFP 52 Bias High Alarm threshold         = 10.000mA
QSFP 52 RX Power High Alarm threshold     = 2.188mW
QSFP 52 Temp Low Alarm threshold          = -5.000C
QSFP 52 Voltage Low Alarm threshold       = 2.970V
QSFP 52 Bias Low Alarm threshold          = 0.500mA
QSFP 52 RX Power Low Alarm threshold      = 0.045mW
================================
QSFP 52 Temp High Warning threshold       = 70.000C
QSFP 52 Voltage High Warning threshold    = 3.465V
QSFP 52 Bias High Warning threshold       = 9.500mA
QSFP 52 RX Power High Warning threshold   = 1.738mW
QSFP 52 Temp Low Warning threshold        = 0.000C
QSFP 52 Voltage Low Warning threshold     = 3.135V
QSFP 52 Bias Low Warning threshold        = 1.000mA
QSFP 52 RX Power Low Warning threshold    = 0.112mW
================================
QSFP 52 Temperature                       = 30.602C
QSFP 52 Voltage                           = 3.311V
QSFP 52 TX1 Bias Current                  = 0.000mA
QSFP 52 TX2 Bias Current                  = 0.000mA
QSFP 52 TX3 Bias Current                  = 0.000mA
QSFP 52 TX4 Bias Current                  = 0.000mA
QSFP 52 RX1 Power                         = 0.000mW
QSFP 52 RX2 Power                         = 0.000mW
QSFP 52 RX3 Power                         = 0.000mW
QSFP 52 RX4 Power                         = 0.000mW
```

Recognize an Overtemperature Condition

An overtemperature condition occurs, for one of two reasons: the card genuinely is too hot or a sensor has malfunctioned.

Inspect cards adjacent to the one reporting the condition to discover the cause.

- If directly adjacent cards are not normal temperature, suspect a genuine overheating condition.
- If directly adjacent cards are normal temperature, suspect a faulty sensor.

When the system detects a genuine over-temperature condition, it powers off the card. To recognize this condition, look for the following system messages:

```
CHMGR-2-MAJOR_TEMP: Major alarm: chassis temperature high (temperature reaches or
exceeds threshold of
[value]C)
CHMGR-2-TEMP_SHUTDOWN_WARN: WARNING! temperature is [value]C; approaching shutdown
threshold of [value]C
```

To view the programmed alarm thresholds levels, including the shutdown value, use the `show alarms threshold` command.

Example of the `show alarms threshold` Command

```
DellEMC#show alarms threshold

--  Temperature Limits (deg C)  --
-----------------------------------------------------------------
          Minor Off     Minor   Major Off    Major   Shutdown
Unit2        55         60      75           80      85
-----------------------------------------------------------------
```

```
           Minor Off     Minor   Major Off    Major   Shutdown
Unit3         55         60      75           80      85
```

Troubleshoot an Over-temperature Condition

To troubleshoot an over-temperature condition, use the following information.

1. Use the `show environment` commands to monitor the temperature levels.
2. Check air flow through the system. Ensure that the air ducts are clean and that all fans are working correctly.
3. After the software has determined that the temperature levels are within normal limits, you can re-power the card safely. To bring back the line card online, use the `power-on` command in EXEC mode.

In addition, to control airflow for adequate system cooling, Dell EMC Networking requires that you install blanks in all slots without a line card.

NOTE: Exercise care when removing a card; if it has exceeded the major or shutdown thresholds, the card could be hot to the touch.

Recognize an Under-Voltage Condition

If the system detects an under-voltage condition, it sends an alarm.

To recognize this condition, look for the following system message: `%CHMGR-1-CARD_SHUTDOWN: Major alarm: stack unit 2 down - auto-shutdown due to under voltage`.

This message indicates that the specified card is not receiving enough power. In response, the system first shuts down Power over Ethernet (PoE). If the under-voltage condition persists, line cards are shut down, then the RPMs.

Troubleshoot an Under-Voltage Condition

To troubleshoot an under-voltage condition, check that the correct number of power supplies are installed and their Status light emitting diodes (LEDs) are lit.

The following table lists information for SNMP traps and OIDs on the environmental monitoring hardware and hardware components.

Table 134. SNMP Traps and OIDs

OID String	OID Name	Description
Receiving Power		
.1.3.6.1.4.1.6027.3.10.1.2.5.1.6	chSysPortXfpRecvPower	OID displays the receiving power of the connected optics.
Transmitting power		
.1.3.6.1.4.1.6027.3.10.1.2.5.1.8	chSysPortXfpTxPower	OID displays the transmitting power of the connected optics.
Temperature		
.1.3.6.1.4.1.6027.3.10.1.2.5.1.7	chSysPortXfpRecvTemp	OID displays the temperature of the connected optics. **NOTE:** These OIDs only generate if you enable the `enable optic-info-update-interval is enabled` command.
Hardware MIB Buffer Statistics		
.1.3.6.1.4.1.6027.3.27.1.4	dellNetFpPacketBufferTable	View the modular packet buffers details per stack unit and the mode of allocation.

Table 134. SNMP Traps and OIDs (continued)

OID String	OID Name	Description
.1.3.6.1.4.1.6027.3.27.1.5	dellNetFpStatsPerPortTable	View the forwarding plane statistics containing the packet buffer usage per port per stack unit.
.1.3.6.1.4.1.6027.3.27.1.6	dellNetFpStatsPerCOSTable	View the forwarding plane statistics containing the packet buffer statistics per COS per port.

Buffer Tuning

Buffer Tuning allows you to modify the way your switch allocates buffers from its available memory and helps prevent packet drops during a temporary burst of traffic.

Using a PreDefined Buffer Profile

Dell EMC Networking OS provides two predefined buffer profiles, one for single-queue (for example, non-quality-of-service [QoS]) applications, and one for four-queue (for example, QoS) applications.

You must reload the system for the global buffer profile to take effect, a message similar to the following displays: `% Info: For the global pre-defined buffer profile to take effect, please save the config and reload the system..`

Dell EMC Networking OS Behavior: After you configure buffer-profile global 1Q, the message displays during every bootup. Only one reboot is required for the configuration to take effect; afterward you may ignore this bootup message.

Dell EMC Networking OS Behavior: If you configure 1Q, save the running-config to the `startup-config`, and then delete the `startup-config` and reload the chassis. The only way to return to the default buffer profile is to remove the 1Q profile configured and then reload the chassis.

If you have already applied a custom buffer profile on an interface, the `buffer-profile global` command fails and a message similar to the following displays: `% Error: User-defined buffer profile already applied. Failed to apply global pre-defined buffer profile. Please remove all user-defined buffer profiles.`

Similarly, when you configure `buffer-profile global`, you cannot not apply a buffer profile on any single interface. A message similar to the following displays: `% Error: Global pre-defined buffer profile already applied. Failed to apply user-defined buffer profile on interface Te 1/1. Please remove global pre-defined buffer profile.`

To apply a predefined buffer profile, use the following command:

- Apply one of the predefined buffer profiles for all port pipes in the system.

  CONFIGURATION mode

  `buffer-profile global [1Q|4Q]`

If the default buffer profile `dynamic` is active, Dell EMC Networking OS displays an error message instructing you to remove the default configuration using the `no buffer-profile global` command.

Troubleshooting Packet Loss

The `show hardware stack-unit` command is intended primarily to troubleshoot packet loss.

To troubleshoot packet loss, use the following commands.

- `show hardware stack-unit` *`stack-unit-number`* `cpu data-plane statistics`
- `show hardware stack-unit` *`stack-unit-number`* `cpu party-bus statistics`
- `show hardware stack-unit` *`stack-unit-number`* `drops unit` *`unit-number`*
- `show hardware stack-unit` *`stack-unit-number`* `unit` *`unit-number`* `{counters | details | port-stats [detail] | register | ipmc-replication | table-dump}`
- `show hardware {ip | ipv6 | mac} {eg-acl | in-acl} stack-unit` *`stack-unit-number`* `port-set 0 pipeline 0-3`
- `show hardware ip qos stack-unit` *`stack-unit-number`* `port-set` *`0`*

- `show hardware system-flow layer2 stack-unit stack-unit-number port-set 0 {counters | pipeline 0-3}`
- `show hardware drops interface interface`
- `show hardware buffer-stats-snapshot resource interface interface`
- `show hardware buffer inteface interface{priority-group { id | all } | queue { id| all} } buffer-info`
- `show hardware buffer-stats-snapshot resource interface interface{priority-group { id | all } | queue { ucast{id | all}{ mcast {id | all} | all}`
- `show hardware drops interface interface`
- `clear hardware stack-unit stack-unit-number counters`
- `clear hardware stack-unit stack-unit-number unit 0-1 counters`
- `clear hardware stack-unit stack-unit-number cpu data-plane statistics`
- `clear hardware stack-unit stack-unit-number cpu party-bus statistics`

Displaying Drop Counters

To display drop counters, use the following commands.

- Identify which stack unit and port pipe is experiencing internal drops.

  `show hardware stack-unit stack-unit-number drops [unit unit-number]`
- Identify which interface is experiencing internal drops.

  `show hardware drops interface interface`

```
DellEMC#show hardware stack-unit 0 drops
UNIT No: 0
Total Ingress Drops :0
Total IngMac Drops :0
Total Mmu Drops :0
Total EgMac Drops :0
Total Egress Drops :0
UNIT No: 1
Total Ingress Drops :0
Total IngMac Drops :0
Total Mmu Drops :0
Total EgMac Drops :0
Total Egress Drops :0

DellEMC#show hardware stack-unit 0 drops unit 0
Port# :Ingress Drops :IngMac Drops :Total Mmu Drops :EgMac Drops :Egress
Drops
1 0 0 0 0 0
2 0 0 0 0 0
3 0 0 0 0 0
4 0 0 0 0 0
5 0 0 0 0 0
6 0 0 0 0 0
7 0 0 0 0 0
8 0 0 0 0 0
```

Example of show hardware drops interface *interface*

```
DellEMC#show hardware drops interface tengigabitethernet 2/1

Drops in Interface Te 2/1:
 --- Ingress Drops      ---
Ingress Drops                 : 0
IBP CBP Full Drops            : 0
PortSTPnotFwd Drops           : 0
IPv4 L3 Discards              : 0
Policy Discards               : 0
Packets dropped by FP         : 0
(L2+L3) Drops                 : 0
Port bitmap zero Drops        : 0
Rx VLAN Drops                 : 0
 --- Ingress MAC counters---
Ingress FCSDrops              : 0
```

```
Ingress MTUExceeds              : 0
 --- MMU Drops           ---
Ingress MMU Drops               : 0
HOL DROPS(TOTAL)                : 0
HOL DROPS on COS0               : 0
HOL DROPS on COS1               : 0
HOL DROPS on COS2               : 0
HOL DROPS on COS3               : 0
HOL DROPS on COS4               : 0
HOL DROPS on COS5               : 0
HOL DROPS on COS6               : 0
HOL DROPS on COS7               : 0
HOL DROPS on COS8               : 0
HOL DROPS on COS9               : 0
HOL DROPS on COS10              : 0
HOL DROPS on COS11              : 0
HOL DROPS on COS12              : 0
HOL DROPS on COS13              : 0
HOL DROPS on COS14              : 0
HOL DROPS on COS15              : 0
HOL DROPS on COS16              : 0
HOL DROPS on COS17              : 0
TxPurge CellErr                 : 0
Aged Drops                      : 0
 --- Egress MAC counters---
Egress FCS Drops                : 0
 --- Egress FORWARD PROCESSOR Drops   ---
IPv4 L3UC Aged & Drops          : 0
TTL Threshold Drops             : 0
INVALID VLAN CNTR Drops         : 0
L2MC Drops                      : 0
PKT Drops of ANY Conditions     : 0
Hg MacUnderflow                 : 0
TX Err PKT Counter              : 0
 --- Error counters---
Internal Mac Transmit Errors    : 0
Unknown Opcodes                 : 0
Internal Mac Receive Errors     : 0
```

Dataplane Statistics

The `show hardware stack-unit cpu data-plane statistics` command provides insight into the packet types coming to the CPU. The `show hardware stack-unit cpu party-bus statistics` command displays input and output statistics on the party bus, which carries inter-process communication traffic between CPUs.

The command output in the following example has been augmented, providing detailed RX/ TX packet statistics on a per-queue basis. The objective is to see whether CPU-bound traffic is internal (so-called *party bus* or IPC traffic) or network control traffic, which the CPU must process.

Example of Viewing Dataplane Statistics

```
DellEMC#show hardware stack-unit 1 cpu data-plane statistics

bc pci driver statistics for device:
 rxHandle         :773
 noMhdr           :0
 noMbuf           :0
 noClus           :0
 recvd            :773
 dropped          :0
 recvToNet        :773
 rxError          :0
 rxFwdError       :0
 rxDatapathErr    :0
 rxPkt(COS0 )     :0
 rxPkt(COS1 )     :0
 rxPkt(COS2 )     :0
 rxPkt(COS3 )     :0
 rxPkt(COS4 )     :0
 rxPkt(COS5 )     :0
```

```
 rxPkt(COS6 )     :0
 rxPkt(COS7 )     :0
 rxPkt(COS8 )     :773
 rxPkt(COS9 )     :0
 rxPkt(COS10)    :0
 rxPkt(COS11)    :0
 rxPkt(UNIT0)     :773
 transmitted      :12698
 txRequested      :12698
 noTxDesc         :0
 txError          :0
 txReqTooLarge    :0
 txInternalError :0
 txDatapathErr    :0
 txPkt(COS0 )     :0
 txPkt(COS1 )     :0
 txPkt(COS2 )     :0
 txPkt(COS3 )     :0
 txPkt(COS4 )     :0
 txPkt(COS5 )     :0
 txPkt(COS6 )     :0
 txPkt(COS7 )     :0
 txPkt(COS8 )     :0
 txPkt(COS9 )     :0
 txPkt(COS10)    :0
 txPkt(COS11)    :0
 txPkt(UNIT0)     :0
```

Example of Viewing Party Bus Statistics

```
DellEMC#sh hardware stack-unit 1 cpu party-bus statistics
Input Statistics:
  27550 packets, 2559298 bytes
  0 dropped, 0 errors
Output Statistics:
  1649566 packets, 1935316203 bytes
  0 errors
```

Display Stack Port Statistics

The `show hardware stack-unit stack-port` command displays input and output statistics for a stack-port interface.

Example of Viewing Stack Unit Statistics

```
DellEMC#show hardware stack-unit 2 stack-port 49
Input Statistics:
  27629 packets, 3411731 bytes
  0 64-byte pkts, 27271 over 64-byte pkts, 207 over 127-byte pkts
  17 over 255-byte pkts, 56 over 511-byte pkts, 78 over 1023-byte pkts
  0 Multicasts, 5 Broadcasts
  0 runts, 0 giants, 0 throttles
  0 CRC, 0 overrun, 0 discarded
Output Statistics:
  1649714 packets, 1948622676 bytes, 0 underruns
  0 64-byte pkts, 27234 over 64-byte pkts, 107970 over 127-byte pkts
  34 over 255-byte pkts, 504838 over 511-byte pkts, 1009638 over 1023-byte pkts
  0 Multicasts, 0 Broadcasts, 1649714 Unicasts
  0 throttles, 0 discarded, 0 collisions
Rate info (interval 45 seconds):
  Input 00.00 Mbits/sec,     2 packets/sec, 0.00% of line-rate
  Output 00.06 Mbits/sec,    8 packets/sec, 0.00% of line-rate
DellEMC#
```

Display Stack Member Counters

You can use the `show hardware` command to display internal receive and transmit statistics, based on the selected command option.

The following example is a sample of the output for the `counters` option.

Example of Displaying Counter Values for all Interface in the Selected Stack-Member and Port-Pipe

```
DellEMC# show hardware stack-unit 1 unit 0 counters

Interface Te 0/0 :
Description                                     Value
RX - IPV4 L3 Unicast Frame Counter                0
RX - IPV4 L3 routed multicast Packets             0
RX - IPV6 L3 Unicast Frame Counter                0
RX - IPV6 L3 routed multicast Packets             0
RX - Unicast Packet Counter                       0
RX - 64 Byte Frame Counter                        0
RX - 64 to 127 Byte Frame Counter                 0
RX - 128 to 255 Byte Frame Counter                0
RX - 256 to 511 Byte Frame Counter                0
RX - 512 to 1023 Byte Frame Counter               0
RX - 1024 to 1518 Byte Frame Counter              0
RX - 1519 to 1522 Byte Good VLAN Frame Counter    0
RX - 1519 to 2047 Byte Frame Counter              0
RX - 2048 to 4095 Byte Frame Counter              0
RX - 4096 to 9216 Byte Frame Counter              0
RX - Good Packet Counter                          0
RX - Packet/frame Counter                         0
RX - Unicast Packet Counter                       0
RX - Multicast Packet Counter                     0
RX - Broadcast Frame Counter                      0
RX - Byte Counter                                 0
RX - Control frame counter                        0
RX - PAUSE frame counter                          0
RX - Oversized frame counter                      0
RX - Jabber frame counter                         0
RX - VLAN tag frame counter                       0
RX - Double VLAN tag frame counter                0
RX - RUNT frame counter                           0
RX - Fragment counter                             0
RX - VLAN tagged packets                          0
TX - 64 Byte Frame Counter                        46
TX - 64 to 127 Byte Frame Counter                 0
TX - 128 to 255 Byte Frame Counter                0
TX - 256 to 511 Byte Frame Counter                0
TX - 512 to 1023 Byte Frame Counter               0
TX - 1024 to 1518 Byte Frame Counter              0
TX - 1519 to 1522 Byte Good VLAN Frame Counter    0
TX - 1519 to 2047 Byte Frame Counter              0
TX - 2048 to 4095 Byte Frame Counter              0
TX - 4096 to 9216 Byte Frame Counter              0
TX - Good Packet Counter                          46
TX - Packet/frame Counter                         46
TX - Unicast Packet Counter                       0
TX - Multicast Packet Counter                     46
TX - Broadcast Frame Counter                      0
TX - Byte Counter                                 2944
TX - Control frame counter                        0
TX - Pause control frame counter                  0
TX - Over size packet counter                     0
TX - Jabber counter                               0
TX - VLAN tag frame counter                       0
TX - Double VLAN tag frame counter                0
TX - RUNT frame counter                           0
TX - Fragment counter                             0

Interface Te 0/1 :
Description                                     Value
RX - IPV4 L3 Unicast Frame Counter                0
RX - IPV4 L3 routed multicast Packets             0
```

```
RX - IPV6 L3 Unicast Frame Counter                  0
-----
-----
-----
-----
-----
Interface Fo 0/60 :
Description                                       Value
RX - IPV4 L3 Unicast Frame Counter                  0
RX - IPV4 L3 routed multicast Packets               0
RX - IPV6 L3 Unicast Frame Counter                  0
RX - IPV6 L3 routed multicast Packets               0
RX - Unicast Packet Counter                         0
RX - 64 Byte Frame Counter                          0
RX - 64 to 127 Byte Frame Counter                   0
RX - 128 to 255 Byte Frame Counter                  0
RX - 256 to 511 Byte Frame Counter                  0
RX - 512 to 1023 Byte Frame Counter                 0
RX - 1024 to 1518 Byte Frame Counter                0
RX - 1519 to 1522 Byte Good VLAN Frame Counter      0
RX - 1519 to 2047 Byte Frame Counter                0
RX - 2048 to 4095 Byte Frame Counter                0
RX - 4096 to 9216 Byte Frame Counter                0
RX - Good Packet Counter                            0
RX - Packet/frame Counter                           0
RX - Unicast Packet Counter                         0
RX - Multicast Packet Counter                       0
RX - Broadcast Frame Counter                        0
RX - Byte Counter                                   0
RX - Control frame counter                          0
RX - PAUSE frame counter                            0
RX - Oversized frame counter                        0
RX - Jabber frame counter                           0
RX - VLAN tag frame counter                         0
RX - Double VLAN tag frame counter                  0
RX - RUNT frame counter                             0
RX - Fragment counter                               0
RX - VLAN tagged packets                            0
TX - 64 Byte Frame Counter                          0
TX - 64 to 127 Byte Frame Counter                   0
TX - 128 to 255 Byte Frame Counter                  0
TX - 256 to 511 Byte Frame Counter                  0
TX - 512 to 1023 Byte Frame Counter                 0
TX - 1024 to 1518 Byte Frame Counter                0
TX - 1519 to 1522 Byte Good VLAN Frame Counter      0
TX - 1519 to 2047 Byte Frame Counter                0
TX - 2048 to 4095 Byte Frame Counter                0
TX - 4096 to 9216 Byte Frame Counter                0
TX - Good Packet Counter                            0
TX - Packet/frame Counter                           0
TX - Unicast Packet Counter                         0
TX - Multicast Packet Counter                       0
TX - Broadcast Frame Counter                        0
TX - Byte Counter                                   0
TX - Control frame counter                          0
TX - Pause control frame counter                    0
TX - Over size packet counter                       0
TX - Jabber counter                                 0
TX - VLAN tag frame counter                         0
TX - Double VLAN tag frame counter                  0
TX - RUNT frame counter                             0
TX - Fragment counter                               0
```

Example of Displaying Counter Information for a Specific Interface

```
DellEMC#show hardware counters interfac tengigabitethernet 5/1
unit: 0 port: 2 (interface Te 5/1)
Description                                       Value

RX - IPV4 L3 Unicast Frame Counter                  0
RX - IPV4 L3 Routed Multicast Packets               0
RX - IPV6 L3 Unicast Frame Counter                  0
RX - IPV6 L3 Routed Multicast Packets               0
```

```
RX - Unicast Packet Counter                          0
RX - 64 Byte Frame Counter                           0
RX - 65 to 127 Byte Frame Counter                    0
RX - 128 to 255 Byte Frame Counter                   0
RX - 256 to 511 Byte Frame Counter                   0
RX - 512 to 1023 Byte Frame Counter                  0
RX - 1024 to 1518 Byte Frame Counter                 0
RX - 1519 to 1522 Byte Good VLAN Frame Counter       0
RX - 1519 to 2047 Byte Frame Counter                 0
RX - 2048 to 4095 Byte Frame Counter                 0
RX - 4096 to 9216 Byte Frame Counter                 0
RX - Good Packet Counter                             0
RX - Packet/Frame Counter                            0
RX - Unicast Frame Counter                           0
RX - Multicast Frame Counter                         0
RX - Broadcast Frame Counter                         0
RX - Byte Counter                                    0
RX - Control Frame Counter                           0
RX - Pause Control Frame Counter                     0
RX - Oversized Frame Counter                         0
RX - Jabber Frame Counter                            0
RX - VLAN Tag Frame Counter                          0
RX - Double VLAN Tag Frame Counter                   0
RX - RUNT Frame Counter                              0
RX - Fragment Counter                                0
RX - VLAN Tagged Packets                             0
RX - Ingress Dropped Packet                          0
RX - MTU Check Error Frame Counter                   0
RX - PFC Frame Priority 0                            0
RX - PFC Frame Priority 1                            0
RX - PFC Frame Priority 2                            0
RX - PFC Frame Priority 3                            0
RX - PFC Frame Priority 4                            0
RX - PFC Frame Priority 5                            0
RX - PFC Frame Priority 6                            0
RX - PFC Frame Priority 7                            0
RX - Debug Counter 0                                 0
RX - Debug Counter 1                                 0
RX - Debug Counter 2                                 0
<output truncated for brevity>
```

Enabling Application Core Dumps

Application core dumps are disabled by default.

A core dump file can be very large. Due to memory requirements the file can only be sent directly to an FTP server; it is not stored on the local flash.

To enable full kernel core dumps, use the following command.

- Enable stack unit kernel full core dumps.

  CONFIGURATION mode

  ```
  logging coredump server
  ```

To undo this command, use the `no logging coredump server` command.

Mini Core Dumps

Dell EMC Networking OS supports mini core dumps on the application and kernel crashes. The mini core dump applies to Master, Standby, and Member units.

Application and kernel mini core dumps are always enabled. The mini core dumps contain the stack space and some other minimal information that you can use to debug a crash. These files are small files and are written into flash until space is exhausted. When the flash is full, the write process is stopped.

A mini core dump contains critical information in the event of a crash. Mini core dump files are located in *flash:/ (root dir)*. The application mini core filename format is *f10StkUnit<Stack_unit_no>.<Application name>.acore.mini.txt*. The kernel mini core filename format is *f10StkUnit<Stack_unit_no>.kcore.mini.txt*. The following are sample filenames.

When a member or standby unit crashes, the mini core file gets uploaded to master unit. When the master unit crashes, the mini core file is uploaded to new master.

The panic string contains key information regarding the crash. Several panic string types exist, and they are displayed in regular English text to enable easier understanding of the crash cause.

Example of Application Mini Core Dump Listings

```
DellEMC#dir
Directory of flash:

1  drw-  16384 Jan 01 1980 00:00:00 +00:00 .
2  drwx  1536  Sep 03 2009 16:51:02 +00:00 ..
3  drw-  512   Aug 07 2009 13:05:58 +00:00 TRACE_LOG_DIR
4  d---  512   Aug 07 2009 13:06:00 +00:00 ADMIN_DIR
5  -rw-  8693  Sep 03 2009 16:50:56 +00:00 startup-config
6  -rw-  8693  Sep 03 2009 16:44:22 +00:00 startup-config.bak
7 -rw- 156  Aug 28 2009 16:16:10 +00:00 f10StkUnit0.mrtm.acore.mini.txt
8 -rw- 156  Aug 28 2009 17:17:24 +00:00 f10StkUnit0.vrrp.acore.mini.txt
9 -rw- 156  Aug 28 2009 18:25:18 +00:00 f10StkUnit0.sysd.acore.mini.txt
10 -rw- 156  Aug 28 2009 19:07:36 +00:00 f10StkUnit0.frrp.acore.mini.txt
11 -rw- 156  Aug 31 2009 16:18:50 +00:00 f10StkUnit2.sysd.acore.mini.txt
12 -rw- 156  Aug 29 2009 14:28:34 +00:00 f10StkUnit0.ipm1.acore.mini.txt
13 -rw- 156  Aug 31 2009 16:14:56 +00:00 f10StkUnit0.acl.acore.mini.txt

flash: 3104256 bytes total (2959872 bytes free)
DellEMC#
```

Example of a Mini Core Text File

```
VALID MAGIC
-----------------PANIC STRING -----------------
panic string is :<null>
---------------STACK TRACE START---------------
0035d60c <f10_save_mmu+0x120>:
00274f8c <panic+0x144>:
0024e2b0 <db_fncall+0x134>:
0024dee8 <db_command+0x258>:
0024d9c4 <db_command_loop+0xc4>:
002522b0 <db_trap+0x158>:
0026a8d0 <mi_switch+0x1b0>:
0026a00c <bpendtsleep>:
----------------STACK TRACE END----------------

--------------------FREE MEMORY---------------
uvmexp.free = 0x2312
```

Enabling TCP Dumps

A TCP dump captures CPU-bound control plane traffic to improve troubleshooting and system manageability. When you enable TCP dump, it captures all the packets on the local CPU, as specified in the CLI.

You can save the traffic capture files to flash, FTP, SCP, or TFTP. The files saved on the flash are located in the *flash://TCP_DUMP_DIR/Tcpdump_<time_stamp_dir>/* directory and labeled *tcpdump_*.pcap*. There can be up to 20 *Tcpdump_<time_stamp_dir>* directories. The 21st file overwrites the oldest saved file. The maximum file size for a TCP dump capture is 1MB. When a file reaches 1MB, a new file is created, up to the specified total number of files.

Maximize the number of packets recorded in a file by specifying the snap-length to capture the file headers only.

The `tcpdump` command has a finite run process. When you enable the `tcpdump` command, it runs until the capture-duration timer and/or the packet-count counter threshold is met. If you do not set a threshold, the system uses a default of a 5 minute capture-duration and/or a single 1k file as the stopping point for the dump.

You can use the capture-duration timer and the packet-count counter at the same time. The TCP dump stops when the first of the thresholds is met. That means that even if the duration timer is 9000 seconds, if the maximum file count parameter is met first, the dumps stop.

To enable a TCP dump, use the following command.

- Enable a TCP dump for CPU bound traffic.

  CONFIGURATION mode

  `tcpdump cp [capture-duration` *`time`* `| filter` *`expression`* `| max-file-count` *`value`* `| packet-count` *`value`* `| snap-length` *`value`* `| write-to` *`path`*`]`

65

Standards Compliance

This chapter describes standards compliance for Dell EMC Networking products.

NOTE: Unless noted, when a standard cited here is listed as supported by the Dell EMC Networking OS, the system also supports predecessor standards. One way to search for predecessor standards is to use the http://tools.ietf.org/ website. Click "Browse and search IETF documents," enter an RFC number, and inspect the top of the resulting document for obsolescence citations to related RFCs.

Topics:

- IEEE Compliance
- RFC and I-D Compliance
- MIB Location

IEEE Compliance

The following is a list of IEEE compliance.

802.1AB	LLDP
802.1D	Bridging, STP
802.1p	L2 Prioritization
802.1Q	VLAN Tagging, Double VLAN Tagging, GVRP
802.1s	MSTP
802.1w	RSTP
802.1X	Network Access Control (Port Authentication)
802.3ab	Gigabit Ethernet (1000BASE-T)
802.3ac	Frame Extensions for VLAN Tagging
802.3ad	Link Aggregation with LACP
802.3ae	10 Gigabit Ethernet (10GBASE-W, 10GBASE-X)
802.3af	Power over Ethernet
802.3ak	10 Gigabit Ethernet (10GBASE-CX4)
802.3az	Energy Efficient Ethernet (EEE)
802.3i	Ethernet (10BASE-T)
802.3u	Fast Ethernet (100BASE-FX, 100BASE-TX)
802.3x	Flow Control
802.3z	Gigabit Ethernet (1000BASE-X)
ANSI/TIA-1057	LLDP-MED
Force10	FRRP (Force10 Redundant Ring Protocol)
Force10	PVST+
SFF-8431	SFP+ Direct Attach Cable (10GSFP+Cu)
MTU	12,000 bytes

RFC and I-D Compliance

Dell EMC Networking OS supports the following standards. The standards are grouped by related protocol. The columns showing support by platform indicate which version of Dell EMC Networking OS first supports the standard.

General Internet Protocols

The following table lists the Dell EMC Networking OS support per platform for general internet protocols.

Table 135. General Internet Protocols

RFC #	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
768	User Datagram Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
793	Transmission Control Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
854	Telnet Protocol Specification	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
959	File Transfer Protocol (FTP)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1321	The MD5 Message-Digest Algorithm	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1350	The TFTP Protocol (Revision 2)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1661	The Point-to-Point Protocol (PPP)		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1989	PPP Link Quality Monitoring		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1990	The PPP Multilink Protocol (MP)		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1994	PPP Challenge Handshake Authentication Protocol (CHAP)		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
24	International ization of the File	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Table 135. General Internet Protocols (continued)

RFC #	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
60	Transfer Protocol						
2474	Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2615	PPP over SONET/SDH		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2698	A Two Rate Three Color Marker		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3164	The BSD syslog Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
draft-ietf-bfd-base-03	Bidirectional Forwarding Detection		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

General IPv4 Protocols

The following table lists the Dell EMC Networking OS support per platform for general IPv4 protocols.

Table 136. General IPv4 Protocols

RFC#	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
791	Internet Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
792	Internet Control Message Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
826	An Ethernet Address Resolution Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1027	Using ARP to Implement Transparent Subnet Gateways	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Table 136. General IPv4 Protocols (continued)

RFC#	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1035	DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION (client)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1191	Path MTU Discovery	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1305	Network Time Protocol (Version 3) Specification, Implementation and Analysis	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1519	Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1542	Clarifications and Extensions for the Bootstrap Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1812	Requirements for IP Version 4 Routers	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2131	Dynamic Host Configuration Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2338	Virtual Router Redundancy Protocol (VRRP)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3021	Using 31-Bit Prefixes on IPv4 Point-to-Point Links	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3046	DHCP Relay Agent Information Option	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3069	VLAN Aggregation for Efficient IP Address Allocation	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3128	Protection Against a Variant of the Tiny Fragment Attack	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

General IPv6 Protocols

The following table lists the Dell EMC Networking OS support per platform for general IPv6 protocols.

Table 137. General IPv6 Protocols

RFC #	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1886	DNS Extensions to support IP version 6	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1981 (Partial)	Path MTU Discovery for IP version 6	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2460	Internet Protocol, Version 6 (IPv6) Specification	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Table 137. General IPv6 Protocols (continued)

RFC #	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
2462 (Partial)	IPv6 Stateless Address Autoconfiguration	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2464	Transmission of IPv6 Packets over Ethernet Networks	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2675	IPv6 Jumbograms	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2711	IPv6 Router Alert Option	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3587	IPv6 Global Unicast Address Format	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4007	IPv6 Scoped Address Architecture	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4291	Internet Protocol Version 6 (IPv6) Addressing Architecture	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4443	Internet Control Message Protocol (ICMPv6) for the IPv6 Specification	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4861	Neighbor Discovery for IPv6	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4862	IPv6 Stateless Address Autoconfiguration	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
5175	IPv6 Router Advertisement Flags Option	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Border Gateway Protocol (BGP)

The following table lists the Dell EMC Networking OS support per platform for BGP protocols.

Table 138. Border Gateway Protocol (BGP)

RFC#	Full Name	S-Series/Z-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1997	BGP ComAmtturnibituitees	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2385	Protection of BGP Sessions via the TCP MD5 Signature Option	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2439	BGP Route Flap Damping	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2545	Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2796	BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Table 138. Border Gateway Protocol (BGP) (continued)

RFC#	Full Name	S-Series/Z-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
2842	Capabilities Advertisement with BGP-4	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2858	Multiprotocol Extensions for BGP-4	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2918	Route Refresh Capability for BGP-4	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3065	Autonomous System Confederations for BGP	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4360	BGP Extended Communities Attribute	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4893	BGP Support for Four-octet AS Number Space	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
5396	Textual Representation of Autonomous System (AS) Numbers	8.1.2	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
draft-ietf-idrbgp4- 20	A Border Gateway Protocol 4 (BGP-4)	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
draft-ietf-idrrestart- 06	Graceful Restart Mechanism for BGP	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Open Shortest Path First (OSPF)

The following table lists the Dell EMC Networking OS support per platform for OSPF protocol.

Table 139. Open Shortest Path First (OSPF)

RFC #	Full Name	S-Series/Z-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1587	The OSPF Not-So-Stubby Area (NSSA) Option	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2154	OSPF with Digital Signatures	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2370	The OSPF Opaque LSA Option	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2740	OSPF for IPv6	9.1(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3623	Graceful OSPF Restart	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4222	Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Intermediate System to Intermediate System (IS-IS)

The following table lists the Dell EMC Networking OS support per platform for IS-IS protocol.

Table 140. Intermediate System to Intermediate System (IS-IS)

RFC#	Full Name	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1142	OSI IS-IS Intra-Domain Routing Protocol (ISO DP 10589)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2763	Dynamic Hostname Exchange Mechanism for IS-IS	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2966	Domain-wide Prefix Distribution with Two-Level IS-IS	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3373	Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3567	IS-IS ACruythpetongtircaapthioic n	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3784	Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
5120	MT-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
5306	Restart Signaling for IS-IS	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
5308	Routing IPv6 with IS-IS	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
draft-ietf-isis-igpp2p-over-lan-06	Point-to-point operation over LAN in link-state routing protocols	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
draft-kaplan-isis-e xt-eth-02	Extended Ethernet Frame Size Support	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Routing Information Protocol (RIP)

The following table lists the Dell EMC Networking OS support per platform for RIP protocol.

Table 141. Routing Information Protocol (RIP)

RFC#	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1058	Routing Information Protocol	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
2453	RIP Version	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
4191	Default Router Preferences and More-Specific Routes	8.3.12.0	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Multicast

The following table lists the Dell EMC Networking OS support per platform for Multicast protocol.

Table 142. Multicast

RFC#	Full Name	S-Series	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1112	Host Extensions for IP Multicasting	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2236	Internet Group Management Protocol, Version 2	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3376	Internet Group Management Protocol, Version 3	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3569	An Overview of Source-Specific Multicast (SSM)	7.8.1 SSM for IPv4	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
3618	Multicast Source Discovery Protocol (MSDP)		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)
draft-ietf-pim-sm-v2-new- 05	Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)	7.8.1 PIM-SM for IPv4	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.10(0.1)	9.10(0.1)

Network Management

The following table lists the Dell EMC Networking OS support per platform for network management protocol.

Table 143. Network Management

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1155	Structure and Identification of Management Information for TCP/IP-based Internets	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)

Table 143. Network Management (continued)

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
1156	Management Information Base for Network Management of TCP/IP-based internets	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1157	A Simple Network Management Protocol (SNMP)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1212	Concise MIB Definitions	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1215	A Convention for Defining Traps for use with the SNMP	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1493	Definitions of Managed Objects for Bridges [except for the dot1dTpLearnedEntryDiscards object]	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1724	RIP Version 2 MIB Extension		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1850	OSPF Version 2 Management Information Base	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
1901	Introduction to Community-based SNMPv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2011	SNMPv2 Management Information Base for the Internet Protocol using SMIv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2012	SNMPv2 Management Information Base for the Transmission Control Protocol using SMIv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2013	SNMPv2 Management Information Base for the User Datagram Protocol using SMIv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2024	Definitions of Managed Objects for Data Link Switching using SMIv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2096	IP Forwarding Table MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2558	Definitions of Managed Objects for the Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Interface Type		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2570	Introduction and Applicability Statements for Internet Standard Management Framework	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2571	An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2572	Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2574	User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)

Table 143. Network Management (continued)

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
2575	View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2576	Coexistence Between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2578	Structure of Management Information Version 2 (SMIv2)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2579	Textual Conventions for SMIv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2580	Conformance Statements for SMIv2	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2618	RADIUS Authentication Client MIB, except the following four counters: radiusAuthClientInvalidServerAddresses radiusAuthClientMalformedAccessResponses radiusAuthClientUnknownTypes radiusAuthClientPacketsDropped	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2698	A Two Rate Three Color Marker	9.5.(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
3635	Definitions of Managed Objects for the Ethernet-like Interface Types	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2674	Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2787	Definitions of Managed Objects for the Virtual Router Redundancy Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2819	Remote Network Monitoring Management Information Base: Ethernet Statistics Table, Ethernet History Control Table, Ethernet History Table, Alarm Table, Event Table, Log Table	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2863	The Interfaces Group MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
2865	Remote Authentication Dial In User Service (RADIUS)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
3273	Remote Network Monitoring Management Information Base for High Capacity Networks (64 bits): Ethernet Statistics High-Capacity Table, Ethernet History High-Capacity Table	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
3416	Version 2 of the Protocol Operations for the Simple	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)

Table 143. Network Management (continued)

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
	Network Management Protocol (SNMP)						
3418	Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
3434	Remote Monitoring MIB Extensions for High Capacity Alarms, High-Capacity Alarm Table (64 bits)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
3580	IEEE 802.1X Remote Authentication Dial In User Service (RADIUS) Usage Guidelines	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
3815	Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
4001	Textual Conventions for Internet Network Addresses	8.3.12	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
4292	IP Forwarding Table MIB	9.5.(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
4750	OSPF Version 2 Management Information Base	9.5.(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
4502	RMON v2 MIB	9.5(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
5060	Protocol Independent Multicast MIB	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
ANSI/TIA-1057	The LLDP Management Information Base extension module for TIA-TR41.4 Media Endpoint Discovery information	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
draft-grant-tacacs-02	The TACACS+ Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
draft-ietf-idr-bgp4 -mib-06	Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
draft-ietf-isis-wgmib - 16	Management Information Base for Intermediate System to Intermediate System (IS-IS): isisSysObject (top level scalar objects) isisISAdjTable isisISAdjAreaAddrTable isisISAdjIPAddrTable isisISAdjProtSuppTable		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
draft-ietf-netmod-	Defines a YANG data model for the configuration of network interfaces. Used in	9.2(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)

Table 143. Network Management (continued)

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
interfaces-cfg-03	the Programmatic Interface RESTAPI feature.						
IEEE 802.1AB	Management Information Base module for LLDP configuration, statistics, local system data and remote systems data components.	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
IEEE 802.1AB	The LLDP Management Information Base extension module for IEEE 802.1 organizationally defined discovery information. (LLDP DOT1 MIB and LLDP DOT3 MIB)	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
IEEE 802.1AB	The LLDP Management Information Base extension module for IEEE 802.3 organizationally defined discovery information. (LLDP DOT1 MIB and LLDP DOT3 MIB)	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
ruzin-mstp-mib-02 (Traps)	Definitions of Managed Objects for Bridges with Multiple Spanning Tree Protocol	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
sFlow.org	sFlow Version 5	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
sFlow.org	sFlow Version 5 MIB	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-BGP4-V2-MIB	Force10 BGP MIB (draft-ietf-idr-bgp4-mibv2-05)	7.8.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
f10–bmp-mib	Force10 Bare Metal Provisioning MIB	9.2(0.0)	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-FIB-MIB	Force10 CIDR Multipath Routes MIB (The IP Forwarding Table provides information that you can use to determine the egress port of an IP packet and troubleshoot an IP reachability issue. It reports the autonomous system of the next hop, multiple next hop support, and policy routing support)		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-CS-CHASSIS-MIB	Force10 C-Series Enterprise Chassis MIB		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)

Table 143. Network Management (continued)

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
FORCE10-IF-EXTENSION-MIB	Force10 Enterprise IF Extension MIB (extends the Interfaces portion of the MIB-2 (RFC 1213) by providing proprietary SNMP OIDs for other counters displayed in the "show interfaces" output)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-LINKAGG-MIB	Force10 Enterprise Link Aggregation MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-CHASSIS-MIB	Force10 E-Series Enterprise Chassis MIB		9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-COPY-CONFIG-MIB	Force10 File Copy MIB (supporting SNMP SET operation)	7.7.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-MONMIB	Force10 Monitoring MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-PRODUCTS-MIB	Force10 Product Object Identifier MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-SS-CHASSIS-MIB	Force10 S-Series Enterprise Chassis MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-SMI	Force10 Structure of Management Information	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-SYSTEM-COMPONENT-MIB	Force10 System Component MIB (enables the user to view CAM usage information)	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-TC-MIB	Force10 Textual Convention	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)
FORCE10-TRAP-	Force10 Trap Alarm MIB	7.6.1	9.8(0.0P2)	9.8(0.0P5)	9.8(1.0)	9.8(1.0)	9.8(1.0)

Table 143. Network Management (continued)

RFC#	Full Name	S4810	S3048–ON	S4048–ON	Z9100–ON	S4048T-ON	S6010–ON
ALAR M-MIB							

MIB Location

You can find Force10 MIBs under the Force10 MIBs subhead on the Documentation page of iSupport:

https://www.force10networks.com/CSPortal20/KnowledgeBase/Documentation.aspx

You also can obtain a list of selected MIBs and their OIDs at the following URL:

https://www.force10networks.com/CSPortal20/Main/Login.aspx

Some pages of iSupport require a login. To request an iSupport account, go to:

https://www.force10networks.com/CSPortal20/AccountRequest/AccountRequest.aspx

If you have forgotten or lost your account information, contact Dell TAC for assistance.

66

X.509v3

supports X.509v3 standards.

Topics:

- Introduction to X.509v3 certificates
- X.509v3 support in
- Information about installing CA certificates
- Information about Creating Certificate Signing Requests (CSR)
- Information about installing trusted certificates
- Transport layer security (TLS)
- Online Certificate Status Protocol (OSCP)
- Verifying certificates
- Event logging

Introduction to X.509v3 certificates

X.509v3 is a standard for public key infrastructure (PKI) to manage digital certificates and public key encryption.

The X.509v3 standard specifies a format for public-key certificates or digital certificates.

Transport Layer Security (TLS) relies on public key certificates to work.

X.509v3 certificates

A X.509v3 or digital certificate is an electronic document used to prove ownership of a public key. It contains information about the key's identity, information about the key's owner, and the digital signature of an entity that has verified the certificate's content as correct.

Certificate authority (CA)

The entity that verifies the contents of the digital certificate and signs it indicating that the certificate is valid and correct is called the Certificate Authority (CA).

Certificate signing requests (CSR)

In an X.509v3 system, an entity that wants a signed certificate or a digital certificate requests one through a Certificate Signing Request (CSR).

How certificates are requested

The following enumeration describes the generic steps that are involved in issuing a digital certificate:

1. An entity or organization that wants a digital certificate requests one through a CSR.
2. To request a digital certificate through a CSR, a key pair is generated and the CSR is signed using the secret private key. The CSR contains information identifying the applicant and the applicant's public key. This public key is used to verify the signature of the CSR and the Distinguished Name (DN).
3. This CSR is sent to a Certificate Authority (CA). The CA verifies the certificate and signs it using the CA's own private key.
4. The CA then issues the certificate by binding a public key to a particular distinguished name (DN). This certificate becomes the entity's trusted root certificate.

Advantages of X.509v3 certificates

Public key authentication is preferred over password-based authentication, although both may be used in conjunction, for various reasons. Public-key authentication provides the following advantages over normal password-based authentication:

- Public-key authentication avoids the human problems of low-entropy password selection and provides more resistance to brute-force attacks than password-based authentication.
- It facilitates trusted, provable identities—when using certificates signed by trusted CAs.
- It also provides integrity and confidentiality in addition to authentication.

X.509v3 support in

supports X.509v3 standards.

Many organizations or entities need to let their customers know that the connection to their devices and network is secure. These organizations pay an internationally trusted Certificate Authorities (CAs) such as VeriSign, DigiCert, and so on, to sign a certificate for their domain.

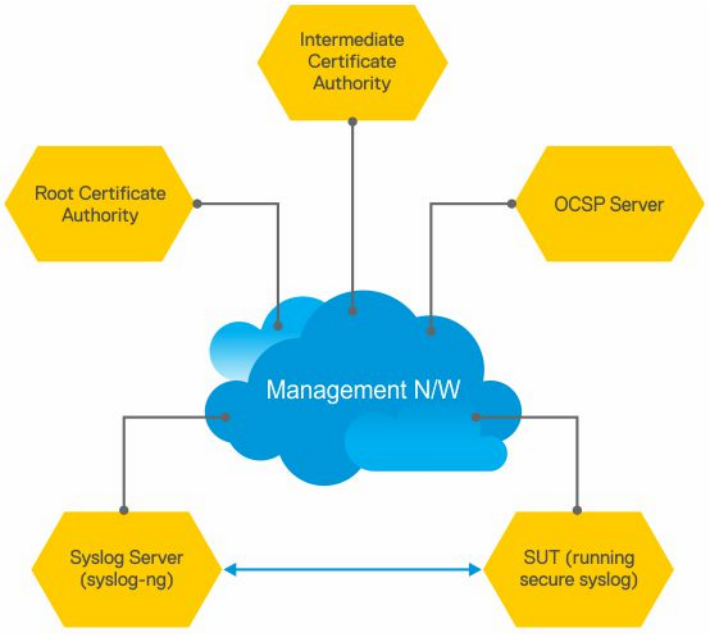
To implement a X.509v3 infrastructure, recommends you to act as your own CA. Common use cases for acting as your own CA include issuing certificates to clients to allow them to authenticate to a server. For example, Apache, OpenVPN, and so on.

Acting as a certificate authority (CA) means dealing with cryptographic pairs of private keys and public certificates. The first cryptographic pair you create is the root pair. This root pair consists of the root key (ca.key.pem) and root certificate—ca.cert.pem. This pair forms the identity of your CA.

Typically, a root CA does not sign server or client certificates directly. The root CA is only ever used to create one or more intermediate CAs. These intermediate CAs are trusted by the root CA to sign certificates on their behalf. This is the best practice. It allows the root key to be kept offline and used to a minimal extent, as any compromise of the root key is disastrous.

For more generic information on setting up your own Certificate Authority (CA), see https://jamielinux.com/docs/openssl-certificate-authority/index.html#

The following figure illustrates a sample network topology in which a simple X.509v3 infrastructure is implemented:



The Root CA generates a private key and a self-signed CA certificate.

The Intermediate CA generates a private key and a Certificate Signing Request (CSR).

Using its private key, the root CA signs the intermediate CA's CSR generating a CA certificate for the Intermediate CA. This intermediate CA can then sign certificates for hosts in the network and also for further intermediate CAs. These CA certificates (root CA and any intermediate CAs), but not the corresponding private keys, are made publicly available on the network.

NOTE: CA certificates may also be bundled together for ease of installation. Their .PEM files are concatenated in order from the "lowest" ranking CA certificate to the Root CA certificate. handles installation of bundled certificate files.

The other hosts on the network, such as the SUT switch, syslog server, and OCSP server, generate private keys and create Certificate Signing Requests (CSRs). The hosts then upload the CSRs to the Intermediate CA or make the CSRs available for the Intermediate CA to download. generates a CSR using the crypto cert generate request command.

The hosts on the network (SUT, syslog, OCSP…) also download and install the CA certificates from the Root and Intermediate CAs. By installing these CA certificates, the hosts trust any certificates signed by these CAs.

NOTE: You can download and install CA certificates in one step using the crypto ca-cert install command.

The intermediate CA signs the CSRs and makes the resulting certificates available for download through FTP root or otherwise.

Alternatively, the Intermediate CA can also generate private keys and certificates for the hosts. The CA then makes the private key or certificate pairs available for each host to download. You can password-encrypt the private key for additional security and then decrypt it with a password using the crypto cert install command.

The hosts on the network (SUT, syslog, OCSP…) download and install their corresponding signed certificates. These hosts can also verify whether they have their own certificates using the private key that they have previously generated.

NOTE: When you use the crypto cert install command to download and install certificates, automatically verifies whether a device has its own certificate.

Now that the X.509v3 certificates are installed on the SUT and Syslog server, these certificates can be used during TLS protocol negotiations so that the devices can verify each other's trustworthiness and exchange session keys to protect session data. The devices verify each other's certificates using the CA certificates they installed earlier. The SUT enables Syslog-over-TLS by configuring the `secure` keyword in the logging configuration. For example, logging 10.11.178.1 secure 6514.

During the initial TLS protocol negotiation, both participating parties also check to see if the other's certificate is revoked by the CA. To do this check, the devices query the CA's designated OCSP responder on the network. The OCSP responder information is included in the presented certificate, the Intermediate CA inserts the info upon signing it, or it may be statically configured on the host.

Information about installing CA certificates

Dell EMC Networking OS enables you to download and install X.509v3 certificates from Certificate Authorities (CAs).

In a data center environment, CA certificates are created by trusted hosts on the network. By digitally signing devices' certificates with the CA's private key, trust can be established among all devices in a network. These CA certificates, installed on each of the devices, are used to verify certificates presented by clients and servers such as the Syslog servers.

Dell EMC Networking OS allows you to download CA certificates using the `crypto ca-cert install` command. In this command, you can specify:

- That the certificate is a CA certificate
- The location from which to download the certificate and the protocol to use. For example, `tftp://192.168.1.100/certificates/CAcert.pem`. Locations can be usbflash, built-in flash, TFTP, FTP, or SCP hosts.

After you download a CA certificate, the system verifies the following aspects of the CA certificate:

- The system checks if "CA:TRUE" is specified in the certificate's extensions section and the `keyCertSign bit (bit 5)` is set in the `KeyUsage` bit string extension. If these extensions are not set, the system does not install the certificate.
- The system checks if the Issuer and Subject fields are the same. If these fields are the same, then the certificate is a self-signed certificate. These certificates are also called the root CA certificates, as they are not signed by another CA. The system verifies the certificate with its own public key and install the certificate.
- If the Issuer and Subjects fields differ, then the certificate is signed by another CA farther up the chain. These certificates are also called intermediate certificates. If a higher CA certificate is installed on the switch, then the system verifies the downloaded certificate with the CA's public key. The system repeats this process until the root certificate is reached. The certificate is rejected if the signature verification fails.
- If a higher CA certificate is not installed on the switch, the system rejects the intermediate CA certificate and logs the attempt. The system also displays a message indicating the reason for the failure of CA certificate installation. The system checks the "not before" and "not after" fields against the current system date to ensure that the certificate has not expired.

The verified CA certificate is installed on the switch by adding it to an existing file that contains trusted certificates. The certificate is inserted into the certificate file that stores certificates in a root-last order. Meaning, the downloaded certificate is fit into the file before its own issuer but following any certificates that it may have issued. This way, the system ensures that the CA certificates file is kept in a root-last order. The file may contain multiple certificates in PEM format concatenated together. This file is stored in a private and persistent location on the device such as the `flash://ADMIN_DIR` folder.

After the CA certificate is installed, the system can secure communications with TLS servers by verifying certificates that are signed by the CA.

Installing CA certificate

To install a CA certificate, enter the `crypto ca-cert install {path}` command in Global Configuration mode.

Information about Creating Certificate Signing Requests (CSR)

Certificate Signing Request (CSR) enables a device to get a X.509v3 certificate from a CA.

In order for a device to get a X.509v3 certificate, the device first requests a certificate from a CA through a Certificate Signing Request (CSR). While creating a CSR, you need to provide the information about the certificate and the private key details. Dell EMC Networking OS enable you to create a private key and a CSR for a device using a single command.

If you do not specify the cert-file option, the system prompts you to enter metadata information related to the CSR as follows:

```
You are about to be asked to enter information that will be incorporated into your
certificate request.
What you are about to enter is what is called a Distinguished Name or a DN.
There are quite a few fields but you can leave some blank.
For some fields there will be a default value; if you enter '.', the field will be left
blank.
-----

Country Name (2 letter code) [US]:
State or Province Name (full name) [Some-State]:California
Locality Name (eg, city) []:San Francisco
Organization Name (eg, company) []:Starfleet Command
Organizational Unit Name (eg, section) []:NCC-1701A
Common Name (eg, YOUR name) [hostname]:S4810-001
Email Address []:scotty@starfleet.com
```

The switch uses SHA-256 as the digest algorithm and the public key algorithm is RSA with a 2048-bit modulus. The **KeyUsage** bits of the certificate assert **keyEncipherment (bit 2)** and **keyAgreement (bit 4)**. The **keyCertSign bit (bit 5)** is NOT be set. The **ExtendedKeyUsage** fields indicate **serverAuth** and **clientAuth**.

The attribute **CA:FALSE** is set in the Extensions section of the certificate. The certificate is NOT used to validate other certificates. The CSR is then copied out to the CA server. It can be copied from flash to a destination like usbflash, TFTP, FTP, or SCP.

The CA server signs the CSR with its private key. The CA server then makes the signed certificate available for the requesting device to download and install.

Creating Certificate Signing Requests (CSR)

To create a private key and CSR, perform the following step:

In global configuration mode, enter the following command:
`crypto cert generate {self-signed | request} [cert-file cert-path key-file {private | key-path}] [country 2-letter code] [state state] [locality city] [organization organization-name] [orgunit unit-name] [cname common-name] [email email-address] [validity days] [length length] [altname alt-name]`

You must specify the following parameters for this command:

- Certificate File
- Private Key
- Country Name
- State or Province Name
- Locality Name
- Organization Name
- Organization Unit Name

- Common Name
- Email address
- Validity
- Length
- Alternate Name

NOTE: The command contains multiple options with the Common Name being a required field and blanks being filled in for unspecified fields.

Information about installing trusted certificates

Dell EMC Networking OS also enables you to install a trusted certificate. The system can then present this certificate for authentication to clients such as SSH and HTTPS.

This trusted certificate is also presented to the TLS server implementations that require client authentication such as Syslog. The certificate is digitally signed with the private key of a CA server.

You can download the trusted certificate for a device from flash, usbflash, tftp, ftp, or scp. This certificate is stored in the BSD file system and can be used to authenticate the switch to clients.

Installing trusted certificates

To install a trusted certificate, perform the following step:

In global configuration mode, enter the following command:

```
crypto cert inatall {path}
```

Transport layer security (TLS)

Transport Layer Security (TLS) provides cryptographic protection for TCP-based application protocols.

In Dell EMC Networking OS, TLS already protects secure HTTP for the REST and HTTPD server implementations.

NOTE: There are three modern versions of the TLS protocol: 1.0, 1.1, and 1.2. Older versions are called SSL v1, v2, and v3, and are not supported.

The TLS protocol implementation in Dell EMC Networking OS takes care of the following activities:

- Session negotiation and shutdown
  - Protocol Version
  - Cryptographic algorithm selection
- Session resumption and renegotiation
- Certificate revocation checking, which may be accomplished through OCSP

When operating in FIPS mode, the system is restricted to only the TLS 1.2 protocol version and support the following cipher suites in line with the NIST SP800-131A Rev 1 policy document—published July 2015:

```
TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA384
TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256
TLS_DHE_RSA_WITH_AES_256_CBC_SHA256
TLS_DHE_RSA_WITH_AES_128_CBC_SHA256
TLS_ECDH_RSA_WITH_AES_256_CBC_SHA256
TLS_ECDH_RSA_WITH_AES_128_CBC_SHA256
TLS_DH_RSA_WITH_AES_256_CBC_SHA256
TLS_DH_RSA_WITH_AES_128_CBC_SHA256
TLS_RSA_WITH_AES_128_CBC_SHA256
TLS_RSA_WITH_AES_256_CBC_SHA256
```

When not operating in FIPS mode, the system may support TLS 1.0 up to 1.2, and older ciphers and hashes:

```
TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA
TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA
TLS_DHE_RSA_WITH_AES_256_CBC_SHA
```

```
TLS_DHE_RSA_WITH_AES_128_CBC_SHA
TLS_DHE_RSA_WITH_3DES_EDE_CBC_SHA
TLS_ECDHE_RSA_WITH_3DES_EDE_CBC_SHA
TLS_ECDH_RSA_WITH_AES_256_CBC_SHA
TLS_ECDH_RSA_WITH_AES_128_CBC_SHA
TLS_DH_RSA_WITH_AES_256_CBC_SHA
TLS_DH_RSA_WITH_AES_128_CBC_SHA
```

TLS compression is disabled by default. TLS session resumption is also supported to reduce processor and traffic overhead due to public key cryptographic operations and handshake traffic. However, the maximum time allowed for a TLS session to resume without repeating the TLS authentication or handshake process is configurable with a default of 1 hour. You can also disable session resumption.

Syslog over TLS

Syslog over TLS mandates that a client certificate must be presented, to ensure that all Syslog entries written to the server are from a trusted client.

Online Certificate Status Protocol (OSCP)

Use the Online Certificate Status Protocol (OSCP) to obtain the revocation status of a X.509v3 certificate.

A device or a Certificate Authority (CAs) can check the status of a X.509v3 certificate by sending an OCSP request to an OCSP server or responder. An OCSP responder, a server typically run by the certificate issuer, returns a signed response signifying that the certificate specified in the request is 'good', 'revoked', or 'unknown'. The OCSP response indicates whether the presented certificate is valid.

OCSP provides a way for Certificate Authorities to revoke signed certificates before the expiration date. In a CA certificate, OCSP Responder information is specified in the `authorityInfoAccess` extension.

A CA can verify the revocation status of a certificate with multiple OCSP responders. When multiple OCSP responders exist, you can configure the order or preference the CA takes while contacting various OCSP responders for verification.

Upon receiving a presented certificate, the system sends an OCSP request to an OCSP responder through HTTP. The system then verifies the OCSP response using either a trusted public key or the OCSP responder's own self-signed certificate. This self-signed certificate installs on the device's trusted location even before an OCSP request is made. The system accepts or rejects the presented certificate based on the OCSP response.

In a scenario where all OCSP responders are unreachable, the switch accepts the certificate. This action is the default behavior. You can also configure an alternate system behavior when all OCSP responders are unreachable. However, the switch may become vulnerable to denial-of-service attack if you configure the system to deny the certificate when OCSP responders are not reachable. For more information, see Configuring Revocation Behavior.

The system creates logs for the following events:

- Failures to reach OCSP responders
- Invalid OCSP responses—for example, cannot verify the signed response with an installed CA certificate.
- Rejection of a certificate due to OCSP

Configuring OCSP setting on CA

You can configure the CA to contact multiple OCSP servers.

To configure OCSP server for a CA, perform the following step:

In the certificate mode, enter the following command:
`ocsp-server` *`URL`* `[nonce] [sign-requests]`

NOTE: If you have an IPv6 address in the URL, then enclose this address in square brackets. For example, http://[1100::203]:6514.

Configuring OCSP behavior

You can configure how the OCSP requests and responses are signed when the CA or the device contacts the OCSP responders.

To configure this behavior, follow this step:

In CONFIGURATION mode, enter the following command:

```
crypto x509 ocsp {[nonce] [sign-request]}
```

Both the `none` and `sign-request` parameters are optional. The default behavior is to not use these two options. If your OCSP responder uses pre-computed responses, you cannot use the `none` feature in the switch's communcations with the responder. If your OCSP responder requires signed requests, you can use the `sign-requests` option.

Configuring Revocation Behavior

You can configure the system behavior if an OCSP responder fails.

By default, when all the OCSP responders fail to send a response to an OSCP request, the system accepts the certificate and logs the event. However, you can configure the system to reject the certificate in case OCSP responders fail.

To configure OCSP revocation settings:

In CONFIGURATION mode, enter the following command:

```
crypto x509 revocation ocsp [accept | reject]
```

The default behavior is to accept certificates if either an OCSP responder is unavailable or if no responder is identified.

Configuring OSCP responder preference

You can configure the preference or order that the CA or a device follows while contacting multiple OCSP responders.

Enter the following command in Certificate mode:

```
ocsp-server prefer
```

Verifying certificates

A CA certificate's public key is used to decrypt a presented certificate's signature to obtain a hash value.

The rest of the presented certificate is also hashed and if the two hashes match then the certificate is considered valid.

During verification, the system checks the presented certificates for revocation information. The system also enables you to configure behavior in case a certificate's revocation status cannot be verified; for example, when the OCSP responder is unreachable you can alter system behavior to accept or reject the certificate depending on configuration. The default behavior is to accept the certificates. The system also logs the events where the OSCP responders fail or invalid OSCP responses are received.

NOTE: A CA certificate can also be revoked.

Verifying Server certificates

Verifying server certificates is mandatory in the TLS protocol.

As a result, all TLS-enabled applications require certificate verification, including Syslog servers. The system checks the Server certificates against installed CA certificates.

NOTE: As part of the certificate verification, the hostname or IP address of the server is verified against the hostname or IP address specified in the application. For example, when using SYSLOG over TLS, the hostname or IP address specified in the `logging syslog-server secure port port-number` command is compared against the SubjectAltName or Common Name field in the server certificate.

Verifying Client Certificates

Verifying client certificates is optional in the TLS protocol and is not explicitly required by Common Criteria.

However, TLS-protected Syslog and RADIUS protocols mandate that certificate-based mutual authentication be performed.

Event logging

The system logs the following events:

- A CA certificate is installed or deleted.
- A self-signed certificate and private key are generated.
- An existing host certificate, a private key, or both are deleted.
- A host certificate and private key are installed successfully.
- An installed certificate (host certificate or CA certificate) is within seven days of expiration. This alert is repeated periodically.
- An OCSP request is not answered with an OCSP response.
- A secure session negotiation fails due to invalid, expired, or revoked certificate.