Dell PowerEdge FN I/O Module

Configuration Guide

9.9(0.0)

Notes, cautions, and warnings

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

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.

Copyright © 2015 Dell Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. Dell™ and the Dell logo are trademarks of Dell Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

2015 - 09

Rev. A00

Contents

1 About this Guide..... 33

Audience..... 33
Conventions..... 33
Information Symbols..... 34
Related Documents..... 34

2 Before You Start..... 35

Operational Modes..... 35
Standalone mode..... 35
VLT mode..... 35
Programmable MUX mode..... 35
Stacking mode..... 36
Full Switch mode..... 36
Default Settings..... 36
Other Auto-Configured Settings..... 37
Data Center Bridging Support..... 37
FCoE Connectivity and FIP Snooping..... 38
iSCSI Operation..... 38
Link Aggregation..... 38
Link Tracking..... 38
Configuring VLANs..... 39
Uplink LAG..... 39
Server-Facing LAGs..... 39
Where to Go From Here..... 39

3 Configuration Fundamentals..... 40

Accessing the Command Line..... 40
CLI Modes..... 40
Navigating CLI Modes..... 41
The do Command..... 42
Undoing Commands..... 43
Obtaining Help..... 43
Entering and Editing Commands..... 44
Command History..... 45
Filtering show Command Outputs..... 45
Multiple Users in Configuration Mode..... 46

4 Getting Started..... 47

Console Access....48
Serial Console....48
External Serial Port with a USB Connector....50
Accessing the CLI Interface and Running Scripts Using SSH....50
Entering CLI commands Using an SSH Connection....50
Executing Local CLI Scripts Using an SSH Connection....50
Boot Process....51
Default Configuration....53
Configuring a Host Name....53
Configuring a Host Name....53
Accessing the System Remotely....54
Accessing the Switch Remotely....54
Configure the Management Port IP Address....54
Configure a Management Route....54
Configuring a Username and Password....55
Configuring the Enable Password....55
Configuration File Management....56
Copy Files to and from the System....56
Save the Running-Configuration....57
Viewing Files....58
Managing the File System....59
View the Command History....60
Using HTTP for File Transfers....61
Upgrading and Downgrading the Dell Networking OS....61
Using Hashes to Verify Software Images Before Installation....62
Deploying FN I/O Module....63
Ensuring the FN IOM system is in Standalone Mode....64
Configuration Examples....64
Configure Interfaces and Port Channel....66
Verifying the Configurations....67

5 Management....68

Configuring Privilege Levels....68
Creating a Custom Privilege Level....68
Customizing a Privilege Level....69
Applying a Privilege Level to a Username....70
Applying a Privilege Level to a Terminal Line....71
Configuring Logging....71
Audit and Security Logs....71
Configuring Logging Format....73
Setting Up a Secure Connection to a Syslog Server....73
Display the Logging Buffer and the Logging Configuration....75

Log Messages in the Internal Buffer ... 75
Configuration Task List for System Log Management ... 76
Disabling System Logging ... 76
Sending System Messages to a Syslog Server ... 76
Configuring a UNIX System as a Syslog Server ... 76
Changing System Logging Settings ... 77
Display the Logging Buffer and the Logging Configuration ... 78
Configuring a UNIX Logging Facility Level ... 78
Synchronizing Log Messages ... 80
Enabling Timestamp on Syslog Messages ... 80
File Transfer Services ... 81
Configuration Task List for File Transfer Services ... 81
Enabling the FTP Server ... 81
Configuring FTP Server Parameters ... 81
Configuring FTP Client Parameters ... 82
Terminal Lines ... 82
Denying and Permitting Access to a Terminal Line ... 82
Configuring Login Authentication for Terminal Lines ... 83
Setting Time Out of EXEC Privilege Mode ... 84
Using Telnet to get to Another Network Device ... 85
Lock CONFIGURATION Mode ... 85
Viewing the Configuration Lock Status ... 86
Limit Concurrent Login Sessions ... 86
Restrictions for Limiting the Number of Concurrent Sessions ... 87
Configuring Concurrent Session Limit ... 87
Enabling the System to Clear Existing Sessions ... 87
Track Login Activity ... 88
Restrictions for Tracking Login Activity ... 88
Configuring Login Activity Tracking ... 88
Display Login Statistics ... 89
Recovering from a Forgotten Password ... 90
Recovering from a Forgotten Enable Password ... 91
Recovering from a Failed Start ... 91

6 802.1X ... 93

The Port-Authentication Process ... 95
EAP over RADIUS ... 96
Configuring 802.1X ... 97
Related Configuration Tasks ... 97
Important Points to Remember ... 97
Enabling 802.1X ... 98
Configuring Request Identity Re-Transmissions ... 99

Configuring a Quiet Period after a Failed Authentication....100
Forcibly Authorizing or Unauthorizing a Port....101
Re-Authenticating a Port....102
Configuring Timeouts....103
Configuring Dynamic VLAN Assignment with Port Authentication....104
Guest and Authentication-Fail VLANs....105
Configuring a Guest VLAN....106
Configuring an Authentication-Fail VLAN....106

7 Access Control List (ACL) VLAN Groups and Content Addressable Memory (CAM)....108
Optimizing CAM Utilization During the Attachment of ACLs to VLANs....108
Guidelines for Configuring ACL VLAN groups....109
Configuring ACL VLAN Groups and Configuring FP Blocks for VLAN Parameters....110
Configuring ACL VLAN Groups....110
Configuring FP Blocks for VLAN Parameters....111
Viewing CAM Usage....112
Allocating FP Blocks for VLAN Processes....113

8 Access Control Lists (ACLs)....115
IP Access Control Lists (ACLs)....115
Implementing ACL on the Dell Networking OS....116
ACLs and VLANs....116
ACL Optimization....116
Determine the Order in which ACLs are Used to Classify Traffic....116
Example of the order Keyword to Determine ACL Sequence....117
IP Fragment Handling....117
IP Fragments ACL Examples....117
Layer 4 ACL Rules Examples....118
Configure a Standard IP ACL....119
Configuring a Standard IP ACL Filter....120
Configure an Extended IP ACL....121
Configuring Filters with a Sequence Number....121
Configuring Filters Without a Sequence Number....122
Established Flag....122
Configure Layer 2 and Layer 3 ACLs....123
Assign an IP ACL to an Interface....123
Applying an IP ACL....124
Counting ACL Hits....124
Configure Ingress ACLs....125
Configure Egress ACLs....125
Applying Egress Layer 3 ACLs (Control-Plane)....126

IP Prefix Lists....127
Implementation Information....127
Configuration Task List for Prefix Lists....127
Creating a Prefix List....128
Creating a Prefix List Without a Sequence Number....129
Viewing Prefix Lists....129
Applying a Prefix List for Route Redistribution....130
Applying a Filter to a Prefix List (OSPF)....131
ACL Resequencing....131
Resequencing an ACL or Prefix List....132
Route Maps....133
Implementation Information....133
Important Points to Remember....134
Configuration Task List for Route Maps....134
Creating a Route Map....134
Configure Route Map Filters....136
Configuring Match Routes....137
Configuring Set Conditions....137
Configure a Route Map for Route Redistribution....138
Configure a Route Map for Route Tagging....139
Continue Clause....139
Logging of ACL Processes....140
Guidelines for Configuring ACL Logging....141
Configuring ACL Logging....141
Flow-Based Monitoring Support for ACLs....142
Behavior of Flow-Based Monitoring....142
Enabling Flow-Based Monitoring....144

9 Bidirectional Forwarding Detection (BFD)....146

How BFD Works....146
BFD Packet Format....147
BFD Sessions....149
BFD Three-Way Handshake....149
Session State Changes....150
Important Points to Remember....151
Configure BFD....151
Configure BFD for Physical Ports....152
Enabling BFD Globally....152
Establishing a Session on Physical Ports....153
Changing Physical Port Session Parameters....154
Disabling and Re-Enabling BFD....155
Configure BFD for Static Routes....155

Related Configuration Tasks....156
Establishing Sessions for Static Routes....156
Changing Static Route Session Parameters....157
Disabling BFD for Static Routes....157
Configure BFD for OSPF....157
Related Configuration Tasks....157
Establishing Sessions with OSPF Neighbors....158
Changing OSPF Session Parameters....159
Disabling BFD for OSPF....159
Configure BFD for OSPFv3....160
Related Configuration Tasks....160
Establishing Sessions with OSPFv3 Neighbors....160
Changing OSPFv3 Session Parameters....160
Disabling BFD for OSPFv3....161
Configure BFD for BGP....161
Prerequisites....161
Establishing Sessions with BGP Neighbors....162
Disabling BFD for BGP....164
Use BFD in a BGP Peer Group....164
Displaying BFD for BGP Information....164
Configure BFD for VRRP....168
Related Configuration Tasks....169
Establishing Sessions with All VRRP Neighbors....169
Establishing VRRP Sessions on VRRP Neighbors....169
Changing VRRP Session Parameters....170
Disabling BFD for VRRP....171
Configure BFD for VLANs....171
Related Configuration Task....172
Establish Sessions with VLAN Neighbors....172
Changing VLAN Session Parameters....172
Disabling BFD for VLANs....173
Configure BFD for Port-Channels....173
Related Configuration Tasks....173
Establish Sessions on Port-Channels....174
Changing Physical Port Session Parameters....174
Disabling BFD for Port-Channels....175
Configuring Protocol Liveness....175
Troubleshooting BFD....175

10 Border Gateway Protocol IPv4 (BGPv4)....177

Autonomous Systems (AS)....177
Sessions and Peers....179

Establish a Session 180
Route Reflectors 181
Communities 181
BGP Attributes 182
Best Path Selection Criteria 182
Weight 184
Local Preference 184
Multi-Exit Discriminators (MEDs) 185
Origin 186
AS Path 187
Next Hop 187
Multiprotocol BGP 187
Implement BGP with the Dell Networking OS 188
Additional Path (Add-Path) Support 188
Advertise IGP Cost as MED for Redistributed Routes 188
Ignore Router-ID for Some Best-Path Calculations 189
Four-Byte AS Numbers 189
AS4 Number Representation 190
AS Number Migration 191
BGP4 Management Information Base (MIB) 193
Important Points to Remember 193
Configuration Information 194
BGP Configuration 194
Enabling BGP 195
Enabling MBGP Configurations 228
BGP Regular Expression Optimization 229
Debugging BGP 229
Storing Last and Bad PDUs 230
PDU Counters 231
Sample Configurations 231

11 Content Addressable Memory (CAM) 241
CAM Allocation 241
Test CAM Usage 242
View CAM-ACL Settings 243
CAM Optimization 243

12 Control Plane Policing (CoPP) 244
Configure Control Plane Policing 245
Configuring CoPP for Protocols 246
Configuring CoPP for CPU Queues 248
Show Commands 249

13 Data Center Bridging (DCB)....251
Supported Modes....251
Ethernet Enhancements in Data Center Bridging....251
Priority-Based Flow Control....252
Enhanced Transmission Selection....253
Data Center Bridging Exchange Protocol (DCBx)....255
Creating a DCB Map....255
Important Points to Remember....256
Applying a DCB Map on Server-Facing Ethernet Ports....256
Data Center Bridging: Default Configuration....256
Data Center Bridging in a Traffic Flow....257
Data Center Bridging: Auto-DCB-Enable Mode....257
Configuring Priority-Based Flow Control....259
How Priority-Based Flow Control is Implemented....261
Configuring Lossless Queues....262
Configuring the PFC Buffer in a Switch Stack....263
Configuring Enhanced Transmission Selection....264
How Enhanced Transmission Selection is Implemented....264
ETS Operation with DCBx....265
Hierarchical Scheduling in ETS Output Policies....266
DCBx Operation....266
DCBx Operation....267
DCBx Port Roles....267
DCB Configuration Exchange....268
Configuration Source Election....269
Propagation of DCB Information....269
Auto-Detection of the DCBx Version....270
DCBx Example....270
DCBx Prerequisites and Restrictions....271
DCBx Error Messages....272
Debugging DCBx on an Interface....272
Verifying the DCB Configuration....273
QoS dot1p Traffic Classification and Queue Assignment....283
Troubleshooting PFC, ETS, and DCBx Operation....283

14 Dynamic Host Configuration Protocol (DHCP)....286
Supported Modes....286
Assigning an IP Address using DHCP....286
Debugging DHCP Client Operation....288
DHCP Client....291
How DHCP Client is Implemented....291

DHCP Client on a Management Interface....292
DHCP Client on a VLAN....292
DHCP Packet Format and Options....292
Option 82....294
Releasing and Renewing DHCP-based IP Addresses....295
Viewing DHCP Statistics and Lease Information....295

15 Equal Cost Multi-Path (ECMP)....297
ECMP for Flow-Based Affinity....297
Enabling Deterministic ECMP Next Hop....297
Link Bundle Monitoring....297
Managing ECMP Group Paths....298

16 FC FPORT....299
FC FPORT....299
Configuring Switch Mode to FCF Port Mode....300
Name Server....300
FCoE Maps....301
Creating an FCoE Map....302
Zoning....303
Creating Zone and Adding Members....304
Creating Zone Alias and Adding Members....304
Creating Zonesets....304
Activating a Zoneset....305
Displaying the Fabric Parameters....305

17 FCoE Transit....309
Supported Modes....309
Fibre Channel over Ethernet....309
Ensure Robustness in a Converged Ethernet Network....309
FIP Snooping on Ethernet Bridges....311
FIP Snooping in a Switch Stack....313
Using FIP Snooping....313
Important Points to Remember....313
Enabling the FCoE Transit Feature....314
Enable FIP Snooping on VLANs....314
Configure the FC-MAP Value....314
Configure a Port for a Bridge-to-Bridge Link....314
Configure a Port for a Bridge-to-FCF Link....315
Impact on Other Software Features....315
FIP Snooping Prerequisites....315
FIP Snooping Restrictions....316

Configuring FIP Snooping....316
Displaying FIP Snooping Information....317
FCoE Transit Configuration Example....322

18 FIPS Cryptography....325
Preparing the System....325
Enabling FIPS Mode....325
Generating Host-Keys....326
Monitoring FIPS Mode Status....326
Disabling FIPS Mode....327

19 Force10 Resilient Ring Protocol (FRRP)....328
Protocol Overview....328
Ring Status....329
Multiple FRRP Rings....330
Important FRRP Points....331
Important FRRP Concepts....332
Implementing FRRP....333
FRRP Configuration....333
Creating the FRRP Group....334
Configuring the Control VLAN....334
Configuring and Adding the Member VLANs....335
Setting the FRRP Timers....336
Clearing the FRRP Counters....337
Viewing the FRRP Configuration....337
Viewing the FRRP Information....337
Troubleshooting FRRP....338
Configuration Checks....338
Sample Configuration and Topology....338

20 GARP VLAN Registration Protocol (GVRP)....341
Important Points to Remember....341
Configure GVRP....341
Related Configuration Tasks....342
Enabling GVRP Globally....342
Enabling GVRP on a Layer 2 Interface....343
Configure GVRP Registration....343
Configure a GARP Timer....344

21 FIP Snooping....345
Supported Modes....345
Fibre Channel over Ethernet....345

Ensuring Robustness in a Converged Ethernet Network....345
FIP Snooping on Ethernet Bridges....347
How FIP Snooping is Implemented....348
FIP Snooping on VLANs....349
FC-MAP Value....349
Bridge-to-FCF Links....349
Impact on other Software Features....349
FIP Snooping Prerequisites....349
FIP Snooping Restrictions....350
Configuring FIP Snooping....350
Displaying FIP Snooping Information....351
FIP Snooping Example....357
Debugging FIP Snooping....357

22 Internet Group Management Protocol (IGMP)....359
IGMP Overview....359
IGMP Version 2....359
Joining a Multicast Group....360
Leaving a Multicast Group....360
IGMP Version 3....360
Joining and Filtering Groups and Sources....361
Leaving and Staying in Groups....362
IGMP Snooping....363
How IGMP Snooping is Implemented on an Aggregator....363
Disabling Multicast Flooding....364
Displaying IGMP Information....364

23 Interfaces....366
Basic Interface Configuration....366
Advanced Interface Configuration....366
Interface Types....367
View Basic Interface Information....367
Configuring the Default Interface....369
Enabling a Physical Interface....370
Physical Interfaces....370
Configuration Task List for Physical Interfaces....370
Overview of Layer Modes....371
Configuring Layer 2 (Data Link) Mode....371
Configuring Layer 2 (Interface) Mode....372
Configuring Layer 3 (Network) Mode....372
Configuring Layer 3 (Interface) Mode....373
Management Interfaces....374

Configuring Management Interfaces on the MXL Switch 374
VLAN Interfaces 375
Loopback Interfaces 376
Null Interfaces 377
VLAN Membership 377
Default VLAN 378
Port-Based VLANs 378
VLANs and Port Tagging 378
Configuring VLAN Membership 379
Displaying VLAN Membership 379
Adding an Interface to a Tagged VLAN 380
Adding an Interface to an Untagged VLAN 380
VLAN Configuration on Physical Ports and Port-Channels 381
Port Channel Interfaces 382
Port Channel Definition and Standards 383
Port Channel Benefits 383
Port Channel Implementation 383
100/1000/10000 Mbps Interfaces in Port Channels 384
Configuration Tasks for Port Channel Interfaces 384
Creating a Port Channel 384
Adding a Physical Interface to a Port Channel 385
Reassigning an Interface to a New Port Channel 387
Configuring the Minimum Oper Up Links in a Port Channel 387
Adding or Removing a Port Channel from a VLAN 388
Assigning an IP Address to a Port Channel 388
Deleting or Disabling a Port Channel 388
Load Balancing through Port Channels 389
Changing the Hash Algorithm 389
Server Ports 390
Default Configuration without Start-up Config 390
Bulk Configuration 391
Interface Range 391
Bulk Configuration Examples 392
Defining Interface Range Macros 393
Define the Interface Range 393
Choosing an Interface-Range Macro 394
Monitoring and Maintaining Interfaces 394
Maintenance Using TDR 395
Splitting QSFP Ports to SFP+ Ports 396
Merging SFP+ Ports to QSFP 40G Ports 396
Configure the MTU Size on an Interface 397
Converting a QSFP or QSFP+ Port to an SFP or SFP+ Port 397

Important Points to Remember ... 398
Example Scenarios ... 398
Layer 2 Flow Control Using Ethernet Pause Frames ... 399
Enabling Pause Frames ... 399
Configure MTU Size on an Interface ... 400
Port-Pipes ... 401
Auto-Negotiation on Ethernet Interfaces ... 401
Setting the Speed and Duplex Mode of Ethernet Interfaces ... 401
View Advanced Interface Information ... 403
Configuring the Interface Sampling Size ... 404
Dynamic Counters ... 405
Enhanced Control of Remote Fault Indication Processing ... 406

24 Internet Protocol Security (IPSec) ... 408
Configuring IPSec ... 408

25 IPv4 Routing ... 410
IP Addresses ... 410
Implementation Information ... 410
Configuration Tasks for IP Addresses ... 411
IPv4 Path MTU Discovery Overview ... 413
Using the Configured Source IP Address in ICMP Messages ... 414
Configuring the ICMP Source Interface ... 414
Configuring the Duration to Establish a TCP Connection ... 415
Enabling Directed Broadcast ... 415
Resolution of Host Names ... 416
Enabling Dynamic Resolution of Host Names ... 416
Specifying the Local System Domain and a List of Domains ... 416
Configuring DNS with Traceroute ... 417
ARP ... 418
Configuration Tasks for ARP ... 418
ARP Learning via Gratuitous ARP ... 420
ARP Learning via ARP Request ... 420
Configuring ARP Retries ... 421
ICMP ... 422
Configuration Tasks for ICMP ... 422
UDP Helper ... 422
Configure UDP Helper ... 422
Important Points to Remember ... 422
Enabling UDP Helper ... 423
Configurations Using UDP Helper ... 423
UDP Helper with Broadcast-All Addresses ... 423

UDP Helper with Subnet Broadcast Addresses 424
UDP Helper with Configured Broadcast Addresses 425
UDP Helper with No Configured Broadcast Addresses 425
Troubleshooting UDP Helper 426

26 IPv6 Routing 427
Protocol Overview 427
Extended Address Space 427
Stateless Autoconfiguration 427
IPv6 Header Fields 429
Version (4 bits) 429
Traffic Class (8 bits) 429
Flow Label (20 bits) 429
Payload Length (16 bits) 429
Next Header (8 bits) 429
Hop Limit (8 bits) 430
Source Address (128 bits) 430
Destination Address (128 bits) 430
Extension Header Fields 430
Hop-by-Hop Options Header 431
Addressing 431
Link-local Addresses 432
Static and Dynamic Addressing 432
Implementing IPv6 with the Dell Networking OS 433
ICMPv6 435
Path MTU Discovery 436
IPv6 Neighbor Discovery 436
IPv6 Neighbor Discovery of MTU Packets 437
Configuration Task List for IPv6 RDNSS 437
Configuring the IPv6 Recursive DNS Server 437
Displaying IPv6 RDNSS Information 438
Debugging IPv6 RDNSS Information Sent to the Host 438
IPv6 Multicast 439
Configuration Task List for IPv6 439
Adjusting Your CAM-Profile 439
Assigning an IPv6 Address to an Interface 440
Assigning a Static IPv6 Route 441
Configuring Telnet with IPv6 441
SNMP over IPv6 442
Showing IPv6 Information 442
Showing an IPv6 Interface 442
Showing IPv6 Routes 443

Showing the Running-Configuration for an Interface....444
Clearing IPv6 Routes....444
Secure Shell (SSH) Over an IPv6 Transport....445

27 iSCSI Optimization....446
Supported Modes....446
iSCSI Optimization Overview....446
Monitoring iSCSI Traffic Flows....448
Information Monitored in iSCSI Traffic Flows....448
Synchronizing iSCSI Sessions Learned on VLT-Lags with VLT-Peer....448
iSCSI Optimization: Operation....448
Configuring iSCSI Optimization....449
Displaying iSCSI Optimization Information....451

28 Intermediate System to Intermediate System....453
IS-IS Protocol Overview....453
IS-IS Addressing....453
Multi-Topology IS-IS....454
Transition Mode....455
Interface Support....455
Adjacencies....455
Graceful Restart....455
Timers....456
Implementation Information....456
Configuration Information....457
Configuration Tasks for IS-IS....457
IS-IS Metric Styles....473
Configure Metric Values....473
Maximum Values in the Routing Table....473
Change the IS-IS Metric Style in One Level Only....473
Leaks from One Level to Another....475
Sample Configurations....476

29 Isolated Networks for Aggregators....481
Configuring and Verifying Isolated Network Settings....481

30 Link Aggregation....482
Supported Modes....482
How the LACP is Implemented on an Aggregator....482
Uplink LAG....483
Server-Facing LAGs....483
LACP Modes....483

Auto-Configured LACP Timeout....483
Link Aggregation Control Protocol (LACP)....484
Configuration Tasks for Port Channel Interfaces....484
Creating a Port Channel....484
Adding a Physical Interface to a Port Channel....484
Reassigning an Interface to a New Port Channel....487
Configuring the Minimum Oper Up Links in a Port Channel....487
Configuring VLAN Tags for Member Interfaces....488
Deleting or Disabling a Port Channel....488
Configuring Auto LAG....489
Configuring the Minimum Number of Links to be Up for Uplink LAGs to be Active....491
Optimizing Traffic Disruption Over LAG Interfaces On IOA Switches in VLT Mode....492
Preserving LAG and Port Channel Settings in Nonvolatile Storage....492
Enabling the LACP link fallback member....492
Enabling the Verification of Member Links Utilization in a LAG Bundle....493
Monitoring the Member Links of a LAG Bundle....493
Verifying LACP Operation and LAG Configuration....494
Multiple Uplink LAGs with 10G Member Ports....496

31 Layer 2....499
Supported Modes....499
Managing the MAC Address Table....499
Clearing the MAC Address Entries....499
Displaying the MAC Address Table....500
Network Interface Controller (NIC) Teaming....500
MAC Address Station Move....501
MAC Move Optimization....502

32 Link Layer Discovery Protocol (LLDP)....503
Supported Modes....503
Protocol Data Units....503
Configure LLDP....505
Related Configuration Tasks....505
Important Points to Remember....505
CONFIGURATION versus INTERFACE Configurations....505
Enabling LLDP....506
Disabling and Undoing LLDP....506
Advertising TLVs....506
Optional TLVs....508
Management TLVs....508
IEEE Organizationally Specific TLVs....508
LLDP-MED Capabilities TLV....510

LLDP-MED Network Policies TLV....511
Extended Power via MDI TLV....512
LLDP Operation....513
Viewing the LLDP Configuration....513
Viewing Information Advertised by Adjacent LLDP Agents....514
Configuring LLDPDU Intervals....515
Configuring a Time to Live....515
Clearing LLDP Counters....516
Debugging LLDP....516
Relevant Management Objects....517

33 Microsoft Network Load Balancing....523
NLB Unicast Mode Scenario....523
NLB Multicast Mode Scenario....524
Limitations With Enabling NLB on Switches....524
Benefits and Working of Microsoft Clustering....524
Enable and Disable VLAN Flooding....525
Configuring a Switch for NLB....525
Multicast NLB Mode....525

34 Multicast Source Discovery Protocol (MSDP)....526
Protocol Overview....526
Anycast RP....528
Implementation Information....528
Configure the Multicast Source Discovery Protocol....528
Related Configuration Tasks....529
Enabling MSDP....533
Manage the Source-Active Cache....534
Viewing the Source-Active Cache....534
Limiting the Source-Active Cache....535
Clearing the Source-Active Cache....535
Enabling the Rejected Source-Active Cache....535
Accept Source-Active Messages that Fail the RFP Check....535
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....542
Logging Changes in Peership States....542
Terminating a Peership....543
Clearing Peer Statistics....543
Debugging MSDP....544

MSDP with Anycast RP..... 544
Configuring Anycast RP.....545
Reducing Source-Active Message Flooding.....546
Specifying the RP Address Used in SA Messages.....546
MSDP Sample Configurations..... 549

35 Multiple Spanning Tree Protocol (MSTP).....552
Protocol Overview..... 552
Spanning Tree Variations.....553
Implementation Information..... 553
Configure Multiple Spanning Tree Protocol.....553
Related Configuration Tasks.....553
Enable Multiple Spanning Tree Globally..... 554
Creating Multiple Spanning Tree Instances..... 554
Influencing MSTP Root Selection.....555
Interoperate with Non-Dell Networking OS Bridges..... 556
Changing the Region Name or Revision.....556
Modifying Global Parameters..... 557
Enable BPDU Filtering Globally..... 558
Modifying the Interface Parameters.....559
Configuring an EdgePort.....560
Flush MAC Addresses after a Topology Change.....561
MSTP Sample Configurations..... 561
Router 1 Running-ConfigurationRouter 2 Running-ConfigurationRouter 3 Running-ConfigurationSFTOS Example Running-Configuration.....561
Debugging and Verifying MSTP Configurations.....564

36 Multicast Features.....567
Enabling IP Multicast.....567
Implementation Information..... 567
First Packet Forwarding for Lossless Multicast.....568
Multicast Policies..... 568
IPv4 Multicast Policies..... 568
Limiting the Number of Multicast Routes.....568
Preventing a Host from Joining a Group.....569
Rate Limiting IGMP Join Requests.....572
Preventing a PIM Router from Forming an Adjacency..... 572
Preventing a Source from Registering with the RP.....572
Preventing a PIM Router from Processing a Join..... 575

37 Open Shortest Path First (OSPFv2 and OSPFv3).....576
Protocol Overview..... 576

Autonomous System (AS) Areas 576
Area Types 577
Networks and Neighbors 578
Router Types 578
Link-State Advertisements (LSAs) 580
Router Priority and Cost 581
OSPF with the Dell Networking OS 582
Graceful Restart 583
Fast Convergence (OSPFv2, IPv4 Only) 584
Processing SNMP and Sending SNMP Traps 584
OSPF ACK Packing 584
Setting OSPF Adjacency with Cisco Routers 584
Configuration Information 585
Configuration Task List for OSPFv2 (OSPF for IPv4) 585
Troubleshooting OSPFv2 598
Configuration Task List for OSPFv3 (OSPF for IPv6) 601
Enabling IPv6 Unicast Routing 602
Assigning IPv6 Addresses on an Interface 602
Assigning Area ID on an Interface 603
Assigning OSPFv3 Process ID and Router ID Globally 603
Configuring Stub Areas 604
Configuring Passive-Interface 604
Redistributing Routes 604
Configuring a Default Route 605
Enabling OSPFv3 Graceful Restart 605
Displaying Graceful Restart 606
OSPFv3 Authentication Using IPsec 607

38 Policy-based Routing (PBR) 617
Overview 617
Implementing Policy-based Routing with Dell Networking OS 619
Configuration Task List for Policy-based Routing 619
PBR Exceptions (Permit) 622
Sample Configuration 624
Create the Redirect-List GOLDAssign Redirect-List GOLD to Interface 2/11View Redirect-List GOLD 625

39 PIM Sparse-Mode (PIM-SM) 629
Implementation Information 629
Protocol Overview 629
Requesting Multicast Traffic 629
Refuse Multicast Traffic 630

Send Multicast Traffic....630
Configuring PIM-SM....631
Related Configuration Tasks....631
Enable PIM-SM....631
Configuring S,G Expiry Timers....632
Configuring a Static Rendezvous Point....633
Overriding Bootstrap Router Updates....634
Configuring a Designated Router....634
Creating Multicast Boundaries and Domains....635
Enabling PIM-SM Graceful Restart....635

40 PIM Source-Specific Mode (PIM-SSM)....636
Configure PIM-SMM....636
Related Configuration Tasks....636
Implementation Information....636
Important Points to Remember....637
Enabling PIM-SSM....637
Use PIM-SSM with IGMP Version 2 Hosts....637
Configuring PIM-SSM with IGMPv2....638

41 Port Monitoring....640
Supported Modes....640
Configuring Port Monitoring....640
Important Points to Remember....641
Port Monitoring....642

42 Private VLANs (PVLAN)....644
Private VLAN Concepts....644
Using the Private VLAN Commands....645
Configuration Task List....646
Private VLAN Configuration Example....650

43 Per-VLAN Spanning Tree Plus (PVST+)....654
Protocol Overview....654
Implementation Information....655
Configure Per-VLAN Spanning Tree Plus....655
Related Configuration Tasks....655
Enabling PVST+....655
Disabling PVST+....656
Influencing PVST+ Root Selection....656
Modifying Global PVST+ Parameters....658
Modifying Interface PVST+ Parameters....659

Configuring an EdgePort.... 660
PVST+ in Multi-Vendor Networks.... 661
Enabling PVST+ Extend System ID.... 661
PVST+ Sample Configurations.... 662
Enable BPDU Filtering globally.... 663

44 Quality of Service (QoS).... 665
Implementation Information.... 667
Port-Based QoS Configurations.... 667
Setting dot1p Priorities for Incoming Traffic.... 667
Honoring dot1p Priorities on Ingress Traffic.... 668
Configuring Port-Based Rate Policing.... 669
Configuring Port-Based Rate Shaping.... 669
Guidelines for Configuring ECN for Classifying and Color-Marking Packets.... 669
Sample configuration to mark non-ecn packets as "yellow" with Multiple traffic class.... 670
Classifying Incoming Packets Using ECN and Color-Marking.... 670
Sample configuration to mark non-ecn packets as "yellow" with single traffic class.... 673
Policy-Based QoS Configurations.... 674
DSCP Color Maps.... 675
Classify Traffic.... 677
Create a QoS Policy.... 681
Create Policy Maps.... 684
Enabling QoS Rate Adjustment.... 688
Enabling Strict-Priority Queueing.... 689
Weighted Random Early Detection.... 689
Creating WRED Profiles.... 690
Applying a WRED Profile to Traffic.... 690
Displaying Default and Configured WRED Profiles.... 691
Displaying WRED Drop Statistics.... 691
Displaying egress-queue Statistics.... 691
Classifying Layer 2 Traffic on Layer 3 Interfaces.... 692
Classifying Packets Based on a Combination of DSCP Code Points and VLAN IDs.... 693

45 Routing Information Protocol (RIP).... 695
Protocol Overview.... 695
RIPv1.... 695
RIPv2.... 695
Implementation Information.... 696
Configuration Information.... 696
Configuration Task List.... 696
RIP Configuration Example.... 703

46 Remote Monitoring (RMON)....708
Implementation Information....708
Fault Recovery....708
Setting the rmon Alarm....709
Configuring an RMON Event....710
Configuring RMON Collection Statistics....710
Configuring the RMON Collection History....711
Enabling an RMON MIB Collection History Group....712

47 Rapid Spanning Tree Protocol (RSTP)....713
Protocol Overview....713
Configuring Rapid Spanning Tree....713
Related Configuration Tasks....713
Important Points to Remember....714
Configuring Interfaces for Layer 2 Mode....714
Enabling Rapid Spanning Tree Protocol Globally....714
Adding and Removing Interfaces....717
Modifying Global Parameters....717
Enable BPDU Filtering Globally....719
Modifying Interface Parameters....719
Configuring an EdgePort....720
Influencing RSTP Root Selection....721
SNMP Traps for Root Elections and Topology Changes....721
Configuring Fast Hellos for Link State Detection....721

48 Security....723
Supported Modes....723
Understanding Banner Settings....723
Accessing the I/O Aggregator Using the CMC Console Only....723
AAA Accounting....724
Configuration Task List for AAA Accounting....724
AAA Authentication....726
Configuration Task List for AAA Authentication....726
AAA Authorization....729
Privilege Levels Overview....729
Configuration Task List for Privilege Levels....730
RADIUS....734
RADIUS Authentication and Authorization....734
Configuration Task List for RADIUS....735
TACACS+....738
Configuration Task List for TACACS+....738

TACACS+ Remote Authentication....740
Enabling SCP and SSH....741
Using SCP with SSH to Copy a Software Image....742
Secure Shell Authentication....743
Telnet....743
VTY Line and Access-Class Configuration....744
VTY Line Local Authentication and Authorization....744
VTY Line Remote Authentication and Authorization....745

49 Service Provider Bridging....746
VLAN Stacking....746
Important Points to Remember....747
Configure VLAN Stacking....747
Creating Access and Trunk Ports....748
Enable VLAN-Stacking for a VLAN....749
Configuring the Protocol Type Value for the Outer VLAN Tag....749
Configuring Options for Trunk Ports....749
Debugging VLAN Stacking....750
VLAN Stacking in Multi-Vendor Networks....751
VLAN Stacking Packet Drop Precedence....755
Enabling Drop Eligibility....755
Honoring the Incoming DEI Value....756
Marking Egress Packets with a DEI Value....756
Dynamic Mode CoS for VLAN Stacking....757
Mapping C-Tag to S-Tag dot1p Values....758
Layer 2 Protocol Tunneling....759
Implementation Information....761
Enabling Layer 2 Protocol Tunneling....762
Specifying a Destination MAC Address for BPDUs....762
Setting Rate-Limit BPDUs....762
Debugging Layer 2 Protocol Tunneling....763
Provider Backbone Bridging....763

50 sFlow....764
Overview....764
Implementation Information....764
Important Points to Remember....765
Enabling and Disabling sFlow....765
Enabling and Disabling sFlow on an Interface....765
Enabling sFlow Max-Header Size Extended....765
sFlow Show Commands....767
Displaying Show sFlow Global....767

Displaying Show sFlow on an Interface ... 767
Displaying Show sFlow on a Stack Unit ... 768
Configuring Specify Collectors ... 768
Changing the Polling Intervals ... 768
Changing the Sampling Rate ... 769
Sub-Sampling ... 769
Back-Off Mechanism ... 770
sFlow on LAG ports ... 770
Enabling Extended sFlow ... 770

51 Simple Network Management Protocol (SNMP) ... 772
Supported Modes ... 772
Implementation Information ... 772
Configuring the Simple Network Management Protocol ... 772
Important Points to Remember ... 772
Setting up SNMP ... 773
Creating a Community ... 773
Reading Managed Object Values ... 773
Displaying the Ports in a VLAN using SNMP ... 774
Fetching Dynamic MAC Entries using SNMP ... 776
Deriving Interface Indices ... 777
Monitor Port-Channels ... 778
Entity MIBS ... 779
Example of Sample Entity MIBS outputs ... 779
SNMP Traps for Link Status ... 780
Standard VLAN MIB ... 780
Enhancements ... 780
Fetching the Switchport Configuration and the Logical Interface Configuration ... 780
MIB Support to Display the Available Memory Size on Flash ... 782
Viewing the Available Flash Memory Size ... 782
MIB Support to Display the Software Core Files Generated by the System ... 782
Viewing the Software Core Files Generated by the System ... 783

52 Stacking ... 784
Supported Modes ... 784
Configuring a Switch Stack ... 784
Stacking Prerequisites ... 784
Master Selection Criteria ... 785
Configuring Priority and stack-group ... 785
Cabling the Switch Stack ... 786
Accessing the CLI ... 786
Configuring and Bringing Up a Stack ... 786

Adding a Stack Unit....787
Resetting a Unit on a Stack....788
Removing an Aggregator from a Stack....788
Merging Two Operational Stacks....789
Verifying a Stack Configuration....789
Using Show Commands....789
Troubleshooting a Switch Stack....790
Failure Scenarios....790
Upgrading a Switch Stack....792
Upgrading a Single Stack Unit....793

53 Storm Control....795
Configure Storm Control....795
Configuring Storm Control from INTERFACE Mode....796
Configuring Storm Control from CONFIGURATION Mode....796

54 Broadcast Storm Control....797
Supported Modes....797
Disabling Broadcast Storm Control....797
Displaying Broadcast-Storm Control Status....797
Configuring Storm Control....797

55 Spanning Tree Protocol (STP)....798
Protocol Overview....798
Configure Spanning Tree....798
Related Configuration Tasks....798
Important Points to Remember....799
Configuring Interfaces for Layer 2 Mode....799
Enabling Spanning Tree Protocol Globally....800
Adding an Interface to the Spanning Tree Group....803
Removing an Interface from the Spanning Tree Group....803
Modifying Global Parameters....803
Modifying Interface STP Parameters....804
Enabling Port Fast....805
Prevent Network Disruptions with BPDU Guard....805
Global BPDU Filtering....808
Interface BPDU Filtering....808
Selecting STP Root....809
STP Root Guard....810
Root Guard Scenario....810
Configuring Root Guard....811
SNMP Traps for Root Elections and Topology Changes....812

Displaying STP Guard Configuration....812

56 System Time and Date....813

Supported Modes....813
Setting the Time for the Software Clock....813
Setting the Timezone....813
Setting Daylight Savings Time....814
Setting Daylight Saving Time Once....814
Setting Recurring Daylight Saving Time....815

57 Tunneling....817

Configuring a Tunnel....817
Configuring Tunnel keepalive....818
Configuring the ip and ipv6 unnumbered....819
Configuring the Tunnel allow-remote....819
Configuring the Tunnel Source Anylocal....820

58 Uplink Failure Detection (UFD)....821

Supported Modes....821
Feature Description....821
How Uplink Failure Detection Works....822
UFD and NIC Teaming....824
Important Points to Remember....824
Uplink Failure Detection (SMUX mode)....825
Configuring Uplink Failure Detection (PMUX mode)....825
Clearing a UFD-Disabled Interface (in PMUX mode)....827
Displaying Uplink Failure Detection....829
Sample Configuration: Uplink Failure Detection....830

59 PMUX Mode of the IO Aggregator....832

I/O Aggregator (IOA) Programmable MUX (PMUX) Mode....832
Configuring and Changing to PMUX Mode....832
Configuring the Commands without a Separate User Account....833
Virtual Link Trunking (VLT)....833
Overview....833
Setting up VLT....834
VLT Terminology....835
Configure Virtual Link Trunking....835
Verifying a VLT Configuration....840
VLT Sample Configurations....843
Troubleshooting VLT....845

60 NPIV Proxy Gateway 847
NPIV Proxy Gateway Configuration 847
NPIV Proxy Gateway Operations and Capabilities 847
NPIV Proxy Gateway Operation 847
NPIV Proxy Gateway: Protocol Services 848
NPIV Proxy Gateway Functionality 848
NPIV Proxy Gateway: Terms and Definitions 849
Configuring an NPIV Proxy Gateway 851
Enabling Fibre Channel Capability on the Switch 852
Creating a DCB Map 852
Applying a DCB Map on Server-facing Ethernet Ports 853
Creating an FCoE VLAN 854
Creating an FCoE Map 854
Applying an FCoE Map on Server-facing Ethernet Ports 855
Applying an FCoE Map on Fabric-facing FC Ports 856
Sample Configuration 857
Displaying NPIV Proxy Gateway Information 858
show interfaces status Command Example 858
show fcoe-map Command Examples 859
show qos dcb-map Command Examples 860
show npiv devices brief Command Example 861
show npiv devices Command Example 862
show fc switch Command Example 863

61 Upgrade Procedures 864
Get Help with Upgrades 864

62 Virtual LANs (VLANs) 865
Default VLAN 865
Port-Based VLANs 866
VLANs and Port Tagging 867
Configuration Task List 867
Configuring Native VLANs 871
Enabling Null VLAN as the Default VLAN 872

63 Virtual Link Trunking (VLT) 873
Overview 873
Multi-domain VLT 874
VLT Terminology 875
Configure Virtual Link Trunking 875
Important Points to Remember 875

Configuration Notes ... 876
RSTP and VLT ... 880
VLT Bandwidth Monitoring ... 880
VLT and IGMP Snooping ... 881
VLT Port Delayed Restoration ... 881
PIM-Sparse Mode Support on VLT ... 881
VLT Multicast ... 883
VLT Unicast Routing ... 884
Non-VLT ARP Sync ... 885
RSTP Configuration ... 885
Preventing Forwarding Loops in a VLT Domain ... 886
Sample RSTP Configuration ... 886
Configuring VLT ... 887
Configuring a VLT Interconnect ... 887
Configuring a VLT Backup Link ... 888
Configuring a VLT Port Delay Period ... 888
Reconfiguring the Default VLT Settings (Optional) ... 888
Connecting a VLT Domain to an Attached Access Device (Switch or Server) ... 889
Configuring a VLT VLAN Peer-Down (Optional) ... 890
Configure Multi-domain VLT (mVLT) (Optional) ... 891
Verifying a VLT Configuration ... 893
Connecting a VLT Domain ... 897
PVST+ Configuration ... 901
Sample PVST+ Configuration ... 901
mVLT Configuration Example ... 902
In Domain 1, configure the VLT domain and VLTi on Peer 1Configure mVLT on Peer 1Add links to the mVLT port-channel on Peer 1Next, configure the VLT domain and VLTi on Peer 2Configure mVLT on Peer 2Add links to the mVLT port-channel on Peer 2In Domain 2, configure the VLT domain and VLTi on Peer 3Configure mVLT on Peer 3Add links to the mVLT port-channel on Peer 3Configure the VLT domain and VLTi on Peer 4Configure mVLT on Peer 4Add links to the mVLT port-channel on Peer 4 ... 903
PIM-Sparse Mode Configuration Example ... 905
Additional VLT Sample Configurations ... 905
Configuring Virtual Link Trunking (VLT Peer 1)Configuring Virtual Link Trunking (VLT Peer 2)Verifying a Port-Channel Connection to a VLT Domain (From an Attached Access Switch) ... 905
Troubleshooting VLT ... 907
Specifying VLT Nodes in a PVLAN ... 909
Association of VLTi as a Member of a PVLAN ... 910
MAC Synchronization for VLT Nodes in a PVLAN ... 910
PVLAN Operations When One VLT Peer is Down ... 911
PVLAN Operations When a VLT Peer is Restarted ... 911

Interoperation of VLT Nodes in a PVLAN with ARP Requests....911
Scenarios for VLAN Membership and MAC Synchronization With VLT Nodes in PVLAN....912
Configuring a VLT VLAN or LAG in a PVLAN....913
Creating a VLT LAG or a VLT VLAN....913
Associating the VLT LAG or VLT VLAN in a PVLAN....914
Proxy ARP Capability on VLT Peer Nodes....915
Working of Proxy ARP for VLT Peer Nodes....916
Configuring VLAN-Stack over VLT....917

64 Virtual Router Redundancy Protocol (VRRP)....920
VRRP Overview....920
VRRP Benefits....921
VRRP Implementation....921
VRRP Configuration....922
Configuration Task List....922
Setting VRRP Initialization Delay....931
Sample Configurations....932
VRRP for an IPv4 Configuration....932

65 Debugging and Diagnostics....934
Supported Modes....934
Debugging Aggregator Operation....934
All interfaces on the Aggregator are operationally down....934
Broadcast, unknown multicast, and DLF packets switched at a very low rate....935
Flooded packets on all VLANs are received on a server....935
Software show Commands....936
Offline Diagnostics....937
Important Points to Remember....937
Running Offline Diagnostics....937
Trace Logs....938
Auto Save on Crash or Rollover....939
Using the Show Hardware Commands....939
Environmental Monitoring....941
Recognize an Over-Temperature Condition....941
Troubleshoot an Over-Temperature Condition....942
Recognize an Under-Voltage Condition....942
Troubleshoot an Under-Voltage Condition....943
Buffer Tuning....943
Deciding to Tune Buffers....945
Sample Buffer Profile Configuration....948
Troubleshooting Packet Loss....949
Displaying Drop Counters....949

Dataplane Statistics..... 951
Displaying Stack Port Statistics..... 952
Enabling Buffer Statistics Tracking 952
Restoring the Factory Default Settings..... 954
Important Points to Remember..... 954

66 Standards Compliance..... 956
IEEE Compliance..... 956
RFC and I-D Compliance..... 956
General Internet Protocols..... 957
General IPv4 Protocols..... 957
Network Management..... 958
MIB Location..... 960

67 FC Flex IO Modules..... 962
FC Flex IO Modules..... 962
Understanding and Working of the FC Flex IO Modules..... 962
FC Flex IO Modules Overview..... 962
FC Flex IO Module Capabilities and Operations..... 963
Guidelines for Working with FC Flex IO Modules..... 964
Processing of Data Traffic..... 966
Installing and Configuring the Switch..... 966
Interconnectivity of FC Flex IO Modules with Cisco MDS Switches..... 969
Data Center Bridging (DCB)..... 970
Ethernet Enhancements in Data Center Bridging..... 970
Enabling Data Center Bridging..... 978
QoS dot1p Traffic Classification and Queue Assignment..... 979
Configure Enhanced Transmission Selection..... 980
Configure a DCBx Operation..... 982
Verifying the DCB Configuration..... 992
PFC and ETS Configuration Examples..... 1001
Using PFC and ETS to Manage Data Center Traffic..... 1001
Fibre Channel over Ethernet for FC Flex IO Modules..... 1004
NPIV Proxy Gateway for FC Flex IO Modules..... 1005
NPIV Proxy Gateway Configuration on FC Flex IO Modules 1005
NPIV Proxy Gateway Operations and Capabilities..... 1005
Configuring an NPIV Proxy Gateway..... 1009
Displaying NPIV Proxy Gateway Information..... 1015

1

About this Guide

This guide describes the supported protocols and software features, and provides configuration instructions and examples, for the Dell Networking Operating System (OS).

Dell Networking FN IOM is available with running Dell Networking OS version 9.9(0.0).

The FN IOM is installed in a Dell PowerEdge FX2 server chassis. For information about how to install and perform the initial switch configuration, refer to the *Getting Started Guides* on the Dell Support website at http://www.dell.com/support/manuals

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 Networking systems. For complete information about protocols, refer to other documentation, including IETF requests for comment (RFCs). The instructions in this guide cite relevant RFCs, and Standards Compliance contains a complete list of the supported RFCs and management information base files (MIBs).

NOTE: You can perform some of the configuration tasks described in this document by using either the Dell command line or the chassis management controller (CMC) graphical interface. Tasks supported by the CMC interface are shown with the CMC icon: CMC

Audience

This document is intended for system administrators who are responsible for configuring and maintaining networks and assumes knowledge in Layer 2 and Layer 3 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.

Information Symbols

This book uses the following information symbols.

NOTE: The Note icon signals important operational information.

CAUTION: The Caution icon signals information about situations that could result in equipment damage or loss of data.

WARNING: The Warning icon signals information about hardware handling that could result in injury.

* (Exception). This symbol is a note associated with additional text on the page that is marked with an asterisk.

Related Documents

For more information about the Dell PowerEdge FN I/O Aggregator, refer to the following documents:

- *Dell PowerEdge FN I/O Aggregator Command Line Reference Guide*
- *Dell PowerEdge FN I/O Aggregator Getting Started Guide*
- *Release Notes for the Dell PowerEdge FN I/O Aggregator*

2

Before You Start

To install the FN IOM in a Dell FX2 server chassis, use the instructions in the *Dell Networking FN IOM Getting Started Guide* that is shipped with the product. The FN IOM installs with zero-touch configuration. After you power it on, it boots up with default settings and auto-configures with software features enabled. This topic describes the default settings and software features that are automatically configured at startup. To reconfigure the FN IOM for customized network operation, use the tasks described in the other topics.

Operational Modes

FN IOM supports five operational modes. Select the operational mode that meets your deployment needs. To enable a new operational mode, reload the switch.

Standalone mode

stack-unit *unit* **iom-mode standalone**

CONFIGURATION mode

```
Dell(conf)#stack-unit 0 iom-mode standalone
```

This mode is the default mode for IOA. It is a fully automated zero-touch mode that allows you to configure VLAN memberships. (Supported in CMC)

VLT mode

stack-unit *unit* **iom-mode vlt**

CONFIGURATION mode

```
Dell(conf)#stack-unit 0 iom-mode vlt
```

Select this mode to multihome server interfaces to different IOA modules. This mode is a low-touch mode where all configuration except VLAN membership is automated. To enable VLAN, configure the VLANs at the server port level. In this mode, port 9 link, which is associated with LAG-127, is dedicated to VLT interconnect.

Programmable MUX mode

stack-unit *unit* **iom-mode programmable-mux**

CONFIGURATION mode

```
Dell(conf)#stack-unit 0 iom-mode programmable-mux
```

Select this mode to configure PMUX mode CLI commands.

Standalone mode is the zero-touch auto configuration default mode. If you want the flexibility to configure different settings, change the FN I/O Module to PMUX mode. PMUX mode provides additional CLI commands to customize the software configuration, as needed.

You can configure any of the external Ethernet ports to operate as stack links.

For more information on the PMUX mode, see PMUX Mode of the IO Aggregator.

Stacking mode

stack-unit *unit* **iom-mode stack**

CONFIGURATION mode

```
Dell(conf)#stack-unit 0 iom-mode stack
```

Select this mode to configure Stacking mode CLI commands.

The lower two external Ethernet ports (ports 9 and 10) operate as stack links.

For more information on the Stacking mode, see .

Full Switch mode

stack-unit *unit* **iom-mode full-switch**

CONFIGURATION mode

```
Dell(conf)#stack-unit 0 iom-mode full-switch
```

Select this mode to configure full switch CLI commands. Full-Switch mode offers Layer 2 or Layer 3 switching functionalities on the Dell FX2 chassis.

For more information on the Full Switch mode, see *Dell Networking OS Command Line Reference Guide*.

Default Settings

The I/O Aggregator provides zero-touch configuration with the following default configuration settings:

- default user name (**root**)
- password (**calvin**)
- VLAN (vlan1) and IP address for in-band management (**DHCP**)
- IP address for out-of-band (OOB) management (**DHCP**)
- read-only SNMP community name (**public**)
- broadcast storm control (**enabled in Standalone mode and disabled in VLT mode**)
- IGMP multicast flooding (**enabled**)
- VLAN configuration (**in Standalone mode, all ports belong to all VLANs**)

You can change any of these default settings using the CLI. Refer to the appropriate chapter for details.

NOTE: You can also change many of the default settings using the chassis management controller (CMC) interface. For information about how to access the CMC to configure the aggregator, refer to the *Dell Chassis Management Controller (CMC) User's Guide* on the Dell Support website at http://support.dell.com/

Other Auto-Configured Settings

After the Aggregator powers on, it auto-configures and is operational with software features enabled, including:

- Ports: Ports are administratively up and auto-configured to operate as hybrid ports to transmit tagged and untagged VLAN traffic.For more information about how ports are numbered, refer to Port Numbering.
- Link aggregation: All uplink ports are configured in a single LAG (LAG 128).
- VLANs: All ports are configured as members of all (4094) VLANs. All VLANs are up and can send or receive layer 2 traffic. For more information, refer to VLAN Membership.
- Data center bridging capability exchange protocol (DCBx): Server-facing ports auto-configure in auto-downstream port roles; uplink ports auto-configure in auto-upstream port roles.
- Fibre Channel over Ethernet (FCoE) connectivity and FCoE initiation protocol (FIP) snooping: The uplink port channel (LAG 128) is enabled to operate in Fibre channel forwarder (FCF) port mode.
- Link layer discovery protocol (LLDP): Enabled on all ports to advertise management TLV and system name with neighboring devices.
- Internet small computer system interface (iSCSI)optimization.
- Internet group management protocol (IGMP) snooping.
- Jumbo frames: Ports are set to a maximum MTU of 12,000 bytes by default.
- Link tracking: Uplink-state group 1 is automatically configured. In uplink state-group 1, server-facing ports auto-configure as downstream interfaces; the uplink port-channel (LAG 128) auto-configures as an upstream interface. Server-facing links are auto-configured to be brought up only if the uplink port-channel is up.
- In VLT mode, port 9 is automatically configured as VLT interconnect ports. VLT domain configuration is automatic. This includes peer-link, configured MAC, backup link and setting every port channel as VLT port-channel.

Data Center Bridging Support

To eliminate packet loss and provision links with required bandwidth, Data Center Bridging (DCB) enhancements for data center networks are supported.

The aggregator provides zero-touch configuration for DCB. The aggregator auto-configures DCBX port roles as follows:

- Server-facing ports are configured as auto-downstream interfaces.
- Uplink ports are configured as auto-upstream interfaces.

In operation, DCBx auto-configures uplink ports to match the DCB configuration in the ToR switches to which they connect.

The Aggregator supports DCB only in standalone mode.

FCoE Connectivity and FIP Snooping

Many data centers use Fiber Channel (FC) in storage area networks (SANs). Fiber Channel over Ethernet (FCoE) encapsulates Fiber Channel frames over Ethernet networks.

On an Aggregator, the internal ports support FCoE connectivity and connects to the converged network adapter (CNA) in servers. FCoE allows Fiber Channel to use 10-Gigabit Ethernet networks while preserving the Fiber Channel protocol.

The Aggregator also provides zero-touch configuration for FCoE connectivity. The Aggregator auto-configures to match the FCoE settings used in the switches to which it connects through its uplink ports.

FIP snooping is automatically configured on an Aggregator. The auto-configured port channel (LAG 128) operates in FCF port mode.

iSCSI Operation

Support for iSCSI traffic is turned on by default when the Aggregator powers up. No configuration is required.

When an aggregator powers up, it monitors known TCP ports for iSCSI storage devices on all interfaces. When a session is detected, an entry is created and monitored as long as the session is active.

The Aggregator also detects iSCSI storage devices on all interfaces and autoconfigures to optimize performance. Performance optimization operations, such as Jumbo frame size support and disabling storm control on interfaces connected to an iSCSI equallogic (EQL) storage device, are applied automatically.

Link Aggregation

All uplink ports are configured in a single LAG (LAG 128). Server-facing ports are auto-configured as part of link aggregation groups if the corresponding server is configured for LACP-based network interface controller (NIC) teaming. Static LAGs are not supported.

NOTE: The recommended LACP timeout is Long-Timeout mode.

Link Tracking

By default, all server-facing ports are tracked by the operational status of the uplink LAG. If the uplink LAG goes down, the aggregator loses its connectivity and is no longer operational; all server-facing ports are brought down after the specified defer-timer interval, which is 10 seconds by default. If you have configured VLAN, you can reduce the defer time by changing the defer-timer value or remove it by using the `no defer-timer` command.

NOTE: If installed servers do not have connectivity to a switch, check the Link Status LED of uplink ports on the aggregator. If all LEDs are on, to ensure the LACP is correctly configured, check the LACP configuration on the ToR switch that is connected to the aggregator .

Configuring VLANs

By default, in Standalone mode, all aggregator ports belong to all 4094 VLANs and are members of untagged VLAN 1. To configure only the required VLANs on a port, use the CLI or CMC interface.

You can configure VLANs only on server ports. The uplink LAG will automatically get the VLANs, based on the server ports VLAN configuration.

Uplink LAG

The untagged VLAN used for the uplink LAG is always the default VLAN.

Server-Facing LAGs

The tagged VLAN membership of a server-facing LAG is automatically configured based on the server-facing ports that are members of the LAG.

The untagged VLAN of a server-facing LAG is configured based on the untagged VLAN to which the lowest numbered server-facing port in the LAG belongs.

NOTE: Dell Networking recommends configuring the same VLAN membership on all LAG member ports.

Where to Go From Here

You can customize the Aggregator for use in your data center network as necessary. To perform additional switch configuration, do one of the following:

- For remote out-of-band management, enter the OOB management interface IP address into a Telnet or SSH client and log in to the switch using the user ID and password to access the CLI.
- For local management using the CLI, use the attached console connection.
- For remote in-band management from a network management station, enter the IP address of the default VLAN and log in to the switch to access the CLI.

In case of a Dell upgrade, you can check to see that an Aggregator is running the latest Dell version by entering the `show version`command. To download Dell version, go to http://support.dell.com

For detailed information about how to reconfigure specific software settings, refer to the appropriate chapter.

3

Configuration Fundamentals

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

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 Dell Networking OS, after you enable a command, it is entered into 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. For more information, refer to Save the Running-Configuration.

NOTE: You can use the chassis management controller (CMC) out-of-band management interface to access and manage an Aggregator using the Dell Networking OS command-line reference. For more information about how to access the CMC to configure an Aggregator, refer to the Dell Chassis Management Controller (CMC) User's Guide on the Dell Support website at *http://support.dell.com/support/edocs/systems/pem/en/index.htm*.

NOTE: You can also perform some of the configurations using the Web GUI - Dell Blade IO Manager. For more information, see the Dell Blade IO Manager Online Help.

Accessing the Command Line

Access the command line through a serial console port or a Telnet session (Logging into the System using Telnet). When the system successfully boots, enter the command line in EXEC mode.

Logging into the System using Telnet

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

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).

The Dell 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.
- CONFIGURATION mode allows you to configure security features, time settings, set logging and SNMP functions, 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 protocols and IP services specific to an interface. An interface can be physical (10 Gigabit Ethernet) or logical (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
            INTERFACE
              10 GIGABIT ETHERNET
              INTERFACE RANGE
              MANAGEMENT ETHERNET
            LINE
              CONSOLE
              VIRTUAL TERMINAL
                        MONITOR SESSION
```

Navigating CLI Modes

The Dell 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 Command Modes

CLI Command Mode	Prompt	Access Command
EXEC	`Dell>`	Access the router through the console or Telnet.
EXEC Privilege	`Dell#`	• From EXEC mode, enter the `enable` command. • From any other mode, use the `end` command.

CLI Command Mode	Prompt	Access Command
CONFIGURATION	`Dell(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.		
10 Gigabit Ethernet Interface	`Dell(conf-if-te-0/1)#`	`interface` (INTERFACE modes)
Interface Range	`Dell(conf-if-range)#`	`interface` (INTERFACE modes)
Management Ethernet Interface	`Dell(conf-if-ma-0/0)#`	`interface` (INTERFACE modes)
MONITOR SESSION	`Dell(conf-mon-sess)#`	`monitor session`
CONSOLE	`Dell(config-line-console)#`	`line` (LINE Modes)
VIRTUAL TERMINAL	`Dell(config-line-vty)#`	`line` (LINE Modes)

The following example shows how to change the command mode from CONFIGURATION mode to INTERFACE configuration mode.

Example of Changing Command Modes

```
Dell(conf)#interface tengigabitethernet 0/2
Dell(conf-if-te-0/2)#
```

The do Command

You can enter an EXEC mode command from any CONFIGURATION mode (CONFIGURATION, INTERFACE, 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.

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

Stack MAC : 00:1e:c9:de:03:7b

--  Stack Info  --
Unit  UnitType     Status         ReqTyp          CurTyp          Version
Ports
-------------------------------------------------------------------------------
----
  0   Management   online         PE-FN-410S-IOA  PE-FN-410S-IOA  1-0(0-1864)
12
  1   Member       not present
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
```

```
Dell#
```

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

```
Dell(conf)# interface managementethernet 0/0
Dell(conf-if-ma-0/0)# ip address 192.168.5.6/16
Dell(conf-if-ma-0/0)#
Dell(conf-if-ma-0/0)#
Dell(conf-if-ma-0/0)#show config
!
interface ManagementEthernet 0/0
ip address 192.168.5.6/16
no shutdown
Dell(conf-if-ma-0/0)#
Dell(conf-if-ma-0/0)# no ip address
Dell(conf-if-ma-0/0)#
Dell(conf-if-ma-0/0)# show config
!
interface ManagementEthernet 0/0
  no ip address
  no shutdown
Dell(conf-if-ma-0/0)#
```

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 prompt lists all of the available keywords. The output of this command is the same for the help command.

  ```
  Dell#?
  start            Start Shell
  capture         Capture Packet
  cd            Change current directory
  clear            Reset functions
  clock             Manage the system clock
  configure         Configuring from terminal
  copy             Copy from one file to another
  --More--
  ```

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

  ```
  Dell(conf)#cl?
  clock
  Dell(conf)#cl
  ```

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

  ```
  Dell(conf)#clock ?
  summer-time      Configure summer (daylight savings) time
  ```

```
timezone          Configure time zone
Dell(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.
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-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

Dell 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 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 0/1`.
- `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 linecard all` command.

```
Dell(conf)#do show stack-unit all stack-ports all pfc details | grep 0
stack unit 0 stack-port all
    0 Pause Tx pkts, 0 Pause Rx pkts
    0 Pause Tx pkts, 0 Pause Rx pkts
    0 Pause Tx pkts, 0 Pause Rx pkts
    0 Pause Tx pkts, 0 Pause Rx pkts
    0 Pause Tx pkts, 0 Pause Rx pkts
    0 Pause Tx pkts, 0 Pause Rx pkts
```

NOTE: Dell 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 linecard all` command.

Example of the `except` Keyword

```
Dell(conf)#do show stack-unit all stack-ports all pfc details | except 0

     Admin mode is On
     Admin is enabled
     Local is enabled
     Link Delay 65535 pause quantum
Dell(conf)#
```

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 linecard all` command.

Example of the `find` Keyword

```
Dell(conf)#do show stack-unit all stack-ports all pfc details | find 0
 stack unit 0 stack-port all
     Admin mode is On
     Admin is enabled
     Local is enabled
     Link Delay 65535 pause quantum
     0 Pause Tx pkts, 0 Pause Rx pkts
Dell(conf)#
```

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: `Dell# command | grep regular-expression | except regular-expression | grep other-regular-expression | find regular-expression | save`.

Multiple Users in Configuration Mode

Dell 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 Networking recommends coordinating with the users listed in the message so that you do not unintentionally overwrite each other's configuration changes.

4

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) during which the route processor module (RPM), switch fabric module (SFM), and line card status light emitting diodes (LEDs) blink green. The system then loads the Dell 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 RPM and line card 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.

Console Access

The switch has two management ports available for system access: a serial console port and an out-of-bounds (OOB) port.

Serial Console

A universal serial bus (USB) (A-Type) connector is located at the front panel. The USB can be defined as an External Serial Console (RS-232) port, and is labeled on the chassis. The USB is present on the lower side, as you face the I/O side of the chassis, as shown.

Serial Console

External Serial Port with a USB Connector

The following table list the pin assignments.

Table 2. Pin Assignments

USB Pin Number	Signal Name
Pin 1	RTS
Pin 2	RX
Pin 3	TX
Pin 4	CTS
Pin 5, 6	GND
RxD	Chassis GND

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 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.

Boot Process

After you follow the *Installation Procedure* in the *Getting Started Guide*, the switch boots up.

The switch with the Dell Networking OS version 8.3.16.1 requires boot flash version 4.0.1.0 and boot selector version 4.0.0.0. The following example shows the completed boot process.

```
syncing disks... done
unmounting file systems...
unmounting /f10/flash (/dev/ld0e)...
unmounting /usr (mfs:31)...
unmounting /lib (mfs:23)...
unmounting /f10 (mfs:20)...
unmounting /tmp (mfs:15)...
unmounting /kern (kernfs)...
unmounting / (/dev/md0a)... done
rebooting...

NetLogic XLP Stage 1 Loader
Built by build at tools-sjc-01 on Thu May 31 23:53:38 2012
IOM Boot Selector Label 4.0.0.0

Nodes online: 1
  GPIO 22 init'ed as an output
  GPIO 23 init'ed as an output
I2C0 speed = 30 KHz, prescaler = 0x0377.
Initialized I2C0 Controller.
I2C1 speed = 100 KHz, prescaler = 0x0109.
Initialized I2C1 Controller.
DDR SPD: Node 0 Channel 0 Mem size = 2048 MB
DDR SPD: Node 0 DRAM frequency 666 MHz
DDR SPD: Node 0 CPU frequency 1200 MHz
RTT Norm:44
NBU0 DRAM BAR0 base: 00000000 limit: 0013f000 xlate: 00000001 node: 00000000
( 0 MB -> 320 MB
, size: 320 MB)
NBU0 DRAM BAR1 base: 001d0000 limit: 0088f000 xlate: 00090001 node: 00000000
( 464 MB -> 2192 MB
, size: 1728 MB)
```

```
Modifying Default Flash Address map..Done
Initialized eMMC Host Controller
Detected SD Card
BLC is 1 (preset 10)
Hit any key to stop autoboot: 0
Boot Image selection
Reading the Boot Block Info...Passed !!
Images are OK A:0x0 B:0x0
Boot Selector set to Bootflash Partition A image...
Verifying Copyright Information..success for Image - 0
Boot Selector: Booting Bootflash Partition A image...
Copying stage-2 loader from 0xb6120000 to 0x8c100000(size = 0x100000)
Boot Image selection DONE.
## Starting application at 0x8C100000 ...

U-Boot 2010.03-rc1(Dell Force10)
Built by build at tools-sjc-01 on Thu May 31 23:53:38 2012
IOM Boot Label 4.0.1.0

DRAM: 2 GB
Initialized CPLD on CS3
Detected [XLP308 (Lite+) Rev A0]
Initializing I2C0: speed = 30 KHz, prescaler = 0x0377 -- done.
Initializing I2C1: speed = 100 KHz, prescaler = 0x0109 -- done.
Initialized eMMC Host Controller
Detected SD Card
Now running in RAM - U-Boot [N64 ABI, Big-Endian] at: ffffffff8c100000
Flash: 256 MB
PCIE (B0:D01:F0) : Link up.
PCIE (B0:D01:F1) : No Link.
In:  serial
Out: serial
Err: serial
Net: nae-0: PHY is Broadcom BCM54616S

--More--

SOFTWARE IMAGE HEADER DATA :
----------------------------

--More--

Starting Dell Networking application

Welcome to Dell Easy Setup Wizard

The setup wizard guides you through the initial switch configuration, and gets
you up and running as quickly as possible. You can skip the setup wizard, and
enter CLI mode to manually configure the switch. You must respond to the next
question to run the setup wizard within 60 seconds, otherwise the system will
continue with normal operation using the default system configuration.
Note: You can exit the setup wizard at any point by entering [ctrl+c].

Would you like to run the setup wizard (you must answer this question within
60 seconds)? [Y/N]: N
00:00:40: %STKUNIT0-M:CP %IFMGR-5-ASTATE_UP: Changed interface Admin state to
up: Vl 1
00:00:42: %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
00:00:42: %STKUNIT0-M:CP %SEC-5-LOGIN_SUCCESS: Login successful for user on
line console
Dell>en
Password:
```

Default Configuration

A version of the Dell Networking OS is pre-loaded onto the chassis; however, the system is not configured when you power up for the first time (except for the default hostname, which is `Dell`). 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 following command.

- Create a host name.
  CONFIGURATION mode

  `hostname` *`name`*

Example of the `hostname` Command

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

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 following command.

- Create a host name.
  CONFIGURATION mode

  `hostname` *`name`*

Example of the `hostname` Command

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

Accessing the System Remotely

You can configure the system to access it remotely by Telnet or SSH.

The switch has a dedicated management port and a management routing table that is separate from the IP routing table.

Accessing the Switch Remotely

Configuring the system for Telnet 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
   ```

   - *slot*: the range is 0.
   - *port*: the range is 0.
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 username password [encryption-type] password
  ```

  - *encryption-type*: specifies how you are inputting the password, is 0 by default, and is not required.
    * 0 is for inputting the password in clear text.
    * 7 is for inputting a password that is already encrypted using a Type 7 hash. Obtaining the encrypted password from the configuration of another Dell Networking system.

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 two types of enable passwords:

- `enable password` stores the password in the running/startup configuration using a DES encryption method.
- `enable secret` is stored in the running/startup configuration in using a stronger, MD5 encryption method.

Dell Networking recommends using the `enable secret` password.

To configure an enable password, use the following command.

- Create a password to access EXEC Privilege mode.
  CONFIGURATION mode

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

  - *level*: is the privilege level, is 15 by default, and is not required
  - *encryption-type*: specifies how you are inputting the password, is 0 by default, and is not required.
    * 0 is for inputting the password in clear text.
    * 7 is for inputting a password that is already encrypted using a DES hash. Obtain the encrypted password from the configuration file of another Dell Networking system. You can only use this for the `enable` password.
    * 5 is for inputting a password that is already encrypted using an MD5 hash. Obtain the encrypted password from the configuration file of another Dell Networking system. You can only use this for the `enable secret` password.

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.

NOTE: Using flash memory cards in the system that have not been approved by Dell Networking can cause unexpected system behavior, including a reboot.

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 Networking OS Command Line Reference Guide*.

- 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 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
Internal flash: `flash`	`copy flash://filename`	`flash://filename`
USB flash: `usbflash`	`usbflash://filename`	`usbflash://filename`
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 you configured a domain name server (DNS) server.

NOTE: If all of the following conditions are true, the Portmode Hybrid configuration is not applied, because of the configuration process for server ports as switch ports by default:

- The running configuration is saved in flash.
- The startup configuration is deleted.
- The switch is reloaded.
- The saved configuration is copied to the running configuration.

To avoid this scenario, delete the switch port configuration from the running configuration before copying the saved configuration to the running configuration.

Example of Copying a File to an FTP Server

Example of Importing a File to the Local System

The bold flash shows the local location and the bold ftp shows the remote location.

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

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

Save the Running-Configuration

The running-configuration contains the current system configuration. Dell Networking recommends coping your running-configuration to the startup-configuration.
The system uses the startup-configuration during boot-up to configure the system. The startup-configuration is stored in the internal flash on the IOM by default, but you can save it to a USB flash device or a remote server.
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-config* and *running-config*. These commands assume that current directory is the internal flash, which is the system default.

- Save the running-config to the startup-configuration on the internal flash.
  EXEC Privilege mode

  ```
  copy running-config startup-config
  ```
- Save the running-configuration on the IOM.
  EXEC Privilege mode

  ```
  copy running-config usbflash://filename
  ```
- 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, you can only use a host name if you have configured a DNS server.
- Save the running-configuration to the startup-configuration on the internal flash of the primary RPM. Then copy the new startup-config file to the external flash of the primary RPM.
  EXEC Privilege mode

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

Dell Networking OS Behavior: If you create a startup-configuration on an RPM and then move the RPM to another chassis, the startup-configuration is stored as a backup file (with the extension *.bak*), and a new, empty startup-configuration file is created. To restore your original startup-configuration in this situation, overwrite the new startup-configuration with the original one using the `copy startup-config.bak startup-config` command.

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 a list of files on the usbflash.
  EXEC Privilege mode

  ```
  dir usbflash:
  ```
- View the contents of a file in the internal flash.
  EXEC Privilege mode

  ```
  show file flash://filename
  ```
- View the contents of a file in the usb flash.
  EXEC Privilege mode

  ```
  show file usbflash://filename
  ```
- View the running-configuration.
  EXEC Privilege mode

  ```
  show running-config
  ```
- View the startup-configuration.
  EXEC Privilege mode

  ```
  show startup-config
  ```

Example of the `dir` Command

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

```
Dell#dir
Directory of flash:

1 drwx       4096 Jan 01 1980 00:00:00 +00:00 .
2 drwx       2048 May 10 2011 14:45:15 +00:00 ..
3 drwx       4096 Feb 17 2011 00:28:00 +00:00 TRACE_LOG_DIR
4 drwx       4096 Feb 17 2011 00:28:02 +00:00 CORE_DUMP_DIR
5 d---       4096 Feb 17 2011 00:28:02 +00:00 ADMIN_DIR
6 -rwx       1272 Apr 29 2011 16:15:14 +00:00 startup-config
7 -rwx      10093 Feb 17 2011 20:48:02 +00:00 abhi-jan26.cfg
8 -rwx     217155 Feb 22 2011 23:14:34 +00:00 show-tech-cfg.txt
9 -rwx       5162 Mar 02 2011 04:02:58 +00:00 runn-feb6
10 -rwx     10507 Mar 03 2011 01:17:16 +00:00 abhi-feb7.cfg
11 -rwx         4 May 06 2011 22:05:06 +00:00 dhcpBindConflict
12 -rwx      6900 Feb 17 2011 04:43:12 +00:00 startup-config.bak
13 -rwx   1244038 Feb 13 2011 04:27:16 +00:00 f10cp_sysd_110213042625.acore.gz

flash: 2143281152 bytes total (2123755520 bytes free)
--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 will not be preserved after a system reboot.

Example of the `show running-config` Command

```
Dell#show running-config
Current Configuration ...
Current Configuration ...
! Version E8-3-16-0
! Last configuration change at Tue Mar 6 11:51:50 2012 by default
! Startup-config last updated at Tue Mar 6 07:41:23 2012 by default
!
boot system stack-unit 5 primary tftp://10.11.200.241/dt-m1000e-3-a2
boot system stack-unit 5 secondary system: B:
boot system stack-unit 5 default tftp://10.11.200.241/dt-m1000e-3-b2
boot system gateway 10.11.209.254
--More--
```

Managing the File System

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

The system stores files on the internal Flash by default but you can configure the system 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.

```
Dell#show file-systems
Size(b)     Free(b)    Feature Type    Flags Prefixes
 2143281152 2000785408 FAT32 USERFLASH rw flash:
15848660992  831594496 FAT32 USBFLASH  rw usbflash:
           -          -     -  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
  ```

You can change the default storage location to the USB Flash, as shown. File management commands then apply to the USB Flash rather than the internal Flash. The bold lines show that no file system is specified and that the file is saved to an USB Flash.

```
Dell#cd usbflash:
Dell#copy running-config test
!
3998 bytes successfully copied

DellS#dir
Directory of usbflash:

1 drwx 4096 Jan 01 1980 00:00:00 +00:00 .
2 drwx 2048 May 02 2012 07:05:06 +00:00 ..
3 -rwx 1272 Apr 29 2011 16:15:14 +00:00 startup-config
4 -rwx 3998 May 11 2011 23:36:12 +00:00 test
```

View the 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.

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

Example of the `show command-history` Command

```
Dell#show command-history
[5/18 21:58:32]: CMD-(TEL0):[enable]by admin from vty0 (10.11.68.5)
[5/18 21:58:48]: CMD-(TEL0):[configure]by admin from vty0 (10.11.68.5)
    - Repeated 1 time.
[5/18 21:58:57]: CMD-(TEL0):[interface port-channel 1]by admin from vty0
(10.11.68.5)
[5/18 21:59:9]: CMD-(TEL0):[show config]by admin from vty0 (10.11.68.5)
```

```
[5/18 22:4:32]: CMD-(TEL0):[exit]by admin from vty0 (10.11.68.5)
[5/18 22:4:41]: CMD-(TEL0):[show interfaces port-channel brief]by admin from
vty0
(10.11.68.5)
```

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. Use the **copy source-file-url http://host[:port]/file-path** command to transfer files to an external server. 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 external FLASH, enter usbflash:// followed by the filename.

In the Dell Networking OS release 9.8(0.0), HTTP services are enhanced to 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

  Dell(conf)#ip http vrf {management | <*vrf-name*>}

Upgrading and Downgrading the Dell Networking OS

NOTE: To upgrade the Dell Networking OS, refer to the Release Notes for the version you want to load on the system.

Using Hashes to Verify Software Images Before Installation

You can use the MD5 message-digest algorithm or SHA256 Secure Hash Algorithm to validate the software image on the flash drive, after the image has been transferred to the system, but before the image has been installed. The validation calculates a hash value of the downloaded image file on system's flash drive, and**,** optionally**,** compares it to a Dell 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, can prevent 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 i-Support. Optionally, the published hash can be included in the **verify** {**md5** | **sha256**} command, which will display whether it matches the calculated hash of the indicated file.

To validate a software image:

1. Download Dell Networking OS software image file from the iSupport page to the local (FTP or TFTP) server. The published hash for that file is displayed next to the software image file on the iSupport page.
2. Go on to the Dell 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 has been transferred to the system, but before the image has been installed, 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 is to use the flash drive. You can just enter the image file name.
- *hash-value*: (Optional). Specify the relevant hash published on i-Support.
- *img-file*: Enter the name **of** the Dell Networking **software** image file to validate

Examples: Without Entering the Hash Value for Verification

MD5

```
Dell# verify md5 flash://FTOS-SE-9.5.0.0.bin
MD5 hash for FTOS-SE-9.5.0.0.bin: 275ceb73a4f3118e1d6bcf7d75753459
```

SHA256

```
Dell# verify sha256 flash://FTOS-SE-9.5.0.0.bin
SHA256 hash for FTOS-SE-9.5.0.0.bin:
e6328c06faf814e6899ceead219afbf9360e986d692988023b749e6b2093e933
```

Examples: Entering the Hash Value for Verification

MD5

```
Dell# verify md5 flash://FTOS-SE-9.5.0.0.bin 275ceb73a4f3118e1d6bcf7d75753459
MD5 hash VERIFIED for FTOS-SE-9.5.0.0.bin
```

SHA256

```
Dell# verify sha256 flash://FTOS-SE-9.5.0.0.bin
e6328c06faf814e6899ceead219afbf9360e986d692988023b749e6b2093e933
SHA256 hash VERIFIED for FTOS-SE-9.5.0.0.bin
```

Deploying FN I/O Module

This section provides design and configuration guidance for deploying the Dell PowerEdge FN I/O Module (FN IOM).

By default the FN IOM is in Standalone Mode. It consists of the following default settings:

Mode	Default Settings
Chassis Management Controller Configuration	Yes
Stacking and VLT	Disabled
Data Center Bridging Exchange Protocol (DCBx)	Enabled
FIP Snooping Bridge (FSB)	Enabled on FN 410S & FN 410T; Disabled on FN 2210S
NPIV Proxy Gateway Mode (NPG)	Enabled in FN 2210S; Disabled on FN 410S & FN 410T
iSCSI Optimization	Enabled
Broadcast Storm Control	Enabled
Internet Group Management Protocol (IGMP) Flood Restrict	Enabled
Network Time Protocol (NTP)	Enabled
Enable and Disable ports	All port enabled by default. Disable ports through Command Line Interface

Mode	Default Settings
Uplink Failure Detection (UFD)	Enabled
TACACS+/RADIUS (AAA Auth/Acct; Auth coming up)	Disabled
DHCP	Enabled n client only mode
Internet Group Management Protocol (IGMP) v2 & v3	Enabled
Auto LACP LAG	Enabled on all ports (Uplink created by default based on LACP PDU)
All VLANS	Available on all server and uplink ports

Ensuring the FN IOM system is in Standalone Mode

To ensure that the FN IOM is in Standalone Mode, use the `show system stack-unit 0 iom-mode` command.

```
Dell#show system stack-unit 0 iom-mode
Unit    Boot-Mode              Next-Boot
------------------------------------------------
0        standalone
```

If the FN IOM system is not in the Standalone Mode, reset the system to its factory default state. To reset the system, use the `restore factory-defaults stack-unit 0 clear-all` command. It brings the system in standalone (preconfigured) mode and sets all configuration settings to a known state.

```
Dell#restore factory-defaults stack-unit 0 clear-all
Proceed with factory settings? Confirm [yes/no]:y
```

Configuration Examples

Following are the configuration examples for common upstream switches. To establish a functional uplink and bring the network ports on PowerEdge FC-Series servers up (FC830, FC630, FC430, FC 620, and FC420) use the following steps:

1. Ensure that the FN IOM system is in Standalone Mode.
2. Create a LACP LAG on the upstream top of rack switch.
3. Verify the connection.

By default the network ports on the PowerEdge FC-Series servers installed in the FX2 chassis is down, until the uplink port channel is operational on the FN IOM system. It is due to an Uplink Failure Detection, by that when upstream connectivity fails; the FN IOM disables the downstream links.

You can disable the UFD on a preconfigured FN IOM using the following commands:

```
Dell#configure
Dell(conf)#uplink-state-group 1
Dell(conf-uplink-state-group-1)#no enable

Dell#show uplink-state-group
Uplink State Group: 1   Status: Disabled, Down
```

To re-enable UFD, use the following commands:

```
Dell#configure
Dell(conf)#uplink-state-group 1
Dell(conf-uplink-state-group-1)#enable

Dell#show uplink-state-group
Uplink State Group: 1   Status: Enabled, Up
```

Sample FN IOM commands and outputs

In the following, port channel 128 is up. The port channel consists of ports TenGigabitEthernet 0/11 and TenGigabitEthernet 0/12.

```
Dell#show interfaces port-channel 128
Port-channel 128 is up, line protocol is up
Created by LACP protocol
Hardware address is f8:b1:56:6e:fc:59, Current address is f8:b1:56:6e:fc:59
Interface index is 1258356736
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 :f8b1566efc59
MTU 12000 bytes, IP MTU 11982 bytes
LineSpeed 20000 Mbit
Members in this channel:  Te 0/11(U) Te 0/12(U)
ARP type: ARPA, ARP Timeout 04:00:00
Queueing strategy: fifo
Input Statistics:
     194 packets, 26975 bytes
     40 64-byte pkts, 47 over 64-byte pkts, 83 over 127-byte pkts
     24 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     160 Multicasts, 34 Broadcasts, 0 Unicasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     1619 packets, 212604 bytes, 0 underruns
     0 64-byte pkts, 201 over 64-byte pkts, 1376 over 127-byte pkts
     42 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     1531 Multicasts, 88 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,          1 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:13:45
```

In the following, port channel 128 is down. As a result, the downstream ports (TenGigabitEthernet 0/1 – 8) are disabled by UFD.

```
Dell#show interfaces tengigabitethernet 0/1
TenGigabitEthernet 0/1 is up, line protocol is down(error-disabled[UFD])
Hardware is DellEth, address is f8:b1:56:6e:fc:59
    Current address is f8:b1:56:6e:fc:59
Server Port AdminState is N/A
Pluggable media not present
Interface index is 1048580
Internet address is not set
Mode of IPv4 Address Assignment : NONE
DHCP Client-ID :f8b1566efc59
MTU 12000 bytes, IP MTU 11982 bytes
LineSpeed auto
Flowcontrol rx off tx off
```

```
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 01:26:42
Queueing strategy: fifo
Input Statistics:
     941 packets, 98777 bytes
     83 64-byte pkts, 591 over 64-byte pkts, 267 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     694 Multicasts, 247 Broadcasts, 0 Unicasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     1474 packets, 238933 bytes, 0 underruns
     0 64-byte pkts, 204 over 64-byte pkts, 1071 over 127-byte pkts
     175 over 255-byte pkts, 24 over 511-byte pkts, 0 over 1023-byte pkts
     1198 Multicasts, 276 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: 00:00:17
```

NOTE: For more information on Uplink Failure Detection and all other configuration settings, see the Uplink Failure Detection (UFD) chapter.

Configure Interfaces and Port Channel

This section provides the commands for configuring port channels on common upstream switches that are connected to the FN IOM system.

Interfaces and Port Channels

The previous illustration is applicable to the FN 410S and FN 410T servers. If you are using the FN 2210S server in its default configuration, you can only use the TenGigabitethernet 0/11 and TenGigabitethernet 0/12 in the uplink port channel. Ports TenGigabitethernet 0/9 and TenGigabitethernet 0/10 are Fibre Channel ports.

By default on the FN IOM, the external Ethernet ports are preconfigured in port channel 128 with LACP enabled. Port channel 128 is in hybrid (trunk) mode.

To bring up the downstream (server) ports on the FN IOM, port channel 128 must be up. When the Port channel 128 is up, it is connected to a configured port channel on an upstream switch. To enable the Port channel 128, connect any combination of the FN IOM's external Ethernet ports (ports TenGigabitethernet 0/9-12) to the upstream switch. The port channel may have a minimum of one and a maximum of four links.

Following example configures a port channel on an upstream switch:

- No configuration is required on the FN IOM system. All configurations are done on the upstream switch.
- In the example, four interfaces are configured in the port channel. Configure as many interfaces as you have connected. Change the port numbers to match your switch.

- Any valid ID number can be used for the port channel number on the upstream switch.

```
Dell(conf)#interface range tengigabitethernet 0/1-4
Dell(conf-if-te-0/1)#no shut
Dell(conf-if-te-0/1)#port-channel-protocol lacp
Dell(conf-if-te-0/1-lacp)#port-channel 128 mode active

Dell(conf)#interface port-channel 128
Dell(conf-if-po-128)#portmode hybrid
Dell(conf-if-po-128)#switchport
Dell(conf-if-po-128)#no shut
```

Verifying the Configurations

When the port channel of the upstream switch is configured, the port channel on the FN IOM is up. A series of messages is logged on the system indicating port channel 128 and the ports connected to the servers downstream are up.

```
Changed interface state to up: Po 128

Downstream interface cleared from UFD error-disabled: Te 0/1
Downstream interface cleared from UFD error-disabled: Te 0/2
Downstream interface cleared from UFD error-disabled: Te 0/3
…etc.
```

To verify the status, run the following commands:

```
Dell#show interfaces port-channel 128
Port-channel 128 is up, line protocol is up

Dell#show uplink-state-group
Uplink State Group: 1   Status: Enabled, Up
```

5

Management

Dell Networking OS supports management.

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

Configuring Privilege Levels

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

There are 15 privilege levels, of which two are pre-defined. The default privilege level is **1**.

- **Level 1** — Access to the system begins at EXEC mode, and EXEC mode commands are limited to basic commands, some of which are `enable`, `disable`, and `exit`.
- **Level 15** — To access all commands, enter EXEC Privilege mode. Normally, enter a password to enter this mode.

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:

- removing commands from the EXEC mode commands
- moving commands from EXEC Privilege mode to EXEC mode
- allowing access to CONFIGURATION mode commands
- allowing access to INTERFACE, LINE, ROUTE-MAP, and ROUTER mode commands

You can access all commands at your privilege level and below.

Moving a Command from EXEC Privilege Mode to EXEC Mode

Remove a command from the list of available commands in EXEC mode for a specific privilege level using 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 restricted command.

Moving a Command from EXEC Privilege Mode to EXEC Mode

Move a command from EXEC Privilege to EXEC mode for a privilege level using 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

Allow access to CONFIGURATION mode using the `privilege exec level` *`level`* command configure from CONFIGURATION mode. A user that enters CONFIGURATION mode remains at his privilege level, and has access to only two commands, end and exit. Individually specify each

CONFIGURATION mode command to which you want to allow access using the `privilege configure level level` command. 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 INTERFACE, LINE, ROUTE-MAP, and ROUTER Mode

1. Similar to allowing access to CONFIGURATION mode, to allow access to INTERFACE, LINE, ROUTE-MAP, and ROUTER modes, first allow access to the command that enters you into the mode. For example, allow a user to enter INTERFACE mode using the `privilege configure level level interface gigabitethernet` command.
2. Then, 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 keywords in the command to which you want to allow access.

Customizing a Privilege Level

to customize a privilege level, use the following commands.

1. Remove a command from the list of available commands in EXEC mode.
   CONFIGURATION mode

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

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

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

   ```
   privilege configure level level {interface | line | route-map | router}
   {command-keyword ||...|| command-keyword}
   ```
5. 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}
   ```

Create a Custom Privilege Level Apply a Privilege Level to a Username

The following configuration 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 and LINE modes with the `no` command

```
Dell(conf)#do show run privilege
!
Dell(conf)#privilege exec level 3 capture
Dell(conf)#privilege exec level 3 configure
Dell(conf)#privilege exec level 4 resequence
Dell(conf)#privilege exec level 3 clear arp-cache
Dell(conf)#privilege exec level 3 clear arp-cache max-buffer-size
Dell(conf)#privilege configure level 3 line
Dell(conf)#privilege configure level 3 interface
Dell(conf)#do telnet 10.11.80.201
[telnet output omitted]
Dell#show priv
Current privilege level is 3.
Dell#?
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]
Dell#config
[output omitted]
Dell(conf)#do show priv
Current privilege level is 3.
Dell(conf)#?
end                Exit from configuration mode
exit               Exit from configuration mode
interface          Select an interface to configure
Dell(conf)#interface ?
loopback           Loopback interface
managementethernet Management Ethernet interface
null               Null interface
port-channel       Port-channel interface
range              Configure interface range
tengigabitethernet TenGigabit Ethernet interface
vlan               VLAN interface
Dell(conf)#interface tengigabitethernet 1/1
Dell(conf-if-te-1/1)#?
end                Exit from configuration mode
exit               Exit from interface configuration mode
Dell(conf-if-te-1/1)#exit
Dell(conf)#line ?
console Primary terminal line
vty Virtual terminal
Dell(conf)#line vty 0
Dell(conf-line-vty)#?
exit               Exit from line configuration mode
Dell(conf-line-vty)#
```

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 terminal line.
  Line mode

  ```
  privilege levellevel
  ```

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 Networking operating system tracks changes in the system using event and error messages. By default, the system 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.

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

```
Dell(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.

Example of the `show logging auditlog` Command

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

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

Example of the `show logging` Command for Security

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

```
Dell#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

```
Dell# 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

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

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.

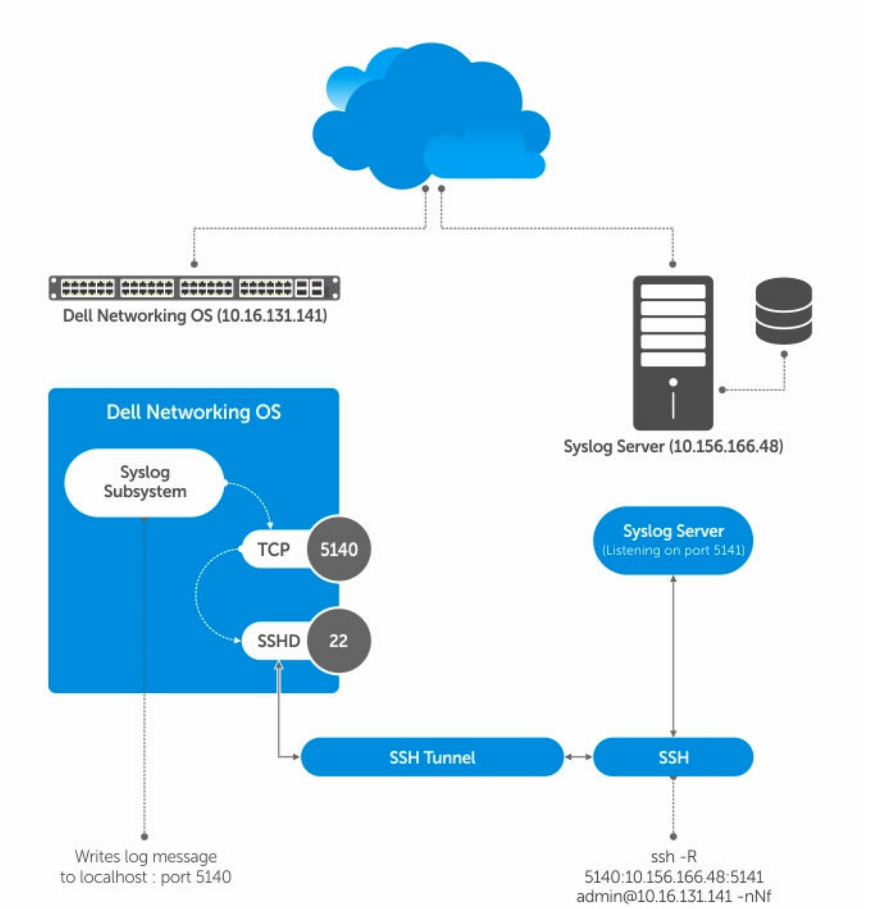


Figure 1. Setting Up a Secure Connection to a Syslog Server

Pre-requisites

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

1. On the switch, enable the SSH server

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

2. On the syslog server, create a reverse SSH tunnel from the syslog server to FTOS 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.

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

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

```
Dell#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.

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:

- Disabling System Logging
- Sending 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}}}
  ```

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/log7.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.

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, the system deletes all messages stored in the buffer. Increasing the buffer size does not affect messages in the buffer.
- Specify the number of messages that the system 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 Configuring 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 the system administrator can view the security logs.

Example of the `show logging` Command

```
Dell#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, as shown in the example for Configuring a UNIX Logging Facility Level.

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)

Example of the `show running-config logging` Command

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

```
Dell#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
Dell#
```

Synchronizing Log Messages

You can configure the system 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]}
   ```

   Configure the following parameters for the virtual terminal lines:

   - *number*: the range is from zero (0) to 9.
   - *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 do not include a time/date stamp 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]
  | uptime]
  ```

  Specify the following optional parameters:

  – `datetime`: You can add the keyword `localtime` to include the `localtime`, `msec`, and `show-timezone`. If you do not add the keyword `localtime`, the time is UTC.

  – `uptime`: To view time since last boot.

  If you do not specify a parameter, the system configures `uptime`.

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.

File Transfer Services

With the Dell 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.

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

Configuration Task List for File Transfer Services

The configuration tasks for file transfer services are:

- Enabling the FTP Server (mandatory)
- Configuring FTP Server Parameters (optional)
- Configuring 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
  ```

Example of Viewing FTP Configuration

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

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 slot/port or number information:
  - For a Loopback interface, enter the keyword `loopback` then a number between 0 and 16383.
  - For a port channel interface, enter the keywords `port-channel` then a number from 1 to 128.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - 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 Enabling the 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 virtual terminal lines (VTYs) connect you through Telnet to the system.

Denying and Permitting Access to a Terminal Line

Dell 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 is a generic way of configuring access restrictions.
- To be able to filter access exclusively using either IPv4 or IPv6 rules, you must 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

  ```
  ip access-class access-list [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.

Example of an ACL that Permits Terminal Access

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

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

Dell OS Behavior: Prior to Dell OS version 7.4.2.0, in order to deny access on a VTY line, apply an ACL and accounting, authentication, and authorization (AAA) to the line. Then users are denied access only after they enter a username and password. Beginning in Dell OS version 7.4.2.0, only an ACL is required, and users are denied access before they are prompted for a username and password.

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, the system 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 `default` keyword. 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
   ```

Example of Terminal Line Authentication

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

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

Setting Time Out of EXEC Privilege Mode

EXEC time-out is a basic security feature that returns the Dell Networking OS to EXEC mode after a period of inactivity on the terminal lines.
To set time out, 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 time-out period to `0`.

  LINE mode

  ```
  exec-timeout minutes [seconds]
  ```

- Return to the default time-out values.

  LINE mode

  ```
  no exec-timeout
  ```

Example of Setting the Time Out Period for EXEC Privilege Mode

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

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

Using Telnet to get to Another Network Device

To telnet to another device, use the following commands.

- Telnet to the stack-unit. You do not need to configure the management port on the stack-unit to be able to telnet to it.

  EXEC Privilege mode

  ```
  telnet-peer-stack-unit
  ```
- Telnet to a device with an IPv4 address.

  EXEC Privilege

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

  If you do not enter an IP address, the system enters a Telnet dialog that prompts you for one.

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

Example of the `telnet` Command for Device Access

```
Dell# telnet 10.11.80.203
Trying 10.11.80.203...
Connected to 10.11.80.203.
Exit character is '^]'.
Login:
Login: admin
Password:
Dell>exit
Dell#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
Dell#
```

Lock CONFIGURATION Mode

The systems 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 locks: 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

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

Dell#config
! Locks configuration mode exclusively.
Dell(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.

Limit Concurrent Login Sessions

Dell Networking OS enables you to limit the number of concurrent login sessions of users on VTY, Aux, 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 Aux 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 all users.
  CONFIGURATION mode

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

Example of Configuring Concurrent Session Limit

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

```
Dell(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
  ```

Example of Enabling the System to Clear Existing Sessions

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

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

Example of Clearing Existing Sessions

When you try to login, 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 VTY 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]:
```

Track Login Activity

Dell 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. 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
```

Example of Configuring Login Activity Tracking

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

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

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

```
Dell(config)#login statistics enable
Dell(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.

```
Dell#show login statistics

------------------------------------------------------------------
User: admin
Last login time: Mon Feb 16 04:40:00 2015
Last login location: Line vty0 ( 10.14.1.97 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 30 day(s): 3
------------------------------------------------------------------
```

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.

```
Dell#show login statistics all

------------------------------------------------------------------
User: admin
Last login time: Mon Feb 16 04:40:00 2015
Last login location: Line vty0 ( 10.14.1.97 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 7 day(s): 3
------------------------------------------------------------------

------------------------------------------------------------------
User: secadm
Last login time: Mon Feb 16 04:45:29 2015
Last login location: Line vty0 ( 10.14.1.97 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 7 day(s): 0
------------------------------------------------------------------
```

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.

```
Dell#show login statistics user admin
```

```
---------------------------------------------------------------------
User: admin
Last login time: Mon Feb 16 04:40:00 2015
Last login location: Line vty0 ( 10.14.1.97 )
Unsuccessful login attempt(s) since the last successful login: 0
Unsuccessful login attempt(s) in last 11 day(s): 3
---------------------------------------------------------------------
```

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
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
4. Set the system parameters to ignore the enable password when the system reloads.
   uBoot mode

   ```
   setenv enablepwdignore true
   ```
5. Reload the system.
   uBoot mode

   ```
   reset
   ```
6. Configure a new enable password.
   CONFIGURATION mode

   ```
   enable {secret | password}
   ```
7. 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, refer to 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
   ```

6

802.1X

802.1X is a method of port security.

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 can be verified (through a username and password, for example). This feature is named for its IEEE specification.

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 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 Networking operating system 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.

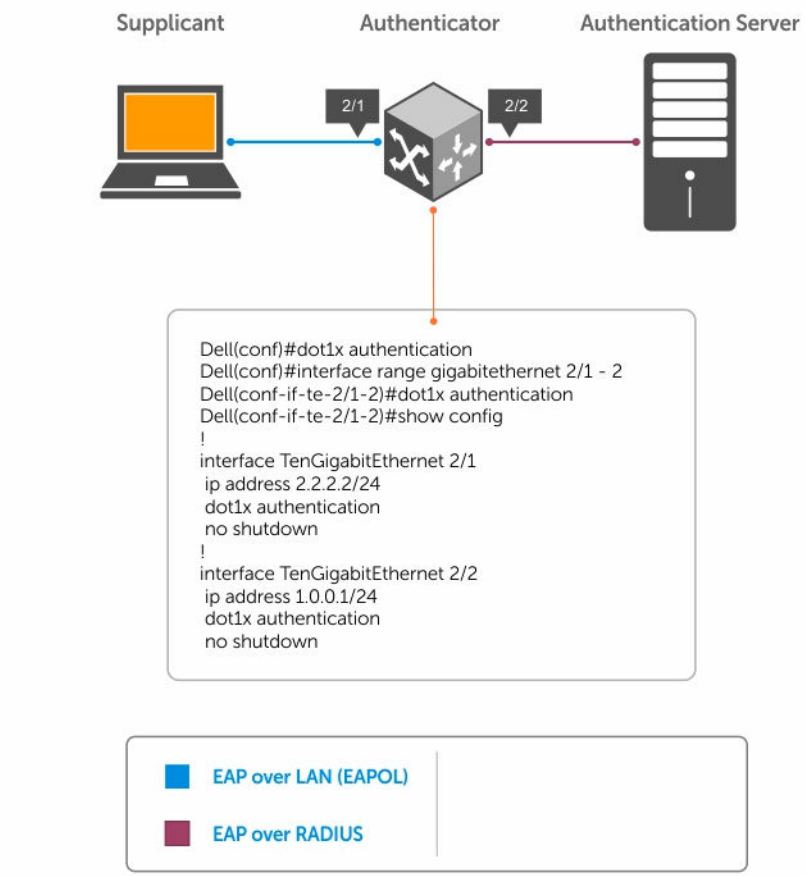


Figure 2. EAP Frames Encapsulated in Ethernet and RADUIS

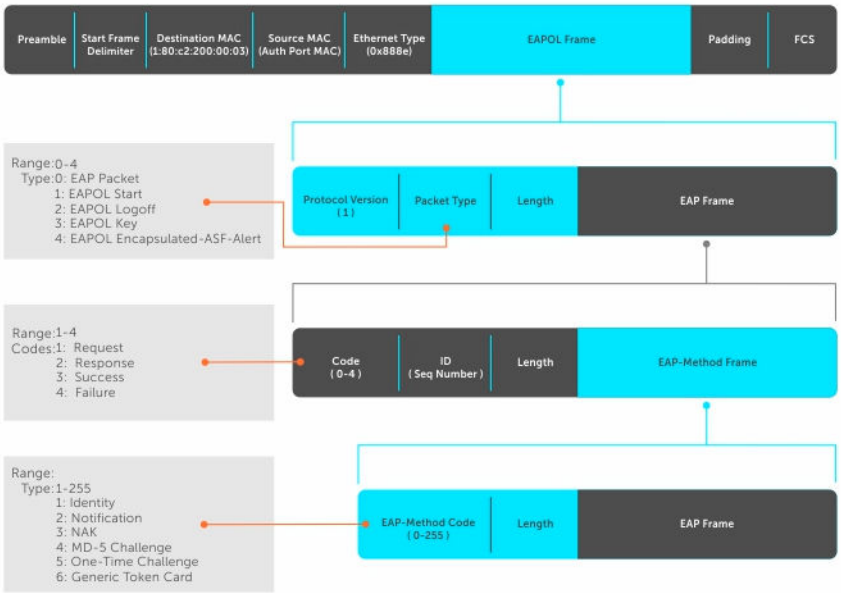


Figure 3. EAP Frames Encapsulated in Ethernet and RADUIS

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 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 Networking switches place 802.1X-enabled ports in the unauthorized state by default.

The 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 that the supplicant 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.

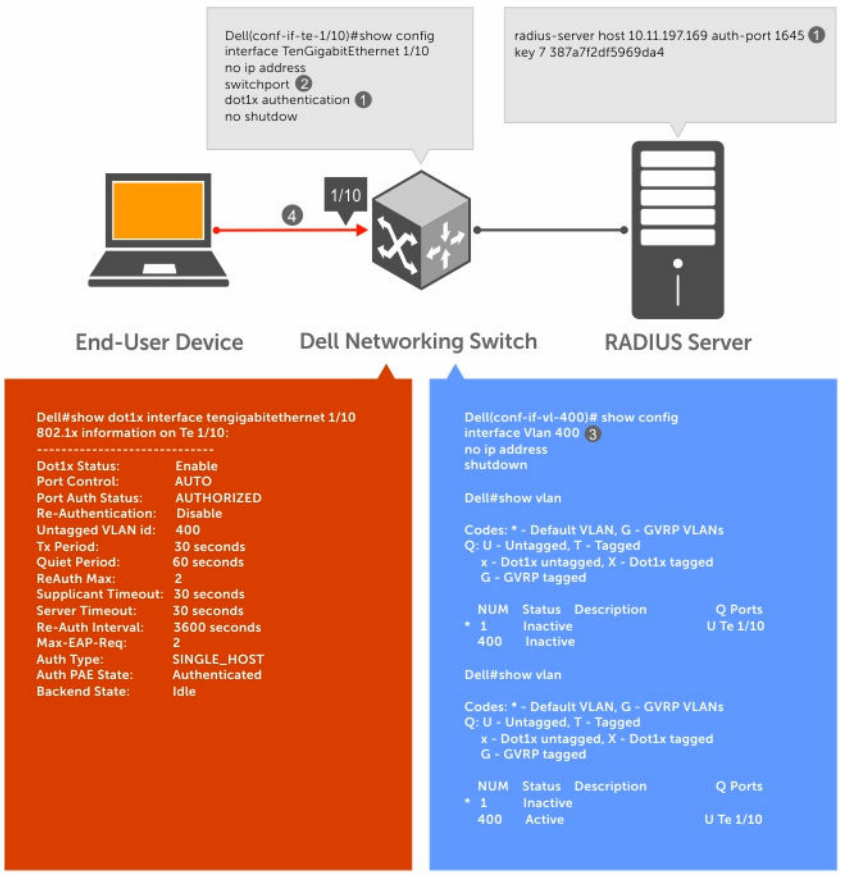


Figure 4. EAP Port-Authentication

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.

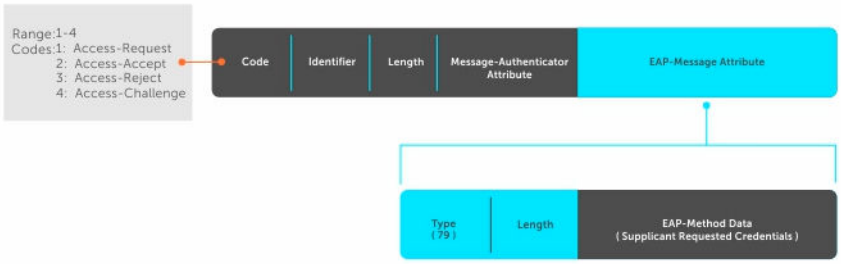


Figure 5. EAP Over RADIUS

RADIUS Attributes for 802.1 Support

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

Attribute 5	**NAS-Port**: the physical port number by which the authenticator is connected to the supplicant.
Attribute 31	**Calling-station-id**: relays the supplicant MAC address to the authentication server.
Attribute 41	**NAS-Port-Type**: NAS-port physical port type. 5 indicates Ethernet.
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 two-step process.

1. Enable 802.1X globally (refer to Enabling 802.1X).
2. Enable 802.1X on an interface (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

- The Dell Networking OS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.
- 802.1X is not supported on port-channels or port-channel members.

Enabling 802.1X

Enable 802.1X globally and at a interface level.

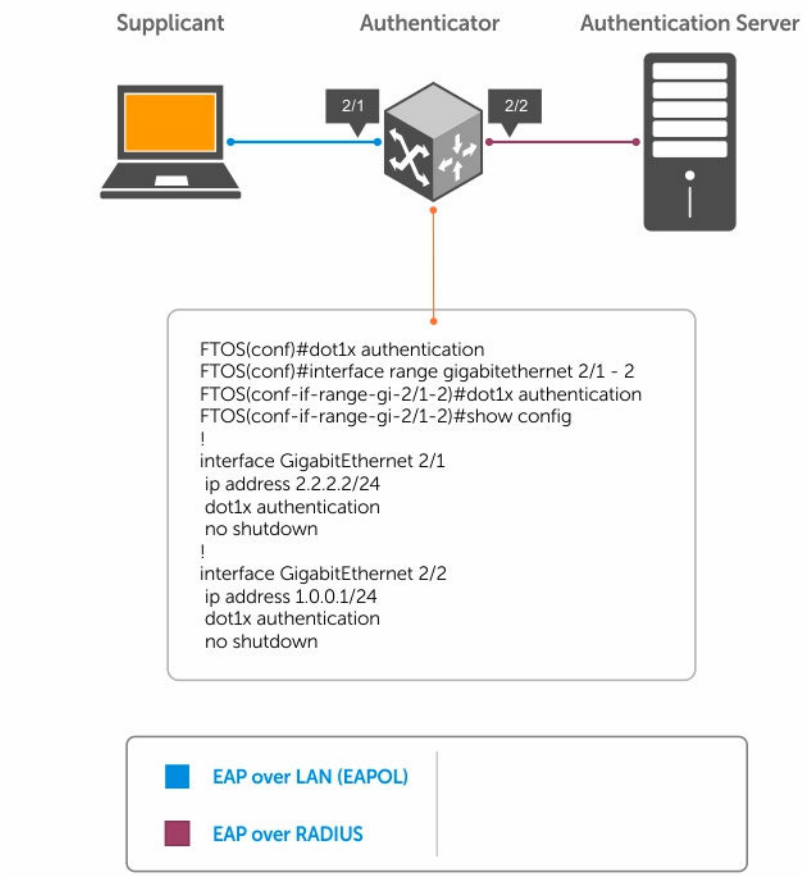


Figure 6. 802.1X Enabled

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 an interface or a range of interfaces.
   INTERFACE mode

   ```
   dot1x authentication
   ```

Example of Verifying that 802.1X is Enabled Globally

Example of Verifying 802.1X is Enabled on an Interface

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

The bold lines show that 802.1X is enabled.

```
Dell#show running-config | find dot1x
dot1x authentication
!
[output omitted]
!
interface GigabitEthernet 2/1
  ip address 2.2.2.2/24
  dot1x authentication
  no shutdown
!
interface GigabitEthernet 2/2
  ip address 1.0.0.1/24
  dot1x authentication
  no shutdown
--More--
```

View 802.1X configuration information for an interface using the `show dot1x interface` command.

The bold lines show that 802.1X is enabled on all ports unauthorized by default.

```
Dell#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 Request Identity Re-Transmissions

If the authenticator sends a Request Identity frame, but the supplicant does not respond, the authenticator waits 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 are configurable.

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 a 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 a maximum of 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, but you can configure this period.

NOTE: The quiet period (`dot1x quiet-period`) is a transmit interval for 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**.

Example of Configuring and Verifying Port Authentication

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.

```
Dell(conf-if-range-Te-0/0)#dot1x tx-period 90
Dell(conf-if-range-Te-0/0)#dot1x max-eap-req 10
Dell(conf-if-range-Te-0/0)#dot1x quiet-period 120
Dell#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

IEEE 802.1X requires that a port can be manually placed into any of 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 Force Authorized, Force Unauthorized, or Auto state.
  INTERFACE mode

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

  The default state is **auto**.

Example of Placing a Port in Force-Authorized State and Viewing the Configuration

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

The bold line shows the new port-control state.

```
Dell(conf-if-gi-2/1)#dot1x port-control force-authorized
Dell(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1
```

```
802.1x information on Gi 2/1:
-----------------------------
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, but you can configure this interval. You can configure a 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 65535.

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

  ```
  dot1x reauth-max number
  ```

  The range is from 1 to 10.

  The default is **2**.

Example of Re-Authenticating a Port and Verifying the Configuration

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

```
Dell(conf-if-gi-2/1)#dot1x reauthentication interval 7200
Dell(conf-if-gi-2/1)#dot1x reauth-max 10
Dell(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

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

```
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
```

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**.

Example of Viewing Configured Server Timeouts

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.

```
Dell(conf-if-gi-2/1)#dot1x port-control force-authorized
Dell(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
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

The system 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 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 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.

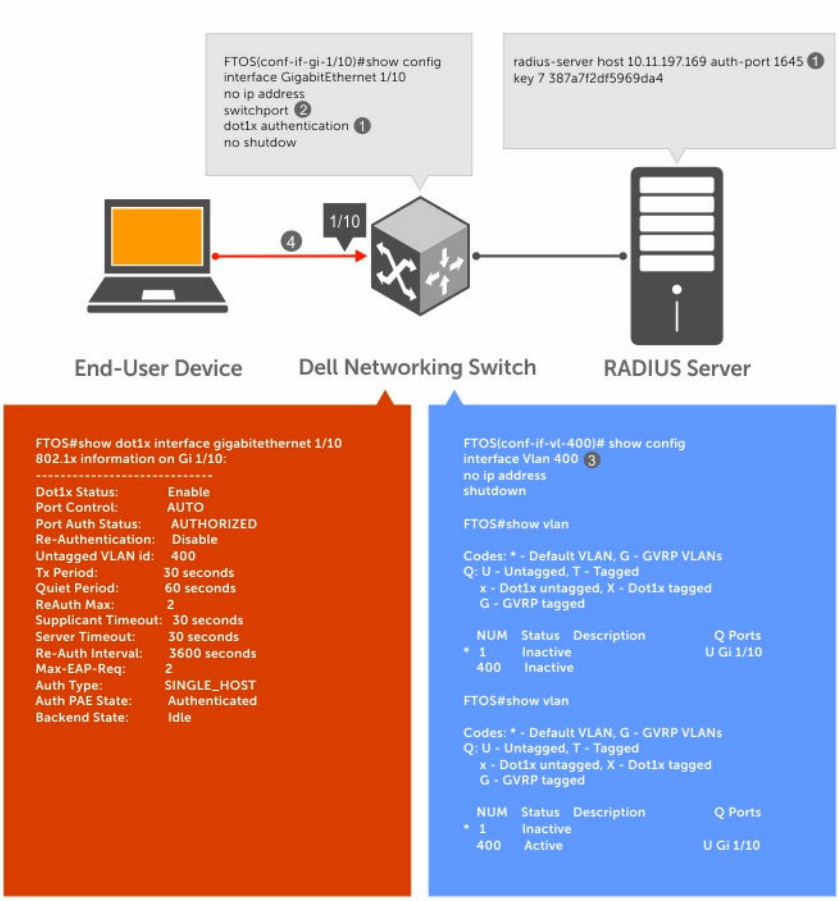


Figure 7. Dynamic VLAN Assignment

1. Configure 8021.x globally (refer to Enabling 802.1X) along with relevant RADIUS server configurations (refer to the illustration in Dynamic 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 Networking 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

```
Dell(conf-if-gi-1/2)#dot1x guest-vlan 200
Dell(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
   switchport
   dot1x guest-vlan 200
   no shutdown
Dell(conf-if-gi-1/2)#
```

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

Example of Viewing Configured Authentication

```
Dell(conf-if-gi-1/2)#dot1x auth-fail-vlan 100 max-attempts 5
Dell(conf-if-gi-1/2)#show config
!
```

```
interface GigabitEthernet 1/2
  switchport
  dot1x guest-vlan 200
  dot1x auth-fail-vlan 100 max-attempts 5
  no shutdown
Dell(conf-if-gi-1/2)#
```

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.

```
Dell(conf-if-gi-2/1)#dot1x port-control force-authorized
Dell(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Enable
Guest VLAN id:             200
Auth-Fail VLAN:            Enable
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
```

7

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

This chapter describes the access control list (ACL) VLAN group and content addressable memory (CAM) enhancements.

Optimizing CAM Utilization During the Attachment of ACLs to VLANs

You can enable and configure the ACL CAM optimization functionality to minimize the number of entries in CAM while ACLs are applied on a VLAN or a set of VLANs, and also while ACLs are applied on a set of ports. This capability enables the effective usage of the CAM space when Layer 3 ACLs are applied to a set of VLANs and when Layer 2 or Layer 3 ACLs are applied on a set of ports.

In releases of Dell Networking OS that do not support the CAM optimization functionality, when an ACL is applied on a VLAN, the ACL rules are configured with the rule-specific parameters and the VLAN as additional attributes in the ACL region. When the ACL is applied on multiple VLAN interfaces, the consumption of the CAM space increases proportionally. For example, when an ACL with 'n' number of rules is applied on 'm' number of VLAN interfaces, a total of n*m entries are configured in the CAM region that is allocated for ACLs. Similarly, when an L2 or L3 ACL is applied on a set of ports, a large portion of the CAM space gets used because a port is saved as a parameter in CAM.

To avoid excessive consumption of the CAM space, configure ACL VLAN groups, which combine all the VLANs that are applied with the same ACL, 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 significantly and saves memory space by using the class ID as a filtering criterion in CAM instead of the VLAN ID.

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. If you apply an ACL separately on the VLAN interface, each ACL has a mapping with the VLAN and increased CAM space utilization occurs. Attaching an ACL individually to VLAN interfaces is similar to the behavior of ACL-VLAN mapping storage in CAM prior to the implementation of the ACL VLAN group functionality.

The ACL manager application on router processor (RP1) contains all the state information about all the ACL VLAN groups that are present. The ACL handler on control processor (CP) and the ACL agent on line cards do not contain any stateful information about the group. The ACL manager application performs the validation after you enter the `acl-vlan-group` command. 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 appropriate error message is displayed. The ACL manager application verifies the following parameters when you enter the `acl-vlan-group` command:

- Whether the CAM profile is set in VFP
- Whether the maximum number of groups in the system has exceeded
- Whether the maximum number of VLAN numbers permitted per ACL group has 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 as 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 any 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, to which the ACL VLAN group is applied, function as restricted interfaces. The ACL VLAN group name is used to identify the group of VLANs that is used to perform 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, a validation is performed to determine whether an 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) can be allocated virtual flow processing slices at a time.
- The maximum number of VLANs that you can configure as a member of ACL VLAN groups is limited to 512 on the switch if two slices are allocated. If only one virtual flow processing slice is allocated,

the maximum number of VLANs that you can configure as a member of an ACL VLAN group is 256 for the switch.

- Port ACL optimization is applicable only for ACLs that are applied without the VLAN range.
- You cannot view the statistical details of ACL rules per VLAN and per interface if you enable the ACL VLAN group capability. You can 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 effectively saving CAM space. The optimization is enabled only if you specify the optimized option with the `ip access-group` command. This option is not valid for VLAN and LAG interfaces.

Configuring ACL VLAN Groups and Configuring FP Blocks for VLAN Parameters

This section describes how to optimize the utilization of CAM blocks by configuring ACL VLAN groups that you can attach to VLAN interfaces and also 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. If you apply an ACL separately on the VLAN interface, each ACL has a mapping with the VLAN and increases the CAM space utilization. Attaching an ACL individually to VLAN interfaces is similar to the behavior of ACL-VLAN mapping storage in CAM prior to the implementation of the ACL VLAN group functionality.

1. Create an ACL VLAN group
   CONFIGURATION mode

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

   You can have up to eight different ACL VLAN groups at any given time.
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}
```

```
Dell#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
Dell#
```

Configuring FP Blocks for VLAN Parameters

Use the `cam-acl-vlan` command to allocate the number of FP blocks for the various VLAN processes on the system. You can use the **no** version of this command to reset the number of FP blocks to default. By default, 0 groups are allocated for the ACL in VCAP. ACL VLAN groups or CAM optimization is not enabled by default, and you need to allocate the slices for CAM optimization.

1. Allocate the number of FP blocks for VLAN Open Flow 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 feature.
   CONFIGURATION mode

```
cam-acl-vlan vlanaclopt <0-2>
```

4. View the number of flow processor (FP) blocks that is allocated for the different VLAN services.
   EXEC Privilege mode

```
Dell#show cam-usage switch

Linecard|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available
CAM
========|========|================|=============|=============|
```

```
==============
   11   |   0    | IN-L2 ACL       |      7152   |          0   |
7152
        |        | IN-L2 FIB       |     32768   |       1081   |
31687
        |        | OUT-L2 ACL      |         0   |          0   |
0
   11   |   1    | IN-L2 ACL       |      7152   |          0   |
7152
        |        | IN-L2 FIB       |     32768   |       1081   |
31687
        |        | OUT-L2 ACL      |         0   |          0   |          0
```

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 sample output shows the consumption of CAM blocks for Layer 2 and Layer 3 ACLs, in addition to other processes that use CAM space:

```
Dell#show cam-usage
Linecard|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available CAM
========|========|=================|=============|=============|==============
    1   |   0    | IN-L2 ACL       |      1008   |       320   |       688
        |        | IN-L2 FIB       |     32768   |      1132   |     31636
        |        | IN-L3 ACL       |     12288   |         2   |     12286
        |        | IN-L3 FIB       |    262141   |        14   |    262127
        |        | IN-L3-SysFlow   |      2878   |        45   |      2833
        |        | IN-L3-TrcList   |      1024   |         0   |      1024
        |        | IN-L3-McastFib  |      9215   |         0   |      9215
        |        | IN-L3-Qos       |      8192   |         0   |      8192
        |        | IN-L3-PBR       |      1024   |         0   |      1024
        |        | IN-V6 ACL       |         0   |         0   |         0
        |        | IN-V6 FIB       |         0   |         0   |         0
        |        | IN-V6-SysFlow   |         0   |         0   |         0
        |        | IN-V6-McastFib  |         0   |         0   |         0
        |        | OUT-L2 ACL      |      1024   |         0   |      1024
        |        | OUT-L3 ACL      |      1024   |         0   |      1024
        |        | OUT-V6 ACL      |         0   |         0   |         0
    1   |   1    | IN-L2 ACL       |       320   |         0   |       320
        |        | IN-L2 FIB       |     32768   |      1136   |     31632
        |        | IN-L3 ACL       |     12288   |         2   |     12286
        |        | IN-L3 FIB       |    262141   |        14   |    262127
        |        | IN-L3-SysFlow   |      2878   |        44   |      2834
--More--
```

The following sample output displays the CAM space utilization when Layer 2 and Layer 3 ACLs are configured:

```
Dell#show cam-usage acl
Linecard|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available CAM
========|========|=================|=============|=============|============
   11   |   0    | IN-L2 ACL       |      1008   |         0    |      1008
```

```
        |         | IN-L3 ACL       |      12288   |          2    |      12286
        |         | OUT-L2 ACL      |       1024   |          2    |       1022
        |         | OUT-L3 ACL      |       1024   |          0    |       1024
```

The following sample output displays the CAM space utilization for Layer 2 ACLs:

```
Dell#show cam-usage switch

Linecard|Portpipe| CAM Partition    | Total CAM   |  Used CAM   |Available CAM
========|========|==================|=============|=============|==============
   11   |   0    | IN-L2 ACL        |       7152  |          0  |       7152
        |        | IN-L2 FIB        |      32768  |       1081  |      31687
        |        | OUT-L2 ACL       |          0  |          0  |          0
   11   |   1    | IN-L2 ACL        |       7152  |          0  |       7152
        |        | IN-L2 FIB        |      32768  |       1081  |      31687
        |        | OUT-L2 ACL       |          0  |          0  |          0
```

The following sample output displays the CAM space utilization for Layer 3 ACLs:

```
Dell#show cam-usage router
Linecard|Portpipe| CAM Partition    | Total CAM   |  Used CAM   |Available CAM
========|========|==================|=============|=============|==============
   11   |   0    | IN-L3 ACL        |       8192  |          3  |       8189
        |        | IN-L3 FIB        |     196607  |          1  |     196606
        |        | IN-L3-SysFlow    |       2878  |          0  |       2878
        |        | IN-L3-TrcList    |       1024  |          0  |       1024
        |        | IN-L3-McastFib   |       9215  |          0  |       9215
        |        | IN-L3-Qos        |       8192  |          0  |       8192
        |        | IN-L3-PBR        |       1024  |          0  |       1024
        |        | OUT-L3 ACL       |      16384  |          0  |      16384
   11   |   1    | IN-L3 ACL        |       8192  |          3  |       8189
        |        | IN-L3 FIB        |     196607  |          1  |     196606
        |        | IN-L3-SysFlow    |       2878  |          0  |       2878
        |        | IN-L3-TrcList    |       1024  |          0  |       1024
        |        | IN-L3-McastFib   |       9215  |          0  |       9215
        |        | IN-L3-Qos        |       8192  |          0  |       8192
        |        | IN-L3-PBR        |       1024  |          0  |       1024
        |        | OUT-L3 ACL       |      16384  |          0  |      16384
```

Allocating FP Blocks for VLAN Processes

The VLAN ContentAware Processor (VCAP) application is a preingress CAP that modifies the VLAN settings before packets are forwarded. To support the ACL CAM optimization functionality, the CAM carving feature is enhanced. A total of four VACP groups are present, of which two are for fixed groups and the other two are for dynamic groups. Out of the total of two dynamic groups, you can allocate zero, one, or two FP blocks to iSCSI Counters, OpenFlow 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 feature, 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, and you need to allocate the slices for CAM optimization.

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

8

Access Control Lists (ACLs)

This chapter describes access control lists (ACLs), prefix lists, and route-maps.

At their simplest, 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 the Content Addressable Memory (CAM) chapter.

IP Access Control Lists (ACLs)

In Dell 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 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 Networking operating system 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.

NOTE: Hot lock ACLs are supported for Ingress ACLs only.

Implementing ACL on the Dell Networking OS

You can assign one IP ACL per interface with the Dell Networking OS. If you do not assign an IP ACL to an interface, it is not used by the software in any other capacity.

The number of entries allowed per ACL is hardware-dependent. For detailed specification on entries allowed per ACL, refer to your line card documentation.

If you enable counters on IP ACL rules that are already configured, those counters are reset when a new rule is inserted or prepended. If a rule is appended, the existing counters are not affected. This is applicable to the following features:

- L2 Ingress Access list
- L2 Egress Access list
- L3 Ingress Access list
- L3 Egress Access list

NOTE: IP ACLs are supported over VLANs in the Dell Networking OS version 6.2.1.1 and higher.

ACLs and VLANs

There are some differences when assigning ACLs to a VLAN rather than a physical port.

For example, when using a single port-pipe, if you apply an ACL to a VLAN, one copy of the ACL entries is installed in the ACL CAM on the port-pipe. The entry looks for the incoming VLAN in the packet. Whereas if 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.

ACL Optimization

If an access list contains duplicate entries, the system 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 whether it is identified as 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, the system 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 such as these, 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. The Dell 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

```
Dell(conf)#ip access-list standard acl1
Dell(config-std-nacl)#permit 20.0.0.0/8
Dell(config-std-nacl)#exit
Dell(conf)#ip access-list standard acl2
Dell(config-std-nacl)#permit 20.1.1.0/24 order 0
Dell(config-std-nacl)#exit
Dell(conf)#class-map match-all cmap1
Dell(conf-class-map)#match ip access-group acl1
Dell(conf-class-map)#exit
Dell(conf)#class-map match-all cmap2
Dell(conf-class-map)#match ip access-group acl2
Dell(conf-class-map)#exit
Dell(conf)#policy-map-input pmap
Dell(conf-policy-map-in)#service-queue 7 class-map cmap1
Dell(conf-policy-map-in)#service-queue 4 class-map cmap2
Dell(conf-policy-map-in)#exit
Dell(conf)#interface gig 1/0
Dell(conf-if-gi-1/0)#service-policy input pmap
```

IP Fragment Handling

The Dell Networking OS supports a configurable option to explicitly deny IP fragmented packets, especially 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, the system always applies implicit deny. You do not have to configure it.
- For IP ACL, the system applies implicit permit for second and subsequent fragment 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.

IP Fragments ACL Examples

The following examples show how you can use ACL commands with the `fragment` keyword to filter fragmented packets.

Example of Permitting All Packets on an Interface

Example of Denying Second and Subsequent Fragments

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.

```
Dell(conf)#ip access-list extended ABC
Dell(conf-ext-nacl)#permit ip any 10.1.1.1/32
Dell(conf-ext-nacl)#deny ip any 10.1.1.1./32 fragments
Dell(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.

```
Dell(conf)#ip access-list extended ABC
Dell(conf-ext-nacl)#deny ip any 10.1.1.1/32 fragments
Dell(conf-ext-nacl)#permit ip any 10.1.1.1/32
Dell(conf-ext-nacl)
```

Layer 4 ACL Rules Examples

The following examples show the ACL commands for Layer 4 packet filtering.

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.

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.

Example of Layer 4 ACL Rules

Example of TCP Packets

In this first example, fragments or non-fragmented TCP packets from 10.1.1.1 with TCP destination port equal to 24 are permitted. All other fragments are denied.

```
Dell(conf)#ip access-list extended ABC
Dell(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
Dell(conf-ext-nacl)#deny ip any any fragment
Dell(conf-ext-nacl)
```

In the following example, 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.

```
Dell(conf)#ip access-list extended ABC
Dell(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
Dell(conf-ext-nacl)#permit tcp host 10.1.1.1 any fragment
Dell(conf-ext-nacl)#deny ip any any fragment
Dell(conf-ext-nacl)
```

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 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]] [order] [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.

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.

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.

Example of Viewing the Rules of a Specific ACL on an Interface

Example of the `seq` Command to Order Filters

```
Dell#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
Dell#
```

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.

```
Dell(conf-std-nacl)#seq 25 deny ip host 10.5.0.0 any
Dell(conf-std-nacl)#seq 15 permit tcp 10.3.0.0 /16 any
Dell(conf-std-nacl)#show config
!
ip access-list standard dilling
  seq 15 permit tcp 10.3.0.0/16 any
Dell(conf-std-nacl)#
```

To delete a filter, use the `no seq sequence-number` command in IP ACCESS LIST mode.

Configuring a Standard IP ACL Filter

If you are creating a standard ACL with only one or two filters, you can let the system ssign 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]] [order] [fragments]`

The following example shows a standard IP ACL in which the system 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.

Example of Viewing Filter Sequence for a Specified Standard ACL

Example of Viewing Standard ACL Filter Sequence for an Interface

```
Dell(config-route-map)#ip access standard kigali
Dell(config-std-nacl)#permit 10.1.0.0/16
Dell(config-std-nacl)#show config
!
ip access-list standard kigali
  seq 5 permit 10.1.0.0/16
Dell(config-std-nacl)#
```

To view all configured IP ACLs, use the `show ip accounting access-list` command in EXEC Privilege mode.

```
Dell#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
  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
  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.

Because traffic passes through the filter in the order of the filter's sequence, 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 port [port]] [count [byte]] [order] [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 might need to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five or another number.

The following examples 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.

```
Dell(conf-ext-nacl)#seq 15 deny ip host 112.45.0.0 any
Dell(conf-ext-nacl)#seq 5 permit tcp 12.1.3.45 255.255.0.0 any
Dell(conf-ext-nacl)#show config
!
ip access-list extended dilling
  seq 5 permit tcp 12.1.0.0 255.255.0.0 any
  seq 15 deny ip host 112.45.0.0 any
Dell(conf-ext-nacl)#
```

Configuring Filters Without a Sequence Number

If you are creating an extended ACL with only one or two filters, you can let the system assign a sequence number based on the order in which the filters are configured. The system 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] [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] [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] [fragments]`

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.

Example of Viewing Filter Sequence for a Specified Extended ACL

```
Dell(config-ext-nacl)#deny tcp host 123.55.34.0 any
Dell(config-ext-nacl)#permit udp 154.44.123.34 0.0.255.255 host 34.6.0.0
Dell(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
Dell(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 Configuring a Standard IP ACL Filter.

Established Flag

To obtain the functionality of **est**, use the following ACLs:

- `permit tcp any any rst`
- `permit tcp any any ack`

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 the system routes the packets, only the L3 ACL governs them because they are not filtered against an L2 ACL.
- When the system switches the packets, first the L3 ACL filters them, then the L2 ACL filters them.
- When the system switches the packets, the egress L3 ACL does not filter 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
- L3 ingress access list

If a rule is simply appended, existing counters are not affected.

Table 4. 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.

For more information about Layer-3 interfaces, refer to Interfaces.

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 | out} [implicit-permit] [vlan vlan-range]
   ```

   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.
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.

Example of Viewing ACLs Applied to an Interface

```
Dell(conf-if)#show conf
!
interface GigabitEthernet 0/0
  ip address 10.2.1.100 255.255.255.0
  ip access-group nimule in
  no shutdown
Dell(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.
You can configure either count (packets) or count (bytes). However, for an ACL with multiple rules, you can configure some ACLs with count (packets) and others as count (bytes) at any given time.

1. Create an ACL that uses rules with the count option. Refer to Configuring a Standard IP ACL Filter.
2. Apply the ACL as an inbound or outbound ACL on an interface. Refer to Assign an IP ACL to an Interface.
3. View the number of packets matching the ACL.

EXEC Privilege mode

`show ip accounting access-list`

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.

```
Dell(conf)#interface tengig 0/0
Dell(conf-if-tengig0/0)#ip access-group abcd in
Dell(conf-if-tengig0/0)#show config
!
tengigethernet 0/0
  no ip address
  ip access-group abcd in
  no shutdown
Dell(conf-if-tengig0/0)#end
Dell#configure terminal
Dell(conf)#ip access-list extended abcd
Dell(conf-ext-nacl)#permit tcp any any
Dell(conf-ext-nacl)#deny icmp any any
Dell(conf-ext-nacl)#permit 1.1.1.2
Dell(conf-ext-nacl)#end
Dell#show ip accounting access-list
!
Extended Ingress IP access list abcd on tengigethernet 0/0
  seq 5 permit tcp any any
  seq 10 deny icmp any any
    seq 15 permit 1.1.1.2
```

Configure Egress ACLs

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 direct operating system (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.

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.

```
Dell(conf)#interface tengig 0/0
Dell(conf-if-tengig0/0)#ip access-group abcd out
Dell(conf-if-tengig0/0)#show config
!
tengigethernet 0/0
  no ip address
  ip access-group abcd out
  no shutdown
Dell(conf-if-tengig0/0)#end
Dell#configure terminal
Dell(conf)#ip access-list extended abcd
Dell(conf-ext-nacl)#permit tcp any any
Dell(conf-ext-nacl)#deny icmp any any
Dell(conf-ext-nacl)#permit 1.1.1.2
Dell(conf-ext-nacl)#end
Dell#show ip accounting access-list
!
Extended Ingress IP access list abcd on tengigethernet 0/0
  seq 5 permit tcp any any
  seq 10 deny icmp any any
     seq 15 permit 1.1.1.2
```

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.

1. Apply Egress ACLs to IPv4 system traffic.
   CONFIGURATION mode

   ```
   ip control-plane [egress filter]
   ```

2. 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
   ```

Dell 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, the system 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 the Dell 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: Dell Networking OS does not support all protocols. It is important to know which protocol you are supporting 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 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).

Example of Assigning Sequence Numbers to Filters

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.

```
Dell(conf-nprefixl)#seq 20 permit 0.0.0.0/0 le 32
Dell(conf-nprefixl)#seq 12 deny 134.23.0.0 /16
Dell(conf-nprefixl)#seq 15 deny 120.23.14.0 /8 le 16
Dell(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
Dell(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 the system assign a sequence number based on the order in which the filters are configured. The system 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 (from 0 to 32).
   - `le max-prefix-length`: is the maximum prefix length to be matched (from 0 to 32).

Example of Creating a Filter with a Dell Networking OS-Assigned Sequence Numbers

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 the two filters with the sequence numbers 5 and 10.

```
Dell(conf-nprefixl)#permit 123.23.0.0 /16
Dell(conf-nprefixl)#deny 133.24.56.0 /8
Dell(conf-nprefixl)#show conf
!
ip prefix-list awe
  seq 5 permit 123.23.0.0/16
  seq 10 deny 133.0.0.0/8
Dell(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]
  ```

Example of the `show ip prefix-list detail` Command

Example of the `show ip prefix-list summary` Command

```
Dell>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)

Dell>
Dell>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
Dell>
```

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]
  ```

Example of Viewing Configured Prefix Lists (ROUTER RIP mode)

To view the configuration, use the `show config` command in ROUTER RIP mode, or the `show running-config rip` command in EXEC mode.

```
Dell(conf-router_rip)#show config
!
router rip
  distribute-list prefix juba out
  network 10.0.0.0
Dell(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]
  ```

Example of Viewing Configured Prefix Lists (ROUTER OSPF mode)

To view the configuration, use the `show config` command in ROUTER OSPF mode, or the `show running-config ospf` command in EXEC mode.

```
Dell(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
Dell(conf-router_ospf)#
```

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 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 5. 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
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 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.

- Resequence an IPv4 or MAC ACL.
  EXEC mode

  `resequence access-list {ipv4 | mac} {`*`access-list-name StartingSeqNum Step-to-Increment`*`}`
- Resequence an IPv4 prefix-list.
  EXEC mode

  `resequence prefix-list {ipv4} {`*`prefix-list-name StartingSeqNum Step-to-Increment`*`}`

Example of Resequencing ACLs When Remarks and Rules Have the Same Number

Example of Resequencing ACLs When Remarks and Rules Have Different Numbers

The example shows the resequencing of an IPv4 access-list beginning with the number 2 and incrementing by 2.

Remarks and rules that originally have the same sequence number have the same sequence number after you apply the `resequence` command.

```
Dell(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
Dell# end
Dell# resequence access-list ipv4 test 2 2
Dell# 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.

```
Dell(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
Dell# end
Dell# resequence access-list ipv4 test 2 2
Dell# 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

Similar to ACLs and prefix lists, route maps are composed of a series of commands that contain a matching criterion and an action; however, route maps can change the packets meeting the criterion.

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.

Implementation Information

The Dell Networking OS 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 specify 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 action 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.

Example of Viewing a Configured Route Map

Example of Multiple Instances of a Route-Map

Example of Deleting One Instance of a Route Map

Example of Viewing All Instances of a Specified Route Map

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.

```
Dell(config-route-map)#show config
!
route-map dilling permit 10
Dell(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. The system 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.

```
Dell#show route-map
route-map zakho, permit, sequence 10
  Match clauses:
  Set clauses:
route-map zakho, permit, sequence 20
  Match clauses:
    interface TenGigabitEthernet 0/1
  Set clauses:
    tag 35
    level stub-area
Dell#
```

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.

```
Dell(conf)#no route-map zakho 10
Dell(conf)#end
Dell#show route-map
route-map zakho, permit, sequence 20
  Match clauses:
    interface TenGigabitEthernet 0/1
  Set clauses:
    tag 35
    level stub-area
Dell#
```

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.

```
Dell#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
Dell#
```

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, the system does a match between all of those `match` commands. If there are multiple `match` commands with different parameters, the system 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

Example of the `match` Command to Match All Specified Values

```
Dell(conf)#route-map force permit 10
Dell(config-route-map)#match tag 1000
Dell(config-route-map)#match tag 2000
Dell(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.

```
Dell(conf)#route-map force permit 10
Dell(config-route-map)#match tag 1000
Dell(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, the system 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

```
Dell(conf)#route-map force permit 10
Dell(config-route-map)#match tag 1000

Dell(conf)#route-map force deny 20
Dell(config-route-map)#match tag 1000

Dell(conf)#route-map force deny 30
Dell(config-route-map)#match tag 1000
```

Configuring Match Routes

To configure match criterion for a route map, use the following commands.

- Match routes whose next hop is a specific interface.
  CONFIG-ROUTE-MAP mode

  ```
  match interface interface
  ```

  The parameters are:

  - For a Loopback interface, enter the keyword `loopback` then a number between zero (0) and 16383.
  - For a 10-Gigabit Ethernet interface, enter the keyword `tengigabitEthernet` then the slot/port information.
  - For a VLAN, 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 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 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 routes with a specific value.
  CONFIG-ROUTE-MAP mode

  ```
  match metric metric-value
  ```
- 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.

- 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 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
  ```
- Specify a tag for the redistributed routes.
  CONFIG-ROUTE-MAP mode

  ```
  set tag tag-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. To apply a route map to traffic, you must call or include that route map in a command such as the redistribute or default-information originate commands in OSPF and BGP.

Route redistribution occurs when the system 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 Gigabitethernet interface 0/0 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 GigabitEthernet 0/0
  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

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 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

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

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 Support for ACLs

Flow-based monitoring conserves bandwidth by monitoring only the specified traffic instead of all traffic on the interface. It is available for Layer 2 and Layer 3 ingress traffic. You can specify traffic using standard or extended access-lists. This feature copies all incoming packets on one 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).

The port mirroring application maintains and performs all the monitoring operations on the chassis. ACL information is sent to the ACL manager, which in turn notifies the ACL agent to add entries in the CAM area. Duplicate entries in the ACL are not saved.

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.

When a stack unit is reset or a stack unit undergoes a failure, the ACL agent registers with the port mirroring application. The port mirroring utility downloads the monitoring configuration to the ACL agent. The interface manager notifies the port mirroring application about the removal of an interface when an ACL entry associated with that interface to is deleted.

Behavior of Flow-Based Monitoring

Activate flow-based monitoring for a monitoring session by entering the `flow-based enable` command in the Monitor Session mode. When you enable this capability, traffic with particular flows that are traversing through the ingress interfaces are examined, and appropriate ACLs can be applied in the ingress direction. By default, flow-based monitoring is not enabled.

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 will be 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]] [monitor]
```

If the number of monitoring sessions increases, inter-process communication (IPC) bandwidth utilization will be high. The ACL manager might require a large bandwidth when you assign an ACL, with many entries, to an interface.

The ACL agent module saves monitoring details in its local database and also in the CAM region to monitor packets that match the specified criterion. The ACL agent maintains data on the source port, the destination port, and the endpoint to which the packet must be forwarded when a match occurs with the ACL entry.

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. Flow-based monitoring is supported only for ingress traffic and not for egress packets.

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 has been enhanced to display the Type field in the output, which indicates whether a particular session is enabled for flow-monitoring.

Example Output of the `show` Command

```
#show running-config monitor session
!
monitor session 11
 flow-based enable
 source GigabitEthernet 13/0 destination GigabitEthernet 13/1 direction both
```

The `show running-config monitor session` command displays whether flow-based monitoring is enabled for a particular session.

The `show config` command has been modified to display monitoring configuration in a particular session.

Example Output of the `show` Command

```
(conf-mon-sess-11)#show config
    !
    monitor session 11
                          flow-based enable
      source GigabitEthernet 13/0 destination GigabitEthernet 13/1 direction
both
```

The `show ip | mac | ipv6 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

```
Dell# show ip accounting access-list
!
Extended Ingress IP access list kar on GigabitEthernet 10/0
Total cam count 1
   seq 5 permit ip 192.168.20.0/24 173.168.20.0/24 monitor

Dell#show mac accounting access-list kar in gi 10/0 out
Egress Extended mac access-list kar on GigabitEthernet 10/0
 seq 5  permit host  11:11:11:11:11:11 host  22:22:22:22:22:22 monitor
     seq 10  permit host  22:22:22:22:22:22 any monitor
     seq 15  permit host  00:0f:fe:1e:de:9b host  0a:0c:fb:1d:fc:aa monitor

Dell#show ipv6 accounting access-list
    !
    Ingress IPv6 access list kar on GigabitEthernet 10/0
    Total cam count 1
     seq 5 permit ipv6 22::/24 33::/24 monitor
```

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 2 and Layer 3 ingress and egress 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
   ```

Example of the `flow-based enable` Command

To view an access-list that you applied to an interface, use the `show ip accounting access-list` command from EXEC Privilege mode.

```
Dell(conf)#monitor session 0
Dell(conf-mon-sess-0)#flow-based enable
Dell(conf)#ip access-list ext testflow
Dell(config-ext-nacl)#seq 5 permit icmp any any count bytes monitor
Dell(config-ext-nacl)#seq 10 permit ip 102.1.1.0/24 any count bytes monitor
Dell(config-ext-nacl)#seq 15 deny udp any any count bytes
Dell(config-ext-nacl)#seq 20 deny tcp any any count bytes
Dell(config-ext-nacl)#exit
```

```
Dell(conf)#interface gig 1/1
Dell(conf-if-gi-1/1)#ip access-group testflow in
Dell(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
  ip address 10.11.1.254/24
  ip access-group testflow in
  shutdown
Dell(conf-if-gi-1/1)#exit
Dell(conf)#do show ip accounting access-list testflow
!
Extended Ingress IP access list testflow on GigabitEthernet 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)
Dell(conf)#do show monitor session 0
SessionID Source Destination Direction Mode      Type
--------- ------ ----------- --------- ----      ----
0         Gi 1/1 Gi 1/2      rx        interface Flow-based
```

9
Bidirectional Forwarding Detection (BFD)

Bidirectional forwarding detection (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 Networking routers, BFD agents maintain sessions that reside on the line card, which frees resources on the route processor module (RPM). Only session state changes are reported to the BFD Manager (on the RPM), 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 Networking has implemented BFD at Layer 3 and with user datagram protocol (UDP) encapsulation. BFD functionality will be implemented in phases. OSPF, IS-IS, VRRP, VLANs, LAGs, static routes, and physical ports support BFD, based on the IETF internet *draft draft-ietf-bfd-base-03*.

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 Networking operating system 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.

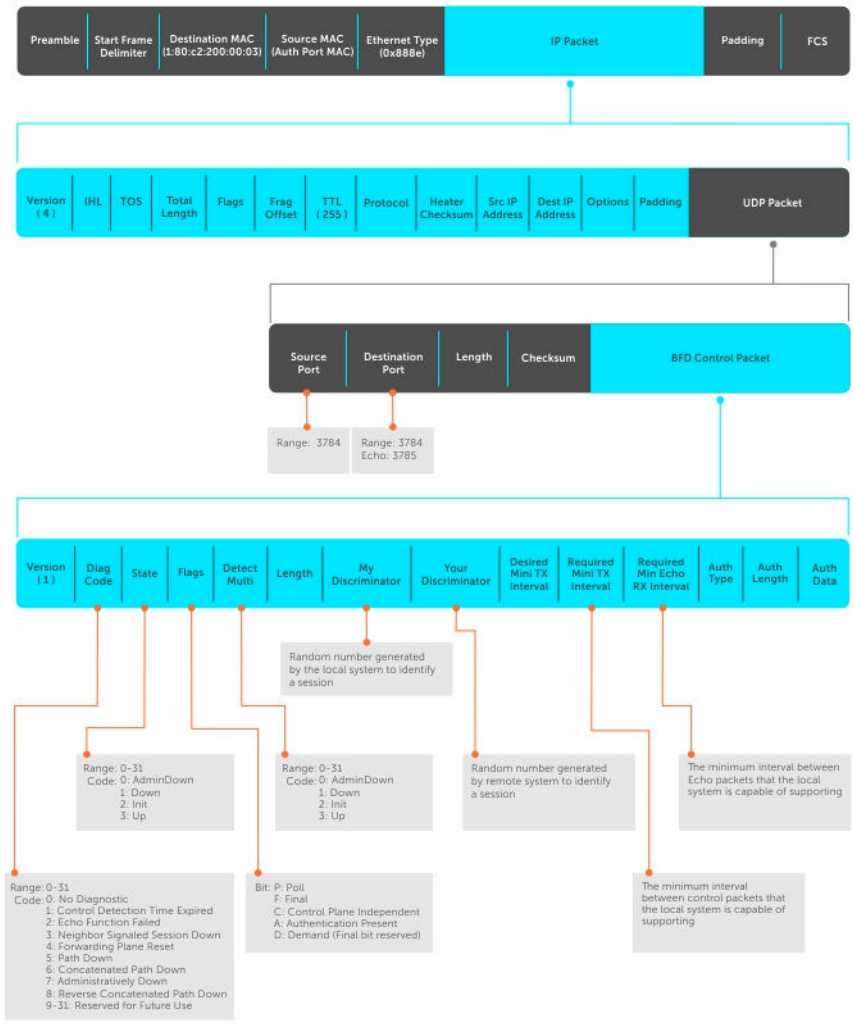


Figure 8. BFD in IPv4 Packet Format

Field	Description
Diagnostic Code	The reason that the last session failed.
State	The current local session state. Refer to BFD Sessions.

Field	Description
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:** The Dell 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.
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:** The Dell Networking OS does not currently support the echo function.
Authentication Type, Authentication Length, Authentication Data	An optional method for authenticating control packets. **NOTE:** The Dell 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

You must enable BFD 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: The Dell Networking OS supports Asynchronous mode only.

A session can have four states: Administratively Down, Down, Init, and Up.

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.
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.

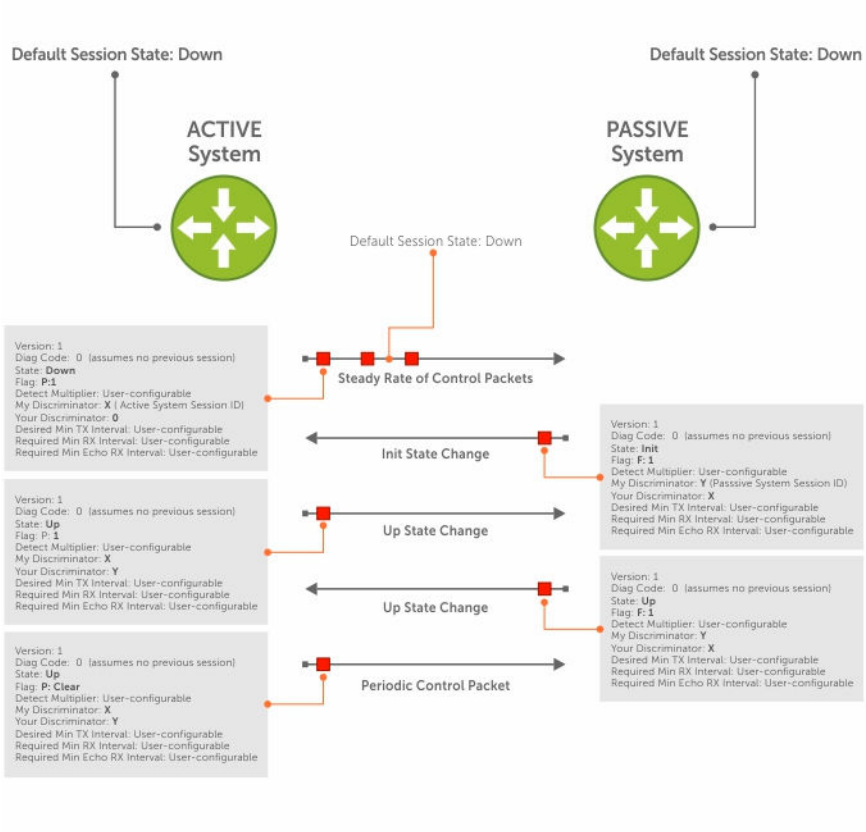


Figure 9. BFD Three-Way Handshake State Changes

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.

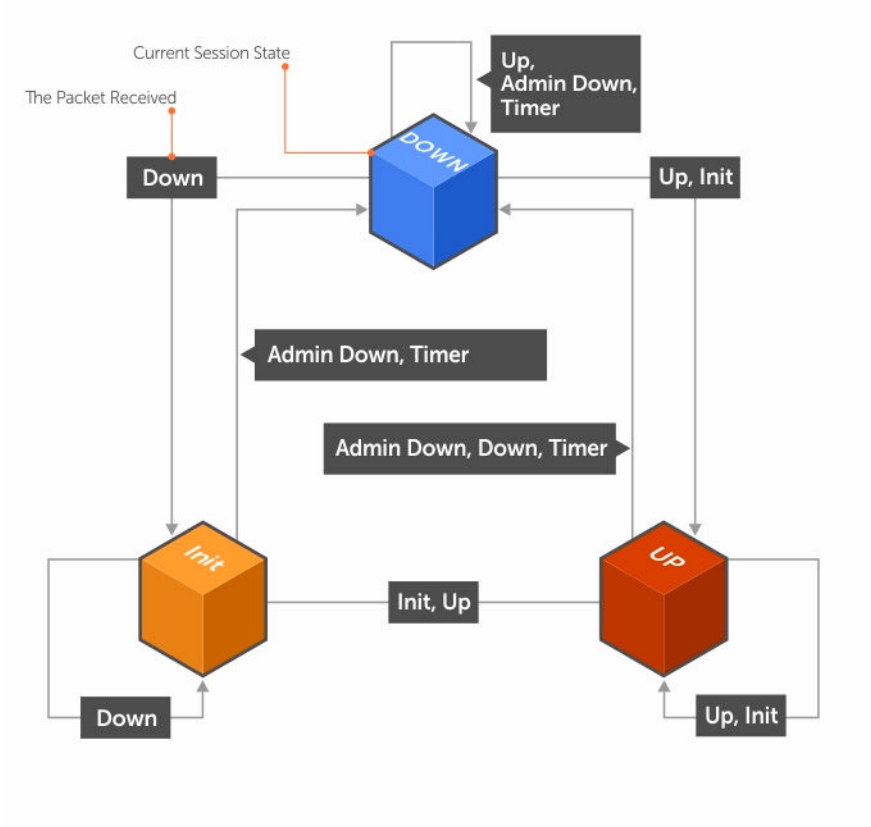


Figure 10. Session State Changes

Important Points to Remember

- BFD for line card ports is hitless, but is not hitless for VLANs because they are instantiated on the RPM.
- The Dell Networking OS supports a maximum of 100 sessions per BFD agent. Each linecard processor has a BFD Agent, so the limit translates to 100 BFD sessions per linecard.
- 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 and virtual links.
- Protocol Liveness is supported for routing protocols only.
- The Dell Networking OS supports only OSPF, OSPFv3, BGP, and VRRP protocols as BFD clients.

Configure BFD

This section contains the following procedures.

- Configure BFD for Physical Ports
- Configure BFD for Port-Channels
- Configure BFD for Static Routes

- Configure BFD for OSPF
- Configure BFD for OSPFv3
- Configure BFD for BGP
- Configure BFD for VRRP
- Configure BFD for VLANs
- Configuring Protocol Liveness
- Troubleshooting BFD

Configure BFD for Physical Ports

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. Refer to Enabling BFD Globally.
2. Establish a session with a next-hop neighbor.

Related Configuration Tasks

- Changing 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
  ```

Example of Verifying BFD is Enabled

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)#
```

Establishing a Session on Physical Ports

To establish a session, enable BFD at the interface level on both ends of the link, as shown in the following illustration. The configuration parameters do not need to match.

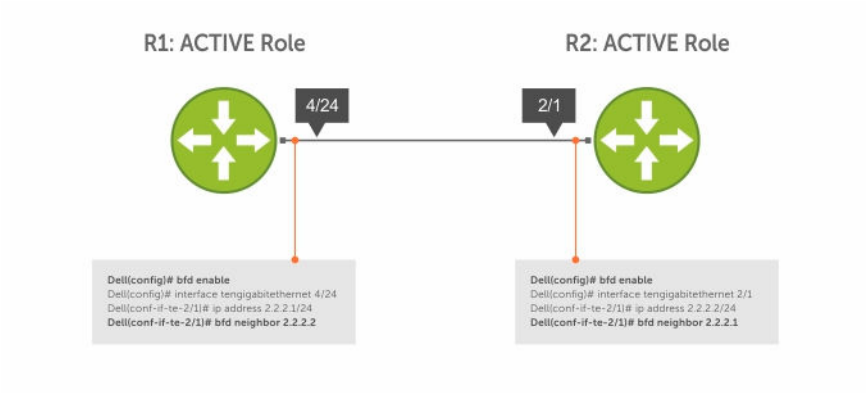


Figure 11. Establishing a BFD Session on Physical Ports

1. Enter interface mode.
   CONFIGURATION mode

   ```
   interface
   ```
2. Assign an IP address to the interface if one is not already assigned.
   INTERFACE mode

   ```
   ip address ip-address
   ```
3. Identify the neighbor that the interface participates with the BFD session.
   INTERFACE mode

   ```
   bfd neighbor ip-address
   ```

Example of Verifying Session Creation

Example of the Viewing Detailed BFD Session Information

To verify that the session is established, use the `show bfd neighbors` command.

The bold line shows the BFD session.

```
R1(conf-if-gi-4/24)#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     Gi 4/24    Up     200     200     3     C
```

To view specific information about BFD sessions, use the `show bfd neighbors detail` command.

```
R1(conf-if-gi-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: GigabitEthernet 4/24
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: False
Client Registered: CLI
Uptime: 00:03:57
Statistics:
  Number of packets received from neighbor: 1775
  Number of packets sent to neighbor: 1775
  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
```

Log messages display when you configure both interfaces for BFD.

```
     R1(conf-if-gi-4/24)#00:36:01: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE:
Changed session state to
Down for neighbor 2.2.2.2 on interface Gi 4/24 (diag: 0)
     00:36:02: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to
Up for neighbor
2.2.2.2 on interface Gi 4/24 (diag: 0)
```

Changing Physical Port Session Parameters

Configure BFD sessions with default intervals and a default role (active).
The parameters that you can configure are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. Configure these parameters per interface; if you change a parameter, the change affects all physical port sessions on that interface.

NOTE: Dell Networking recommends maintaining the default values.

Change session parameters for all sessions on an interface.

INTERFACE mode

```
bfd interval milliseconds min_rx milliseconds multiplier value role [active
| passive]
```

Changing Session Parameters for Physical Ports

View session parameters using the `show bfd neighbors detail` command.

The bold line shows the parameter changes.

```
R1(conf-if-gi-4/24)#bfd interval 200 min_rx 200 multiplier 4 role passive
R1(conf-if-gi-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: GigabitEthernet 4/24
```

```
State: Up
Configured parameters:
  TX: 200ms, RX: 200ms, Multiplier: 4
Neighbor parameters:
  TX: 200ms, RX: 200ms, Multiplier: 3
Actual parameters:
  TX: 200ms, RX: 200ms, 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-gi-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 Gi 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 Gi 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. Refer to Enabling BFD Globally.

2. On the local system, establish a session with the next hop of a static route. Refer to Establishing Sessions for Static Routes.
3. On the remote system, establish a session with the physical port that is the origin of the static route. Refer to Establishing a Session on Physical Ports.

Related Configuration Tasks

- Changing Static Route Session Parameters
- Disabling BFD for Static Routes

Establishing Sessions for Static Routes

Sessions are established for all neighbors that are the next hop of a static route.

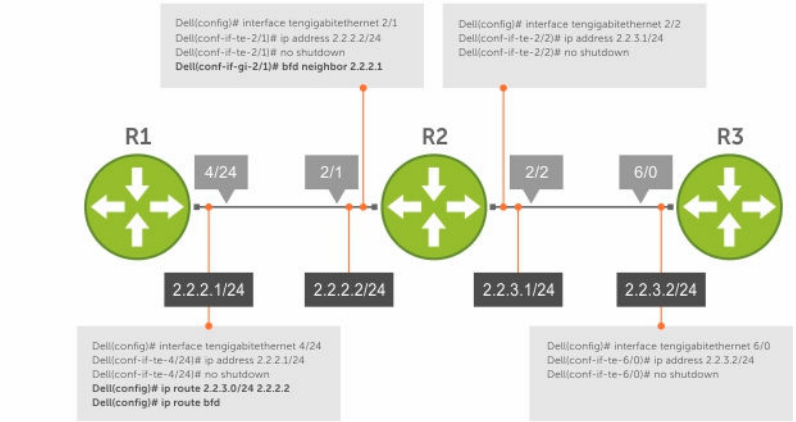


Figure 12. Establishing Sessions for Static Routes

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
  ```

Example of the `show bfd neighbors` Command to Verify Static Routes

To verify that sessions have been created for static routes, use the `show bfd neighbors` command. The bold line shows BFD for static routes is enabled.

```
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    Gi 4/24   Up    200    200    4    R
```

To view detailed session information, use the `show bfd neighbors detail` command, as shown in the examples in Disabling BFD for BGP.

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 interval milliseconds min_rx milliseconds multiplier value role
  [active | passive]
  ```

To view session parameters, use the `show bfd neighbors detail` command, as shown in the examples in Displaying BFD for BGP Information.

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
  ```

Configure BFD for OSPF

When using BFD with OSPF, the OSPF protocol registers with the BFD manager on the RPM.

BFD sessions are established with all neighboring interfaces participating in OSPF. If a neighboring interface fails, the BFD agent on the line card notifies the BFD manager, which in turn notifies the OSPF protocol that a link state change occurred.

Configuring BFD for OSPF is a two-step process:

1. Enable BFD globally. Refer to Enabling BFD Globally.
2. Establish sessions with OSPF neighbors. Refer to Establishing Sessions with OSPF Neighbors.

Related Configuration Tasks

- Changing OSPF Session Parameters
- Disabling BFD for OSPF

Establishing Sessions with OSPF Neighbors

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.

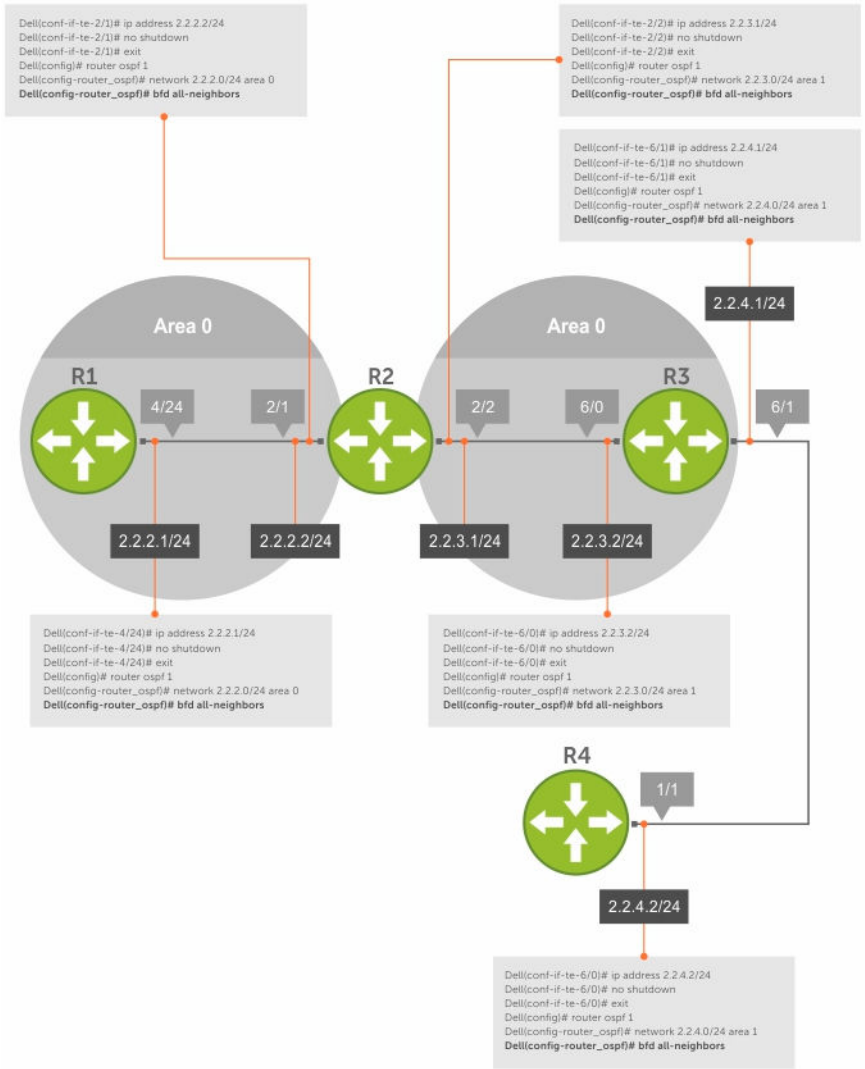


Figure 13. Establishing Sessions with OSPF Neighbors

To establish BFD with all OSPF neighbors or with OSPF neighbors on a single interface, use the following commands.

- 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
```

Example of Verifying Sessions with OSPF 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    Gi 2/1    Up    200    200    3    O
* 2.2.3.1  2.2.3.2    Gi 2/2    Up    200    200    3    O
```

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, as shown in the example in Displaying BFD for BGP Information.

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.

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, as shown in the example in Displaying BFD for BGP Information.

- 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 BGP

In a BGP core network, bidirectional forwarding detection (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 1GE, 10GE, 40GE, port-channel, and VLAN interfaces. BFD for BGP does not support IPv6 and the BGP multihop feature.

Prerequisites

Before configuring BFD for BGP, you must first configure the following settings:

1. Configure BGP on the routers that you want to interconnect, as described in Border Gateway Protocol IPv4 (BGPv4).
2. Enable fast fall-over for BGP neighbors to reduce convergence time (the `neighbor fall-over` command), as described in Configuring BGP Fast Fail-Over.

Establishing Sessions with BGP Neighbors

Before configuring BFD for BGP, you must first configure BGP on the routers that you want to interconnect.
For more information, refer to Border Gateway Protocol IPv4 (BGPv4).
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.

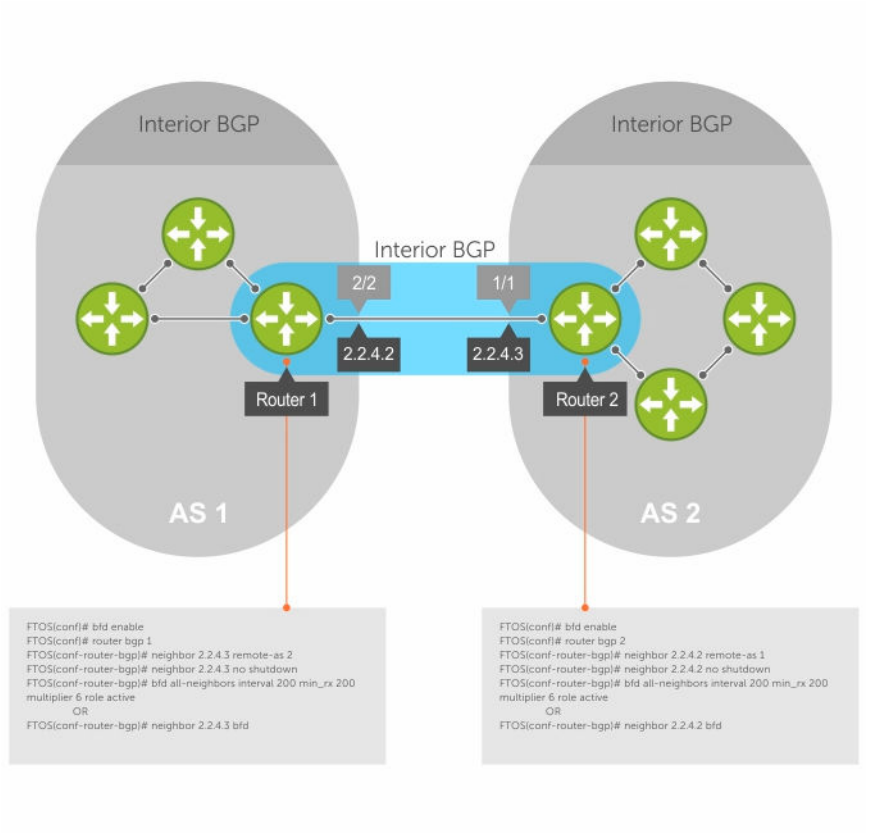


Figure 14. Establishing Sessions with BGP Neighbors

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 only in BGP IPv4 networks.

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 BGP 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.

You can configure BFD for BGP on the following types of interfaces: physical port (10GE or 40GE), port channel, and VLAN.

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. 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 [interval millisecs min_rx millisecs multiplier value role
   {active | passive}]
   ```

   OR

   ```
   neighbor {ip-address | peer-group-name} bfd
   ```

   NOTES:
   - When you establish a BFD session with a specified BGP neighbor or peer group using the `neighbor bfd` command, the default BFD session parameters are used (interval: 100 milliseconds, min_rx: 100 milliseconds, multiplier: 3 packets, and role: active).
   - When you explicitly enable or disable a BGP neighbor for a BFD session with the `neighbor bfd` or `neighbor bfd disable` commands, 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. Also, the neighbor only inherits the global timer values configured with the `bfd all-neighbors` command (interval, min_rx, and multiplier).
6. Repeat Steps 1 to 5 on each BGP peer participating in a BFD session.

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 `no neighbor {ip-address | peer-group-name} bfd disable` command in ROUTER BGP configuration mode.
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 | 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 | 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 Configuring 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]
  ```
- Display BFD packet counters for sessions with BGP neighbors.

  EXEC Privilege mode

  ```
  show bfd counters bgp [interface]
  ```
- 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]
  ```

Example of Verifying BGP Configuration

Example of Viewing All BFD Neighbors

Example of Viewing BFD Neighbor Detail

Example of Viewing Configured BFD Counters

Example of Viewing BFD Summary Information

Example of Viewing BFD Information for a Specified Neighbor

```
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

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 6/0    Up    200    200    3    B
* 2.2.2.3  2.2.2.2    Te 6/1    Up    200    200    3    B
* 3.3.3.3  3.3.3.2    Te 6/2    Up    200    200    3    B
```

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 6/0
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 6/1
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: 00:02:22
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

R2# show bfd counters bgp

Interface TenGigabitEthernet 6/0

Protocol BGP
Messages:
Registration    : 5
De-registration : 4
```

```
Init            : 0
Up              : 6
Down            : 0
Admin Down      : 2

Interface TenGigabitEthernet 6/1

Protocol BGP
Messages:
Registration    : 5
De-registration : 4
Init            : 0
Up              : 6
Down            : 0
Admin Down      : 2

Interface TenGigabitEthernet 6/2

Protocol BGP
Messages:
Registration    : 1
De-registration : 0
Init            : 0
Up              : 1
Down            : 0
Admin Down      : 2
```

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 bold lines show the message displayed when you enable a BFD session with different configurations:

- Message displayed 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 displayed when you enable a BFD session with a BGP neighbor using the `neighbor` *`ip-address`* `bfd` command.
- Message displayed 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_ExaScale#

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. Refer to Establishing VRRP Sessions on VRRP 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.

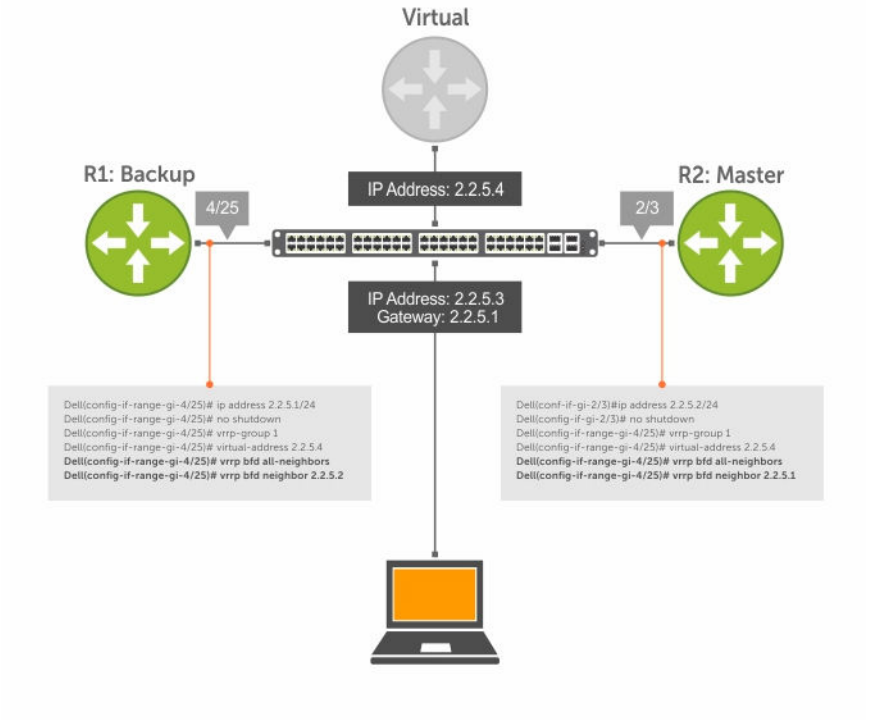


Figure 15. Establishing Sessions with All VRRP Neighbors

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
  ```

Example of Viewing Sessions with VRRP Neighbors

Example of Viewing VRRP Session State Information

To view the established sessions, use the `show bfd neighbors` command.

The bold line shows that VRRP BFD sessions are enabled.

```
R1(conf-if-gi-4/25)#vrrp bfd all-neighbors
R1(conf-if-gi-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     Gi 4/25  Down  200   200   3    V
```

To view session state information, use the `show vrrp` command.

The bold line shows the VRRP BFD session.

```
R1(conf-if-gi-4/25)#do show vrrp
------------------
GigabitEthernet 4/1, VRID: 1, Net: 2.2.5.1
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, as shown in the example in *Verifying BFD Sessions with BGP Neighbors Using the `show bfd neighbors` command* example in Displaying BFD for BGP Information.

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
```

Configure BFD for VLANs

BFD on Dell Networking systems is a Layer 3 protocol.

Use BFD with routed virtual local area networks (VLANs). BFD on VLANs is analogous to BFD on physical ports. If you enable the no routing protocol, and a remote system fails, the local system does not remove the connected route until the first failed attempt to send a packet. If you enable BFD, the local system removes the route when it stops receiving periodic control packets from the remote system.

There is one BFD agent for VLANs and port-channels that resides on RP2, as opposed to the other agents that are on the line card. Therefore, the 100 total possible sessions that this agent can maintain is shared for VLANs and port-channels.

Configuring BFD for VLANs is a two-step process:

1. Enable the BFD globally. Refer to Enabling BFD Globally.
2. Establish sessions with VLAN neighbors. Refer to Establish Sessions with VLAN Neighbors.

Related Configuration Task

- Changing VLAN Session Parameters.
- Disabling BFD for VLANs.

Establish Sessions with VLAN Neighbors

To establish a session, enable BFD at interface level on both ends of the link, as shown in the following illustration. The session parameters do not need to match.

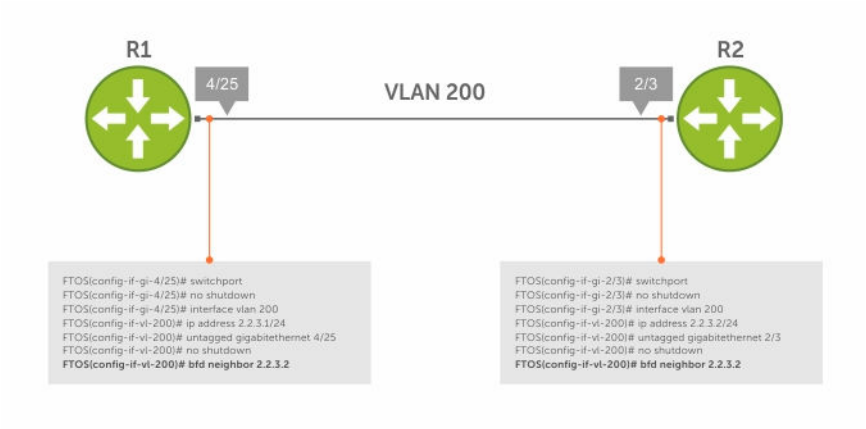


Figure 16. Establishing Sessions with VLAN Neighbors

To establish a BFD session with a VLAN neighbor, follow this step.

- Establish sessions with a VLAN neighbor.
  INTERFACE VLAN mode

  ```
  bfd neighbor ip-address
  ```

View the established sessions using the `show bfd neighbors` command, as shown in the following example.

```
R2(conf-if-vl-200)#bfd neighbor 2.2.3.2
R2(conf-if-vl-200)#do show bfd neighbors

* - 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
* 2.2.3.2 2.2.3.1    Vl 200    Up    200    200    3    C
```

Changing VLAN 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 per interface, if you make a configuration change, the change affects all sessions on that interface.

CAUTION: When configuring BFD on VLAN or LAG interfaces, Dell Networking recommends a minimum value of 500 milliseconds for both the transmit and minimum receive time, which yields a final detection time of (500ms *3) 1500 milliseconds.

To change parameters for a session, use the following commands.

- Change session parameters for all sessions on an interface.
  INTERFACE VLAN mode

  ```
  bfd interval milliseconds min_rx milliseconds multiplier value role [active | passive]
  ```

To view session parameters, use the `show bfd neighbors` command, as shown in the example Changing Physical Port Session Parameters.

Disabling BFD for VLANs

If you disable BFD on an interface, sessions on the interface 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 BFD on a VLAN interface, use the following command.

- Disable all sessions on a VLAN interface.
  INTERFACE VLAN mode

  ```
  no bfd enable
  ```

Configure BFD for Port-Channels

BFD on port-channels is analogous to BFD on physical ports.

If you enable the no routing protocol, and a remote system fails, the local system does not remove the connected route until the first failed attempt to send a packet. If you enable BFD, the local system removes the route when it stops receiving periodic control packets from the remote system.

There is one BFD agent for VLANs and port-channels that resides on RP2, as opposed to the other agents that are on the line card. Therefore, the 100 total possible sessions that this agent can maintain is shared for VLANs and port-channels.

Configuring BFD for port-channels is a two-step process:

- Enable BFD globally. Refer to Enabling BFD Globally.
- Establish sessions on port-channels. Refer to Establish Sessions on Port-Channels.

Related Configuration Tasks

- Changing Port-Channel Session Parameters.
- Disabling BFD for Port-Channels.

Establish Sessions on Port-Channels

To establish a session, you must enable BFD at interface level on both ends of the link, as shown in the following example. The session parameters do not need to match.

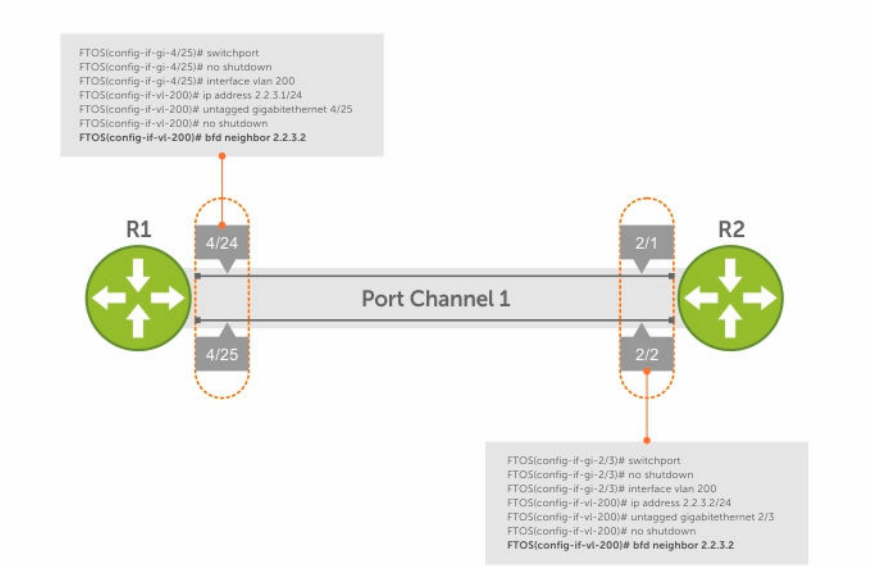


Figure 17. Establishing Sessions on Port-Channels

To establish a session on a port-channel, use the `bfd neighbor` *`ip-address`* command in INTERFACE PORT-CHANNEL mode.

View the established sessions using the `show bfd neighbors` command, as shown in Changing Port-Channel Session Parameters.

Viewing Established Sessions for VLAN Neighbors

```
R2(conf-if-po-1)#bfd neighbors 2.2.2.1
R2(conf-if-po-1)#do show bfd neighors
* - 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
* 2.2.2.2   2.2.2.1    Po 1      Up    200    200    3     C
```

Changing Physical Port 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 per interface; if you change a parameter, the change affects all physical port sessions on that interface.

CAUTION: When configuring BFD on VLAN or LAG interfaces, Dell Networking recommends a minimum value of 500 milliseconds for both the transmit and minimum receive time, which yields a final detection time of (500ms *3) 1500 milliseconds.

Change session parameters for all sessions on an interface.

INTERFACE PORT-CHANNEL mode

```
bfd interval milliseconds min_rx milliseconds multiplier value role [active
| passive]
```

View session parameters using the `show bfd neighbors detail` command.

Disabling BFD for Port-Channels

If you disable BFD on an interface, sessions on the interface are torn down.
A final Admin Down control packet is sent to all neighbors, and sessions on the remote system are placed in a Down state.

To disable BFD for a port-channel, use the following command.

- Disable BFD for a port-channel.
  INTERFACE PORT-CHANNEL mode

  ```
  no bfd enable
  ```

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
  ```

Example of Output from the `debug bfd detail` Command

Example of Output from the `debug bfd packet` Command

The following example shows a three-way handshake using the `debug bfd detail` command.

```
     R1(conf-if-gi-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 Gi 4/24 (diag: 0)
     00:54:38 : Sent packet for session with neighbor 2.2.2.2 on Gi 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 Gi 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 Gi 4/24 (diag: 0)
```

The following example displays 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 Gi 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 Gi 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 Gi 4/24
TX packet dump:
```

The output for the `debug bfd event` command is the same as the log messages that appear on the console by default.

10

Border Gateway Protocol IPv4 (BGPv4)

This chapter provides a general description of BGPv4 as it is supported in the Dell Networking operating system.

BGP protocol standards are listed in the Standards Compliance chapter.

BGP is an external gateway protocol that transmits interdomain routing information within and between autonomous systems (AS). 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.

Autonomous Systems (AS)

BGP autonomous systems (ASs) are a collection of nodes under common administration with common network routing policies.

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.

When BGP operates inside an AS (AS1 or AS2, as seen in the following illustration), it is referred to as Internal BGP (IBGP Interior Border Gateway Protocol). When BGP operates between ASs (AS1 and AS2), it is called External BGP (EBGP Exterior Border Gateway Protocol). 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.

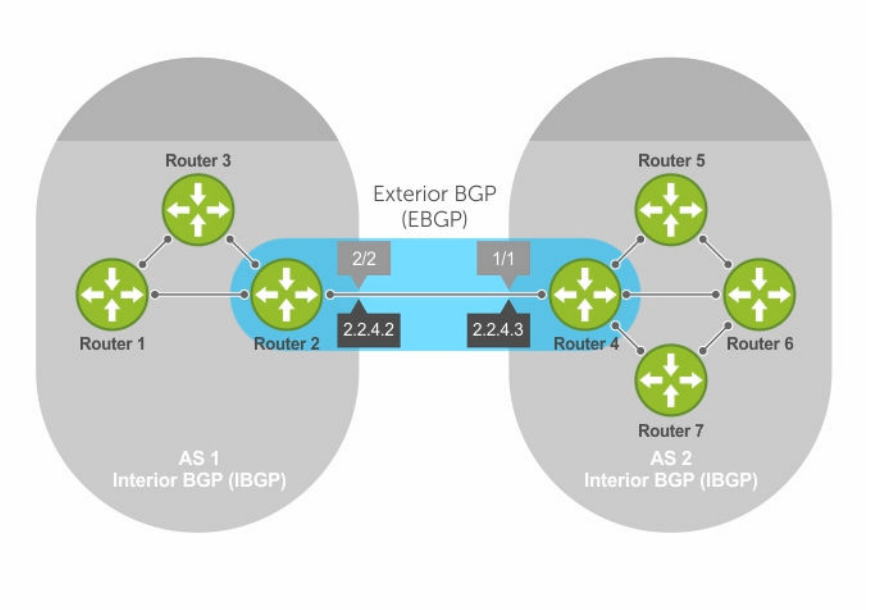


Figure 18. Interior BGP

BGP version 4 (BGPv4) supports classless interdomain routing 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.

Because each BGP router talking to another router is a session, a BGP 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 following illustration, 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.

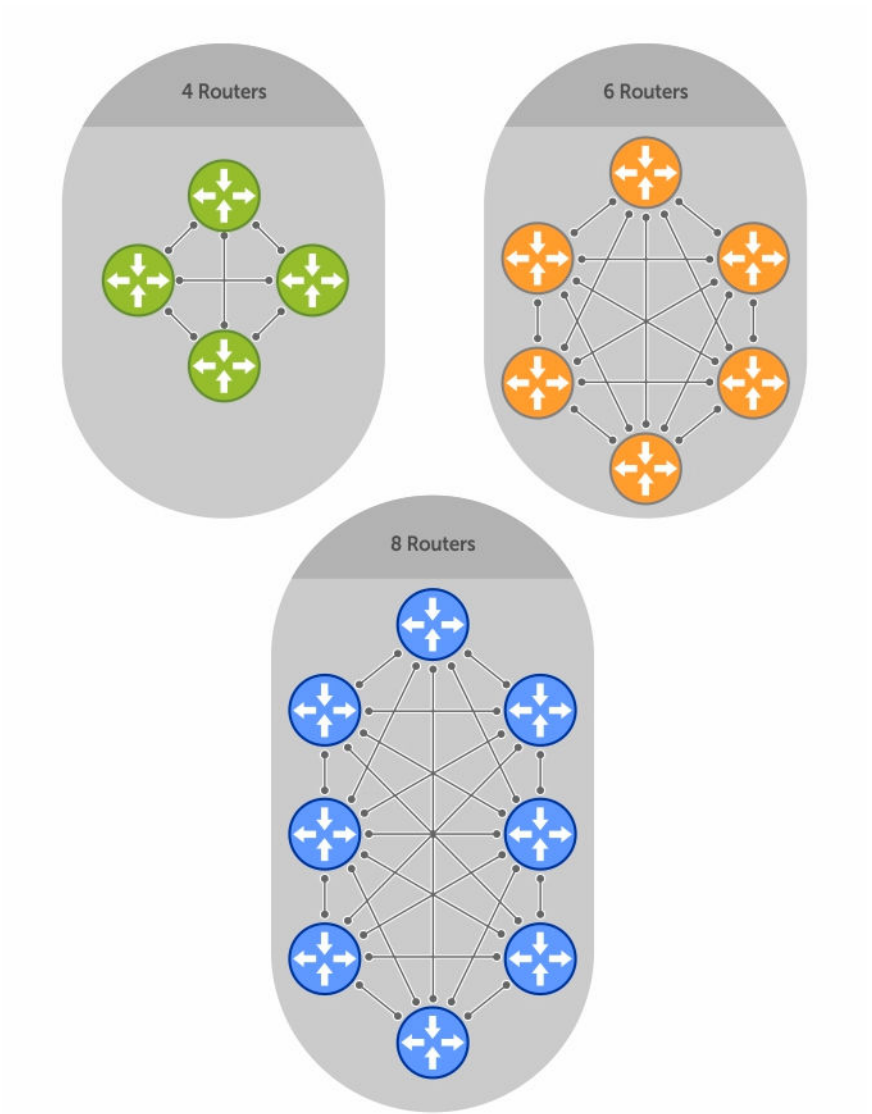


Figure 19. BGP Routers in Full Mesh

The number of BGP speakers each BGP peer must maintain increases exponentially. Network management quickly becomes impossible.

Sessions and Peers

When two routers communicate using the BGP protocol, a BGP session is started. The two end-points of that session are Peers. A Peer is also called a Neighbor.

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.
Keepalive and 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 Ggroups 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.

Route Reflectors

Route reflectors (RR) reorganize the iBGP core into a hierarchy and allow some route advertisement rules.

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.

NOTE: Address-family specific RR configurations are not supported.

- 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.

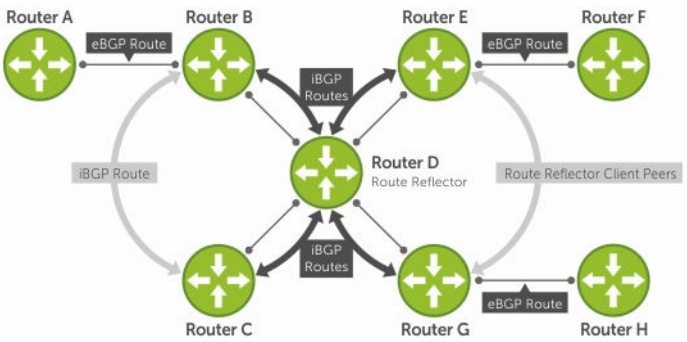


Figure 20. BGP Router Rules

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.

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.

BGP Attributes

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

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 are 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 the system 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, the system compares MED between the adjacent paths within an AS group because all paths in the AS group are from the same AS.

NOTE: In the Dell Networking OS version 8.3.11.4, 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.

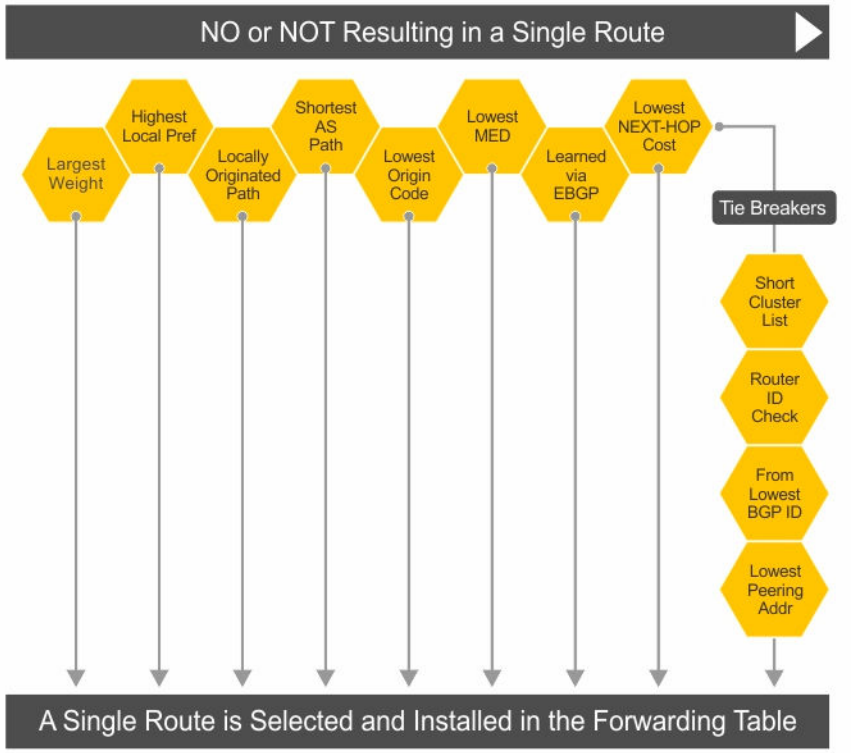


Figure 21. BGP Best Path Selection

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. The system 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 are 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 the system 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, the system 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.

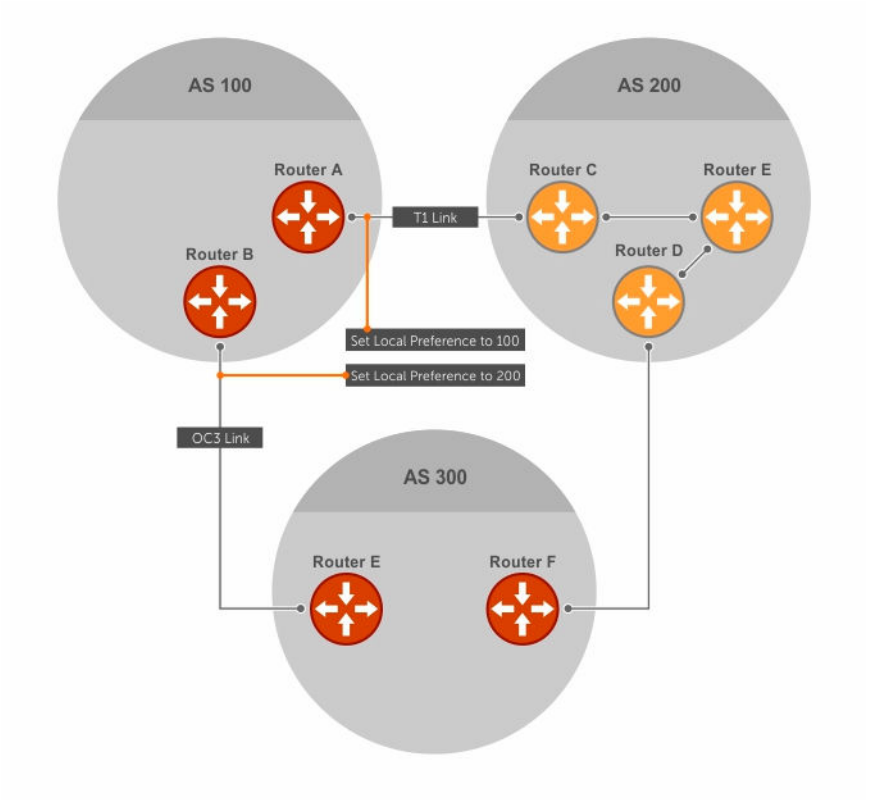


Figure 22. BGP Local Preference

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.

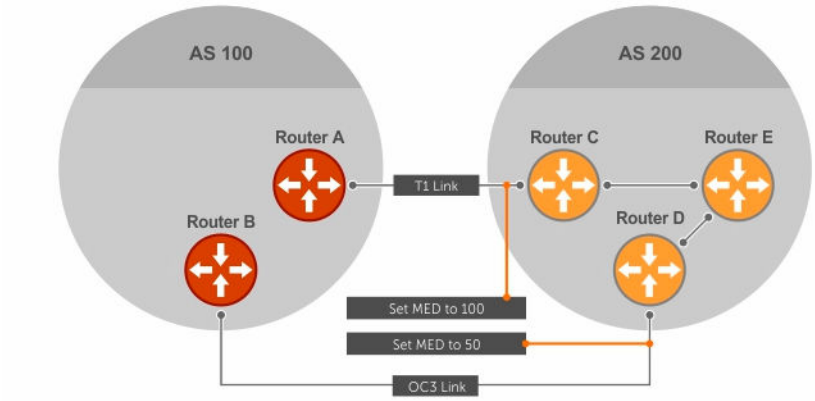


Figure 23. Multi-Exit Discriminators

NOTE: With the Dell Networking OS version 8.3.1.0, 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.

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 the Dell 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

```
Dell#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.

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

```
Dell#show ip bgp paths
Total 30655 Paths
Address    Hash Refcount Metric Path
0x4014154  0    3       18508   701 3549 19421 i
0x4013914  0    3       18508   701 7018 14990 i
0x5166d6c  0    3       18508   209 4637 1221 9249 9249 i
0x5e62df4  0    2       18508   701 17302 i
0x3a1814c  0    26      18508   209 22291 i
0x567ea9c  0    75      18508   209 3356 2529 i
0x6cc1294  0    2       18508   209 1239 19265 i
0x6cc18d4  0    1       18508   701 2914 4713 17935 i
0x5982e44  0    162     18508   209 i
0x67d4a14  0    2       18508   701 19878 ?
0x559972c  0    31      18508   209 18756 i
0x59cd3b4  0    2       18508   209 7018 15227 i
0x7128114  0    10      18508   209 3356 13845 i
0x536a914  0    3       18508   209 701 6347 7781 i
0x2ffe884  0    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.

The system allows you to set the next hop attribute in the CLI. Setting the next hop attribute lets you determine a router as the next hop for a BGP neighbor.

Multiprotocol BGP

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.

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.

Implement BGP with the Dell Networking OS

The following sections describe how to implement BGP on the Dell 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 reduces the time taken for BGP convergence by advertising multiple paths to its peers for the same address prefix without new paths implicitly replacing the existing paths. An iBGP speaker that receives multiple paths from its peers should calculate the best path in its own. BGP add-path helps switchover to next new best path based on IGP convergence time when best path becomes 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.

The Dell Networking OS version 8.3.1.0 and later support configuring 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 6. 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 for Some Best-Path Calculations

The Dell Networking OS version 8.3.1.0 and later allows you to 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.

Four-Byte AS Numbers

The Dell Networking OS version 7.7.1 and later supports 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 `four-octet-support` command.

AS4 Number Representation

The Dell Networking OS version 8.2.1.0 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 method the Dell Networking OS has used for all previous Dell Networking OS versions. ASPLAIN remains the default method with the Dell Networking OS version 8.2.1.0 and later. 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

The Dell Networking OS version 8.3.1.0 applies the ASN notation type change dynamically to the running-config statements.

When you apply or change an asnotation, 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

Example of the Running Configuration When AS Notation is Disabled

```
ASDOT
Dell(conf-router_bgp)#bgp asnotation asdot
Dell(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>

Dell(conf-router_bgp)#do show ip bgp
BGP table version is 24901, local router ID is 172.30.1.57
<output truncated>

ASDOT+
Dell(conf-router_bgp)#bgp asnotation asdot+
```

```
Dell(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>

Dell(conf-router_bgp)#do show ip bgp
BGP table version is 24901, local router ID is 172.30.1.57
<output truncated>

AS-PLAIN
Dell(conf-router_bgp)#bgp asnotation asplain+
Dell(conf-router_bgp)#sho conf
!
router bgp 100
  bgp four-octet-asdot+
  bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

Dell(conf-router_bgp)#do sho ip bgp
BGP table version is 31571, local router ID is 172.30.1.57
<output truncated>

AS-PLAIN
Dell(conf-router_bgp)#bgp asnotation asplain
Dell(conf-router_bgp)#sho conf

AS NOTATION DISABLED
Dell(conf-router_bgp)#no bgp asnotation
Dell(conf-router_bgp)#sho conf
!
router bgp 100
bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

Dell(conf-router_bgp)#do sho ip bgp
BGP table version is 28093, local router ID is 172.30.1.57

AS4 SUPPORT DISABLED
Dell(conf-router_bgp)#no bgp four-octet-as-support
Dell(conf-router_bgp)#sho conf
!
router bgp 100
```

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 24. 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 Dell Networking OS BGP management information base (MIB) support 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

- 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 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 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.

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)

BGP Configuration

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.

By default, BGP is disabled.

By default, the system compares the MED attribute on different paths from within the same AS (the `bgp always-compare-med` command is not enabled).

NOTE: In the Dell Networking OS, all newly configured neighbors and peer groups are disabled. To enable a neighbor or peer group, enter the `neighbor {ip-address | peer-group-name} no shutdown` command.

The following table displays the default values for BGP.

Table 7. BGP Default Values

Item	Default
BGP Neighbor Adjacency changes	All BGP neighbor changes are logged.
Fast External Fall over 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

Enabling BGP

By default, BGP is not enabled on the system. The Dell Networking OS supports one autonomous system (AS) and assigns the AS number (ASN).
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 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.

NOTE: Find Sample Configurations for enabling BGP routers at the end of this chapter.

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.

   a. Enable 4-Byte support for the BGP process.

      NOTE: This command is OPTIONAL. Enable if you want to use 4-Byte AS numbers or if you support AS4 number representation.

      CONFIG-ROUTER-BGP mode

      ```
      bgp four-octet-as-support
      ```

      NOTE: Use it only if you support 4-Byte AS numbers or if you support AS4 number representation. If you are supporting 4-Byte ASNs, enable this command.

      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.

      Disabling 4-Byte AS numbers also disables ASDOT and ASDOT+ number representation. All AS numbers are displayed in ASPLAIN format.

   b. Enable IPv4 multicast or IPv6 mode.
      CONFIG-ROUTER-BGP mode

      ```
      address-family [ipv4 | ipv6}
      ```

      Use this command to enter BGP for IPv6 mode (CONF-ROUTER_BGPv6_AF).

2. Add a neighbor as a remote AS.
   CONFIG-ROUTER-BGP mode

   ```
   neighbor {ip-address | peer-group name} remote-as as-number
   ```

   - *peer-group name*: 16 characters
   - *as-number*: from 0 to 65535 (2 Byte) or from 1 to 4294967295 (4 Byte) or 0.1 to 65535.65535 (Dotted format)

   Formats: IP Address A.B.C.D

   You must use Configuring Peer Groups *before* assigning them a remote AS.

3. Enable the BGP neighbor.
   CONFIG-ROUTER-BGP mode

```
neighbor {ip-address | peer-group-name} no shutdown
```

Example of the `show ip bgp summary` Command (2-Byte AS number displayed)

Example of the `show ip bgp summary` Command (4-Byte AS number displayed)

Example of the `show ip bgp neighbors` Command

Example of Verifying BGP Configuration

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 first example shows the summary with a 2-byte AS number displayed (in bold); the second example shows that the summary with a 4-byte AS number using the `show ip bgp summary` command (displays a 4–byte AS number in bold).

```
R2#show ip bgp summary
BGP router identifier 192.168.10.2, local AS number 65123
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
1 paths using 72 bytes of memory
BGP-RIB over all using 73 bytes of memory
1 BGP path attribute entrie(s) using 72 bytes of memory
1 BGP AS-PATH entrie(s) using 47 bytes of memory
5 neighbor(s) using 23520 bytes of memory

Neighbor     AS    MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/Pfx

10.10.21.1   65123 0       0       0      0    0 never    Active
10.10.32.3   65123 0       0       0      0    0 never    Active

R2#show ip bgp summary
BGP router identifier 192.168.10.2, local AS number 48735.59224
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
1 paths using 72 bytes of memory
BGP-RIB over all using 73 bytes of memory
1 BGP path attribute entrie(s) using 72 bytes of memory
1 BGP AS-PATH entrie(s) using 47 bytes of memory
5 neighbor(s) using 23520 bytes of memory

Neighbor     AS    MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/Pfx

10.10.21.1   65123 0       0       0      0   0    never   Active
10.10.32.3   65123 0       0       0      0   0    never   Active
```

For the router's identifier, the system 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 Networking OS Command Line Interface Reference Guide*.

```
Dell#show ip bgp neighbors

BGP neighbor is 10.114.8.60, remote AS 18508, external link
  BGP version 4, remote router ID 10.20.20.20
  BGP state ESTABLISHED, in this state for 00:01:58
  Last read 00:00:14, hold time is 90, keepalive interval is 30 seconds
  Received 18552 messages, 0 notifications, 0 in queue
  Sent 11568 messages, 0 notifications, 0 in queue
  Received 18549 updates, Sent 11562 updates
  Minimum time between advertisement runs is 30 seconds

  For address family: IPv4 Unicast
  BGP table version 216613, neighbor version 201190
  130195 accepted prefixes consume 520780 bytes
  Prefix advertised 49304, rejected 0, withdrawn 36143

  Connections established 1; dropped 0
  Last reset never
Local host: 10.114.8.39, Local port: 1037
Foreign host: 10.114.8.60, Foreign port: 179

BGP neighbor is 10.1.1.1, remote AS 65535, internal link
  Administratively shut down
  BGP version 4, remote router ID 10.0.0.0
  BGP state IDLE, in this state for 17:12:40
  Last read 17:12:40, hold time is 180, keepalive interval is 60 seconds
  Received 0 messages, 0 notifications, 0 in queue
  Sent 0 messages, 0 notifications, 0 in queue
  Received 0 updates, Sent 0 updates
  Minimum time between advertisement runs is 5 seconds

  For address family: IPv4 Unicast
  BGP table version 0, neighbor version 0
  0 accepted prefixes consume 0 bytes
  Prefix advertised 0, rejected 0, withdrawn 0

  Connections established 0; dropped 0
  Last reset never
  No active TCP connection
Dell#

R2#show running-config bgp
!
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 ISP1in
  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 no shutdown
neighbor 192.168.10.1 remote-as 65123
neighbor 192.168.10.1 update-source Loopback 0
```

Configuring AS4 Number Representations

Enable one type of AS number representation: ASPLAIN, ASDOT+, or ASDOT.

Term	Description
ASPLAIN	the method Dell Networking OS used for all previous Dell Networking OS versions. It remains the default method with the Dell Networking OS version 8.2.1.0 and later. With the ASPLAIN notation, a 32–bit binary AS number is translated into a decimal value.
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>.
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 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 the system 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+
  ```

Example of the `bgp asnotation asplain` Command

Example of the `bgp asnotation asdot` Command

Example of the `bgp asnotation asdot+` Command

```
Dell(conf-router_bgp)#bgp asnotation asplain
Dell(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

Dell(conf-router_bgp)#bgp asnotation asdot
Dell(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

Dell(conf-router_bgp)#bgp asnotation asdot+
Dell(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
```

Configuring Peer Groups

To configure multiple BGP neighbors at one time, create and populate a BGP peer group.
An advantage of peer groups is that members of a peer group inherit the configuration properties of the group and share same update policy.

A maximum of 256 peer groups are allowed on the system.

Create a peer group by assigning it a name, then adding members to the peer group. 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.

NOTE: Find Sample Configurations for enabling peer groups at the end of this chapter.

1. Create a peer group by assigning a name to it.
   CONFIG-ROUTERBGP mode

   ```
   neighbor peer-group-name peer-group
   ```
2. Enable the peer group.
   CONFIG-ROUTERBGP mode

   ```
   neighbor peer-group-name no shutdown
   ```

   By default, all peer groups are disabled.
3. Create a BGP neighbor.
   CONFIG-ROUTERBGP mode

   ```
   neighbor ip-address remote-as as-number
   ```
4. Enable the neighbor.

CONFIG-ROUTERBGP mode

`neighbor ip-address no shutdown`

5. Add an enabled neighbor to the peer group.
   CONFIG-ROUTERBGP mode

   `neighbor ip-address peer-group peer-group-name`
6. Add a neighbor as a remote AS.
   CONFIG-ROUTERBGP mode

   `neighbor {ip-address | peer-group name} remote-as as-number`

   Formats: IP Address A.B.C.D

   - *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.

Example of Viewing a Newly Created Peer Group

Example of Enabling a Peer Group

Example of the `show ip bgp peer-group` 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).

```
Dell(conf-router_bgp)#neighbor zanzibar peer-group
Dell(conf-router_bgp)#show conf
!
router bgp 45
  bgp fast-external-fallover
  bgp log-neighbor-changes
  neighbor zanzibar peer-group
  neighbor zanzibar shutdown
  neighbor 10.1.1.1 remote-as 65535
  neighbor 10.1.1.1 shutdown
```

To enable a peer group, use the `neighbor peer-group-name no shutdown` command in CONFIGURATION ROUTER BGP mode (shown in bold).

```
Dell(conf-router_bgp)#neighbor zanzibar no shutdown
Dell(conf-router_bgp)#show config
!
router bgp 45
  bgp fast-external-fallover
  bgp log-neighbor-changes
  neighbor zanzibar peer-group
  neighbor zanzibar no shutdown
  neighbor 10.1.1.1 remote-as 65535
  neighbor 10.1.1.1 shutdown
```

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.

```
Dell>show ip bgp peer-group

Peer-group zanzibar, remote AS 65535
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 26
Peer-group members (* - outbound optimized):
  10.68.160.1
  10.68.161.1
  10.68.162.1
  10.68.163.1
  10.68.164.1
  10.68.165.1
  10.68.166.1
  10.68.167.1
  10.68.168.1
  10.68.169.1
  10.68.170.1
  10.68.171.1
  10.68.172.1
  10.68.173.1
```

```
  10.68.174.1
  10.68.175.1
  10.68.176.1
  10.68.177.1
  10.68.178.1
```

Configuring BGP Fast Fail-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 fail-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 fail-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 fail-over feature is configured on a per-neighbor or peer-group basis and is disabled by default.

To enable the BGP fast fail-over feature, use the following command.

To disable fast fail-over, use the `[no] neighbor [neighbor | peer-group] fail-over` command in CONFIGURATION ROUTER BGP mode.

- Enable BGP Fast Fail-Over.
  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | peer-group-name} fail-over
  ```

Example of Verifying that Fast Fail-Over is Enabled on a BGP Neighbor

Example of Verifying that Fast Fail-Over is Enabled on a Peer-Group

To verify fast fail-over is enabled on a particular BGP neighbor, use the `show ip bgp neighbors` command. Because fast fail-over is disabled by default, it appears only if it has been enabled (shown in bold).

```
Dell#sh ip bgp neighbors

BGP neighbor is 100.100.100.100, remote AS 65517, 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)
```

```
fail-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
```

To verify that fast fail-over is enabled on a peer-group, use the `show ip bgp peer-group` command (shown in bold).

```
Dell#sh ip bgp peer-group

Peer-group test
  fail-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):
    100.100.100.100*

Dell#

router bgp 65517
  neighbor test peer-group
  neighbor test fail-over
  neighbor test no shutdown
```

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, the system 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 256. 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 Configuring Peer Groups.

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 | 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: IP Address: A.B.C.D.

  You must use Configuring Peer Groups *before* assigning it to an AS. This feature is not supported on passive peer groups.

Example of the Verifying that Local AS Numbering is Disabled

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
```

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 | peer-group-name} allowas-in number`

  – *Peer Group Name*: 16 characters.
  – *Number*: 1 through 10.

  Format: IP Address: A.B.C.D.

  You must use Configuring Peer Groups'before assigning it to an AS.

Example of Viewing AS Numbers in AS Paths

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
```

Enabling Graceful Restart

Use this feature to lessen the negative effects of a BGP restart.
The Dell 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, The Dell 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, the system 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 `configure router bgp graceful-restart` command.

- Enable graceful restart for the BGP node.

  CONFIG-ROUTER-BGP mode

  ```
  bgp graceful-restart
  ```
- Set maximum restart time for all peers.

  CONFIG-ROUTER-BGP mode

  ```
  bgp graceful-restart [restart-time time-in-seconds]
  ```

  The default is **120 seconds**.
- Set maximum time 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**.
- Local router supports graceful restart as a receiver only.

  CONFIG-ROUTER-BGP mode

  ```
  bgp graceful-restart [role receiver-only]
  ```

Enabling Neighbor Graceful Restart

BGP graceful restart is active only when the neighbor becomes established. Otherwise, it is disabled. Graceful-restart applies to all neighbors with established adjacency.

With the graceful restart feature, the system enables the receiving/restarting mode by default. In Receiver-Only mode, graceful restart saves the advertised routes of peers that support this capability

when they restart. This option provides support for remote peers for their graceful restart without supporting the feature itself.

You can implement BGP graceful restart either by neighbor or by BGP peer-group. For more information, refer to the *Dell Networking OS Command Line Interface Reference Guide*.

- Add graceful restart to a BGP neighbor or peer-group.
  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | peer-group-name} graceful-restart
  ```
- Set the maximum restart time for the neighbor or peer-group.
  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | peer-group-name} graceful-restart [restart-time time-in-seconds]
  ```

  The default is **120 seconds**.
- Local router supports graceful restart for this neighbor or peer-group as a receiver only.
  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | peer-group-name} graceful-restart [role receiver-only]
  ```
- Set the maximum time to retain the restarting neighbor's or peer-group's stale paths.
  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | peer-group-name} graceful-restart [stale-path-time time-in-seconds]
  ```

  The default is **360 seconds**.

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 any 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 | 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.

Example of the `show ip bgp paths` Command

To view all BGP path attributes in the BGP database, use the `show ip bgp paths` command in EXEC Privilege mode.

```
Dell#show ip bgp paths
Total 30655 Paths
Address   Hash Refcount Metric Path
0x4014154 0    3        18508 701 3549 19421 i
0x4013914 0    3        18508 701 7018 14990 i
0x5166d6c 0    3        18508 209 4637 1221 9249 9249 i
0x5e62df4 0    2        18508 701 17302 i
0x3a1814c 0    26       18508 209 22291 i
0x567ea9c 0    75       18508 209 3356 2529 i
0x6cc1294 0    2        18508 209 1239 19265 i
0x6cc18d4 0    1        18508 701 2914 4713 17935 i
0x5982e44 0    162      18508 209 i
0x67d4a14 0    2        18508 701 19878 ?
0x559972c 0    31       18508 209 18756 i
0x59cd3b4 0    2        18508 209 7018 15227 i
0x7128114 0    10       18508 209 3356 13845 i
0x536a914 0    3        18508 209 701 6347 7781 i
0x2ffe884 0    1        18508 701 3561 9116 21350 i
0x2ff7284 0    99       18508 701 1239 577 855 ?
0x2ff7ec4 0    4        18508 209 3561 4755 17426 i
0x2ff8544 0    3        18508 701 5743 2648 i
0x736c144 0    1        18508 701 209 568 721 1494 i
0x3b8d224 0    10       18508 209 701 2019 i
0x5eb1e44 0    1        18508 701 8584 16158 i
0x5cd891c 0    9        18508 209 6453 4759 i
--More--
```

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 the Dell 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

```
Dell(config)#router bgp 99
Dell(conf-router_bgp)#neigh AAA peer-group
Dell(conf-router_bgp)#neigh AAA no shut
Dell(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
Dell(conf-router_bgp)#neigh 10.155.15.2 filter-list 1 in
Dell(conf-router_bgp)#ex

Dell(conf)#ip as-path access-list Eagle
Dell(config-as-path)#deny 32$
Dell(config-as-path)#ex
Dell(conf)#router bgp 99
Dell(conf-router_bgp)#neighbor AAA filter-list Eagle in
Dell(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
Dell(conf-router_bgp)#ex
```

Redistributing Routes

In addition to filtering routes, you can add routes from other routing instances or protocols to the BGP process. With the `redistribute` command, you can include ISIS, OSPF, static, or directly connected routes in the BGP process.
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 in BGP.
  ROUTER BGP or CONF-ROUTER_BGPv6_ AF mode

  `redistribute {connected | static} [route-map map-name]`

  Configure the `map-name` parameter to specify the name of a configured route map.
- Include specific ISIS routes in 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:
  - `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**.
  - `map-name`: name of a configured route map.
- Include specific OSPF routes in IS-IS.
  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:
  - `process-id`: the range is from 1 to 65535.
  - `match external`: the range is from 1 or 2.
  - `match internal`

– `metric-type`: external or internal.

– `map-name`: name of a configured route map.

Enabling Additional Paths

The add-path feature is disabled by default.

NOTE: Note: In some cases, while receiving 1K same routes from more than 64 iBGP neighbors, BGP sessions holdtime of 10 seconds may flap. The BGP add-path does not update packets for advertisement and cannot scale to higher numbers. Either reduce the number of routes added or increase the holddown timer value.

To allow multiple paths sent to peers, use the following commands.

1. Allow the advertisement of multiple paths for the same address prefix without the new paths replacing any previous ones.
   CONFIG-ROUTER-BGP mode

   ```
   bgp add-path [send | receive | both] count
   ```

   The range is from 2 to 64.
2. Allow the specified neighbor/peer group to send/ receive multiple path advertisements.
   CONFIG-ROUTER-BGP mode

   ```
   neighbor add-path
   ```
3. Configure the maximum number of parallel routes (multipath support) BGP supports.
   CONFIG-ROUTER-BGP mode

   ```
   max-path number
   ```

   The range is from 2 to 64.

Configuring IP Community Lists

Within the Dell 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. 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 boundary, but are sent to CONFED-EBGP and IBGP peers.

The Dell 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.

Example of the `show ip community-lists` 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.

```
Dell#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
```

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.

Example of the `show ip extcommunity-lists` Command

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.

```
Dell#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
```

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 | 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 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.

Manipulating the COMMUNITY Attribute

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, the system does not send the COMMUNITY attribute.
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 | peer-group-name} send-community
  ```

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 | peer-group-name} route-map map-name {in | out}
   ```

Example of the `show ip bgp community` Command

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.

```
Dell>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
--More--
```

Changing MED Attributes

By default, the system uses the MULTI_EXIT_DISC or MED attribute when comparing EBGP paths from the same AS.
To change how the MED attribute is used, 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`: Treat a path missing an MED as the most preferred one.

To view the nondefault values, use the `show config` command in CONFIGURATION ROUTER BGP mode.

Changing the LOCAL_PREFERENCE Attribute

In the Dell Networking OS, you can change the value of the LOCAL_PREFERENCE attribute.
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**.

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.

A more flexible method for manipulating the LOCAL_PREF attribute value is to use a route map.

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 | 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.

Changing the NEXT_HOP Attribute

You can change how the NEXT_HOP attribute is used.
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 as the next hop for a BGP neighbor.
  CONFIG-ROUTER-BGP mode

  `neighbor {ip-address | peer-group-name} next-hop-self`
- Sets the next hop address.
  CONFIG-ROUTE-MAP mode

  `set next-hop ip-address`

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 | 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.

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 16 parallel paths to a destination.
To allow more than one path, use the following command.

The `show ip bgp network` command includes multipath information for that network.

- Enable multiple parallel paths.
  CONFIG-ROUTER-BGP mode

  `maximum-paths {ebgp | ibgp} number`

The `show ip bgp network` command includes multipath information for that network.

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: The Dell Networking OS supports up to 255 characters in a set community statement inside a route map.

NOTE: With the Dell Networking OS, you can create inbound and outbound policies. Each of the commands used for filtering has `in` and `out` parameters that you must apply. In the system, 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.

To filter routes using prefix lists, 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 be 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 | peer-group-name} distribute-list prefix-list-name {in
| out}
```

Configure the following parameters:

- *ip-address* or *peer-group-name*: enter the neighbor's IP 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.

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, refer to 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 | peer-group-name} route-map map-name {in | out}
   ```

   Configure the following parameters:

- *ip-address* or *peer-group-name*: enter the neighbor's IP 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.

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 BGP Routes Using AS-PATH Information

To filter routes based on AS-PATH information, use these commands.

1. Create a AS-PATH ACL and assign it a name.
   CONFIGURATION mode

   ```
   ip as-path access-list as-path-name
   ```
2. Create a AS-PATH ACL filter with a deny or permit action.
   AS-PATH ACL mode

   ```
   {deny | permit} as-regular-expression
   ```
3. Return to CONFIGURATION mode.
   AS-PATH ACL

   ```
   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 | peer-group-name} filter-list as-path-name {in | out}
   ```

   Configure the following parameters:
   - *ip-address* or *peer-group-name*: enter the neighbor's IP address or the peer group's name.
   - *as-path-name*: enter the name of a configured AS-PATH ACL.
   - in: apply the AS-PATH ACL map to inbound routes.
   - out: apply the AS-PATH ACL to outbound routes.

To view which commands are configured, use the `show config` command in CONFIGURATION ROUTER BGP mode and the `show ip as-path-access-list` command in EXEC Privilege mode.

To forward all routes not meeting the AS-PATH ACL criteria, include the **permit .*** filter in your AS-PATH ACL.

Configuring BGP Route Reflectors

BGP route reflectors are intended for ASs with a large mesh; they reduce the amount of BGP control traffic.

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 an ID to a router reflector cluster.
  CONFIG-ROUTER-BGP mode

  ```
  bgp cluster-id cluster-id
  ```

  You can have multiple clusters in an AS.
- Configure the local router as a route reflector and the neighbor or peer group identified is the route reflector client.
  CONFIG-ROUTER-BGP mode

  ```
  neighbor {ip-address | peer-group-name} route-reflector-client
  ```

When you enable a route reflector, the system 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.

Aggregating Routes

The Dell Networking OS provides multiple ways to aggregate routes in the BGP routing table. 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]
  ```

Example of Viewing Aggregated Routes

In the `show ip bgp` command, aggregates contain an 'a' in the first column (shown in bold) and routes suppressed by the aggregate contain an 's' in the first column.

```
Dell#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 ?
*>a 9.0.0.0/8    192.0.0.0              32768 18508 701 {7018 2686 3786} ?
```

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.

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, 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 dampening [ip-address mask]
  ```
- View all flap statistics or for specific routes meeting the following criteria.
  EXEC or EXEC Privilege mode

`show ip bgp 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 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.

  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.

Example of Configuring a Route for Reuse or Restart

Example of Viewing the Number of Dampened Routes

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.

```
Dell(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>
Dell(conf-router_bgp)#bgp dampening 2 ?
<1-20000>     Value to start reusing a route (default = 750)
Dell(conf-router_bgp)#bgp dampening 2 2000 ?
<1-20000>     Value to start suppressing a route (default = 2000)
Dell(conf-router_bgp)#bgp dampening 2 2000 3000 ?
<1-255>       Maximum duration to suppress a stable route (default = 60)
Dell(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).

```
Dell>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
Dell>
```

To view which routes are dampened (non-active), use the `show ip bgp dampened-routes` command in EXEC Privilege mode.

Changing BGP Timers

To configure BGP timers, use either or both of the following commands.

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`* `|` *`peer-group-name`*`} timers` *`keepalive holdtime`*

  - *`keepalive`*: the range is from 1 to 65535. Time interval, in seconds, between keepalive messages sent to the neighbor routers. The default is **60 seconds**.
  - *`holdtime`*: the range is from 3 to 65536. Time interval, in seconds, between the last keepalive message and declaring the router dead. The default is **180 seconds**.
- Configure timer values for all neighbors.

  CONFIG-ROUTER-BGP mode

  `timers bgp` *`keepalive holdtime`*

  - *`keepalive`*: the range is from 1 to 65535. Time interval, in seconds, between keepalive messages sent to the neighbor routers. The default is **60 seconds**.
  - *`holdtime`*: the range is from 3 to 65536. Time interval, in seconds, between the last keepalive message and declaring the router dead. The default is **180 seconds**.

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.

Enabling BGP Neighbor 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 reconfig allows for policies to be applied to a session without clearing the BGP Session. Soft-reconfig 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 enable soft-reconfiguration for a neighbor and you execute the `clear ip bgp soft in` command, the update database stored in the router is replayed and updates are reevaluated. With this command, the replay and update process is triggered only if a route-refresh request is not negotiated with the peer. If the request is indeed negotiated (after execution of `clear ip bgp soft in`), BGP sends a route-refresh request to the neighbor and receives all of the peer's updates.

To use soft reconfiguration (or soft reset) without preconfiguration, both BGP peers must support the soft route refresh capability, which is advertised in the open message sent when the peers establish a TCP session.

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.

- Clear all information or only specific details.
  EXEC Privilege mode

  `clear ip bgp {* | neighbor-address | AS Numbers | ipv4 | peer-group-name} [soft [in | out]]`
  - `*`: Clears all peers.
  - `neighbor-address`: Clears the neighbor with this IP address.
  - `AS Numbers`: Peers' AS numbers to be cleared.
  - `ipv4`: Clears information for the IPv4 address family.
  - `peer-group-name`: Clears all members of the specified peer group.
- Enable soft-reconfiguration for the BGP neighbor specified.
  CONFIG-ROUTER-BGP mode

  `neighbor {ip-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.

Example of Soft-Reconfigration of a BGP Neighbor

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.

```
Dell>router bgp 100
   neighbor 10.108.1.1 remote-as 200
   neighbor 10.108.1.1 soft-reconfiguration inbound
```

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.

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.

The Dell 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, the system 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 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 Networking OS Command Line Interface Reference Guide*.

- Enables support for the IPv4 multicast family on the BGP node.
  CONFIG-ROUTER-BGP mode

  ```
  address family ipv4 multicast
  ```
- Enable IPv4 multicast support on a BGP neighbor/peer group.
  CONFIG-ROUTER-BGP-AF (Address Family) mode

  ```
  neighbor [ip-address | peer-group-name] activate
  ```

BGP Regular Expression Optimization

The Dell 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
  ```

  In-BGP is shown using the `show ip protocols` command.

The system 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

The system 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

```
Dell(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
```

PDU Counters

The Dell Networking OS version 7.5.1.0 introduces 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.

Sample Configurations

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.

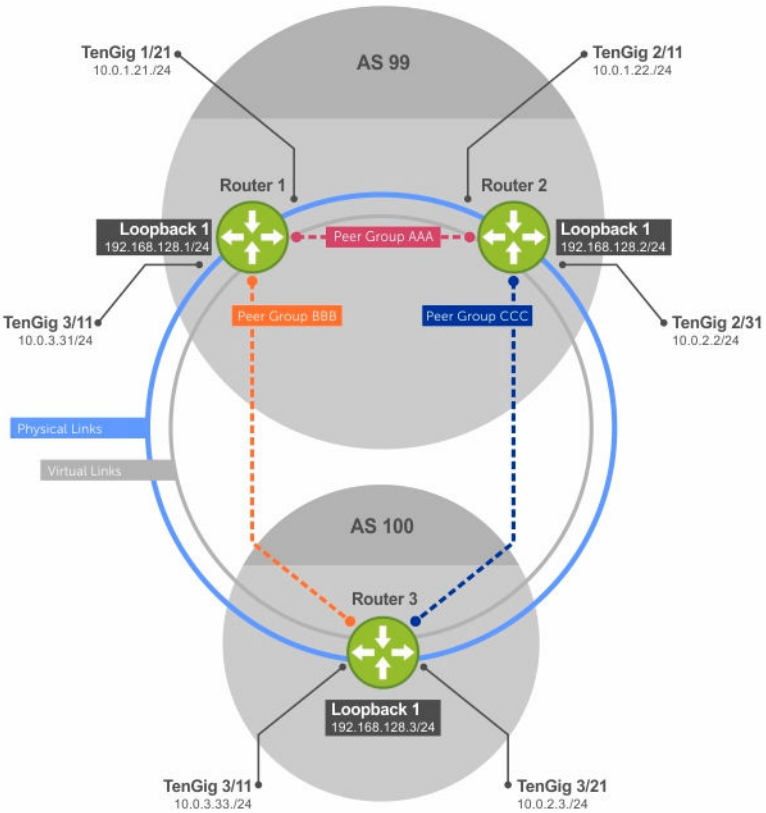


Figure 25. Sample Configurations

Example of Enabling BGP (Router 1)

Example of Enabling BGP (Router 2)

Example of Enabling BGP (Router 3)

Example of Enabling Peer Groups (Router 1)

Example of Enabling Peer Groups (Router 2)

Example of Enabling Peer Groups (Router 3)

```
R1# conf
R1(conf)#int loop 0
R1(conf-if-lo-0)#ip address 192.168.128.1/24
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 gig 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 gig 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)#router bgp 99
R1(conf-router_bgp)#network 192.168.128.0/24
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 192.168.128.3 remote 100
R1(conf-router_bgp)#neighbor 192.168.128.3 no shut
R1(conf-router_bgp)#neighbor 192.168.128.3 update-source loop 0
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 192.168.128.2 no shutdown
  neighbor 192.168.128.3 remote-as 100
  neighbor 192.168.128.3 update-source Loopback 0
  neighbor 192.168.128.3 no shutdown
R1(conf-router_bgp)#end
R1#
R1#show ip bgp summary
BGP router identifier 192.168.128.1, local AS number 99
BGP table version is 4, main routing table version 4
4 network entrie(s) using 648 bytes of memory
6 paths using 408 bytes of memory
BGP-RIB over all using 414 bytes of memory
3 BGP path attribute entrie(s) using 144 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  4       5       4      0   0    00:00:32   1
192.168.128.3 100 5       4       1      0   0    00:00:09   4
R1#

R2# conf
R2(conf)#int loop 0
R2(conf-if-lo-0)#ip address 192.168.128.2/24
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 gig 2/11
R2(conf-if-gi-2/11)#ip address 10.0.1.22/24
R2(conf-if-gi-2/11)#no shutdown
R2(conf-if-gi-2/11)#show config
!
interface GigabitEthernet 2/11
  ip address 10.0.1.22/24
  no shutdown
R2(conf-if-gi-2/11)#int gig 2/31

R2(conf-if-gi-2/31)#ip address 10.0.2.2/24
```

```
R2(conf-if-gi-2/31)#no shutdown
R2(conf-if-gi-2/31)#show config
!
interface GigabitEthernet 2/31
ip address 10.0.2.2/24
no shutdown
R2(conf-if-gi-2/31)#
R2(conf-if-gi-2/31)#router bgp 99
R2(conf-router_bgp)#network 192.168.128.0/24
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 192.168.128.3 remote 100
R2(conf-router_bgp)#neighbor 192.168.128.3 no shut
R2(conf-router_bgp)#neighbor 192.168.128.3 update loop 0
R2(conf-router_bgp)#show config
!
router bgp 99
  bgp router-id 192.168.128.2
  network 192.168.128.0/24
  bgp graceful-restart
  neighbor 192.168.128.1 remote-as 99
  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 update-source Loopback 0
  neighbor 192.168.128.3 no shutdown
R2(conf-router_bgp)#end

R2#show ip bgp summary
BGP router identifier 192.168.128.2, 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 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  40      35      1      0   0    00:01:05   1
192.168.128.3 100 4       4       1      0   0    00:00:16   1
R2#

R3# conf
R3(conf)#
R3(conf)#int loop 0
R3(conf-if-lo-0)#ip address 192.168.128.3/24
R3(conf-if-lo-0)#no shutdown
R3(conf-if-lo-0)#show config
!
  interface Loopback 0
  ip address 192.168.128.3/24
no shutdown
R3(conf-if-lo-0)#int gig 3/11
R3(conf-if-gi-3/11)#ip address 10.0.3.33/24
R3(conf-if-gi-3/11)#no shutdown
R3(conf-if-gi-3/11)#show config
!
interface GigabitEthernet 3/11
  ip address 10.0.3.33/24
  no shutdown

R3(conf-if-lo-0)#int gig 3/21
```

```
R3(conf-if-gi-3/21)#ip address 10.0.2.3/24
R3(conf-if-gi-3/21)#no shutdown
R3(conf-if-gi-3/21)#show config
!
interface GigabitEthernet 3/21
  ip address 10.0.2.3/24
  no shutdown

R3(conf-if-gi-3/21)#
R3(conf-if-gi-3/21)#router bgp 100
R3(conf-router_bgp)#show config
!
router bgp 100
R3(conf-router_bgp)#network 192.168.128.0/24
R3(conf-router_bgp)#neighbor 192.168.128.1 remote 99
R3(conf-router_bgp)#neighbor 192.168.128.1 no shut
R3(conf-router_bgp)#neighbor 192.168.128.1 update-source loop 0
R3(conf-router_bgp)#neighbor 192.168.128.2 remote 99
R3(conf-router_bgp)#neighbor 192.168.128.2 no shut
R3(conf-router_bgp)#neighbor 192.168.128.2 update loop 0
R3(conf-router_bgp)#show config
!
router bgp 100
  network 192.168.128.0/24
  neighbor 192.168.128.1 remote-as 99
  neighbor 192.168.128.1 update-source Loopback 0
  neighbor 192.168.128.1 no shutdown
  neighbor 192.168.128.2 remote-as 99
  neighbor 192.168.128.2 update-source Loopback 0
  neighbor 192.168.128.2 no shutdown
R3(conf)#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  24      25      1      0   0    00:14:20   1
192.168.128.2 99  14      14      1      0   0    00:10:22   1
R3#

R1#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      1      0   (0)  00:00:17   1
192.168.128.3 100 30      29      1      0   (0)  00:00:14   1
!
R1#show ip bgp neighbors

BGP neighbor is 192.168.128.2, remote AS 99, internal link
  Member of peer-group AAA for session parameters
  BGP version 4, remote router ID 192.168.128.2
  BGP state ESTABLISHED, in this state for 00:00:37
  Last read 00:00:36, last write 00:00:36
  Hold time is 180, keepalive interval is 60 seconds
  Received 23 messages, 0 in queue
    2 opens, 0 notifications, 2 updates
    19 keepalives, 0 route refresh requests
  Sent 24 messages, 0 in queue
    2 opens, 1 notifications, 2 updates
    19 keepalives, 0 route refresh requests
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)

  Update source set to Loopback 0
  Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer

  Connections established 2; dropped 1
  Last reset 00:00:57, due to user reset

  Notification History
    'Connection Reset' Sent : 1 Recv: 0
Last notification (len 21) sent 00:00:57 ago
    ffffffff ffffffff ffffffff ffffffff 00150306 00000000
Local host: 192.168.128.1, Local port: 179
Foreign host: 192.168.128.2, Foreign port: 65464
BGP neighbor is 192.168.128.3, remote AS 100, external link
```

```
  Member of peer-group BBB for session parameters
  BGP version 4, remote router ID 192.168.128.3
  BGP state ESTABLISHED, in this state for 00:00:37
  Last read 00:00:36, last write 00:00:36
  Hold time is 180, keepalive interval is 60 seconds
  Received 30 messages, 0 in queue
    4 opens, 2 notifications, 4 updates
    20 keepalives, 0 route refresh requests
  Sent 29 messages, 0 in queue
    4 opens, 1 notifications, 4 updates
    20 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)
Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)
Update source set to Loopback 0
  Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer
Connections established 4; dropped 3
  Last reset 00:00:54, due to user reset
R1#

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

R2#show ip bgp neighbor

BGP neighbor is 192.168.128.1, remote AS 99, internal link
  Member of peer-group AAA for session parameters
  BGP version 4, remote router ID 192.168.128.1
  BGP state ESTABLISHED, in this state for 00:11:42
  Last read 00:00:38, last write 00:00:38
  Hold time is 180, keepalive interval is 60 seconds
  Received 140 messages, 0 in queue
    6 opens, 2 notifications, 19 updates
    113 keepalives, 0 route refresh requests
  Sent 136 messages, 0 in queue
    12 opens, 3 notifications, 6 updates
    115 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 5 seconds
  Minimum time before advertisements start is 0 seconds

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
R3#show ip bgp neighbor

BGP neighbor is 192.168.128.1, remote AS 99, external link
  Member of peer-group BBB for session parameters
  BGP version 4, remote router ID 192.168.128.1
  BGP state ESTABLISHED, in this state for 00:00:21
  Last read 00:00:09, last write 00:00:08
  Hold time is 180, keepalive interval is 60 seconds
  Received 93 messages, 0 in queue
    5 opens, 0 notifications, 5 updates
```

```
    83 keepalives, 0 route refresh requests
  Sent 99 messages, 0 in queue
    5 opens, 4 notifications, 5 updates
    85 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)

Update source set to Loopback 0
Peer active in peer-group outbound optimization

For address family: IPv4 Unicast
BGP table version 1, neighbor version 1
Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
Prefixes advertised 1, denied 0, withdrawn 0 from peer

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)

Update source set to Loopback 0
Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
BGP table version 2, neighbor version 2
Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
Prefixes advertised 1, denied 0, withdrawn 0 from peer

Connections established 6; dropped 5
Last reset 00:12:01, due to Closed by neighbor

Notification History
  'HOLD error/Timer expired' Sent : 1 Recv: 0
  'Connection Reset' Sent : 2 Recv: 2

  Last notification (len 21) received 00:12:01 ago
    ffffffff ffffffff ffffffff ffffffff 00150306 00000000
Local host: 192.168.128.2, Local port: 65464
Foreign host: 192.168.128.1, Foreign port: 179

BGP neighbor is 192.168.128.3, remote AS 100, external link
  Member of peer-group BBB for session parameters
  BGP version 4, remote router ID 192.168.128.3
  BGP state ESTABLISHED, in this state for 00:18:51
  Last read 00:00:45, last write 00:00:44
  Hold time is 180, keepalive interval is 60 seconds
  Received 138 messages, 0 in queue
    7 opens, 2 notifications, 7 updates
    122 keepalives, 0 route refresh requests
  Sent 140 messages, 0 in queue
```

```
    7 opens, 4 notifications, 7 updates
    122 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds
Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)
ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Update source set to Loopback 0
  Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
  BGP table version 2, neighbor version 2
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer
```

11

Content Addressable Memory (CAM)

Content addressable memory (CAM) is a type of memory that stores information in the form of a lookup table.

On Dell Networking systems, CAM stores Layer 2 and Layer 3 forwarding information, access-lists (ACLs), flows, and routing policies.

CAM Allocation

Allocate space for IPV4 ACLs and quality of service (QoS) regions 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.

NOTE: There are 16 FP blocks, but the system flow requires three blocks that cannot be reallocated.

The following table lists the default CAM allocation settings.

Table 8. Default Cam Allocation Settings

CAM Allocation	Setting
L3 ACL (ipv4acl)	2
L2 ACL(l2acl)	2
IPv6 L3 ACL (ipv6acl)	0
L3 QoS (ipv4qos)	2
L2 QoS (l2qos)	0
L2PT (l2pt)	0
MAC ACLs (ipmacacl)	0
ECFMACL (ecfmacl)	0
nlbclusteracl	2
FCOEACL (fcoeacl)	4
ISCSIOPTACL (iscsioptacl)	2
VMAN QoS (vman-qos)	0
VMAN Dual QoS (vman-dual-qos)	0

The `ipv6acl` and `vman-dual-qos` allocations must be entered as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd numbered ranges.

NOTE: There can be only 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.

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.

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 space for the ACLs and QoS regions.
2. Enter the number of FP blocks for each region.
   EXEC Privilege mode

   ```
   l2acl number ipv4acl number ipv6acl number, ipv4qos number l2qos number,
   l2pt number ipmacacl number ecfmacl number nlbcluster number[vman-qos |
   vman-dual-qos number
   ```

3. Reload the system.
   EXEC Privilege mode

   ```
   reload
   ```

4. Verify that the new settings will be written to the CAM on the next boot.
   EXEC Privilege mode

   ```
   show cam-acl
   ```

Test CAM Usage

This command applies to both IPv4 CAM profiles, but is best used when verifying QoS optimization for IPv6 ACLs.

Use this command to determine whether sufficient ACL CAM space is available to enable a service-policy. Create a Class Map with all required ACL rules, then execute the `test cam-usage` command in Privilege mode to verify the actual CAM space required. The Status column in the command output indicates whether or not the policy can be enabled.

Example of the `test cam-usage` Command

```
Dell#test cam-usage service-policy input pmap stack-unit all

Stack-Unit | Portpipe | CAM Partition | Available CAM | Estimated CAM per Port
| Status
--------------------------------------------------------------------------------
----------
         2 |        0 |         L2ACL |            28 |                      1
| Allowed (28)
```

View CAM-ACL Settings

View the current cam-acl settings using the `show cam-acl` command.

Example of Viewing CAM-ACL Settings

```
Dell#show cam-acl

-- Chassis Cam ACL --
         Current Settings(in block sizes)
L2Acl       : 6
Ipv4Acl     : 2
Ipv6Acl     : 0
Ipv4Qos     : 2
L2Qos       : 1
L2PT        : 0
IpMacAcl    : 0
VmanQos     : 0
VmanDualQos : 0
EcfmAcl     : 0
FcoeAcl     : 0
iscsiOptAcl : 2

-- Stack unit 5 --
         Current Settings(in block sizes)
L2Acl       : 6
Ipv4Acl     : 2
Ipv6Acl     : 0
Ipv4Qos     : 2
L2Qos       : 1
L2PT        : 0
IpMacAcl    : 0
VmanQos     : 0
VmanDualQos : 0
EcfmAcl     : 0
FcoeAcl     : 0
iscsiOptAcl : 2
Dell#
```

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 1 FP entry is used).

When you disable this command, the system behaves as described in this chapter. However, enabling CAM optimization would apply a single rate policy FP entry. If the input service policy maps applied to several ports are the same, rate policing is applied to all the ports as a group and not individually.

12

Control Plane Policing (CoPP)

Dell Networking OS supports control plane policing (CoPP).

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.

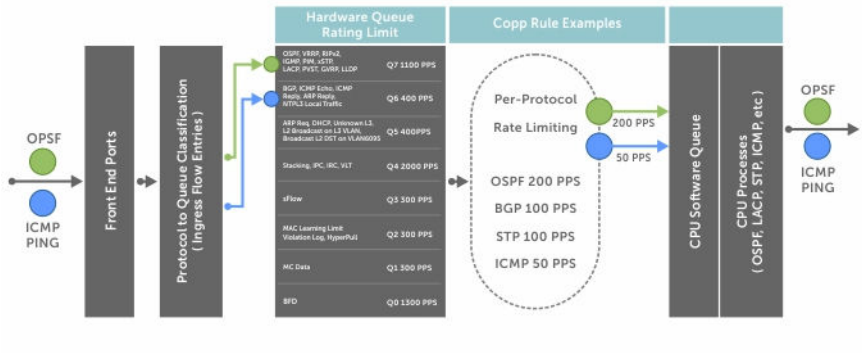


Figure 26. Control Plane Policing

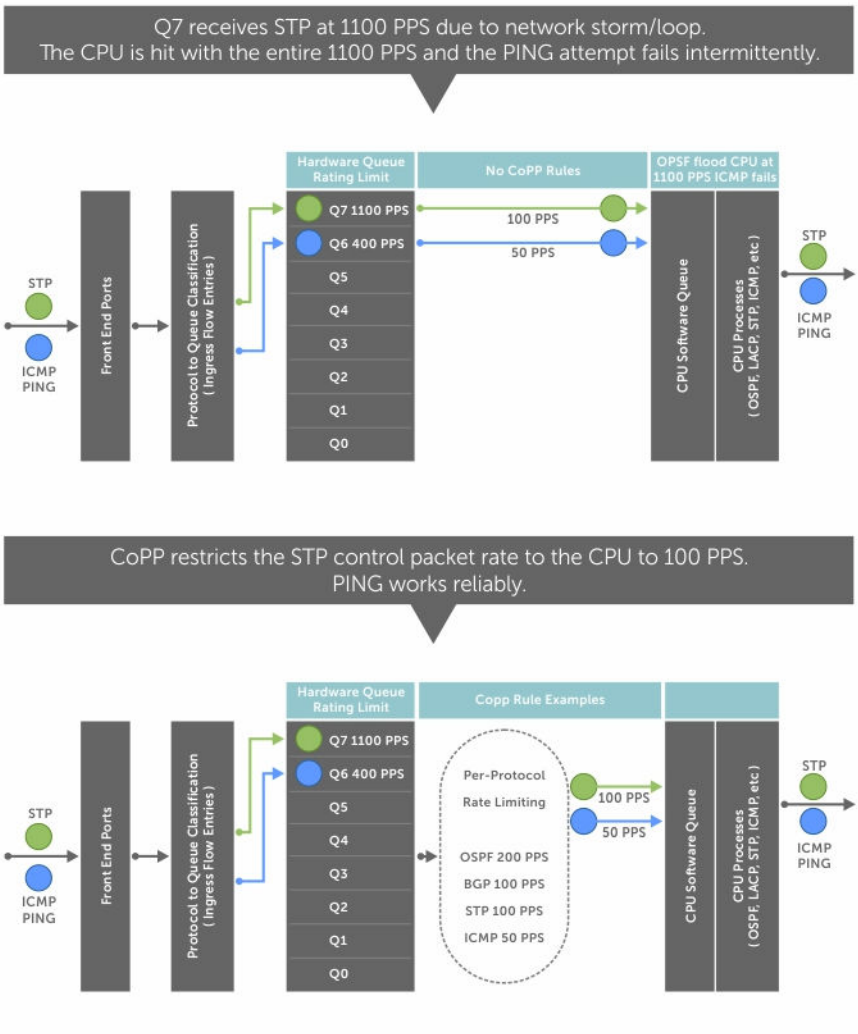


Figure 27. CoPP Implemented Versus CoPP Not Implemented

Configure Control Plane Policing

The switch can process maximum of 4200 PPS (packets per second). 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 applies 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)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 | 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
   ```
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
```

Example of Creating the IP/IPv6/MAC Extended ACL

Example of Creating the QoS Input Policy

Example of Creating the QoS Class Map

Example of Matching the QoS Class Map to the QoS Policy

Example of Creating the Control Plane Service Policy

```
Dell(conf)#ip access-list extended ospf cpu-qos
Dell(conf-ip-acl-cpuqos)#permit ospf
Dell(conf-ip-acl-cpuqos)#exit

Dell(conf)#ip access-list extended bgp cpu-qos
Dell(conf-ip-acl-cpuqos)#permit bgp
Dell(conf-ip-acl-cpuqos)#exit

Dell(conf)#mac access-list extended lacp cpu-qos
Dell(conf-mac-acl-cpuqos)#permit lacp
Dell(conf-mac-acl-cpuqos)#exit

Dell(conf)#ipv6 access-list ipv6-icmp cpu-qos
Dell(conf-ipv6-acl-cpuqos)#permit icmp
Dell(conf-ipv6-acl-cpuqos)#exit

Dell(conf)#ipv6 access-list ipv6-vrrp cpu-qos
Dell(conf-ipv6-acl-cpuqos)#permit vrrp
Dell(conf-ipv6-acl-cpuqos)#exit

Dell(conf)#qos-policy-in rate_limit_200k cpu-qos
Dell(conf-in-qos-policy-cpuqos)#rate-police 200 40 peak 500 40
Dell(conf-in-qos-policy-cpuqos)#exit

Dell(conf)#qos-policy-in rate_limit_400k cpu-qos
Dell(conf-in-qos-policy-cpuqos)#rate-police 400 50 peak 600 50
Dell(conf-in-qos-policy-cpuqos)#exit

Dell(conf)#qos-policy-in rate_limit_500k cpu-qos
Dell(conf-in-qos-policy-cpuqos)#rate-police 500 50 peak 1000 50
Dell(conf-in-qos-policy-cpuqos)#exit

Dell(conf)#class-map match-any class_ospf cpu-qos
Dell(conf-class-map-cpuqos)#match ip access-group ospf
Dell(conf-class-map-cpuqos)#exit

Dell(conf)#class-map match-any class_bgp cpu-qos
Dell(conf-class-map-cpuqos)#match ip access-group bgp
Dell(conf-class-map-cpuqos)#exit

Dell(conf)#class-map match-any class_lacp cpu-qos
Dell(conf-class-map-cpuqos)#match mac access-group lacp
Dell(conf-class-map-cpuqos)#exit

Dell(conf)#class-map match-any class-ipv6-icmp cpu-qos
Dell(conf-class-map-cpuqos)#match ipv6 access-group ipv6-icmp
Dell(conf-class-map-cpuqos)#exit

Dell(conf)#policy-map-input egressFP_rate_policy cpu-qos
Dell(conf-policy-map-in-cpuqos)#class-map class_ospf qos-policy rate_limit_500k
Dell(conf-policy-map-in-cpuqos)#class-map class_bgp qos-policy rate_limit_400k
Dell(conf-policy-map-in-cpuqos)#class-map class_lacp qos-policy rate_limit_200k
```

```
Dell(conf-policy-map-in-cpuqos)#class-map class-ipv6 qos-policy rate_limit_200k
Dell(conf-policy-map-in-cpuqos)#exit

Dell(conf)#control-plane-cpuqos
Dell(conf-control-cpuqos)#service-policy rate-limit-protocols
egressFP_rate_policy
Dell(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 0 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 name
   ```

Example of Creating the QoS Policy

Example of Assigning the QoS Policy to the Queues

Example of Creating the Control Plane Service Policy

```
Dell#conf
Dell(conf)#qos-policy-input cpuq_1
Dell(conf-qos-policy-in)#rate-police 3000 40 peak 500 40
Dell(conf-qos-policy-in)#exit

Dell(conf)#qos-policy-input cpuq_2
Dell(conf-qos-policy-in)#rate-police 5000 80 peak 600 50
Dell(conf-qos-policy-in)#exit

Dell(conf)#policy-map-input cpuq_rate_policy cpu-qos
Dell(conf-qos-policy-in)#service-queue 5 qos-policy cpuq_1
Dell(conf-qos-policy-in)#service-queue 6 qos-policy cpuq_2
Dell(conf-qos-policy-in)#service-queue 7 qos-policy cpuq_1

Dell#conf
Dell(conf)#control-plane
Dell(conf-control-plane)#service-policy rate-limit-cpu-queues cpuq_rate_policy
```

Show Commands

The following section describes the CoPP show commands.

To view the rates for each queue, use the `show cpu-queue rate cp` command.

Example of Viewing Queue Rates

Example of Viewing Queue Mapping

Example of Viewing Queue Mapping for MAC Protocols

Example of Viewing Queue Mapping for IPv6 Protocols

```
Dell#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
Dell#
```

To view the queue mapping for each configured protocol, use the `show ip protocol-queue-mapping` command.

```
Dell#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      _
Dell#
```

To view the queue mapping for the MAC protocols, use the `show mac protocol-queue-mapping` command.

```
Dell#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
```

```
Dell#
```

To view the queue mapping for IPv6 protocols, use the `show ipv6 protocol-queue-mapping` command.

```
Dell#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     _

Dell#
```

13

Data Center Bridging (DCB)

On an I/O Aggregator, data center bridging (DCB) features are auto-configured in standalone mode. You can display information on DCB operation by using **show** commands.

NOTE: DCB features are not supported on an Aggregator in stacking mode.

Supported Modes

Standalone, Stacking, PMUX, VLT, and Full–Switch

Ethernet Enhancements in Data Center Bridging

The following section describes DCB.

- The device supports the following DCB features:
  - Data center bridging exchange protocol (DCBx)
  - Priority-based flow control (PFC)
  - Enhanced transmission selection (ETS)

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, 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 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:

- 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 based on Fibre Channel media uses the 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.

- Servers use 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: In Dell Networking OS version 9.4.0.x, only the PFC, ETS, and DCBx features are supported 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 large amounts of queued LAN traffic do not cause storage traffic to be dropped, and that storage traffic does not result in high latency for high-performance computing (HPC) traffic between servers.

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 3.

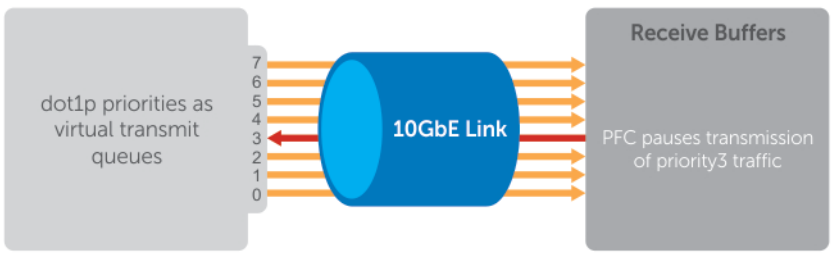


Figure 28. Priority-Based Flow Control

In the system, 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.

- 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 enabled DCB. When you enable DCB globally, you cannot simultaneously enable TX and RX on the interface for flow control and link-level flow control is disabled.
- Buffer space is allocated and de-allocated only when you configure a PFC priority on the port.
- PFC delay constraints place an upper limit on the transmit time of a queue after receiving a message to pause a specified priority.
- By default, PFC is enabled on an interface with no dot1p priorities configured. You can configure the PFC priorities if the switch negotiates with a remote peer using DCBX. During DCBX negotiation with a remote peer:
  - DCBx communicates with the remote peer by link layer discovery protocol (LLDP) type, length, value (TLV) to determine current policies, such as PFC support and enhanced transmission selection (ETS) BW allocation.
  - If the negotiation succeeds and the port is in DCBX Willing mode to receive a peer configuration, PFC parameters from the peer are used to configured PFC priorities on the port. If you enable the link-level flow control mechanism on the interface, DCBX negotiation with a peer is not performed.
  - If the negotiation fails and PFC is enabled on the port, any user-configured PFC input policies are applied. If no PFC dcb-map has been previously applied, the PFC default setting is used (no priorities configured). If you do not enable PFC on an interface, you can enable the 802.3x link-level pause function. By default, the link-level pause is disabled, when you disable DCBx and PFC. If no PFC dcb-map has been applied on the interface, the default PFC settings are used.
- PFC supports buffering to receive data that continues to arrive on an interface while the remote system reacts to the PFC operation.
- PFC uses the DCB MIB IEEE802.1azd2.5 and the PFC MIB IEEE802.1bb-d2.2.

If DCBx negotiation is not successful (for example, due to a version or TLV mismatch), DCBx is disabled and you cannot enable PFC or ETS.

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.

Although you can configure strict-priority queue scheduling for a priority group, ETS introduces flexibility that allows the bandwidth allocated to each priority group to be dynamically managed according to the amount of LAN, storage, and server traffic in a flow. Unused bandwidth is dynamically allocated to prioritized priority groups. Traffic is queued according to its 802.1p priority assignment, while flexible bandwidth allocation and the configured queue-scheduling for a priority group is supported.

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.

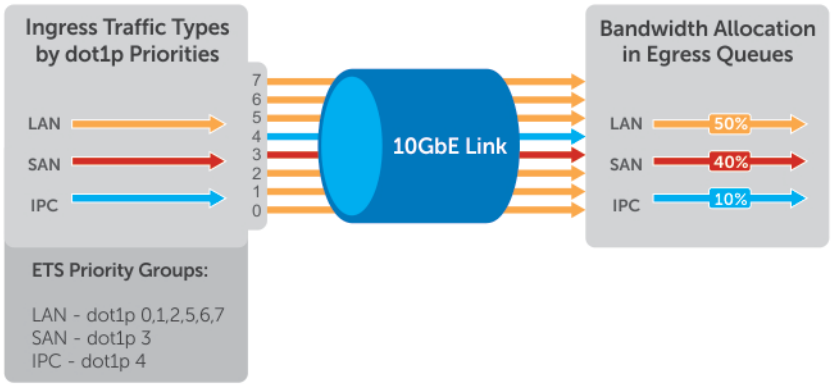


Figure 29. Enhanced Transmission Selection

The following table lists the traffic groupings ETS uses to select multiprotocol traffic for transmission.

Table 9. ETS Traffic Groupings

Traffic Groupings	Description
Priority group	A group of 802.1p priorities used for bandwidth allocation and queue scheduling. All 802.1p priority traffic in a group must have the same traffic handling requirements for latency and frame loss.
Group ID	A 4-bit identifier assigned to each priority group. The range is from 0 to 7.
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 the Dell 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
- Bandwidth allocated by the ETS algorithm is made available after strict-priority groups are serviced. If a priority group does not use its allocated bandwidth, the unused bandwidth is made available to other priority groups so that the sum of the bandwidth use is 100%. If priority group bandwidth use exceeds 100%, all configured priority group bandwidth is decremented based on the configured percentage ratio until all priority group bandwidth use is 100%. If priority group bandwidth usage is less than or equal to 100% and any default priority groups exist, a minimum of 1% bandwidth use is assigned by decreasing 1% of bandwidth from the other priority groups until priority group bandwidth use is 100%.
- For ETS traffic selection, an algorithm is applied to priority groups using:
  - Strict priority shaping

  - ETS shaping
  - (Credit-based shaping is not supported)
- ETS uses the DCB MIB IEEE 802.1azd2.5.

Data Center Bridging Exchange Protocol (DCBx)

The data center bridging exchange (DCBx) protocol is disabled by default on any switch on which PFC or ETS are enabled.

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. For more information, refer to Link Layer Discovery Protocol (LLDP). 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.

Creating a DCB Map

Configure the priority-based flow control (PFC) and enhanced traffic selection (ETS) settings in a DCB map before you can apply them on downstream server-facing ports.

NOTE: This feature is supported only on PMUX mode.

1. Create a DCB map to specify the PFC and ETS settings for groups of dot1p priorities.
   CONFIGURATION mode

   ```
   dcb-map name
   ```
2. Configure the PFC setting (on or off) and the ETS bandwidth percentage allocated to traffic in each priority group or whether priority group traffic should be handled with strict priority scheduling.
   DCB MAP mode

   ```
   priority-group group_num {bandwidth percentage | strict-priority} pfc {on | off}
   ```
3. Specify the priority group ID number to handle VLAN traffic for each dot1p class-of-service: 0 through 7. Leave a space between each priority group number.
   DCB MAP mode

   ```
   priority-pgid dot1p0_group-num dot1p1_group-num dot1p2_group-num dot1p3_group-num dot1p4_group-num dot1p5_group-num dot1p6_group-num dot1p7_group-num
   ```

Important Points to Remember

- If you remove a dot1p priority-to-priority group mapping from a DCB map (the `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.
- 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.

Applying a DCB Map on Server-Facing Ethernet Ports

You can apply a DCB map only on a physical Ethernet interface and can apply only one DCB map per interface.

1. Enter Interface Configuration mode on a server-facing port to apply a DCB map.
   CONFIGURATION mode

   ```
   interface tengigabitEthernet slot/port
   ```
2. Apply the DCB map on an Ethernet port. Repeat this step to apply a DCB map to more than one port.
   INTERFACE mode

   ```
   dcb-map name
   ```

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:

```
Dell(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
Dell(conf)#
```

NOTE: In Dell Networking OS we support 4 data queues.

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.

Data Center Bridging in a Traffic Flow

The following figure shows how DCB handles a traffic flow on an interface.

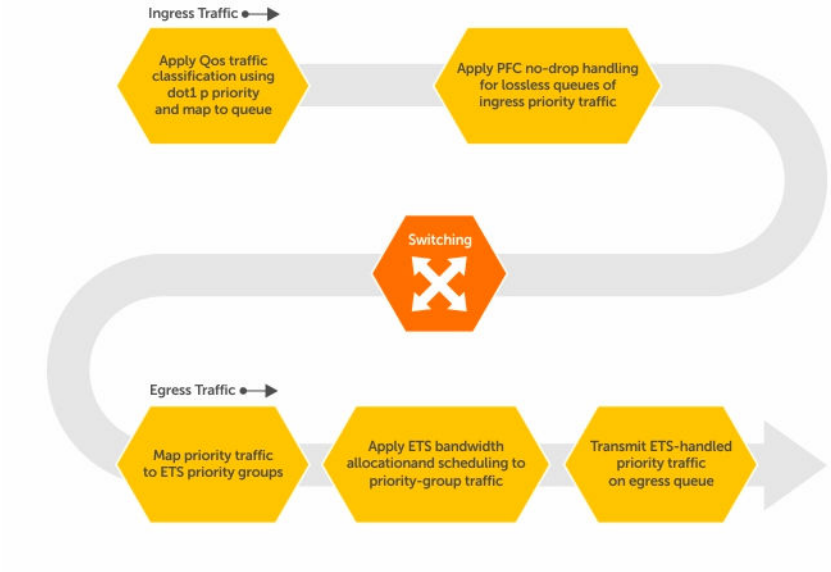


Figure 30. DCB PFC and ETS Traffic Handling

Data Center Bridging: Auto-DCB-Enable Mode

On an Aggregator in standalone or VLT modes, the default mode of operation for data center bridging on Ethernet ports is auto-DCB-enable mode. In this mode, Aggregator ports detect whether peer devices support CEE or not, and enable ETS and PFC or link-level flow control accordingly:

- Interfaces come up with DCB disabled and link-level flow control enabled to control data transmission between the Aggregator and other network devices (see Flow Control Using Ethernet Pause Frames). When DCB is disabled on an interface, PFC, and ETS are also disabled.
- When DCBx protocol packets are received, interfaces automatically enable DCB and disable link-level flow control.

DCB is required for PFC, ETS, DCBx, and FCoE initialization protocol (FIP) snooping to operate.

NOTE: Normally, interfaces do not flap when DCB is automatically enabled.

DCB processes VLAN-tagged packets and dot1p priority values. Untagged packets are treated with a dot1p priority of 0.

For DCB to operate effectively, ingress traffic is classified according to its dot1p priority so that it maps to different data queues. The dot1p-queue assignments used on an Aggregator are shown in Table 6-1 in dcb enable auto-detect on-next-reload Command Example QoS dot1p Traffic Classification and Queue Assignment.

When DCB is Disabled (Default) By default, Aggregator interfaces operate with DCB disabled and link-level flow control enabled. When an interface comes up, it is automatically configured with:

- Flow control enabled on input interfaces.
- A DCB-MAP policy is applied with PFC disabled.

The following example shows a default interface configuration with DCB disabled and link-level flow control enabled.

show interfaces Command Example: DCB disabled and Flow Control enabled

```
Dell#show running-config interface te 0/4
!
interface TenGigabitEthernet 0/4
 mtu 12000
 portmode hybrid
 switchport
 auto vlan
 flowcontrol rx on tx off
 dcb-map DCB_MAP_PFC_OFF
!
 protocol lldp
  advertise management-tlv management-address system-name
  dcbx port-role auto-downstream
 no shutdown
Dell#
```

When DCB is Enabled When an interface receives a DCBx protocol packet, it automatically enables DCB and disables link-level flow control. The dcb-map and flow control configurations are removed as shown in the following example.

show interfaces Command Example: DCB enabled and Flow Control disabled

```
Dell#show running-config interface te 0/3
!
interface TenGigabitEthernet 0/3
 mtu 12000
 portmode hybrid
 switchport
 auto vlan
!
 protocol lldp
  advertise management-tlv management-address system-name
  dcbx port-role auto-downstream
 no shutdown
Dell#
```

When no DCBx TLVs are received on a DCB-enabled interface for 180 seconds, DCB is automatically disabled and flow control is re-enabled.

Lossless Traffic Handling In auto-DCB-enable mode, Aggregator ports operate with the auto-detection of DCBx traffic. At any moment, some ports may operate with link-level flow control while others operate with DCB-based PFC enabled.

As a result, lossless traffic is ensured only if traffic ingresses on a PFC-enabled port and egresses on another PFC-enabled port.

Lossless traffic is not guaranteed when it is transmitted on a PFC-enabled port and received on a link-level flow control-enabled port, or transmitted on a link-level flow control-enabled port and received on a PFC-enabled port.

Enabling DCB on Next Reload To configure the Aggregator so that all interfaces come up with DCB enabled and flow control disabled, use the dcb enable on-next-reload command. Internal PFC buffers are automatically configured.

1. Globally enable DCB on all interfaces after next switch reload.

   CONFIGURATION mode

   ```
   dcb enable on-next-reload
   ```

To reconfigure the Aggregator so that all interfaces come up with DCB disabled and link-level flow control enabled, use the `no dcb enable on-next-reload` command. PFC buffer memory is automatically freed.

Enabling Auto-DCB-Enable Mode on Next Reload To configure the Aggregator so that all interfaces come up in auto-DCB-enable mode with DCB disabled and flow control enabled, use the `dcb enable aut-detect on-next-reload` command.

1. Globally enable auto-detection of DCBx and auto-enabling of DCB on all interfaces after switch reload.

   CONFIGURATION mode

   ```
   dcb enable auto-detect on-next-reload
   ```

Enabling DCB To configure the Aggregator so that all interfaces are DCB enabled and flow control disabled, use the `dcb enable` command.

Disabling DCB To configure the Aggregator so that all interfaces are DCB disabled and flow control enabled, use the `no dcb enable` command.

dcb enable auto-detect on-next-reload Command Example

```
Dell#dcb enable auto-detect on-next-reload
```

Configuring 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 ensure complete no-drop service, apply the same dcb-map on all PFC and ETS enabled interfaces.

1. Create a DCB map to apply priority based flow control or enhanced transmission selection for specified priority groups and priorities.
   CONFIGURATION mode

   ```
   dcb-map map-name
   ```

   The maximum is 32 alphanumeric characters.
2. Configure the priority group with PGID, bandwidth percentage or strict priority for ETS and PFC mode.
   DCB-MAP mode

   ```
   priority-group pg_num [bandwidth percentage | strict-priority] pfc [on | off]
   ```

   pg_num range is from 0 to 7.

   `bandwidth` *percentage* range is from 1 to 100.

   Either strict-priority or bandwidth percentage can be set for ETS on the priority group.

   PFC can either be enabled or disabled for the priority group.
3. Configure the priorities to priority group.
   DCB-MAP mode

   ```
   priority-pgid <pgid> <pgid> <pgid> <pgid> <pgid> <pgid> <pgid> <pgid>
   ```

   pgid range is from 0 to 7.

   Configure priority to priority group mapping from priority 0 to priority 7 in order.
4. Exit the DCB MAP configuration mode.
   DCB-MAP mode

   ```
   exit
   ```
5. Enter interface configuration mode.
   CONFIGURATION mode

   ```
   interface type slot/port
   ```
6. Apply the dcb-map with PFC and ETS configurations to both ingress and egress interfaces.
   INTERFACE mode

   ```
   dcb-map map-name
   ```
7. Repeat steps 1 to 6 on all PFC and ETS enabled interfaces to ensure lossless traffic service.

   NOTE: All these configurations are available only in PMUX mode and you cannot perform these configurations in Standalone mode.

Dell Networking OS Behavior: As soon as you apply a DCB MAPwith 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.

By applying a DCB MAP with PFC enabled, you enable PFC operation on ingress port traffic. To achieve complete lossless handling of traffic, also enable PFC on all DCB egress ports or configure the dot1p priority-queue assignment of PFC priorities to lossless queues.

To remove a DCB MAP, including the PFC and ETS configurations it contains, use the `no dcb-map` *`map-name`* command in INTERFACE Configuration mode.

You can enable any number of 802.1p priorities for PFC. Queues to which PFC priority traffic is mapped are lossless by default. Traffic may be interrupted due to an interface flap (going down and coming up) when you reconfigure the lossless queues for no-drop priorities in a PFC dcb-map and reapply the policy to an interface.

To apply PFC, a PFC peer must support the configured priority traffic (as detected by DCBx).

To honor a PFC pause frame multiplied by the number of PFC-enabled ingress ports, the minimum link delay must be greater than the round-trip transmission time the peer requres.

If you apply dcb-map with PFC disabled (`no pfc mode on`):

- You can enable link-level flow control on the interface. To delete the dcb-map, first disable link-level flow control. PFC is then automatically enabled on the interface because an interface is by default PFC-enabled.
- PFC still allows you to configure lossless queues on a port to ensure no-drop handling of lossless traffic.

NOTE: You cannot enable PFC and link-level flow control at the same time on an interface.

When you apply a dcb-map to an interface, an error message displays if:

- The PFC dot1p priorities result in more than two lossless port queues globally on the switch.
- Link-level flow control is already enabled. You cannot be 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.

A DCB MAP for PFC applied to an interface may become invalid if you reconfigure dot1p-queue mapping. This situation occurs when the new dot1p-queue assignment exceeds the maximum number (2) of lossless queues supported globally on the switch. In this case, all PFC configurations received from PFC-enabled peers are removed and resynchronized with the peer devices.

Traffic may be interrupted when you reconfigure PFC no-drop priorities in a dcb-map or reapply the dcb-map to an interface.

NOTE: All these configurations are available only in PMUX mode and you cannot perform these configurations in Standalone mode.

How Priority-Based Flow Control is Implemented

Priority-based flow control provides a flow control mechanism based on the 802.1p priorities in converged Ethernet traffic received on an interface and is enabled by default. As an enhancement to the existing Ethernet pause mechanism, 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 creates zero-loss links for SAN traffic that requires no-drop service, while at the same time retaining packet-drop congestion management for LAN traffic.

PFC is implemented on an Aggregator as follows:

- If DCB is enabled, as soon as a dcb-map with PFC is applied on an interface, DCBx starts exchanging information with PFC-enabled peers. The IEEE802.1Qbb, CEE and CIN versions of PFC TLV are supported. DCBx also validates PFC configurations received in TLVs from peer devices.
- To achieve complete lossless handling of traffic, enable PFC operation on ingress port traffic and on all DCB egress port traffic.
- All 802.1p priorities are enabled for PFC. Queues to which PFC priority traffic is mapped are lossless by default. Traffic may be interrupted due to an interface flap (going down and coming up).
- For PFC to be applied on an Aggregator port, the auto-configured priority traffic must be supported by a PFC peer (as detected by DCBx).
- A dcb-map for PFC applied to an interface may become invalid if dot1p-queue mapping is reconfigured. This situation occurs when the new dot1p-queue assignment exceeds the maximum number (2) of lossless queues supported globally on the switch. In this case, all PFC configurations received from PFC-enabled peers are removed and re-synchronized with the peer devices.
- Dell Networking OS does not support MACsec Bypass Capability (MBC).

Configuring Lossless Queues

DCB also supports the manual configuration of lossless queues on an interface when PFC mode is turned off and priority classes are disabled in a DCB map, apply the map on the interface.
Prerequisite: A DCB input policy 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:

```
Dell(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.
```

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 type slot/port
   ```
2. Configure the port queues that still functions 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 four.

   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 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. Reconfigure the input policy using a smaller number of PFC priorities.

If you configure lossless queues on an interface that already has a DCB input policy with PFC enabled (`pfc mode on`), an error message displays.

Traffic may be interrupted due to an interface flap (going down and coming up) when you reconfigure lossless queues on no-drop priorities in an input policy and re-apply the policy to an interface.

Configuring the PFC Buffer in a Switch Stack

In a switch stack, you must configure all stacked ports with the same PFC configuration. In addition, you must configure a separate buffer of memory allocated exclusively to a service pool accessed by queues on which priority-based control flows are mapped.

These PFC-enabled queues ensure the lossless transmission of storage and server traffic. The buffer required for the PFC service pool is calculated based on the number of ports and port queues used by PFC traffic.

You can configure the size of the PFC buffer for all switches in a stack or all port pipes on a specified stack unit by entering the following commands on the master switch.

- Configure the PFC buffer for all switches in the stack.
  CONFIGURATION mode

  ```
  [no] dcb stack-unit all pfc-buffering pfc-port {1-56} pfc-queues {1-2}
  ```

  By default, the PFC buffer is enabled on all ports on the stack unit.

- Configure the PFC buffer for all port pipes in a specified stack unit by specifying the port-pipe number, number of PFC-enabled ports, and number of configured lossless queues.

  CONFIGURATION mode

  ```
  [no] dcb stack-unit stack-unit-id [port-set port-set-id] pfc-buffering pfc-ports {1-56} pfc-queues {1-2}
  ```

  Valid stack-unit IDs are 0 to 5.

  The only valid port-set ID (port-pipe number) is 0.

Dell Networking OS Behavior:

To achieve lossless PFC operation, the PFC port count and queue number used for the reserved buffer size that is created must be greater than or equal to the buffer size required for PFC-enabled ports and lossless queues on the switch.

For the PFC buffer configuration to take effect, you must reload the stack or a specified stack unit (use the `reload` command at EXEC Privilege level).

If you configure the PFC buffer on all stack units, delete the startup configuration on both the master and standby, and reload the stack, the new master (previously standby) generates the following syslog message for each stack unit when it boots up: `PFC_BUFFER_CONFIG_CHANGED is generated for all stack units`.

Configuring 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.

To configure ETS and apply an ETS dcb-map to an interface, you must follow the steps described in Configuring Priority-Based Flow Control.

How Enhanced Transmission Selection is Implemented

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, groups within an 802.1p priority class are auto-configured to provide 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.

NOTE: The IEEE 802.1Qaz, CEE, and CIN versions of ETS are supported.

ETS is implemented on an Aggregator as follows:

- Traffic in priority groups is assigned to strict-queue or WERR scheduling in a dcb-map and is managed using the ETS bandwidth-assignment algorithm. Dell Networking OS de-qeues 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 scheduling apply only to data queues, not to control queues.
- Dell Networking OS supports hierarchical scheduling on an interface. Dell Networking OS control traffic is redirected to control queues as higher priority traffic with strict priority scheduling. After 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.
- By default, equal bandwidth is assigned to each port queue and each dot1p priority in a priority group.
- By default, equal bandwidth is assigned to each priority group in the dcb-map applied to an egress port. The sum of auto-configured bandwidth allocation to dot1p priority traffic in all ETS priority groups is 100%.
- dot1p priority traffic on the switch is scheduled according to the default dot1p-queue mapping. dot1p priorities within the same queue should have the same traffic properties and scheduling method.
- 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 should be in the same priority group.
  - By default:
    - All 802.1p priorities are grouped in priority group 0.
    - 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 (4) on the port. The 802.1p priorities in a priority group can map to multiple queues.
- A dcb-map is created to associate a priority group with a dcb-map with scheduling and bandwidth configuration, and applied on egress ports.
  - The ETS configuration associated with 802.1p priority traffic in a dcb-map is used in DCBx negotiation with ETS peers.
  - When a dcb-map is applied to an interface, ETS-configured scheduling and bandwidth allocation take precedence over any auto-configured settings in the QoS output policies.
  - ETS is enabled by default with the default ETS configuration applied (all dot1p priorities in the same group with equal bandwidth allocation).

ETS Operation with DCBx

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.
- ETS operational parameters are determined by the DCBX port-role configurations.
- ETS configurations received from TLVs from a peer are validated.
- In case of a hardware limitation or TLV error, the DCBx operation on an ETS port goes down.
- 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: Allows a single priority flow in a priority group to increase its bandwidth usage to the bandwidth total of the priority group. A single flow in a group can use all the bandwidth allocated to the group.
  * Link strict priority: Allows a flow in any priority group to increase to the maximum link bandwidth.

CIN supports only the default dot1p priority-queue assignment in a priority group.

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).

DCBx Operation

The data center bridging exchange protocol (DCBx) is used by DCB devices 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).

The following versions of DCBx are supported on an Aggregator: CIN, CEE, and IEEE2.5.

DCBx requires the LLDP to be enabled 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

The following DCBx port roles are auto-configured on an Aggregator to propagate DCB configurations learned from peer DCBx devices internally to other switch ports:

Auto-upstream The port advertises its own configuration to DCBx peers and receives its configuration from DCBx peers (ToR or FCF device). 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.

Default DCBx port role: Uplink ports are auto-configured in an auto-upstream role. Server-facing ports are auto-configured in an auto-downstream role.

NOTE: You can change the port roles only in the PMUX mode. Use the following command to change the port roles:

```
dcbx port-role {auto-downstream | auto-upstream | config-source | manual}
```

`manual` is the default port role.

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

On an Aggregator, the DCBx protocol supports the exchange and propagation of configuration information for the following DCB features.

- Enhanced transmission selection (ETS)
- Priority-based flow control (PFC)

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: 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 of the DCBx Version

The Aggregator operates in auto-detection mode so that 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).
- The peer times out.
- Multiple peers are detected on the link.

DCBx operations on a port are performed according to the auto-configured DCBx 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.

DCBx Example

The following figure shows how DCBx is used on an Aggregator installed in a Dell PowerEdge FX2 server chassis in which servers are also installed.

The Aggregator ports are numbered 1 to 12. Ports 1 to 8 are internal server-facing interfaces. Ports 9 to 12 are uplink ports. The uplink ports on the base module (ports 9 to 12) are used for uplinks configured as DCBx auto-upstream ports. The Aggregator is connected to third-party, top-of-rack (ToR) switches through the uplinks. The ToR switches are part of a Fibre Channel storage network.

The internal ports (ports 1 to 8) connected to the 10GbE backplane are configured as auto-downstream ports.

On the Aggregator, PFC and ETS use DCBx to exchange link-level configuration with DCBx peer devices.

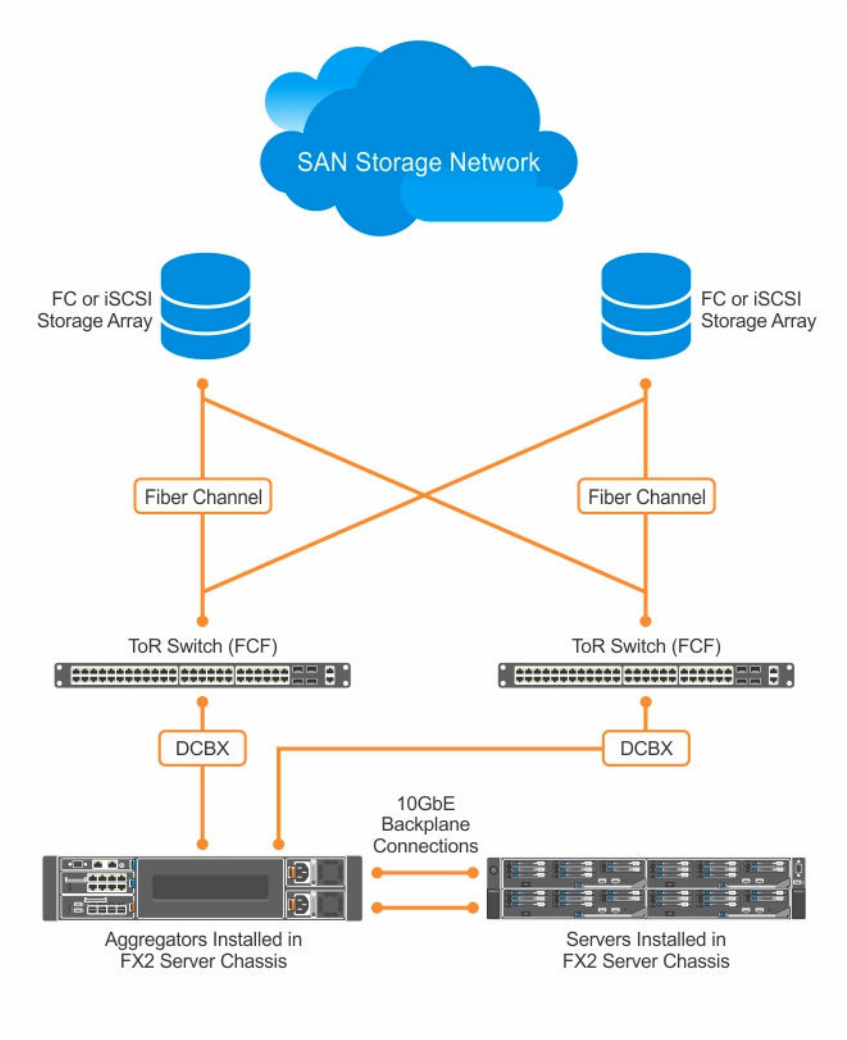


Figure 31. DCBx Sample Topology

DCBx Prerequisites and Restrictions

The following prerequisites and restrictions apply when you configure DCBx operation on a port:

- DCBx requires LLDP in both send (TX) and receive (RX) modes to be enabled on a port interface. 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 (LLD), and network interface virtualization (NIV).

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 10. Displaying DCB Configurations

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 interface port-type slot/port pfc statistics`	Displays counters for the PFC frames received and transmitted (by dot1p priority class) on an interface.
`show interface port-type slot/port 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 {stack-unit unit-number | tengigabitethernet slot/port}` command.
`show interface port-type slot/port 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 stack-unit unit-number` command.

Example of the `show dcb` Command

```
Dell(conf)#show dcb stack-unit 0 port-set 0

DCB Status: Enabled, PFC Queue Count: 2

stack-unit Total Buffer PFC Total Buffer PFC Shared Buffer PFC Available Buffer
      PP            (KB)             (KB)               (KB)                 (KB)
-------------------------------------------------------------------------------
--
0     0    3822         1912             832               450
Dell(conf)#
```

Example of the `show interface pfc statistics` Command

```
Dell#show interfaces tengigabitethernet 0/3 pfc statistics
Interface TenGigabitEthernet 0/3

Priority Rx XOFF Frames Rx Total Frames Tx Total Frames
-----------------------------------------------------
0        0              0               0
1        0              0               0
2        0              0               0
3        0              0               0
4        0              0               0
5        0              0               0
```

```
6        0              0              0
7        0              0              0
```

Example of the `show interfaces pfc summary` Command

```
Dell# show interfaces tengigabitethernet 0/4 pfc summary
Interface TenGigabitEthernet 0/4
    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

Dell# show interfaces tengigabitethernet 0/4 pfc detail
Interface TenGigabitEthernet 0/4
    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
2 Input Appln Priority TLV pkts, 0 Output Appln Priority TLV pkts, 0 Error
Appln Priority TLV Pkts
```

The following table describes the `show interface pfc summary` command fields.

Table 11. `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

Fields	Description
	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 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)	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).
Reason	Reason displayed when the DCBx operational status for PFC on a port is 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.

Fields	Description
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	Priority bitmap received from the remote DCBx port in FCoE advertisements in application priority TLVs.
Application Priority TLV: Remote ISCSI Priority Map	Priority bitmap received from the remote DCBx port in iSCSI advertisements in application priority TLVs.
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.
Input Appln Priority TLV pkts	Number of Application Priority TLVs received.
Output Appln Priority TLV pkts	Number of Application Priority TLVs transmitted.
Error Appln Priority TLV pkts	Number of Application Priority error packets received.

Example of the `show interface ets summary` Command

```
Dell# show interfaces te 0/1 ets summary
Interface TenGigabitEthernet 0/1
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
```

Example of the `show interface ets detail` Command

```
Dell# show interfaces tengigabitethernet 0/4 ets detail
Interface TenGigabitEthernet 0/4
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

Remote Parameters:
------------------
Remote is disabled

Local Parameters :
------------------
Local is enabled
PG-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

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


0 Input Conf TLV Pkts, 0 Output Conf TLV Pkts, 0 Error Conf TLV Pkts
0 Input Reco TLV Pkts, 0 Output Reco TLV Pkts, 0 Error Reco TLV Pkts
```

The following table describes the `show interface ets detail` command fields.

Table 12. `show interface ets detail` Command Description

Field	Description
Interface	Interface type with stack-unit and port number.
Max 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. When on, the scheduling and bandwidth allocation configured in an ETS output policy or received in a DCBx TLV from a peer can take effect on an interface.
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.
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.

Field	Description
	• 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.
Reason	Reason displayed when the DCBx operational status for ETS on a port is down.
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.
Reco TLV Tx Status	Status of ETS Recommendation TLV advertisements: enabled or disabled.
Input Conf TLV pkts, Output Conf TLV pkts, Error Conf TLV pkts	Number of ETS Configuration TLVs received and transmitted, and number of ETS Error Configuration TLVs received.
Input Reco TLV pkts, Output Reco TLV pkts, Error Reco TLV pkts	Number of ETS Recommendation TLVs received and transmitted, and number of ETS Error Recommendation TLVs received.

Example of the `show stack-unit all stack-ports all pfc details` Command

```
Dell# 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
```

Example of the `show stack-unit all stack-ports all ets details` Command

```
Dell# 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                          -            -
```

Example of the `show interface DCBx detail` Command

```
Dell# show interface tengigabitethernet 0/4 dcbx detail
Dell#show interface te 0/4 dcbx detail

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 0/4
  Remote Mac Address 00:00:00:00:00:11
     Port Role is Auto-Upstream
  DCBX Operational Status is Enabled
  Is Configuration Source? TRUE

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: 2
  Acknowledgment Number: 2
  Protocol State: In-Sync

Peer DCBX Status:
----------------
  DCBX Operational Version is 0
  DCBX Max Version Supported is 255
  Sequence Number: 2
  Acknowledgment Number: 2
2 Input PFC TLV pkts, 3 Output PFC TLV pkts, 0 Error PFC pkts, 0 PFC Pause Tx
pkts,
0 Pause Rx pkts
2 Input PG TLV Pkts, 3 Output PG TLV Pkts, 0 Error PG TLV Pkts
2 Input Appln Priority TLV pkts, 0 Output Appln Priority TLV pkts, 0 Error
Appln Priority
TLV Pkts
  Total DCBX Frames transmitted 27
  Total DCBX Frames received 6
  Total DCBX Frame errors 0
  Total DCBX Frames unrecognized 0
```

The following table describes the `show interface DCBx detail` command fields.

Table 13. `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 or auto-downstream.
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).

Field	Description
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.
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.
PFC TLV Statistics: Input PFC TLV pkts	Number of PFC TLVs received.
PFC TLV Statistics: Output PFC TLV pkts	Number of PFC TLVs transmitted.
PFC TLV Statistics: Error PFC pkts	Number of PFC error packets received.
PFC TLV Statistics: PFC Pause Tx pkts	Number of PFC pause frames transmitted.
PFC TLV Statistics: PFC Pause Rx pkts	Number of PFC pause frames received.
PG TLV Statistics: Input PG TLV Pkts	Number of PG TLVs received.
PG TLV Statistics: Output PG TLV Pkts	Number of PG TLVs transmitted.
PG TLV Statistics: Error PG TLV Pkts	Number of PG error packets received.
Application Priority TLV Statistics: Input Appln Priority TLV pkts	Number of Application TLVs received.
Application Priority TLV Statistics: Output Appln Priority TLV pkts	Number of Application TLVs transmitted.
Application Priority TLV Statistics: Error Appln Priority TLV Pkts	Number of Application TLV error packets received.

QoS dot1p Traffic Classification and Queue Assignment

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	dot1p priorities in ingress traffic are used at the port or global switch level.
Layer 2 class maps	dot1p priorities are used to classify traffic in a class map and apply a service policy to an ingress port to map traffic to egress queues.

NOTE: Dell Networking does not recommend mapping all ingress traffic to a single queue when using PFC and ETS. However, Dell 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.

Although the system allows you to change the default dot1p priority-queue assignments, DCB policies applied to an interface may become invalid if you reconfigure dot1p-queue mapping. If the configured dcb-map remains valid, the change in the dot1p-queue assignment is allowed. For DCB ETS enabled interfaces, traffic destined to queue that is not mapped to any dot1p priority are dropped.

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

Troubleshooting PFC, ETS, and DCBx Operation

In the `show interfaces pfc | ets | dcbx` output, the DCBx operational status may be down for any of the reasons described in the following table.

When DCBx is down, the following values display in the `show` output field for DCBx Oper status:

- PFC DCBx Oper status: Down
- ETS DCBx Oper status: Down
- DCBx Oper status: Disabled.

Reason	Description
Port Shutdown	Port is shut down. All other reasons for DCBx inoperation, if any, are ignored.
LLDP Rx/Tx is disabled	LLDP is disabled (Admin Mode set to rx or tx only) globally or on the interface.
Waiting for Peer	Waiting for peer or detected peer connection has aged out.
Multiple Peer Detected	Multiple peer connections detected on the interface.
Version Conflict	DCBx version on peer version is different than the local or globally configured DCBx version.
Config-source Down	Although DCBx parameters match in auto-upstream or auto-downstream port, the configuration source (elected dynamically or CLI-configured) is down.
Unrecognized TLV received	Invalid TLV length TLV has been received. In the case of an invalid PFC or ETS TLV, the error displays in the `show interfaces pfc | ets` output. The `show interfaces dcbx` output displays PFC or ETS as down.
Admin Mode OFF (Local)	Local Admin Mode is disabled.
Admin Mode OFF (Remote)	Remote Admin Mode is disabled.
Waiting for ACK from Peer	For a legacy DCBx version, a peer has not acknowledged the reception of a sent packet. This reason displays only when a remote peer is willing to receive a DCB configuration.
Error Bit set	For a legacy DCBx version, a peer has sent packets with an error bit set. This reason displays only when a remote peer is willing to receive a DCB configuration.
Enabled with ETS Mismatch (**`show interfaces dcbx`** output)	DCBx is enabled but an ETS validation failure error has occurred.
PFC is down (**`show interfaces pfc`** output)	One of the following PFC-specific errors has occurred: • No MBC support. • Configured PFC priorities exceed maximum PFC capability limit. • New dot1p-to-queue mapping violates the allowed system limit for PFC Enable status per priority
ETS is down (**`show interfaces ets`** output)	One of the following ETS-specific errors occurred in ETS validation: • Unsupported PGID • A priority group exceeds the maximum number of supported priorities. • COSQ is mapped to more than one priority group. • COSQ is mapped to more than one priority group. - Invalid or unsupported transmission selection algorithm (TSA). • Bandwidth is configured for an unconfigured priority group. • Total ETS bandwidth mapped in priority groups is not equal to 100%.

Reason	Description
	• Priorities mapped to a queue use different TSAs. • Total bandwidth assigned to priorities in one or more priority groups is not equal to 100%. Or one of the following ETS failure errors occurred: • Incompatible priority group ID (PGID). • Incompatible bandwidth (BW) allocation. • Incompatible TSA. • Incompatible TC BW. • Incompatible TC TSA.

14

Dynamic Host Configuration Protocol (DHCP)

The Aggregator is auto-configured to operate as a dynamic host configuration protocol (DHCP) client. The DHCP server, DHCP relay agent, and secure DHCP features are not supported. The 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.

NOTE: The DHCP server and relay agent features are not supported on an Aggregator.

Supported Modes

Stacking, PMUX, Standalone, VLT

Assigning 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.

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.
DHCPRELEASE	A DHCP client sends this message when it is stopped forcefully to return its IP address to the server.

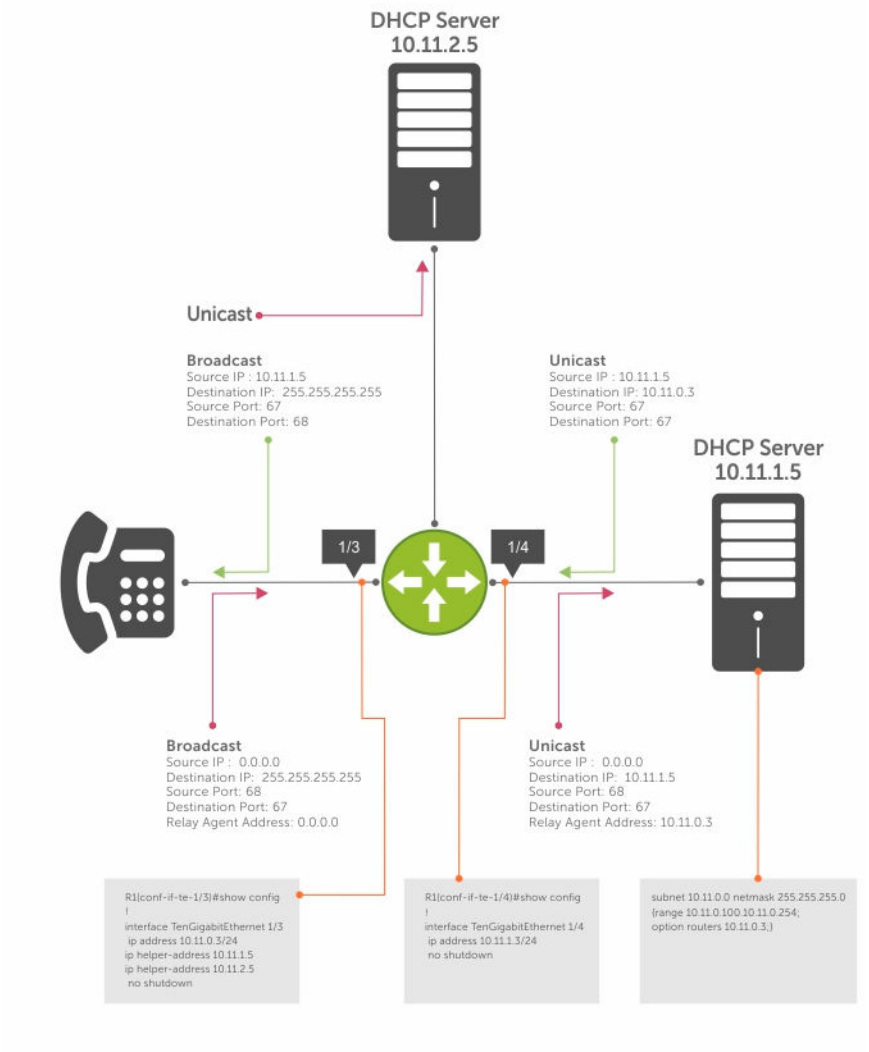


Figure 32. Assigning Network Parameters using DHCP

Dell Networking OS Behavior: DHCP is implemented in Dell Networking OS based on RFC 2131 and 3046.

Debugging DHCP Client Operation

To enable debug messages for DHCP client operation, enter the following **debug** commands:

- Enable the display of log messages for all DHCP packets sent and received on DHCP client interfaces.
  EXEC Privilege

  ```
  [no] debug ip dhcp client packets [interface type slot/port]
  ```
- Enable the display of log messages for the following events on DHCP client interfaces: IP address acquisition, IP address release, Renewal of IP address and lease time, and Release of an IP address.
  EXEC Privilege

  ```
  [no] debug ip dhcp client events [interface type slot/port]
  ```

The following example shows the packet- and event-level debug messages displayed for the packet transmissions and state transitions on a DHCP client interface.

DHCP Client: Debug Messages Logged during DHCP Client Enabling/Disabling

```
Dell (conf-if-Ma-0/0)# ip address dhcp
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :DHCP ENABLE
CMD
Received in state START
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :Transitioned
to state SELECTING
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:  DHCP
DISCOVER sent in
Interface
Ma 0/0
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:  Received
DHCPOFFER packet in
Interface Ma 0/0 with Lease-ip:10.16.134.250, Mask:255.255.0.0,Server-Id:
10.16.134.249
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :Transitioned
to state REQUESTING
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:DHCP REQUEST
sent in Interface
Ma 0/0
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:Received
DHCPACK packet in
Interface
Ma 0/0 with Lease-IP:10.16.134.250, Mask:255.255.0.0,DHCP REQUEST sent in
Interface Ma 0/0
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :Transitioned
to state BOUND



Dell(conf-if-ma-0/0)# no ip address
Dell(conf-if-ma-0/0)#1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG:
DHCLIENT_DBG_EVT: Interface
Ma 0/0 :DHCP DISABLE CMD Received in state SELECTING
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :Transitioned
to state START
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :DHCP DISABLED
CMD
sent to FTOS in state START



Dell# release dhcp int Ma 0/0
Dell#1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT:
Interface Ma 0/0 :DHCP
RELEASE
CMD Received in state BOUND
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT: DHCP
RELEASE sent in Interface
Ma 0/0
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :Transitioned
```

```
to state STOPPED
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :DHCP IP
RELEASED
CMD sent to FTOS in state STOPPED

Dell# renew dhcp int Ma 0/0
Dell#1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT:
Interface Ma 0/0 :DHCP
RENEW
CMD Received in state STOPPED
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT: Interface
Ma 0/0 :Transitioned
to state SELECTING
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT: DHCP
DISCOVER sent in Interface
Ma 0/0
1w2d23h: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT: Received
DHCPOFFER packet in
Interface Ma 0/0 with Lease-Ip:10.16.134.250, Mask:255.255.0.0,Server-Id:
10.16.134.249
```

The following example shows the packet- and event-level debug messages displayed for the packet transmissions and state transitions on a DHCP client interface when you release and renew a DHCP client.

DHCP Client: Debug Messages Logged during DHCP Client Release/Renew

```
Dell# release dhcp interface managementethernet 0/0
May 27 15:55:22: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT:
Interface Ma 0/0 :
DHCP RELEASE CMD Received in state BOUND
May 27 15:55:22: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:
DHCP RELEASE sent in Interface Ma 0/0
May 27 15:55:22: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT:
Interface Ma 0/0 :
Transitioned to state STOPPED
May 27 15:55:22: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT:
Interface Ma 0/0 :
DHCP IP RELEASED CMD sent to FTOS in state STOPPED



Dell# renew dhcp interface tengigabitethernet 0/1
Dell#May 27 15:55:28: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG:
DHCLIENT_DBG_EVT: Interface Ma 0/0 :
DHCP RENEW CMD Received in state STOPPED
May 27 15:55:31: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_EVT:
Interface Ma 0/0 :
Transitioned to state SELECTING
May 27 15:55:31: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:
DHCP DISCOVER sent in Interface Ma 0/0
May 27 15:55:31: %STKUNIT0-M:CP %DHCLIENT-5-DHCLIENT-LOG: DHCLIENT_DBG_PKT:
Received DHCPOFFER packet in Interface Ma 0/0 with Lease-Ip:10.16.134.250,
Mask:255.255.0.0,Server-Id:10.16.134.249
```

DHCP Client

An Aggregator is auto-configured to operate as a DHCP client. The DHCP client functionality is enabled only on the default VLAN and the management interface.

A DHCP client is a network device that requests an IP address and configuration parameters from a DHCP server. On an Aggregator, the DHCP client functionality is implemented as follows:

- The public out-of-band management (OOB) interface and default VLAN 1 are configured, by default, as a DHCP client to acquire a dynamic IP address from a DHCP server.

  You can override the DHCP-assigned address on the OOB management interface by manually configuring an IP address using the CLI or CMC interface. If no user-configured IP address exists for the OOB interface exists and if the OOB IP address is not in the startup configuration, the Aggregator will automatically obtain it using DHCP.

  You can also manually configure an IP address for the VLAN 1 default management interface using the CLI. If no user-configured IP address exists for the default VLAN management interface exists and if the default VLAN IP address is not in the startup configuration, the Aggregator will automatically obtain it using DHCP.
- The default VLAN 1 with all ports configured as members is the only L3 interface on the Aggregator. When the default management VLAN has a DHCP-assigned address and you reconfigure the default VLAN ID number, the Aggregator:
  - Sends a DHCP release to the DHCP server to release the IP address.
  - Sends a DHCP request to obtain a new IP address. The IP address assigned by the DHCP server is used for the new default management VLAN.

How DHCP Client is Implemented

The Aggregator is enabled by default to receive DHCP server-assigned dynamic IP addresses on an interface. This setting persists after a switch reboot. If you enter the `shutdown` command on the interface, DHCP transactions are stopped and the dynamically-acquired IP address is saved. Use the `show interface` *`type slot/port`* command to display the dynamic IP address and DHCP as the mode of IP address assignment. If you later enter the `no shutdown` command and the lease timer for the dynamic IP address has expired, the IP address is unconfigured and the interface tries to acquire a new dynamic address from DHCP server.

If you later enter the `no shutdown` command and the lease timer for the dynamic IP address has expired, the IP address is released.

When you enter the `release dhcp` command, although the IP address that was dynamically-acquired from a DHCP server is released from an interface, the ability to acquire a new DHCP server-assigned address remains in the running configuration for the interface. To acquire a new IP address, enter either the `renew dhcp` command at the EXEC privilege level or the `ip address dhcp` command at the interface configuration level.

If you enter `renew dhcp` command on an interface already configured with a dynamic IP address, the lease time of the dynamically acquired IP address is renewed.

Important: To verify the currently configured dynamic IP address on an interface, enter the `show ip dhcp lease` command. The `show running-configuration` command output only displays ip address dhcp; the currently assigned dynamic IP address is not displayed.

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 of whether the DHCP client and server are on 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.

DHCP Client on a VLAN

The following conditions apply on a VLAN that operates as a DHCP client:

- The default VLAN 1 with all ports auto-configured as members is the only L3 interface on the Aggregator.
- When the default management VLAN has a DHCP-assigned address and you reconfigure the default VLAN ID number, the Aggregator:
  - Sends a DHCP release to the DHCP server to release the IP address.
  - Sends a DHCP request to obtain a new IP address. The IP address assigned by the DHCP server is used for the new default management VLAN.

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.

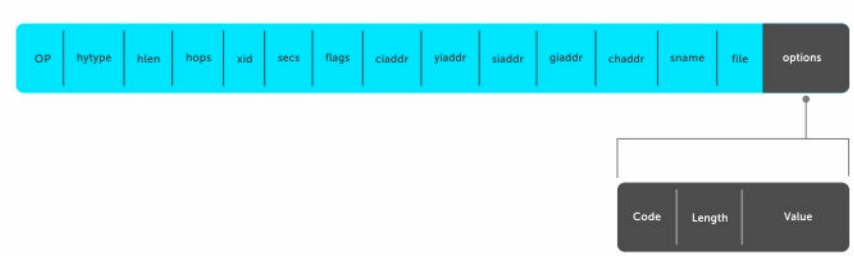


Figure 33. DHCP packet Format

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

Option	Number and Description
	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.
End	Option 255 Signals the last option in the DHCP packet.

Option 82

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 is comprised of two sub-options, 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 sub-option is the MAC address of the relay agent that adds Option 82.

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 can use 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.

To insert Option 82 into DHCP packets, follow this step.

- Insert Option 82 into DHCP packets.
  CONFIGURATION mode

```
int ma 0/0
ip add dhcp relay information-option remote-id
```

For routers between the relay agent and the DHCP server, enter the `trust-downstream` option.

Releasing and Renewing DHCP-based IP Addresses

On an Aggregator configured as a DHCP client, you can release a dynamically-assigned IP address without removing the DHCP client operation on the interface.
To manually acquire a new IP address from the DHCP server, use the following command.

- Release a dynamically-acquired IP address while retaining the DHCP client configuration on the interface.

  EXEC Privilege mode

  ```
  release dhcp interface type slot/port
  ```
- Acquire a new IP address with renewed lease time from a DHCP server.

  EXEC Privilege mode

  ```
  renew dhcp interface type slot/port
  ```

Viewing DHCP Statistics and Lease Information

To display DHCP client information, enter the following **show** commands:

- Display statistics about DHCP client interfaces.

  EXEC Privilege

  ```
  show ip dhcp client statistics interface type slot/port
  ```
- Clear DHCP client statistics on a specified or on all interfaces.

  EXEC Privilege

  ```
  clear ip dhcp client statistics {all | interface type slot/port}
  ```
- Display lease information about the dynamic IP address currently assigned to a DHCP client interface.

  EXEC Privilege

  ```
  show ip dhcp lease [interface type slot/port]
  ```

View the statistics about DHCP client interfaces with the **show ip dhcp client statistics** command and the lease information about the dynamic IP address currently assigned to a DHCP client interface with the **show ip dhcp lease** command.

Example of the show ip dhcp client statistics Command

```
Dell#show ip dhcp client statistics interface managementethernet 0/0
Interface Name          Ma 0/0
Message                     Received
DHCPOFFER                   0
DHCPACK                     0
DHCPNAK                     0
Message                     Sent
DHCPDISCOVER                1626
DHCPREQUEST                 0
DHCPDECLINE                 0
DHCPRELEASE                 0
DHCPREBIND                  0
DHCPRENEW                   0
```

```
DHCPINFORM                  0
Dell#
```

Example of the show ip dhcp lease Command

```
Dell# show ip dhcp
Interface  Lease-IP Def-Router  ServerId  State  Lease Obtnd At       Lease
Expires At
=========  ========  =========   ========  =====  ==============
================ Ma 0/0
0.0.0.0/0 0.0.0.0    0.0.0.0   INIT   -----NA-----         ----NA----

Vl 1   10.1.1.254/24 0.0.0.0   10.1.1.1   BOUND  08-26-2011 04:33:39
08-27-2011 04:33:39

Renew Time          Rebind Time
==========          ========
----NA----           ----NA----

08-26-2011 16:21:50  08-27-2011 01:33:39
```

15

Equal Cost Multi-Path (ECMP)

Dell Networking OS supports equal cost multi-path (ECMP).

ECMP for Flow-Based Affinity

Dell Networking OS supports ECMP for flow-based affinity.

NOTE: IPv6 /128 routes having multiple paths do not form ECMPs. The /128 route is treated as a host entry and finds its place in the host table.

NOTE: Using XOR algorithms results in imbalanced loads across an ECMP/LAG when the number of members in said ECMP/LAG is a multiple of 4.

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 sort 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
  ```

Link Bundle Monitoring

Monitoring linked ECMP 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 causes a syslog to be

sent and an alarm event to be generated. When the deviation clears, another syslog is sent and a clear alarm event is generated.

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, the utilization calculation is performed 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.

To enable the link bundle monitoring feature, for link bundle monitoring with ECMP, use the `ecmp-group` command.

You can configure the ecmp-group with id 2, enabled for link bundle monitoring. 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.

Managing ECMP Group Paths

Configure the maximum number of paths for an ECMP route that the L3 CAM can hold to avoid path degeneration.
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: Save the new ECMP settings to the startup-config (`write-mem`) then reload the system for the new settings to take effect.

- 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
  ```

Example of the `ip ecmp-group maximum-paths` Command

```
Dell(conf)#ip ecmp-group maximum-paths 3
User configuration has been changed. Save the configuration and reload to take
effect
Dell(conf)#
```

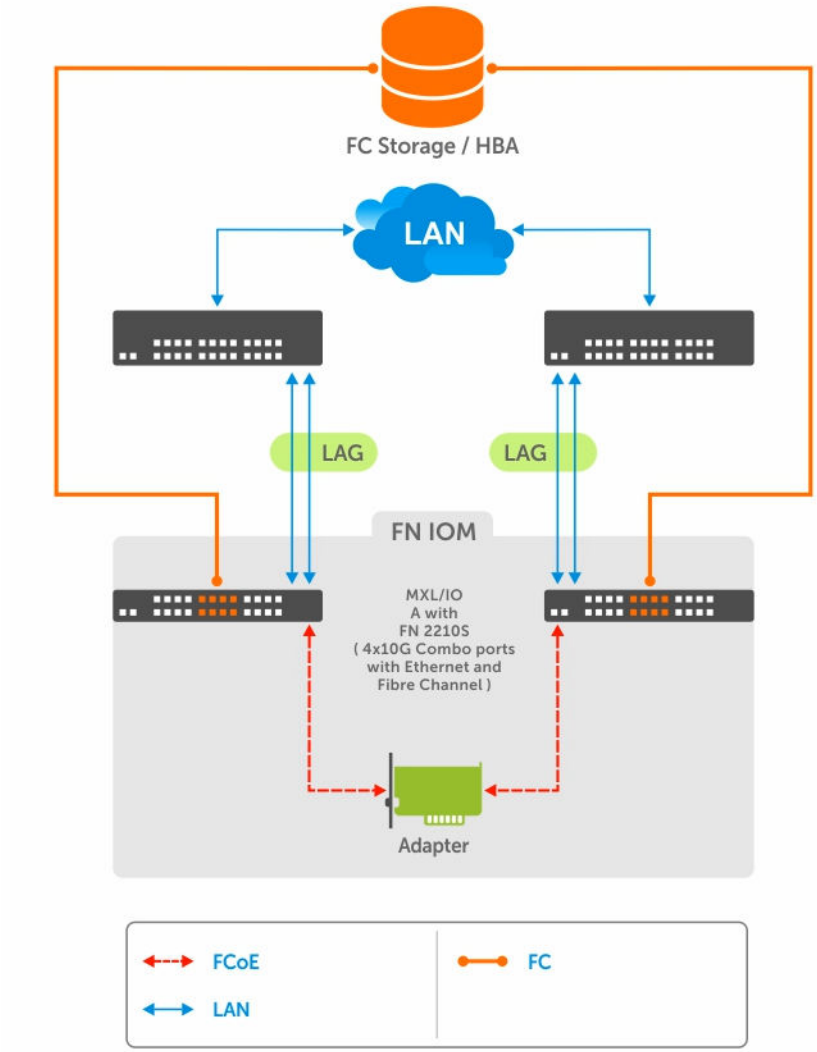
16

FC FPORT

Dell Networking OS supports FC FPort on Combo port card.

FC FPORT

FC FPort can be configured on ports 9 and 10 of Combo port card in Dell FX2 chassis.



Configuring Switch Mode to FCF Port Mode

To configure switch mode to Fabric services, use the following commands.

1. Configure Switch mode to FCF Port.
   CONFIGURATION mode

   ```
   feature fc fport domain id 2
   ```

   NOTE: Enable `remote-fault-signaling rx off` command in FCF FPort mode on interfaces connected to the Compellent and MDF storage devices.
2. Create an FCoE map with the parameters used in the communication between servers and a SAN fabric.
   CONFIGURATION mode

   ```
   fcoe-map map-name
   ```
3. Configure the association between the dedicated VLAN used to carry FCoE traffic between servers and a SAN, and the fabric where the desired storage arrays are installed.
   FCOE MAP mode

   ```
   fabric-id fabric-num vlan vlan-id
   ```
4. Configure the FCoE mapped address prefix (FC-MAP) value which is used to identify FCoE traffic transmitted on the FCoE VLAN for the specified fabric.
   FCOE MAP mode

   ```
   fc-map fc-map-value
   ```
5. Configure the SAN fabric to which the FC port connects by entering the name of the FCoE map applied to the interface.
   INTERFACE mode

   ```
   fcoe-map <fcoe-map-name> tengigabitEthernet slot/port
   ```

   The FCoE map contains FCoE and FC parameter settings (refer to FCoE Maps). Manually apply the `fcoe-map` to any Ethernet ports used for FCoE.

Name Server

Each participant in the FC environment has a unique ID, which is called the World Wide Name (WWN). This WWN is a 64-bit address.

A Fibre Channel fabric uses another addressing scheme to address the ports in the switched fabric. Each port in the switched fabric is assigned a 24-bit address by the FC switch. When the device logs into the switch fabric on a specific port with its WWN, the switch assigns the port address to that port, and the switch also maintains the correlation between the port address and the WWN address of the device on that port. This function of the switch is implemented by using a name server, a database of objects in which the fabric attached device registers its values.

FC-ID: 24-bit port address. Consists of three parts:

- **Domain**: Address of the switch itself. There are only 239 addresses available for switches in your SAN environment.
- **Area**: Identifies a group of F_Ports. Each group of ports has a different area number, even if there is only one port in the group.
- **Port**: Provides 256 addresses for identifying attached N_Ports and NL_Ports.

The following sequence explains the operation for the attached N_Port:

- N_Port sends a Fabric Login (FLOGI) as it requests a unique 24-bit address from the Fabric Login Server.
- N_Port sends FLOGI to address 0xFFFFFE. Upon success, it obtains a valid address (FCID).
- N_Port sends a Port Login (PLOGI) to inform the Fabric Name Server of its personality and capabilities, this includes WWNN, WWPN.
- N_Port sends PLOGI to address 0xFFFFFC to register this address with the name server.

Command	Description
show fc ns switch	Display all the devices in name server database of the switch.
show fc ns switch brief	Displays the local name server entries — brief version.

The following configurations are applicable only after configuring the switch mode to FCF Port mode using the `feature fc fport domain id 2` command. When you set Switch mode to FCF Port mode, any previously configured fcoe-map is removed.

FCoE Maps

To identify the SAN fabric to which FCoE storage traffic is sent, use an FCoE map.

Using an FCoE map, an NPG operates as an FCoE-FC bridge between an FC SAN and FCoE network by providing FCoE-enabled servers and switches with the necessary parameters to log in to a SAN fabric.

An FCoE map applies the following parameters on server-facing Ethernet and fabric-facing FC ports:

- The dedicated FCoE VLAN used to transport FCoE storage traffic.
- The FC-MAP value used to generate a fabric-provided MAC address.
- The association between the FCoE VLAN ID and FC fabric ID where the desired storage arrays are installed. Each Fibre Channel fabric serves as an isolated SAN topology within the same physical network.
- A server uses the priority to select an upstream FCoE forwarder (FCF priority).
- FIP keepalive (FKA) advertisement timeout.

NOTE: In each FCoE map, the fabric ID, FC-MAP value, and FCoE VLAN must be unique. To access one SAN fabric, use one FCoE map.

When you configure a switch as an NPG, FCoE transit with FIP snooping is automatically enabled and configured using the parameters in the FCoE map applied to server-facing Ethernet and fabric-facing FC interfaces.

After you apply an FCoE map on an FC port, when you enable the port (using the `no shutdown` command), the NPG starts sending FIP multicast advertisements on behalf of the FC port to downstream

servers to advertise the availability of a new FCF port on the FCoE VLAN. The FIP advertisement also contains a keepalive message to maintain connectivity between a SAN fabric and downstream servers.

NOTE: After removing and reapplying the fabric map or after modifying the FCoE map, the Fiber Channel (FC) devices do not re-login. To mitigate this issue, you must first run the shut command and then the no shutdown command on each member interface after you alter the FCOE map.

Creating an FCoE Map

An FCoE map consists of the following elements.

- An association between the dedicated VLAN used to carry FCoE traffic and SAN fabric where the storage arrays are installed. Use a separate FCoE VLAN for each fabric to which FCoE traffic is forwarded. Any non-FCoE traffic sent on a dedicated FCoE VLAN is dropped.
- The FC-MAP value used to generate the fabric-provided MAC address (FPMA). The server uses the FPMA to transmit FCoE traffic to the fabric. You can associate an FC-MAP with only one FCoE VLAN and vice versa.
- FCF priority: a CNA server uses the priority to select an upstream FCoE forwarder (FCF).
- FIP keepalive (FKA) advertisement timeout.

The values for the FCoE VLAN, fabric ID, and FC-MAP must be unique. Apply an FCoE map on downstream server-facing Ethernet ports and upstream fabric-facing Fibre Channel ports.

1. Create an FCoE map which contains parameters used in the communication between servers and a SAN fabric.

   CONFIGURATION mode

   ```
   fcoe-map map-name
   ```

2. Configure the association between the dedicated VLAN and the fabric where the desired storage arrays are installed.

   FCoE MAP mode

   ```
   fabric-id fabric-num vlan vlan-id
   ```

   The fabric and VLAN ID numbers must be the same.

   The fabric and VLAN ID range is from 2 to 4094.

   For example:

   ```
   fabric id 10 vlan 10
   ```

3. Add a text description of the settings in the FCoE map.

   FCoE MAP mode

   ```
   description text
   ```

   The maximum is 32 characters.

4. Specify the FC-MAP value used to generate a fabric-provided MAC address, which is required to send FCoE traffic from a server on the FCoE VLAN to the FC fabric specified in Step 2.

   FCoE MAP mode

   ```
   fc-map fc-map-value
   ```

   You must enter a unique MAC address prefix as the FC-MAP value for each fabric.

The range is from 0EFC00 to 0EFCFF.

The default is none.

5. Configure the priority used by a server CNA to select the FCF for a fabric login (FLOGI).
   FCoE MAP mode

   ```
   fcf-priority priority
   ```

   The range is from 1 to 255.

   The default is **128**.
6. Enable the monitoring FIP keep-alive messages (if it is disabled) to detect if other FCoE devices are reachable.
   FCoE MAP mode

   ```
   keepalive
   ```

   The default is FIP keep-alive monitoring is enabled.
7. Configure the time interval (in seconds) used to transmit FIP keepalive advertisements.
   FCoE MAP mode

   ```
   fka-adv-period seconds
   ```

   The range is from 8 to 90 seconds.

   The default is **8 seconds**.

Zoning

Dell Networking OS supports the zoning configurations for Fabric FCF Port mode operation.

In FCF Port mode, the fcoe-map `fabric map-name` has the default Zone mode set to deny. This setting denies all the fabric connections unless included in an active zoneset. To change this setting, use the `default-zone-allow` command. Changing this setting to all allows all the fabric connections without zoning.

Zoning is a mechanism to ensure only the nodes that are part of a zone can communicate with each other. Zoning prevents unauthorized access of storage assets. A zone consists of members which are nodes that the adapter address, fabric address, interface, or alias specifies. S5000 supports using World Wide Port Name (WWPN), Fibre Channel ID (FC-ID), or alias as members of a zone.

- WWPN: End device's port WWN name.
- FC-ID: Switch assigned 24-bit device FC address.
- Alias: User-defined name of a zone member.

NOTE: Dell Networking OS does not support using WWNN or Domain/Port as members of a zone.

Creating Zone and Adding Members

To create a zone and add members to the zone, use the following commands.

1. Create a zone.
   CONFIGURATION mode

   ```
   fc zone zonename
   ```
2. Add members to a zone.
   ZONE CONFIGURATION mode

   ```
   member word
   ```

   The member can be WWPN (00:00:00:00:00:00:00:00), port ID (000000), or alias name (word).

Example of Creating a Zone and Adding Members

```
Dell(conf)#fc zone z1
Dell(conf-fc-zone-z1)#member 11:11:11:11:11:11:11:11
Dell(conf-fc-zone-z1)#member 020202
Dell(conf-fc-zone-z1)#exit
```

Creating Zone Alias and Adding Members

To create a zone alias and add devices to the alias, follow these steps.

1. Create a zone alias name.
   CONFIGURATION mode

   ```
   fc alias ZoneAliasName
   ```
2. Add devices to an alias.
   ALIAS CONFIGURATION mode

   ```
   member word
   ```

   The member can be WWPN (00:00:00:00:00:00:00:00), port ID (000000), or alias name (word).

Example of Creating a Zone Alias and Adding Members

```
Dell(conf)#fc alias al1
Dell(conf-fc-alias-al1)#member 030303
Dell(conf-fc-alias-al1)#exit
Dell(conf)#fc zone z1
Dell(conf-fc-zone-z1)#member al1
Dell(conf-fc-zone-z1)#exit
```

Creating Zonesets

A zoneset is a grouping or configuration of zones.
To create a zoneset and zones into the zoneset, use the following steps.

1. Create a zoneset.
   CONFIGURATION mode

```
fc zoneset zoneset_name
```

2. Add zones into a zoneset.

   ZONESET CONFIGURATION mode

   ```
   member zonename
   ```

Example of Creating Zonesets

```
Dell(conf)#fc zoneset zs1
Dell(conf-fc-zoneset-zs1)#member z1
Dell(conf-fc-zoneset-zs1)#
Dell(conf-fc-zoneset-zs1)#exit
Dell(conf-fc-zoneset-zs1)#
```

Activating a Zoneset

Activating a zoneset makes the zones within it effective.

On a switch, only one zoneset can be active. Any changes in an activated zoneset do not take effect until it is re-activated.

By default, the fcoe-map `fabric map-name`does not have any active zonesets.

1. Enter enter the `fc-fabric` command in fcoe-map to active or de-activate the zoneset.

   ```
   Dell(conf-fcoe-map)#fc-fabric
   ```

   Example:

   ```
   Dell(conf)#fcoe-map map
   Dell(conf-fcoe-map)#fc-fabric
   Dell(conf-fmap-map-fcfabric)#active-zoneset set
   Dell(conf-fmap-map-fcfabric)#no active-zoneset?
   active-zoneset
   Dell(conf-fmap-map-fcfabric)#no active-zoneset ?
   <cr>
   Dell(conf-fmap-map-fcfabric)#no active-zoneset
   ```

2. View the active zoneset.

   ```
   show fc zoneset active
   ```

Displaying the Fabric Parameters

To display information on switch-wide and interface-specific fabric parameters, use the show commands in the following table.

Examples of these `show` commands follow this table.

Command	Description
show config	Displays the fabric parameters.
show fcoe-map	Displays the fcoe-map.
show fc ns switch	Display all the devices in name server database of the switch.
show fc ns switch brief	Display all the devices in name server database of the switch - brief version.

Command	Description
show fc zoneset	Displays the zoneset.
show fc zoneset active	Displays the active zoneset.
show fc zone	Displays the configured zone.
show fc alias	Displays the configured alias.
show fc switch	Displays the FC Switch mode and world wide name.

Example of the `show config` Command

```
Dell(conf-fcoe-SAN_FABRIC)#show config
!
fcoe-map SAN_FABRIC
 description SAN_FABRIC
 fc-map 0efc00
 fabric-id 1002 vlan 1002
!
 fc-fabric
  default-zone-allow all
Dell(conf-fcoe-SAN_FABRIC)#
```

Example of the `show fcoe-map` Command

```
Dell(conf)#do show  fcoe-map

Fabric Name         map

Fabric Type         Fport
Fabric Id           1002
Vlan Id             1002
Vlan priority       3
FC-MAP              0efc00
FKA-ADV-Period      8
Fcf Priority        128
Config-State        ACTIVE
Oper-State          UP
=======================================================
Switch Config Parameters
=======================================================
DomainID            2
=======================================================
Switch Zoning Parameters
=======================================================
Default Zone Mode:      Deny
Active Zoneset:         set
=======================================================
Members
Fc 0/9
Te 0/2
=======================================================
=======================================================
```

Example of the `show fc ns switch` Command

```
Dell(conf)#do show fc ns sw

Total number of devices =   1

Switch Name                     28:4e:55:4c:4c:29:00:00
Domain Id                       2
```

```
Switch Port                     4
FC-Id                           02:04:03
Port Name                       20:01:d4:ae:52:44:37:b2
Node Name                       20:00:d4:ae:52:44:37:b2
Class of Service                8
Symbolic Port Name              Broadcom Port0 pWWN 20:01:d4:ae:52:44:37:b2
Symbolic Node Name              Broadcom BCM57810 FCoE 7.6.3.0 7.6.59.0 WIN-
KBFI7FJ2FUH
Port Type                       N_Port
```

Example of the `show fc ns switch brief` Command

```
Dell#show fc ns switch brief
Total number of devices = 1
Intf#  Domain FC-ID    Enode-WWPN              Enode-WWNN
Fc 0/9 1      01:35:00 10:00:8c:7c:ff:17:f8:01 20:00:8c:7c:ff:17:f8:01
Dell#
```

Example of the `show fc zoneset` Command

```
Dell#show fc zoneset
ZoneSetName ZoneName ZoneMember
=======================================
fcoe_srv_fc_tgt
           brcd_sanb
                    brcd_cna1_wwpn1
                    sanb_p2tgt1_wwpn

Active Zoneset: fcoe_srv_fc_tgt
ZoneName ZoneMember
=======================================
brcd_sanb
         10:00:8c:7c:ff:21:5f:8d
         20:02:00:11:0d:03:00:00
Dell#
```

Example of the `show fc zoneset active` Command

```
Dell#show fc zoneset active
Active Zoneset: fcoe_srv_fc_tgt
ZoneName  ZoneMember
================================
brcd_sanb
          10:00:8c:7c:ff:21:5f:8d
          20:02:00:11:0d:03:00:00
Dell#
```

Example of the `show fc zone` Command

```
Dell#show fc zone
ZoneName  ZoneMember
============================
brcd_sanb
          brcd_cna1_wwpn1
          sanb_p2tgt1_wwpn
Dell#
```

Example of the `show fc alias` Command

```
Dell(conf)#do show fc alias

ZoneAliasName               ZoneMember
=========================================================
test
                                20:02:d4:ae:52:44:38:4f
                                20:34:78:2b:cb:6f:65:57
```

Example of the `show fc switch` Command

```
Dell(conf)#do show  fc switch
Switch Mode : FPORT
Switch WWN  : 10:00:aa:00:00:00:00:ac
Dell(conf)#
```

17

FCoE Transit

Dell Networking OS supports the Fibre Channel over Ethernet (FCoE) Transit feature. When you enable the switch for FCoE transit, the switch functions as a FIP snooping bridge.

NOTE: FCoE transit is not supported on Fibre Channel interfaces.

Supported Modes

Full–Switch

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. 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 14. 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.

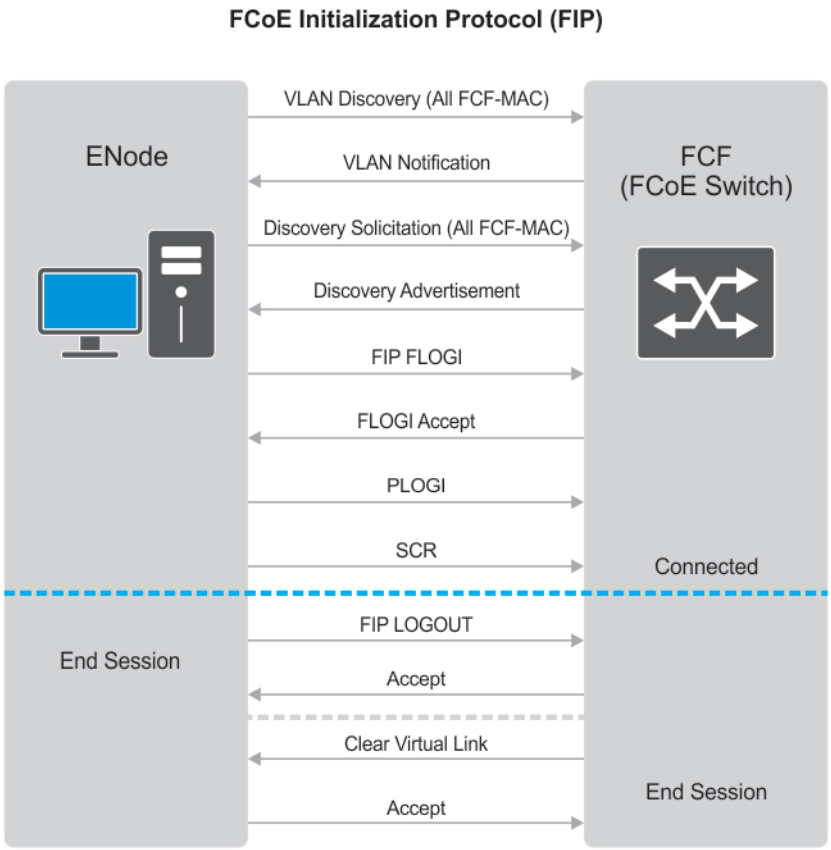


Figure 34. FIP Discovery and Login Between an ENode and an FCF

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 and configure the FIP snooping parameters. 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:

Global ACLs	These are applied on server-facing ENode ports.
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 an FN IOM 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 FN IOM switch. The FN IOM switch operates as a lossless FIP snooping bridge to transparently forward FCoE frames between the ENode servers and the FCF switch.

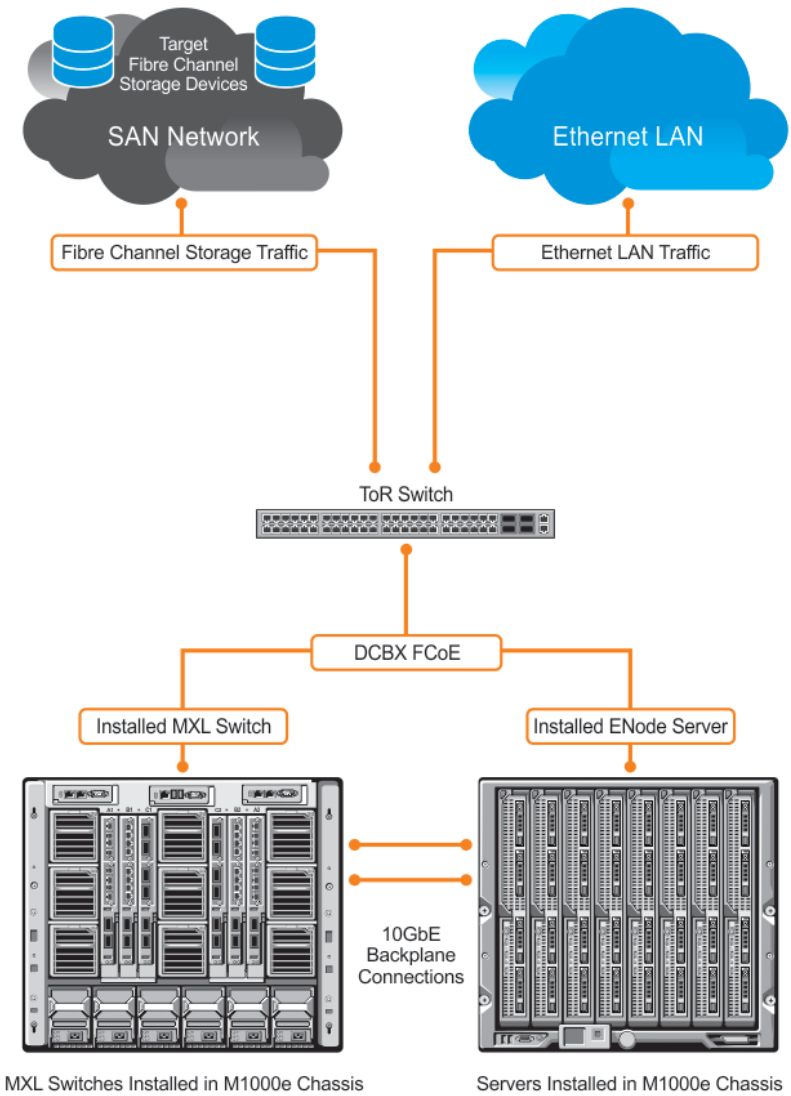


Figure 35. FIP Snooping on an FN IOM Switch

The following sections describe how to configure the FIP snooping feature on a switch that functions as a FIP snooping bridge so that it can perform the following functions:

- 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.
- 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.

NOTE: As a best practice, Dell Networking recommends not configuring FIP Snooping on a stacked switch.

Using FIP Snooping

There are four steps to configure FCoE transit.

1. Enable the FCoE transit feature on a switch to maintain FIP snooping information on the 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 Configuring FIP Snooping.

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.

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 Configuring FIP Snooping procedure.

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 (FIP snooping bridge-bridge) 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 15. 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 deleted. If a port is enabled for FIP snooping in ENode or FCF mode, the ENode/FCF MAC-based ACLs are deleted.

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 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.

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 12.
- Links to other FIP snooping bridges on a FIP snooping-enabled port (bridge-to-bridge links) are not supported on the switch.

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. Enable the FCoE transit feature on a switch.
   CONFIGURATION mode.

   ```
   feature fip-snooping
   ```
2. Enable FIP snooping on all VLANs or on a specified VLAN.
   CONFIGURATION mode or VLAN INTERFACE mode.

   ```
   fip-snooping enable
   ```

   By default, FIP snooping is disabled on all VLANs.
3. Configure the FC-MAP value used by FIP snooping on all VLANs.
   CONFIGURATION VLAN or INTERFACE mode

   ```
   fip-snooping fc-map fc-map-value
   ```

   The default is 0x0EFC00.

   The valid values are from 0EFC00 to 0EFCFF.
4. Enter interface configuration mode to configure the port for FIP snooping links.
   CONFIGURATION mode

   ```
   interface port-type slot/port
   ```

   By default, a port is configured for bridge-to-ENode links.

5. Configure the port for bridge-to-FCF links.
   INTERFACE or CONFIGURATION mode

   ```
   fip-snooping port-mode fcf
   ```

NOTE: To disable the FIP snooping 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 16. 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.

Command	Output
`show fip-snooping vlan`	Displays information on the FCoE VLANs on which FIP snooping is enabled.

Example of the `show fip-snooping sessions` Command

```
Dell#show fip-snooping sessions
Enode MAC          Enode Intf  FCF MAC            FCF Intf  VLAN
aa:bb:cc:00:00:00  Te 0/2      aa:bb:cd:00:00:00  Te 0/11   100
aa:bb:cc:00:00:00  Te 0/2      aa:bb:cd:00:00:00  Te 0/11   100
aa:bb:cc:00:00:00  Te 0/2      aa:bb:cd:00:00:00  Te 0/11   100
aa:bb:cc:00:00:00  Te 0/2      aa:bb:cd:00:00:00  Te 0/11   100
aa:bb:cc:00:00:00  Te 0/2      aa:bb:cd:00:00:00  Te 0/11   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 17. `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.

Example of Viewing FIP Snooping Configuration

```
Dell# 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
```

Example of the `show fip-snooping enode` Command

```
Dell# show fip-snooping enode
Enode MAC          Enode Interface FCF MAC          VLAN  FC-ID
---------          --------------- -------          ----  -----
d4:ae:52:1b:e3:cd  Te 0/11         54:7f:ee:37:34:40 100   62:00:11
```

The following table describes the `show fip-snooping enode` command fields.

Table 18. `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.

Example of the `show fip-snooping fcf` Command

```
Dell# 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 19. `show fip-snooping fcf` Command Description

Field	Description
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.
FC-MAP	FC-Map value advertised by the FCF.
ENode Interface	Slot/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.

Example of the `show fip-snooping statistics interface vlan` (VLAN and Port) Command

```
Dell# 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
Dell(conf)#

Dell# show fip-snooping statistics int tengigabitethernet 0/11
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
```

Example of the `show fip-snooping statistics port-channel` Command

```
Dell# 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 20. `show fip-snooping statistics` Command Description

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.
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.

Field	Description
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.

Example of the `show fip-snooping system` Command

```
Dell# show fip-snooping system
Global Mode                   : Enabled
FCOE VLAN List (Operational) : 1, 100
FCFs                          : 1
Enodes                        : 2
Sessions                      : 17
```

NOTE: NPIV sessions are included in the number of FIP-snooped sessions displayed.

Example of the `show fip-snooping vlan` Command

```
Dell# 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 an FN IOM 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.

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:

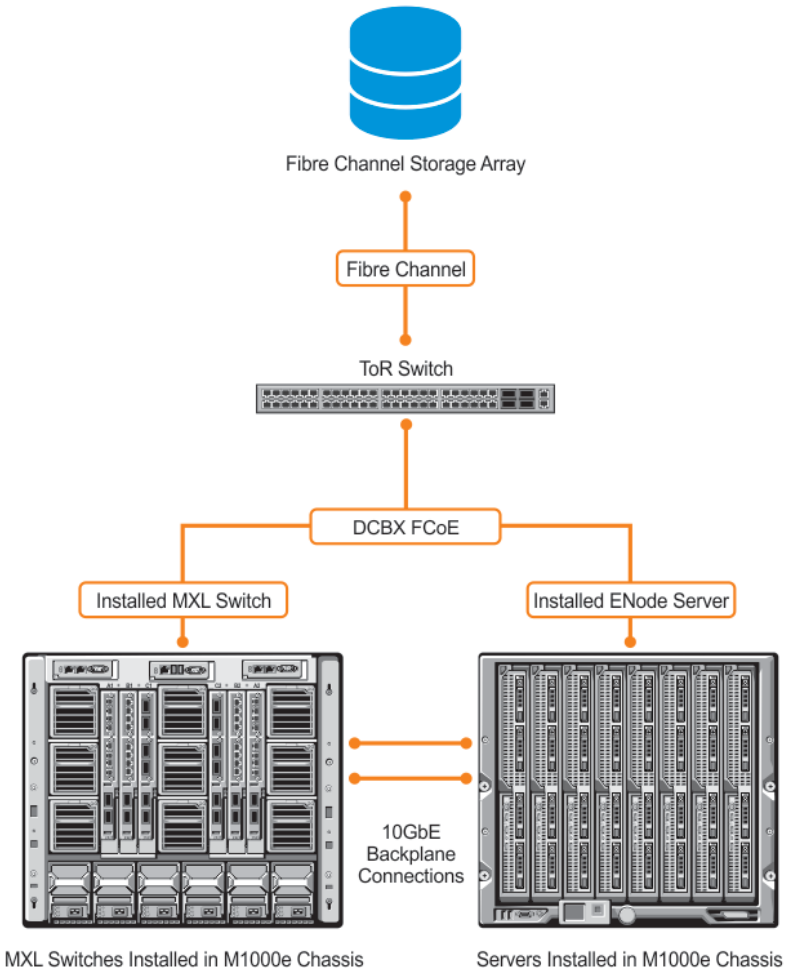


Figure 36. FIP Snooping on an FN IOMSwitch Configuration Example

- 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 (0/50), on an ENode server-facing port (0/1), 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)

```
Dell(conf)# feature fip-snooping
```

Example of Enabling FIP Snooping on the FCoE VLAN

```
Dell(conf)# interface vlan 10
Dell(conf-if-vl-10)# fip-snooping enable
```

Example of Enabling an FC-MAP Value on a VLAN

```
Dell(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

```
Dell(conf)# interface tengigabitethernet 0/1
Dell(conf-if-te-0/1)# portmode hybrid
Dell(conf-if-te-0/1)# switchport
```

NOTE: A port is enabled by default for bridge-ENode links.

Example of Configuring the FCF-Facing Port

```
Dell(conf)# interface tengigabitethernet 0/11
Dell(conf-if-te-0/11)# portmode hybrid
Dell(conf-if-te-0/11)# switchport
Dell(conf-if-te-0/11)# fip-snooping port-mode fcf
```

Example of Configuring FIP Snooping Ports as Tagged Members of the FCoE VLAN

```
Dell(conf)# interface vlan 10
Dell(conf-if-vl-10)# tagged tengigabitethernet 0/11
Dell(conf-if-te-0/11)# no shut
Dell(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.

18

FIPS Cryptography

Dell Networking OS supports federal information processing standard (FIPS) cryptography.

This chapter describes how to enable FIPS cryptography requirements on Dell Networking platforms. This feature 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.

NOTE: For the Dell Networking OS version 8.3.12.0, only the SSH and SCP copy features use FIPS Cryptographic mode to secure management interface user sessions and file transfers. Other features that use cryptographic algorithms do not, or cannot, use FIPS mode. You must configure the management interfaces to limit access to/from the system to SSH alone.

Preparing the System

Before you enable FIPS mode, Dell 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 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`

Example of the `show fips status` Command

Example of the `show system` Command

```
Dell#show fips status
FIPS Mode : Enabled
for the system using the show system command.

Dell#show system

Stack MAC : 00:01:e8:8a:ff:0c

Reload Type : normal-reload [Next boot : normal-reload]
```

```
-- Unit 0 --
Unit Type       : Management Unit
Status          : online
Next Boot       : online
Required Type   : XML - 52-port GE/TE/FG (SE)
Current Type    : XML - 52-port GE/TE/FG (SE)
Master priority : 0
Hardware Rev    : 3.0
Num Ports       : 64
Up Time         : 7 hr, 3 min
Dell Version    : XML-8-3-7-1061
Jumbo Capable   : yes
POE Capable     : no
FIPS Mode       : enabled
Burned In MAC   : 00:01:e8:8a:ff:0c
No Of MACs      : 3
...
```

Disabling FIPS Mode

The following describes 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) ?
  ```

19

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) 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.

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.

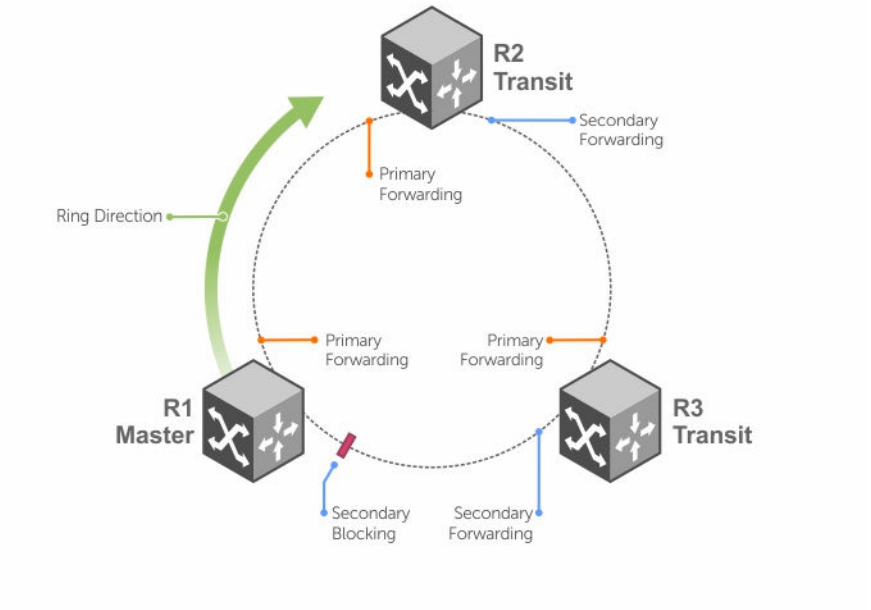


Figure 37. Normal Operating FRRP Topology

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.

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.

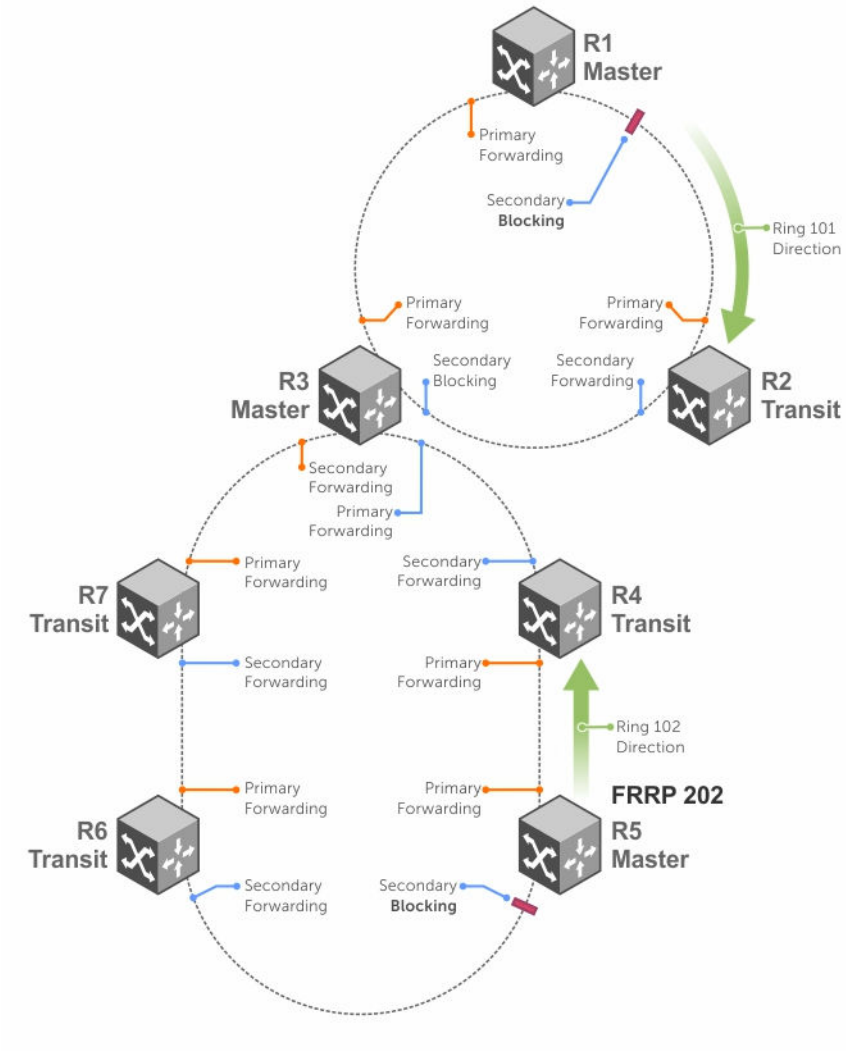


Figure 38. Multiple Rings Connected by a Single Switch Example

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.
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.

Concept	Explanation
	• **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
- Setting the FRRP Timers

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 slot/ port {range}
   ```

   - For a 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

   Slot/Port, Range: Slot and Port ID for the interface. Range is entered Slot/Port-Port.

3. Assign the Primary and Secondary ports and the control VLAN for the ports on the ring.
   CONFIG-FRRP mode.

   ```
   interface primary int slot/port secondary int slot/port control-vlan vlan id
   ```

   - For a 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a SONET interface, enter the keyword `sonet` then the slot/port information.
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

   Slot/Port, Range: Slot and Port ID for the interface. Range is entered Slot/Port-Port.

   VLAN ID: The VLAN identification of the control VLAN.
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 slot/port {range}
   ```

   - For a 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a SONET interface, enter the keyword `sonet` then the slot/port information.
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - *`Slot/Port, range`*: Slot and Port ID for the interface. The range is entered Slot/Port-Port.

3. Assign the Primary and Secondary ports and the Control VLAN for the ports on the ring.

   CONFIG-FRRP mode.

   ```
   interface primary int slot/port secondary int slot/port control-vlan vlan id
   ```

   - For a 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information.
   - For a SONET interface, enter the keyword `sonet` then the slot/port information.
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

   `Slot/Port, Range`: Slot and Port ID for the interface. Range is entered Slot/Port-Port.

   `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.

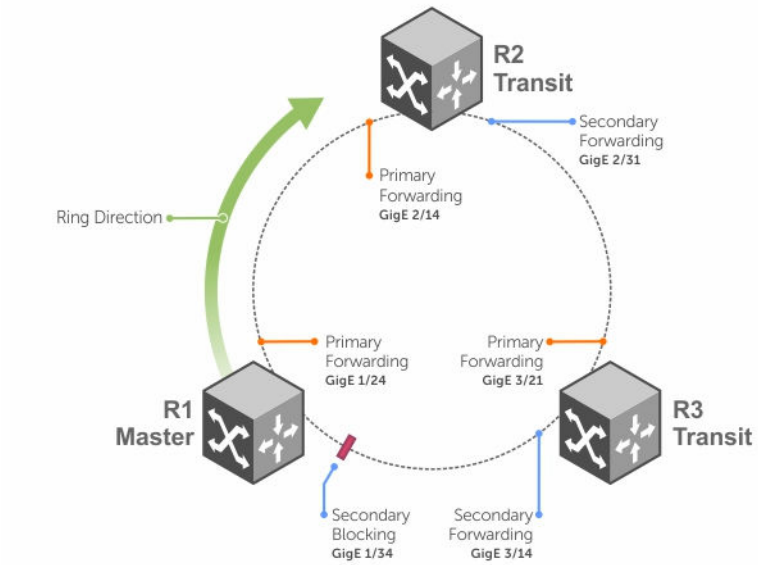


Figure 39. Basic Topology and CLI Commands

Example of R1 MASTER

Example of R2 TRANSIT

Example of R3 TRANSIT

```
interface GigabitEthernet 1/24
  no ip address
  switchport
```

```
  no shutdown
!
interface GigabitEthernet 1/34
  no ip address
  switchport
  no shutdown
!
interface Vlan 101
  no ip address
  tagged GigabitEthernet 1/24,34
  no shutdown
!
interface Vlan 201
  no ip address
  tagged GigabitEthernet 1/24,34
  no shutdown

!
protocol frrp 101
  interface primary GigabitEthernet 1/24
secondary GigabitEthernet 1/34 control-vlan 101
  member-vlan 201
  mode master
  no disable

interface GigabitEthernet 2/14
  no ip address
  switchport
  no shutdown
!
interface GigabitEthernet 2/31
  no ip address
  switchport
  no shutdown
!
interface Vlan 101
  no ip address
  tagged GigabitEthernet 2/14,31
  no shutdown
!
interface Vlan 201
  no ip address
  tagged GigabitEthernet 2/14,31
  no shutdown
!
protocol frrp 101
  interface primary GigabitEthernet 2/14 secondary GigabitEthernet 2/31 control-
vlan 101
  member-vlan 201
  mode transit
  no disable

interface GigabitEthernet 3/14
  no ip address
  switchport
  no shutdown
!
interface GigabitEthernet 3/21
  no ip address
  switchport
  no shutdown
!
interface Vlan 101
  no ip address
```

```
  tagged GigabitEthernet 3/14,21
  no shutdown
!
interface Vlan 201
  no ip address
  tagged GigabitEthernet 3/14,21
  no shutdown

!
protocol frrp 101
  interface primary GigabitEthernet 3/21
secondary GigabitEthernet 3/14 control-vlan 101
  member-vlan 201
  mode transit
  no disable
```

20
GARP VLAN Registration Protocol (GVRP)

Dell Networking OS supports GARP VLAN registration protocol (GVRP).

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.
- On the Switch, you cannot enable per-VLAN spanning tree+ (PVST+) and GVRP at the same time. If spanning tree and GVRP are both required, implement either rapid spanning tree protocol (RSTP), spanning tree protocol (STP), or multiple spanning tree protocol (MSTP). Dell Networking OS does support enabling GVRP and MSTP at the same time.

```
Dell(conf)#protocol spanning-tree pvst
Dell(conf-pvst)#no disable
% Error: GVRP running. Cannot enable PVST.

.........
Dell(conf)#protocol spanning-tree mstp
Dell(conf-mstp)#no disable
% Error: GVRP running. Cannot enable MSTP.
```

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, that type of port is referred to as a VLAN trunk port, but it is not necessary to specifically identify to the Dell Networking operating system (OS) that the port is a trunk port.

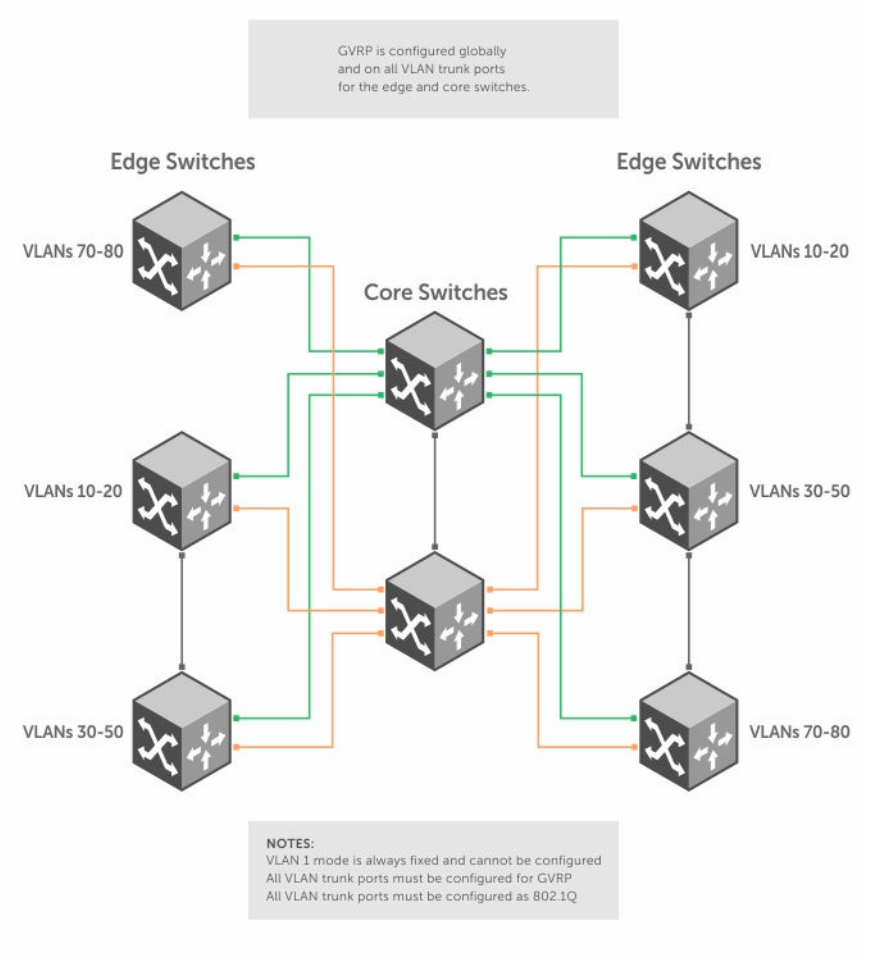


Figure 40. Global GVRP Configuration Example

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
```

Example of Configuring GVRP

```
Dell(conf)#protocol gvrp
Dell(config-gvrp)#no disable
Dell(config-gvrp)#show config
!
protocol gvrp
  no disable
Dell(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
  ```

Example of Enabling GVRP on an Interface

```
Dell(conf-if-gi-1/21)#switchport
Dell(conf-if-gi-1/21)#gvrp enable
Dell(conf-if-gi-1/21)#no shutdown
Dell(conf-if-gi-1/21)#show config
!
interface GigabitEthernet 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 three GVRP registration modes:

- **Normal Registration** — Allows dynamic creation, registration, and de-registration of VLANs (if you enabled dynamic VLAN creation). By default, the registration mode is set to Normal when you enable GVRP on a port. This default mode enables the port to dynamically register and de-register VLANs, and to propagate both dynamic and static VLAN information.
- **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 1/21 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

```
Dell(conf-if-gi-1/21)#gvrp registration fixed 34,35
Dell(conf-if-gi-1/21)#gvrp registration forbidden 45,46
Dell(conf-if-gi-1/21)#show conf
!
interface GigabitEthernet 1/21
  no ip address
  switchport
  gvrp enable
  gvrp registration fixed 34-35
  gvrp registration forbidden 45-46
  no shutdown
Dell(conf-if-gi-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 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 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 default is **10000ms**.

Example of the `garp timer` Command

```
Dell(conf)#garp timer leav 1000
Dell(conf)#garp timers leave-all 5000
Dell(conf)#garp timer join 300

Verification:

Dell(conf)#do show garp timer
GARP Timers Value (milliseconds)
----------------------------------------
Join Timer      300
Leave Timer     1000
LeaveAll Timer  5000
Dell(conf)#
```

Dell displays this message if an attempt is made to configure an invalid GARP timer: `Dell(conf)#garp timers join 300 % Error: Leave timer should be >= 3*Join timer.`

21

FIP Snooping

This chapter describes about the FIP snooping concepts and configuration procedures.

Supported Modes

Standalone, PMUX, VLT

Fibre Channel over Ethernet

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 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. For more information, refer to the Data Center Bridging (DCB) chapter.

Ensuring Robustness in a Converged Ethernet Network

Fibre Channel networks used for SAN traffic employ switches that operate as trusted devices. End devices log into the switch to which they are attached in order to communicate with the other end devices attached to the Fibre Channel network. 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 attacks.

To ensure similar Fibre Channel robustness and security with FCoE in an Ethernet cloud network, the Fibre Channel over Ethernet initialization protocol (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 access control list (ACLs) that can emulate a point-to-point link by providing the traffic enforcement required to create a Fibre Channel-level of robustness. 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, and use the FIP snooping data to dynamically configure ACLs on the bridge to only permit traffic authorized by the FCF.

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. 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 a functionality for discovering and logging in to 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 below illustration about FIP discovery, depicts the communication that occurs between an ENode server and an FCoE switch (FCF).

FIP performs the following functions:

- FIP virtual local area network (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 perform fabric login (FLOGI) and fabric discovery (FDISC) to 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 link termination logout (LOGO) functions properly.

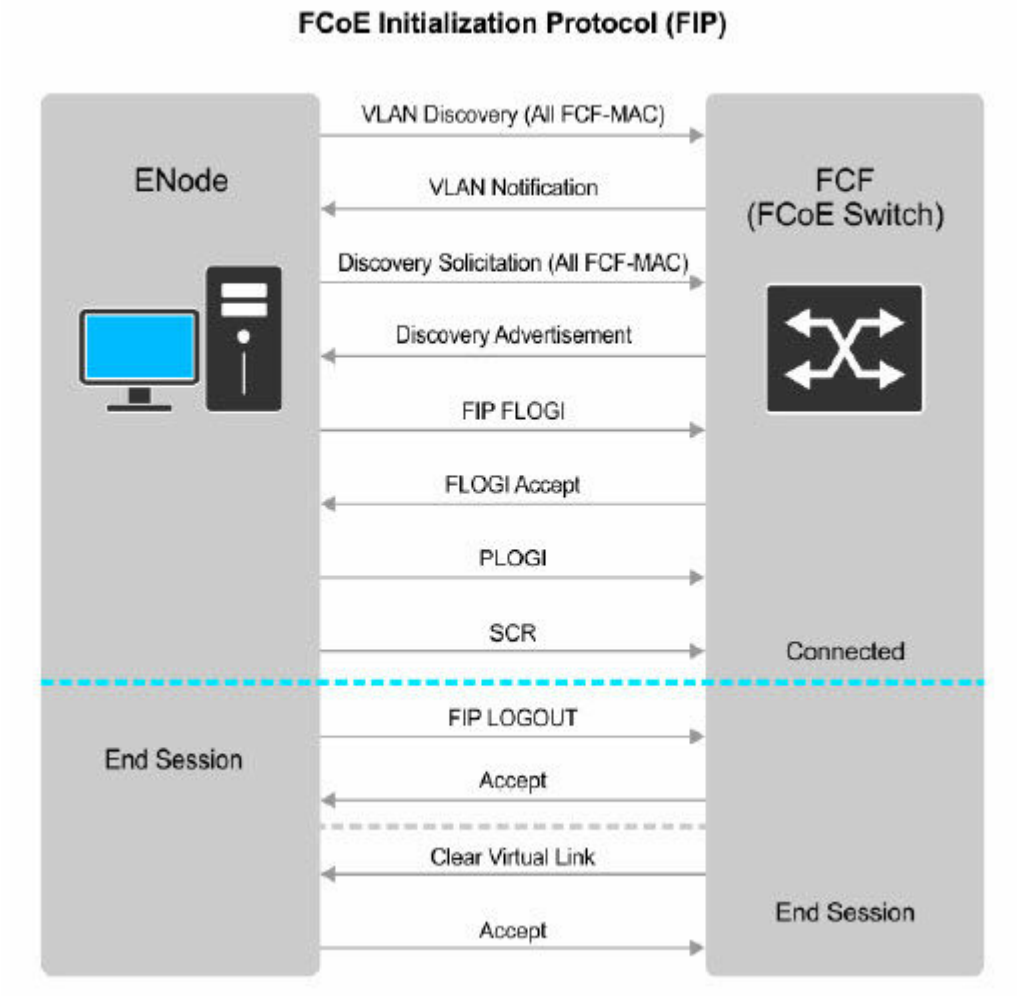


Figure 41. FIP Discovery and Login between an ENode and an FCF

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 the following port modes:

- ENode mode for server-facing ports
- FCF mode for a trusted port directly connected to an FCF

You must enable FIP snooping on an Aggregator and configure the FIP snooping parameters. When you enable FIP snooping, all ports on the switch by default become ENode ports.

Dynamic ACL generation on an Aggregator operating as a FIP snooping bridge functions as follows:

- Global ACLs are applied on server-facing ENode ports.
- Port-based ACLs are applied on ports directly connected to an FCF and on server-facing ENode ports.
- Port-based ACLs take precedence over global ACLs.
- FCoE-generated ACLs take precedence over user-configured ACLs. A user-configured ACL entry cannot deny FCoE and FIP snooping frames.

The below illustration depicts an Aggregator used as a FIP snooping bridge in a converged Ethernet network. The ToR switch operates as an FCF for FCoE traffic. Converged LAN and SAN traffic is transmitted between the ToR switch and an Aggregator. The Aggregator operates as a lossless FIP snooping bridge to transparently forward FCoE frames between the ENode servers and the FCF switch.

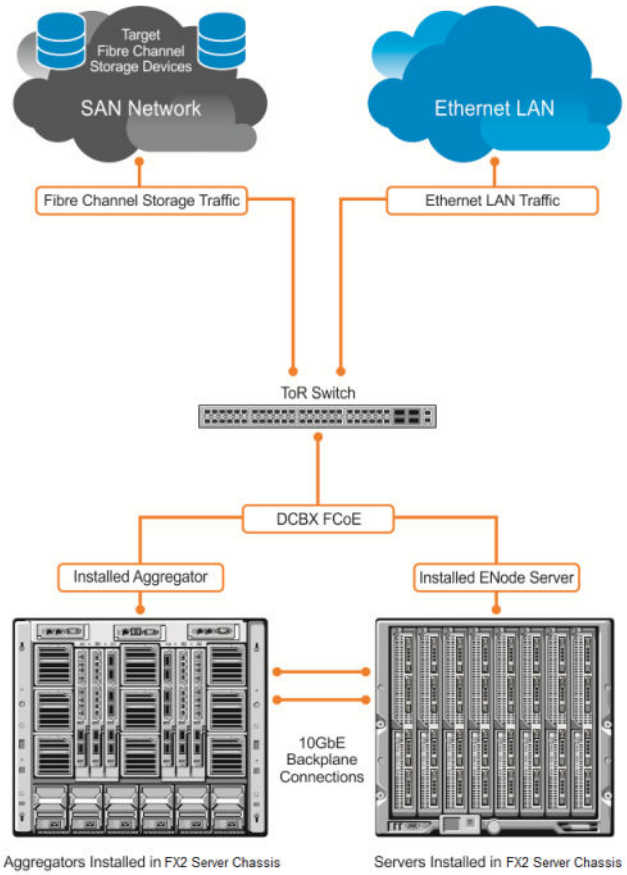


Figure 42. FIP Snooping on an Aggregator

The following sections describes how to configure the FIP snooping feature on a switch that functions as a FIP snooping bridge so that it can perform the following functions:

- Performs FIP snooping (allowing and parsing FIP frames) globally on all VLANs or on a per-VLAN basis.
- Set the FCoE MAC address prefix (FC-MAP) value used by an FCF to assign a MAC address to an ECoE end-device (server ENode or storage device) after a server successfully logs in.
- Set the FCF mode to provide additional port security on ports that are directly connected to an FCF.
- Check FIP snooping-enabled VLANs to ensure that they are operationally active.
- 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.

How FIP Snooping is Implemented

As soon as the Aggregator is activated in an Dell PowerEdge FX2 server chassis as a switch-bridge, existing VLAN—specific and FIP snooping auto-configurations are applied. The Aggregator snoops FIP packets on VLANs enabled for FIP snooping and allows legitimate sessions. By default, all FCoE and FIP frames are dropped unless specifically permitted by existing FIP snooping-generated ACLs.

FIP Snooping on VLANs

FIP snooping is enabled globally on an Aggregator on all 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.
- Atleast one interface is auto-configured for FCF (FIP snooping bridge — FCF) mode on a FIP snooping-enabled VLAN. Multiple FCF trusted interfaces are auto-configured in a VLAN.
- A maximum of eight VLANs are supported for FIP snooping on an Aggregator. FIP snooping processes FIP packets in traffic only from the first eight incoming VLANs.

FC-MAP Value

The FC-MAP value that is applied globally by the Aggregator on all FCoE VLANs to authorize FCoE traffic is auto-configured.

The 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 values 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.

Bridge-to-FCF Links

A port directly connected to an FCF is auto-configured in FCF mode. Initially, all FCoE traffic is blocked; only FIP frames are allowed to pass.

FCoE traffic is allowed on the port only after a successful FLOGI request/response and confirmed use of the configured FC-MAP value for the VLAN.

Impact on other Software Features

FIP snooping affects other software features on an Aggregator as follows:

- MAC address learning: MAC address learning is not performed on FIP and FCoE frames, which are denied by ACLs dynamically created by FIP snooping in 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 the 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).

FIP Snooping Prerequisites

On an Aggregator, FIP snooping requires the following conditions:

- A FIP snooping bridge requires DCBX and PFC to be enabled on the switch for lossless Ethernet connections (refer to ***Data Center Bridging (DCB)***). Dell recommends that you also enable ETS; ETS is recommended but not required. DCBX and PFC mode are auto-configured on Aggregator ports and 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.
- VLAN membership:
  - The Aggregator auto-configures the VLANs which handle FCoE traffic. You can reconfigure VLAN membership on a port (`vlan tagged` command).

– Each FIP snooping port is auto-configured to operate in Hybrid mode so that it accepts both tagged and untagged VLAN frames.
– Tagged VLAN membership is auto-configured on each FIP snooping port that sends and receives FCoE traffic and has links with an FCF, ENode server or another FIP snooping bridge.
– The default VLAN membership of the port should continue to operate with untagged frames. FIP snooping is not supported on a port that is configured for non-default untagged VLAN membership.

FIP Snooping Restrictions

The following restrictions apply to FIP snooping on an Aggregator:

- The maximum number of FCoE VLANs supported on the Aggregator 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. This is configurable only in PMUX mode.
- In a full FCoE N port ID virtualization (NPIV) configuration, 16 sessions (one FLOGI + 15 NPIV sessions) are supported per ENode. In an FCoE NPV confguration, only one session is supported per ENode.
- The maximum number of FCFs supported per FIP snooping-enabled VLAN is 12.
- Links to other FIP snooping bridges on a FIP snooping-enabled port (bridge-to-bridge links) are not supported on the Aggregator.

Configuring FIP Snooping

FIP snooping is auto-configured on an Aggregator in standalone mode. You can display information on FIP snooping operation and statistics by entering `show` commands. 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. Enable the FCoE transit feature on a switch.
   CONFIGURATION mode.

   ```
   feature fip-snooping
   ```
2. Enable FIP snooping on all VLANs or on a specified VLAN.
   CONFIGURATION mode or VLAN INTERFACE mode.

   ```
   fip-snooping enable
   ```

   By default, FIP snooping is disabled on all VLANs.
3. Configure the FC-MAP value used by FIP snooping on all VLANs.
   CONFIGURATION VLAN or INTERFACE mode

   ```
   fip-snooping fc-map fc-map-value
   ```

   The default is 0x0EFC00.

   The valid values are from 0EFC00 to 0EFCFF.
4. Enter interface configuration mode to configure the port for FIP snooping links.
   CONFIGURATION mode

```
interface port-type slot/port
```

By default, a port is configured for bridge-to-ENode links.

5. Configure the port for bridge-to-FCF links.
   INTERFACE or CONFIGURATION mode

```
fip-snooping port-mode fcf
```

NOTE: All these configurations are available only in PMUX mode.

NOTE: To disable the FIP snooping 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 `show` commands from the table below, to display information on FIP snooping.

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). Information on NPIV sessions is also displayed.
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	Displays statistics on the FIP packets snooped on all interfaces, including VLANs, physical ports, and port channels.

Command	Output
***port-type port/slot* | interface port-channel *port-channel-number*]**	
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	Display 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	Display information on the FCoE VLANs on which FIP snooping is enabled.

`show fip-snooping sessions` **Command Example**

```
Dell#show fip-snooping sessions
Enode MAC         Enode Intf      FCF MAC    FCF Intf VLAN FCoE MAC        FC-ID
00:0e:1e:0c:54:a6 Te 0/1  00:05:73:f2:4f:ae Po128 100 0e:fc:00:9a:00:27 9a:
00:27 20:01:00:0e:1e:0c:54:a6
00:0e:1e:06:01:5e Te 0/3  00:05:73:f2:4f:af Po128 100 0e:fc:00:9a:01:18 9a:
01:18 20:01:00:0e:1e:06:01:5
Port WWNN
20:00:00:0e:1e:0c:54:a6
20:00:00:0e:1e:0c:54:a6
```

`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.

`show fip-snooping config` **Command Example**

```
Dell#show fip-snooping config
FIP Snooping Feature enabled Status: Enabled
FIP Snooping Global enabled Status: Enabled
Global FC-MAP Value: 0X0EFC00
Maximum Sessions Per Enode Mac: 32
Dell#
```

`show fip-snooping enode` **Command Example**

```
Dell# show fip-snooping enode
Enode MAC              Enode Interface          FCF MAC
VLAN             FC-ID
---------              ---------------          -------
----               ------
d4:ae:52:1b:e3:cd      Te 0/1                   54:7f:ee:37:34:40
100            62:00:11
```

`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.

`show fip-snooping fcf` **Command Example**

```
Dell# 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
```

`show fip-snooping fcf` **Command Description**

Field	Description
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.
FC-MAP	FC-Map value advertised by the FCF.
ENode Interface	Slot/ 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.

`show fip-snooping statistics` **(VLAN and port) Command Example**

```
Dell# 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
Dell(conf)#

Dell# show fip-snooping statistics int tengigabitethernet 0/1
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
```

`show fip-snooping statistics` **(port channel) Command Example**

```
Dell# 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
```

`show fip-snooping statistics` **Command Description**

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.
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.

Field	Description
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.

`show fip-snooping system` **Command Example**

```
Dell# show fip-snooping system
Global Mode                     : Enabled
FCOE VLAN List (Operational)    : 1, 100
FCFs                            : 1
Enodes                          : 2
Sessions                        : 17
```

NOTE: NPIV sessions are included in the number of FIP-snooped sessions displayed.

`show fip-snooping vlan` **Command Example**

```
Dell# show fip-snooping vlan
* = Default VLAN

VLAN    FC-MAP          FCFs    Enodes  Sessions
----    ------          ----    ------  --------
*1      -               -       -       -
100     0X0EFC00        1       2       17
```

NOTE: NPIV sessions are included in the number of FIP-snooped sessions displayed.

FIP Snooping Example

The following figure shows an Aggregator 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.

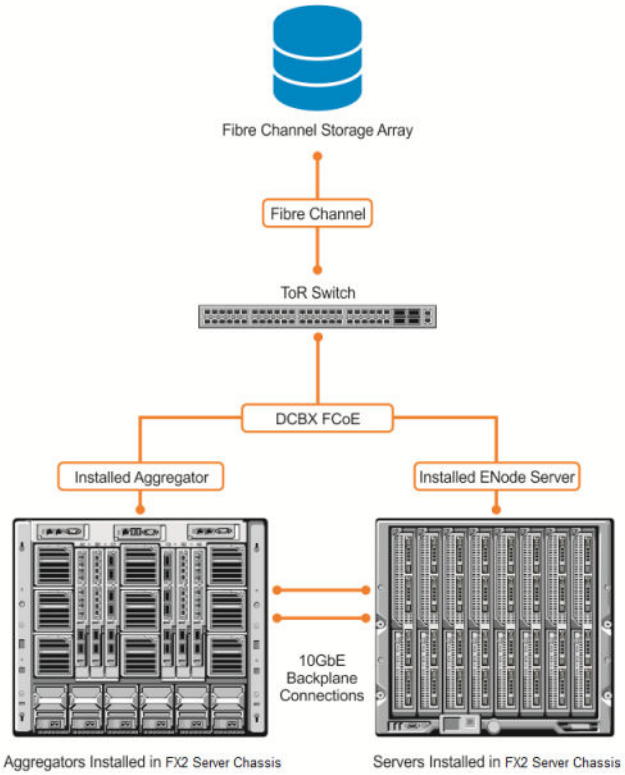


Figure 43. FIP Snooping on an Aggregator

In tbe above figure, DCBX and PFC are enabled on the Aggregator (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 FIP Snooping

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.

Debugging FIP Snooping

To enable debug messages for FIP snooping events, enter the `debug fip-snooping` command..

1. Enable FIP snooping debugging on for all or a specified event type, where:

- `all` enables all debugging options.
- `acl` enables debugging only for ACL-specific events.
- `error` enables debugging only for error conditions.
- `ifm` enables debugging only for IFM events.
- `info` enables debugging only for information events.
- `ipc` enables debugging only for IPC events.
- `rx` enables debugging only for incoming packet traffic.

EXEC PRIVILEGE mode

```
debug fip-snooping [all | acl | error | ifm | info | ipc | rx]
```

To turn off debugging event messages, enter the `no debug fip-snooping` command.

22

Internet Group Management Protocol (IGMP)

On an Aggregator, IGMP snooping is auto-configured. You can display information on IGMP by using `show ip igmp` command.

Multicast is based on identifying many hosts by a single destination IP address. Hosts represented by the same IP address are a multicast group. The internet group management protocol (IGMP) is a Layer 3 multicast protocol that hosts use to join or leave 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.

This chapter contains the following sections:

- IGMP Overview
- IGMP Snooping

IGMP 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 upon 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 which is as illustrated below:

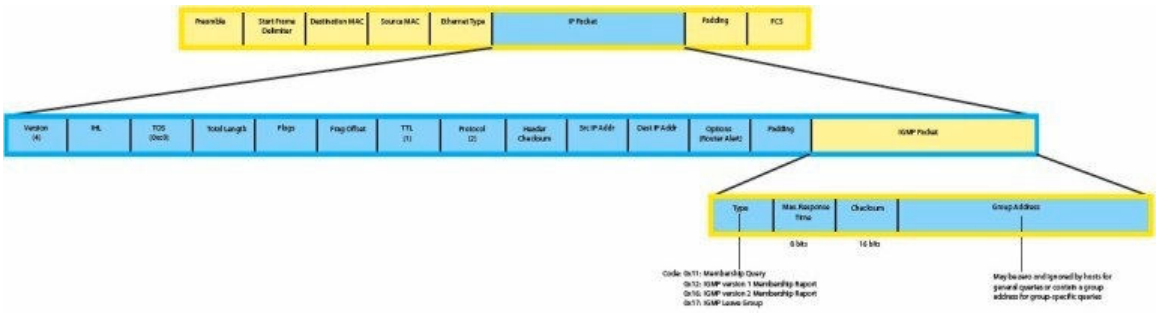


Figure 44. IGMP Version 2 Packet Format

Joining 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.
  - 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.
  - 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 IGMP Snooping).
  - 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

- 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.
- 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.
- Any remaining hosts respond to the query according to the delay timer mechanism (refer to IGMP Snooping). 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 the 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 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.

To accommodate these protocol enhancements, the IGMP version 3 packet structure is different from version 2. Queries (shown below in query packet format) are still sent to the all-systems address 224.0.0.1, but reports (shown below in report packet format) are sent to all the IGMP version 3 — capable multicast routers address 224.0.0.22.

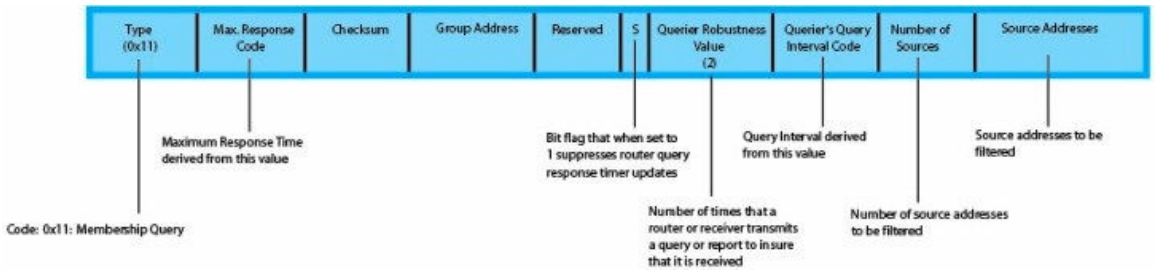


Figure 45. IGMP version 3 Membership Query Packet Format

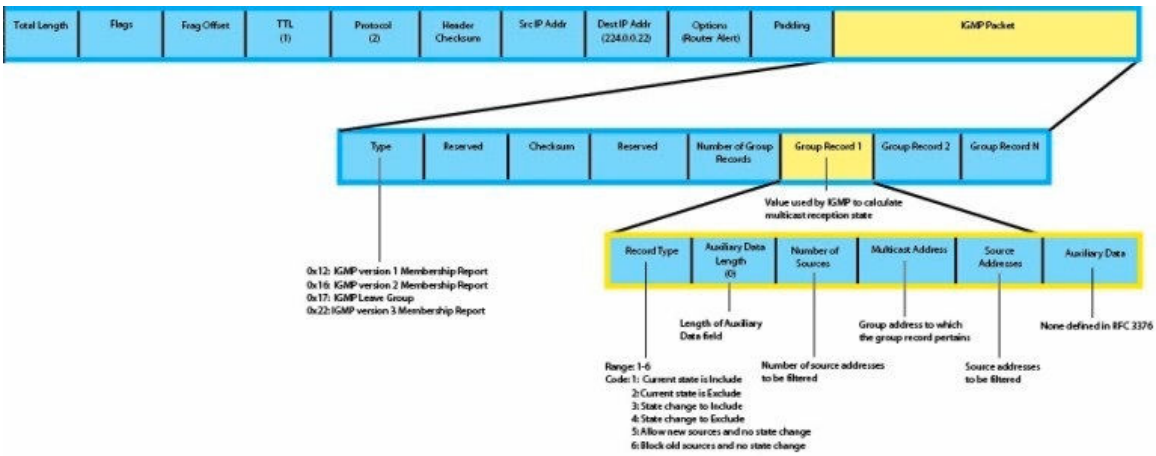


Figure 46. IGMP version 3 Membership Report Packet Format

Joining and Filtering Groups and Sources

The below illustration shows how multicast routers maintain the group and source information from unsolicited reports.

- The first unsolicited report from the host indicates that it wants to receive traffic for group 224.1.1.1.
- 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 prevent traffic from all other sources in the group from reaching the subnet, so 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.

- 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.

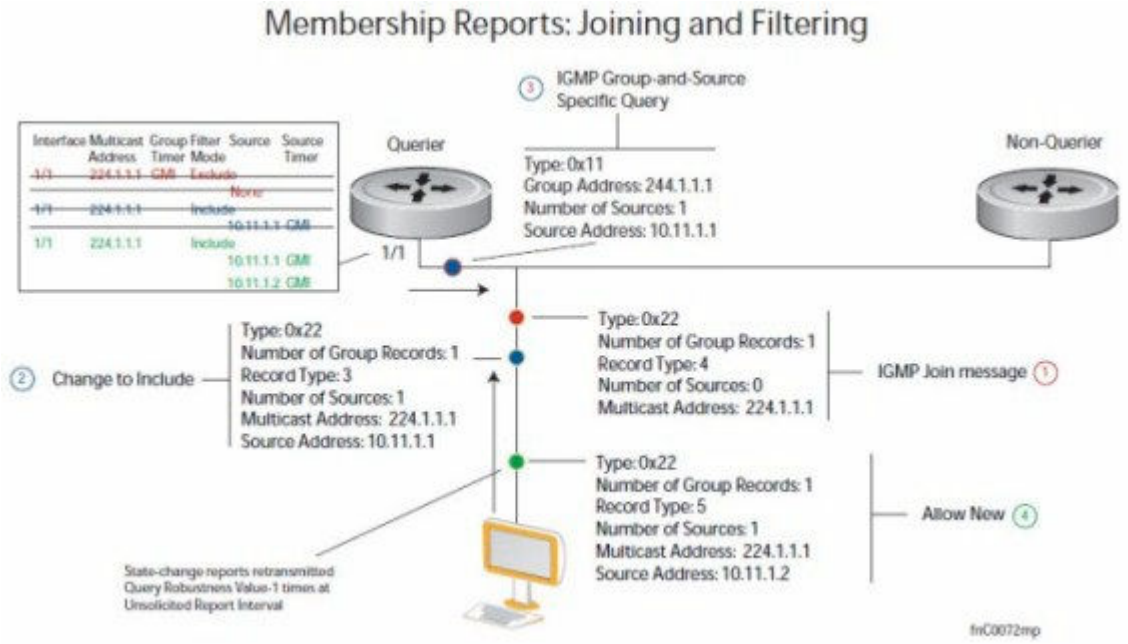


Figure 47. IGMP Membership Reports: Joining and Filtering

Leaving and Staying in Groups

The below illustration shows how multicast routers track and refreshes the state change in response to group-and-specific and general queries.

- 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.
- 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 interested, they respond with their current state information and the querier refreshes the relevant state information.
- Separately in the below figure, the querier sends a general query to 224.0.0.1.
- Host 2 responds to the periodic general query so the querier refreshes the state information for that group.

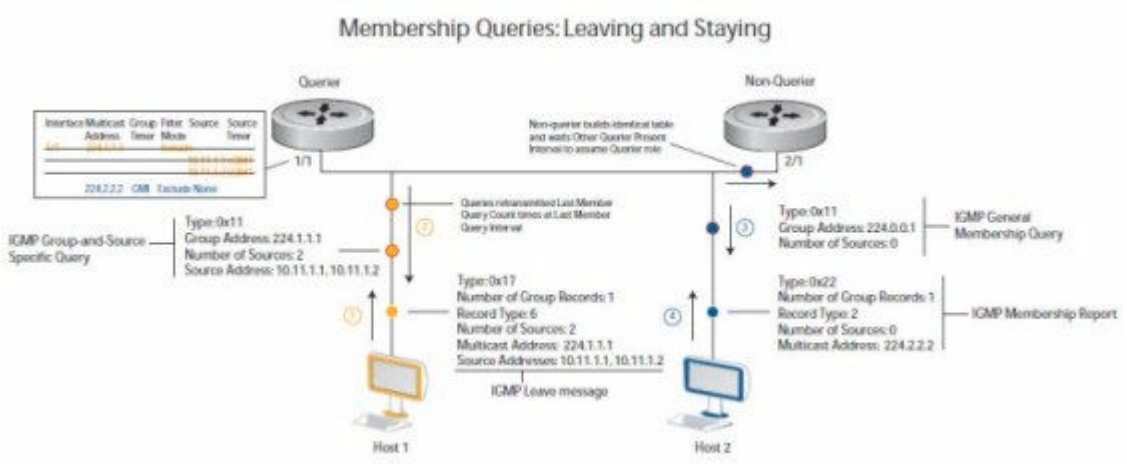


Figure 48. IGMP Membership Queries: Leaving and Staying in Groups

IGMP Snooping

IGMP snooping is auto-configured on an Aggregator.

Multicast packets are addressed with multicast MAC addresses, which represents a group of devices rather than one unique device. Switches forward multicast frames out of all ports in a VLAN by default, even if there are only a small number of interested hosts, resulting in a waste of bandwidth. IGMP snooping enables switches to use information in IGMP packets to generate a forwarding table that associate ports with multicast groups, so that the received multicast frames are forwarded only to interested receivers.

How IGMP Snooping is Implemented on an Aggregator

- IGMP snooping is enabled by default on the switch.
- Dell Networking OS supports version 1, version 2 and version 3 hosts.
- Dell Networking OS — IGMP snooping is based on the IP multicast address (not on the Layer 2 multicast MAC address). IGMP snooping entries are stored in the Layer 3 flow table instead of in the Layer 2 forwarding information base (FIB).
- Dell Networking OS — IGMP snooping is based on draft-ietf-magma-snoop-10.
- A maximum of 2k groups and 4k virtual local area networks (VLAN) are supported.
- IGMP snooping is not supported on the default VLAN interface.
- Flooding of unregistered multicast traffic is enabled by default.
- Queries are not accepted from the server side ports and are only accepted from the uplink LAG.
- Reports and Leaves are flooded by default to the uplink LAG irrespective of whether it is an mrouter port or not.

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. To disable multicast flooding on all VLAN ports, enter the **no ip igmp snooping flood** command in global configuration mode.

When multicast flooding is disabled, unregistered multicast data traffic is forwarded to only multicast router ports on all VLANs. If there is no multicast router port in a VLAN, unregistered multicast data traffic is dropped.

Displaying IGMP Information

Use the `show` commands from the below table, to display information on IGMP. If you specify a group address or interface:

- Enter a group address in dotted decimal format; for example, 225.0.0.0.
- Enter an interface in one of the following formats: `tengigabitethernet` *`slot/port`*, `port-channel` *`port-channel-number`*, or `vlan` *`vlan-number`*.

Displaying IGMP Information

Command	Output
show ip igmp snooping groups [*group-address* [detail] | detail | *interface* [*group-address* [detail]]	Displays information on IGMP groups.
show ip igmp snooping interface [*interface*]	Displays IGMP information on IGMP-enabled interfaces.
show ip igmp snooping mrouter [vlan *vlan-number*]	Displays information on IGMP-enabled multicast router (mrouter) interfaces.
clear ip igmp snooping groups [*group-address* | *interface*]	Clears IGMP information for group addresses and IGMP-enabled interfaces.

`show ip igmp snooping groups` **Command Example**

```
Dell# show ip igmp snooping groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address    Interface                Mode          Uptime    Expires
Last Reporter
226.0.0.1        Vlan 1500                INCLUDE       00:00:19  Never
```

```
1.1.1.2
226.0.0.1        Vlan 1600                INCLUDE       00:00:02  Never
1.1.1.2
Dell#show ip igmp snooping groups detail

Interface             Vlan 1500
Group                 226.0.0.1
Uptime                00:00:21
Expires               Never
Router mode           INCLUDE
Last reporter         1.1.1.2
Last reporter mode    INCLUDE
Last report received  IS_INCL
Group source list
Source address                    Uptime      Expires
1.1.1.2                           00:00:21    00:01:48
  Member Ports: Po 1

Interface             Vlan 1600
Group                 226.0.0.1
Uptime                00:00:04
Expires               Never
Router mode           INCLUDE
Last reporter         1.1.1.2
Last reporter mode    INCLUDE
Last report received  IS_INCL
Group source list
Source address                    Uptime      Expires
1.1.1.2                           00:00:04    00:02:05
  Member Ports: Po 1
Dell#
```

`show ip igmp snooping interface` **Command Example**

```
Dell# show ip igmp snooping interface

Vlan 2 is up, line protocol is down
  Inbound IGMP access group is not set
  Interface IGMP group join rate limit is not set
  IGMP snooping is enabled on interface
  IGMP Snooping query interval is 60 seconds
  IGMP Snooping querier timeout is 125 seconds
  IGMP Snooping last member query response interval is 1000 ms
  IGMP snooping fast-leave is disabled on this interface
  IGMP snooping querier is disabled on this interface
Vlan 3 is up, line protocol is down
  Inbound IGMP access group is not set
  Interface IGMP group join rate limit is not set
  IGMP snooping is enabled on interface
  IGMP Snooping query interval is 60 seconds
  IGMP Snooping querier timeout is 125 seconds
  IGMP Snooping last member query response interval is 1000 ms
  IGMP snooping fast-leave is disabled on this interface
  IGMP snooping querier is disabled on this interface
--More--
```

`show ip igmp snooping mrouter` **Command Example**

```
Dell# show ip igmp snooping mrouter
Interface Router Ports
Vlan 1000 Po 128
Dell#
```

23

Interfaces

This chapter describes 100/1000/10000 Mbps Ethernet, 10 Gigabit Ethernet, and 40 Gigabit Ethernet interface types, both physical and logical, and how to configure them with the Dell Networking operating software (OS).

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
- Server Ports

Advanced Interface Configuration

- Bulk Configuration
- Define the Interface Range
- Monitoring and Maintaining Interfaces
- Splitting QSFP Ports to SFP+ Ports
- Configure MTU Size on an Interface
- Layer 2 Flow Control Using Ethernet Pause Frames
- Configure the MTU Size on an Interface
- Port-Pipes
- Auto-Negotiation on Ethernet Interfaces
- View Advanced Interface Information

NOTE: You can also perform some of the configurations using the Web GUI - Dell Blade IO Manager. For more information, see the Dell Blade IO Manager Online Help.

Interface Types

The following table describes different interface types.

Table 21. Interface Types

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 - No Shutdown (enabled) L3 - Shutdown (disabled)

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 the system 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.

Example of the `show interfaces` Command

Example of the `show ip interfaces brief` Command

Example of the `show running-config` Command to View Physical Interfaces

The following example shows the configuration and status information for one interface.

```
Dell#show interfaces tengigabitethernet 0/16
TenGigabitEthernet 0/16 is up, line protocol is up
Hardware is DellForce10Eth, address is 00:1e:c9:f1:00:05
    Current address is 00:1e:c9:f1:00:05
Server Port AdminState is Up
Pluggable media not present
Interface index is 38080769
Internet address is not set
Mode of IP Address Assignment : NONE
DHCP Client-ID :tenG145001ec9f10005
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 5d1h18m
Queueing strategy: fifo
Input Statistics:
    34561 packets, 6266197 bytes
    38 64-byte pkts, 4373 over 64-byte pkts, 21491 over 127-byte pkts
    8659 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    21984 Multicasts, 12577 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics:
    44329 packets, 4722779 bytes, 0 underruns
    0 64-byte pkts, 44329 over 64-byte pkts, 0 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    44329 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: 4d0h28m
    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
Dell
```

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, GigabitEthernet interface 1/5 is in Layer 3 mode because an IP address has been assigned to it and the interface's status is operationally up.

```
Dell#show ip interface brief
Interface            IP-Address  OK? Method  Status                Protocol
GigabitEthernet 1/0  unassigned  NO  Manual  administratively down  down
GigabitEthernet 1/1  unassigned  NO  Manual  administratively down  down
GigabitEthernet 1/2  unassigned  YES Manual  up                    up
GigabitEthernet 1/3  unassigned  YES Manual  up                    up
```

```
GigabitEthernet 1/4  unassigned  YES Manual  up                    up
GigabitEthernet 1/5  10.10.10.1  YES Manual  up                    up
GigabitEthernet 1/6  unassigned  NO  Manual  administratively down  down
GigabitEthernet 1/7  unassigned  NO  Manual  administratively down  down
GigabitEthernet 1/8  unassigned  NO  Manual  administratively down  down
```

To view only configured interfaces, use the `show interfaces configured` command in the EXEC Privilege mode. In the previous example, GigabitEthernet interface 1/5 is in Layer 3 mode because an IP address has been assigned to it and the interface's status is operationally up.

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 line cards.

```
Dell#show running
Current Configuration ...
!
interface GigabitEthernet 9/6
  no ip address
  shutdown
!
interface GigabitEthernet 9/7
  no ip address
  shutdown
!
interface GigabitEthernet 9/8
  no ip address
  shutdown
!
interface GigabitEthernet 9/9
  no ip address
  shutdown
```

Configuring the Default Interface

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
   ```

   ```
   Dell(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 the interface to the factory default state.
   CONFIGURATION mode

   ```
   default interface interface-type slot/port
   ```

   ```
   Dell(conf)#default interface tengigabitethernet 1/5
   ```
3. Verify the configuration.

INTERFACE mode

```
show config
```

```
Dell(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 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 slot/port` command.

1. Enter the keyword `interface` then the type of interface and slot/port information.

   CONFIGURATION mode

   ```
   interface interface-type
   ```

   - For the Management interface on the RPM, enter the keyword `ManagementEthernet` then the 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.
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 switch interfaces support Layer 2 and Layer 3 traffic over the 100/1000/10000, 10-Gigabit, 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 Physical Interfaces.

Dell Networking OS Behavior: The MXL 10/40GbE switch systems use 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
- Configuring Layer 3 (Interface) Mode
- Configuring Layer 3 (Network) Mode
- Management Interfaces
- Auto-Negotiation on Ethernet Interfaces
- Adjusting the Keepalive Timer
- Clearing Interface Counters

Overview of Layer Modes

On all systems running the Dell 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 22. Layer modes

Type of Interface	Possible Modes	Requires Creation	Default State
10/100/1000 Ethernet, Gigabit Ethernet, 10 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
  ```

Example of a Basic Layer 2 Interface Configuration

```
Dell(conf-if)#show config
!
interface Port-channel 1
  no ip address
```

```
  switchport
  no shutdown
Dell(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
  ```

For information about enabling and configuring the Spanning Tree Protocol, refer to Spanning Tree Protocol (STP).

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
  ```
- Enable the interface.
  INTERFACE mode

  ```
  no shutdown
  ```

```
Dell(conf-if)#show config
!
interface TenGigabitEthernet 1/5
  ip address 10.10.10.1 /24
  no shutdown
Dell(conf-if)#
```

Example of Error Due to Issuing a Layer 3 Command on a Layer 2 Interface

If an interface is in the incorrect layer mode for a given command, an error message displays (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.

```
Dell(conf-if)#show config
!
interface GigabitEthernet 1/2
  no ip address
```

```
  switchport
  no shutdown
Dell(conf-if)#ip address 10.10.1.1 /24
% Error: Port is in Layer 2 mode Gi 1/2.
Dell(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.

Example of the `show ip interface` Command

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.

```
Dell(conf-if-vl-10)#do sh int vl 10
Vlan 10 is up, line protocol is up
Address is 00:1e:c9:f1:03:38, Current address is 00:1e:c9:f1:03:38
Interface index is 1107787786
Internet address is 5.5.5.1/24
Mode of IP Address Assignment : MANUAL
DHCP Client-ID: vlan10001ec9f10338
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:01:09
Queueing strategy: fifo
Time since last interface status change: 00:00:46
```

Management Interfaces

The IOM management interface has both a public IP and private IP address on the internal fabric D interface.

The public IP address is exposed to the outside world for Web GUI configurations/WSMAN and other proprietary traffic. You can statically configure the public IP address or obtain the IP address dynamically using the dynamic host configuration protocol (DHCP).

NOTE: When you shut down a management interface, connectivity to the interface's private IP address is disabled.

You can access the full switch using:

- Internal RS-232 using the chassis management controller (CMC). Telnet into CMC and do a `connect -b switch-id` to get console access to corresponding IOM.
- External serial port with a universal serial bus (USB) connector (front panel): connect using the IOM front panel USB serial line to get console access (Labeled as USB B).
- Telnet/others using the public IP interface on the fabric D interface.
- CMC through the private IP interface on the fabric D interface.

The MXL switch system supports the management Ethernet interface as well as the standard interface on any front-end port. You can use either method to connect to the system.

Configuring Management Interfaces on the MXL Switch

On the MXL Switch IO Module, the dedicated management interface provides management access to the system.
You can configure this interface with the Dell Networking OS, 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 the Dell Networking OS. In addition, proxy ARP is not supported on this interface.

To configure a management interface, use the following commands.

For additional management access, IOM supports the default VLAN (VLAN 1) L3 interface in addition to the public fabric D management interface. You can assign the IP address for the VLAN 1 default management interface using the setup wizard (or) through the CLI.

If you do not configure the VLAN 1 default using the wizard or CLI presented in startup-config, by default, the VLAN 1 management interface gets its IP address using DHCP.

There is only one management interface for the whole stack.

You can manage the MXL Switch from any port. Configure an IP address for the port using the `ip address` command. Enable the IP address for the port using the `no shutdown` command. You can use the `description` command from INTERFACE mode to note that the interface is the management interface. There is no separate management routing table, so you must configure all routes in the IP routing table (use the `ip route` command).

- Enter the slot and the port (0) to configure a Management interface.
  CONFIGURATION mode

  `interface managementethernet` *interface*

  The slot range is 0–0.

- 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).

Example of the `show interface` Command

To display the configuration for a given port, use the `show interface` command from EXEC Privilege mode, as shown in the following example.

To display the routing table for a given port, use the `show ip route` command from EXEC Privilege mode.

```
Dell#show int tengig 0/16
TenGigabitEthernet 0/16 is up, line protocol is down
Hardware is DellForce10Eth, address is 00:1e:c9:bb:02:c2
    Current address is 00:1e:c9:bb:02:c2
Server Port AdminState is Down
Pluggable media not present
Interface index is 38080769
Internet address is not set
Mode of IP Address Assignment : NONE
DHCP Client-ID :tenG145001ec9bb02c2
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
Flowcontrol rx off tx off
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 2w4d2h
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: 2w4d2h
Dell#
```

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, refer to 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.

The Dell 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.

Example of a Configuration for a VLAN Participating in an OSPF Process

```
interface Vlan 10
  ip address 1.1.1.2/24
  tagged GigabitEthernet 2/2-13
  tagged TenGigabitEthernet 5/0
  ip ospf authentication-key Dell 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 same commands found in the physical interface are also found in the Loopback interfaces.
For more information, refer to Access Control Lists (ACLs).

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.

VLAN Membership

A virtual LAN (VLANs) is a logical broadcast domain or logical grouping of interfaces in a LAN in which all data received is kept locally and broadcast to all members of the group. In Layer 2 mode, VLANs move traffic at wire speed and can span multiple devices. Dell Networking OS supports up to 4093 port-based VLANs and one default VLAN, as specified in IEEE 802.1Q.
VLAN provide the following benefits:

- Improved security because you can isolate groups of users into different VLANs.
- Ability to create one VLAN across multiple devices.

On an Aggregator in standalone mode, all ports are configured by default as members of all (4094) VLANs, including the default VLAN. All VLANs operate in Layer 2 mode. You can reconfigure the VLAN membership for individual ports by using the `vlan tagged` or `vlan untagged` commands in INTERFACE configuration mode (Configuring VLAN Membership). Physical Interfaces and port channels can be members of VLANs.

NOTE: You can assign a static IP address to default VLAN 1 using the `ip address` command. To assign a different VLAN ID to the default VLAN, use the `default vlan-id vlan-id` command.

Following table lists out the VLAN defaults in Dell Networking OS:

Table 23. VLAN Defaults

Feature	Default
Mode	Layer 2 (no IP address is assigned)
Default VLAN ID	VLAN 1

Default VLAN

When an Aggregator boots up, all interfaces are up in Layer 2 mode and placed in the default VLAN as untagged interfaces. Only untagged interfaces can belong to the default VLAN.

By default, VLAN 1 is the default VLAN. To change the default VLAN ID, use the `default vlan-id <1-4094>` command in CONFIGURATION mode. You cannot delete the default VLAN.

Port-Based VLANs

Port-based VLANs are a broadcast domain defined by different ports or interfaces. In Dell Networking OS, a port-based VLAN can contain interfaces from different stack units within the chassis. Dell 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, it must be in Layer 2 mode. After you place an interface in Layer 2 mode, it is automatically placed in the default VLAN. Dell Networking OS 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. The information that is preserved as the frame moves through the network. The below figure shows the structure of a frame with a tag header. The VLAN ID is inserted in the tag header.

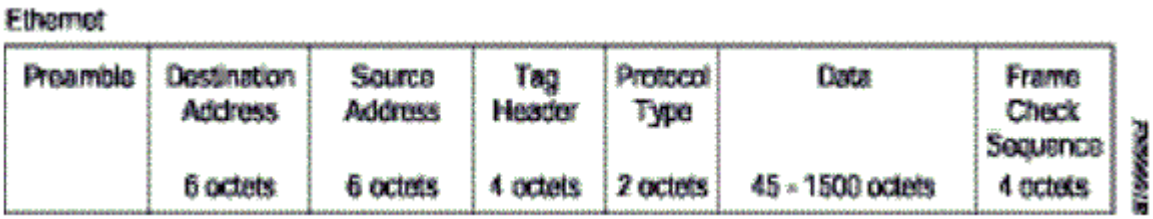


Figure 49. Tagged Frame Format

The tag header contains some key information used by Dell Networking OS:

- 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 1518 bytes 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.

Configuring VLAN Membership

By default, all Aggregator ports are member of all (4094) VLANs, including the default untagged VLAN 1. You can use the CLI or CMC interface to reconfigure VLANs only on server-facing interfaces (1–8) so that an interface has membership only in specified VLANs.

To assign an Aggregator interface in Layer 2 mode to a specified group of VLANs, use the `vlan tagged` and `vlan untagged` commands. To view which interfaces are tagged or untagged and to which VLAN they belong, use the show vlan command (Displaying VLAN Membership).

To reconfigure an interface as a member of only specified tagged VLANs, enter the `vlan tagged` command in INTERFACE mode:

1. Add the interface as a tagged member of one or more VLANs, where: *vlan-id* specifies a tagged VLAN number. Range: 2-4094.

   INTERFACE mode

   ```
   vlan tagged {vlan-id }
   ```

To reconfigure an interface as a member of only specified untagged VLANs, enter the `vlan untagged` command in INTERFACE mode:

1. Add the interface as an untagged member of one or more VLANs, where: *vlan-id* specifies an untagged VLAN number. Range: 2-4094.

   INTERFACE mode

   ```
   vlan untagged {vlan-id}
   ```

If you configure additional VLAN membership and save it to the startup configuration, the new VLAN configuration takes place immediately.

Dell Networking OS Behavior: When two or more server-facing ports with VLAN membership are configured in a LAG based on the NIC teaming configuration in connected servers learned via LACP, the resulting LAG is a tagged member of all the configured VLANs and an untagged member of the VLAN to which the port with the lowest port ID belongs. For example, if port 0/3 is an untagged member of VLAN 2 and port 0/4 is an untagged member of VLAN 3, the resulting LAG consisting of the two ports is an untagged member of VLAN 2 and a tagged member of VLAN 3.

Displaying VLAN Membership

To view the configured VLANs, enter the show vlan command in EXEC privilege mode:

```
Dell#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, H - VSN tagged
```

```
   i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

    NUM    Status    Description                    Q Ports
*   1      Inactive                                 U Te 0/1-8
Dell#
```

NOTE: A VLAN is active only if the VLAN contains interfaces and those interfaces are operationally up. In the above example, VLAN 1 is inactive because it does not contain any interfaces. The other VLANs listed contain enabled interfaces and are active. 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.

Adding an Interface to a Tagged VLAN

The following example shows you how to add a tagged interface (Te 0/2) to the VLANs.

Enter the `vlan tagged` command to add interface Te 0/2 to VLANs 2 - 4, which is as shown below.

Enter the `show config` command to verify that interface Te 0/2 is a tagged member of the VLANs.

```
Dell(conf-if-te-0/2)#vlan tagged ?
VLAN-RANGE        Comma/Hyphen separated VLAN ID set
Dell(conf-if-te-0/2)#vlan tagged 2,3-4
Dell(conf-if-te-0/2)#show config
!
interface TenGigabitEthernet 0/2
mtu 12000
vlan tagged 2-4
!
port-channel-protocol LACP
port-channel 1 mode active
!
protocol lldp
advertise management-tlv system-name
dcbx port-role auto-downstream
no shutdown
Dell(conf-if-te-0/2)#
```

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.

NOTE: When you remove a tagged interface from a VLAN (using the `no vlan tagged` command), it remains tagged only if it is a tagged interface in another VLAN. If you remove the tagged interface from the only VLAN to which it belongs, the interface is placed in the default VLAN as an untagged interface.

Adding an Interface to an Untagged VLAN

To move an untagged interfaces from the default VLAN to another VLAN, use the `vlan untagged` command as shown in the below figure.

```
Dell(conf)# interface tengigabit 0/2
Dell(conf-if-te-0/2)#vlan untagged ?
<1-4094> Untagged VLAN id
Dell(conf-if-te-0/2)#
Dell(conf-if-te-0/2)#vlan untagged 4094
Dell(conf-if-te-0/2)#show config
!
interface TenGigabitEthernet 0/2
mtu 12000
vlan untagged 4094
```

```
!
port-channel-protocol LACP
port-channel 1 mode active
!
protocol lldp
advertise management-tlv system-name
dcbx port-role auto-downstream
no shutdown
Dell(conf-if-te-0/2)#
```

VLAN Configuration on Physical Ports and Port-Channels

Unlike other Dell Networking OS platforms, IOA allows VLAN configurations on port and port-channel levels. This allows you to assign VLANs to a port/port-channel.

NOTE: In PMUX mode, in order to avoid loops, only disjoint VLANs are allowed between the uplink ports/uplink LAGs and uplink-to-uplink switching is disabled.

1. Initialize the port with configurations such as admin up, portmode, and switchport.

```
Dell#configure
Dell(conf)#int tengigabitethernet 0/1
Dell(conf-if-te-0/1)#no shutdown
Dell(conf-if-te-0/1)#portmode hybrid
Dell(conf-if-te-0/1)#switchport
```

2. Configure the tagged VLANs 10 through 15 and untagged VLAN 20 on this port.

```
Dell(conf-if-te-0/1)#vlan tagged 10-15
Dell(conf-if-te-0/1)#vlan untagged 20
Dell(conf-if-te-0/1)#
```

3. Show the running configurations on this port.

```
Dell(conf-if-te-0/1)#show config
!
interface TenGigabitEthernet 0/1
portmode hybrid
switchport
vlan tagged 10-15
vlan untagged 20
no shutdown
Dell(conf-if-te-0/1)#end
Dell#
```

4. Initialize the port-channel with configurations such as admin up, portmode, and switchport.

```
Dell#configure
Dell(conf)#int port-channel 128
Dell(conf-if-po-128)#portmode hybrid
Dell(conf-if-po-128)#switchport
```

5. Configure the tagged VLANs 10 through 15 and untagged VLAN 20 on this port-channel.

```
Dell(conf-if-po-128)#vlan tagged 10-15
Dell(conf-if-po-128)#
Dell(conf-if-po-128)#vlan untagged 20
```

6. Show the running configurations on this port-channel.

```
Dell(conf-if-po-128)#show config
!
interface Port-channel 128
portmode hybrid
switchport
vlan tagged 10-15
vlan untagged 20
```

```
shutdown
Dell(conf-if-po-128)#end
Dell#
```

7. Show the VLAN configurations.

```
Dell#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, H - VSN tagged
   i - Internal untagged, I - Internal tagged, v - VLT
untagged, V - VLT tagged

    NUM     Status      Description     Q     Ports
* 1        Active                        U Te 0/3
  10        Active                          T Po128(Te 0/4-5)
                                              T Te 0/1
  11        Active                          T Po128(Te 0/4-5)
  12     Active                           T Po128(Te 0/4-5)
                                              T Te 0/1
  13        Active                         T Po128(Te 0/4-5)
                                              T Te 0/1
  14     Active                           T Po128(Te 0/4-5)
                                              T Te 0/1
  15     Active                           T Po128(Te 0/4-5)
                                              T Te 0/1
  20     Active                           U Po128(Te 0/4-5)
                                              U Te 0/1
Dell#
```

You can remove the inactive VLANs that have no member ports using the following command:

```
Dell#configure
Dell(conf)#no interface vlan vlan-id
```

vlan-id — Inactive VLAN with no member ports

You can remove the tagged VLANs using the `no vlan tagged vlan-range` command. You can remove the untagged VLANs using the `no vlan untagged` command in the physical port/port-channel.

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 the Dell 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 40-Gigabit interface by aggregating four 10-Gigabit Ethernet interfaces together. If one of the five interfaces fails, traffic is redistributed across the three remaining interfaces.

Port Channel Implementation

The Dell 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, refer to Link Aggregation Control Protocol (LACP).

There are 128 port-channels with 16 members per channel.

As soon as you configure a port channel, the system 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 line card resets and chassis 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 100, 1000, or 10000 Mbps Ethernet interfaces and TenGigabit Ethernet interfaces. The interface speed (100, 1000, or 10000 Mbps) the port channel uses is determined by the first port channel member that is physically up. The system disables the interfaces that do match the interface speed that the first channel member sets. That first interface may be the first interface that is physically brought up 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 Gigabit Ethernet interface, all interfaces at 1000 Mbps are kept up, and all 100/1000/10000 interfaces that are not set to 1000 speed or auto negotiate are disabled.

100/1000/10000 Mbps Interfaces in Port Channels

When both 100/1000/10000 interfaces and TenGigabitEthernet 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. At that time, 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, the system disables them.

For example, if four interfaces (TenGig 0/0, 0/1, 0/2, and 0/3) in which TenGig 0/0 and TenGig 0/3 are set to speed 100 Mb/s and the others are set to 10000 Mb/s, with all interfaces enabled, and you add them to a port channel by entering `channel-member tengigabitethernet 0/0-3` while in port channel interface mode, and the system determines if the first interface specified (TenGig 0/0) is up. After it is up, the common speed of the port channel is 100 Mb/s. The system disables those interfaces configured with speed 1000 Mb/s or whose speed is 1000 Mb/s as a result of auto-negotiation.

In this example, you can change the common speed of the port channel by changing its configuration so the first enabled interface referenced in the configuration is a 1000 Mb/s speed interface. You can also change the common speed of the port channel here by setting the speed of the TenGig 0/0 interface to 1000 Mb/s.

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)

Creating a Port Channel

You can create up to 128 port channels with 16 port members per group on an MXL switch.
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

You can add any physical interface to a port channel if the interface configuration is minimal.

NOTE: Port channels can contain a mix of 100/1000/10000 Ethernet interfaces and 10 Gigabit Ethernet interface, but the Dell Networking OS disables the interfaces that are not the same speed of the first channel member in the port channel (refer to 100/1000/10000 Mbps Interfaces in Port Channels).

You can configure only the following commands on an interface if it is a member of a port channel:

- `description`
- `shutdown/no shutdown`
- `mtu`
- `ip mtu` (if the interface is on a Jumbo-enabled by default)

NOTE: The MXL switch supports jumbo frames by default (the default maximum transmission unit [MTU] is 1554 bytes) You can configure the MTU using 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.

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
   ```

Example of the `show interfaces port-channel brief` Command

Example of the `show interface port-channel` Command

Example of Error Due to an Attempt to Configure an Interface that is Part of a Port Channel

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.

```
Dell#show int port brief
Codes: L - LACP Port-channel

  LAG Mode Status Uptime   Ports
  1   L3   down   00:00:00 Te 0/16 (Down)
Dell#
```

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.

```
Dell#show int port-channel
Port-channel 1 is down, line protocol is down
Hardware address is 00:1e:c9:f1:00:05, Current address is 00:1e:c9:f1:00:05
Interface index is 1107755009
Minimum number of links to bring Port-channel up is 1
Internet address is not set
Mode of IP Address Assignment : NONE
DHCP Client-ID :lag1001ec9f10005
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
Members in this channel: Te 0/16(D)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:05:44
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: 00:05:44
```

When more than one interface is added to a Layer 2-port channel, the system 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 GigabitEthernet 1/6 is part of port channel 1, which is in Layer 2 mode, and an error message appeared when an IP address was configured.

```
Dell(conf-if-po-1)#show config
!
interface Port-channel 1
  no ip address
  channel-member TenGigabitEthernet 0/16
  shutdown
Dell(conf-if-po-1)#
Dell(conf-if-po-1)#int tengig 1/6
Dell(conf-if)#ip address 10.56.4.4 /24
% Error: Te 1/6 Port is part of a LAG.
Dell(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.

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
   ```

Example of Moving an Interface to a New Port Channel

The following example shows moving the TenGigabitEthernet 1/8 interface from port channel 4 to port channel 3.

```
Dell(conf-if-po-1)#show config
!
interface Port-channel 1
  no ip address
  channel-member TenGigabitEthernet 0/16
  shutdown
Dell(conf-if-po-1)#no chann tengig 1/8
Dell(conf-if-po-1)#int port 5
Dell(conf-if-po-5)#channel tengig 1/8
Dell(conf-if-po-5)#show conf
!
interface Port-channel 5
  no ip address
  channel-member TenGigabitEthernet 1/8
  shutdown
Dell(conf-if-po-5)#
```

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

```
Dell#config t
Dell(conf)#int po 1
```

```
Dell(conf-if-po-1)#minimum-links 5
Dell(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`

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 Networking OS uses hash algorithms for distributing traffic evenly over channel members in a port channel (LAG). The hash algorithm distributes traffic among 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 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. Their syntax and implementation are somewhat different between the E-Series and the C-Series and S-Series

NOTE: Hash-based load-balancing on MPLS does not work when packet-based hashing (load-balance ip-selection packet-based) is enabled.

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.

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 Networking OS Command Reference Guide*.
- Change the Hash algorithm seed value to get better hash value

Hash seed is used to compute the hash value. By default hash seed is chassis MAC 32 bits. we can also change the hash seed by the following command.

CONFIGURATION mode

```
hash-algorithm seed {seed value}
```

- Change to another algorithm.

  CONFIGURATION mode

Example of the `hash-algorithm` Command

```
Dell(conf)#hash-algorithm ecmp xor1 lag crc16
Dell(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` —always 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.

Server Ports

By default, the MXL switch allows the server ports to come up as switch ports in `no shut` mode, ready to switch traffic.

Default Configuration without Start-up Config

This feature is enabled by default and can be enabled on reload by deleting the start-up config file.

On reload, all the server ports (1-32) come up as switch ports in No Shut mode. Uplinks remain in Shut mode ensuring that there are no network loops.

With this feature, you can install servers and test their connectivity by running applications on the servers, even before configuring VLAN membership, STP on all interfaces or uplinks.

NOTE: This feature does not impact BMP mode. It always applies when reloading in Normal mode.

Important Points to Remember

- On a new MXL switch running the Dell Networking OS version 9.2(0.0), with no saved startup configuration, the switch comes up with all server ports as switch ports in No Shut state. When you configure STP, the switch brings up the uplink and saves the running configuration to the startup-config file. All the server ports without any specific configuration have the default configuration of Layer2 switch port and No Shut mode saved.
- On an existing MXL switch with a saved startup configuration, running an older Dell Networking OS version, an upgrade to a new version does not change the current behavior. This is because the start-up config file in older Dell Networking OS versions have the default configuration of Shut mode for all the server ports without any specific configuration. To enable this feature after upgrading a switch with a saved startup configuration, delete the start-up config file and reboot the switch. This allows all the server ports to come as Layer2 switch ports in No Shut state.
- In a stacked configuration of MXL switches, the behavior is similar to a standalone configuration. If a start-up config file is detected at bootup, the entire stack reboots using the saved configuration. If no start-up config file is detected at restart, the entire logical switch, including master unit, standby master, and any stack units restart with all server ports as Layer2 switch ports in No Shut mode.
- If a new stack unit is added to an existing stack, by default, the server side interfaces always start in Shut mode. If the startup configuration is deleted after a stack unit was added to a stack and the stack is reloaded, on reboot the entire logical switch comes up with all server ports as Layer2 switch ports in No Shut mode.

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.

To display all interfaces that have been validated under the interface range context, use the `show range` command in Interface Range mode.

To display the running configuration only for interfaces that are part of interface range, use the `show configuration` command in Interface Range mode.

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,gi 3/0,fa 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)

```
Dell(conf)# interface range tengigabitethernet 5/1 - 23
Dell(conf-if-range-te-5/1-23)# no shutdown
Dell(conf-if-range-te-5/1-23)#
```

Create a Multiple-Range

The following is an example of multiple range.

Example of the `interface range` Command (Multiple Ranges)

```
Dell(conf)#interface range tengigabitethernet 3/0 , tengigabitethernet 2/1 -
47 , vlan 1000
Dell(conf-if-range-te-2/1-47)#
```

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

```
Dell(conf)#interface range vlan 1 , vlan 1 , vlan 3 , vlan 3
Dell(conf-if-range-vl-1,vl-3)#
Dell(conf)#interface range tengigabitethernet 2/0 - 23 , tengigabitethernet 2/0
- 23 , tengigab 2/0 - 23
Dell(conf-if-range-te-2/0-23)#
```

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

```
Dell(conf)#interface range tengigabitethernet 2/0 - 23 , tengigab 2/1 - 10
Dell(conf-if-range-te-2/0-23)#
```

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

```
Dell(conf)#inte ra tengig 2/1 - 11 , tengig 2/1 - 23
Dell(conf-if-range-te-2/1-23)#
```

Commas

The following is an example of how to use commas to add different interface types to the range, enabling all Gigabit Ethernet interfaces in the range 5/1 to 5/23 and both Ten Gigabit Ethernet interfaces 1/1 and 1/2.

Example of Multiple-Range Bulk Configuration Gigabit Ethernet and Ten-Gigabit Ethernet

```
Dell(conf-if)# interface range tengigabitethernet 5/1 - 23, tengigabitethernet
1/1 - 2
Dell(conf-if-range-te-5/1-23)# no shutdown
Dell(conf-if-range-te-5/1-23)#
```

Add Ranges

The following example shows how to use commas to add VLAN and port-channel interfaces to the range.

Example of Multiple-Range Bulk Configuration with VLAN and Port-channel

```
Dell(conf-ifrange-te-5/1-23-te-1/1-2)# interface range Vlan 2 - 100 , Port 1 -
25
Dell(conf-if-range-te-5/1-23-te-1/1-2-vl-2-100-po-1-25)# no shutdown
Dell(conf-if-range)#
```

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` *`interface-range macro_name`* {vlan *`vlan_ID - vlan_ID`*} | {{tengigabitethernet | fortyGigE} *`slot/ interface - interface`*} [, {vlan *`vlan_ID - vlan_ID`*} {{tengigabitethernet | fortyGigE} *`slot/interface - interface`*}]

Define the Interface Range

The following example shows how to define an interface-range macro named “test” to select Fast Ethernet interfaces 5/1 through 5/4.

Example of the `define interface-range` Command for Macros

```
Dell(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
  ```

Example of Using a Macro to Change the Interface Range Configuration Mode

The following example shows how to change to the interface-range configuration mode using the interface-range macro named "test."

```
Dell(config)# interface range macro test
Dell(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

  ```
  monitor interface interface
  ```

  Enter the type of interface and slot/port information:
  - For a 100/1000/10000 Ethernet interface, enter the keyword `TenGigabitEthernet` then the 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.

Example of the `monitor interface` Command

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

```
Dell#monitor interface tengig 3/1

Dell Networking uptime is 1 day(s), 4 hour(s), 31 minute(s)
  Monitor time: 00:00:00 Refresh Intvl.: 2s
```

```
Interface: TenGig 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

Dell
```

Maintenance Using TDR

The time domain reflectometer (TDR) is supported on all Dell Networking switch/routers.
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 the condition of cables on 100/1000/10000 BASE-T modules, 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>
```

Splitting QSFP Ports to SFP+ Ports

The MXL 10/40GbE switch supports splitting a 40GbE port on the base module or a 2-Port 40GbE QSFP+ module into four 10GbE SFP+ ports using a 4x10G breakout cable.

NOTE: By default, the 40GbE ports on a 2-Port 40GbE QSFP+ module come up in 4x10GbE (quad) mode as eight 10GbE ports. On the base module, you must convert the 40GbE ports to 4x10GbE mode as described in the following section.

NOTE: When you split a 40G port (such as fo 0/4) into four 10G ports, the 40G interface configuration is available in the startup configuration when you save the running configuration by using the write memory command. When a reload of the system occurs, the 40G interface configuration is not applicable because the 40G ports are split into four 10G ports after the reload operation. While the reload is in progress, you might see error messages when the configuration file is being loaded. You can ignore these error messages. Similarly, such error messages are displayed during a reload after you configure the four individual 10G ports to be stacked as a single 40G port.

- Split a single 40G port into 4-10G ports.
  CONFIGURATION mode

  ```
  stack-unit port number portmode quad
  ```

  - `stack-unit`: Enter the stack member unit identifier of the stack member to reset. The range is from 0 to 5.
  - `port <port number>` : Enter the port number of the 40G port to be split. The valid values on base module: 33 or 37; OPTM SLOT 0: 41 or 45; OPTM SLOT 1: 49 or 53.
  - `portmode quad`: Identifies the uplink port as a split 10GbE SFP+ port.

  To display the stack-unit number, enter the `show system brief` command.
- Save the configuration and reload the switch.
  CONFIGURATION mode

  ```
  write memory

  reload
  ```

Merging SFP+ Ports to QSFP 40G Ports

To remove FANOUT mode in 40G QSFP Ports, use the following commands.

1. Merge 4-10G ports to a single 40G port.
   CONFIGURATION mode

   ```
   no stack-unit port number portmode quad
   ```

   - *stack-unit*: Enter the stack member unit identifier of the stack member to reset. The range is from 0 to 5.
   - `port <port number>`: Enter the port number of the 40GbE QSFP+ port. Valid values on base module: 33 or 37; OPTM SLOT 0: 41 or 45; OPTM SLOT 1: 49 or 53.

- `portmode quad`: Identifies the uplink port as a split 10GbE SFP+ port.

2. Save the configuration and reload the switch.
   CONFIGURATION mode

   ```
   write memory

   reload
   ```

Important Points to Remember

- You cannot use split ports as stack-link to stack an MXL Switch.
- Split ports cannot be a part of any stacked system.
- The quad port must be in a default configuration before it can be split into 4x10G ports.
- The 40G port is lost in the configuration when the port is split; be sure the port is also removed from other L2/L3 feature configurations.
- The system must be reloaded after issuing the CLI for the change to take effect.

Configure the MTU Size on an Interface

The link MTU is the frame size of a packet. The IP MTU size is used for IP fragmentation.

If the system determines that the IP packet must be fragmented as it leaves the interface, the system divides the packet into fragments no bigger than the size set in the `ip mtu` command.

In the Dell Networking OS, MTU is defined as the entire Ethernet packet (Ethernet header + FCS + payload).

Because different networking vendors define MTU differently, check their documentation when planning MTU sizes across a network.

The following table lists the various Layer 2 overheads found in the Dell Networking OS and the number of bytes.

Table 24. Layer 2 Overhead

Transmission Media	MTU Range (in bytes)
Ethernet	594-12000 = link MTU 576-11982 = IP MTU

Converting a QSFP or QSFP+ Port to an SFP or SFP+ Port

You can convert a QSFP or QSFP+ port to an SFP or SFP+ port using the Quad to Small Form Factor Pluggable Adapter (QSA).

QSA provides smooth connectivity between devices that use Quad Lane Ports (such as the 40 Gigabit Ethernet adapters) and 10 Gigabit hardware that uses SFP+ based cabling. Using this adapter, you can effectively use a QSFP or QSFP+ module to connect to a lower-end switch or server that uses an SFP or SFP+ based module.

When connected to a QSFP or QSFP+ port on a 40 Gigabit adapter, QSA acts as an interface for the SFP or SFP+ cables. This interface enables you to directly plug in an SFP or SFP+ cable originating at a 10 Gigabit Ethernet port on a switch or server.

You can use QSFP optical cables (without a QSA) to split a 40 Gigabit port on a switch or a server into four 10 Gigabit ports. To split the ports, .enable the fan-out mode.

Similarly, you can enable the fan-out mode to configure the QSFP port on a device to act as an SFP or SFP+ port. As the QSA enables a QSFP or QSFP+ port to be used as an SFP or SFP+ port, Dell Networking OS does not immediately detect the QSA after you insert it into a QSFP port cage.

After you insert an SFP or SFP+ cable into a QSA connected to a 40 Gigabit port, Dell Networking OS assumes that all the four fanned-out 10 Gigabit ports have plugged-in SFP or SFP+ optical cables. However, the link UP event happens only for the first 10 Gigabit port and you can use only that port for data transfer. As a result, only the first fanned-out port is identified as the active 10 Gigabit port with a speed of 10G or 1G depending on whether you insert an SFP+ or SFP cable respectively.

NOTE: Although it is possible to configure the remaining three 10 Gigabit ports, the Link UP event does not occur for these ports leaving the lanes unusable. Dell Networking OS perceives these ports to be in a Link Down state. You must not try to use these remaining three 10 Gigabit ports for actual data transfer or for any other related configurations.

NOTE: Trident2 chip sets do not work at 1G speeds with auto-negotiation enabled. As a result, when you peer any device using SFP, the link does not come up if auto-negotiation is enabled. Therefore, disable auto-negotiation on platforms that currently use Trident2 chip sets (S6000 and Z9000). This limitation applies only when you convert QSFP to SFP using the QSA. This constraint does not apply for QSFP to SFP+ conversions using the QSA.

Important Points to Remember

- Before using the QSA to convert a 40 Gigabit Ethernet port to a 10 Gigabit SFP or SFP+ port, enable 40 G to 4*10 fan-out mode on the device.
- When you insert a QSA into a 40 Gigabit port, you can use only the first 10 Gigabit port in the fan-out mode to plug-in SFP or SFP+ cables. The remaining three 10 Gigabit ports are perceived to be in Link Down state and are unusable.
- You cannot use QSFP Optical cables on the same port where QSA is used.
- When you remove the QSA module alone from a 40 Gigabit port, without connecting any SFP or SFP + cables; Dell Networking OS does not generate any event. However, when you remove a QSA module that has SFP or SFP+ optical cables plugged in, Dell Networking OS generates an SFP or SFP+ Removed event.
- The QSA module does not have a designated EEPROM. To recognize a QSA, Dell Networking OS reads the EEPROM corresponding to an SFP+ or SFP module that is plugged into QSA. The access location of this EEPROM is different from the EEPROM location of the QSFP+ module.
- The diagnostics application is capable of detecting insertion or removal of both the QSA as well as the SFP+ or SFP optical cables plugged into the QSA. In addition, the diagnostic application is also capable of reading the DDS and Vendor information from the EEPROM corresponding to SFP+ or SFP optical cables. As a result, no separate detection of QSA is required.

Example Scenarios

Consider the following scenarios:

- QSFP port 0 is connected to a QSA with SFP+ optical cables plugged in.
- QSFP port 4 is connected to a QSA with SFP optical cables plugged in.
- QSFP port 8 in fanned-out mode is plugged in with QSFP optical cables.
- QSFP port 12 in 40 G mode is plugged in with QSFP optical cables.

For these configurations, the following examples show the command output that the `show interfaces tengigbitethernet transceiver`, `show interfaces tengigbitethernet`, and `show inventory media` commands displays:

NOTE: In the following `show interfaces tengigbitethernet` commands, the ports 1,2, and 3 are inactive and no physical SFP or SFP+ connection actually exists on these ports. However, Dell Networking OS still perceives these ports as valid and the output shows that pluggable media (optical cables) is inserted into these ports. This is a software limitation for this release.

Layer 2 Flow Control Using Ethernet Pause Frames

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.

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 a 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. The only configuration applicable to half duplex ports is `rx off tx off`.

NOTE: If a port is over-subscribed, Ethernet Pause Frame flow control does not ensure no loss behavior.

The following error message appears when trying to enable flow control when you already configured half duplex: `Can't configure flowcontrol when half duplex is configure, config ignored.`

The following error message appears when trying to enable half duplex and flow control configuration is on: `Can't configure half duplex when flowcontrol is on, config ignored.`

Enabling Pause Frames

Enable Ethernet pause frames flow control on all ports on a chassis. If not, the system may exhibit unpredictable behavior.

NOTE: If you disable `rx flow control`, Dell 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. (also refer to iSCSI Optimization: Operation).

NOTE: After you disable DCB, if link-level flow control is not automatically enabled on an interface, to enable flow control, manually shut down the interface (`shutdown` command) and re-enable it (`no shutdown` command).

To enable pause frames, use the following command.

- Control how the system responds to and generates 802.3x pause frames on 10 Gig 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.

  The default is **rx off**.

Configure 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 MTU range is from 592 to 12000, with a default of 1500. IP MTU automatically configures.

The following table lists the various Layer 2 overheads found in the Dell Networking OS and the number of bytes.

Table 25. Difference between Link MTU and IP MTU

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.

Port-Pipes

A high-speed data bus connection used to switch traffic between front-end ports is known as the port pipe. A port pipe is a Dell Networking-specific term for the hardware path that packets follow through a system.

The MXL switch supports single port pipe only.

Auto-Negotiation on Ethernet Interfaces

By default, auto-negotiation of speed and duplex mode is enabled on 100/1000/10000 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 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.

For 100/1000/10000 Ethernet interfaces, the `negotiation auto` command is tied to the `speed` command. Auto-negotiation is always enabled when the `speed` command is set to 1000 in IOS.

Setting the Speed and Duplex Mode 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] 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 slot/port
   ```
5. Set the local port speed.
   INTERFACE mode

   ```
   speed {100 | 1000 | 10000 | auto}
   ```
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
   ```

Example of the `show interfaces status` Command to View Link Status

Example of Setting Port Speed and Disabling Auto-Negotiation

NOTE: The `show interfaces status` command displays link status, but not administrative status. For both link and administrative status, use the `show ip interface [interface | brief] [configuration]` command.

```
Dell#show interfaces status
Port Description Status Speed Duplex Vlan
Te 0/1           Down   Auto  Auto   --
Te 0/2           Down   Auto  Auto   --
Te 0/3           Down   Auto  Auto   --
Te 0/4           Down   Auto  Auto   --
Te 0/5           Down   Auto  Auto   --
Te 0/6           Down   Auto  Auto   --
Te 0/7           Down   Auto  Auto   --
Te 0/8           Down   Auto  Auto   --
Te 0/9           Down   Auto  Auto   --
Te 0/10          Down   Auto  Auto   --
Te 0/11          Down   Auto  Auto   --
Te 0/12          Down   Auto  Auto   --
Te 0/13          Down   Auto  Auto   --
[output omitted]
```

In the previous example, several ports display “Auto” in the Speed field, including port 0/1. In the following example, the speed of port 0/1 is set to 100Mb and then its auto-negotiation is disabled.

```
Dell#configure
Dell(config)#interface tengig 0/1
Dell(conf-if-te-0/1)#speed 100
Dell(conf-if-te-0/1)#duplex full
Dell(conf-if-te-0/1)#no negotiation auto
Dell(conf-if-te-0/1)#show config
```

```
!
interface TenGigabitEthernet 0/1
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 after you enable auto-negotiation.

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

```
Dell(conf)# int tengig 0/0
Dell(conf-if)#neg auto
Dell(conf-if-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
Dell(conf-if-autoneg)#mode ?
forced-master    Force port to master mode
forced-slave     Force port to slave mode
Dell(conf-if-autoneg)#
```

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.
When you use the `configured` keyword, only interfaces that have non-default configurations display.
Example of `show` Commands that Use the `configured` Keyword

Example of the `show interfaces switchport` Command

The following example lists the possible `show` commands that have the configured keyword available:

```
Dell#show interfaces configured
Dell#show interfaces tengigabitEthernet 0 configured
```

```
Dell#show ip interface configured
Dell#show ip interface tengigabitEthernet 1 configured
Dell#show interfaces fortygigabitEthernet 0 configured
Dell#show ip interface fortygigabitEthernet 1 configured
Dell#show ip interface brief configured
Dell#show running-config interfaces configured
Dell#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.

```
Dell#show interfaces switchport
Name: TenGigabitEthernet 13/0
802.1QTagged: True
Vlan membership:
Vlan 2

Name: TenGigabitEthernet 13/1
802.1QTagged: True
Vlan membership:
Vlan 2

Name: TenGigabitEthernet 13/2
802.1QTagged: True
Vlan membership:
Vlan 2

Name: TenGigabitEthernet 13/3
802.1QTagged: True
Vlan membership:
Vlan 2
--More--
```

Configuring the Interface Sampling Size

You can enter any value between five and 299 seconds (the default). If you enter 1 to 5 seconds, software polling is done at 5 second intervals. If you enter 6 to 10 seconds, software polling is done at 10 second interval. For any other value, 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
  ```

Example of the `rate-interval` Command

The bold lines shows the default value of 299 seconds, the change-rate interval of 100, and the new rate interval set to 100.

```
Dell#show interfaces
TenGigabitEthernet 10/0 is down, line protocol is down
Hardware is Dell Force10Eth, 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

Dell(conf)#interface tengigabitethernet 10/0
Dell(conf-if-te-10/0)#rate-interval 100

Dell#show interfaces
TenGigabitEthernet 10/0 is down, line protocol is down
Hardware is Dell Force10Eth, 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 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
```

Dynamic Counters

By default, counting is enabled for IPFLOW, IPACL, L2ACL, L2FIB.

For the remaining applications, the system 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 the Dell 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 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 from 1 to 128.
  - 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, 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`.

Example of the `clear counters` Command

When you enter this command, confirm that you want the Dell Networking OS to clear the interface counters for that interface.

```
Dell#clear counters tengig 0/0
Clear counters on TenGigabitEthernet 0/0 [confirm]
Dell#
```

Enhanced Control of Remote Fault Indication Processing

By default, MXL module processes RFI errors transmitted by remote peers and brings down the interface when an RFI error is detected.
You must enter the interface configuration mode before configuring Remote Fault Indication (RFI).
You can use the following CLI commands to enable or disable processing of received RFI events:

```
Dell(conf-if-te-1/3)#remote-fault-signaling rx ?
on Enable
off Disable
```

The default is "remote-fault-signaling rx on".

24

Internet Protocol Security (IPSec)

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 file transfer protocols (FTPs) and can operate in Transport mode. In Transport mode, IPSec encrypts only the packet payload; the IP header is unchanged. This is the default mode.

NOTE: Due to performance limitations on the control processor, you cannot enable IPSec on all packets in a communication session.

IPSec uses the following protocols:

- **Authentication Headers (AH)** — Disconnected integrity and origin authentication for IP packets
- **Encapsulating Security (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 and 3DES
  - MD5 and CBC
  - MD5 and DES
  - SHA1 and 3DES
  - SHA1 and CBC
  - SHA1 and DES

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 21

match 1 tcp a::1 /128 21 a::2 /128 0

match 2 tcp 1.1.1.1 /32 0 1.1.1.2 /32 21

match 3 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
```

25

IPv4 Routing

The Dell Networking 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 Networking operating system (OS).

IP Feature	Default
DNS	Disabled
Directed Broadcast	Disabled
Proxy ARP	Enabled
ICMP Unreachable	Disabled
ICMP Redirect	Disabled

IP Addresses

The Dell Networking OS supports IP version 4, as described in RFC 791. It also supports 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

In the Dell Networking OS, you can configure any IP address as a static route except IP addresses already assigned to interfaces.

NOTE: The Dell Networking OS versions 7.7.1.0 and later support 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. The system supports RFC 3021 with ARP.

NOTE: Even though Dell 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 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.
In the system, 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 interface
   ```

   - For a Loopback interface, enter the keyword `loopback` then a number from 0 to 16383.
   - For the Management interface on the RPM, enter the keyword `ManagementEthernet` then the slot/port information. The slot range is from 0 to 1. The port range is 0/0.
   - For a port channel interface, enter the keywords `port-channel` then a number from 1 to 128.
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - 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.

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.

Example the `show config` Command

```
Dell(conf-if-te-0/16)#show conf
!
interface TenGigabitEthernet 0/16
```

```
  no ip address
  shutdown
Dell(conf-if-te-0/16)#
```

Example of the show ip interface Command

```
Dell#show ip interface tengig 0/16
TenGigabitEthernet 0/16 is down, line protocol is down
Internet address is not set
IP MTU is 1500 bytes
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
Dell#
```

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` *`ip-address mask`* {*`ip-address`* | *`interface`* [*`ip-address`*]} [*`distance`*] [permanent] [tag *`tag-value`*]

  Use the following required and optional parameters:
  - *`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)
  - `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)

Example of the `show ip route static` Command

To view the configured routes, use the `show ip route static` command.

```
Dell#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 5/0      1/0       00:02:30
S 6.1.2.2/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.3/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.4/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.5/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.6/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.7/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.8/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
S 6.1.2.9/32   via 6.1.20.2, Te 5/0      1/0       00:02:30
```

```
S 6.1.2.10/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.11/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.12/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.13/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.14/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.15/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.16/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 6.1.2.17/32  via 6.1.20.2, Te 5/0      1/0      00:02:30
S 11.1.1.0/24  Direct, Nu 0              0/0      00:02:30
               Direct, Lo 0
--More--
```

The system installs a next hop that is on the directly connected subnet of current IP address on the interface (for example, if `interface tengig 0/0` is on 172.31.5.0 subnet, the system installs the static route).

The system 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. For example, if `tengig 0/0` has ip address on subnet 2.2.2.0 and if 172.31.5.43 recursively resolves to 2.2.2.0, the system installs the static route.

- When the interface goes down, the system withdraws the route.
- When the interface comes up, the system re-installs the route.
- When the recursive resolution is "broken," the system withdraws the route.
- When the recursive resolution is satisfied, the system re-installs the route.

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}
  ```

Example of the `show ip route static` Command

To view the configured static routes for the management port, use the `show ip management-route` command in EXEC privilege mode.

```
Dell#show ip management-route all

Destination    Gateway                State
-----------    -------                -----
1.1.1.0/24     172.31.1.250           Active
172.16.1.0/24  172.31.1.250           Active
172.31.1.0/24  ManagementEthernet 1/0 Connected

Dell#
```

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.

PMTD is enabled by default on the switches that support this capability. To enable PMTD to function correctly, you must enter the `ip unreachables` command on a VLAN interface to enable the generation of ICMP unreachable messages. PMTD is supported on all the layer 3 VLAN interfaces. Because all of the Layer 3 interfaces are mapped to the VLAN ID of 4095 when VLAN sub-interfaces are configured on it, it is not possible to configure unique layer 3 MTU values for each of the layer 3 interfaces. If a VLAN interface contains both IPv4 and IPv6 addresses configured on it, both the IPv4 and IPv6 traffic are applied the same MTU size; you cannot specify different MTU values for IPv4 and IPv6 packets.

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 amount of time 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

   ```
   Dell(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

   ```
   Dell>show ip tcp reduced-syn-ack-wait
   ```

Enabling Directed Broadcast

By default, the system drops directed broadcast packets destined for an interface. This default setting provides some protection against denial of service (DoS) attacks.
To enable the system 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 such commands 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.

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

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.

Example of the `show hosts` Command

To view current bindings, use the `show hosts` command.

```
Dell>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
Dell>
```

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, the system 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. The Dell 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 the system cannot resolve the

domain, it tries the domain name assigned to the local system. If that does not resolve the partial domain, the system 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. The Dell 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.

Example of the `traceroute` Command

The following text is example output of DNS using the `traceroute` command.

```
Dell#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.force10networks.com (10.11.84.18) 000.000 ms 000.000 ms 000.000 ms
Dell#
```

ARP

The Dell 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, the system 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 the Dell 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 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 ip-address mac-address interface
  ```

  – *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.

Example of the `show arp` Command

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.

```
Dell#show arp

Protocol Address       Age(min) Hardware Address   Interface VLAN CPU
-------------------------------------------------------------------------------
---------
Internet 10.11.68.14   94       00:01:e9:45:00:03  Ma 0/0    -    CP
Internet 10.11.209.254 0        00:01:e9:45:00:03  Ma 0/0    -    CP
Dell#
```

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 port channel interface, enter the keywords `port-channel` then a number from 1 to 128.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a VLAN interface, enter the keyword `vlan` then a number between 1 and 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 the system, 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. Gratuitous ARP can:

- 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. When a gratuitous ARP is received, the system installs an ARP entry on the CPU.

To enable ARP learning via gratuitous ARP, use the `arp learn-enable` command in CONFIGURATION mode.

ARP Learning via ARP Request

In the Dell Networking OS versions prior to 8.3.1.0, the system 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.

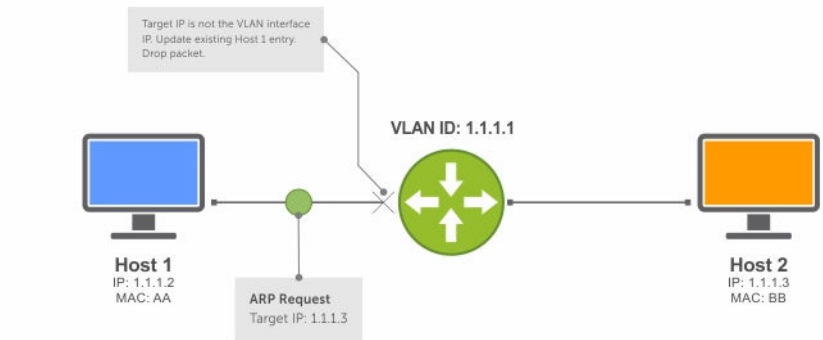


Figure 50. ARP Learning via ARP Request

Beginning with the Dell 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.

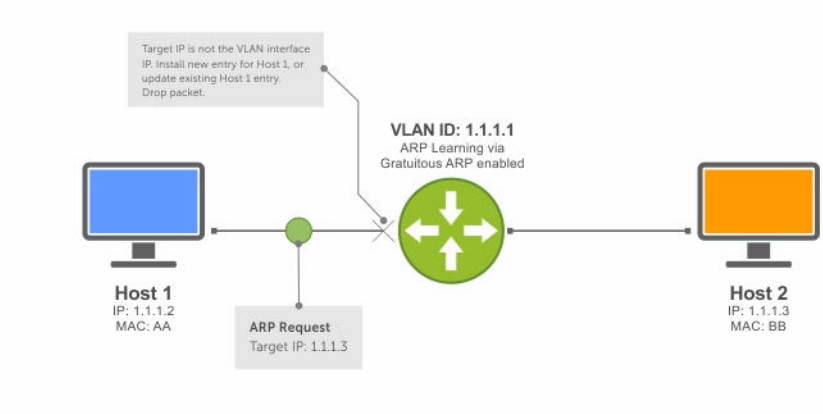


Figure 51. ARP Learning via ARP Request with ARP Learning via Gratuitous ARP Enabled

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

In the Dell Networking OS versions prior to 8.3.1.0, the number of ARP retries is set to five and is not configurable. After five retries, the system backs off for 20 seconds before it sends a new request. Beginning with the Dell Networking OS version 8.3.1.0, the number of ARP retries is configurable. The default backoff interval remains at 20 seconds. On the switch using the Dell Networking OS version 8.3.8.0 and later, the time between ARP re-send is configurable. This timer is an exponential backoff timer. Over the specified period, the time between ARP requests increases. This 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 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.

- Set the system 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.

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

Configuring the system to direct UDP broadcast is a one-step process:

1. Enable UDP helper and specify the UDP ports for which traffic is forwarded. Refer to 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`

Example of Enabling UDP Helper

Example of the `show ip udp-helper` Command

```
Dell(conf-if-te-1/1)#ip udp-helper udp-port 1000
Dell(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.

```
Dell#show ip udp-helper
--------------------------
Port          UDP port list
--------------------------
TenGig 1/1    1000
```

Configurations Using UDP Helper

When you enable UDP helper and the destination IP address of an incoming packet is a broadcast address, the system 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, the system 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.

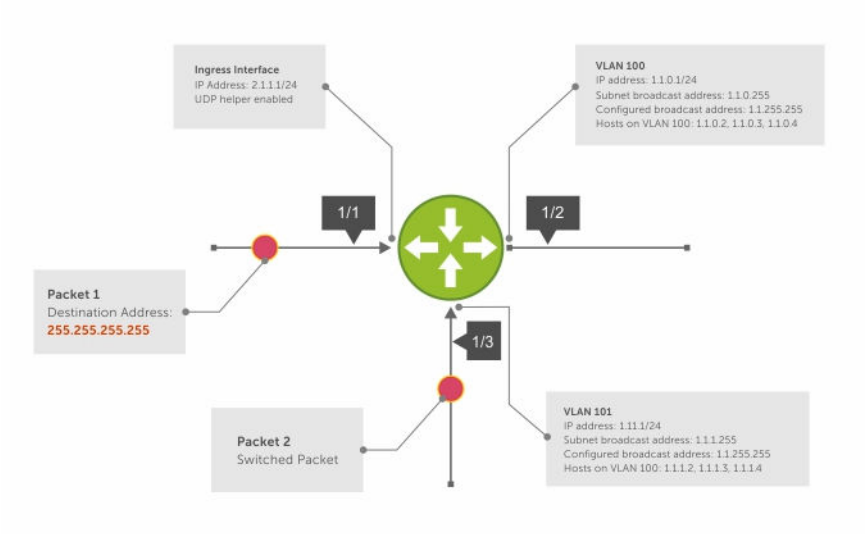


Figure 52. UDP Helper with Broadcast-All Addresses

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.

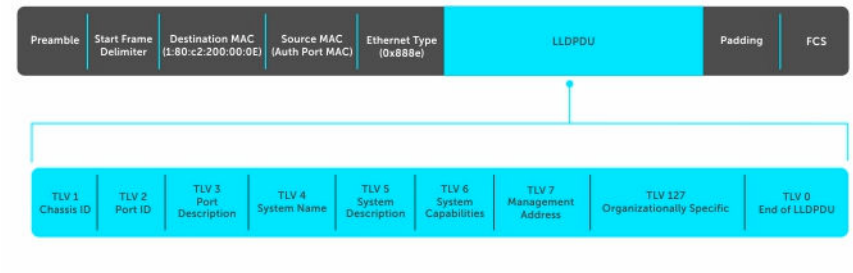


Figure 53. UDP Helper with Subnet Broadcast Addresses

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.

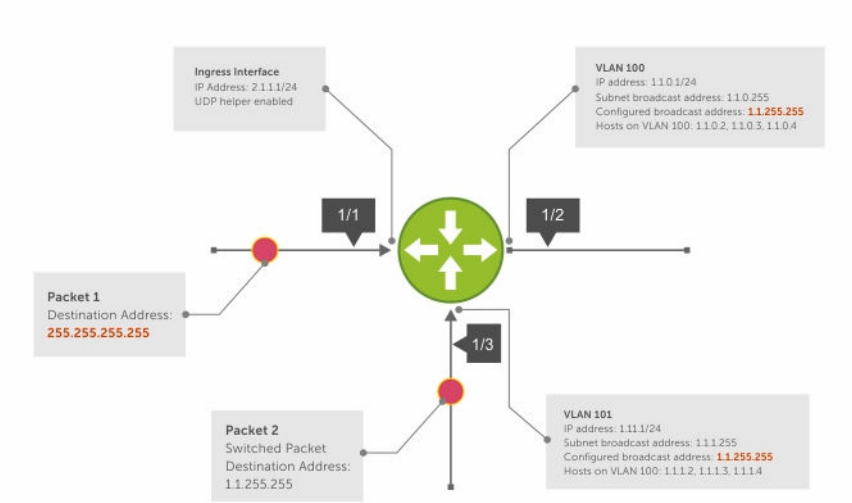


Figure 54. UDP Helper with Configured Broadcast Addresses

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

```
Dell(conf)# debug ip udp-helper
01:20:22: Pkt rcvd on TenGig 5/0 with IP DA (0xffffffff) will be sent on TenGig
5/1 TenGig 5/
2 Vlan 3
01:44:54: Pkt rcvd on TenGig 7/0 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

2005-11-05 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

2005-11-05 11:59:35 %RELAY-I-BOOTREQUEST, Forwarded BOOTREQUEST for 00:02:2D:8D:
46:DC to
137.138.17.6

2005-11-05 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

2005-07-05 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
```

26

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 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 Networking Operating System (OS) version supporting which features and platforms, refer to Implementing IPv6 with Dell Networking OS.

NOTE: Even though Dell 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.

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: The Dell 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.

The Dell 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 55. 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.

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 the Dell Networking OS

The Dell Networking OS supports both IPv4 and IPv6 and both may be used simultaneously in your system.

The following table lists the Dell 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 26. Feature Details

Feature and Functionality	Dell Networking OS Release Introduction	Documentation and Chapter Location
	FN IOM	
Basic IPv6 Commands	9.9(0.0)	IPv6 Basic Commands in the *Dell Networking OS Command Line Interface Reference Guide*.
IPv6 Basic Addressing		
IPv6 address types: Unicast	9.9(0.0)	Extended Address Space in this chapter
IPv6 neighbor discovery	9.9(0.0)	IPv6 Neighbor Discovery in this chapter
IPv6 stateless autoconfiguration	9.9(0.0)	Stateless Autoconfiguration in this chapter
IPv6 MTU path discovery	9.9(0.0)	Path MTU Discovery in this chapter
IPv6 ICMPv6	9.9(0.0)	ICMPv6 in this chapter
IPv6 ping	9.9(0.0)	ICMPv6 in this chapter
IPv6 traceroute	9.9(0.0)	ICMPv6 in this chapter
IPv6 Routing		
Static routing	9.9(0.0)	Assigning a Static IPv6 Route in this chapter
Route redistribution	9.9(0.0)	OSPF, IS-IS, and IPv6 BGP chapters in the *Dell Networking OS Command Line Reference Guide*.
Multiprotocol BGP extensions for IPv6	9.9(0.0)	IPv6 BGP in the *Dell Networking OS Command Line Reference Guide*.
IPv6 BGP MD5 Authentication	9.9(0.0)	IPv6 BGP in the *Dell Networking OS Command Line Reference Guide*.

Feature and Functionality	Dell Networking OS Release Introduction	Documentation and Chapter Location
IS-IS for IPv6	9.9(0.0)	Intermediate System to Intermediate System (IS-IS) IPv6 IS-IS in the *Dell Networking OS Command Line Reference Guide*.
IS-IS for IPv6 support for redistribution	9.9(0.0)	Intermediate System to Intermediate System (IS-IS) IPv6 IS-IS in the *Dell Networking OS Command Line Reference Guide*.
ISIS for IPv6 support for distribute lists and administrative distance	9.9(0.0)	Intermediate System to Intermediate System (IS-IS) IPv6 IS-IS in the *Dell Networking OS Command Line Reference Guide*.
OSPF for IPv6 (OSPFv3)	9.9(0.0)	OSPFv3 in the *Dell Networking OS Command Line Reference Guide*.
Equal Cost Multipath for IPv6	9.9(0.0)	
IPv6 Services and Management	9.9(0.0)	
Telnet client over IPv6 (outbound Telnet)	9.9(0.0)	Configuring Telnet with IPv6 in this chapter Control and Monitoring in the *Dell Networking OS Command Line Reference Guide*.
Telnet server over IPv6 (inbound Telnet)	9.9(0.0)	Configuring Telnet with IPv6 in this chapter Control and Monitoring in the *Dell Networking OS Command Line Reference Guide*.
Secure Shell (SSH) client support over IPv6 (outbound SSH) Layer 3 only	9.9(0.0)	Secure Shell (SSH) Over an IPv6 Transport in this chapter
Secure Shell (SSH) server support over IPv6 (inbound SSH) Layer 3 only	9.9(0.0)	Secure Shell (SSH) Over an IPv6 Transport in this chapter
IPv6 Access Control Lists	9.9(0.0)	IPv6 Access Control Lists in the *Dell Networking OS Command Line Reference Guide*.
IPv6 Multicast		

Feature and Functionality	Dell Networking OS Release Introduction	Documentation and Chapter Location
PIM-SM for IPv6	N/A	IPv6 Multicast in this chapter IPv6 PIM in the *Dell Networking OS Command Line Reference Guide*.
PIM-SSM for IPv6	N/A	IPv6 Multicast in this chapter IPv6 PIM in the *Dell Networking OS Command Line Reference Guide*.
MLDv1/v2	N/A	IPv6 Multicast in this chapter Multicast IPv6 in the *Dell Networking OS Command Line Reference Guide*.
MLDv1 Snooping	N/A	IPv6 Multicast in this chapter Multicast IPv6 in the *Dell Networking OS Command Line Reference Guide*.
MLDv2 Snooping	N/A	IPv6 Multicast in this chapter Multicast IPv6 in the *Dell Networking OS Command Line Reference Guide*.
IPv6 QoS		
trust DSCP values	N/A	IPv6 Multicast in this chapter

ICMPv6

Dell Networking OS supports ICMPv6.

ICMP for IPv6 combines the roles of ICMP, IGMP and ARP in IPv4. Similar to IPv4, it provides functions for reporting delivery and forwarding errors, and provides a simple echo service for troubleshooting. The Dell 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 `ping` and `traceroute` commands extend to support IPv6 addresses. These commands use ICMPv6 Type-2 messages.

Path MTU Discovery

Dell Networking OS supports IPv6 path maximum transmission unit (MTU) discovery.

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 processing efficiency because each packet carries more data while protocol overheads (for example, headers) or underlying per-packet delays remain fixed.

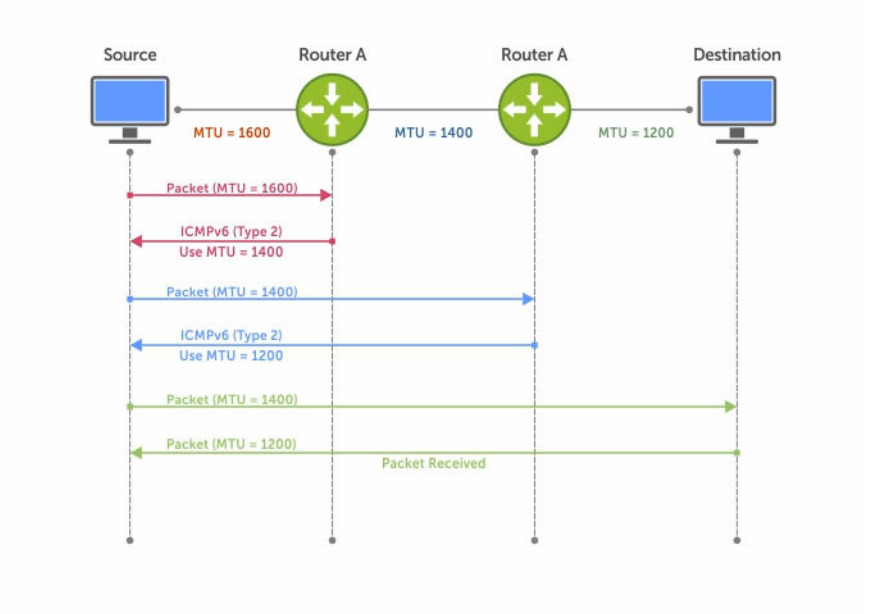


Figure 56. Path MTU Discovery Process

IPv6 Neighbor Discovery

Dell Networking OS supports IPv6 neighbor discovery protocol (NDP).

NDP is a top-level protocol for neighbor discovery on an IPv6 network. In lieu 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 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.

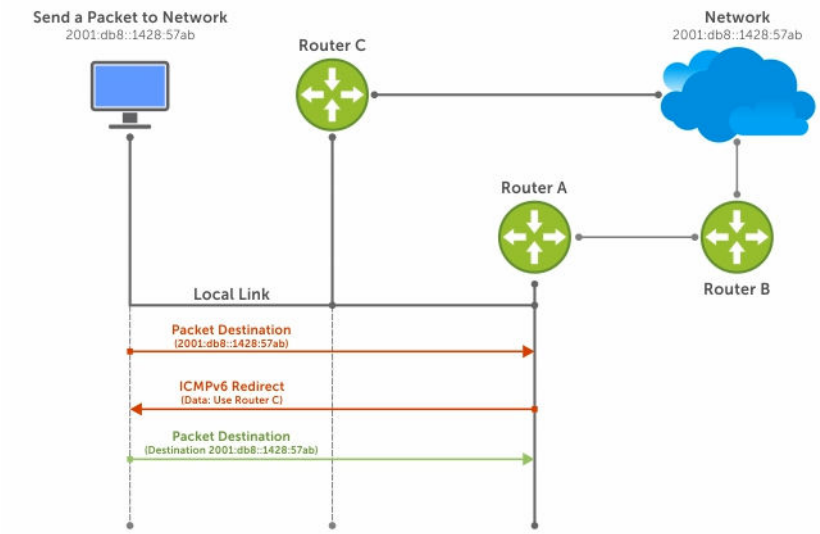


Figure 57. NDP Router Redirect

IPv6 Neighbor Discovery of MTU Packets

With the Dell Networking OS version 8.3.1.0, 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.

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`.

To display IPv6 RDNSS information, use the `show configuration` command in INTERFACE CONFIG mode.

The following example uses the `show configuration` command to display IPv6 RDNSS information.

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.

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.

IPv6 Multicast

Dell Networking OS supports IPv6 multicast.

The Dell Networking OS supports the following protocols to implement IPv6 multicast routing:

- **Multicast listener discovery protocol (MLD)** — MLD on a multicast router sends out periodic general MLD queries that the switch forwards through all ports in the VLAN. There are two versions of MLD: MLD version 1 is based on version 2 of the Internet group management protocol (IGMP) for IPv4; MLD version 2 is based on version 3 of the IGMP for IPv4. IPv6 multicast for the Dell Networking OS supports versions 1 and 2.
- **Protocol-independent multicast-sparse mode (PIM-SM)** — PIM-SM is a multicast protocol in which multicast receivers explicitly join to receive multicast traffic. The protocol uses a router as the root or rendezvous point (RP) of the share tree distribution tree to distribute multicast traffic to a multicast group. Messages to join the multicast group (Join messages) are sent towards the RP and data is sent from senders to the RP so receivers can discover who are the senders and begin receiving traffic destined to the multicast group.

For more information, refer to the *Neighbor Discovery Protocol (NDP)*, *Multicast IPv6*, and *Protocol Independent Multicast (IPv6)* chapters in the *Dell Networking OS Command Line Interface Reference Guide*.

Configuration Task List 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

Dell Networking OS supports the `cam-acl` command.
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

Dell Networking OS supports IPv6 addresses.
Essentially, IPv6 is enabled in the Dell 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.
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

Dell Networking OS supports IPv6 static routes.
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` *`prefix type {slot/port} forwarding router tag`*

  - *`prefix`*: IPv6 route prefix
  - *`type {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 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/ port information.
  - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/ port information.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a Loopback interface, enter the keyword `loopback` then the loopback number.
  - For a port-channel interface, enter the keywords `port-channel` then the port-channel number.
  - For a VLAN interface, enter the keyword `vlan` then the VLAN ID.
  - For a Null interface, enter the keyword `null` then the Null interface number.

Configuring Telnet with IPv6

Dell Networking OS supports IPv6 telnet.
The Telnet client and server in the Dell 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: Dell Networking OS supports Telnet to link local addresses.

- Enter the IPv6 Address for the device.
  EXEC mode or EXEC Privileged mode

  `telnet` *`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

Dell Networking OS supports the simple network management protocol (SNMP).

You can configure SNMP over IPv6 transport so that an IPv6 host can perform SNMP queries and receive SNMP notifications from a device running IPv6. The Dell 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 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`

Showing IPv6 Information

Dell Networking OS supports all of the following `show` commands.
View specific IPv6 configuration with the following commands.

- List the IPv6 show options.
  EXEC mode or EXEC Privileged mode

  ```
  show ipv6 ?
  ```

Example of `show ipv6` Command Options

```
Dell#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
Dell#
```

Showing an IPv6 Interface

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 type {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 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/ port information.
- For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/ port information.
- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
- For a loopback interface, enter the keyword `loopback` then the loopback number.
- For a port-channel interface, enter the keywords `port-channel` then the port-channel number.
- For a VLAN interface, enter the keyword `vlan` then the VLAN ID.

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` *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.

Example of the `show ipv6 route summary` Command

Example of the `show ipv6 route` Command

Example of the `show ipv6 route static` Command

```
Dell#show ipv6 route summary

Route Source Active Routes Non-active Routes
connected    5             0
static       0             0
Total        5             0

Dell#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   2001::/64 [0/0]
     Direct, Gi 1/1, 00:28:49
C   2002::/120 [0/0]
     Direct, Gi 1/1, 00:28:49
C   2003::/120 [0/0]
     Direct, Gi 1/1, 00:28:49

Dell#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, Gi 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 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` then the slot/ port information.
  - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/ port information.
  - For the Management interface on the RPM, enter the keyword `ManagementEthernet` then the slot/port information.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

Example of the `show running-config interface` Command

```
Dell#show run int gi 2/2
!
interface GigabitEthernet 2/2
  no ip address
  ipv6 address 3:4:5:6::8/24
  shutdown
Dell#
```

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 {* |` *ipv6 address prefix-length*}
  - `*`: 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.

Secure Shell (SSH) Over an IPv6 Transport

Dell Networking OS supports IPv6 secure shell (SSH).

The Dell 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 Networking OS Command Line Interface Reference Guide*.

27

iSCSI Optimization

An Aggregator enables internet small computer system interface (iSCSI) optimization with default iSCSI parameter settings(Default iSCSI Optimization Values) and is auto-provisioned to support:

iSCSI Optimization: Operation

To display information on iSCSI configuration and sessions, use `show` commands.

iSCSI optimization enables quality-of-service (QoS) treatment for iSCSI traffic.

Supported Modes

Stacking, VLT, Standalone, PMUX, and Full–Switch

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 self-configuration of several key network configurations that enables optimization of the network for better storage traffic throughput.

iSCSI optimization provides a means of monitoring iSCSI sessions and applying 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 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 VLT peers.

iSCSI optimization functions as follows:

- Manual configuration to detect Compellent storage arrays where auto-detection is not supported.
- Jumbo frames — Ports are set to a maximum transmission unit (MTU) of 12,000 bytes.
- If you configured flow-control, iSCSI uses the current configuration. If you did not configure flow-control, iSCSI auto-configures flow control.
- 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.

The following figure shows iSCSI optimization between servers in a server enclosure and a storage array in which an Aggregator connects installed servers (iSCSI initiators) to a storage array (iSCSI targets) in a SAN network. iSCSI optimization running on the Aggregator is configured to use dot1p priority-queue assignments to ensure that iSCSI traffic in these sessions receives priority treatment when forwarded on Aggregator hardware.

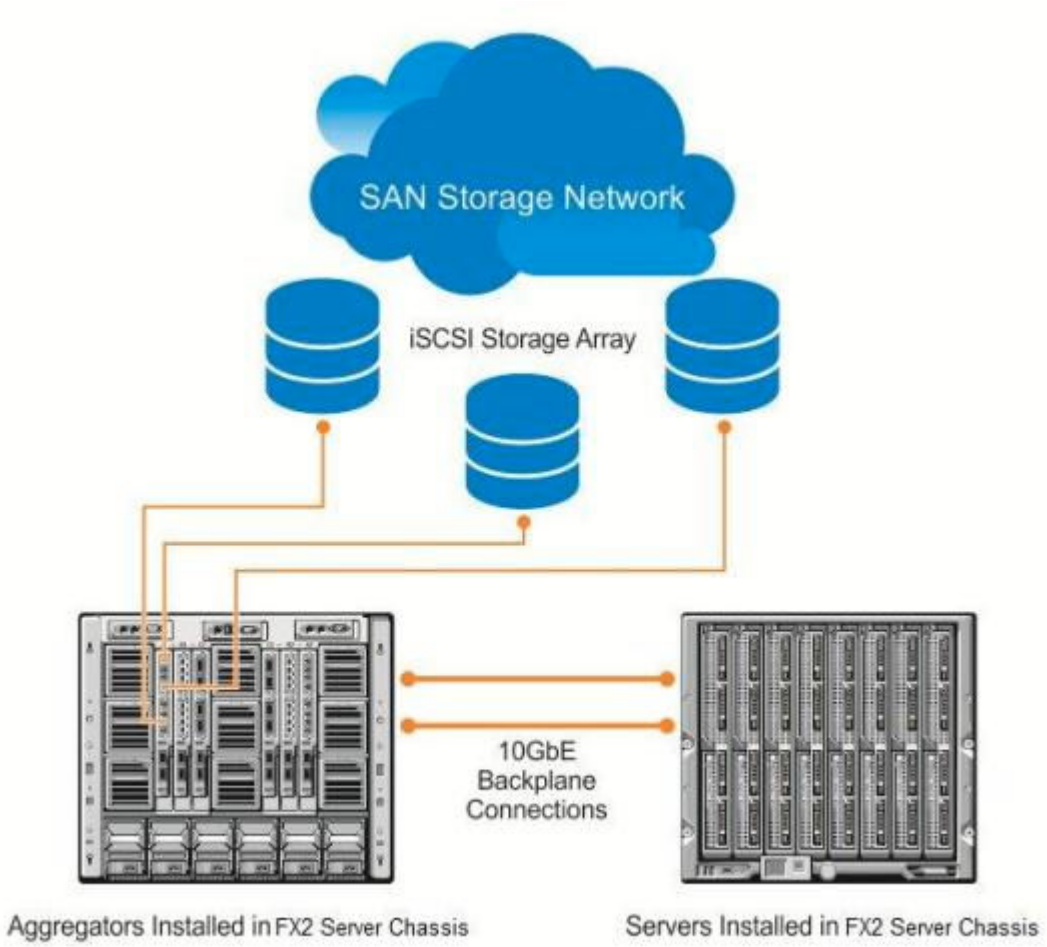


Figure 58. iSCSI Optimization Example

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.

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

If no iSCSI traffic is detected for a session during a user-configurable aging period, the session data clears.

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.

iSCSI Optimization: Operation

When the Aggregator auto-configures with iSCSI enabled, the following occurs:

- Link-level flow control is enabled on PFC disabled interfaces.
- iSCSI session snooping is enabled.

- iSCSI LLDP monitoring starts to automatically detect EqualLogic arrays.

iSCSI optimization requires LLDP to be enabled. LLDP is enabled by default when an Aggregator auto-configures.

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.
```

Configuring iSCSI Optimization

To configure iSCSI optimization, use the following commands.

1. **For a non-DCB environment:** Enable iSCSI.
   CONFIGURATION mode

   ```
   iscsi enable
   ```
2. (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.
3. (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.

4. (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**.

5. (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**.

6. (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).

7. (Optional) Enter interface configuration mode to configure the auto-detection of Dell Compellent disk arrays.

   CONFIGURATION mode

   ```
   interface port-type slot/port
   ```

8. (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.

   NOTE: All these configurations are available only in PMUX mode.

Displaying iSCSI Optimization Information

To display information on iSCSI optimization, use the show commands detailed in the below table:

Table 27. Displaying iSCSI Optimization Information

Command	Output
show iscsi	Displays the currently configured iSCSI settings.
show iscsi sessions	Displays information on active iSCSI sessions on the switch that have been established since the last reload.
show iscsi sessions detailed [session *isid*]	Displays detailed information on active iSCSI sessions on the switch. To display detailed information on specified iSCSi session, enter the session's iSCSi ID.
show run iscsi	Displays all globally-configured non-default iSCSI settings in the current Dell Networking OS session.

show iscsi Command Example

```
Dell# show iscsi
iSCSI is enabled
iSCSI session monitoring is enabled
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
```

show iscsi sessions Command Example

```
Dell# show iscsi sessions
Session 0:
-------------------------------------------------------------------------------
---------
Target: iqn.2001-05.com.equallogic:0-8a0906-0e70c2002-10a0018426a48c94-iom010
Initiator: iqn.1991-05.com.microsoft:win-x918v27yajg
ISID: 400001370000
Session 1:
-------------------------------------------------------------------------------
---------
Target: iqn.2001-05.com.equallogic:0-8a0906-0f60c2002-0360018428d48c94-iom011
Initiator: iqn.1991-05.com.microsoft:win-x918v27yajg
ISID: 400001370000.
```

show iscsi sessions detailed Command Example

```
Dell# show iscsi sessions 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
Session 1 :
-----------------------------------------------------------------------------
Target:iqn.2010-11.com.ixia:ixload:iscsi-TG1
Initiator:iqn.2010-11.com.ixia.ixload:initiator-iscsi-35
Up Time:00:00:01:22(DD:HH:MM:SS)
Time for aging out:00:00:09:31(DD:HH:MM:SS)
ISID:806978696102
Initiator      Initiator    Target        Target   Connection ID
IP Address     TCP Port     IP Address    TCPPort
10.10.0.53     33432        10.10.0.101   3260      0
```

28

Intermediate System to Intermediate System

Dell Networking OS supports intermediate system to intermediate system (IS-IS).

- The IS-IS protocol is an interior gateway protocol (IGP) that uses a shortest-path-first algorithm. Dell Networking supports both IPv4 and IPv6 versions of IS-IS.
- The IS-IS protocol standards are listed in the Standards Compliance chapter.

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 Networking implementation supports only IP traffic.

IS-IS uses the following management information base (MIB): `draft-ietf-isis-wg-mib-16 and f10-isis`.

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.

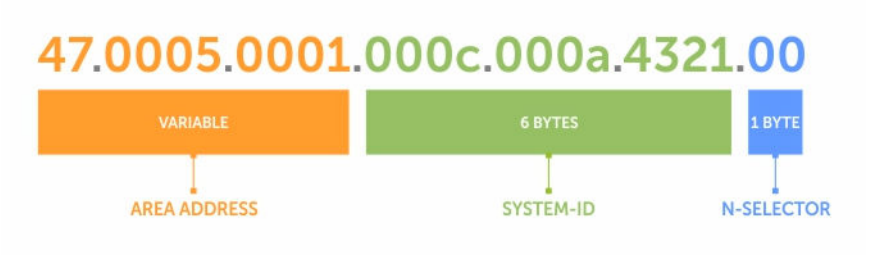


Figure 59. ISO Address Format

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

Dell Networking OS supports Graceful Restart for both Helper and Restart modes.

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 the fact 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) and are still resident. 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, the system 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. The system does not support ISO CLNS routing; however, the ISO NET format is supported for addressing.

To support IPv6, the Dell 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 28. 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
- Configuring Multi-Topology IS-IS (MT IS-IS)
- Configuring IS-IS Graceful Restart
- Changing LSP Attributes
- Configuring the IS-IS Metric Style
- Configuring the 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 the Loopback interface on the RPM, enter the keyword `loopback` then a number from 0 to 16383.
   - For a port channel, enter the keywords `port-channel` then a number from 1 to 255.
   - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
   - For a VLAN, 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.

Example of Viewing IS-IS Configuration (EXEC Privilege Mode)

```
Dell#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
  GigabitEthernet 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
Dell#
```

To view IS-IS protocol statistics, use the `show isis traffic` command in EXEC Privilege mode.

Example of the `show isis traffic` Command

```
Dell#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
Dell#
```

You can assign more NET addresses, but the System ID portion of the NET address must remain the same. The Dell 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 20 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**.

Example of the `show isis graceful-restart detail` Command

Example of the `show isis interface` Command

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.

```
Dell#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 GigabitEthernet 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
Dell#
```

To view all interfaces configured with IS-IS routing along with the defaults, use the `show isis interface` command in EXEC Privilege mode.

```
Dell#show isis interface G1/34
GigabitEthernet 2/10 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
Dell#
```

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**.

Example of Viewing IS-IS Configuration (ROUTER ISIS Mode)

To view the configuration, use the `show config` command in ROUTER ISIS mode or the `show running-config isis` command in EXEC Privilege mode.

```
Dell#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
Dell#
```

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. The system supports five different metric styles: narrow, wide, transition, narrow transition, and wide transition.

By default, the system 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.

The Dell Networking OS supports the following IS-IS metric styles.

Table 29. 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

```
Dell#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
    GigabitEthernet 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
Dell#
```

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.

  The default is **10**.
- 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}
  ```

Example of the `show isis database` Command to View Level 1-2 Link State Databases

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, so 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.

```
Dell#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
Dell.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
Dell.00-00     0x00000004    0xCDA9       1107         0/0/0

Dell#
```

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 the Loopback interface on the RPM, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel, enter the keywords `port-channel` then a number from 1 to 255.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/ port information.
  - For a VLAN, 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 the Dell 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 slot/port information:
  - For the Loopback interface on the RPM, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel, enter the keywords `port-channel` then a number from 1 to 255.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.

  - For a VLAN, 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 slot/port information:
  - For the Loopback interface on the RPM, enter the keyword `loopback` then a number from 0 to 16383.
  - For a port channel, enter the keywords `port-channel` then a number from 1 to 255.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a VLAN, 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**.
  - `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 Networking 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 Networking 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, the system 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
  ```

Example of Viewing the Overload Bit Setting

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.

```
Dell#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
Dell.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
Dell.00-00     0x00000004   0xCDA9       1093         0/0/0
Dell#
```

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.

The system 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

The Dell 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 30. 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
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.

Beginning Metric Style	Final Metric Style	Resulting IS-IS Metric Value
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 31. 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 32. 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
narrow	transition	original value
wide	narrow	truncated value
wide	narrow transition	truncated value
wide	wide transition	original value
wide	transition	truncated value

Level-1 Metric Style	Level-2 Metric Style	Resulting Metric 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:

```
Dell#clear isis *
% ISIS not enabled.
Dell#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.

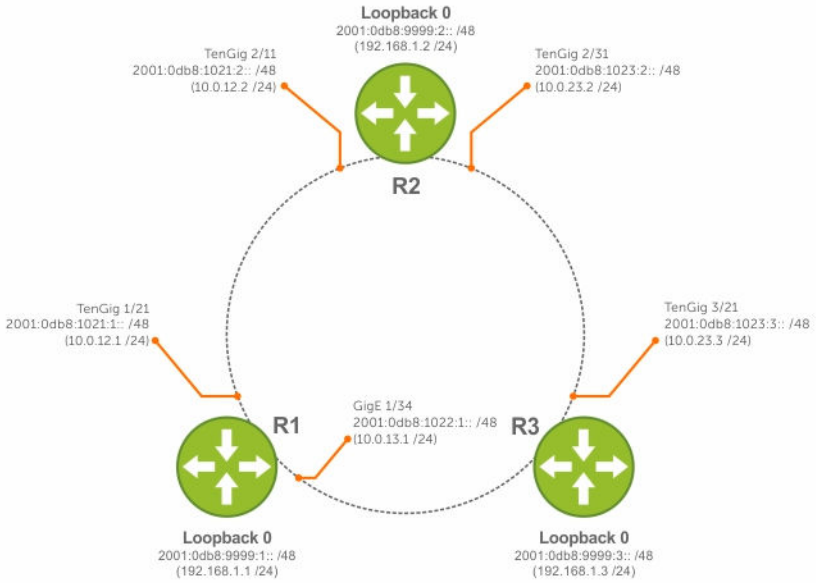


Figure 60. IPv6 IS-IS Sample Topography

IS-IS Sample Configuration — Router 1

IS-IS Sample Configuration — Router 2

IS-IS Sample Configuration — Router 3

The following is a sample configuration for enabling IPv6 IS-IS.

```
R1(conf)#interface Loopback 0
R1(conf-if-lo-0)#ip address 192.168.1.1/24
R1(conf-if-lo-0)#ipv6 address 2001:db8:9999:1::/48

R1(conf-if-lo-0)#ip router isis 9999
R1(conf-if-lo-0)#no shutdown
R1(conf-if-lo-0)#router isis 9999
R1(conf-router_isis)#is-type level-1
R1(conf-router_isis)#net FF.F101.0002.0C00.1111.00
R1(conf-router_isis)#ipv6 route 2001:db8:9999:2::/128 2001:db8:1021:2::
R1(conf)#ipv6 route 2001:db8:9999:3::/128 2001:db8:1022:3::
R1(conf)#ip route 192.168.1.2/32 10.0.12.2
R1(conf)#ip route 192.168.1.3/32 10.0.13.3

R1(conf)#interface GigabitEthernet 1/21
R1(conf-if-gi-1/21)#ip address 10.0.12.1/24
R1(conf-if-gi-1/21)#ipv6 address 2001:db8:1022:1::/48
R1(conf-if-gi-1/21)#isis circuit-type level-1
R1(conf-if-gi-1/21)#isis network point-to-point
R1(conf-if-gi-1/21)#ip router isis 9999
R1(conf-if-gi-1/21)#no shutdown

R1(conf-if-gi-1/21)#interface GigabitEthernet 1/34
R1(conf-if-gi-1/34)# ip address 10.0.13.1/24
R1(conf-if-gi-1/34)#ipv6 address 2001:db8:1021:1::/48
R1(conf-if-gi-1/34)#ip router isis 9999
R1(conf-if-gi-1/34)#no shutdown
R1(conf-if-gi-1/34)#end
```

```
R1#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 10.0.12.0/24   Direct, Gi 1/21        0/0   00:00:57
C 192.168.1.0/24 Direct, Lo 0           0/0   00:04:19
S 192.168.1.2/32 via 10.0.12.2, Gi 1/21 1/0   00:00:57

R1#show isis data
IS-IS Level-1 Link State Database
LSPID LSP Seq Num       LSP Checksum LSP Holdtime ATT/P/OL
R1.00-00  * 0x0000000F 0x3A6C   1176              0/0/0
R1.02-00  * 0x00000002 0x90AC   1076              0/0/0
R1.03-00  * 0x00000002 0x67C3   1176              0/0/0
R2.00-00  0x0000000C   0x5418   1183              0/0/0
R2.00-00  0x00000009   0x1E39   1183              0/0/0
R2.03-00  0x00000002   0x589D   1180              0/0/0

R1#show isis neigh
System Id Interface State Type Priority Uptime   Circuit Id
R2        Gi 1/21   Up    L1   64       00:02:28 R1.02
R2        Gi 1/34   Up    L1   64       00:00:42 R1.03
R1#

R2(conf)#interface Loopback 0
R2(conf-if-lo-0)#ip address 192.168.1.1/24
R2(conf-if-lo-0)#ipv6 address 2001:db8:9999:1::/48
R2(conf-if-lo-0)#ip router isis 9999
R2(conf-if-lo-0)#no shutdown

R2(conf-if-lo-0)#router isis 9999
R2(conf-router_isis)#int gi 2/11
R2(conf-if-gi-2/11)#ip address 10.0.12.2/24
R2(conf-if-gi-2/11)#ipv6 address 2001:db8:9999:2::/48
R2(conf-if-gi-2/11)#ip router isis 9999
R2(conf-if-gi-2/11)#isis network point-to-point
R2(conf-if-gi-2/11)#no shutdown
R2(conf-if-gi-2/11)#int gi 2/31
R2(conf-if-gi-2/31)#ip address 10.0.23.2/24
R2(conf-if-gi-2/31)#ipv6 address 2001:db8:1021:2::/48
R2(conf-if-gi-2/31)#ip router isis 9999
R2(conf-if-gi-2/31)#isis network point-to-point
R2(conf-if-gi-2/31)#no shutdown
R2(conf-if-gi-2/31)#router isis 9999

R2(conf-router_isis)#ipv6 route 2001:db8:9999:1::/128 2001:db8:1021:1::
R2(conf)#ipv6 route 2001:db8:9999:3::/128 2001:db8:1023:3::
R2(conf)#ip route 192.168.1.1/32 10.0.12.1
R2(conf)#ip route 192.168.1.3/32 10.0.23.3
R2(conf)#ex

R2#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 172.21.212.1 to network 0.0.0.0

   Destination     Gateway              Dist/Metric Last Change
-----------        -------              -----------   -----------
*S 0.0.0.0/0       via 172.21.212.1, Vl 212 1/0    00:59:50
C  10.0.12.0/24    Direct, Gi 2/11          0/0    00:02:25
C  10.0.23.0/24    Direct, Gi 2/31          0/0    00:01:53
C  10.10.92.0/24   Direct, Po 4             0/0    6d9h
C  172.21.212.0/24 Direct, Vl 212           0/0    2d20h
C  192.168.1.0/24  Direct, Lo 0             0/0    01:11:48
S  192.168.1.1/32  via 10.0.12.1, Gi 2/11   1/0    00:00:51
S  192.168.1.3/32  via 10.0.23.3, Gi 2/31   1/0    00:00:39

R2#show isis data
IS-IS Level-1 Link State Database
LSPID    LSP Seq Num   LSP Checksum LSP Holdtime ATT/P/OL
R2.00-00 * 0x0000000F  0x0174       1088         0/0/0

R2#show isis neigh
A2#show isis neigh
System Id Interface State Type Priority Uptime    Circuit Id
R1        Gi 2/11 Up      L1   64       00:02:19  A102.02
R3        Gi 2/31 Up      L1   64       00:00:25  A121.03
R2#

R3(conf)#interface Loopback 0
R3(conf-if-lo-0)#ip address 192.168.1.3/24
R3(conf-if-lo-0)#ipv6 address 2001:db8:9999:3::/48
R3(conf-if-lo-0)#ip router isis 9999
R3(conf-if-lo-0)#no shutdown
R3(conf-if-lo-0)#router isis 9999
R3(conf-router_isis)#net FF.F101.0002.0C00.1133.00
R3(conf-router_isis)#ipv6 route 2001:db8:9999:1::/128 2001:db8:1022:1::
R3(conf)#ipv6 route 2001:db8:9999:2::/128 2001:db8:1023:2::
R3(conf)#ip route 192.168.1.1/32 10.0.13.1
R
3(conf)#interface GigabitEthernet 3/14
R3(conf-if-gi-3/14)#ip address 10.0.13.3/24
R3(conf-if-gi-3/14)#ipv6 address 2001:db8:1022:3::/48
R3(conf-if-gi-3/14)#ip router isis 9999
R3(conf-if-gi-3/14)#isis circuit-type level-1
R3(conf-if-gi-3/14)#isis network point-to-point
R3(conf-if-gi-3/14)#no shutdown
R3(conf-if-gi-3/14)#interface GigabitEthernet 3/21
R3(conf-if-gi-3/21)#ip address 10.0.23.3/24
R3(conf-if-gi-3/21)#ipv6 address 2001:db8:1023:3::/48
R3(conf-if-gi-3/21)#ip router isis 9999
R3(conf-if-gi-3/21)#isis circuit-type level-1
R3(conf-if-gi-3/21)#isis network point-to-point
R3(conf-if-gi-3/21)#no shutdown
R3(conf-if-gi-3/21)#end

R3#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 10.0.13.0/24   Direct, Gi 3/14         0/0    00:00:10
C 10.0.23.0/24   Direct, Gi 3/21         0/0    00:00:03
C 192.168.1.0/24 Direct, Lo 0            0/0    00:00:32
S 192.168.1.1/32 via 10.0.13.1, Gi 3/14  1/0    00:00:10
S 192.168.1.2/32 via 10.0.23.2, Gi 3/21  1/0    00:00:03

R2#show isis data
IS-IS Level-1 Link State Database
LSPID      LSP Seq Num   LSP Checksum LSP Holdtime ATT/P/OL
R1.00-00   0x0000000F    0x3A6C      1198         0/0/0
R1.03-00   0x00000001    0x69C2      1193         0/0/0
R2.00-00   * 0x00000007  0x51F6      1198         0/0/0
R2.03-00   * 0x00000001  0x5A9C      1200         0/0/0

IS-IS Level-2 Link State Database
LSPID    LSP Seq Num    LSP Checksum LSP Holdtime ATT/P/OL
R3.00-00 * 0x00000008   0xC09C       1199         0/0/0
R3#show isis neigh
System Id Interface State Type Priority Uptime  Circuit Id
R1        Gi 3/14   Init  L1   64       00:00:02  R1.03
R2        Gi 3/21   Up    L1   64       00:00:14  A101.03
```

29

Isolated Networks for Aggregators

An Isolated Network is an environment in which servers can only communicate with the uplink interfaces and not with each other even though they are part of same VLAN. If the servers in the same chassis need to communicate with each other, it requires a non-isolated network connectivity between them or it needs to be routed in the TOR.
Isolated Networks can be enabled on per VLAN basis. If a VLAN is set to be isolated, all the packets of originating from the server ports for that VLAN (Isolated Network) will be redirected to uplink LAG, including the packets destined for the server ports on the same blade.

ToR applies required ACLs and other necessary actions before sending the packet to destination. If the packet is destined to server on the same IOA blade, it is routed back on the uplink lag where it was received. Traffic that hits at the uplink ports are regularly switched based on the L2 MAC lookup. Unknown Unicast and Multicast packets from Uplink Port towards server port on an isolated network enabled VLAN, is dropped.

The isolated network feature is supported only in the standalone mode.

Isolated network is currently not supported in the following modes:

- VLT mode
- Stacking mode
- PMUX mode

NOTE: Isolated Networks is not enabled for FCOE VLANs and on default VLAN. It can be managed via CLI or AFM. For more information, refer to AFM user manual.

Configuring and Verifying Isolated Network Settings

Enable the isolated-network functionality for a particular VLAN or a set of VLANs using below command:

```
Dell(conf)#io-aggregator isolated-network vlan <vlan-range>
```

To disable the isolated-network functionality, use the no form of command.

```
Dell(conf)#no io-aggregator isolated-network vlan <vlan-range>
```

To display the VLANs that are configured to be part of an isolated network on the Aggregator, use the below command.

```
Dell#show io-aggregator isolated-networks
Isolated Network Enabled VLANs : 5-10
```

30

Link Aggregation

Unlike IOA Automated modes (Standalone and VLT modes), the IOA Programmable MUX (PMUX) can support multiple uplink LAGs. You can provision multiple uplink LAGs. The I/O Aggregator auto-configures with link aggregation groups (LAGs) as follows:

- All uplink ports are automatically configured in a single port channel (LAG 128).
- Server-facing LAGs are automatically configured if you configure server for link aggregation control protocol (LACP)-based NIC teaming (Network Interface Controller (NIC) Teaming).

No manual configuration is required to configure Aggregator ports in the uplink or a server-facing LAG.

NOTE: Static LAGs are not supported on the SMUX Aggregator.

NOTE: In order to avoid loops, only disjoint VLANs are allowed between the uplink ports/uplink LAGs and uplink-to-uplink switching is disabled.

Supported Modes

Standalone, VLT, PMUX, Stacking

How the LACP is Implemented on an Aggregator

The LACP provides a means for two systems (also called partner systems) to exchange information through dynamic negotiations to aggregate two or more ports with common physical characteristics to form a link aggregation group.

NOTE: A link aggregation group is referred to as a *port channel* by the Dell Networking OS.

A LAG provides both load-sharing and port redundancy across stack units. An Aggregator supports LACP for auto-configuring dynamic LAGs. Use CLI commands to display LACP information, clear port-channel counters, and debug LACP operation for auto-configured LAG on an Aggregator.

The Dell 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 you configure as LACP-capable.

NOTE: In Standalone, VLT, and Stacking modes, you can configure a maximum of 16 members in port-channel 128. In PMUX mode, you can have multiple port-channels with up to 16 members per channel.

Uplink LAG

When the Aggregator power is on, all uplink ports are configured in a single LAG (LAG 128).

Server-Facing LAGs

Server-facing ports are configured as individual ports by default. If you configure a server NIC in standalone, stacking, or VLT mode for LACP-based NIC teaming, server-facing ports are automatically configured as part of dynamic LAGs. The LAG range 1 to 127 is reserved for server-facing LAGs.

After the Aggregator receives LACPDU from server-facing ports, the information embedded in the LACPDU (remote-system ID and port key) is used to form a server-facing LAG. The LAG/port-channel number is assigned based on the first available number in the range from 1 to 127. For each unique remote system-id and port-key combination, a new LAG is formed and the port automatically becomes a member of the LAG.

All ports with the same combination of system ID and port key automatically become members of the same LAG. Ports are automatically removed from the LAG if the NIC teaming configuration on a server-facing port changes or if the port goes operationally down. Also, a server-facing LAG is removed when the last port member is removed from the LAG.

The benefit of supporting a dynamic LAG is that the Aggregator's server-facing ports can toggle between participating in the LAG or acting as individual ports based on the dynamic information exchanged with a server NIC. LACP supports the exchange of messages on a link to allow their LACP instances to:

- Reach agreement on the identity of the LAG to which the link belongs.
- Attach the link to that LAG.
- Enable the transmission and reception functions in an orderly manner.
- Detach the link from the LAG if one of the partner stops responding.

LACP Modes

The Aggregator supports only LACP active mode as the default mode of operation. In active mode, a port interface is considered to be not part of a LAG but rather in an active negotiating state.

A port in active mode automatically initiates negotiations with other ports by sending LACP packets. If you configure server-facing ports for LACP-based NIC teaming, LACP negotiations take place to aggregate the port in a dynamic LAG. If you do not configure server-facing ports for LACP-based NIC teaming, a port is treated as an individual port in active negotiating state.

Auto-Configured LACP Timeout

LACP PDUs are exchanged between port channel (LAG) interfaces to maintain LACP sessions. LACP PDUs are transmitted at a slow or fast transmission rate, depending on the LACP timeout value configured on the partner system.

The timeout value is the amount of time that a LAG interface waits for a PDU from the partner system before bringing the LACP session down. The default timeout is long-timeout (30 seconds) and is not user-configurable on the Aggregator.

Link Aggregation Control Protocol (LACP)

The commands for Dell Networks's implementation of the link aggregation control protocol (LACP) for creating dynamic link aggregation groups (LAGs) — known as port-channels in the Dell Networking OS — are provided in the following sections.

NOTE: For static LAG commands, refer to the chapter), 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.

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. In VLT mode, port channel configurations are allowed in the startup configuration.

These are the mandatory and optional configuration tasks:

- Reassigning an Interface to a New Port Channel (optional)
- Configuring VLAN Tags for Member Interfaces (optional)
- Deleting or Disabling a Port Channel (optional)

Creating a Port Channel

You can create up to 128 port channels with four port members per group on the Aggregator.
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 3 mode. To 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 Gigabit Ethernet and 10/100/1000 Ethernet interfaces, but Dell Networking OS disables the interfaces that are not the same speed of the first channel member in the port channel.

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`
- `ip mtu` (if the interface is on a Jumbo-enabled by default)

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.

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 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
   ```

   This command is applicable only in PMUX mode.

   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.

Example of the `show interfaces port-channel brief` Command

```
Dell#sh int port-channel brief
Codes: L - LACP Port-channel
       O - OpenFlow Controller Port-channel

    LAG  Mode  Status       Uptime      Ports
L   1    L2    up           00:00:19    Te 0/7     (Up)
                                        Te 0/8     (Up)
L   128  L2    up           00:00:36    Te 0/9     (Up)
                                        Te 0/10    (Up)
                                        Te 0/11    (Up)
Dell#
```

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.

Example of the `show interface port-channel` Command

```
Dell#show interface port-channel 1
Port-channel 1 is down, line protocol is down
Hardware address is 00:1e:c9:de:04:9c, Current address is 00:1e:c9:de:04:9c
```

```
Interface index is 1107492865
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 :001ec9de049c
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
Members in this channel:
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 02:57:05
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: 02:57:05

Dell#
```

When more than one interface is added to a Layer 2-port channel, Dell 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.

Example of Error Due to an Attempt to Configure an Interface that is Part of a Port Channel

```
Dell(conf)#int port-channel 128
Dell(conf-if-po-128)#show config
!
interface Port-channel 128
mtu 12000
portmode hybrid
switchport
fip-snooping port-mode fcf
no shutdown
link-bundle-monitor enable
Dell(conf-if-po-128)#
```

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 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
   ```

   This command is applicable only in PMUX mode.
3. Add the interface to the second port channel.
   INTERFACE PORT-CHANNEL mode

   ```
   channel-member interface
   ```

Example of Moving an Interface to a New Port Channel

The following example shows moving the TenGigabitEthernet 0/8 interface from port channel 4 to port channel 3.

```
Dell(conf-if-po-4)#show config
!
interface Port-channel 4
channel-member TenGigabitEthernet 0/8
no shutdown
Dell(conf-if-po-4)#no chann tengi 0/8
Dell(conf-if-po-4)#int port 3
Dell(conf-if-po-3)#channel tengi 0/8
Dell(conf-if-po-3)#sho conf
!
interface Port-channel 3
channel-member TenGigabitEthernet 0/8
shutdown
Dell(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

```
Dell#config t
Dell(conf)#int po 128
Dell(conf-if-po-128)#minimum-links 5
Dell(conf-if-po-128)#
```

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

   ```
   Dell(conf-if-te-0/2)#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

   ```
   Dell(conf-if-te-0/2)#switchport
   ```

   This `switchport` configuration is allowed only in PMUX mode. In all other modes, it is automatically configured.
3. Verify the manually configured VLAN membership (**show interfaces switchport** *interface* command).
   EXEC mode

   ```
   Dell(conf)# interface tengigabitethernet 0/1
   Dell(conf-if-te-0/1)#switchport
   Dell(conf-if-te-0/1)# vlan tagged 2-5,100,4010
   Dell#show interfaces switchport te 0/1

   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 0/1
   802.1QTagged: True
   Vlan membership:
   Q       Vlans
   T       2-5,100,4010

   Dell#
   ```

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.

Configuring Auto LAG

You can enable or disable auto LAG on the server-facing interfaces. By default, auto LAG is enabled. This functionality is supported on the Aggregator in Standalone, Stacking, and VLT modes.
To configure auto LAG, use the following commands:

1. Enable the auto LAG on all the server ports.
   CONFIGURATION mode

   ```
   io-aggregator auto-lag enable

   Dell(config)# io-aggregator auto-lag enable
   ```

   To disable the auto LAG on all the server ports, use the `no io-aggregator auto-lag enable` command. When disabled, all the server ports associated in a LAG are removed and the LAG itself gets removed. Any LACPDUs received on the server ports are discarded.

   In VLT mode, the global auto LAG is automatically synced to the peer VLT through ICL message.
2. Enable the auto LAG on a specific server port.
   Interface Configuration mode

   ```
   auto-lag enable

   Dell(conf-if-te-0/1)# auto-lag enable
   ```

   To disable the auto LAG, use the `no auto-lag enable` command. When disabled, the server port is removed from the LAG and if the server port is the last member of the LAG, the LAG itself gets removed. Any LACPDUs received on the server port are discarded.

   In VLT mode, the interface level auto LAG configuration is not synced to the peer. Only the global auto LAG is synced to the peer.
3. View the auto LAG configurations.
   EXEC mode

   ```
   show io-aggregator auto-lag status

   Dell# show io-aggregator auto-lag status

   Auto LAG creation on server port(s) is enabled
   ```

   For the interface level auto LAG configurations, use the `show interface` command.

   ```
   Dell(conf-if-te-0/1)#no auto-lag enable
   Dell(conf-if-te-0/1)#
   Dell(conf-if-te-0/1)#do show interface tengigabitethernet 0/1
   TenGigabitEthernet 0/1 is up, line protocol is down(error-disabled[UFD])
   Hardware is DellEth, address is f8:b1:56:07:1d:8e
       Current address is f8:b1:56:07:1d:8e
   Server Port AdminState is Up
   Pluggable media not present
   Interface index is 15274753
   ```

```
Internet address is not set
Mode of IPv4 Address Assignment : NONE
DHCP Client-ID :f8b156071d8e
MTU 12000 bytes, IP MTU 11982 bytes
LineSpeed auto
Auto-lag is disabled
Flowcontrol rx on tx off
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:12:53
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: 00:11:36

Dell(conf-if-te-0/1)#
```

Sample Configuration

```
Dell# config terminal
Dell(conf)# no io-aggregator auto-lag enable
Dell(conf)# end
Dell# show io-aggregator auto-lag status

Auto LAG creation on server port(s) is disabled

Dell#

Dell# config terminal
Dell(config)# interface tengigabitethernet 0/1
Dell(config-if-te-0/1)# no auto-lag enable
Dell(config-if-te-0/1)# show config
!
interface TenGigabitEthernet 0/1
mtu 12000
portmode hybrid
switchport
no auto-lag enable
!
protocol lldp
  advertise management-tlv management-address system-name
  dcbx port-role auto-downstream
 no shutdown
Dell#
```

Configuring the Minimum Number of Links to be Up for Uplink LAGs to be Active

You can activate the LAG bundle for uplink interfaces or ports (the uplink port-channel is LAG 128) on the I/O Aggregator only when a minimum number of member interfaces of the LAG bundle is up. For example, based on your network deployment, you may want the uplink LAG bundle to be activated only if a certain number of member interface links is also in the up state. If you enable this setting, the uplink LAG bundle is brought up only when the specified minimum number of links are up and the LAG bundle is moved to the down state when the number of active links in the LAG becomes less than the specified number of interfaces. By default, the uplink LAG 128 interface is activated when at least one member interface is up.
To configure the minimum number of member links that must be up for a LAG bundle to be fully up, perform the following steps:

Specify the minimum number of member interfaces of the uplink LAG 128 bundle that must be up for the LAG bundle to be brought up. The default minimum number of member links that must be active for the uplink LAG to be active is 1. Enter the `minimum-links number` command in the Port Channel Interface 128 Configuration mode to specify this value.

```
Dell(conf)#interface port-channel 128

Dell(conf-if-po-128)#minimum-links 4
```

Use the `show interfaces port-channel` command to view information regarding the configured LAG or port channel settings. The **Minimum number of links to bring Port-channel up is** field in the output of this command displays the configured number of active links for the LAG to be enabled.

```
Dell#show interfaces port-channel 128
Port-channel 128 is up, line protocol is down(minimum links not up)
Created by LACP protocol
Hardware address is 00:01:02:03:04:05, Current address is 00:01:02:03:04:05
Interface index is 1107492992
Minimum number of links to bring Port-channel up is 4
Internet address is not set
Mode of IPv4 Address Assignment : NONE
DHCP Client-ID :000102030405
MTU 12000 bytes, IP MTU 11982 bytes
LineSpeed auto
Members in this channel:
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 05:22:24
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: 05:22:28
```

Optimizing Traffic Disruption Over LAG Interfaces On IOA Switches in VLT Mode

When you use the `write memory` command while an Aggregator operates in VLT mode, the VLT LAG configurations are saved in nonvolatile storage (NVS).

By restoring the settings saved in NVS, the VLT ports come up quicker on the primary VLT node and traffic disruption is reduced. The delay in restoring the VLT LAG parameters is reduced (90 seconds by default) on the secondary VLT peer node before it becomes operational. This makes sure that the configuration settings of the primary VLT node are synchronized with the secondary VLT peer node before the secondary VLT mode is up. The traffic outage is less than 200 millisconds during the restart or switchover of the VLT peer nodes from primary to secondary.

Preserving LAG and Port Channel Settings in Nonvolatile Storage

Use the `write memory` command on an I/O Aggregator, which operates in either standalone or PMUX modes, to save the LAG port channel configuration parameters. This behavior enables the port channels to be brought up because the configured interface attributes are available in the system database during the booting of the device. With the reduction in time for the port channels to become active after the switch is booted, the loss in the number of packets that are serviced by these interfaces is minimized.

Enabling the LACP link fallback member

By default, the `lacp link-fallback member-independent port-channel 128` command is enabled on the Standalone mode. This command is not available in the VLT mode. To disable the `LACP link fallback member`, use the `no lacp link-fallback member-independent port-channel 128` command.

The following log message appears when LACP link fallback is enabled:

```
Feb 26 15:53:32: %STKUNIT0-M:CP IFMGR-5-NO_LACP_PDU_RECEIVED_FROM_PEER:
Connectivity to PEER is restricted because LACP PDU's are not received. Verify
the LACP configurations on PEER to bring up the Uplink LAG to attain better
bandwidth
```

The following log message appears when LACP link fallback is removed:

```
Feb 26 15:53:32: %STKUNIT0-M:CP %SMUX-5-SMUX_LACP_PDU_RECEIVED_FROM_PEER: LACP
PDU received from PEER and connectivity to PEER will be restored to Uplink Port-
channel 128.
```

Enabling the Verification of Member Links Utilization in a LAG Bundle

To examine the working efficiency of the LAG bundle interfaces, perform the following steps:

1. The functionality to detect the working efficiency of the LAG bundle interfaces is automatically activated on all the port channels, except the port channel that is configured as a VLT interconnect link, during the booting of the switch.
2. Use the `show link-bundle-distribution port-channel` *`interface-number`* command to display the traffic handling and utilization of the member interfaces of the port channel. The following sample output is displayed when you enter this `show` command.

   EXEC

```
Dell#show link-bundle-distribution port-channel
Dell#show link-bundle-distribution port-channel 1
Link-bundle trigger threshold - 60

LAG bundle - 1      Utilization[In Percent] - 0      Alarm State - Inactive

Interface                Line Protocol   Utilization[In Percent]
Te 0/5                   Up              0
Te 0/7                   Up              0
```

Monitoring the Member Links of a LAG Bundle

You can examine and view the operating efficiency and the traffic-handling capacity of member interfaces of a LAG or port channel bundle. This method of analyzing and tracking the number of packets processed by the member interfaces helps you manage and distribute the packets that are handled by the LAG bundle. The functionality to detect the working efficiency of the LAG bundle interfaces is automatically activated on all the port channels, except the port channel that is configured as a VLT interconnect link, during the booting of the switch. This functionality is supported on I/O Aggregators in PMUX, Standalone, and VLT modes. By default, this capability is enabled on all of the port channels set up on the switch.

You can use the `show link-bundle-distribution port-channel` *`interface-number`* command to display the traffic-handling and utilization of the member interfaces of the port channel. The following table describes the output fields of this show command.

Table 33. Output Field Descriptions for show link-bundle-distribution port-channel Command

Field	Description
Link-bundle trigger threshold	Threshold value that is the checkpoint, exceeding which the link bundle is marked as being overutilized and alarm is generated
LAG bundle number	Number of the LAG bundle
Utilization (In Percent)	Traffic usage in percentage of the packets processed by the port channel
Alarm State	Indicates whether an alarm is generated if over-utilization of the port channel occurred. The value, Active, is displayed for this field.

Field	Description
Interface	Slot and port number, and the type of the member interface of the port channel
Line Protocol	Indicates whether the interface is administratively up or down
Utilization (In Percent)	Traffic usage in percentage of the packets processed by the particular member interface

You can also use the `show running-configuration interface port-channel` command in EXEC Privilege mode to view whether the mechanism to evaluate the utilization of the member interfaces of the LAG bundle is enabled. The following sample output illustrates the portion of this `show` command:

```
Dell#show running-config int port-channel 1
!
interface Port-channel 1
 mtu 12000
 portmode hybrid
 switchport
 vlt-peer-lag port-channel 1
 no shutdown
 link-bundle-monitor enable
```

Verifying LACP Operation and LAG Configuration

To verify the operational status and configuration of a dynamically created LAG, and LACP operation on a LAG on an Aggregator, enter the `show interfaces port-channel` *`port-channel-number`* and `show lacp` *`port-channel-number`* commands.

The `show` outputs in this section for uplink LAG 128 and server-facing LAG 1 refer to the *LACP Operation on an Aggregator* figure

```
show interfaces port-channel 128
```

Command Example

```
Dell# show interfaces port-channel 128

Port-channel 128 is up, line protocol is up
Created by LACP protocol
Hardware address is 00:01:e8:e1:e1:c1, Current address is 00:01:e8:e1:e1:c1
Interface index is 1107755136
Minimum number of links to bring Port-channel up is 1
Internet address is not set
Mode of IP Address Assignment : NONE
DHCP Client-ID :lag1280001e8e1e1c1
MTU 12000 bytes, IP MTU 11982 bytes
LineSpeed 40000 Mbit
Members in this channel:  Te0/9  Te0/10 Te 0/11 Te0/12
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:11:50
Queueing strategy: fifo
Input Statistics:
     182 packets, 17408 bytes
     92 64-byte pkts, 0 over 64-byte pkts, 90 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     182 Multicasts, 0 Broadcasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
```

```
     2999 packets, 383916 bytes, 0 underruns
     5 64-byte pkts, 214 over 64-byte pkts, 2727 over 127-byte pkts
     53 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     2904 Multicasts, 95 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,          4 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:11:42
```

`show lacp 128` **Command Example**

```
Dell# show lacp 128

Port-channel 128 admin up, oper up, mode lacp
Actor   System ID:  Priority 32768, Address 0001.e8e1.e1c3
Partner System ID:  Priority 32768, Address 0001.e88b.253d
Actor Admin Key 128, Oper Key 128, Partner Oper Key 128, VLT Peer Oper Key 128
LACP LAG 128 is an aggregatable link
LACP LAG 128 is a normal LAG

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 0/9 is enabled, LACP is enabled and mode is lacp
Port State: Bundle
  Actor   Admin: State ADEHJLMP Key 128 Priority 32768
           Oper: State ADEGIKNP Key 128 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ACEGIKNP Key 128 Priority 32768

Port Te 0/10 is enabled, LACP is enabled and mode is lacp
Port State: Bundle
  Actor   Admin: State ADEHJLMP Key 128 Priority 32768
           Oper: State ADEGIKNP Key 128 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ACEGIKNP Key 128 Priority 32768

Port Te 0/11 is enabled, LACP is enabled and mode is lacp
Port State: Bundle
  Actor   Admin: State ADEHJLMP Key 128 Priority 32768
           Oper: State ADEGIKNP Key 128 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ACEGIKNP Key 128 Priority 32768
```

`show interfaces port-channel 1` **Command Example**

```
Dell# show interfaces port-channel 1

Port-channel 1 is up, line protocol is up
Created by LACP protocol
Hardware address is 00:01:e8:e1:e1:c1, Current address is 00:01:e8:e1:e1:c1
Interface index is 1107755009
Minimum number of links to bring Port-channel up is 1
Internet address is not set
Mode of IP Address Assignment : NONE
DHCP Client-ID :lag10001e8e1e1c1
MTU 12000 bytes, IP MTU 11982 bytes
LineSpeed 10000 Mbit
Members in this channel:  Te 0/12(U)
```

```
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:12:41
Queueing strategy: fifo
Input Statistics:
     112 packets, 18161 bytes
     0 64-byte pkts, 46 over 64-byte pkts, 37 over 127-byte pkts
     29 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     59 Multicasts, 53 Broadcasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     135 packets, 19315 bytes, 0 underruns
     0 64-byte pkts, 79 over 64-byte pkts, 32 over 127-byte pkts
     24 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     93 Multicasts, 42 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: 00:12:38
```

`show lacp 1` **Command Example**

```
Dell# show lacp 1

Port-channel 1 admin up, oper up, mode lacp
Actor   System ID:  Priority 32768, Address 0001.e8e1.e1c3
Partner System ID:  Priority 65535, Address 24b6.fd87.d8ac
Actor Admin Key 1, Oper Key 1, Partner Oper Key 33, VLT Peer Oper Key 1
LACP LAG 1 is an aggregatable link
LACP LAG 1 is a normal LAG

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 0/12 is enabled, LACP is enabled and mode is lacp
Port State: Bundle
  Actor   Admin: State ADEHJLMP Key 1 Priority 32768
           Oper: State ADEGIKNP Key 1 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ADEGIKNP Key 33 Priority 255
```

Multiple Uplink LAGs with 10G Member Ports

The following sample commands configure multiple dynamic uplink LAGs with 10G member ports based on LACP.

1. Bring up all the ports.

   ```
   Dell#configure
   Dell(conf)#int range tengigabitethernet 0/1 - 12
   Dell(conf-if-range-te-0/1-12)#no shutdown
   ```

2. Associate the member ports into LAG-10 and 11.

   ```
   Dell#configure
   Dell(conf)#int range tengigabitethernet 0/4 - 5
   Dell(conf-if-range-te-0/4-5)#port-channel-protocol lacp
   Dell(conf-if-range-te-0/4-5-lacp)#port-channel 10 mode active
   ```

```
Dell(conf-if-range-te-0/4-5-lacp)#end
Dell#
Dell#configure
Dell(conf)#int tengigabitethernet 0/6
Dell(conf-if-te-0/6)#port-channel-protocol lacp
Dell(conf-if-te-0/6-lacp)#port-channel 11 mode active
Dell(conf-if-te-0/6-lacp)#end
Dell#
```

3. Show the LAG configurations and operational status.

```
Dell#show interface port-channel brief Codes: L - LACP Port-channel

Dell#show interface port-channel brief Codes: L - LACP Port-channel
O - OpenFlow Controller Port-channel LAG  Mode  Status      Uptime     Ports
L   10   L3    up          00:01:00   Te 0/4   (Up)
                                                        Te 0/5   (Up)
L   11   L3    up          00:00:01   Te 0/6   (Up)
Dell#
```

4. Configure the port mode, VLAN, and so forth on the port-channel.

```
Dell#configure
Dell(conf)#int port-channel 10
Dell(conf-if-po-10)#portmode hybrid
Dell(conf-if-po-10)#switchport
Dell(conf-if-po-10)#vlan tagged 1000
Dell(conf-if-po-10)#link-bundle-monitor enable

Dell#configure
Dell(conf)#int port-channel 11
Dell(conf-if-po-11)#portmode hybrid
Dell(conf-if-po-11)#switchport
Dell(conf-if-po-11)#vlan tagged 1000
% Error: Same VLAN cannot be added to more than one uplink port/LAG.
Dell(conf-if-po-11)#vlan tagged 1001
Dell(conf-if-po-11)#link-bundle-monitor enable

Dell#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, H - VSN tagged
   i - Internal untagged, I - Internal tagged, v - VLT
untagged, V - VLT tagged

    NUM     Status      Description     Q     Ports
* 1        Active                        U Po10(Te 0/4-5)
                                           U Po11(Te 0/6)
  1000  Active                          T Po10(Te 0/4-5)
  1001  Active                          T Po11(Te 0/6)
Dell#
```

5. Show LAG member ports utilization.

```
Dell#show link-bundle-distribution
Link-bundle trigger threshold - 60
LAG bundle - 10 Utilization[In Percent] - 0 Alarm State - Inactive
Interface     Line Protocol     Utilization[In Percent]
Te 0/4                Up                            0
Te 0/5                Up                            0
LAG bundle - 11 Utilization[In Percent] - 0 Alarm State - Inactive
Interface     Line Protocol     Utilization[In Percent]
```

```
Te 0/6                  Up                              0
Dell#
```

31

Layer 2

The Aggregator supports CLI commands to manage the MAC address table:

- Clearing the MAC Address Entries
- Displaying the MAC Address Table

The Aggregator auto-configures with support for Network Interface Controller (NIC) Teaming.

NOTE: On an Aggregator, all ports are configured by default as members of all (4094) VLANs, including the default VLAN. All VLANs operate in Layer 2 mode. You can reconfigure the VLAN membership for individual ports by using the `vlan tagged` or `vlan untagged` commands in INTERFACE configuration mode. See VLAN Membership for more information.

Supported Modes

Standalone, PMUX, VLT, Stacking

NOTE: You can also perform some of the configurations using the Web GUI - Dell Blade IO Manager. For more information, see the Dell Blade IO Manager Online Help.

Managing the MAC Address Table

On an Aggregator in VLT and PMUX modes, you can manage the MAC address table by:

- Displaying the MAC Address Table

In the Standalone mode, use the `show cam mac stack-unit 0 port-set 0` command to view the mac-addresses.

The Aggregator auto-configures with support for Network Interface Controller (NIC) Teaming.

Clearing the MAC Address 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 in PMUX mode and 300 seconds in Standalone and VLT modes.

You can manually clear the MAC address table of dynamic entries.

To clear a MAC address table, use the following command:

1. Clear a MAC address table of dynamic entries. EXEC Privilege mode

`clear mac-address-table dynamic {all | interface {tengigabitethernet <0-5> | SLOT/PORT} }`

- `all`: deletes all dynamic entries.
- `interface`: deletes all entries for the specified interface.

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

  NOTE: This command is available only in PMUX 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.

Network Interface Controller (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.

Support for NIC teaming is auto-configured on the Aggregator, including support for:

- MAC Address Station Move
- MAC Move Optimization

The below fig 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.

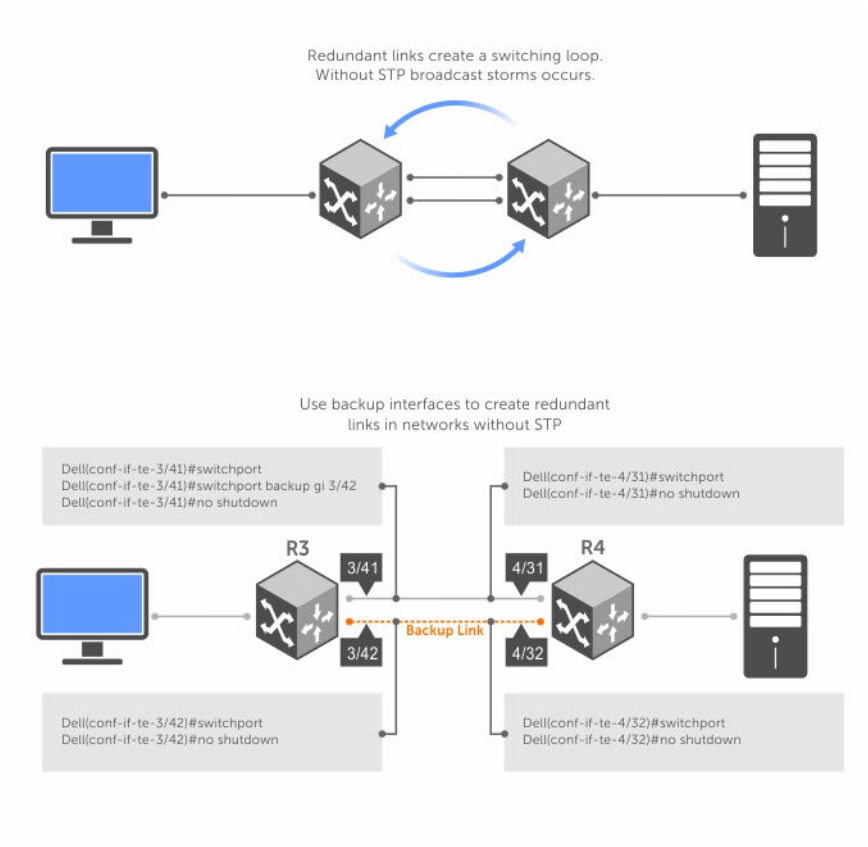


Figure 61. Redundant NOCs with NIC Teaming

MAC Address Station Move

When you use NIC teaming, consider that the server MAC address is originally learned on Port 0/1 of the switch (see figure below). If the NIC fails, the same MAC address is learned on Port 0/5 of the switch. The MAC address is disassociated with one port and re-associated with another in the ARP table; in other words, the ARP entry is "moved". The Aggregator is auto-configured to support MAC Address station moves.

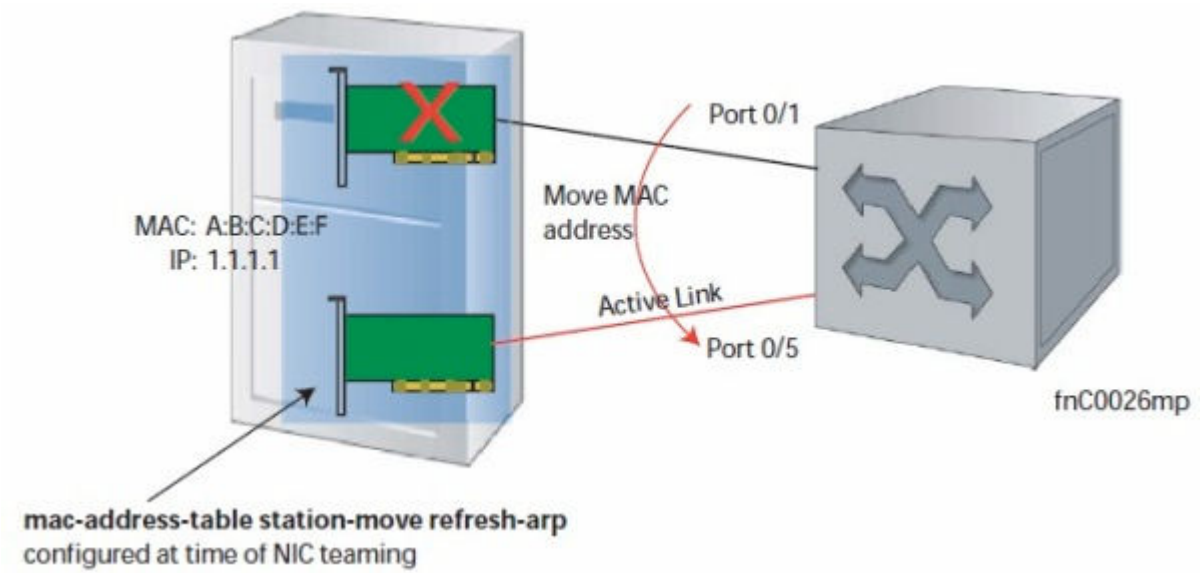


Figure 62. MAC Address Station Move

MAC Move Optimization

Station-move detection takes 5000ms because this is the interval at which the detection algorithm runs.

32

Link Layer Discovery Protocol (LLDP)

Link layer discovery protocol (LLDP) advertises connectivity and management from the local station to the adjacent stations on an IEEE 802 LAN.

LLDP facilitates multi-vendor interoperability by using standard management tools to discover and make available a physical topology for network management. The Dell Networking operating software implementation of LLDP is based on IEEE standard 801.1ab.

The starting point for using LLDP is invoking LLDP with the protocol lldp command in either CONFIGURATION or INTERFACE mode.

The information LLDP distributes is stored by its recipients in a standard management information base (MIB). You can access the information by a network management system through a management protocol such as simple network management protocol (SNMP).

An Aggregator auto-configures to support the link layer discovery protocol (LLDP) for the auto-discovery of network devices. You can use CLI commands to display acquired LLDP information, clear LLDP counters, and debug LACP operation.

Supported Modes

Standalone, PMUX, VLT, Stacking

Protocol Data Units

Configuration information is exchanged in the form of type, length, value (TLV) segments. The below figure shows the chassis ID TLV.

- **Type** — Indicates the type of field that a part of the message represents.
- **Length** — Indicates the size of the value field (in byte).
- **Value** — Indicates the data for this part of the message.

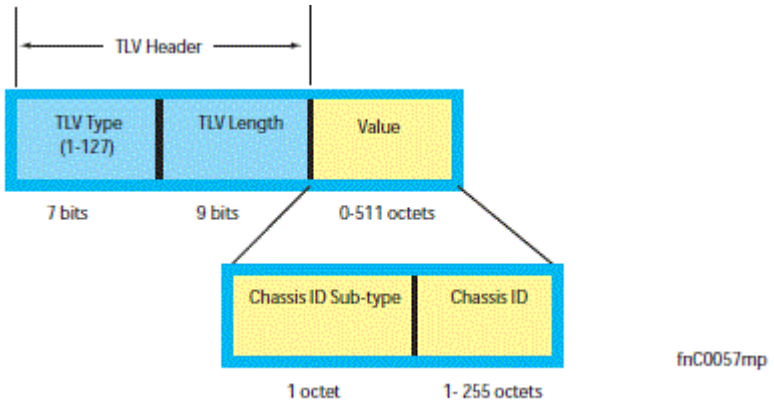


Figure 63. Type, Length, Value (TLV) Segment

TLVs are encapsulated in a frame called an LLDP data unit (LLDPDU), 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 (as shown in the below table). All types are mandatory in the construction of an LLDPDU except Optional TLVs. You can configure the inclusion of individual Optional TLVs.

Table 34. Type, Length, Value (TLV) Types

Type	TLV	Description
0	End of LLDPDU	Marks the end of an LLDPDU.
1	Chassis ID	The Chassis ID TLV is a mandatory TLV that identifies the chassis containing the IEEE 802 LAN station associated with the transmitting LLDP agent.
2	Port ID	The Port ID TLV is a mandatory TLV that identifies the port component of the MSAP identifier associated with the transmitting LLDP agent.
3	Time to Live	The Time To Live TLV indicates the number of seconds that the recipient LLDP agent considers the information associated with this MSAP identifier to be valid.
—	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.

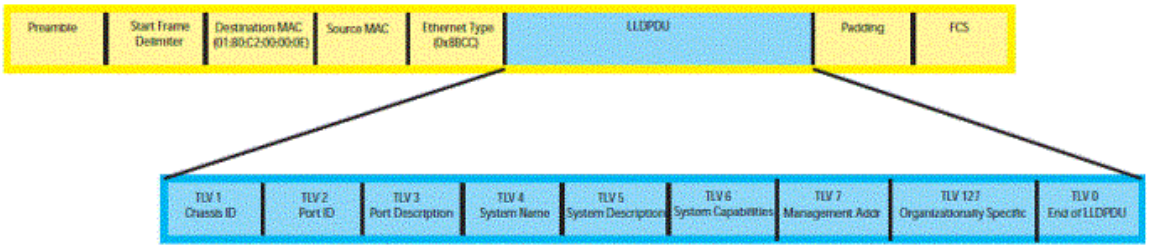


Figure 64. LLDPDU Frame

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
- Configuring LLDPDU Intervals
- Configuring a Time to Live
- Debugging LLDP

Important Points to Remember

- LLDP is enabled by default.
- Dell Networking systems support up to eight neighbors per interface.
- Dell 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.

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)

```
Dell(conf)#protocol lldp
Dell(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


Dell(conf-lldp)#exit
Dell(conf)#interface tengigabitethernet 0/3
Dell(conf-if-te-0/3)#protocol lldp
Dell(conf-if-te-0/3-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
Dell(conf-if-te-0/3-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`.
   - 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.

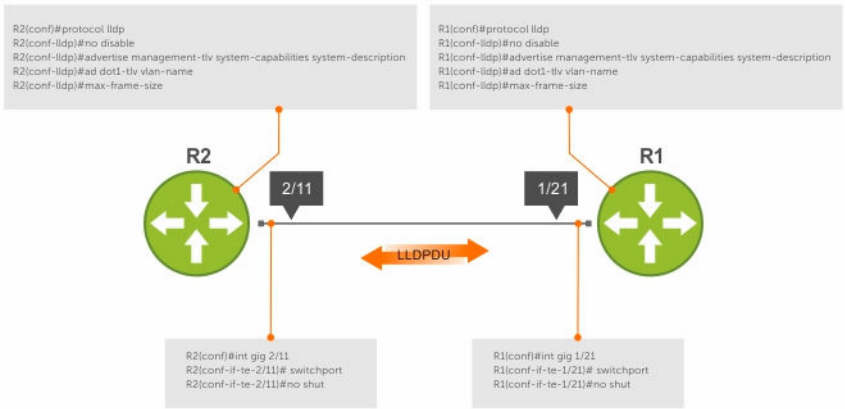


Figure 65. Configuring LLDP

Optional TLVs

The Dell Networking Operating System (OS) supports the following 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.

- Organizationally Unique Identifier (OUI)—a unique number assigned by the IEEE to an organization or vendor.
- OUI Sub-type—These sub-types indicate the kind of information in the following data field. The sub-types are determined by the owner of the OUI.

Figure 66. Organizationally Specific TLVs

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 Networking system to advertise any or all of these TLVs.

Table 35. Optional TLV Types

Type	TLV	Description
Optional TLVs		
4	Port description	A user-defined alphanumeric string that describes the port. The Dell 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,

Type	TLV	Description
		Router, Telephone, DOCSIS cable device, end station only, or other.
8	Management address	Indicates the network address of the management interface. The Dell Networking OS does not currently support this TLV.
IEEE 802.1 Organizationally Specific TLVs		
127	Port-VLAN ID	On Dell Networking systems, indicates the untagged VLAN to which a port belongs.
127	Port and Protocol VLAN ID	On Dell 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	VLAN Name	Indicates the user-defined alphanumeric string that identifies the VLAN.
127	Protocol Identity	Indicates the protocols that the port can process. The Dell 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 Networking OS implementation of LLDP, but is available and mandatory (non-configurable) in the LLDP-MED implementation.
127	Power via MDI	Dell Networking supports the LLDP-MED protocol, which recommends that Power via MDI TLV be not implemented, and therefore Dell Networking implements Extended Power via MDI TLV only.

Type	TLV	Description
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. The Dell Networking OS does not currently support this TLV.
127	Maximum Frame Size	Indicates the maximum frame size capability of the MAC and PHY.

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 Networking system is a network connectivity device, which is Type 4.

When you enable LLDP-MED in Dell Networking OS (using the `advertise med` command), the system begins transmitting this TLV.

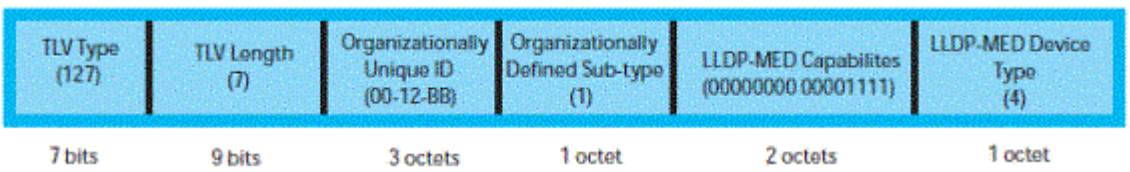


Figure 67. LLDP-MED Capabilities TLV

Table 36. Dell Networking OS LLDP-MED Capabilities

Bit Position	TLV	Dell 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 37. 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 CLI (XXAdvertising 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 38. Network Policy Applications

Type	Application	Description
0	Reserved	—
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.

Type	Application	Description
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	—

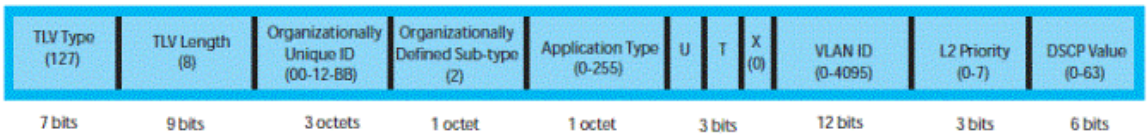


Figure 68. LLDP-MED Policies TLV

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 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 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 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 Networking also honors the power priority value the powered device sends; however, the CLI configuration takes precedence.
- **Power Value** — Dell 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 Networking also honors the power value (power requirement) the powered device sends when the port is configured for `power inline auto`.

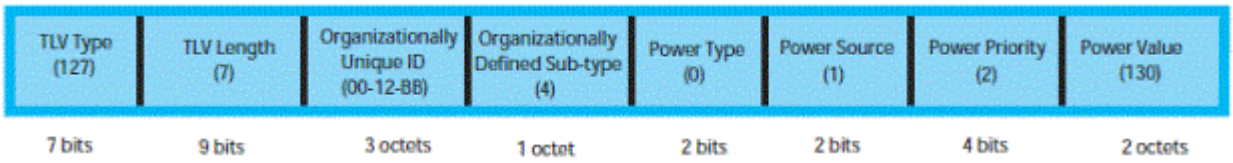


Figure 69. Extended Power via MDI TLV

LLDP Operation

On an Aggregator, LLDP operates as follows:

- LLDP is enabled by default.
- LLDPDUs are transmitted and received by default. LLDPDUs are transmitted periodically. The default interval is 30 seconds.
- LLDPDU information received from a neighbor expires after the default Time to Live (TTL) value: 120 seconds.
- Dell Networking OS supports up to eight neighbors per interface.
- Dell Networking OS supports 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.
- LLDP is not hitless.

Viewing the LLDP Configuration

To view the LLDP configuration, use the following command.

- Display the LLDP configuration.
  CONFIGURATION or INTERFACE mode

  ```
  show config
  ```

Example of Viewing LLDP Global Configurations

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
  advertise dot3-tlv max-frame-size
  advertise management-tlv system-capabilities system-description
  hello 10
  no disable
R1(conf-lldp)#
```

Example of Viewing LLDP Interface Configurations

```
R1(conf-lldp)#exit
R1(conf)#interface tengigabitethernet 0/3
R1(conf-if-te-0/3)#show config
!
interface tengigabitEthernet 0/3
  switchport
  no shutdown
R1(conf-if-te-0/3)#protocol lldp
R1(conf-if-te-0/3-lldp)#show config
!
```

```
  protocol lldp
R1(conf-if-te-0/3-lldp)#
```

Viewing Information Advertised by Adjacent LLDP Agents

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`

Example of Viewing Brief Information Advertised by Neighbors

```
Dell(conf-if-te-0/3-lldp)#end
Dell(conf-if-te-0/3)#do show lldp neighbors
Loc PortID   Rem Host Name  Rem Port Id               Rem Chassis Id
-------------------------------------------------------------------------
Te 0/2       -              TenGigabitEthernet 0/7   00:01:e8:06:95:3e
Te 0/3       -              TenGigabitEthernet 0/8   00:01:e8:09:c2:4a
```

Example of Viewing Details Advertised by Neighbors

```
Dell#show lldp neighbors detail
=============================================================================
Local Interface Te 0/2 has 1 neighbor
  Total Frames Out: 6547
  Total Frames In: 4136
  Total Neighbor information Age outs: 0
  Total Frames Discarded: 0
  Total In Error Frames: 0
  Total Unrecognized TLVs: 0
  Total TLVs Discarded: 0
  Next packet will be sent after 7 seconds
  The neighbors are given below:
------------------------------------------------------------------------------

  Remote Chassis ID Subtype: Mac address (4)
  Remote Chassis ID: 00:01:e8:06:95:3e
  Remote Port Subtype: Interface name (5)
  Remote Port ID: TenGigabitEthernet 0/7
  Local Port ID: TenGigabitEthernet 0/2
  Locally assigned remote Neighbor Index: 4
  Remote TTL: 120
  Information valid for next 120 seconds
  Time since last information change of this neighbor: 01:50:16
  Remote MTU: 1554
  Remote System Desc: Dell Real Time Operating System Software
     Dell Operating System Version: 2.0.
     Dell Application Software Version:  9-4(0-180).
     Copyright (c) 1999-2014 by Dell Inc. All Rights Reserved.
  Existing System Capabilities: Repeater Bridge Router
  Enabled System Capabilities: Repeater Bridge Router
  Remote Port Vlan ID: 1
  Port and Protocol Vlan ID: 1, Capability: Supported, Status: Enabled
  ------------------------------------------------------------------------------

==============================================================================
```

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
  ```

Example of Viewing LLDPDU Intervals

```
Dell#conf
Dell(conf)#protocol lldp
Dell(conf-lldp)#show config
!
protocol lldp
Dell(conf-lldp)#hello ?
<5-180>                 Hello interval in seconds (default=30)
Dell(conf-lldp)#hello 10
Dell(conf-lldp)#show config
!
protocol lldp
 hello 10
Dell(conf-lldp)#
Dell(conf-lldp)#no hello
Dell(conf-lldp)#show config
!
protocol lldp
Dell(conf-lldp)#
```

Configuring a Time to Live

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
  ```

Example of the `multiplier` Command to Configure Time to Live

```
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)#
```

Clearing LLDP Counters

You can clear LLDP statistics that are maintained on an Aggregator for LLDP counters for frames transmitted to and received from neighboring devices on all or a specified physical interface.

To clear LLDP counters, enter the `clear lldp` counters command.

1. Clear counters for LLDP frames sent to and received from neighboring devices on all Aggregator interfaces or on a specified interface.

   interface specifies a 10GbE uplink port in the format: **tenGigabitEthernet** *slot*/*port*.

   EXEC Privilege mode

   `clear lldp counters [interface]`

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.
  `debug lldp detail`

Figure 70. The debug lldp detail Command — LLDPDU Packet Dissection

Relevant Management Objects

Dell Networkings OS supports all IEEE 802.1AB MIB objects.

The following tables list the objects associated with:

- Received and transmitted TLVs
- LLDP configuration on the local agent
- IEEE 802.1AB Organizationally Specific TLVs
- Received and transmitted LLDP-MED TLVs

Table 39. 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.

MIB Object Category	LLDP Variable	LLDP MIB Object	Description
	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 40. 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
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

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
		interface numbering subtype	Local	lldpLocManAddrIfSubtype
			Remote	lldpRemManAddrIfSubtype
		interface number	Local	lldpLocManAddrIfId
			Remote	lldpRemManAddrIfId
		OID	Local	lldpLocManAddrOID
			Remote	lldpRemManAddrOID

Table 41. 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
		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 42. 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
			Remote	lldpXMedRemMediaPolicyPriority
		DSCP Value	Local	lldpXMedLocMediaPolicyDscp
			Remote	lldpXMedRemMediaPolicyDscp

TLV Sub-Type	TLV Name	TLV Variable	System	LLDP-MED MIB Object
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

33

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. 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 for the transmission of network traffic. At the same time, it also uses a single virtual IP address for all clients as the destination IP address, which enables servers to join the same multicast group in a way 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. The NLB functionality enables flooding of traffic over the VLAN ports (for unicast mode) or a subset of ports in a VLAN (for multicast mode) to avoid overloading and effective performance of the servers for optimal processing of data packets.

NLB functions in two modes, namely unicast mode and multicast mode. The cluster IP address and the associated cluster MAC address are configured in the NLB application running on the Windows Server. In the unicast mode, when the server IP address is attempted to be resolved to the MAC address using the ARP application, the switch determines whether the ARP reply, obtained from the server, is of 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 is mapped to a cluster multicast MAC address that is configured using a static ARP CLI configuration command. After the NLB entry is learned, the traffic is forwarded to all the servers in the VLAN corresponding to the cluster virtual IP address.

NLB Unicast Mode Scenario

Consider a sample topology in which four servers, namely S1 through S4, are configured as a cluster or a farm. This set of servers is connected to a Layer 3 switch, which in turn is connected 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 of a switch cannot learn a single MAC address, the servers are assigned with 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 the VLAN named 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 in this case, 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 feature enabled, after learning the NLB ARP entry, all the subsequent traffic is flooded on all ports in VLAN1.

With NLB, the data frame is forwarded to all the servers for them to perform load-balancing.

NLB Multicast Mode Scenario

Consider a sample topology in which four servers, namely S1 through S4, are configured as a cluster or a farm. This set of servers is connected to a Layer 3 switch, which in turn is connected to the end-clients. They contain a single multicast MAC address (MAC-Cluster: 03-00-5E-11-11-11).

In the multicast NLB mode, a static ARP configuration command is configured to associate the cluster IP address with a multicast cluster MAC address.

With multicast NLB mode, the data is forwarded to all the servers based on the port specified using the Layer 2 multicast command, which is the `mac-address-table static <multicast_mac> multicast vlan <vlan_id> output-range <port1>, <port2>` command in CONFIGURATION mode.

Limitations With Enabling NLB on Switches

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 the NLB is applicable and the 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.

Benefits and Working of Microsoft Clustering

Microsoft clustering allows multiple servers using Microsoft Windows to be represented by one MAC address and IP address in order to provide transparent failover or balancing. Dell Networking OS does not recognize server clusters by default; it must be configured 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 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, you must configure the `ip vlan-flooding` command on the Dell switch at the time that the Microsoft cluster is configured. 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. Since 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 the feature is enabled, are deleted when the feature is disabled, and RP2 triggers an ARP resolution. The feature is disabled with the `no ip vlan-flooding` command.
- When a port is added to the VLAN, the port automatically receives traffic if the feature is enabled. Old ARP entries are not deleted or updated.
- When a member port is deleted, its ARP entries are also deleted from the CAM.
- Port channels in the VLAN also receive traffic.
- There is no impact on the configuration from saving the configuration.
- The feature, if enabled, is displayed in the `show running-config` command output that displays the `ip vlan-flooding` CLI configuration. Apart from it, there is no indication of the enabling of this capability.

Configuring a Switch for NLB

To enable a switch for unicast NLB mode of functioning, 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, needs to be flooded across all the member ports of that VLAN. 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.

CONFIGURATION mode

```
ip vlan-flooding
```

Multicast NLB Mode

To enable a switch for multicast NLB mode of functioning, perform the following steps:

1. In the multicast mode of NLB, add a static ARP entry by entering the `arp ip-address multicast-mac-address` command in Global configuration mode to associate an IP address with a multicast MAC address in the switch. This setting causes the multicast MAC address to be mapped to the cluster IP address for the NLB mode of operation of the switch.

   INTERFACE mode

   ```
   arp ip-address multicast-mac-address interface
   ```
2. Associate specific MAC or hardware addresses to VLANs.

   CONFIGURATION mode

   ```
   mac-address-table static multicast-mac-address vlan vlan-id output-range
   interface
   ```

34

Multicast Source Discovery Protocol (MSDP)

Dell Networking OS supports multicast source discovery protocol (MSDP).

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.

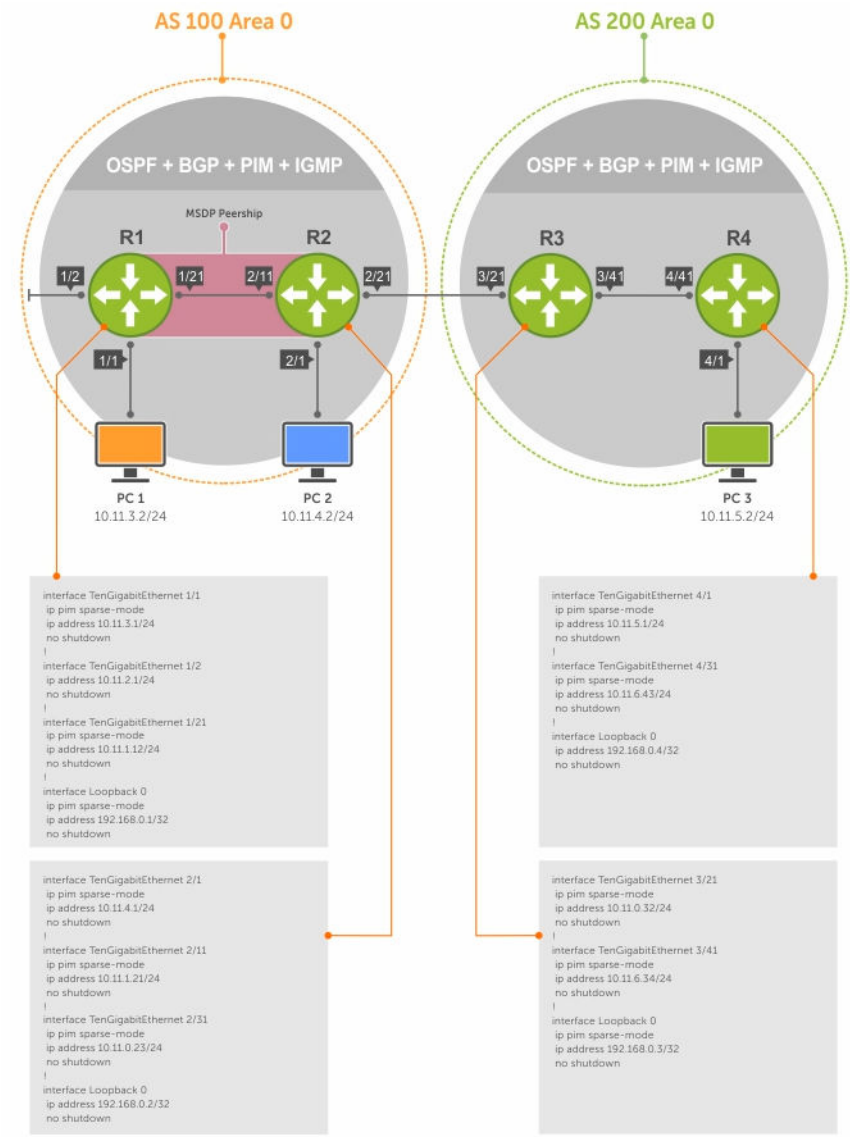


Figure 71. Multicast Source Discovery Protocol (MSDP)

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.

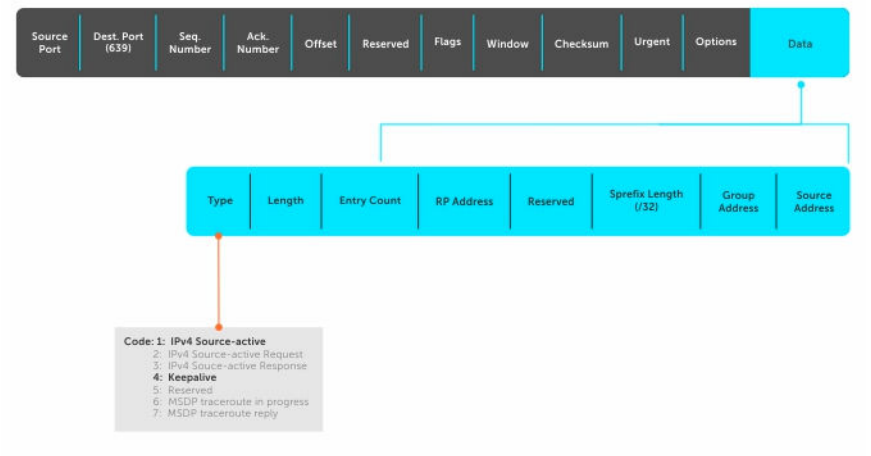


Figure 72. MSDP SA Message Format

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 Networking operating system (OS) implementation of MSDP is in accordance with RFC 3618 and Anycast RP is in accordance with RFC 3446.

Configure the 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 OSPFv3) 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. Enabling MSDP.
4. Peer the RPs in each routing domain with each other. Refer to Enabling MSDP.

Related Configuration Tasks

The following lists related MSDP configuration tasks.

- Enabling MSDP
- Manage the Source-Active Cache
- Accept Source-Active Messages that Fail the RFP Check
- 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
- Terminating a Peership
- Clearing Peer Statistics
- Debugging MSDP
- Anycast RP
- MSDP Sample Configurations

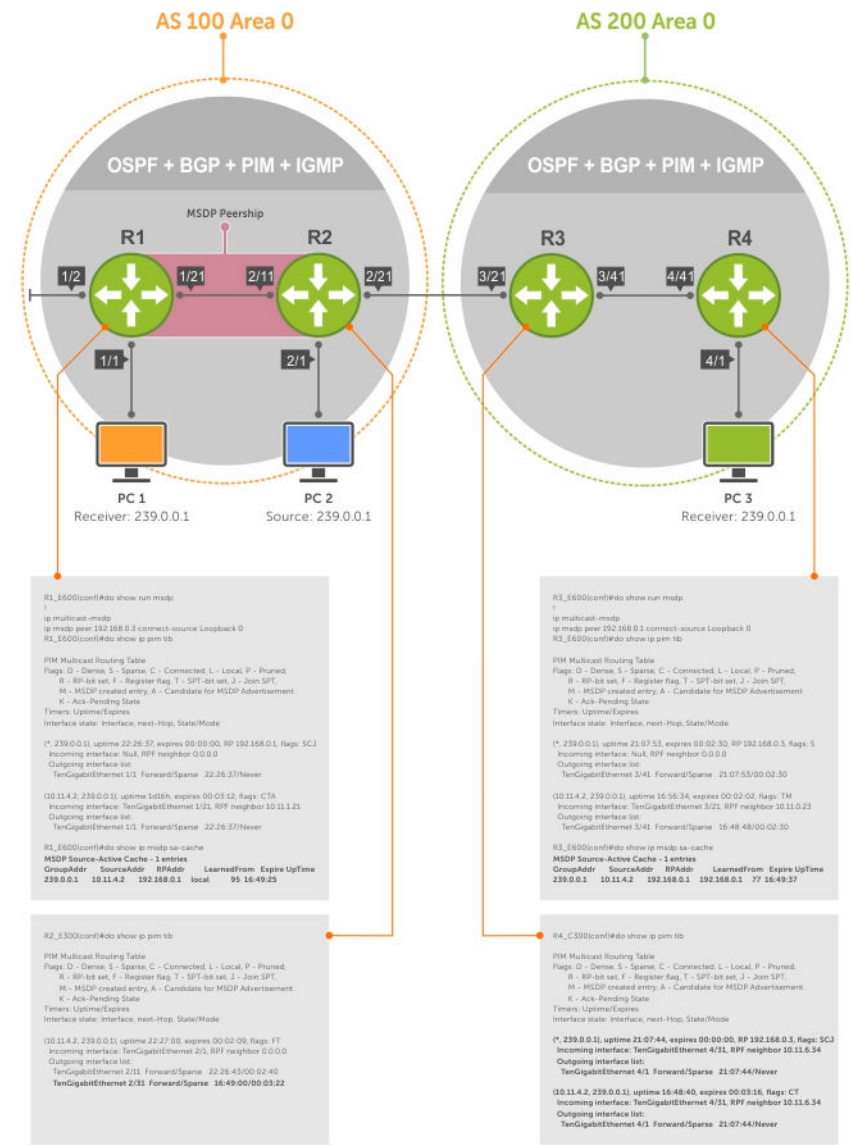


Figure 73. Configuring Interfaces for MSDP

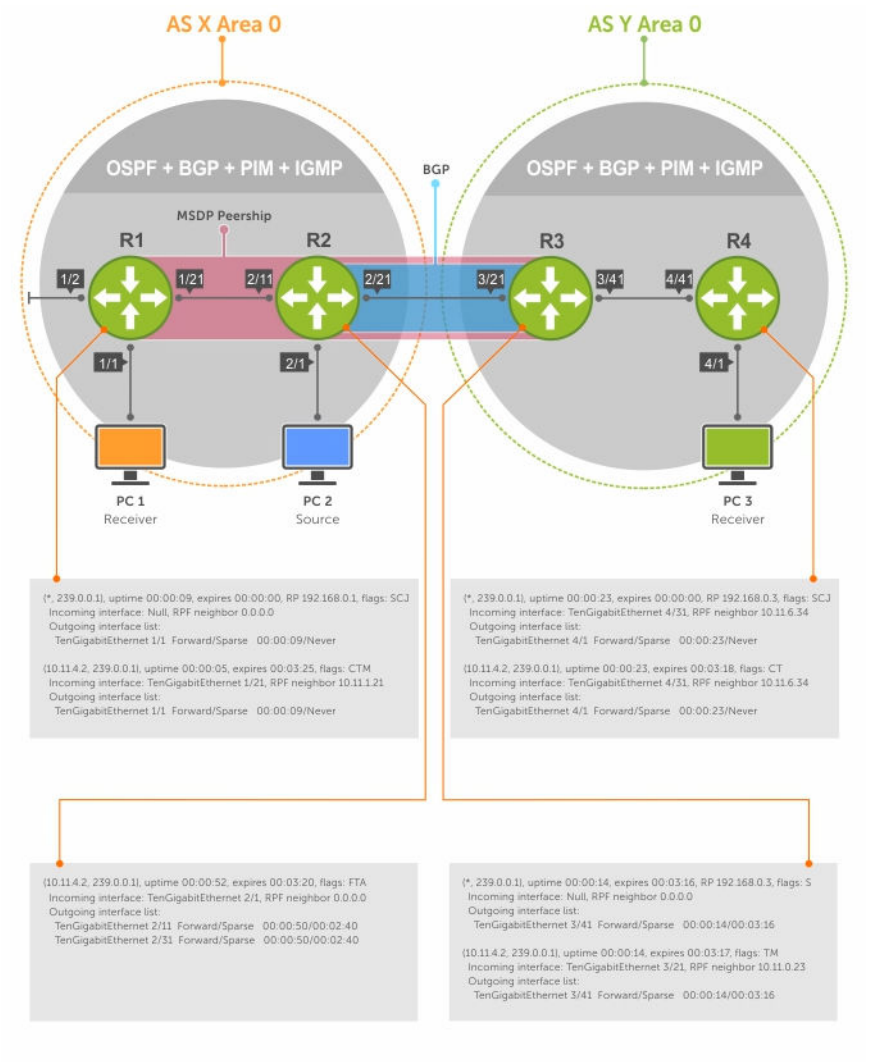


Figure 74. Configuring OSPF and BGP for MSDP

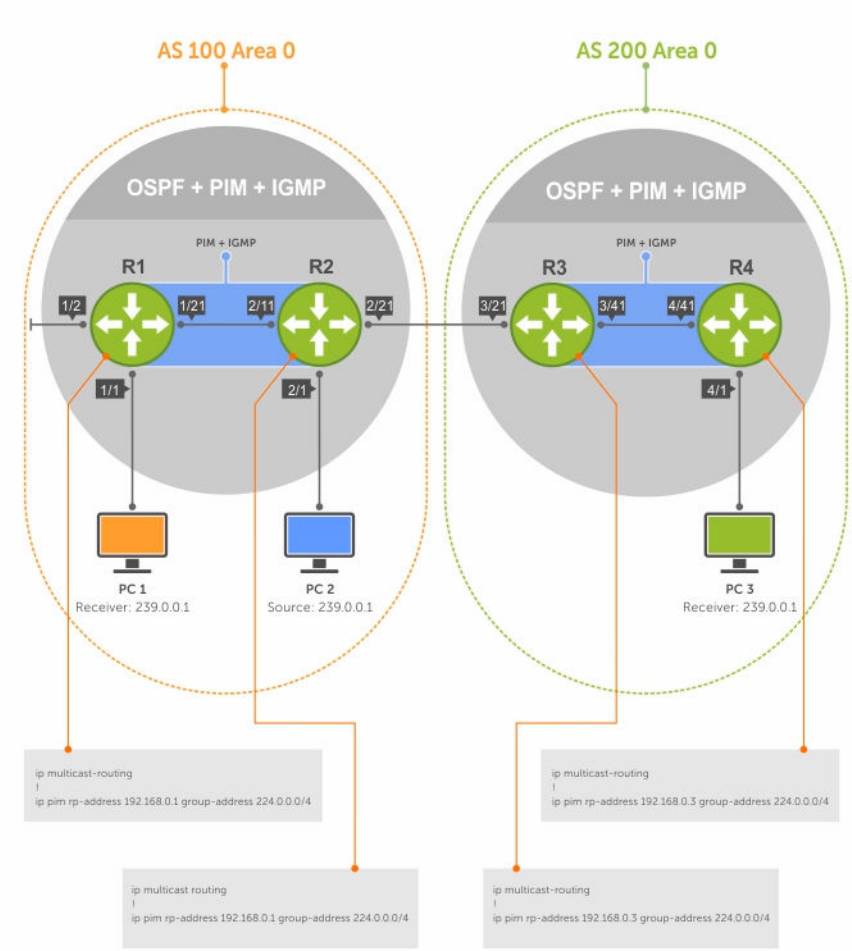


Figure 75. Configuring PIM in Multiple Routing Domains

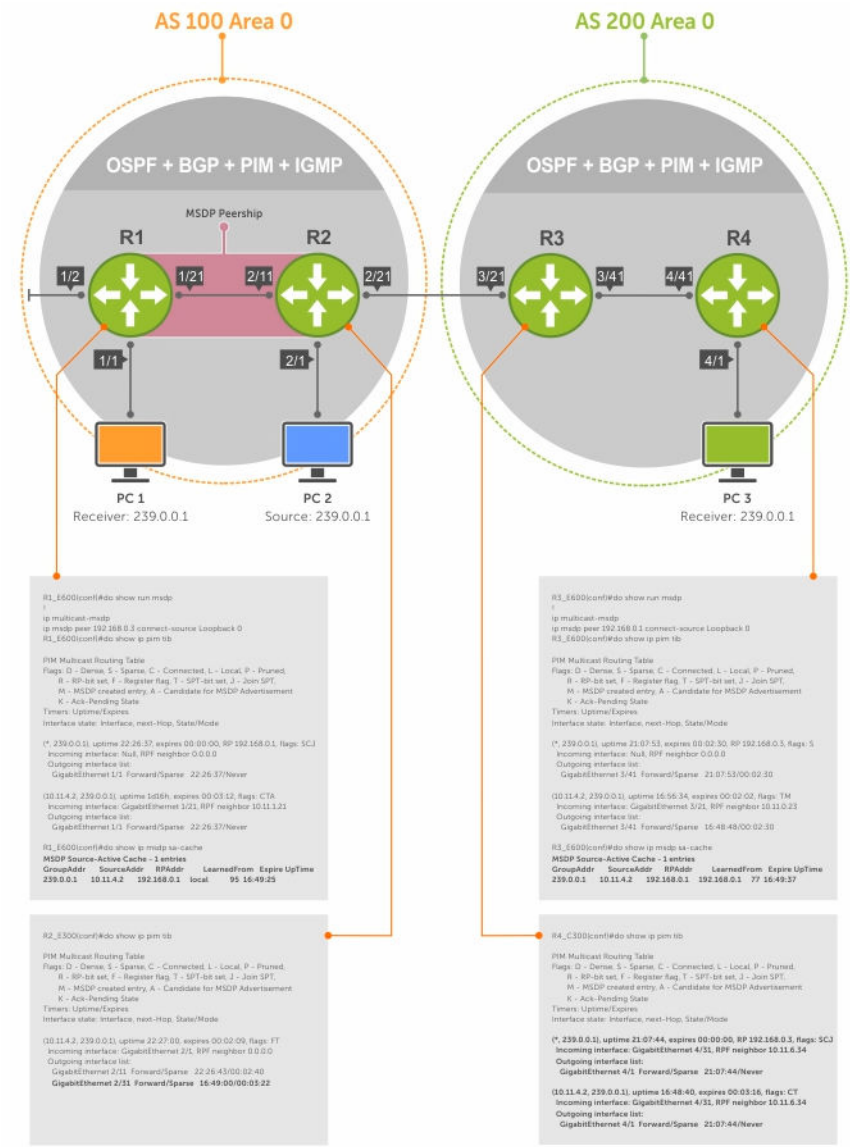


Figure 76. Configuring MSDP

Enabling 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
   ```

Example of Configuring MSDP

Example of Viewing Peer Information

```
R3_E600(conf)#ip multicast-msdp
R3_E600(conf)#ip msdp peer 192.168.0.1 connect-source Loopback 0
R3_E600(conf)#do show ip msdp summary

Peer Addr   Local Addr   State     Source   SA   Up/Down Description
192.168.0.1 192.168.0.3 Established Lo 0    1   00:05:29
```

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_E600#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
  ```

Example of the `show ip msdp sa-cache` Command

```
R3_E600#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 Networking operating system 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 the system 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.

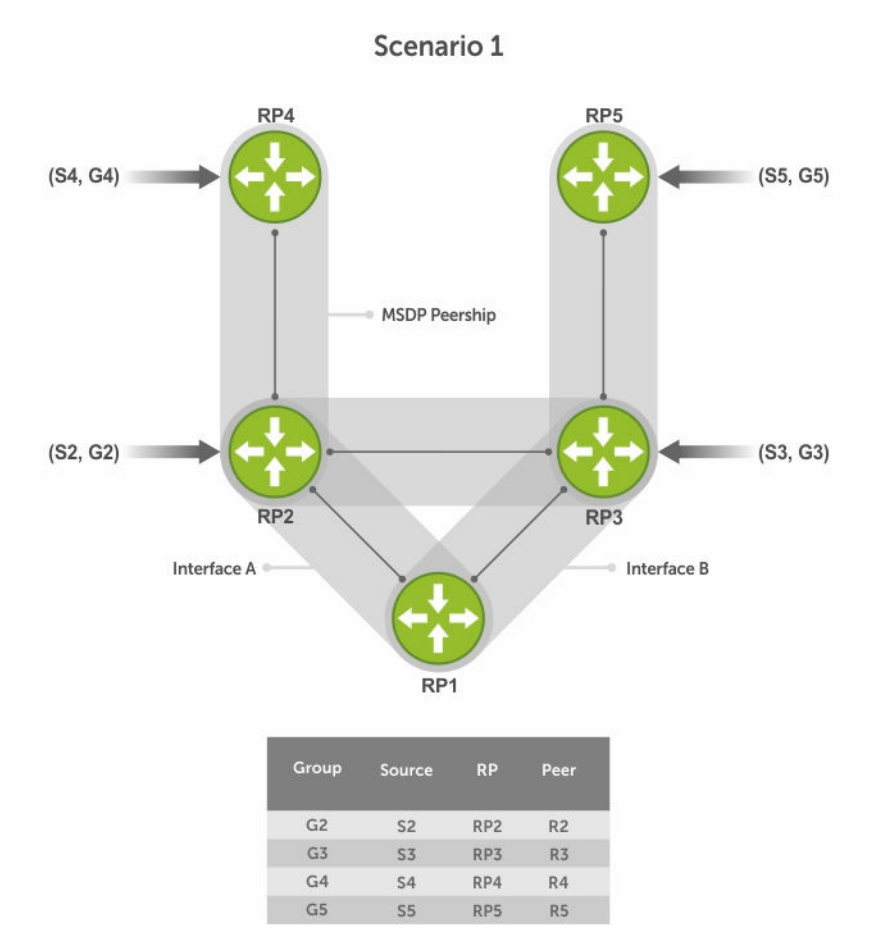


Figure 77. MSDP Default Peer, Scenario 1

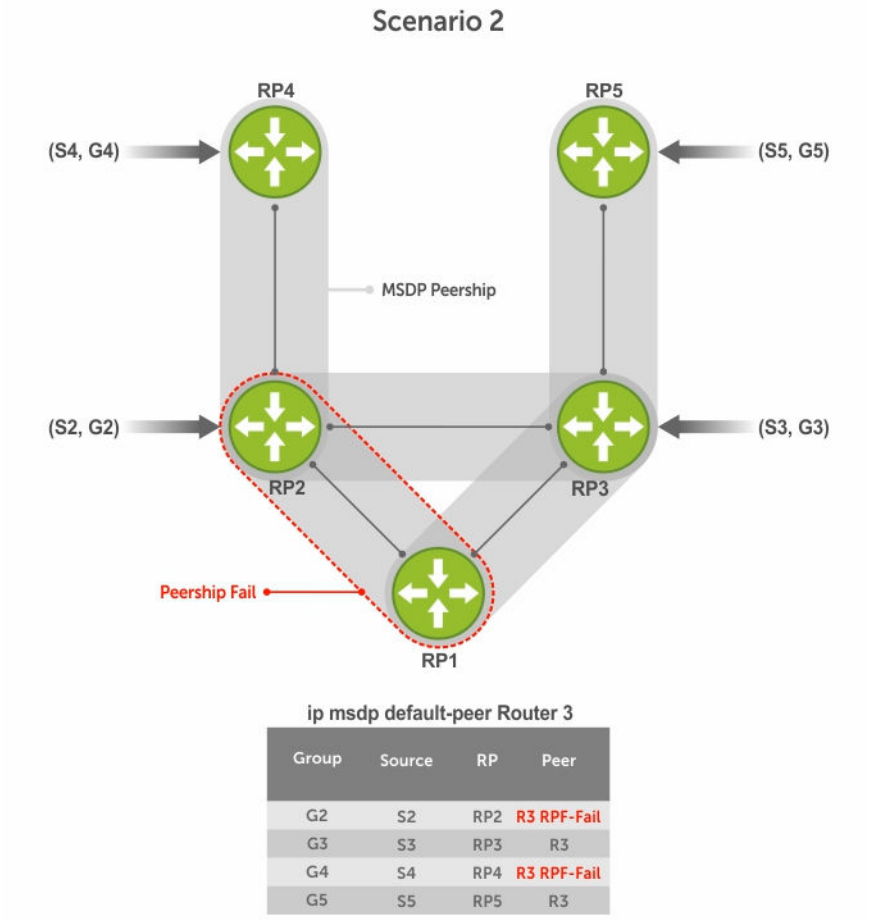


Figure 78. MSDP Default Peer, Scenario 2

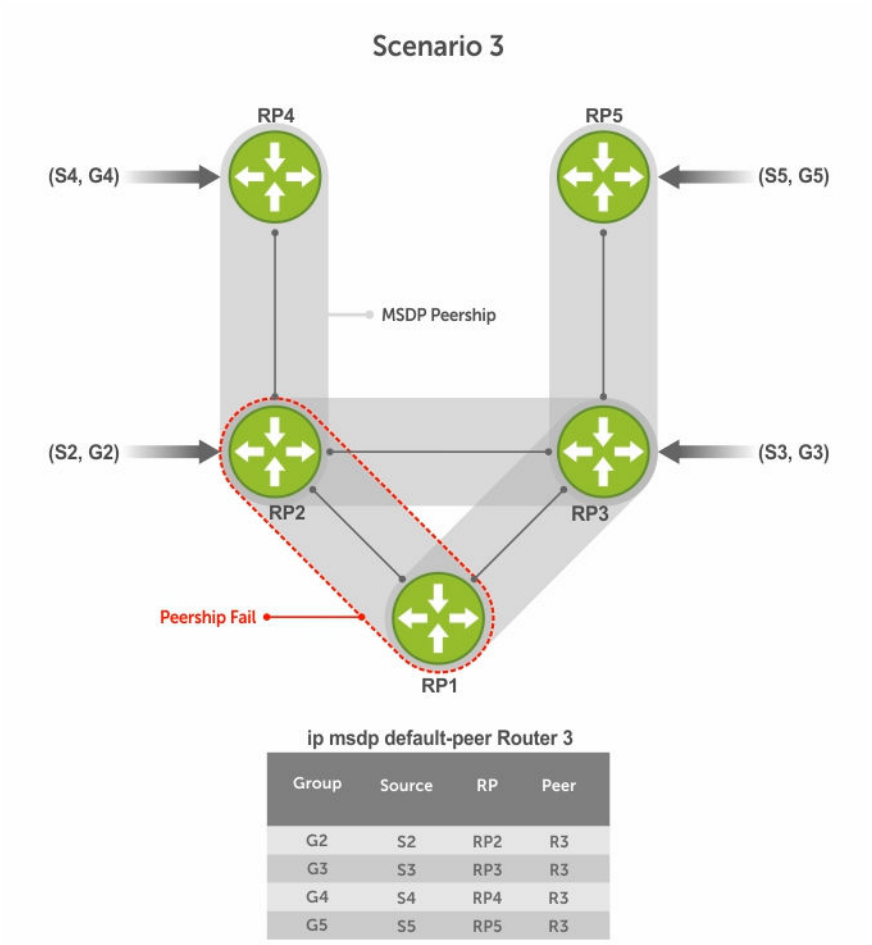


Figure 79. MSDP Default Peer, Scenario 3

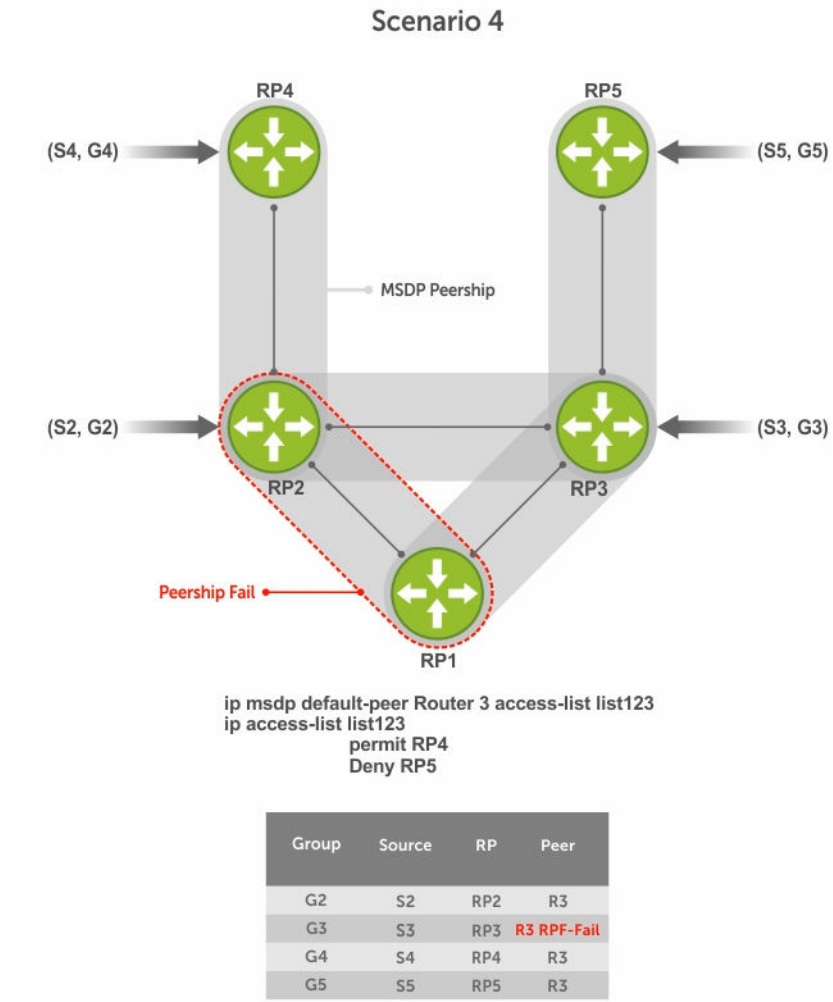


Figure 80. MSDP Default Peer, Scenario 4

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.

Example of the `ip msdp default-peer` Command and Viewing Denied Sources

```
Dell(conf)#ip msdp peer 10.0.50.2 connect-source Vlan 50
Dell(conf)#ip msdp default-peer 10.0.50.2 list fifty
```

```
Dell(conf)#ip access-list standard fifty
Dell(conf)#seq 5 permit host 200.0.0.50

Dell#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

Dell#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
   ```

Example of Verifying the System is not Caching Local Sources

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_E600(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
   ```

Example of Verifying the System is not Caching Remote Sources

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_E600(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_E600(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_E600(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_E600(conf)#do show ip msdp sa-cache
```

```
R3_E600(conf)#
R3_E600(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
  ```

Example of Verifying the System is not Advertising Local Sources

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_E600(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_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
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_E600(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_E600(conf)#do show ip msdp sa-cache
R3_E600(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
```

Example of the Verifying that Peering State is Disabled

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_E600(conf)#ip msdp shutdown 192.168.0.1
R3_E600(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_E600(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
```

Example of the `clear ip msdp peer` Command and Verifying Statistics are Cleared

```
R3_E600(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_E600(conf)#do clear ip msdp peer 192.168.0.1
R3_E600(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
  ```

Example of the `debug ip msdp` Command

```
R1_E600(conf)#do debug ip msdp
All MSDP debugging has been turned on
R1_E600(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.

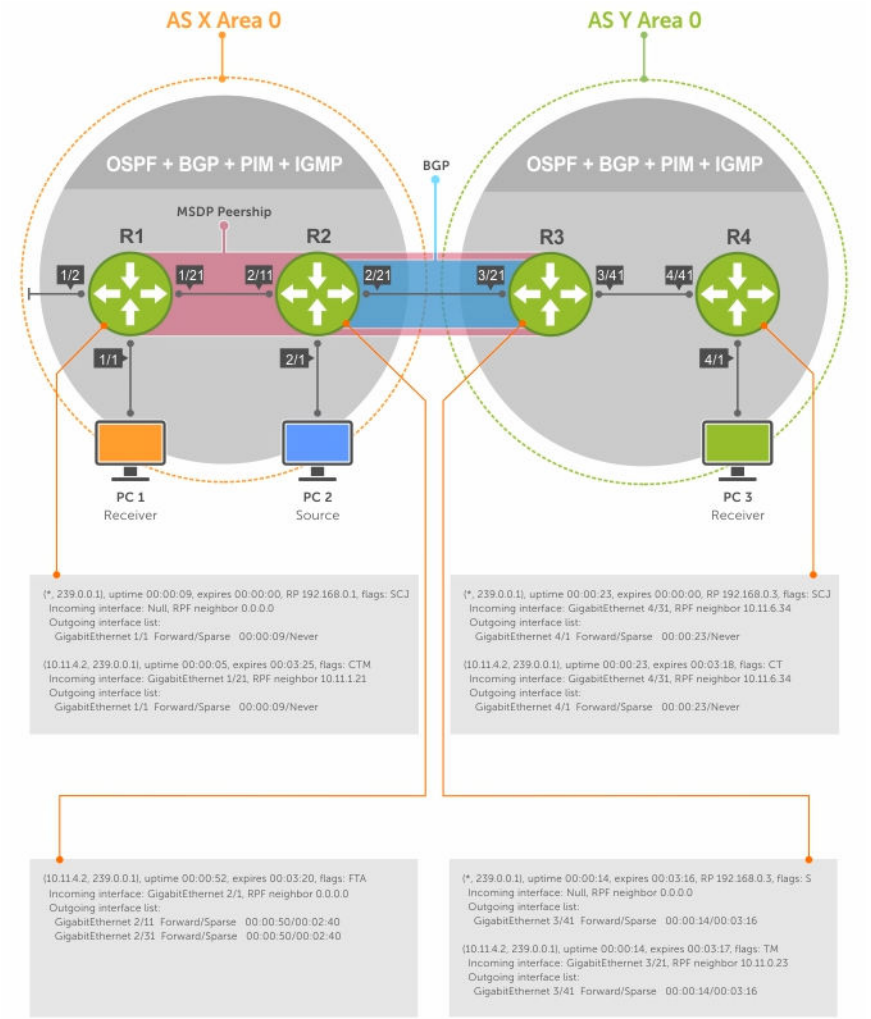


Figure 81. MSDP with Anycast RP

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
```

Example of R1 Configuration for MSDP with Anycast RP

Example of R2 Configuration for MSDP with Anycast RP

Example of R3 Configuration for MSDP with Anycast RP

```
ip multicast-routing
!
interface GigabitEthernet 1/1
```

```
  ip pim sparse-mode
  ip address 10.11.3.1/24
  no shutdown
!
interface GigabitEthernet 1/2
  ip address 10.11.2.1/24
  no shutdown
!
interface GigabitEthernet 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
```

```
ip multicast-routing
!
interface GigabitEthernet 2/1
  ip pim sparse-mode
  ip address 10.11.4.1/24
  no shutdown
!
interface GigabitEthernet 2/11
  ip pim sparse-mode
  ip address 10.11.1.21/24
  no shutdown
!
interface GigabitEthernet 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

ip multicast-routing
!
interface GigabitEthernet 3/21
  ip pim sparse-mode
  ip address 10.11.0.32/24
  no shutdown

interface GigabitEthernet 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, refer to the illustrations in the Related Configuration Tasks section.

MSDP Sample Configuration: R1 Running-Config

MSDP Sample Configuration: R2 Running-Config

MSDP Sample Configuration: R3 Running-Config

MSDP Sample Configuration: R4 Running-Config

```
ip multicast-routing
!
interface GigabitEthernet 1/1
  ip pim sparse-mode
  ip address 10.11.3.1/24
  no shutdown
!
interface GigabitEthernet 1/2
  ip address 10.11.2.1/24
  no shutdown
!
interface GigabitEthernet 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
```

```
ip multicast-routing
!
interface GigabitEthernet 2/1
  ip pim sparse-mode
  ip address 10.11.4.1/24
  no shutdown
!
interface GigabitEthernet 2/11
  ip pim sparse-mode
  ip address 10.11.1.21/24
  no shutdown
!
interface GigabitEthernet 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
```

```
ip multicast-routing
!
interface GigabitEthernet 3/21
  ip pim sparse-mode
  ip address 10.11.0.32/24
  no shutdown
!
interface GigabitEthernet 3/41
  ip pim sparse-mode
  ip address 10.11.6.34/24
  no shutdown
!
interface ManagementEthernet 0/0
  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

ip multicast-routing
!
interface GigabitEthernet 4/1
  ip pim sparse-mode
  ip address 10.11.5.1/24
  no shutdown
!
interface GigabitEthernet 4/22
  ip address 10.10.42.1/24
  no shutdown
!
interface GigabitEthernet 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
```

35

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 on per-VLAN spanning tree plus (PVST+).

Protocol Overview

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.

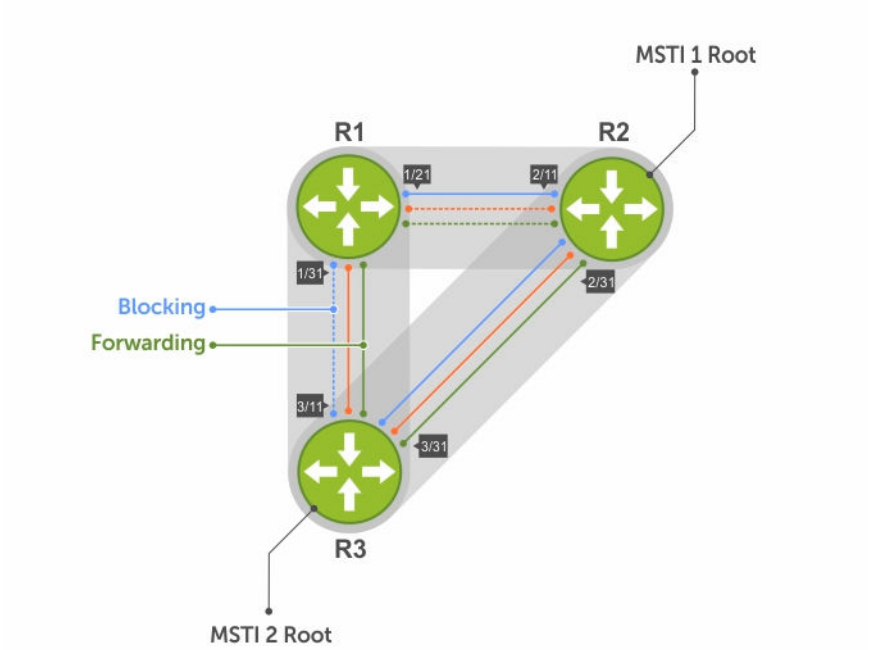


Figure 82. MSTP with Three VLANs Mapped to TWO Spanning Tree Instances

Spanning Tree Variations

The Dell Networking operating system (OS) supports four variations of spanning tree, as shown in the following table.

Table 43. Spanning Tree Variations

Dell Networking Term	IEEE Specification
Spanning Tree Protocol (STP)	802 .1d
Rapid Spanning Tree Protocol (RSTP)	802 .1w
Multicast Source Discovery Protocol (MSDP)	802 .1s
Per-VLAN Spanning Tree Plus (PVST+)	Third Party

Implementation Information

The following describes the MSTP implementation information.

- The Dell 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.
- The Dell Networking OS supports only one MSTP region.
- When you enable MSTP, all ports in Layer 2 mode participate in MSTP.
- On the switch, 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
- Influencing MSTP Root Selection
- Interoperate with Non-Dell Networking OS Bridges
- Enable BPDU Filtering Globally
- Modifying Global Parameters
- Configuring an EdgePort
- Flush MAC Addresses after a Topology Change
- Debugging and Verifying MSTP Configurations
- Prevent Network Disruptions with BPDU Guard

- SNMP Traps for Root Elections and Topology Changes

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
   ```

Example of Verifying MSTP is Enabled

To verify that MSTP is enabled, use the `show config` command in PROTOCOL MSTP mode.

```
Dell(conf)#protocol spanning-tree mstp
Dell(config-mstp)#show config
!
protocol spanning-tree mstp
no disable
Dell#
```

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.

Example of the `msti` Command

Example of Viewing MSTP Port States

```
Dell(conf)#protocol spanning-tree mstp
Dell(conf-mstp)#msti 1 vlan 100
Dell(conf-mstp)#msti 2 vlan 200-300
Dell(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.

To view the forwarding/discarding state of the ports participating in an MSTI, use the `show spanning-tree msti` command from EXEC Privilege mode.

```
Dell#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 TenGig 1/21
Port 374 (TenGigabitEthernet 1/21) is root Forwarding
Port path cost 2000, 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 2000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 93671, received 46843
The port isnot in the Edge port mode, bpdu filter is disabled
Port 384 (TenGigabitEthernet 1/31) is alternate Discarding
Port path cost 2000, 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 2000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 39291, received 7547
The port isnot in the Edge port mode, bpdu filter is disabled
```

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**.

Example of Assigning and Verifying the Root Bridge Priority

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.

```
Dell(conf-mstp)#msti 2 bridge-priority 0
```

```
Dell(conf-mstp)#show config
!
protocol spanning-tree mstp
  MSTI 2 bridge-priority 0
Dell(conf-mstp)#
```

Interoperate with Non-Dell Networking OS Bridges

The Dell 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 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 name and revision match on all Dell Networking OS equipment. If you have non-Dell Networking OS equipment that participates in MSTP, ensure these values match on all the equipment.

NOTE: Some non-Dell Networking OS equipment 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`*

  The range is from 0 to 65535.

  The default is **0**.

Example of the `name` Command

To view the current region name and revision, use the `show spanning-tree mst configuration` command from EXEC Privilege mode.

```
Dell(conf-mstp)#name my-mstp-region
Dell(conf-mstp)#exit
Dell(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 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 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**.

Example of the `forward-delay` Parameter

To view the current values for MSTP parameters, use the `show running-config spanning-tree mstp` command from EXEC privilege mode.

```
Dell(conf-mstp)#forward-delay 16
Dell(conf-mstp)#exit
Dell(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
Dell(conf)#
```

Enable BPDU Filtering Globally

The enabling of BPDU Filtering stops transmitting of BPDUs on the operational port fast enabled ports by default.
When BPDUs are received, the spanning tree is automatically prepared. By default global bpdu filtering is disabled.

Enable BPDU Filter globally to filter transmission of BPDU port fast enabled interfaces.

PROTOCOL MSTP mode

```
edge-port bpdu filter default
```

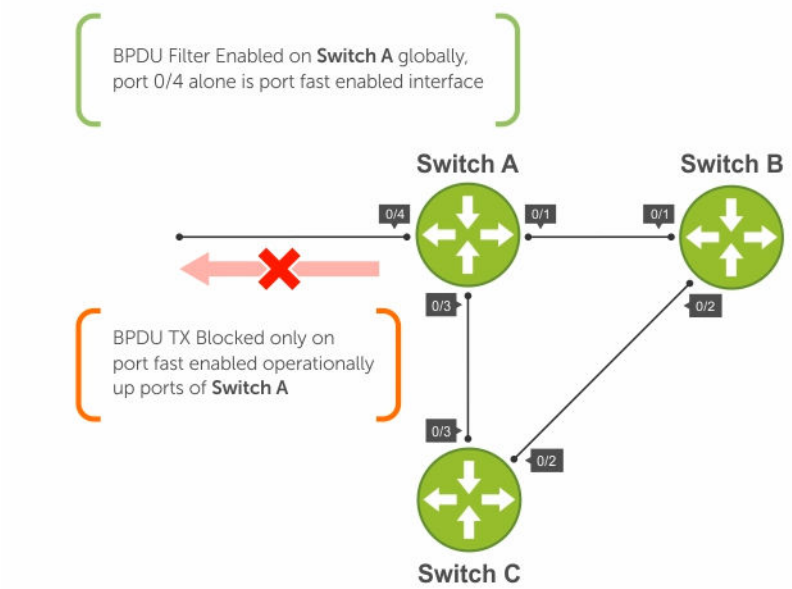


Figure 83. BPDU Filtering Enabled Globally

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 44. Default Values for Port Costs by Interface

Port Cost	Default Value
1000-Mb/s Ethernet interfaces	20000
40-Gigabit Ethernet interfaces	1400
10-Gigabit Ethernet interfaces	2000
Port Channel with one 10–Gigabit Ethernet interface	2000
Port Channel with one 40-Gigabit Ethernet interface	1400
Port Channel with two 10-Gigabit Ethernet interfaces	1800
Port Channel with two 40-Gigabit Ethernet interfaces	600

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.

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 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 mstp edge-port [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).

Example of Enabling an EdgePort on an Interface

To verify that EdgePort is enabled, use the `show config` command from INTERFACE mode.

```
Dell(conf-if-gi-3/41)#spanning-tree mstp edge-port
Dell(conf-if-gi-3/41)#show config
!
interface GigabitEthernet 3/41
  no ip address
  switchport
  spanning-tree mstp edge-port
  spanning-tree MSTI 1 priority 144
  no shutdown
Dell(conf-if-gi-3/41)#
```

Flush MAC Addresses after a Topology Change

The Dell 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 Networking OS systems.

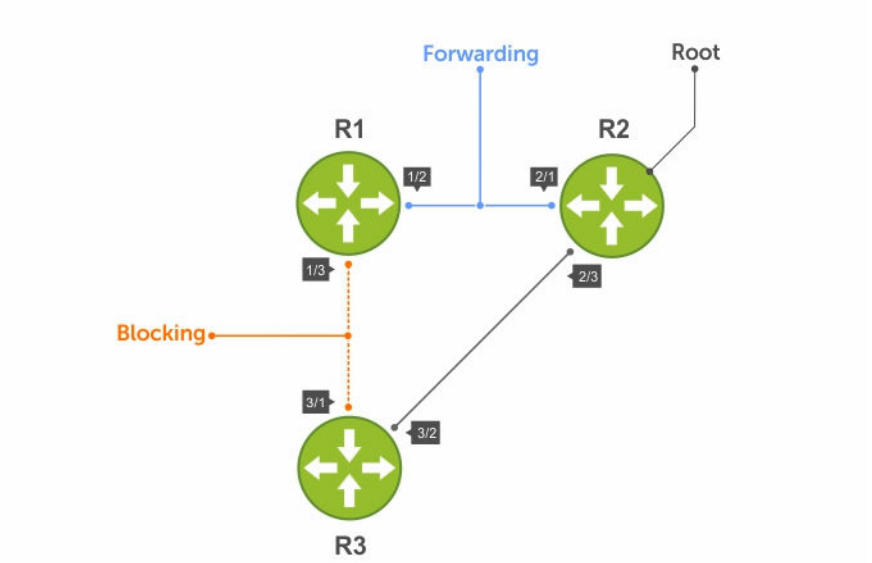


Figure 84. MSTP with Three VLANs Mapped to Two Spanning Tree Instances

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 GigabitEthernet 1/21
  no ip address
  switchport
  no shutdown
!
interface GigabitEthernet 1/31
  no ip address
  switchport
  no shutdown
!
(Step 3)
interface Vlan 100
  no ip address
  tagged GigabitEthernet 1/21,31
  no shutdown
!
interface Vlan 200
  no ip address
  tagged GigabitEthernet 1/21,31
  no shutdown
!
interface Vlan 300
  no ip address
  tagged GigabitEthernet 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 GigabitEthernet 2/11
  no ip address
  switchport
  no shutdown
!
interface GigabitEthernet 2/31
  no ip address
  switchport
  no shutdown
!
(Step 3)
interface Vlan 100
  no ip address
  tagged GigabitEthernet 2/11,31
 no shutdown
!
interface Vlan 200
  no ip address
```

```
  tagged GigabitEthernet 2/11,31
  no shutdown
!
interface Vlan 300
  no ip address
  tagged GigabitEthernet 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 GigabitEthernet 3/11
  no ip address
  switchport
  no shutdown
!
interface GigabitEthernet 3/21
  no ip address
  switchport
  no shutdown
!
(Step 3)
interface Vlan 100
  no ip address
  tagged GigabitEthernet 3/11,21
  no shutdown
!
interface Vlan 200
  no ip address
  tagged GigabitEthernet 3/11,21
  no shutdown
!
interface Vlan 300
  no ip address
  tagged GigabitEthernet 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
  ```

Example of Viewing MSTP Configuration

Example of Viewing the Debug Log (Successful MSTP Configuration)

Example of Viewing the Debug Log (Unsuccessful MSTP Configuration)

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.

```
Dell#show run spanning-tree mstp
!
protocol spanning-tree mstp
  name Tahiti
  revision 123
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300
no disable
```

"Same Region," shown in bold in the following example shows that the MSTP routers are in a single region.

```
Dell#debug spanning-tree mstp bpdu
MSTP debug bpdu is ON
Dell#
4w0d4h : MSTP: Sending BPDU on Tengig 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 Tengig 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78Same 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, Int Root Path Cost: 0
Rem Hops: 19, Bridge Id: 32768:0001.e8d5.cbbd
4w0d4h : INST 1: Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
    Brg/Port Prio: 32768/128, Rem Hops: 19
    INST 2: Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
    Brg/Port Prio: 32768/128, Rem Hops: 19
```

The bold line in the following example shows that the MSTP routers are in different regions and are not communicating with each other.

```
4w0d4h : MSTP: Received BPDU on TenGig 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78Different 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, 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
```

36

Multicast Features

Dell Networking OS supports multicast features.

The Dell Networking OS supports the following multicast protocols:

- PIM Sparse-Mode (PIM-SM)
- PIM Source-Specific Mode (PIM-SSM)
- Internet Group Management Protocol (IGMP)
- Multicast Source Discovery Protocol (MSDP)

Enabling IP Multicast

Prior to enabling any multicast protocols, you must enable multicast routing.

- Enable multicast routing.
  CONFIGURATION mode

  ```
  ip multicast-routing
  ```

Implementation Information

Because protocol control traffic in the 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 Networking OS might forward data traffic with certain MAC addresses to the CPU in addition to control traffic.

As the upper 5 bits of an IP Multicast address are dropped in the translation, 32 different multicast group IDs all 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 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
RIP	01:00:5e:00:00:09
NTP	01:00:5e:00:01:01

Protocol	Ethernet Address
VRRP	01:00:5e:00:00:12
PIM-SM	01:00:5e:00:00:0d

- The Dell Networking OS implementation of MTRACE is in accordance with IETF draft *draft-fenner-traceroute-ipm*.
- Multicast is not supported on secondary IP addresses.
- Egress L3 ACL is not applied to multicast data traffic if you enable multicast routing.

First Packet Forwarding for Lossless Multicast

Beginning with the Dell Networking OS version version 8.3.1.0, all initial multicast packets are forwarded to receivers to achieve lossless multicast.

In previous versions, when the Dell Networking system is an RP, all initial packets are dropped until PIM creates an (S,G) entry. When the system is an RP and a Source DR, these initial packet drops represent a loss of native data, and when the system is an RP only, the initial packets drops represent a loss of register packets.

Both scenarios might be unacceptable depending on the multicast application. Beginning with the Dell Networking OS versions noted here, when the system is the RP, and has receivers for a group G, it forwards all initial multicast packets for the group based on the (*,G) entry rather than discarding them until the (S,G) entry is created, making Dell Networking systems suitable for applications sensitive to multicast packet loss.

NOTE: When a source begins sending traffic, the Source DR forwards the initial packets to the RP as encapsulated registered packets. These packets are forwarded via the soft path at a maximum rate of 70 packets/second. Incoming packets beyond this rate are dropped.

Multicast Policies

The Dell Networking OS offers parallel 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
- Rate Limiting IGMP Join Requests
- 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 Networking OS does not process any IGMP or multicast listener discovery protocol (MLD) joins to 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 MSDP.

- If the limit is increased after it is reached, join subsequent join requests are accepted. In this case, increase the limit by at least 10% for IGMP and MLD to resume.
- If the limit is decreased after it is reached, the system does not clear the existing sessions. Entries are cleared after a timeout (you may also clear entries using `clear ip mroute`).

NOTE: The Dell 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 system displays the following:

```
      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 if from 1 to 50000.

  The default is **15000**.

NOTE: The IN-L3-McastFib CAM partition is used to store multicast routes and is a separate hardware limit that exists per port-pipe. Any software-configured limit may supersede by this hardware space limitation. The opposite is also true, the CAM partition might not be exhausted at the time the system-wide route limit the `ip multicast-limit` command sets is reached.

Preventing a Host from Joining a Group

You can prevent a host from joining a particular group by blocking specific IGMP reports. Create 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 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 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 system clears 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 system from clearing the routing table by entering a *permit any* rule with 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.

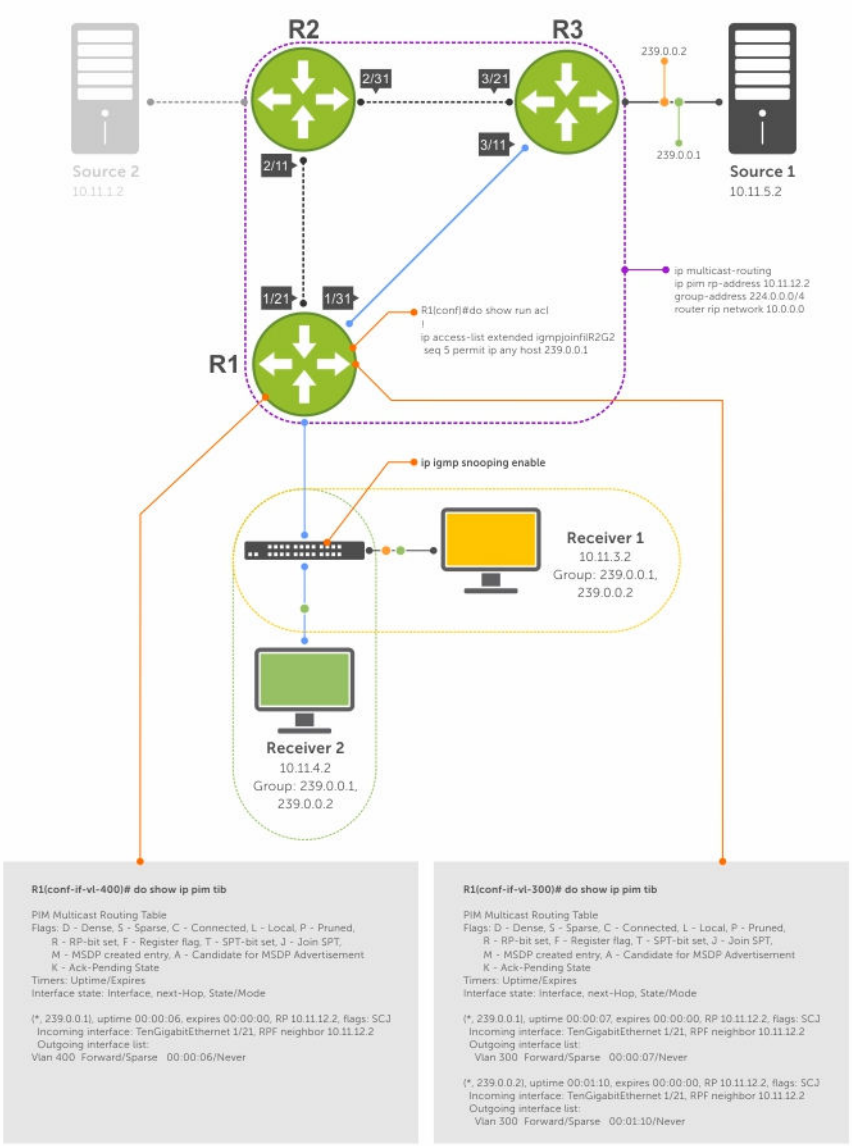


Figure 85. Preventing a Host from Joining a Group

Table 45. Preventing a Host from Joining a Group — Description

Location	Description
1/21	• Interface GigabitEthernet 1/21 • ip pim sparse-mode • ip address 10.11.12.1/24

Location	Description
	• no shutdown
1/31	• Interface GigabitEthernet 1/31 • ip pim sparse-mode • ip address 10.11.13.1/24 • no shutdown
2/1	• Interface GigabitEthernet 2/1 • ip pim sparse-mode • ip address 10.11.1.1/24 • no shutdown
2/11	• Interface GigabitEthernet 2/11 • ip pim sparse-mode • ip address 10.11.12.2/24 • no shutdown
2/31	• Interface GigabitEthernet 2/31 • ip pim sparse-mode • ip address 10.11.23.1/24 • no shutdown
3/1	• Interface GigabitEthernet 3/1 • ip pim sparse-mode • ip address 10.11.5.1/24 • no shutdown
3/11	• Interface GigabitEthernet 3/11 • ip pim sparse-mode • ip address 10.11.13.2/24 • no shutdown
3/21	• Interface GigabitEthernet 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 GigabitEthernet 1/1 • no shutdown
Receiver 2	• Interface VLAN 400 • ip pim sparse-mode • ip address 10.11.4.1/24 • untagged GigabitEthernet 1/2 • **ip igmp access-group igmpjoinfilR2G2**

Location	Description
	• no shutdown

Rate Limiting IGMP Join Requests

If you expect a burst of IGMP Joins, protect the IGMP process from overload by limiting that rate at which new groups can be joined.
Hosts whose IGMP requests are denied will use the retry mechanism built-in to IGMP so that they're membership is delayed rather than permanently denied.

- Limit the rate at which new groups can be joined.
  INTERFACE mode

  ```
  ip igmp group-join-limit
  ```

To view the enable status of this feature, use the `show ip igmp interface` command from EXEC Privilege mode.

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 protocol independent multicast (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 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.

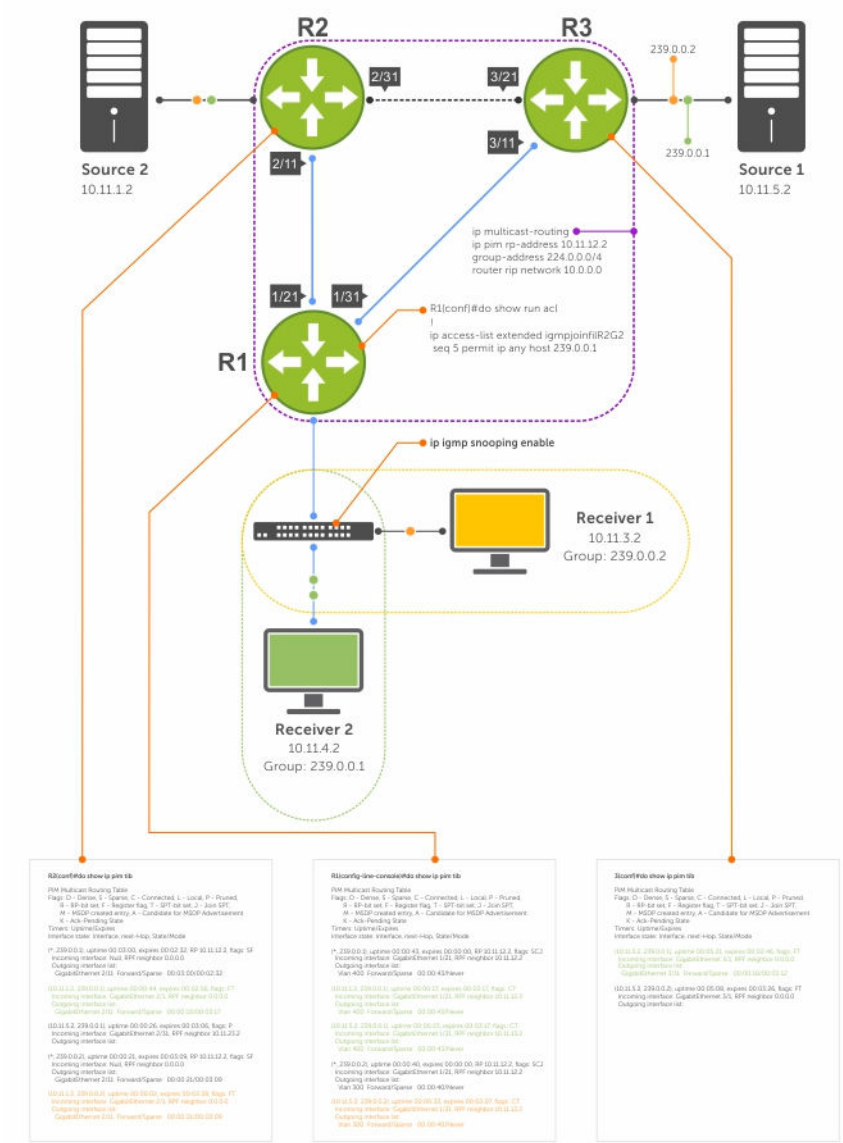


Figure 86. Preventing a Source from Transmitting to a Group

Table 46. Preventing a Source from Transmitting to a Group — Description

Location	Description
1/21	• Interface GigabitEthernet 1/21 • ip pim sparse-mode • ip address 10.11.12.1/24 • no shutdown
1/31	• Interface GigabitEthernet 1/31 • ip pim sparse-mode • ip address 10.11.13.1/24

Location	Description
	• no shutdown
2/1	• Interface GigabitEthernet 2/1 • ip pim sparse-mode • ip address 10.11.1.1/24 • no shutdown
2/11	• Interface GigabitEthernet 2/11 • ip pim sparse-mode • ip address 10.11.12.2/24 • no shutdown
2/31	• Interface GigabitEthernet 2/31 • ip pim sparse-mode • ip address 10.11.23.1/24 • no shutdown
3/1	• Interface GigabitEthernet 3/1 • ip pim sparse-mode • ip address 10.11.5.1/24 • no shutdown
3/11	• Interface GigabitEthernet 3/11 • ip pim sparse-mode • ip address 10.11.13.2/24 • no shutdown
3/21	• Interface GigabitEthernet 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 GigabitEthernet 1/1 • no shutdown
Receiver 2	• Interface VLAN 400 • ip pim sparse-mode • ip address 10.11.4.1/24 • untagged GigabitEthernet 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 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 is generated 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 UDP-encapsulated registration messages between the DR and RP routers which are being sent to the CPU.

- Prevent the PIM SM router from creating state based on multicast source and/ or group.
  INTERFACE

  ```
  ip pim join-filter
  ```

37

Open Shortest Path First (OSPFv2 and OSPFv3)

Dell Networking OS supports open shortest path first (OSPFv2 for IPv4) and OSPF version 3 (OSPF for IPv6).

This chapter provides a general description of OSPFv2 (OSPF for IPv4) and OSPFv3 (OSPF for IPv6) as supported in the Dell 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.

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.

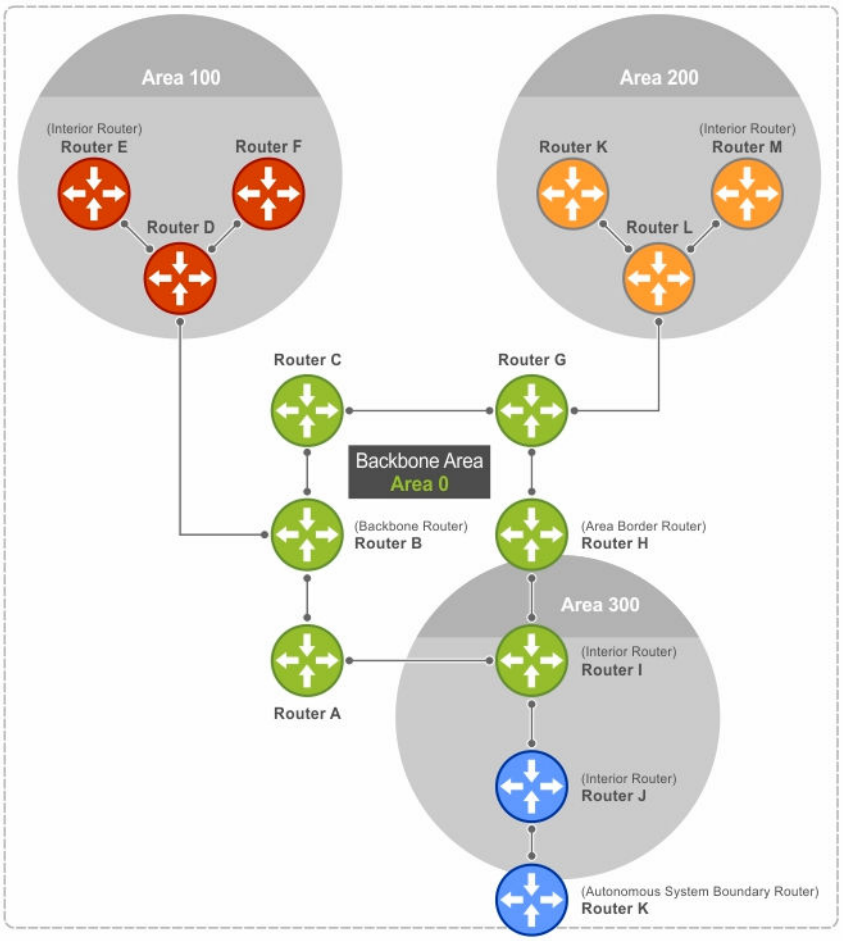


Figure 87. Autonomous System Areas

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. Areas can be defined in such a way that the backbone is not contiguous. In this case, backbone connectivity must be restored through virtual links. Virtual links are configured between any backbone routers that share a link to a non-backbone area and function as if they were direct links.

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.

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 virtual link cannot traverse stubby areas.
- 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. However, a virtual link can traverse it.
- Totally stubby areas are referred to as no summary areas in the Dell 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 Networking recommends that the router ID and the router's IP address reflect each other.

The following example shows different router designations.

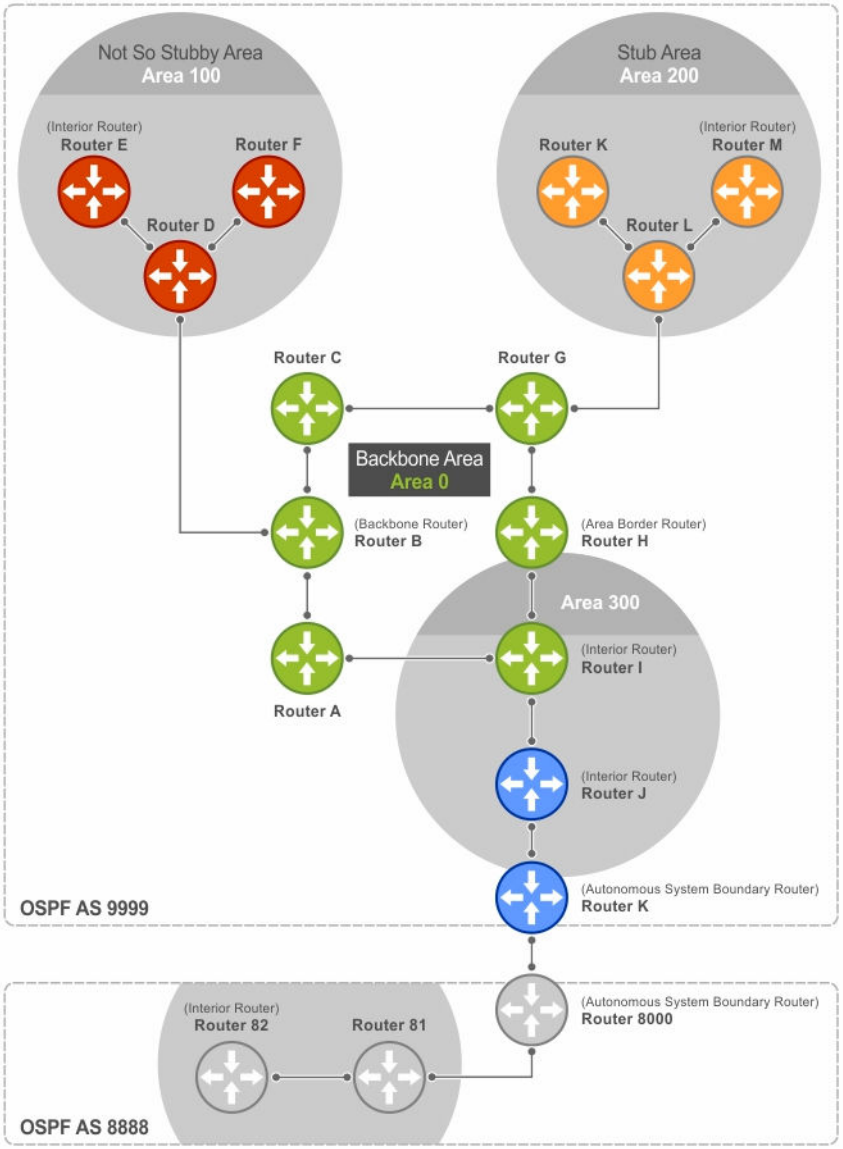


Figure 88. OSPF Routing Examples

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.

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 previous example.

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 the Dell Networking OS. If you do not define DR or BDR in the Dell Networking OS, 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 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: External 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: 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.
- 4: virtual link neighboring router ID.

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
```

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.

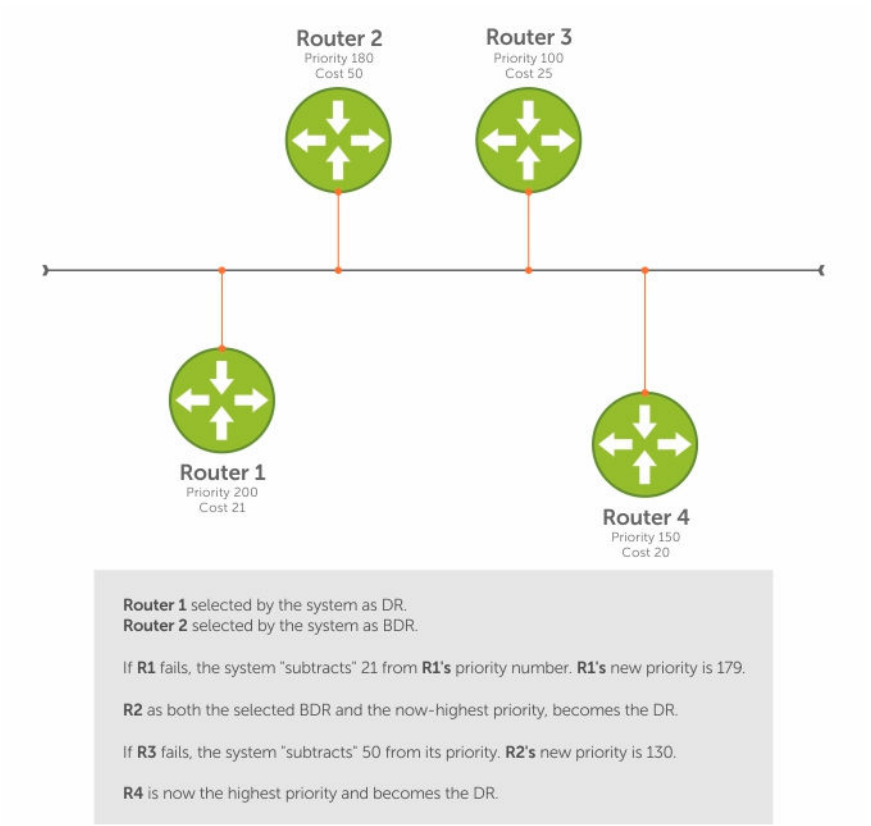


Figure 89. Priority and Cost Examples

OSPF with the Dell Networking OS

The Dell Networking OS supports up to 10,000 OSPF routes for OSPFv2. Within that 10,000 routes, you can designate up to 8,000 routes as external and up to 2,000 as inter/intra area routes.

The Dell Networking OS version 7.8.1.0 and later supports multiple OSPF processes (OSPF MP). The FN IOM switch supports up to 16 processes simultaneously.

On OSPFv3, the system supports only one process at a time for all platforms.

Prior to the Dell Networking OS version 7.8.1.0, the system supported one OSPFv2 and one OSPFv3 process ID per system. OSPFv2 and OSPFv3 can coexist but you must configure them individually.

The Dell 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 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 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 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.

The Dell Networking OS allows you to accept and originate LSAa 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.

Processing SNMP and Sending SNMP Traps

Only the process in `default vrf` can process the SNMP requests and send SNMP traps.

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 Networking and Cisco routers, the hello interval and dead interval must be the same on both routers.
In Dell 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. 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.

For more information regarding this functionality or for assistance, go to http://www.dell.com/support/my-support/.

To ensure equal intervals between the routers, use the following command.

- Manually set the dead interval of the Dell Networking router to match the Cisco configuration.

INTERFACE mode

```
ip ospf dead-interval <x>
```

Example of Setting a Dead Interval

Example of Viewing Dead Interval Setting

In the following example, the dead interval is set at 4x the hello interval (shown in bold).

```
Dell(conf)#int tengig 2/2
Dell(conf-if-te-2/2)#ip ospf hello-interval 20
Dell(conf-if-te-2/2)#ip ospf dead-interval 80

Dell(conf-if-te-2/2)#
```

In the following example, the dead interval is set at 4x the hello interval (shown in bold).

```
Dell (conf-if-te-2/2)#ip ospf dead-interval 20
Dell (conf-if-te-2/2)#do show ip os int tengig 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)
Dell (conf-if-gi-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.

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)

Dell Networking OS supports open shortest path first version 2 (OSPF for IPv4).
The following configuration tasks include two mandatory tasks and several optional tasks:

- Enabling OSPFv2 (mandatory)
- Assigning an OSPFv2 Area (mandatory)
- Enable OSPFv2 on Interfaces
- Configuring Stub Areas
- Configuring LSA Throttling Timers
- Enabling Passive Interfaces

- Enabling Fast-Convergence
- Changing OSPFv2 Parameters on Interfaces
- Enabling OSPFv2 Authentication
- 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`

   Example

```
Dell#
Dell#conf
Dell(conf)#router ospf 1
Dell(conf-router_ospf-1)#timer spf 2 5
Dell(conf-router_ospf-1)#
Dell(conf-router_ospf-1)#show config
!
router ospf 1
timers spf 2 5
Dell(conf-router_ospf-1)#
Dell(conf-router_ospf-1)#end
Dell#
```

For a complete list of the OSPF commands, refer to the *OSPF* section in the *Dell 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.

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`

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.

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:

```
Dell(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 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*

Example of Viewing the Current OSPFv2 Status

```
Dell#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
Dell#
```

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 following example, 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

Example of Viewing Active Interfaces and Assigned Areas

Example of Viewing OSPF Status on a Loopback Interface

```
Dell#(conf)#int tengig 4/44
Dell(conf-if-te-4/44)#ip address 10.10.10.10/24
Dell(conf-if-te-4/44)#no shutdown
Dell(conf-if-te-4/44)#ex
Dell(conf)#router ospf 1
Dell(conf-router_ospf-1)#network 1.2.3.4/24 area 0
Dell(conf-router_ospf-1)#network 10.10.10.10/24 area 1
Dell(conf-router_ospf-1)#network 20.20.20.20/24 area 2
Dell(conf-router_ospf-1)#
Dell#
```

Dell 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.

```
Dell>show ip ospf 1 interface

TenGigabitEthernet 12/17 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 12/21 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)
Dell>
```

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.

```
Dell#show ip ospf 1 int

TenGigabitEthernet 13/23 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.
Dell#
```

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 database database-summary
```

2. Enter CONFIGURATION mode.
   EXEC Privilege mode

   ```
   configure
   ```

3. Enter ROUTER OSPF mode.
   CONFIGURATION mode

   ```
   router ospf process-id
   ```

   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.

Example of the `show ip ospf database database-summary` Command

To view which LSAs are transmitted, use the `show ip ospf database process-id database-summary` command in EXEC Privilege mode.

```
Dell#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
Dell#
```

To view information on areas, use the `show ip ospf process-id` command in EXEC Privilege mode.

Configuring LSA Throttling Timers

Configured LSA timers replace the standard transmit and acceptance times for LSAs.
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 continues to transmit at the max-interval. If the system is stable for twice the maximum interval time, the system reverts to the start-interval timer and the cycle begins again.

1. Specify the interval times for all LSA transmissions.
   CONFIG-ROUTEROSPF- id mode

   ```
   timers throttle lsa all {start-interval | hold-interval | max-interval}
   ```

   - `start-interval`: set the minimum interval between the initial sending and resending the same LSA. The range is from 0 to 600,000 milliseconds.

- `hold-interval`: set the next interval to send the same LSA. This interval is the time between sending the same LSA after the start-interval has been attempted. The range is from 1 to 600,000 milliseconds.
- `max-interval`: set the maximum amount of time the system waits before sending the LSA. The range is from 1 to 600,000 milliseconds.

2. Specify the interval for LSA acceptance.

   CONFIG-ROUTEROSPF- id mode

   ```
   timers throttle lsa arrival arrival-time
   ```

   - *`arrival-time`*: set the interval between receiving the same LSA repeatedly, to allow sufficient time for the system to accept the LSA. The range is from 0 to 600,000 milliseconds.

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 port channel, enter the keywords `port-channel` then a number from 1 to 128.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information (for example, `passive-interface ten 2/3`).
  - For a VLAN, 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.

Example of Viewing Passive Interfaces

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).

```
Dell#show ip ospf 34 int

TenGigabitEthernet 0/0 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 0/1 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.

Example of the `fast-converge` Command

Example of Disabling Fast-Convergence

In the examples below, `Convergence Level` shows the fast-converge parameter setting and `Min LSA origination` shows the LSA parameters (shown in bold).

```
Dell(conf-router_ospf-1)#fast-converge 2
Dell(conf-router_ospf-1)#ex
Dell(conf)#ex
Dell#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
Dell#

Dell#(conf-router_ospf-1)#no fast-converge
Dell#(conf-router_ospf-1)#ex
Dell#(conf)#ex
Dell##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 0 msec, Min LSA arrival 1000 msec
Number of area in this router is 0, normal 0 stub 0 nssa 0
Dell#
```

Changing OSPFv2 Parameters on Interfaces

In the Dell 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.

Example of Changing and Verifying the `cost` Parameter and Viewing Interface Status

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.

```
Dell(conf-if)#ip ospf cost 45
Dell(conf-if)#show config
!
interface TenGigabitEthernet 0/0
  ip address 10.1.2.100 255.255.255.0
  no shutdown
  ip ospf cost 45
Dell(conf-if)#end

Dell#show ip ospf 34 interface
  GigabitEthernet 0/0 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
Dell#
```

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.

  During the `auth-change-wait-time`, OSPF sends out packets with both the new and old authentication schemes.

  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.
For more information, refer to Graceful Restart.

The Dell 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]
   ```

   The system 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.

For more information about OSPF graceful restart, refer to the *Dell Networking OS Command Line Reference Guide*.

Example of the `show run ospf` Command

When you configure a graceful restart on an OSPFv2 router, the `show run ospf` command displays information similar to the following.

```
Dell#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
Dell#
```

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 | ospf <process if> | 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 | rip | ospf <process id> | static} [metric` *`metric-value`* `| metric-type` *`type-value`*`] [route-map` *`map-name`*`] [tag` *`tag-value`*`]`

Configure the following required and optional parameters:

- `bgp, connected, ospf, 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.

Example of Viewing OSPF Configuration after Redistributing Routes

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.

```
Dell(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
Dell(conf-router_ospf)#
```

Troubleshooting OSPFv2

The Dell Networking OS has several tools to make troubleshooting easier.
Be sure to check the following, as these questions represent typical issues that interrupt an OSPFv2 process.

NOTE: The following is not a comprehensive list, just 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.

Example of Viewing OSPF Configuration

```
Dell#show run ospf
!
router ospf 3
!
router ospf 4
  router-id 4.4.4.4
  network 4.4.4.0/28 area 1
!
router ospf 5
!
router ospf 6
!
router ospf 7
  mib-binding
!
router ospf 8
!
router ospf 90
  area 2 virtual-link 4.4.4.4
  area 2 virtual-link 90.90.90.90 retransmit-interval 300
!
ipv6 router ospf 999
```

```
  default-information originate always
  router-id 10.10.10.10
Dell#
```

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.

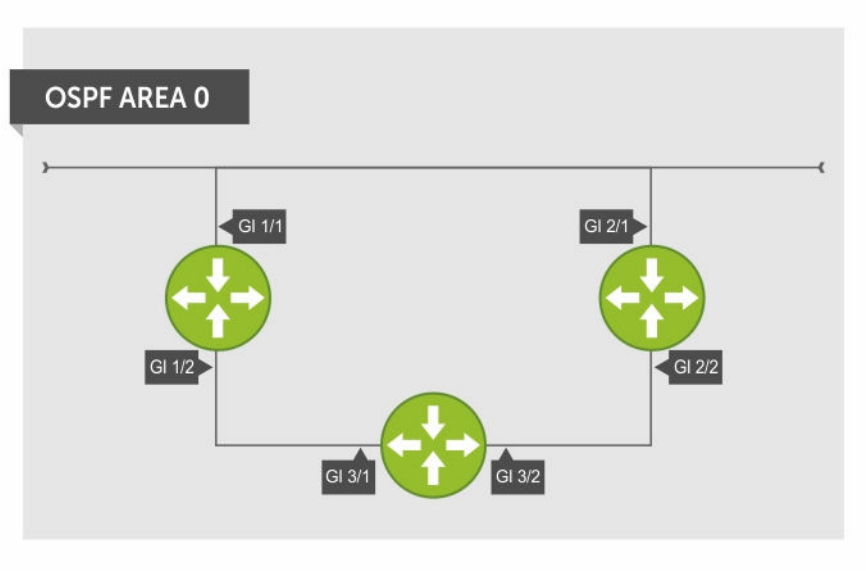


Figure 90. Basic Topology and CLI Commands for OSPFv2

OSPF Area 0 — Gi 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 GigabitEthernet 1/1
  ip address 10.1.11.1/24
  no shutdown
!
interface GigabitEthernet 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 — Gi 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 GigabitEthernet 3/1
  ip address 10.1.13.3/24
  no shutdown
!
interface GigabitEthernet 3/2
  ip address 10.2.13.3/24
  no shutdown
```

OSPF Area 0 — Gl 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 GigabitEthernet 2/1
  ip address 10.2.21.2/24
  no shutdown
!
interface GigabitEthernet 2/2
  ip address 10.2.22.2/24
  no shutdown
```

Configuration Task List for OSPFv3 (OSPF for IPv6)

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.

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`

Example

```
Dell#
Dell#conf
Dell(conf)#ipv6 router ospf 1
Dell(conf-ipv6-router_ospf)#timer spf 2 5
Dell(conf-ipv6-router_ospf)#
Dell(conf-ipv6-router_ospf)#show config
!
ipv6 router ospf 1
timers spf 2 5
Dell(conf-ipv6-router_ospf)#
Dell(conf-ipv6-router_ospf)#end
Dell#
```

NOTE: IPv6 and OSPFv3 do not support Multi-Process OSPF. You can only enable a single OSPFv3 process.

Enabling IPv6 Unicast Routing

To enable IPv6 unicast routing, use the following command.

- Enable IPv6 unicast routing globally.
  CONFIGURATION mode

  ```
  ipv6 unicast routing
  ```

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
  ```

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 {type slot/port}
  ```

  `Interface`: identifies the specific interface that is passive.
  - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information (for example, `passive-interface gi 2/1`).
  - For a port channel, enter the keywords `port-channel` then a number from 1 to 255 (for example, `passive-interface po 100`)
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information (for example, `passive-interface ten 2/3`).
  - For a VLAN, enter the keyword `vlan` then a number from 1 to 4094 (for example, `passive-interface vlan 2222`).

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 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 Dell Networking OS. 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

Dell Networking OS supports graceful restart for OSPFv3.
For more information about graceful restart, refer to Graceful Restart.
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.

- 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 database grace-lsa
  ```

- Display the currently configured OSPFv3 parameters for graceful restart (shown in the following example).

  EXEC Privilege mode

  ```
  show ipv6 ospf database database-summary
  ```

Example of the `show run ospf` Command

Example of the `show ipv6 ospf database database-summary` Command

Example of the `show ipv6 ospf database grace-lsa` Command

```
Dell#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
```

```
Dell#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
```

```
Dell#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 : Gi 5/3
Restart Interval     : 180
Restart Reason       : Switch to Redundant Processor
```

OSPFv3 Authentication Using IPsec

Dell Networking OS supports OSPFv3 authentication using IP security (IPsec).

Starting in Dell Networking OS version 8.4.2.0, 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: The Dell 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 the Dell 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 OSPFv2 (OSPF for IPv4)).

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 OSPFv2 (OSPF for IPv4)).

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 OSPFv2 (OSPF for IPv4)).

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 Port Channel interface, enter the keywords `port-channel number`.
  - For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information.
  - For a VLAN interface, enter the keywords `vlan vlan-id`. The valid VLAN IDs are 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).

Example of the `show crypto ipsec policy` Command

```
Dell#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
```

Example of the `show crypto ipsec sa ipv6` Command

```
Dell#show crypto ipsec sa ipv6

Interface: TenGigabitEthernet 0/0
  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 0/1
  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 Dell Networking OS has several tools to make troubleshooting easier. Consider the following information as these are typical issues that interrupt the OSPFv3 process.

NOTE: The following troubleshooting section is not meant to be a comprehensive list, only 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 virtual links`
- `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 summary
  ```
- View the summary information for the OSPFv3 database.
  EXEC Privilege mode

  ```
  show ipv6 ospf database
  ```
- View the configuration of OSPFv3 neighbors.
  EXEC Privilege mode

  ```
  show ipv6 ospf neighbor
  ```
- View debug messages for all OSPFv3 interfaces.
  EXEC Privilege mode

  ```
  debug ipv6 ospf [event | packet] {type slot/port}
  ```

  - `event`: View OSPF event messages.
  - `packet`: View OSPF packets.
  - For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` then the slot/port information (for example, `passive-interface gi 2/1`).

- For a port channel, enter the keywords `port-channel` then a number from 1 to 255.
- For a 10-Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` then the slot/port information (for example, `passive-interface ten 2/3`).
- For a VLAN, enter the keyword `vlan` then a number from 1 to 4094 (for example, `passive-interface vlan 2222`). The system supports 4094 VLANs.

38

Policy-based Routing (PBR)

Dell Networking OS supports policy-based routing.

This chapter covers the following topics:

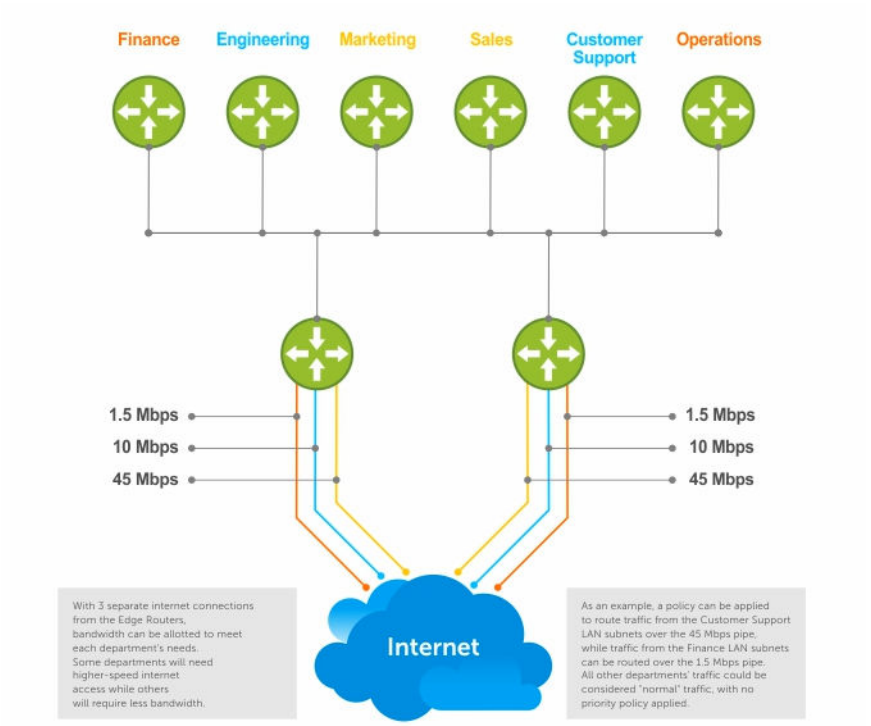
- Overview
- Implementing Policy-based Routing with Dell Networking OS
- Configuration Task List for Policy-based Routing
- Sample Configuration

Overview

Policy-based Routing (PBR) enables you to make routing decisions based on policies applied to a specific interface. When a router receives a packet it normally decides where to forward it 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, etc. For example, a network administrator might want to forward a packet that uses TCP across a different next-hop than packets using ICMP.

Rules for **PBR** can also be a combination of things:

When the packet comes from this source and wants to go to that destination then route it to 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 a PBR, you create a Redirect List. Redirect lists are defined by rules, or routing policies. The following parameters can be defined in the 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

Once a redirect-list is applied to an interface, all traffic passing through it is subjected to the rules defined in the redirect-list.

The traffic is forwarded based on the following:

- Next-hop addresses are verified. If the specified next hop is reachable, then the traffic is forwarded to the specified next-hop.
- If the specified next-hops are not reachable, then the normal routing table is used to forward the traffic.
- Dell Networking OS supports multiple next-hop entries in the redirect lists.
- Redirect-Lists are applied at Ingress.

PBR with Redirect-to-Tunnel Option:

The user can provide a tunnel id for a redirect rule. In this case, the resolved next hop would be the tunnel interface IP. The qualifiers of the rule would be pertaining to the inner IP details. For next hop to be

a tunnel interface user needs to provide tunnel id mandatory. Instead if user provides the tunnel destination IP as next hop, that would be treated as IPv4 next hop and not tunnel next hop.

PBR with Multiple Tacking Option:

Policy based routing with multiple tracking option extends and introduces the capabilities of object tracking to verify the next hop IP address before forwarding the traffic to the next hop. The verification method is made transparent to the user. 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 in this case). It allows 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 reach ability and/or check the Tunnel Interface UP or DOWN status, and then route traffic out to the next-hop and/or Tunnel Interface.

Implementing Policy-based Routing with Dell Networking OS

- 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. Network addresses that are in the middle of a subnet can be included or excluded. Specific bitmasks can be entered using the dotted decimal format.

Non-contiguous bitmask example

```
Dell#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 8/0
```

Hot-Lock PBR

Ingress and egress Hot Lock PBR allow you to add or delete new rules into an existing policy (already written into CAM) without disruption to traffic flow. Existing entries in CAM are adjusted to accommodate the new entries. Hot Lock PBR is enabled by default.

Configuration Task List for Policy-based Routing

To enable the PBR:

- Create a Redirect List
- Create a Rule for a Redirect-list
- Create a Track-id list. For complete tracking information, refer to Object Tracking chapter.
- Apply a Redirect-list to an Interface using a Redirect-group

Create a Redirect List

Use the following command in **CONFIGURATION** mode:

1. Create a redirect list by entering the list name. Format: 16 characters

   CONFIGURATION mode

   ip redirect-list *redirect-list-name*

   Delete the redirect list with the **no ip redirect-list** command.

The following example creates a redirect list by the name of "xyz."

```
Dell(conf)#ip redirect-list ?
WORD   Redirect-list name (max 16 chars)
Dell(conf)#ip redirect-list xyz
```

Create a Rule for a Redirect-list

Use the following command in `CONFIGURATION REDIRECT-LIST` mode to set the rules for the redirect list. 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**.**

1. Configure a rule for the redirect list.

   CONF-REDIRECT-LIST mode

   seq *{number}* **redirect** *{ip-address}{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
   - 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

   Delete a rule with the **no redirect** command.

   The redirect rule supports Non-contiguous bitmasks for PBR in the Destination router IP address

The below step shows a step-by-step example of 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

Creating a Rule Example:

```
Dell(conf-redirect-list)#redirect ?
A.B.C.D                 Forwarding router's address
```

```
Dell(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
Dell(conf-redirect-list)#redirect 3.3.3.3 ip ?
A.B.C.D                 Source address
any                     Any source host
host                    A single source host
Dell(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
Dell(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
Dell(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
Dell(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 /32 77.1.1.1 /32 ?
Dell(conf-redirect-list)#redirect 3.3.3.3 ip 222.1.1.1 /32 77.1.1.1 /32
Dell(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
```

Multiple rules can be applied 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.

Creating multiple rules for a redirect-list:

```
Dell(conf)#ip redirect-list test
Dell(conf-redirect-list)#seq 10 redirect 10.1.1.2 ip 20.1.1.0/24 any
Dell(conf-redirect-list)#seq 15 redirect 10.1.1.3 ip 20.1.1.0/25 any
Dell(conf-redirect-list)#seq 20 redirect 10.1.1.3 ip 20.1.1.128/24 any
Dell(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
Dell(conf-redirect-list)#
```

NOTE: Starting in release 9.4(0.0), Dell Networking OS supports 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. The user can configure multiple recursive routes in a redirect list by entering multiple **seq redirect** statements 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** statement is ignored and the next statement in the list with a different route is used.

PBR Exceptions (Permit)

Use the command **permit** to create an exception to a redirect list. Exceptions are used when a forwarding decision should be based on the routing table rather than a routing policy.

Dell 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. Since the order of rules is important, ensure that you configure any necessary sequence numbers.

The permit statement is never applied because the redirect list covers all source and destination IP addresses.

Ineffective PBR Exception due to Low Sequence Number

```
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 that the permit statement or PBR exception is effective, use a lower sequence number, as shown below:

```
ip redirect-list rcl0
seq 10 permit ip host 3.3.3.3 any
seq 15 redirect 2.2.2.2 ip any any
```

Apply a Redirect-list to an Interface using a Redirect-group

IP redirect lists are supported on physical interfaces as well as 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, on the S-Series, the traffic redirected to the destination port-channel is sometimes switched.

Use the following command in`INTERFACE` mode to apply a redirect list to an interface. 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.

1. Apply a redirect list (policy-based routing) to an interface.

   INTERFACE mode

   ip redirect-group *redirect-list-name*

   redirect-list-name is the name of a redirect list to apply to this interface. FORMAT: up to 16 characters

   Delete the redirect list from this interface with the **[no] ip redirect-group** command.

In this example, the list "xyz" is applied to the tenGigabitEthernet 4/0 interface.

Applying a Redirect-list to an Interface Example:

```
Dell(conf-if-te-4/0)#ip redirect-group xyz
Dell(conf-if-te-4/0)#
```

Applying a Redirect-list to an Interface Example:

```
Dell(conf-if-te-1/0)#ip redirect-group test
Dell(conf-if-te-1/0)#ip redirect-group xyz
Dell(conf-if-te-1/0)#show config
!
interface TenGigabitEthernet 1/0
 no ip address
 ip redirect-group test
 ip redirect-group xyz
 shutdown
Dell(conf-if-te-1/0)#
```

In addition to supporting multiple redirect-lists in a redirect-group, multiple redirect-groups are supported on a single interface. Dell 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 command in EXEC mode:

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 is displayed in dotted format (x.x.x.x). The contiguous mask is displayed in /x format. Some sample outputs are shown below:

```
Dell#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.

```
Dell#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), ARP resolved
Applied interfaces:
 Te 8/0
```

NOTE: If, the redirect-list is applied to an interface, the output of **show ip redirect-list redirect-list-name** command displays reachability and ARP status for the specified next-hop.

Showing CAM PBR Configuration Example :

```
Dell#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

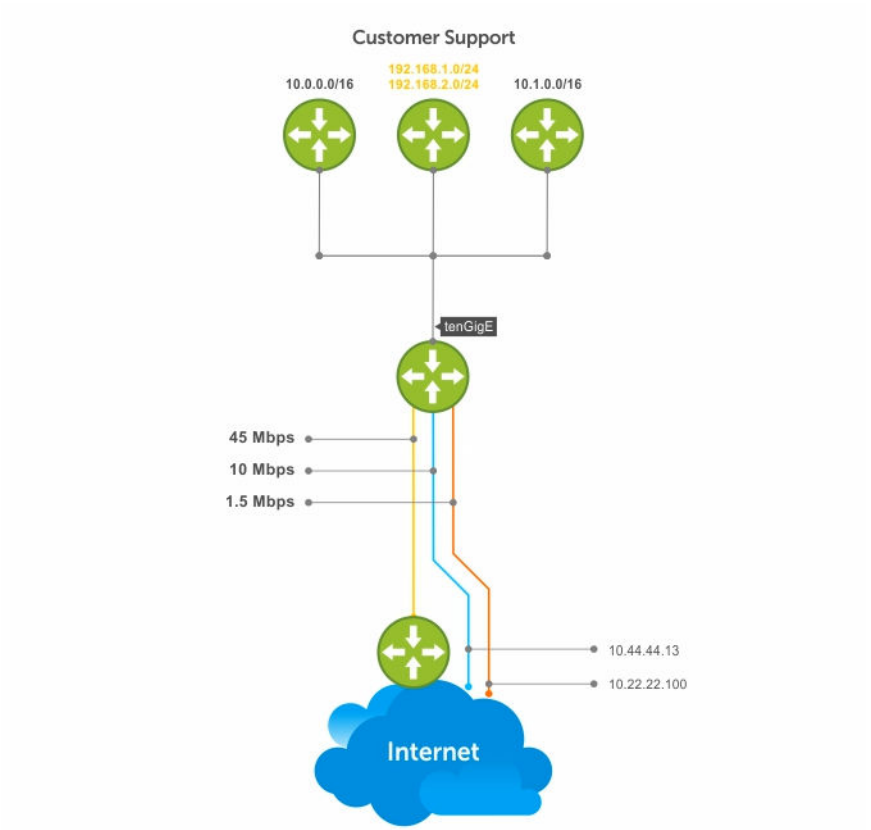
The following configuration is an example for setting up a PBR. These are not comprehensive directions. They are intended to give you a some guidance with typical configurations. You can copy and paste from these examples to your CLI. Be sure you make the necessary changes to support your own IP Addresses, Interfaces, Names, etc.

Graphic illustration of the configuration shown below:

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

PBR Sample Configuration examples are shown below:



Create the Redirect-List GOLD

```
EDGE_ROUTER(conf-if-Te-3/23)#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), ARP resolved
  seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any, Next-hop reachable (via
Te 3/23), ARP resolved
  seq 15 permit ip any any
 Applied interfaces:
  Te 2/11
EDGE_ROUTER#
```

Configuration Tasks for Creating a PBR list using Explicit Track Objects for Redirect IP's

Create Track Objects to track the Redirect IP's:

```
Dell#configure terminal
Dell(conf)#track 3 ip host 42.1.1.2 reachability
Dell(conf-track-3)#probe icmp
Dell(conf-track-3)#track 4 ip host 43.1.1.2 reachability
Dell(conf-track-4)#probe icmp
Dell(conf-track-4)#end
```

Create a Redirect-list with Track Objects pertaining to Redirect-IP's:

```
Dell#configure terminal
Dell(conf)#ip redirect-list redirect_list_with_track
Dell(conf-redirect-list)#redirect 42.1.1.2 track 3 tcp 155.55.2.0/24
222.22.2.0/24
Dell(conf-redirect-list)#redirect 42.1.1.2 track 3 tcp any any
Dell(conf-redirect-list)#redirect 42.1.1.2 track 3 udp 155.55.0.0/16 host
144.144.144.144
Dell(conf-redirect-list)#redirect 42.1.1.2 track 3 udp any host 144.144.144.144
Dell(conf-redirect-list)#redirect 43.1.1.2 track 4 ip host 7.7.7.7 host
144.144.144.144
Dell(conf-redirect-list)#end
```

Verify the Status of the Track Objects (Up/Down):

```
Dell#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:

```
Dell#
Dell(conf)#int TenGigabitEthernet 2/28
Dell(conf-if-te-2/28)#ip redirect-group redirect_list_with_track
Dell(conf-if-te-2/28)#end
```

Verify the Applied Redirect Rules:

```
Dell#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
Dell#
```

Configuration Tasks for Creating a PBR list using Explicit Track Objects for Tunnel Interfaces

Creating steps for Tunnel Interfaces:

```
Dell#configure terminal
Dell(conf)#interface tunnel 1
Dell(conf-if-tu-1)#tunnel destination 40.1.1.2
Dell(conf-if-tu-1)#tunnel source 40.1.1.1
Dell(conf-if-tu-1)#tunnel mode ipip
Dell(conf-if-tu-1)#tunnel keepalive 60.1.1.2
Dell(conf-if-tu-1)#ip address 60.1.1.1/24
Dell(conf-if-tu-1)#ipv6 address 600:10::1/64
Dell(conf-if-tu-1)#no shutdown
Dell(conf-if-tu-1)#end
Dell#

Dell#configure terminal
Dell(conf)#interface tunnel 2
Dell(conf-if-tu-2)#tunnel destination 441:10::2
Dell(conf-if-tu-2)#tunnel source 441:10::1
Dell(conf-if-tu-2)#tunnel mode ipv6
Dell(conf-if-tu-2)#tunnel keepalive 601:10::2
Dell(conf-if-tu-2)#ipv6 address 601:10::1/64
Dell(conf-if-tu-2)#no shutdown
Dell(conf-if-tu-2)#end
Dell#
```

Create Track Objects to track the Tunnel Interfaces:

```
Dell#configure terminal
Dell(conf)#track 1 interface tunnel 1 ip routing
Dell(conf-track-1)#exit
Dell(conf)#track 2 interface tunnel 2 ipv6 routing
Dell(conf-track-2)#end
```

Verify the Status of the Track Objects (Up/Down):

```
Dell#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
Dell#
```

Create a Redirect-list with Track Objects pertaining to Tunnel Interfaces:

```
Dell#configure terminal
Dell(conf)#ip redirect-list explicit_tunnel
Dell(conf-redirect-list)#redirect tunnel 1 track 1 tcp 155.55.2.0/24
222.22.2.0/24
Dell(conf-redirect-list)#redirect tunnel 1 track 1 tcp any any
Dell(conf-redirect-list)#redirect tunnel 1 track 1 udp 155.55.0.0/16 host
144.144.144.144
Dell(conf-redirect-list)#redirect tunnel 2 track 2 tcp 155.55.2.0/24
222.22.2.0/24
Dell(conf-redirect-list)#redirect tunnel 2 track 2 tcp any any
Dell(conf-redirect-list)#end
Dell#
```

Apply the Redirect Rule to an Interface:

```
Dell#configure terminal
Dell(conf)#interface TenGigabitEthernet 2/28
Dell(conf-if-te-2/28)#ip redirect-group  explicit_tunnel
Dell(conf-if-te-2/28)#exit
Dell(conf)#end
```

Verify the Applied Redirect Rules:

```
Dell#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
Dell#
```

39

PIM Sparse-Mode (PIM-SM)

Dell Networking OS supports protocol-independent multicast sparse-mode (PIM-SM).

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.

Implementation Information

Be aware of the following PIM-SM implementation information.

- The Dell Networking implementation of PIM-SM is based on IETF *Internet Draft draft-ietf-pim-sm-v2-new-05*.
- FN IOM supports a maximum of 31 PIM interfaces and 2K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors FN IOM can have.
- 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.
- The Dell Networking operating system (OS) reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.
- The Dell Networking OS supports PIM-SM on physical, virtual local area network (VLAN), and port-channel interfaces.
- The Dell Networking OS supports 2000 IPv6 multicast forwarding entries, with up to 128 PIM-source-specific multicast (SSM) neighbors/interfaces.
- IPv6 Multicast is not supported on synchronous optical network technologies (SONET) interfaces.

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 Networking 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. The Dell 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 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
```

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 multicast routing on the system.
   CONFIGURATION mode

   ```
   ip multicast-routing
   ```
2. Enable PIM-Sparse mode.
   INTERFACE mode

   ```
   ip pim sparse-mode
   ```

Example of Viewing PIM-SM Interfaces

Example of Viewing PIM Neighbors

Example of Viewing the PIM Multicast Routing Table

To display which interfaces are enabled with PIM-SM, use the `show ip pim interface` command from EXEC Privilege mode.

```
Dell#show ip pim interface
Address    Interface VIFindex Ver/  Nbr   Query DR  DR
                             Mode  Count Intvl Prio
189.87.5.6  Gi 4/11  0x2      v2/S  1     30    1   127.87.5.6
189.87.3.2  Gi 4/12  0x3      v2/S  1     30    1   127.87.3.5
189.87.31.6 Gi 7/11  0x0      v2/S  0     30    1   127.87.31.6
```

```
189.87.50.6 Gi 7/13  0x4      v2/S  1     30    1   127.87.50.6
Dell#
```

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` command EXEC Privilege mode.

```
Dell#show ip pim neighbor
Neighbor     Interface Uptime/Expires     Ver  DR
Address                                       Prio/Mode
127.87.5.5   Gi 4/11   01:44:59/00:01:16  v2   1 / S
127.87.3.5   Gi 4/12   01:45:00/00:01:16  v2   1 / DR
127.87.50.5  Gi 7/13   00:03:08/00:01:37  v2   1 / S
Dell#
```

To display the PIM routing table, use the `show ip pim tib` command from EXEC privilege mode.

```
Dell#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: GigabitEthernet 4/12, RPF neighbor 10.87.3.5
  Outgoing interface list:
    GigabitEthernet 4/11
    GigabitEthernet 7/13

(10.87.31.5, 192.1.2.1), uptime 00:01:24, expires 00:02:26, flags: FT
  Incoming interface: GigabitEthernet 7/11, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 4/11
    GigabitEthernet 4/12
    GigabitEthernet 7/13
--More--
```

Configuring S,G Expiry Timers

By default, S, G entries expire in 210 seconds. You can configure a global expiry time (for all [S,G] entries) or configure an expiry time for a particular entry.
If you configure both, the ACL supersedes the global configuration for the specified entries.

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 or to configure the expiry time for a particular (S,G) entry, use the following commands.

1. 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**.

2. Create an extended ACL.
   CONFIGURATION mode

   ```
   ip access-list extended access-list-name
   ```

3. Specify the source and group to which the timer is applied using extended ACLs with permit rules only.
   CONFIG-EXT-NACL mode

   ```
   [seq sequence-number] permit ip source-address/mask | any | host source-address} {destination-address/mask | any | host destination-address}
   ```

4. Set the expiry time for a specific (S,G) entry (as shown in the following example).
   CONFIGURATION mode

   ```
   ip pim sparse-mode sg-expiry-timer seconds sg-list access-list-name
   ```

   The range is from 211 to 86,400 seconds.

   The default is **210**.

Example Configuring an (S,G) Expiry Time

NOTE: The expiry time configuration is nullified and the default global expiry time is used if:

- an ACL is specified in the `ip pim sparse-mode sg-expiry-timer` command, but the ACL has not been created or is a standard ACL.
- if the expiry time is specified for an (S,G) entry in a deny rule.

```
Dell(conf)#ip access-list extended SGtimer
Dell(config-ext-nacl)#permit ip 10.1.2.3/24 225.1.1.0/24
Dell(config-ext-nacl)#permit ip any 232.1.1.0/24
Dell(config-ext-nacl)#permit ip 100.1.1.0/16 any
Dell(config-ext-nacl)#show conf
!
ip access-list extended SGtimer
  seq 5 permit ip 10.1.2.0/24 225.1.1.0/24
  seq 10 permit ip any 232.1.1.0/24
  seq 15 permit ip 100.1.0.0/16 any
Dell(config-ext-nacl)#exit
Dell(conf)#ip pim sparse-mode sg-expiry-timer 1800 sg-list SGtimer
```

To display the expiry time configuration, use the `show running-configuration [acl | pim]` command from EXEC Privilege mode.

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`

Example of Viewing an RP on a Loopback Interface

```
Dell#sh run int loop0
!
```

```
interface Loopback 0
  ip address 1.1.1.1/32
  ip pim sparse-mode
  no shutdown
Dell#sh run pim
!
ip pim rp-address 1.1.1.1 group-address 224.0.0.0/4
```

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 override bootstrap router updates with your static RP configuration.

  `ip pim rp-address`

Example of Viewing the Rendezvous Point (Multicast Group)

Example of Viewing the Rendezvous Point (Multicast Group Range)

To display the assigned RP for a group, use the `show ip pim rp` command from EXEC privilege mode.

```
Dell#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.

```
Dell#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 IP 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 IP 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
  ```

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
  ```

Enabling PIM-SM Graceful Restart

To enable PIM-SM graceful restart, use the following commands.

- Enable PIM-SM graceful restart (non-stop forwarding capability).

  CONFIGURATION mode

  ```
  ip pim graceful-restart nsf
  ```

  - (option) `restart-time`: the time the Dell Networking system requires to restart. The default value is **180 seconds**.
  - (option) `stale-entry-time`: the maximum amount of time that the Dell Networking system preserves entries from a restarting neighbor. The default value is **60 seconds**.
  - (option) `helper-only`: this mode takes precedence over any graceful restart configuration.

40

PIM Source-Specific Mode (PIM-SSM)

Dell Networking OS supports PIM source-specific mode (PIM-SSM).

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 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.

Configure PIM-SMM

Configuring PIM-SSM is a two-step process.

1. Configure PIM-SMM.
2. Enable PIM-SSM for a range of addresses.

Related Configuration Tasks

- Use PIM-SSM with IGMP Version 2 Hosts

Implementation Information

- The Dell Networking implementation of PIM-SSM is based on RFC 3569.
- The Dell Networking operating system (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.

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
   ```

Enabling PIM-SSM

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 the Dell 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, the Dell 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, the system displays an error message. If you apply an extended ACL before you create it, the system accepts the configuration, but when the ACL is later defined, the system 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: Gi 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: Gi 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: Gi 1/2
```

41

Port Monitoring

The Aggregator supports user-configured port monitoring. See `Configuring Port Monitoring` for the configuration commands to use.

Port monitoring copies all incoming or outgoing packets on one 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).

Supported Modes

Standalone, PMUX, VLT, Stacking

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

   Dell(conf-if-te-0/1)#show config
   !
   interface TenGigabitEthernet 0/1
    no shutdown
   ```

2. Create a monitoring session using the command monitor session from CONFIGURATION mode, as shown in the following example.

   CONFIGURATION mode

   ```
   monitor session

   Dell(conf)#monitor session 1
   Dell(conf-mon-sess-1)#source tengig 0/1 destination tengig
   0/8 direction both
   ```

3. Specify the source and destination port and direction of traffic, as shown in the following example.

   MONITOR SESSION mode

   ```
   source

   Dell(conf)#monitor session 1
   Dell(conf-mon-sess-1)#source tengig 0/1 destination tengig
   0/8 direction both
   ```

NOTE: By default, all uplink ports are assigned to port-channel (LAG) 128 and the destination port in a port monitoring session must be an uplink port. When you configure the destination port using the `source` command, the destination port is removed from LAG 128. To display the uplink ports currently assigned to LAG 128, enter the `show lag 128` command.

Example of Viewing Port Monitoring Configuration

To display information on currently configured port-monitoring sessions, use the `show monitor session` command from EXEC Privilege mode.

```
Dell(conf)# monitor session 0
Dell(conf-mon-sess-0)# source tengig 0/1 dest tengig 0/8 direction rx
Dell(conf-mon-sess-0)#exit
Dell(conf)# do show monitor session 0
SessionID Source     Destination   Direction  Mode       Type
--------- ------      -----------  ---------  ----       ----
    0    TenGig 0/1   TenGig 0/8       rx     interface  Port-based
Dell(conf)#
```

In the following example, the host and server are exchanging traffic which passes through the uplink interface 0/1. Port 0/1 is the monitored port and port 0/8 is the destination port, which is configured to only monitor traffic received on tengigabitethernet 0/1 (host-originated traffic).

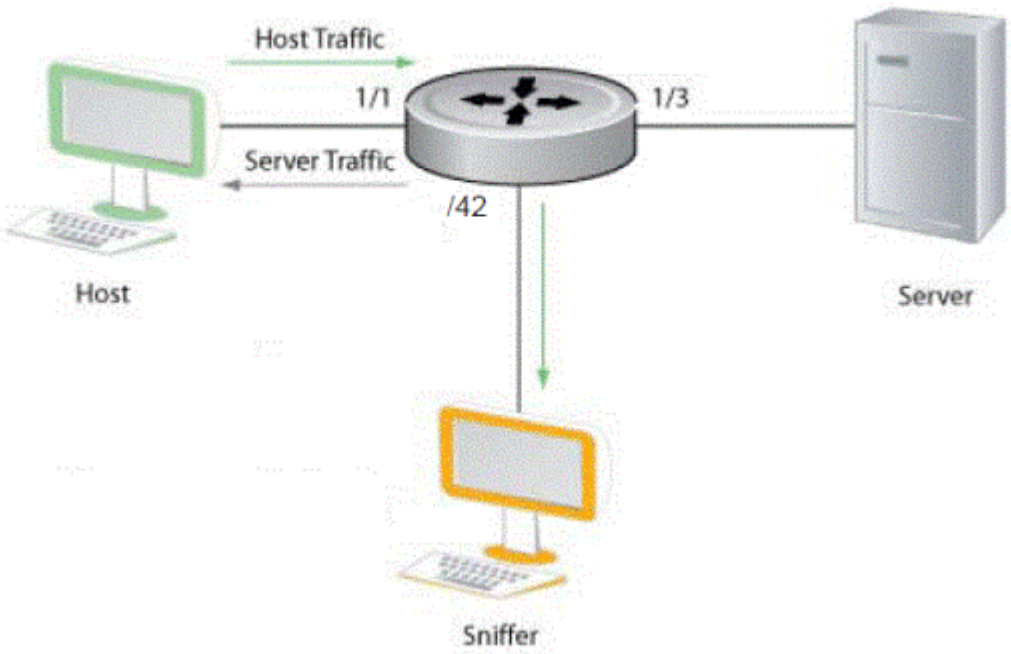


Figure 91. Port Monitoring Example

Important Points to Remember

- Port monitoring is supported on physical ports only; virtual local area network (VLAN) and port-channel interfaces do not support port monitoring.

- The monitored (the source, [MD]) and monitoring ports (the destination, [MG]) must be on the same switch.
- The monitored (source) interface must be a server-facing interface in the format *slot/port*, where the valid slot numbers are 0 and server-facing port numbers are from 1 to 8.
- The destination interface must be an uplink port (ports 9 to 12).
- In general, a monitoring port should have `no ip address` and `no shutdown` as the only configuration; the Dell Networking OS permits a limited set of commands for monitoring ports. You can display these commands using the `?` command.
- A monitoring port may not be a member of a VLAN.
- There may only be one destination port in a monitoring session.
- A source port (MD) can only be monitored by one destination port (MG). If you try to assign a monitored port to more than one monitoring port, the following message displays:

```
Dell(conf)#mon ses 1
Dell(conf-mon-sess-1)#source tengig 0/1 destination tengig 0/8 direction both
Dell(conf-mon-sess-1)#do show monitor session

SessionID   Source       Destination   Direction    Mode       Type
---------   ------       -----------   ---------    ----        ----
      1    TenGig 0/1    TenGig 0/8   both        interface   Port-based

Dell(conf-mon-sess-1)#mon ses 2
Dell(conf-mon-sess-2)#source tengig 0/1 destination tengig 0/8 direction both
% Error: MD port is already being monitored.
```

NOTE: There is no limit to the number of monitoring sessions per system, provided that there are only four destination ports per port-pipe. If each monitoring session has a unique destination port, the maximum number of session is four per port-pipe.

Port Monitoring

The Aggregator supports multiple source-destination statements in a monitor session, but there may only be one destination port in a monitoring session.

There may only be one destination port in a monitoring session (`% Error: Only one MG port is allowed in a session.`).

The number of source ports the Dell Networking OS allows within a port-pipe is equal to the number of physical ports in the port-pipe (n). Multiple source ports may have up to four different destination ports (`Exceeding max MG ports for this MD port pipe.`).

In the following examples, ports 0/1, 0/2, 0/3, and 0/4 all belong to the same port-pipe. These ports mirror traffic to four different destinations (0/9, 0/10, 0/11, and 0/12).

Example of Number of Monitoring Ports

```
Dell#show mon session
SessionID  Source     Destination  Direction Mode       Type
---------  ------     -----------  --------- ----       ----
0        TenGig 0/1   TenGig 0/9     rx     interface  Port-based
10       TenGig 0/2   TenGig 0/10    rx     interface  Port-based
20       TenGig 0/3   TenGig 0/11    rx     interface  Port-based
30       TenGig 0/4   TenGig 0/12    rx     interface  Port-based
Dell(conf)#
```

A source port may only be monitored by one destination port, but a destination port may monitor more than one source port.

Dell Networking OS Behavior: All monitored frames are tagged if the configured monitoring direction is transmit (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 0/1 destination tengig 0/9 direction tx,* if the source port 0/1 is an untagged member of any VLAN, all monitored frames that the destination port 0/9 receives are tagged with the VLAN ID of the source port.

42

Private VLANs (PVLAN)

Dell Networking OS supports private VLAN (PVLAN) feature.

For syntax details about the commands described in this chapter, refer to the Private VLANs commands chapter in the *Dell Networking OS Command Line Reference Guide*.

Private VLANs extend the Dell Networking operating system (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.

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). For more information about port channels, refer to Port Channel Interfaces in the Interfaces chapter.

For an introduction to VLANs, refer to Layer 2.

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 are augmented in the Dell Networking OS version 7.8.1.0 to provide PVLAN data. For more information, refer to the *Dell 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 those 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)

Example of the `switchport mode private-vlan` Command

For interface details, refer to Enabling a Physical Interface in the Interfaces chapter.

NOTE: You cannot add interfaces that are configured as PVLAN host or promiscuous ports to regular VLANs. Conversely, you cannot add "regular" ports (ports not configured as PVLAN ports) to PVLANs as a host or promiscuous member.

The example below shows the `switchport mode private-vlan` command on a port and on a port channel.

```
Dell#conf
Dell(conf)#interface TenGigabitEthernet 2/1
Dell(conf-if-te-2/1)#switchport mode private-vlan promiscuous

Dell(conf)#interface TenGigabitEthernet 2/2
Dell(conf-if-te-2/2)#switchport mode private-vlan host

Dell(conf)#interface TenGigabitEthernet 2/3
Dell(conf-if-te-2/3)#switchport mode private-vlan trunk

Dell(conf)#interface TenGigabitEthernet 2/2
Dell(conf-if-te-2/2)#switchport mode private-vlan host
```

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 are comprised of 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 singly 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.

Example of Configuring Private VLAN Members

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).

```
Dell#conf
Dell(conf)# interface vlan 10
Dell(conf-vlan-10)# private-vlan mode primary
Dell(conf-vlan-10)# private-vlan mapping secondary-vlan 100-101
Dell(conf-vlan-10)# untagged TenGig 2/1
Dell(conf-vlan-10)# tagged TenGig 2/3

Dell(conf)# interface vlan 101
Dell(conf-vlan-101)# private-vlan mode community
Dell(conf-vlan-101)# untagged TenGig 2/10

Dell(conf)# interface vlan 100
Dell(conf-vlan-100)# private-vlan mode isolated
Dell(conf-vlan-100)# untagged Te 2/2
```

Private VLAN Configuration Example

The following example shows a private VLAN topology.

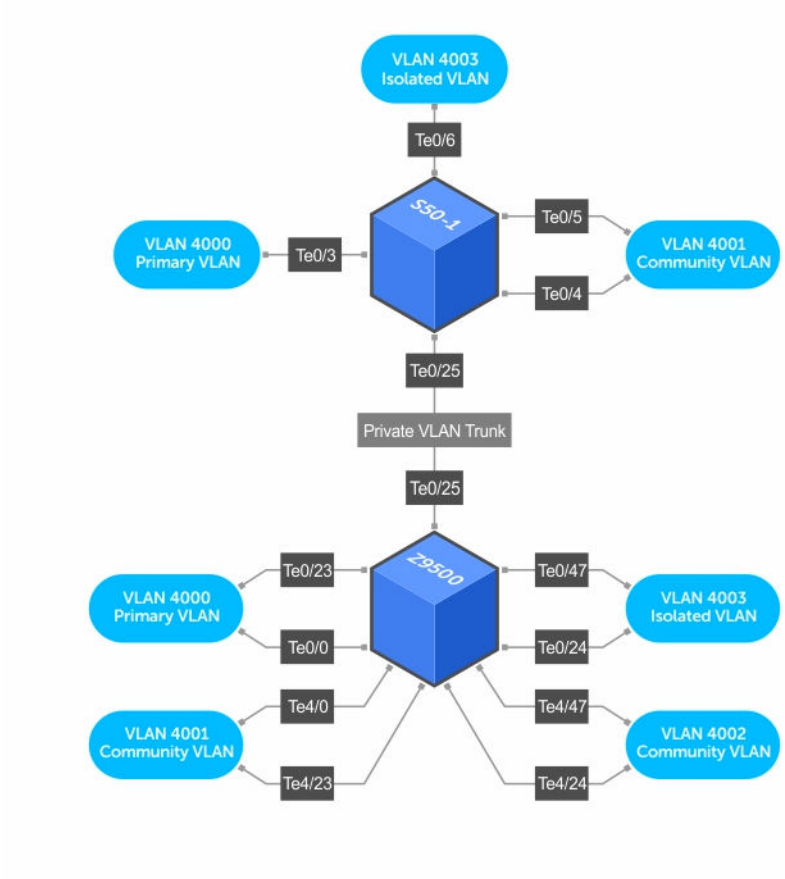


Figure 92. Sample Private VLAN Topology

The following configuration is based on the example diagram for the FN IOM switch:

- TenGig 0/0 and TenGig 0/23 are configured as promiscuous ports, assigned to the primary VLAN, VLAN 4000.
- TenGig 0/25 is configured as a PVLAN trunk port, also assigned to the primary VLAN 4000.
- TenGig 0/24 and TenGig 0/47 are configured as host ports and assigned to the isolated VLAN, VLAN 4003.
- TenGig 4/0 and TenGig 0/23 are configured as host ports and assigned to the community VLAN, VLAN 4001.
- TenGig 4/24 and TenGig 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.

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 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 FN IOM switch in the topology diagram previously shown.
- Display the primary-secondary VLAN mapping. The following example shows the output from the FN IOM switch.
  ```
  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.

Example of Viewing a Private VLAN

```
Dell#show vlan private-vlan
Primary Secondary Type      Active Ports
------- --------- --------- ------ -----------------------------------------
20                Primary   Yes    Te 1/1,5
        30        Community Yes    Te 1/2
```

```
        40       Isolated   Yes    Te 1/3
Dell#
```

Example of 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.

Example of Viewing VLAN Status

```
Dell#show vlan
Codes: * - Default VLAN, G - GVRP VLANs, R - Remote Port Mirroring 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 - VSN tagged
   i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

NUM   Status Description Q Ports
* 1   Active            U Te 5/41
P 20  Active            T Te 1/1,5
C 30  Active            T Te 1/2
I 40  Active            T Te 1/3
Dell#
```

Example of Viewing Private VLAN Configuration

```
Dell#show vlan
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  switchport mode private-vlan promiscuous
  no keepalive
  no shutdown
!
interface TenGigabitEthernet 1/2
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown
!
interface TenGigabitEthernet 1/3
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown
!
interface TenGigabitEthernet 1/5
  no ip address
  switchport
  switchport mode private-vlan trunk
  no shutdown
interface Vlan 20
  private-vlan mode primary
  private-vlan mapping secondary-vlan 30,40
```

```
  no ip address
  tagged TenGigabitEthernet 1/1,5
  shutdown
!
interface Vlan 30
  private-vlan mode community
  no ip address
  tagged TenGigabitEthernet 1/2
  no shutdown
!
```

43

Per-VLAN Spanning Tree Plus (PVST+)

Dell Networking OS supports per-VLAN spanning tree plus (PVST+).

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.

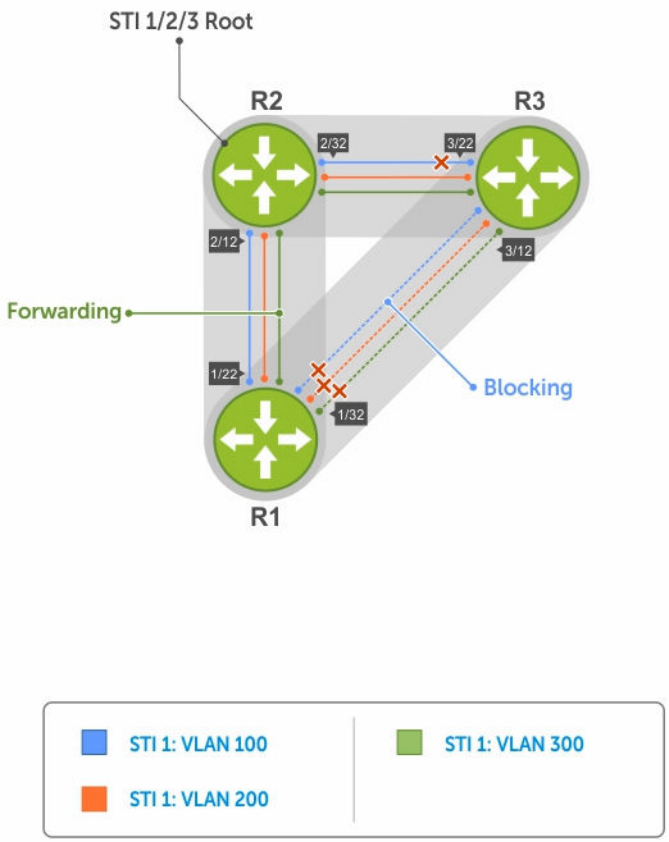


Figure 93. Per-VLAN Spanning Tree

The Dell Networking operating system (OS) supports three other variations of spanning tree, as shown in the following table.

Table 47. Spanning Tree Variations Dell Networking OS Supports

Dell 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 Networking OS implementation of PVST+ is based on IEEE Standard 802.1w.
- The Dell 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 Networking systems in a multivendor network, verify that the costs are values you intended.
- You can enable PVST+ on 254 VLANs.

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
- Enable BPDU Filtering Globally
- Configuring an EdgePort
- Flush MAC Addresses after a Topology Change
- Prevent Network Disruptions with BPDU Guard
- SNMP Traps for Root Elections and Topology Changes
- PVST+ in Multi-Vendor Networks
- Enabling PVST+ Extend System ID
- PVST+ Sample Configurations

Enabling PVST+

When you enable PVST+, the Dell 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
  ```

Example of Viewing PVST+ Configuration

To display your PVST+ configuration, use the `show config` command from PROTOCOL PVST mode.

```
Dell(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.

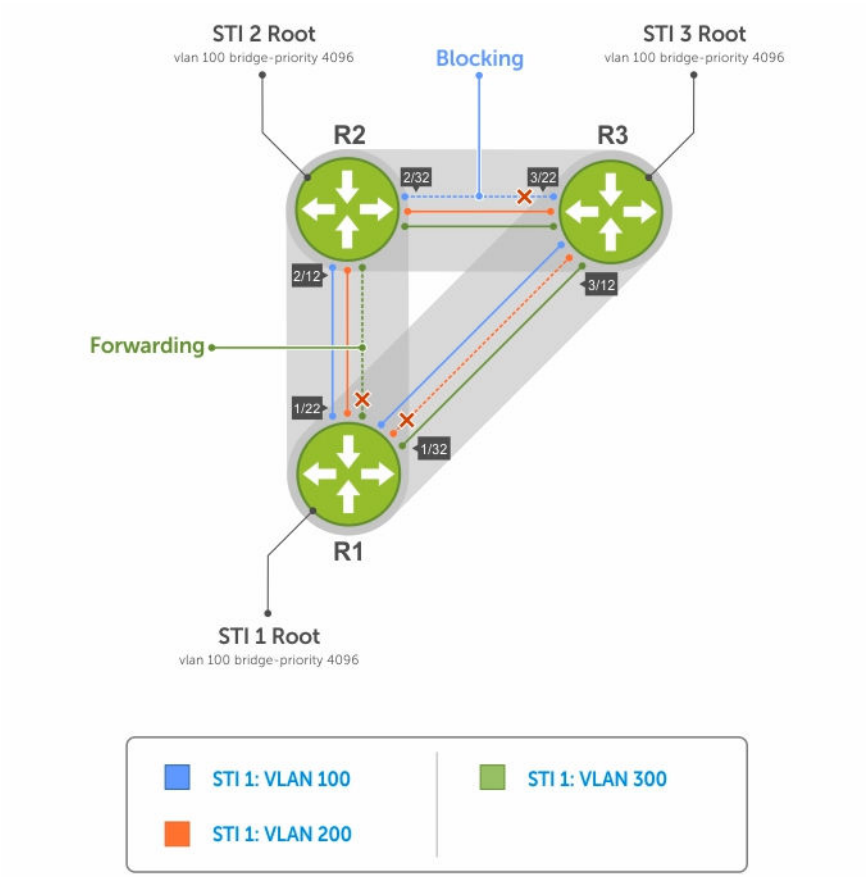


Figure 94. Load Balancing with PVST+

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**.

Example of the `show spanning-tree pvst vlan` Command

To display the PVST+ forwarding topology, use the `show spanning-tree pvst [vlan vlan-id]` command from EXEC Privilege mode.

```
Dell(conf-if-te-5/41)#do show spanning-tree pvst vlan 2
VLAN 2
Root Identifier has priority 32768, Address 001e.c9f1.00f3
```

```
Root Bridge hello time 2, max age 20, forward delay 15
Bridge Identifier has priority 32768, Address 001e.c9f1.00f3
Configured hello time 2, max age 20, forward delay 15
Bpdu filter disabled globally
We are the root of VLAN 2
Current root has priority 32768, Address 001e.c9f1.00f3
Number of topology changes 2, last change occured 00:14:39 ago on Po 23

Port 24 (Port-channel 23) is designated Forwarding
Port path cost 1600, Port priority 128, Port Identifier 128.24
Designated root has priority 32768, address 001e.c9f1.00:f3
Designated bridge has priority 32768, address 001e.c9f1.00:f3
Designated port id is 128.24 , designated path cost 0
Number of transitions to forwarding state 1
BPDU sent 449, received 0
The port is not in the Edge port mode, bpdu filter is disabled

Port 450 (TenGigabitEthernet 5/41) is disabled Discarding
Port path cost 2000, Port priority 128, Port Identifier 128.450
Designated root has priority 32768, address 001e.c9f1.00:f3
Designated bridge has priority 32768, address 001e.c9f1.00:f3
Designated port id is 128.450 , designated path cost 0
Number of transitions to forwarding state 0
BPDU sent 0, received 0
The port is not in the Edge port mode, bpdu filter is disabled

Port 459 (TenGigabitEthernet 5/50) is designated Forwarding
Port path cost 2000, Port priority 128, Port Identifier 128.459
Designated root has priority 32768, address 001e.c9f1.00:f3
Designated bridge has priority 32768, address 001e.c9f1.00:f3
Designated port id is 128.459 , designated path cost 0
Number of transitions to forwarding state 1
BPDU sent 457, received 0
The port is not in the Edge port mode, bpdu filter is disable
```

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 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 48. Default Values for Port Cost

Port Cost	Default Value
1000-Mb/s Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
40-Gigabit Ethernet interfaces	1400
Port Channel with one 10-Gigabit Ethernet interface	2000
Port Channel with one 40-Gigabit Ethernet interface	1400
Port Channel with two 10-Gigabit Ethernet interfaces	1800
Port Channel with two 40-Gigabit Ethernet interfaces	600

NOTE: The Dell 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 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 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 Networking systems which have hybrid ports participating in PVST+ transmit two kinds of BPDUs: an 802.1D BPDU and an untagged PVST+ BPDU.

Dell Networking systems do not expect PVST+ BPDU (tagged or untagged) on an untagged port. If this situation occurs, the Dell Networking OS places the port in an Error-Disable state. This behavior might result in the network not converging. To prevent the system 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 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.

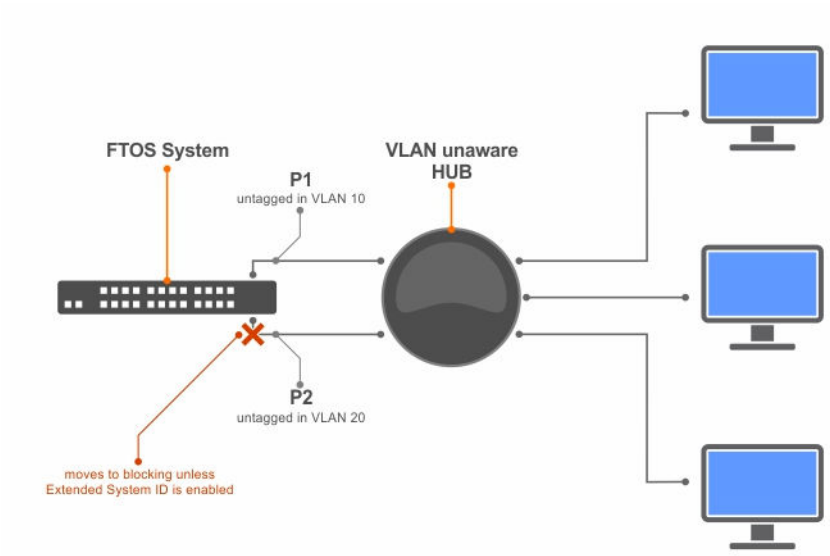


Figure 95. PVST+ with Extend System ID

- Augment the bridge ID with the VLAN ID.
  PROTOCOL PVST mode

```
extend system-id
```

Example of Viewing the Extend System ID in a PVST+ Configuration

```
Dell(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)

Example of PVST+ Configuration (R2)

Example of PVST+ Configuration (R3)

```
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

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

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
```

Enable BPDU Filtering globally

The enabling of BPDU Filtering stops transmitting of BPDUs on the operational port fast enabled ports by default.
When BPDUs are received, the spanning tree is automatically prepared. By default global bpdu filtering is disabled.

Enable BPDU Filter globally to filter transmission of BPDU port fast enabled interfaces.

PROTOCOL PVST mode

```
edge-port bpdu filter default
```

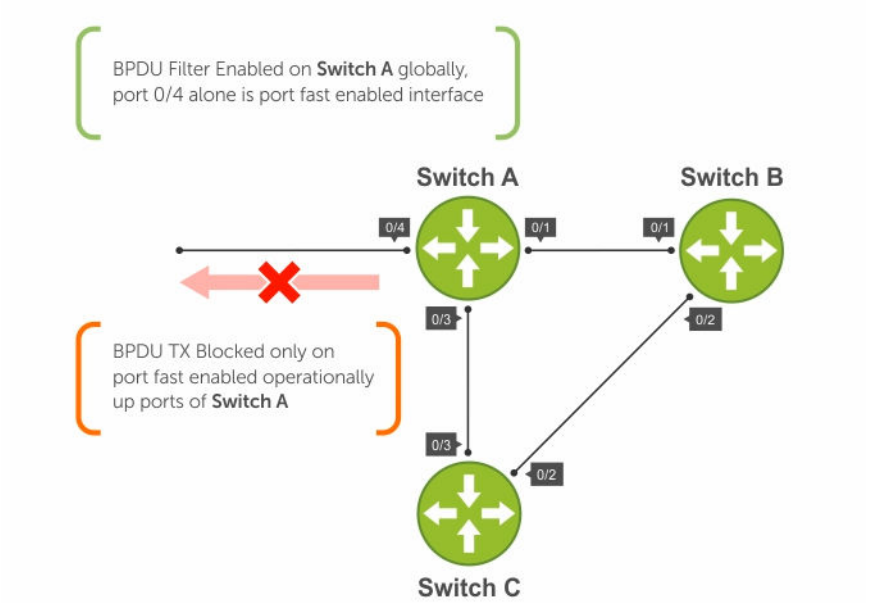


Figure 96. BPDU Filtering Enabled Globally

44

Quality of Service (QoS)

Dell Networking OS supports quality of service (QoS).

Differentiated service is accomplished by classifying and queuing traffic, and assigning priorities to those queues.

The switch traffic has four data queues per port. All queues are serviced using the Weighted Round Robin scheduling algorithm. You can only manage prioritize queuing on egress.

NOTE: When you enable DCB, the egress QoS features in the output QoS policy-map (such as service-class bandwidth-percentage and bandwidth-percentage), the default bandwidth allocation ratio for egress queues and strict-priority may not work as intended.This is to provide compatibility with DCBX. Hence, it is recommended to have the DCB disabled when you wish to apply these features exclusively.

Table 49. Dell Networking OS Support for Port-Based, Policy-Based, and Multicast QoS 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

Feature	Direction
Create Policy Maps	Ingress + Egress
Create Input Policy Maps	Ingress
Honor DSCP Values on Ingress Packets	Ingress
Honoring dot1p Values on Ingress Packets	Ingress
Create Output Policy Maps	Egress
Specify an Aggregate QoS Policy	Egress
QoS Rate Adjustment	
Strict-Priority Queueing	
Weighted Random Early Detection	Egress
Create WRED Profiles	Egress

Figure 97. Dell Networking QoS Architecture

Implementation Information

The Dell 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
- Configuring Port-Based Rate Policing
- Configuring Port-Based Rate Shaping

Setting dot1p Priorities for Incoming Traffic

Change the priority of incoming traffic on the interface using the dot1p-priority command from INTERFACE mode. The Dell 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 50. 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`

Example of Configuring a dot1p Priority on an Interface

NOTE: The `dot1p-priority` command marks all incoming traffic on an interface with a specified dot1p priority and maps all incoming traffic to the corresponding queue. When you enable PFC and/or ETS on an interface, incoming traffic with a specified dot1p priority can be distributed across different queues. Therefore, when you use PFC and ETS to manage data center traffic, it is not recommended that you use the `dot1p-priority` command to set a queue assignment. For more information, refer to Data Center Bridging (DCB).

```
Dell#config
Dell(conf)#interface tengigabitethernet 1/0
Dell(conf-if)#switchport
Dell(conf-if)#dot1p-priority 1
Dell(conf-if)#end
Dell#
```

Honoring dot1p Priorities on Ingress Traffic

By default, the Dell 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
  ```

Example of Configuring an Interface to Honor dot1p Priorities on Ingress Traffic

```
Dell#config t
Dell(conf)#interface tengigabitethernet 1/0
Dell(conf-if)#service-class dynamic dot1p
Dell(conf-if)#end
Dell#
```

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 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
  ```

Example of the `rate police` Command

```
Dell#config t
Dell(conf)#interface tengigabitethernet 1/0
Dell(conf-if)#rate police 100 40 peak 150 50
Dell(conf-if)#end
Dell#
```

Configuring Port-Based Rate Shaping

Rate shaping buffers, rather than drops, traffic that exceeds 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
  ```

Example of `rate—shape` Command

```
Dell#config
Dell(conf)#interface tengigabitethernet 1/0
Dell(conf-if)#rate shape 500 50
Dell(conf-if)#end
Dell#
```

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 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 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 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 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 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 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 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 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
```

Policy-Based QoS Configurations

Policy-based QoS configurations consist of the components shown in the following example.

Figure 98. Constructing Policy-Based QoS Configurations

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 **te 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`

```
Dell(conf)# qos dscp-color-map bat-enclave-map
Dell(conf-dscp-color-map)# dscp yellow 9,10,11,13,15,16
Dell (conf-dscp-color-map)# exit
```

Assign the color map, **bat-enclave-map** to interface .

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.

```
Dell# 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.

```
Dell# 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.

Display summary information about a color policy for a specific interface.

Display detailed information about a color policy for a specific interface

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, the Dell 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

   ```
   match ip
   ```

   After you create a class-map, the Dell 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.
4. Link the class-map to a queue.
   POLICY MAP mode

   ```
   service-queue
   ```

Example of Creating a Layer 3 Class Map

```
Dell(conf)#ip access-list standard acl1
Dell(conf-std-nacl)#permit 20.0.0.0/8
Dell(conf-std-nacl)#exit
Dell(conf)#ip access-list standard acl2
Dell(conf-std-nacl)#permit 20.1.1.0/24 order 0
Dell(conf-std-nacl)#exit
Dell(conf)#class-map match-all cmap1
Dell(conf-class-map)#match ip access-group acl1
Dell(conf-class-map)#exit
Dell(conf)#class-map match-all cmap2
Dell(conf-class-map)#match ip access-group acl2
Dell(conf-class-map)#exit
Dell(conf)#policy-map-input pmap
```

```
Dell(conf-policy-map-in)#service-queue 0 to 3 class-map cmap1
Dell(conf-policy-map-in)#service-queue 1 class-map cmap2
Dell(conf-policy-map-in)#exit
Dell(conf)#interface tengig 1/0
Dell(conf-if-te-1/0)#service-policy input pmap
```

Examples f Creating a Layer 3 IPv6 Class Map

The following example matches IPv6 traffic with a DSCP value of 40.

```
Dell(conf)# class-map match-all test
Dell(conf-class-map)# match ipv6 dscp 40
```

The following example matches IPv4 and IPv6 traffic with a precedence value of 3.

```
Dell(conf)# class-map match-any test1
Dell(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 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

   ```
   match mac
   ```

   After you create a class-map, the system 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, the system 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 4, though you intended for these packets to match positive against *cmap2* and be buffered in queue 1.

In cases such as these, where class-maps with overlapping ACL rules are applied to different queues, use the keyword `order`. The Dell 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**.

Setting DSCP Values for Egress Packets Based on Flow

Match-any Layer 3 flows may have several match criteria. All flows that match at least one of the match criteria are mapped to the same queue because they are in the same class map.
Setting a DSCP value from QOS-POLICY-IN mode (refer to Setting a DSCP Value for Egress Packets) assigns the same DSCP value to all of the matching flows in the class-map.
The flow-based DSCP marking feature allows you to assign *different* DSCP to each match criteria.

- Create matching flows within a class map that have different DSCP values.
  CLASS MAP mode

  ```
  match set-ip-dscp
  ```

  The values you set from CLASS-MAP mode override the value you set in the QoS input policy DSCP value. Packets matching the rule are marked with the specified value.

Example of Marking Flows in the Same Queue with Different DSCP Values

```
Dell#show run class-map
!
class-map match-any example-flowbased-dscp
  match ip access-group test set-ip-dscp 2
  match ip access-group test1 set-ip-dscp 4
  match ip precedence 7 set-ip-dscp 1

Dell#show run qos-policy-input
!
qos-policy-input flowbased
  set ip-dscp 3
```

Displaying Configured Class Maps and Match Criteria

To display all class-maps or a specific class map, use the following command.
Dell 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
```

Examples of Traffic Classifications

The following example shows incorrect traffic classifications.

```
Dell#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
Dell#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

Dell#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

Dell#show cam layer3-qos interface tengigabitethernet 2/49
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.

```
Dell#show cam layer3-qos interface tengigabitethernet 2/49
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. There are two types of input QoS policies: Layer 3 and Layer 2.

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 rule 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 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 DSCP Value for Egress Packets

   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 DSCP Value for Egress Packets

Set the DSCP value for egress packets based on ingress QOS classification.
The 6 bits that are used for DSCP are also used to identify the queue in which traffic is buffered.

Example of Setting a DSCP Value for Egress Packets

```
Dell#config
Dell(conf)#qos-policy-input my-input-qos-policy
Dell(conf-qos-policy-in)#set ip-dscp 34

Dell(conf-qos-policy-in)#show config
!
qos-policy-input my-input-qos-policy
  set ip-dscp 34
Dell(conf-qos-policy-in)#end
Dell#
```

Setting a dot1p Value for Egress Packets

To set a dot1p value for egress packets, use the following command.

- Set a 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:
   Configuring Policy-Based Rate Shaping

   Allocating Bandwidth to Queue

   Configure a Scheduler 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 Dell Networking recommends pre-calculating your bandwidth requirements before creating them. Make sure you apply the QoS policy to all the four queues and that the sum of the bandwidths allocated through them is exactly 100.

When you apply the QoS policies through output policy map and if the sum of the bandwidth percentages configured is below or above 100, the actual bandwidth is allocated proportionally. If the sum of allocated bandwidth is less than 100, the unused bandwidth is allotted to un-allocated queues. If

the sums of allocated bandwidth exceed 100, then 1% of the bandwidth is derived for unassigned queues from assigned queues.

- Allocate bandwidth to queues.
  QOS-POLICY-OUT mode

  ```
  bandwidth-percentage
  ```

Configure a Scheduler to Queue

By default, the switch schedules packets for egress based on weighted round robin (WRR).

NOTE: The bandwidth and scheduler cannot be configured at the same time. Policy-level scheduler assigned to queue is applied to both unicast and multicast traffic.

Setting DSCP Values for Egress Packets Based on Flow

Match-any Layer 3 flows may have several match criteria. All flows that match at least one of the match criteria are mapped to the same queue because they are in the same class map.
Setting a DSCP value from QOS-POLICY-IN mode (refer to Setting a DSCP Value for Egress Packets) assigns the same DSCP value to all of the matching flows in the class-map.

The flow-based DSCP marking feature allows you to assign *different* DSCP to each match criteria.

- Create matching flows within a class map that have different DSCP values.
  CLASS MAP mode

  ```
  match ip dscp 0-63 set-ip-dscp 0-63
  ```

  The values you set from CLASS-MAP mode override the value you set in the QoS input policy DSCP value. Packets matching the rule are marked with the specified value.

Example of Marking Flows in the Same Queue with Different DSCP Values

```
Dell#show run class-map
!
class-map match-any example-flowbased-dscp
  match ip access-group test set-ip-dscp 2
  match ip access-group test1 set-ip-dscp 4
  match ip precedence 7 set-ip-dscp 1

Dell#show run qos-policy-input
!
qos-policy-input flowbased
  set ip-dscp 3

Dell#
```

Specifying WRED Drop Precedence

- 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

   Enabling Fall Back to Trust Diffserve or dot1p
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-aggregate
  ```

Honoring DSCP Values on Ingress Packets

The Dell 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 the Dell 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` command.

NOTE: Packets with DSCP value of 63 are dropped.

Table 51. 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

The Dell 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 52. Default dot1p to Queue Mapping

dot1p	Queue ID
0	0
1	0
2	0
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
  ```

Enabling Fall Back to Trust Diffserve or dot1p

When using QoS service policies with multiple class maps, you can configure the Dell Networking OS to use the incoming DSCP or dot1p marking as a secondary option for packet queuing if no match occurs in the class maps.

When you use class-maps, traffic is matched against each class-map sequentially from first to last. The sequence is based on the priority of the rules, as follows:

1. Rules with lowest priority, or in the absence of a priority configuration.
2. Rules of the next numerically higher queue.

By default, if no match occurs, the packet is queued to the default queue, Queue 0.

In the following configuration, packets are classified to queues using three class maps:

Example of Viewing Packet Classes Based on DSCP Value

```
!
policy-map-input input-policy
  service-queue 1 class-map qos-BE1
  service-queue 3 class-map qos-AF3
  service-queue 4 class-map qos-AF4
  trust diffserv fallback
!
class-map match-any qos-AF3
  match ip dscp 24
  match ip access-group qos-AF3-ACL
!
class-map match-any qos-AF4
  match ip dscp 32
  match ip access-group qos-AF4-ACL
!
class-map match-all qos-BE1
  match ip dscp 0
  match ip access-group qos-BE1-ACL
```

The packet classification logic for the configuration shown is as follows:

1. Match packets against `match-any qos-AF4`. If a match exists, queue the packet as AF4 in Queue 4, and if no match exists, go to the next class map.
2. Match packets against `match-any qos-AF3`. If a match exists, queue the packet as AF3 in Queue 3, and if no match exists, go to the next class map.
3. Match packets against `match-all qos-BE1`. If a match exists, queue the packet as BE1, and if no match exists, queue the packets to the default queue, Queue 0.
4. You can optionally classify packets using their DSCP marking, instead of placing packets in Queue 0, if no match occurs. In the above configuration, if no match occurs against `match-all qos-BE1`, the classification logic continues.
5. Queue the packet according to the DSCP marking. The DSCP to Queue mapping is as noted in Honoring dot1p Values on Ingress Packets.

The behavior is similar for `trust dot1p fallback` in a Layer2 input policy map; the dot1p-to-queue mapping is noted in Honoring dot1p Values on Ingress Packets.

To enable fall back to trust diffserve or dot1p, use the following command.

- Classify packets according to their DSCP value as a secondary option in case no match occurs against the configured class maps.

POLICY-MAP-IN mode

```
trust {diffserve | dot1p} fallback
```

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 an input Layer 2 QoS policy on an interface you also configure with the `vlan-stack access` command.
- If you apply a service policy that contains an ACL to more than one interface, the system 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; in this case, INTERFACE mode must be in Switchport mode.

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.

Enabling QoS Rate Adjustment

By default, while rate limiting, policing, and shaping, the Dell 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, and `no qos-rate-adjust` is listed in the running-configuration

- Include a specified 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, enter `qos-rate-adjust 8`. For variable length overhead fields, know the number of bytes you want to include.

  The default is disabled.

Enabling Strict-Priority Queueing

Strict-priority means that the Dell Networking OS de-queues all packets from the assigned queue before servicing any other queues.

- The `strict-priority` supersedes `bandwidth-percentage` and `bandwidth-weight percentage` configurations.
- A queue with strict priority can starve other queues in the same port-pipe.
- If more than two strict priority queues are configured, the strict priority queue with a higher queue number is scheduled first.
- Assign strict priority to one unicast queue.
  CONFIGURATION mode

  `strict-priority`

  The range is from 1 to 3.

Weighted Random Early Detection

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.

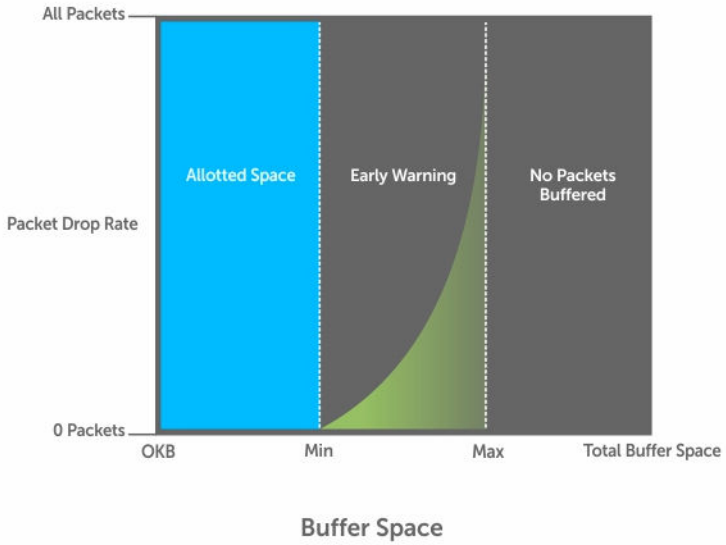


Figure 99. Packet Drop Rate for WRED

You can create a custom WRED profile or use one of the five pre-defined profiles.

Table 53. 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

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 the system should apply the profile. The Dell 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 Networking uses the first 3 bits of this field (DP) to determine the drop precedence.

- DP values of 110, 100, and 101 map to yellow; all other values map to green.
- If you do not configure the system 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
  ```

Displaying WRED Profiles

```
Dell#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
0
Dell#
```

Displaying WRED Drop Statistics

To display WRED drop statistics, use the following command.

- Display the number of packets the system the WRED profile drops.
  EXEC Privilege mode

  ```
  show qos statistics
  ```

Example of the `show qos statistics wred-profile` Command

```
Dell#show qos statistics wred-profile

Interface Te 0/20

Drop-statistic Dropped Pkts
Green          11234
Yellow         12484
Out of Profile 0

Dell#
```

Displaying egress-queue Statistics

To display egress-queue statistics of both transmitted and dropped packets and bytes, 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
```

Example of the `show qos statistics egress-queue` Command

```
Dell#show qos statistics egress-queue

Interface Te 1/1
Unicast/Multicast Egress Queue Statistics
Queue# Q# Type          TxPkts          TxBytes    DroppedPkts     DroppedBytes
-------------------------------------------------------------------------------
                                                                             --
  0      UCAST              0                0              0                0
  1      UCAST              0                0              0                0
  2      UCAST              0                0              0                0
  3      UCAST              0                0              0                0
  4      UCAST              0                0              0                0
  5      UCAST              0                0              0                0
  6      UCAST              0                0              0                0
  7      UCAST              0                0              0                0
  8      UCAST            204            13056              0                0
  9      MCAST              0                0              0                0
 10      MCAST              0                0              0                0
 11      MCAST              0                0              0                0
 12      MCAST              0                0              0                0
 13      MCAST              0                0              0                0
 14      MCAST              0                0              0                0
 15      MCAST              0                0              0                0
 16      MCAST              0                0              0                0
 17      MCAST              0                0              0                0
Dell#
```

Classifying Layer 2 Traffic on Layer 3 Interfaces

To process Layer 3 packets that contain Dot1p — (IEEE 802.1p) Packet classification (Layer 2 headers), configure VLAN tags on a physical Layer 3 interface (that is configured with an IP address and is not associated with any VLAN). You can also configure a VLAN subinterface over a physical underlying interface and classify packets using the dot1p value.

To apply an input policy map to Layer 3 physical interfaces, use the `service-policy input policy-name layer 2` command in Interface Configuration mode.

To apply a Layer 2 policy on Layer 3 interfaces, perform the following:

1. Configure an interface with an IP address or a VLAN subinterface
   CONFIGURATION mode

   ```
   Dell(conf)# int fo 0/0
   ```

   INTERFACE mode

   ```
   Dell(conf-if-fo-0/0)# ip address 90.1.1.1/16
   ```

2. Configure the Layer 2 policy with Layer 2 (Dot1p or source MAC-based) classification rules.
   CONFIGURATION mode

   ```
   Dell(conf)# policy-map-input l2p layer2
   ```

3. Apply the Layer 2 policy on the Layer 3 interface.
   INTERFACE mode

   ```
   Dell(conf-if-fo-0/0)# service-policy input l2p layer2
   ```

Classifying Packets Based on a Combination of DSCP Code Points and VLAN IDs

You can configure a classifier map, which contains both the Differentiated Services Code Point (DSCP) and MAC VLAN IDs as parameters, for filtering packets that are received before they are forwarded or dropped. You can now specify both DSCP-IP packet classification (Layer 3 headers) and Dot1p—(IEEE 802.1p) Packet classification (Layer 2 headers) as match criteria in a Layer 3 class map.

The type of the class map is determined during the creation of a class map. In releases of Dell Networking OS earlier than Release 9.2(0.0), you can configure only the dot1p value as the filter criterion in Layer 2 class maps and the DSCP value as the filter parameter in Layer 3 class maps. It was also possible to classify packets using both the Layer 2 attribute, dot1p value or MAC VLAN, in a Layer 2 class map and the Layer 3 attribute, DSCP value, in a Layer 3 class map. However, it was not possible to configure both dot1p or MAC VLAN, and DSCP values in the same Layer 2 or Layer 3 class map.

All class maps are Layer 3 by default. You can now configure a Layer 3 class map to differentiate traffic according to the IP VLAN value and the DSCP value. You can use the `match ip vlan vlan-id` command in Class Map Input Configuration mode to specify a match criterion for a class map based on a VLAN ID. You can attach this class map with a policy map, and associate the policy map with a service queue. When you link class maps to queues using the `service-queue` command, Dell Networking OS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities).

To create IP VLAN and DSCP values as match criteria in a Layer 3 class map, and to associate the class map with a policy map that is linked to a service queue, perform the following:

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 to be a member of the class. 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. Use a DSCP value as a match criterion.
   CLASS-MAP mode

   ```
   Dell (conf-class-map)#match ip dscp 5
   ```
3. Configure a match criterion for a class map based on the VLAN ID.
   CLASS-MAP mode

   ```
   Dell (conf-class-map)#match ip vlan 5
   ```
4. Create a QoS input policy on the device.
   CONFIGURATION mode

   ```
   Dell(conf)#qos-policy-input pp_qospolicy
   ```
5. Specify the DSCP value to be set on the matched traffic.
   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
```

45

Routing Information Protocol (RIP)

The routing information protocol (RIP) is based on a distance-vector algorithm and tracks distances or hop counts to nearby routers when establishing network connections.

RIP protocol standards are listed in the Standards Compliance chapter.

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

The Dell 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 the system.

Table 54. 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 the system.

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 the 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 Networking OS Command Reference Interface Guide*.

Enabling RIP Globally

By default, RIP is not enabled in the system.
To enable RIP globally, use the following commands.

1. Enter ROUTER RIP mode and enable the RIP process on the system.
   CONFIGURATION mode

   ```
   router rip
   ```

2. Assign an IP network address as a RIP network to exchange routing information.
   ROUTER RIP mode

   ```
   network ip-address
   ```

Example of Verifying RIP is Enabled

Example of Viewing Learned RIP Routes

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 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.

```
Dell(conf-router_rip)#show config
!
router rip
  network 10.0.0.0
Dell(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.

```
Dell#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 0/0
160.160.0.0/16    auto-summary
2.0.0.0/8
    [120/1] via 29.10.10.12, 00:01:22, Fa 0/0
2.0.0.0/8         auto-summary
4.0.0.0/8
    [120/1] via 29.10.10.12, 00:01:22, Fa 0/0
4.0.0.0/8         auto-summary
8.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
8.0.0.0/8         auto-summary
12.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
12.0.0.0/8        auto-summary
20.0.0.0/8
    [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
20.0.0.0/8        auto-summary
29.10.10.0/24 directly connected,Fa 0/0
29.0.0.0/8        auto-summary
31.0.0.0/8
```

```
    [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
31.0.0.0/8        auto-summary
192.162.2.0/24
    [120/1] via 29.10.10.12, 00:01:21, Fa 0/0
192.162.2.0/24    auto-summary
192.161.1.0/24
    [120/1] via 29.10.10.12, 00:00:27, Fa 0/0
192.161.1.0/24    auto-summary
192.162.3.0/24
    [120/1] via 29.10.10.12, 00:01:22, Fa 0/0
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 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
  ```

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 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.

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, the system 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.

Setting the Send and Receive Version

To change the RIP version globally or on an interface in the system, 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]
  ```

Example of an Interface Participating in the RIP Process

Example of Configuring an Interface to Send/Receive Specified Versions of RIP

Example of the `show ip protocols` Command to Verify RIP Versions on an Interface

To see whether the `version` command is configured, use the `show config` command in ROUTER RIP mode. To view the routing protocols configuration, use the `show ip protocols` command in EXEC 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 (GigabitEthernet 0/0) participating in the RIP process is also set to send and receive RIPv2 (shown in bold).

```
Dell#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
       GigabitEthernet  0/0   2     2
  Routing for Networks:
       10.0.0.0

  Routing Information Sources:
  Gateway      Distance          Last Update

  Distance: (default is 120)

Dell#
```

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.

```
Dell(conf-if)#ip rip send version 1 2
Dell(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 the Dell Networking OS does globally (shown in bold).

```
Dell#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
         FastEthernet 0/0   2      1   2
  Routing for Networks:
         10.0.0.0

Routing Information Sources:
  Gateway      Distance        Last Update

  Distance: (default is 120)

Dell#
```

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 the Dell 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.

Example of the `debug ip rip` Command

The following example shows the confirmation when you enable the debug function.

```
Dell#debug ip rip
RIP protocol debug is ON
Dell#
```

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:

- RIP Configuration on Core2
- Core 2 RIP Output
- RIP Configuration on Core3
- Core 3 RIP Output
- RIP Configuration Summary

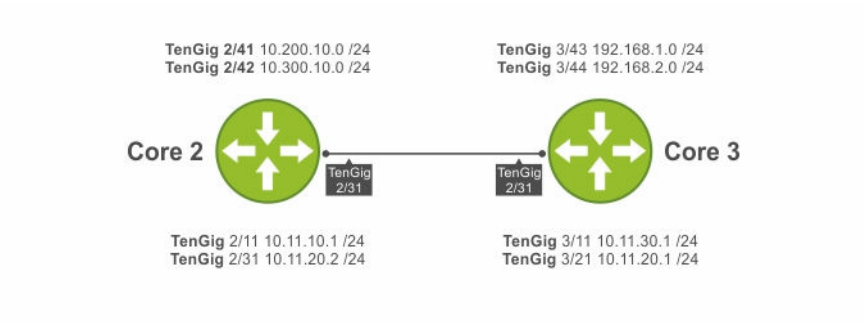


Figure 100. RIP Topology Example

RIP Configuration on Core2

The following example shows how to configure RIPv2 on a host named Core2.

Example of Configuring RIPv2 on Core 2

```
Core2(conf-if-gi-2/31)#
Core2(conf-if-gi-2/31)#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.

Example of the `show ip rip database` Command to View Learned RIP Routes on Core 2

Example of the `show ip route` Command to Show RIP Setup on Core 2

Example of the `show ip protocols` Command to Show RIP Configuration Activity on Core 2

- 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.

```
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 2/31
10.300.10.0/24     directly connected,TenGigabitEthernet 2/42
10.200.10.0/24     directly connected,TenGigabitEthernet 2/41
10.11.20.0/24      directly connected,TenGigabitEthernet 2/31
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 2/31
192.168.1.0/24     auto-summary
192.168.2.0/24
    [120/1] via 10.11.20.1, 00:00:03, TenGigabitEthernet 2/31
192.168.2.0/24     auto-summary
Core2#

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, TenGig 2/11          0/0   00:02:26
C    10.11.20.0/24   Direct, TenGig 2/31          0/0   00:02:02
R    10.11.30.0/24   via 10.11.20.1, TenGig 2/31  120/1 00:01:20
C    10.200.10.0/24  Direct, TenGig 2/41          0/0   00:03:03
C    10.300.10.0/24  Direct, TenGig 2/42          0/0   00:02:42
R    192.168.1.0/24  via 10.11.20.1, TenGig 2/31  120/1 00:01:20
R    192.168.2.0/24  via 10.11.20.1, TenGig 2/31  120/1 00:01:20
Core2#
R    192.168.1.0/24  via 10.11.20.1, TenGig 2/31  120/1 00:05:22
R    192.168.2.0/24  via 10.11.20.1, TenGig 2/31  120/1 00:05:22

Core2#

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 2/42 2 2
      TenGigabitEthernet 2/41 2 2
      TenGigabitEthernet 2/31 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.

Example of Configuring RIPv2 on Core3

```
Core3(conf-if-gi-3/21)#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.

Example of the `show ip rip database` Command to View Learned RIP Routes on Core 3

Example of the `show ip routes` Command to View RIP Setup on Core 3

Example of the `show ip protocols` Command to Show RIP Configuration Activity 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 3/43
192.168.1.0/24     auto-summary
192.168.2.0/24     directly connected,TenGigabitEthernet 3/44
192.168.2.0/24     auto-summary
Core3#

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, TenGig 3/21   120/1      00:01:14
C  10.11.20.0/24  Direct, TenGig 3/21           0/0        00:01:53
C  10.11.30.0/24  Direct, TenGig 3/11           0/0        00:06:00
R  10.200.10.0/24 via 10.11.20.2, TenGig        3/21 120/1 00:01:14
R  10.300.10.0/24 via 10.11.20.2, TenGig        3/21 120/1 00:01:14
C  192.168.1.0/24 Direct, TenGig                3/43 0/0   00:06:53
C  192.168.2.0/24 Direct, TenGig                3/44 0/0   00:06:26
Core3#

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 3/44 2 2
      TenGigabitEthernet 3/43 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

Example of Viewing RIP Configuration on Core 2

Example of Viewing RIP Configuration on Core 3

```
!
interface TenGigabitEthernet 2/11
  ip address 10.11.10.1/24
  no shutdown
!
interface TenGigabitEthernet 2/31
  ip address 10.11.20.2/24
  no shutdown
!
interface TenGigabitEthernet 2/41
  ip address 10.200.10.1/24
  no shutdown
!
interface TenGigabitEthernet 2/42
  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

!
interface TenGigabitEthernet 3/11
  ip address 10.11.30.1/24
  no shutdown

!
interface TenGigabitEthernet 3/21
  ip address 10.11.20.1/24
  no shutdown

!
interface TenGigabitEthernet 3/43
  ip address 192.168.1.1/24
  no shutdown

!
interface TenGigabitEthernet 3/44
  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
```

46

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 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).

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
- Enabling an RMON MIB Collection History Group

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.

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**.

Example of the `rmon alarm` Command

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.

```
Dell(conf)#rmon alarm 10 1.3.6.1.2.1.2.2.1.20.1 20 delta rising-threshold 15 1
falling-threshold 0
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 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**.

Example of the `rmon event` Command

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".

```
Dell(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 owner-string]
  ```

  - `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.
- *owner-string*: (Optional) records the name of the owner of the RMON group of statistics. The default is a **null-terminated string**.

Example of the `rmon collection statistics` Command

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*.

```
Dell(conf-if-te-0/40)#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 owner-string]
  [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**.
  - *owner-string*: (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).

Example of the `rmon collection history` Command

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.

```
Dell(conf-if-mgmt)#rmon collection history controlEntry 20 owner john
```

Enabling an RMON MIB Collection History Group

The `rmon collection history` command enables an RMON MIB collection history group of statistics.

In the following example, the command 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.

Example of the `rmon collection history` Command

```
Dell(conf-if-te-0/40)#rmon collection history controlEntry 20 owner
john
```

47

Rapid Spanning Tree Protocol (RSTP)

Dell Networking OS supports rapid spanning tree protocol (RSTP).

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 operating system (OS) supports three other variations of spanning tree, as shown in the following table.

Table 55. Spanning Tree Variations Dell Networking OS Supports

Dell 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
- Enable BPDU Filtering Globally
- Modifying Interface Parameters
- Configuring an EdgePort
- Prevent Network Disruptions with BPDU Guard
- Influencing RSTP Root Selection
- SNMP Traps for Root Elections and Topology Changes
- Configure Spanning Tree
- Configuring Fast Hellos for Link State Detection
- Flush MAC Addresses after a Topology Change

Important Points to Remember

- RSTP is disabled by default.
- The Dell 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 Networking recommends limiting the range to five ports and 40 VLANs.

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
   ```

Example of Verifying an Interface is in Layer 2 Mode and Enabled

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.

```
Dell(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  no shutdown
Dell(conf-if-te-1/1)#
```

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.

Example of Verifying that RSTP is Enabled

```
Dell(conf-rstp)#show config
!
protocol spanning-tree rstp
no disable
Dell(conf-rstp)#
```

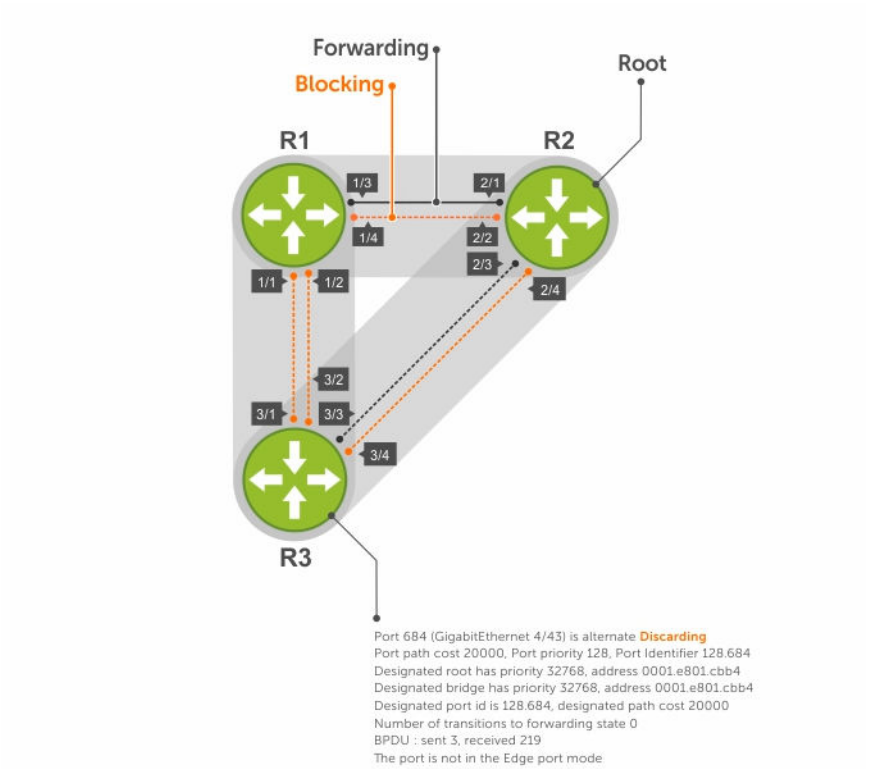


Figure 101. Rapid Spanning Tree Enabled Globally

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.

Example of the `show spanning-tree rstp` Command

```
Dell#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 TenGig 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, bpdu filter is disabled

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, bpdu filter is disabled

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, bpdu filter is disabled

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, bpdu filter is disabled

Dell#
```

To confirm that a port is participating in RSTP, use the `show spanning-tree rstp brief` command from EXEC privilege mode.

Example of the `show spanning-tree rstp brief` Command

```
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
---------- -------- ----   ------- --- ------- ------------- --------
TenGig 3/1  128.681 128  20000  BLK 20000 32768 0001.e80b.88bd  128.469
TenGig 3/2  128.682 128  20000  BLK 20000 32768 0001.e80b.88bd  128.470
TenGig 3/3  128.683 128  20000  FWD 20000 32768 0001.e801.cbb4  128.379
TenGig 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 Bpdu
                                                           Filter
------- --- ------ -------- ---- ------- --- ----------- ----- ----
TenGig 3/1  Altr  128.681 128 20000  BLK 20000  P2P       No    No
TenGig 3/2  Altr  128.682 128 20000  BLK 20000  P2P       No    No
TenGig 3/3  Root  128.683 128 20000  FWD 20000  P2P       No    No
TenGig 3/4  Altr  128.684 128 20000  BLK 20000  P2P       No    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`

  For bridge protocol data units (BPDU) filtering behavior, refer to Removing an Interface from the Spanning Tree Group.

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 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 56. RSTP Default Values

RSTP Parameter	Default Value
Forward Delay	**15 seconds**
Hello Time	**2 seconds**

RSTP Parameter	Default Value
Max Age	**20 seconds**
Port Cost: • 10-Gigabit Ethernet interfaces • 40-Gigabit Ethernet interfaces • Port Channel with two 40-Gigabit Ethernet interfaces • Port Channel with two 10-Gigabit Ethernet interfaces	Port Cost: • **1400** • **2000** • **600** • **1800**
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 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.

Enable BPDU Filtering Globally

The enabling of BPDU Filtering stops transmitting of BPDUs on the operational port fast enabled ports by default.
When BPDUs are received, the spanning tree is automatically prepared. By default global bpdu filtering is disabled.

Enable BPDU Filter globally to filter transmission of BPDU port fast enabled interfaces.

PROTOCOL SPANNING TREE RSTP mode

```
edge-port bpdu filter default
```

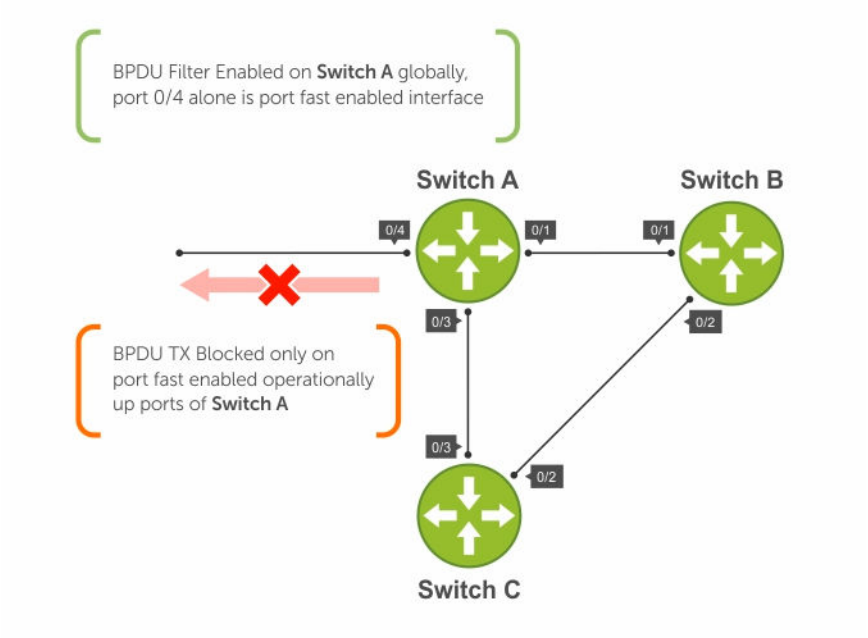


Figure 102. BPDU Filtering Enabled Globally

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 240.

  The default is **128**.

To view the current values for interface parameters, use the `show spanning-tree rstp` command from EXEC privilege mode.

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 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][bpdufilter]
  ```

Example of Verifying an EdgePort is Enabled on an Interface

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 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.

```
Dell(conf-if-te-2/0)#show config
!
interface TenGigabitethernet 2/0
  no ip address
  switchport
  spanning-tree rstp edge-port
  shutdown
Dell(conf-if-te-2/0)#
```

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.

Example of the bridge-priority Command

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).

```
Dell(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
```

SNMP Traps for Root Elections and Topology Changes

To enable SNMP traps for RSTP, MSTP, and PVST+ collectively, use the following command.

Enable SNMP traps for RSTP, MSTP, and PVST+ collectively.

```
snmp-server enable traps xstp
```

Configuring Fast Hellos for Link State Detection

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.

Example of Verifying Hello-Time Interval

```
Dell(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.

48

Security

Security features are supported on the I/O Aggregator.

This chapter describes several ways to provide access security to the Dell Networking system.

For details about all the commands described in this chapter, refer to the *Security* chapter in the *Dell PowerEdge FN I/O Aggregator Command Line Reference Guide*.

Supported Modes

Standalone, PMUX, VLT, Stacking

NOTE: You can also perform some of the configurations using the Web GUI - Dell Blade IO Manager. For more information, see the Dell Blade IO Manager Online Help.

Understanding Banner Settings

This functionality is supported on the Aggregator.

A banner is a note that is displayed when you log in to the system, depending on the privilege level and the command mode into which the you log in. You can specify different banners to be displayed as the message-of-the-day (MOTD), as the opening quote in EXEC mode, or as the beginning message in EXEC Privilege mode. Setting up a banner enables you to display any important information or group-level notification that needs to be communicated to all the users of the system.

A login banner message is displayed only in EXEC Privilege mode after entering the enable command followed by the password. These banners are not displayed to users in EXEC mode. When you connect to a system, the message-of-the-day (MOTD) banner is displayed first, followed by the login banner and prompts. After you log in to the system with valid authentication credentials, the EXEC banner is shown.

You can use the MOTD banner to alert users of critical upcoming events so that they can plan and schedule their accessibility to the device. You can modify the banner messages depending on the requirements or conditions.

Accessing the I/O Aggregator Using the CMC Console Only

This functionality is supported on the Aggregator.

You can enable the option to access and administer an Aggregator only using the chassis management controller (CMC) interface, and prevent the usage of the CLI interface of the device to configure and monitor settings. You can configure the restrict-access session command to disable access of the Aggregator using a Telnet or SSH session; the device is accessible only using the CMC GUI. You can use

the no version of this command to reactivate the Telnet or SSH session capability for the device. Use the show restrict-access command to view whether the access to a device using Telnet or SSH is disabled or not.

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 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 | exec | suppress | system level} {default | name}
  {start-stop | wait-start | stop-only} {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.
  - `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. Currently, Dell Networking OS supports only TACACS+.

Suppressing AAA Accounting for Null Username Sessions

When you activate AAA accounting, the Dell 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 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.

Example of Configuring AAA Accounting to Track EXEC and EXEC Privilege Level Command Use

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 exec default start-stop tacacs+
Dell(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
  ```

Example of Enabling AAA Accounting with a Named Method List

```
Dell(config-line-vty)# accounting commands 15 com15
Dell(config-line-vty)# accounting exec execAcct
```

Monitoring AAA Accounting

Dell 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
  ```

Example of the `show accounting` Command for AAA Accounting

```
Dell#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
Dell#
```

AAA Authentication

Dell 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 Networking implementation, the Dell 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 Networking uses local usernames/passwords (stored on the Dell 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 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.

Configuration Task List for AAA Authentication

The following sections provide the configuration tasks.

- Configure Login Authentication for Terminal Lines
- 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 Networking OS Command Reference Guide*.

Configure Login Authentication for Terminal Lines

You can assign up to five authentication methods to a method list. Dell 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 Networking OS applies the next method list until the user either passes or fails the authentication. If the user fails a method list, Dell 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 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 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` or `enable password` command in CONFIGURATION mode.
   - `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 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.

Example of Enabling Authentication from the RADIUS Server

```
Dell(config)# aaa authentication enable default radius tacacs
Radius and TACACS server has to be properly setup for this.
Dell(config)# radius-server host x.x.x.x key <some-password>
Dell(config)# tacacs-server host x.x.x.x key <some-password>
```

To use local authentication for `enable secret` on the console, while using remote authentication on VTY lines, issue the following commands.

Example of Enabling Local Authentication for the Console and Remote Authentication for VTY Lines

```
Dell(config)# aaa authentication enable mymethodlist radius tacacs
Dell(config)# line vty 0 9
Dell(config-line-vty)# enable authentication mymethodlist
```

Server-Side Configuration

- **TACACS+** — When using TACACS+, Dell Networking OS 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, Dell Networking 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.

AAA Authorization

The Dell 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, the system 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 the Dell Networking OS, you can configure a privilege level for users who need limited access to the system.

Every command in the Dell Networking OS is assigned a privilege level of 0, 1, or 15. You can configure up to 16 privilege levels. The Dell 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 the Dell 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 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. The Dell 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 privilege levels and passwords, refer to the *Security* chapter in the *Dell Networking OS Command Reference Guide*.

Configuring a Username and Password

In the Dell 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*: Restrict access by access-class.
  - `nopassword`: Require password for the user to login.
  - *encryption-type*: Enter 0 for plain text or 7 for encrypted text.
  - *password*: Enter a string. Specify the password for the user.
  - `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 the Dell Networking OS, use the `enable` command to enter EXEC Privilege level 15. After entering the command, the system 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 plain text or 7 for 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, 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 the Dell 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.

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).
   - `access-class` *access-list-name*: enter the name of a configured IP ACL.
   - `privilege` *level*: the range is from 0 to 15.
   - `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 text string.
   - `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 text 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 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.

Example of Configuring a Custom Privilege Level

```
Dell(conf)#username john privilege 8 password john
Dell(conf)#enable password level 8 notjohn
Dell(conf)#privilege exec level 8 configure
Dell(conf)#privilege config level 8 snmp-server
Dell(conf)#end
Dell#show running-config
Current Configuration ...
!
hostname FTOS
!
enable password level 8 notjohn
enable password FTOS
!
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.

Example of Privilege Level Login and Available Commands

```
apollo% telnet 172.31.1.53
Trying 172.31.1.53...
Connected to 172.31.1.53.
Escape character is '^]'.
Login: john
Password:
Dell#show priv
```

```
Current privilege level is 8
Dell#?
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
Dell#confi
Dell(conf)#?
end            Exit from Configuration mode
```

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 25 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, the system 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.

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 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 and Authorization

The Dell 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 Overview

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`

Setting Access to Privilege Levels through RADIUS

To configure a privilege level for the user to enter into when they connect to a session, use the following command.

Configure a privilege level for the user to enter into when they connect to a session through the RADIUS server.

`privilege level`

Configure this value on the client system.

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 Networking OS commands related to RADIUS, refer to the *Security* chapter in the *Dell 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 Networking OS attempts to connect with them in the order in which they were configured. When Dell 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 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
  ```

TACACS+

Dell 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 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}
   ```

Example of a Failed Authentication

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 Networking OS employs the second method (or third method, if necessary) automatically. For example, if the TACACS+ server is reachable, but the server key is invalid, Dell 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.

```
Dell(conf)#
Dell(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+
Dell(conf)#
Dell(conf)#do show run tacacs+
!
tacacs-server key 7 d05206c308f4d35b
tacacs-server host 10.10.10.10 timeout 1
```

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

```
Dell(conf)#
Dell(conf)#aaa authentication login tacacsmethod tacacs+
Dell(conf)#aaa authentication exec tacacsauthorization tacacs+
Dell(conf)#tacacs-server host 25.1.1.2 key Force
Dell(conf)#
Dell(conf)#line vty 0 9
Dell(config-line-vty)#login authentication tacacsmethod
Dell(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.

Example of Connecting with a TACACS+ Server Host

To specify multiple TACACS+ server hosts, configure the `tacacs-server host` command multiple times. If you configure multiple TACACS+ server hosts, Dell 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:
Dell#
Dell#
```

Enabling SCP and SSH

Secure shell (SSH) is a protocol for secure remote login and other secure network services over an insecure network. Dell Networking OS is compatible with SSH versions 1.5 and 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 Networking OS Command Line Interface Reference Guide*.

Dell 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 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 {1 | 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
  ```

Specifying an SSH Version

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.

```
Dell(conf)#ip ssh server version 2
Dell(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
Dell(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:
   ```

Example of Using SCP to Copy from an SSH Server on Another Switch

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 key-size` : configure the size of the server-generated RSA SSHv1 key.
- `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.

```
Dell#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:
```

Secure Shell Authentication

Secure Shell (SSH) is enabled by default using the SSH Password Authentication method.

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.

Example of Using Telnet for Remote Login

```
Dell(conf)#ip telnet server enable
Dell(conf)#no ip telnet server enable
```

VTY Line and Access-Class Configuration

Various methods are available to restrict VTY access in Dell Networking OS. These depend on which authentication scheme you use — line, local, or remote.

Table 57. 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 Dell Networking OS version 5.2.1.0 and later)
RADIUS	YES	NO	YES (with Dell Networking OS version 6.1.1.0 and later)

Dell Networking OS 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

Dell Networking OS 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.

Dell Networking OS can assign different access classes to different users by username. Until users attempt to log in, Dell Networking OS 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, Dell Networking OS retrieves the access class from the local database and applies it. (Dell Networking OS 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)

```
Dell(conf)#user gooduser password abc privilege 10 access-class permitall
Dell(conf)#user baduser password abc privilege 10 access-class denyall
Dell(conf)#
Dell(conf)#aaa authentication login localmethod local
Dell(conf)#
Dell(conf)#line vty 0 9
Dell(config-line-vty)#login authentication localmethod
Dell(config-line-vty)#end
```

VTY Line Remote Authentication and Authorization

Dell Networking OS retrieves the access class from the VTY line.

The Dell Networking OS takes the access class from the VTY line and applies it to ALL users. Dell Networking OS 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, Dell Networking OS immediately applies it. If the access-class is set to deny all or deny for the incoming subnet, Dell Networking OS 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)

```
Dell(conf)#ip access-list standard deny10
Dell(conf-ext-nacl)#permit 10.0.0.0/8
Dell(conf-ext-nacl)#deny any
Dell(conf)#
Dell(conf)#aaa authentication login tacacsmethod tacacs+
Dell(conf)#tacacs-server host 256.1.1.2 key Force10
Dell(conf)#
Dell(conf)#line vty 0 9
Dell(config-line-vty)#login authentication tacacsmethod
Dell(config-line-vty)#
Dell(config-line-vty)#access-class deny10
Dell(config-line-vty)#end
(same applies for radius and line authentication)
```

49

Service Provider Bridging

Dell Networking OS supports service provider 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*.

VLAN stacking 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.

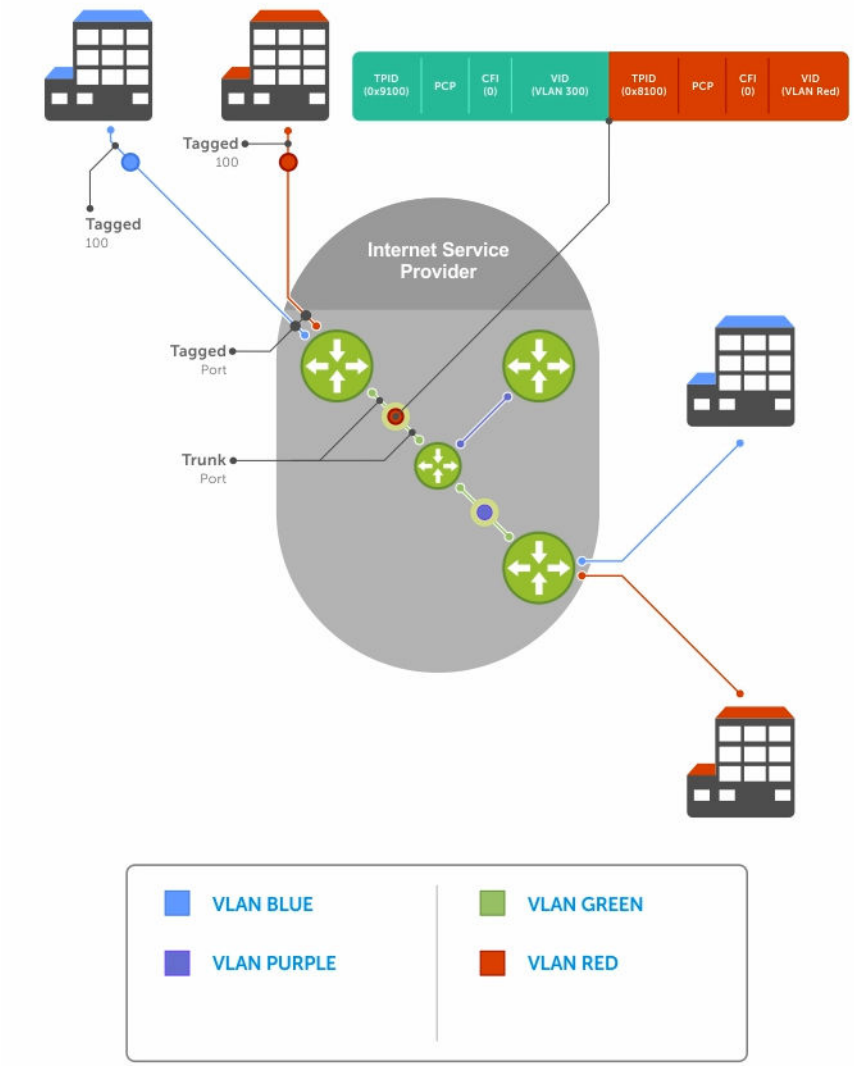


Figure 103. VLAN Stacking in a Service Provider Network

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 Networking cautions against using the same MAC address on different customer VLANs, on the same VLAN-Stack VLAN.

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. Enable VLAN-Stacking for a VLAN.

Related Configuration Tasks

- Configuring the Protocol Type Value for the Outer VLAN Tag
- Configuring 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
   ```

Example of Displaying the VLAN-Stack Configuration for a Switchport

To display the VLAN-Stacking configuration for a switchport, use the `show config` command from INTERFACE mode.

```
Dell#show run interface gi 7/0
!
interface GigabitEthernet 7/0
  no ip address
  switchport
  vlan-stack access
  no shutdown
Dell#show run interface gi 7/12
!
interface GigabitEthernet 7/12
  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
  ```

Example of Viewing VLAN Stack Member Status

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.

```
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

NUM  Status      Q Ports
* 1  Active      U Gi 13/0-5,18
  2  Inactive
  3  Inactive
  4  Inactive
  5  Inactive
  6  Active      M Po1(Gi 13/14-15)
                 M Gi 13/13
Dell#
```

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. The system displays the S-Tag TPID only if it is a non-default value.

Configuring 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]
   ```

Example of Configuring a Trunk Port as a Hybrid Port and Adding it to Stacked VLANs

In the following example, GigabitEthernet 0/1 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.

```
Dell(conf)#int gi 0/1
Dell(conf-if-gi-0/1)#portmode hybrid
Dell(conf-if-gi-0/1)#switchport
Dell(conf-if-gi-0/1)#vlan-stack trunk
Dell(conf-if-gi-0/1)#show config
!
interface GigabitEthernet 0/1
  no ip address
  portmode hybrid
  switchport
  vlan-stack trunk
  shutdown
Dell(conf-if-gi-0/1)#interface vlan 100
Dell(conf-if-vl-100)#untagged gigabitethernet 0/1
Dell(conf-if-vl-100)#interface vlan 101
Dell(conf-if-vl-101)#tagged gigabitethernet 0/1
Dell(conf-if-vl-101)#interface vlan 103
Dell(conf-if-vl-103)#vlan-stack compatible
Dell(conf-if-vl-103-stack)#member gigabitethernet 0/1
Dell(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 Gi 0/1
  101  Inactive               T Gi 0/1
  103  Inactive               M Gi 0/1
```

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
  ```

Example of Debugging a VLAN and its Ports

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)

```
Dell# debug member vlan 603
vlan id  : 603
ports    : Gi 2/47 (MT), Gi 3/1(MU), Gi 3/25(MT), Gi 3/26(MT), Gi 3/27(MU)

Dell#debug member port gigabitethernet 2/47
vlan id  : 603 (MT), 100(T), 101(NU)
Dell#
```

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; the Dell Networking OS uses 0x9100 (shown in the following) while non-Dell 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 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. Beginning with the Dell Networking OS version 8.2.1.0, 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. The Dell 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.

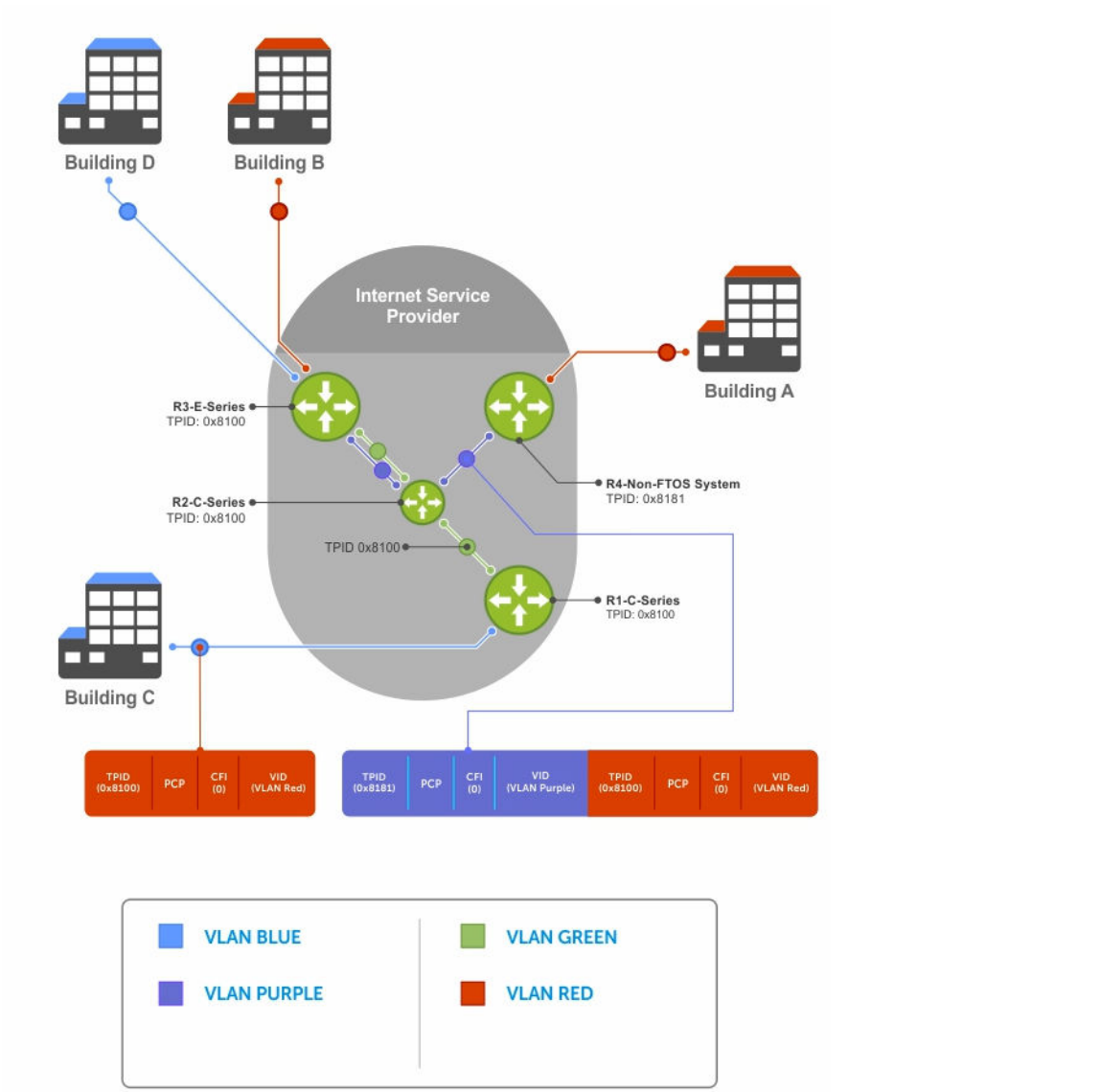


Figure 104. Single and Double-Tag TPID Match

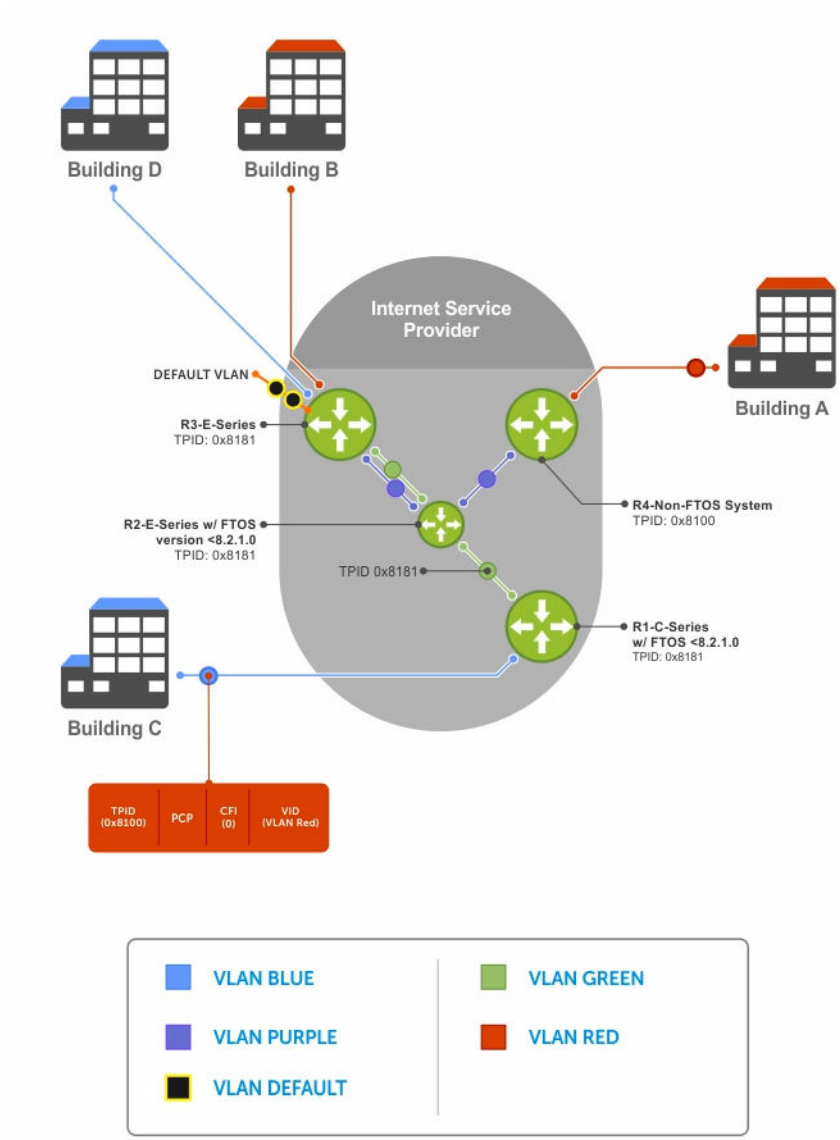


Figure 105. Single and Double-Tag First-byte TPID Match

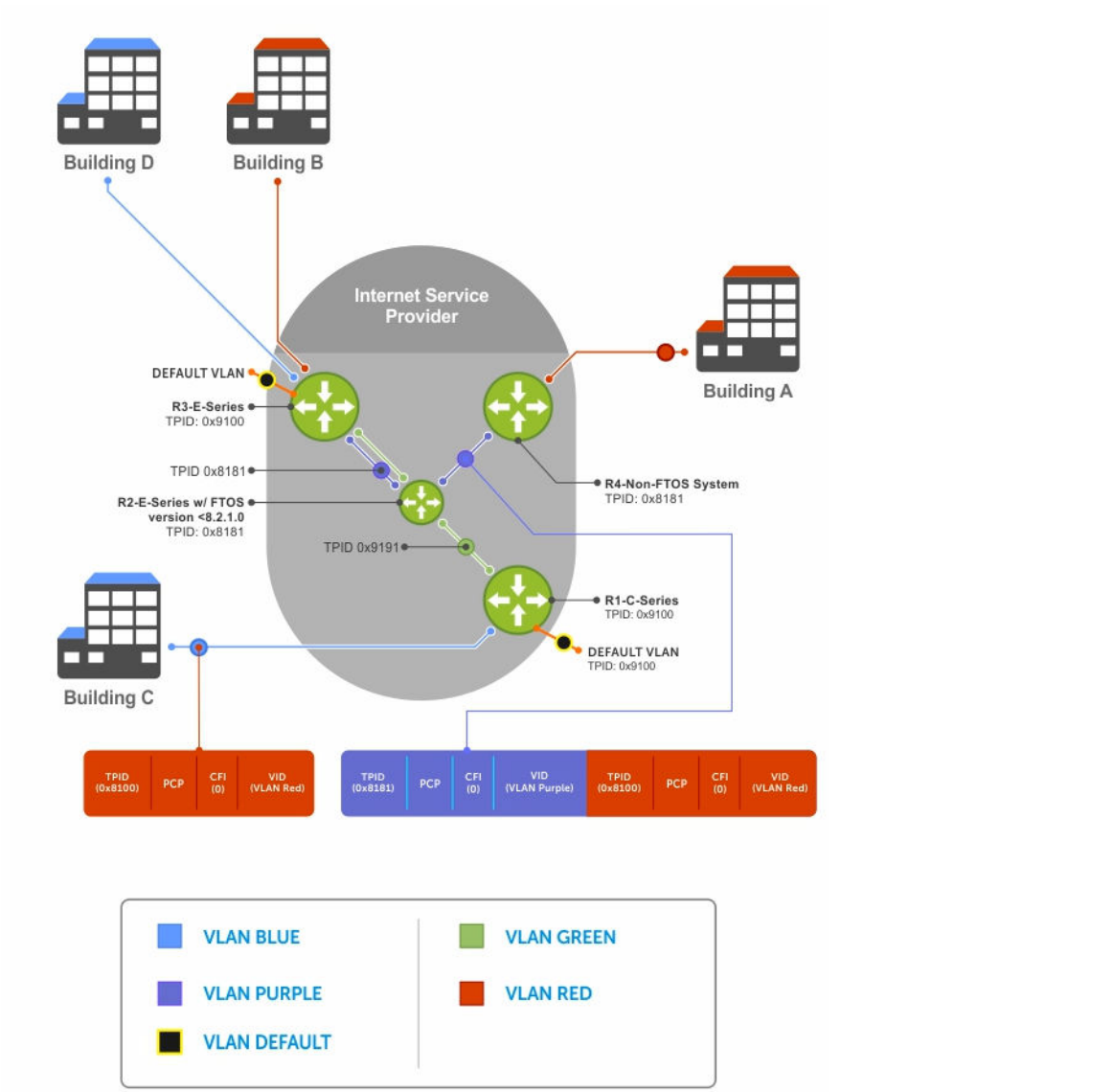


Figure 106. Single and Double-Tag TPID Mismatch

Table 58. 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

Network Position	Incoming Packet TPID	System TPID	Match Type	Pre-Version 8.2.1.0	Version 8.2.1.0+
		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
		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

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 59. 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 a Dell 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 the Dell 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.

Example of Viewing DEI-Honoring Configuration

To display the DEI-honoring configuration, use the `show interface dei-honor [interface slot/port | linecard number port-set number]` in EXEC Privilege mode.

```
Dell#show interface dei-honor

Default Drop precedence: Green
Interface CFI/DEI  Drop precedence
---------------------------------------
Gi 0/1    0        Green
Gi 0/1    1        Yellow
Gi 8/9    1        Red
Gi 8/40   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}
  ```

Example of Viewing DEI-Marking Configuration

To display the DEI-marking configuration, use the `show interface dei-mark [interface slot/port | linecard number port-set number]` in EXEC Privilege mode.

```
Dell#show interface dei-mark
```

```
Default CFI/DEI Marking: 0
Interface Drop precedence CFI/DEI
--------------------------------
Gi 0/1     Green           0
Gi 0/1     Yellow          1
Gi 8/9     Yellow          0
Gi 8/40    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.

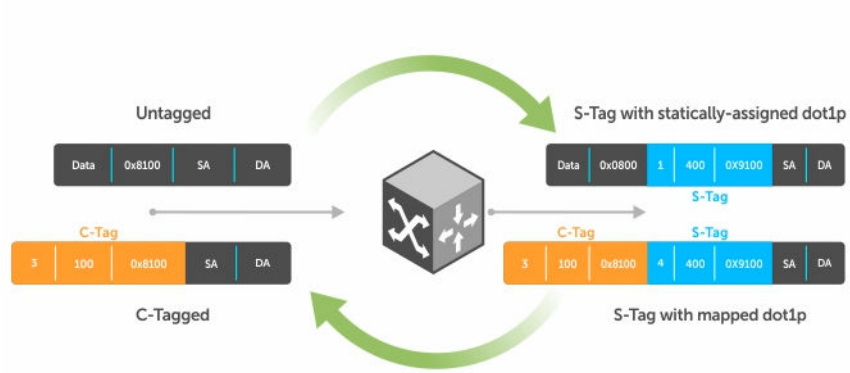


Figure 107. Statically and Dynamically Assigned dot1p for VLAN Stacking

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 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 GigabitEthernet 0/21
  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 reload
   ```

3. 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).

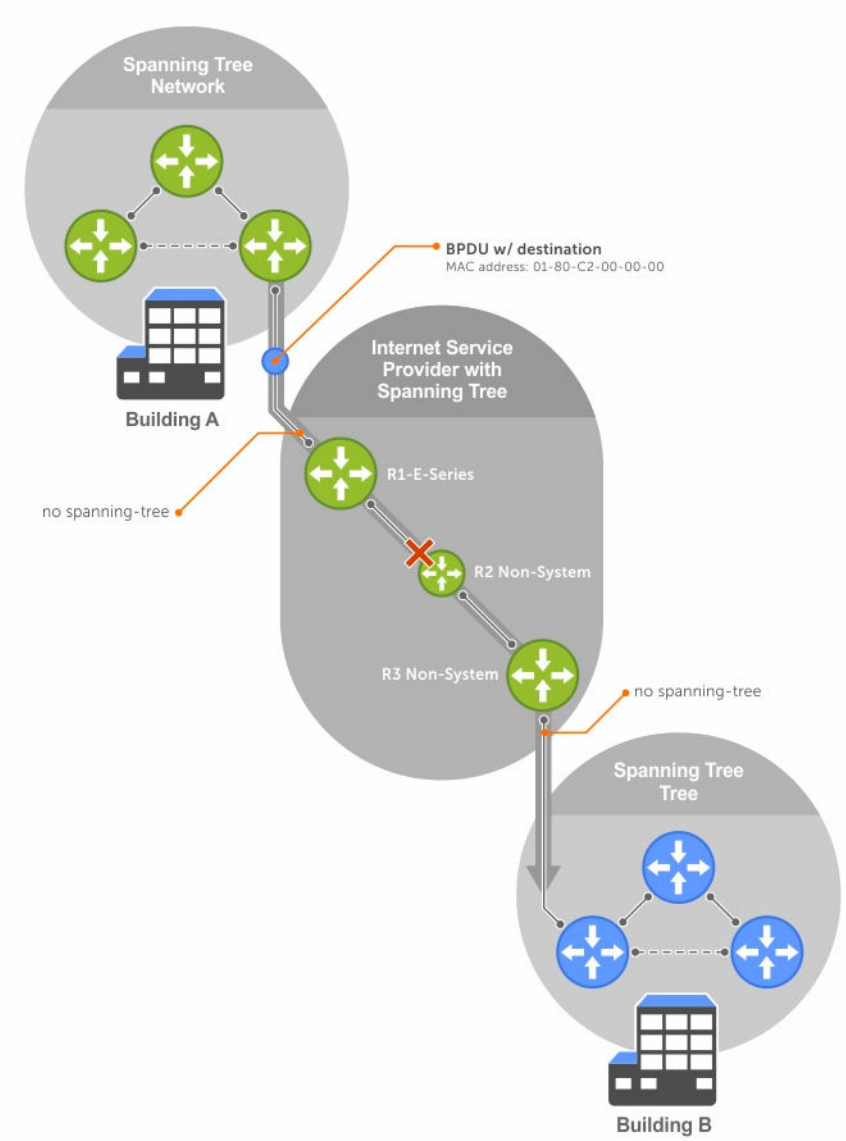


Figure 108. VLAN Stacking without L2PT

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 Networking OS Behavior: In the Dell Networking OS versions prior to 8.2.1.0, the MAC address that Dell Networking systems use to overwrite the Bridge Group Address on ingress was non-configurable. The value of the L2PT MAC address was the Dell Networking-unique MAC address, 01-01-e8-00-00-00. As such, with these versions, Dell Networking systems are required at the egress edge of the intermediate

network because only the Dell Networking OS could recognize the significance of the destination MAC address and rewrite it to the original Bridge Group Address. In the Dell 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 Networking systems can recognize and rewrite the address at egress edge.

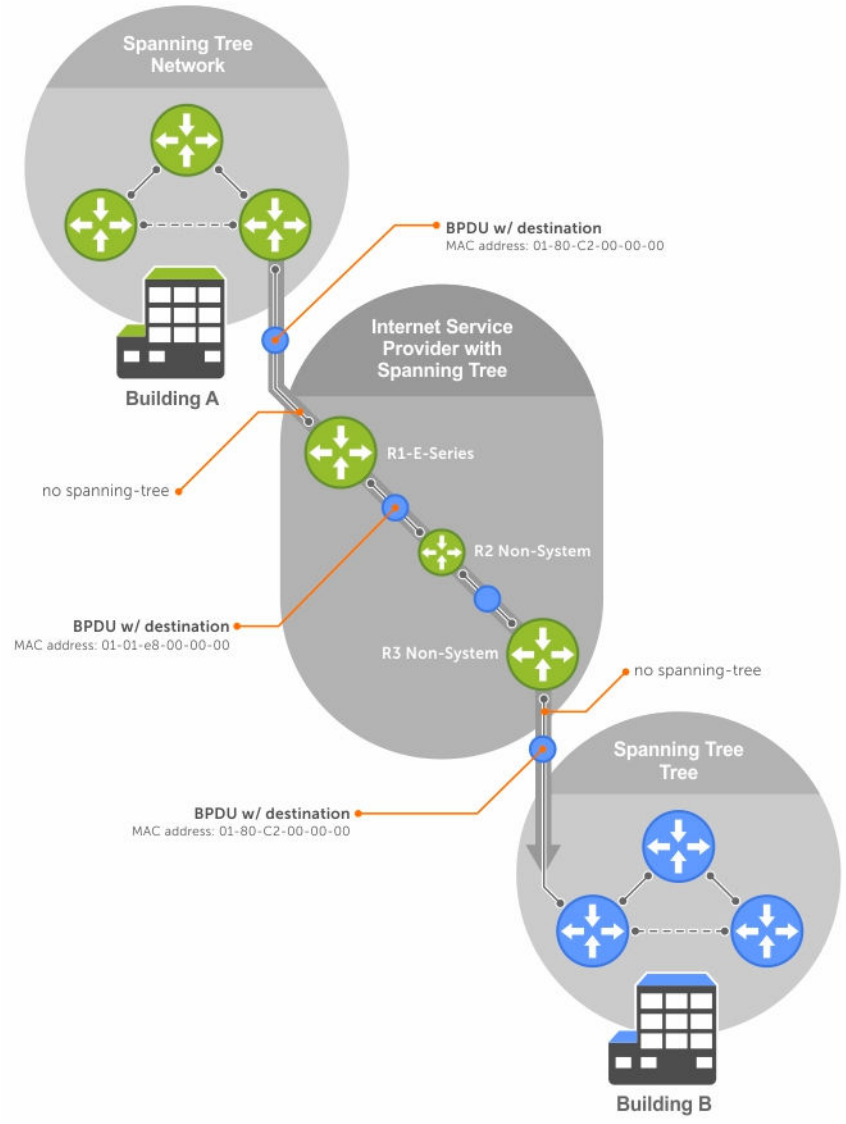


Figure 109. VLAN Stacking with L2PT

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, the Dell Networking OS uses a Dell 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.

50

sFlow

Dell Networking OS supports configuring sFlow.

Overview

The Dell 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 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.

The default global sampling rate is 32768. The sampling rates are determined as follows:

- If the interface stats 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 shutdown, the sampling rate is set using the global sampling rate.
- If the global sample rate is non-default, for example 256 bytes, and if the sampling rate is not configured on an interface, the sampling rate of the interface is the global non-default sampling rate, that is 256 bytes.

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 Networking OS implementation of the sFlow MIB supports sFlow configuration using the `snmpset` command.
- The Dell 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.
- On the switch, up to 700 packets can be sampled and processed per second.

Enabling and Disabling sFlow

By default, sFlow is disabled globally on the system.
Use the following command to enable sFlow globally.

- Enable sFlow globally.
  CONFIGURATION mode

  ```
  [no] sflow enable
  ```

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 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. If the traffic ingresses on an sFlow enabled interface, 256 bytes are copied.

- 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 Viewing the Extended Max Header Size

```
Dell#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 sflow` command

The bold line shows the sFlow default maximum header size:

```
Dell#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   <<< sampling rate based on
line speed if global sampling rate is default
Dell#
```

Example of Viewing the `sFlow max-header-size extended` on an Interface Mode

If you enable sFlow on an interface, the show output displays the following (shown in bold).

```
Dell(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 the `show running-config sflow` Command

```
Dell#show running-config sflow
!
sflow collector 100.1.1.12 agent-addr 100.1.1.1
sflow enable
sflow max-header-size extended


Dell#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

The Dell Networking OS includes the following sFlow display commands.

- Displaying Show sFlow Global
- Displaying Show sFlow on an Interface
- Displaying Show sFlow on a Stack Unit

Displaying Show sFlow Global

To view sFlow statistics, use the following command.

- Display sFlow configuration information and statistics.
  EXEC mode

  ```
  show sflow
  ```

Example of Viewing sFlow Configuration (Global)

The first bold line indicates sFlow is globally enabled.

```
Dell#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
Dell#
```

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
```

Example of Viewing sFlow Configuration (Interface)

```
Dell#show sflow interface tengigabitethernet 1/16
Te 1/16
Configured sampling rate         :8192
Actual sampling rate             :8192
Sub-sampling rate                :2
Counter polling interval         :15
Samples rcvd from h/w            :33
Samples dropped for sub-sampling :6
```

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 unit-number
  ```

Example of Viewing sFlow Configuration (Stack Unit)

```
Dell#show sflow stack-unit 1
Stack-Unit 1
  Samples rcvd from h/w         :0
  Total UDP packets exported    :0
  UDP packets exported via RPM  :0
  UDP packets dropped           :0
Dell#
```

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**.

Changing the Sampling Rate

The `sflow sample-rate` command, when issued in CONFIGURATION mode, changes the default sampling rate.
By default, the sampling rate of an interface is set to the same value as the current global default sampling rate. If the value entered is not a correct power of 2, the command generates an error message with the previous and next power-of-2 value. Select one of these two numbers and re-enter the command. (For more information on values in power-of-2, refer to Sub-Sampling.)
To change the sampling rate either globally or on an interface, use the following command.

- Change the global or interface sampling rate.
  CONFIGURATION mode or INTERFACE mode

  `[no] sflow sample-rate` *`sample-rate`*

  `sample-rate`: The range is from 256 to 8388608 for the C-Series and S-Series. The range is from 2 to 8388608 for the E-Series.

  The rate must be entered in factors of 2 (for example, 4096 or 8192).

Sub-Sampling

The sFlow sample rate is not the frequency of sampling, but the number of packets that are skipped before the next sample is taken.

Therefore, the sFlow agent uses sub-sampling to create multiple sampling rates per port-pipe. To achieve different sampling rates for different ports in a port-pipe, the sFlow agent takes the lowest numerical value of the sampling rate of all the ports within the port-pipe and configures all the ports to this value. The sFlow agent is then able to skip samples on ports where you require a larger sampling rate value.

Sampling rates are configurable in powers of two. This configuration allows the smallest sampling rate possible on the hardware and also allows all other sampling rates to be available through sub-sampling.

For example, if Tengig 1/0 and 1/1 are in a port-pipe, and they are configured with a sampling rate of 4096 on interface Tengig 1/0, and 8192 on Tengig 1/1, the sFlow agent does the following:

1. Configures the hardware to a sampling rate of 4096 for all ports with sFlow enabled on that port-pipe.
2. Configures interface Tengig 1/0 to a sub-sampling rate of 1 to achieve an actual rate of 4096.

3. Configures interface Tengig 1/1 to a sub-sampling rate of 2 to achieve an actual rate of 8192.

NOTE: Sampling rate backoff can change the sampling rate value that is set in the hardware. This equation shows the relationship between actual sampling rate, sub-sampling rate, and the hardware sampling rate for an interface:

Actual sampling rate = sub-sampling rate * hardware sampling rate

Note the absence of a configured rate in the equation. That absence is because when the hardware sampling rate value on the port-pipe exceeds the configured sampling rate value for an interface, the actual rate changes to the hardware rate. The sub-sampling rate never goes below a value of one.

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

Dell Networking OS 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.

- 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
  ```

Example of Verifying Extended sFlow is Enabled

Example of Verifying Extended sFlow Disabled

The bold line shows that extended sFlow settings are enabled on all three types.

```
Dell#show sflow
sFlow services are enabled
```

```
Global default sampling rate: 4096
Global default counter polling interval: 15
Global extended information enabled: switch
1 collectors configured
Collector IP addr: 10.10.10.3, Agent IP addr: 10.10.0.0, UDP port: 6343
77 UDP packets exported
0 UDP packets dropped
165 sFlow samples collected
69 sFlow samples dropped due to sub-sampling
Stackunit 1 Port set 0 H/W sampling rate 8192
Tengig 1/16: configured rate 8192, actual rate 8192, sub-sampling rate 1
Tengig 1/17: configured rate 16384, actual rate 16384, sub-sampling rate 2
Stackunit 3 Port set 1 H/W sampling rate 16384
Tengig 3/40: configured rate 16384, actual rate 16384, sub-sampling rate 1
```

If you did not enable any extended information, the show output displays the following (shown in bold).

```
Dell#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
```

51

Simple Network Management Protocol (SNMP)

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*.

NOTE: An I/O Aggregator supports standard and private SNMP MIBs, including Get operations in supported MIBs.

Supported Modes

Standalone, PMUX, VLT, Stacking

Implementation Information

The Dell Networking OS supports SNMP version 1 as defined by RFC 1155, 1157, and 1212, SNMP version 2c as defined by RFC 1901.

Configuring the Simple Network Management Protocol

NOTE: The configurations in this chapter use a UNIX environment with net-snmp version 5.4. This is only one of many RFC-compliant SNMP utilities you can use to manage the Aggregator using SNMP. Also, these configurations use SNMP version 2c.

Configuring SNMP version 1 or version 2 requires only a single step:

1. Create a community.

NOTE: IOA supports only Read-only mode.

Important Points to Remember

- Typically, 5-second timeout and 3-second retry values on an SNMP server are sufficient for both local area network (LAN) and wide area network (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.

Setting up SNMP

Dell Networking OS supports SNMP version 1 and version 2 which 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).

Creating a Community

For SNMPv1 and SNMPv2, create a community to enable the community-based security in the Dell 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.

The Dell 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 retrieve values in Read-Only mode. Read-write access is not supported.

```
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP
WARM_START.
```

To create an SNMP community:

- Choose a name for the community.
  CONFIGURATION mode

  ```
  snmp-server community name ro
  ```

Example of Creating an SNMP Community

To view your SNMP configuration, use the `show running-config snmp` command from EXEC Privilege mode.

```
Dell(conf)#snmp-server community my-snmp-community ro
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP
WARM_START.
Dell#show running-config snmp
!
snmp-server community mycommunity ro
Dell#
```

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 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.

Example of Reading the Value of a Managed Object

```
> 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
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32856932) 3 days, 19:16:09.32
```

Example of 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
SNMPv2-MIB::sysName.0 = STRING:
```

Example of Reading the Value of Many Managed Objects at Once

```
> snmpwalk -v 2c -c mycommunity 10.16.130.148 .1.3.6.1.2.1.1
SNMPv2-MIB::sysDescr.0 = STRING: Dell Networking OS
Operating System Version: 1.0
Application Software Version: E8-3-17-46
Series: I/O-Aggregator
Copyright (c) 1999-2012 by Dell Inc. All Rights Reserved.
Build Time: Sat Jul 28 03:20:24 PDT 2012
SNMPv2-MIB::sysObjectID.0 = OID: SNMPv2-SMI::enterprises.6027.1.4.2
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (77916) 0:12:59.16
SNMPv2-MIB::sysContact.0 = STRING:
SNMPv2-MIB::sysName.0 = STRING: FTOS
SNMPv2-MIB::sysLocation.0 = STRING:
SNMPv2-MIB::sysServices.0 = INTEGER: 4
```

Displaying the Ports in a VLAN using SNMP

Dell Networking OS identifies VLAN interfaces using an interface index number that is displayed in the output of the `show interface vlan` command.

Example of Identifying the VLAN Interface Index Number

```
Dell(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 in the following example.

Example of Viewing the Ports in a VLAN in SNMP

```
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
```

The table that the Dell 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 on an Aggregator, 12 ports. The last stack unit is assigned eight pairs, the eight pair is unused.

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.

Example of Viewing VLAN Ports Using SNMP (Port Assigned)

```
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 TenGi 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.

Fetching Dynamic MAC Entries using SNMP

The Aggregator supports the RFC 1493 dot1d table for the default VLAN and the dot1q table for all other VLANs.

NOTE: The table contains none of the other information provided by the show vlan command, such as port speed or whether the ports are tagged or untagged.

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 60. 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 0/7, which is a member of the default VLAN, VLAN 1. The SNMP walk returns the values for dot1dTpFdbAddress, dot1dTpFdbPort, and dot1dTpFdbStatus.

Each object is comprised of 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 0/7, the manager returns the integer 118.

Example of Fetching Dynamic MAC Addresses on the Default VLAN

```
----------------------------MAC Addresses on Dell Networking
System-------------------------------
Dell#show mac-address-table
VlanId   Mac Address      Type     Interface      State
1       00:01:e8:06:95:ac Dynamic  Tengig 0/7    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 Dynamic MAC Addresses on a Non-default VLANs

In the following example, TenGigabitEthernet 0/7 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 Dell Networking
System-------------------------------
Dell#show mac-address-table
VlanId  Mac Address        Type     Interface      State
1000    00:01:e8:06:95:ac  Dynamic  Tengig 0/7    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

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 Dell Networking
System-------------------------------
Dell(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
```

Deriving Interface Indices

The Dell Networking OS assigns an interface number to each (configured or unconfigured) physical and logical interface. Display the interface index number using the `show interface`command from EXEC Privilege mode, as shown in the following example.
The interface index is a binary number with bits that indicate the slot number, port number, interface type, and card type of the interface. The Dell Networking OS converts this binary index number to decimal, and displays it in the output of the `show interface` command.

Starting from the least significant bit (LSB):

- the first 14 bits represent the card type
- the next 4 bits represent the interface type
- the next 7 bits represent the port number
- the next 5 bits represent the slot number
- the next 1 bit is 0 for a physical interface and 1 for a logical interface
- the next 1 bit is unused

For example, the index 44634369 is 10101010010001000100000001 in binary. The binary interface index for TenGigabitEthernet 0/4 of an Aggregator. Notice that the physical/logical bit and the final, unused bit

are not given. The interface is physical, so this must be represented by a 0 bit, and the unused bit is always 0. These two bits are not given because they are the most significant bits, and leading zeros are often omitted.

For interface indexing, slot and port numbering begins with binary one. If the Dell Networking system begins slot and port numbering from 0, binary 1 represents slot and port 0. In the Aggregator the first interface is 0/1 and 0/0s Ifindex is unused and Ifindex creation logic is not changed. Because Zero is reserved for logical interfaces, it starts from 1. For the first interface, port number is set to 1. Adding it causes an increment by 1 for the next interfaces, so it only starts from 2. Therefore, the port number is set to 4 for 0/3.

Example of Deriving the Interface Index Number

```
Dell#show interface tengig 0/2
TenGigabitEthernet 0/2 is up, line protocol is up
Hardware is Dell Force10Eth, address is 00:01:e8:0d:b7:4e
    Current address is 00:01:e8:0d:b7:4e
Interface index is 72925242
[output omitted]
```

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.

NOTE: The interface index does not change if the interface reloads or fails over. If the unit is renumbered (for any reason) the interface index changes during a reload.

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.6027.3.2.1.1.1.1.1.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.1.2 = INTEGER: 2
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.2.1 = Hex-STRING: 00 01 E8 13 A5 C7
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.2.2 = Hex-STRING: 00 01 E8 13 A5 C8
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.3.1 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.3.2 = INTEGER: 1107755010
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.4.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.4.2 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.5.1 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.5.2 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.6.1 = STRING: "Tengig 0/4 " << Channel
member for Po1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.6.2 = STRING: "Tengig 0/5 " << 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 you learn the MAC address for the LAG, the LAG status also displays.

```
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.6027.3.2.1.1.4.1.2.1.0.0.0.0.0.0.1.1 = Hex-STRING: 00 00
00 00 00 01
dot3aCurAggIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.3.1.0.0.0.0.0.0.1.1 = INTEGER: 1
dot3aCurAggStatus
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.4.1.0.0.0.0.0.0.1.1 = INTEGER: 1 <<
Status active, 2 – status
inactive
```

For L3 LAG, you do not have this support.

```
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: Tengig 0/1"
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: Tengig 0/1"
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"
```

Entity MIBS

The Entity MIB provides a mechanism for presenting hierarchies of physical entities using SNMP tables. The Entity MIB contains the following groups, which describe the physical elements and logical elements of a managed system. The following tables are implemented for the Aggregator.

- **Physical Entity:** A physical entity or physical component represents an identifiable physical resource within a managed system. Zero or more logical entities may utilize a physical resource at any given time. Determining which physical components are represented by an agent in the EntPhysicalTable is an implementation-specific matter. Typically, physical resources (for example, communications ports, backplanes, sensors, daughter-cards, power supplies, and the overall chassis), which you can manage via functions associated with one or more logical entities, are included in the MIB.
- **Containment Tree:** Each physical component may be modeled as contained within another physical component. A containment-tree is the conceptual sequence of *entPhysicalIndex* values that uniquely specifies the exact physical location of a physical component within the managed system. It is generated by following and recording each *entPhysicalContainedIn* instance up the tree towards the root, until a value of zero indicating no further containment is found.

Example of Sample Entity MIBS outputs

```
Dell#show inventory optional-module
Unit Slot Expected Inserted Next Boot Status/Power(On/Off)
--------------------------------------------------------------------------
1 0 SFP+ SFP+ AUTO Good/On
```

```
1 1 QSFP+ QSFP+ AUTO Good/On
* - Mismatch
Dell#
The status of the MIBS is as follows:
$ snmpwalk -c public -v 2c 10.16.150.162 .1.3.6.1.2.1.47.1.1.1.1.2
SNMPv2-SMI::mib-2.47.1.1.1.1.2.1 = ""
SNMPv2-SMI::mib-2.47.1.1.1.1.2.2 = STRING: "PowerEdge-FN-410S-IOA"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.3 = STRING: "Chassis 0 container"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.4 = STRING: "Module 0"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.5 = STRING: "Unit: 0 Port 1 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.6 = STRING: "Unit: 0 Port 2 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.7 = STRING: "Unit: 0 Port 3 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.8 = STRING: "Unit: 0 Port 4 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.9 = STRING: "Unit: 0 Port 5 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.10 = STRING: "Unit: 0 Port 6 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.11 = STRING: "Unit: 0 Port 7 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.12 = STRING: "Unit: 0 Port 8 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.13 = STRING: "Unit: 0 Port 9 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.14 = STRING: "Unit: 0 Port 10 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.15 = STRING: "Unit: 0 Port 11 10G Level"
SNMPv2-SMI::mib-2.47.1.1.1.1.2.16 = STRING: "Unit: 0 Port 12 10G Level"
shathishmuthu@login-maa-107 ~ $
```

SNMP Traps for Link Status

To enable SNMP traps for link status changes, use the `snmp-server enable traps snmp linkdown linkup` command.

Standard VLAN MIB

When the Aggregator is in Standalone mode, where all the 4000 VLANs are part of all server side interfaces as well as the single uplink LAG, it takes a long time (30 seconds or more) for external management entities to discover the entire VLAN membership table of all the ports. Support for current status OID in the standard VLAN MIB is expected to simplify and speed up this process. This OID provides 4000 VLAN entries with port membership bit map for each VLAN and reduces the scan for (4000 X Number of ports) to 4000.

Enhancements

1. The `dot1qVlanCurrentEgressPorts` MIB attribute has been enhanced to support logical LAG interfaces.
2. Current status OID in standard VLAN MIB is accessible over SNMP.
3. The bitmap supports 42 bytes for physical ports and 16 bytes for the LAG interfaces (up to a maximum of 128 LAG interfaces).
4. A 59 byte buffer bitmap is supported and in that bitmap:
   - First 42 bytes represent the physical ports.
   - Next 16 bytes represent logical ports 1-128.
   - An additional 1 byte is reserved for future.

Fetching the Switchport Configuration and the Logical Interface Configuration

Important Points to Remember

- The SNMP should be configured in the chassis and the chassis management interface should be up with the IP address.
- If a port is configured in a VLAN, the respective bit for that port will be set to 1 in the specific VLAN.
- In the aggregator, all the server ports and uplink LAG 128 will be in switchport. Hence, the respective bits are set to 1.

The following output is for the default VLAN.

Example of dot1qVlanCurrentUntaggedPorts output

```
snmpwalk -Os -c public -v 1 10.16.151.151 1.3.6.1.2.1.17.7.1.4.2.1.5
mib-2.17.7.1.4.2.1.5.0.1107525633 = Hex-STRING: FF FF FF FF 00 00 00 00 00 00
00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 01 00
```

The last byte is free byte . The bit for LAGs starts from 43 byte. If server LAG 1 is created with server ports Te 0/6 and Te 0/7, the respective bit for the ports are unset and the bit for LAG 1 is set in default VLAN. The corresponding output will be as follows:

```
snmpwalk -Os -c public -v 1 10.16.151.151 1.3.6.1.2.1.17.7.1.4.2.1.5
mib-2.17.7.1.4.2.1.5.0.1107525633 = Hex-STRING: F9 FF FF FF 00 00 00 00 00 00
00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 80 00 00 00 00 00
00 00 00 00 00 00 00 00 00 01 00
```

In the above example, the 43rd byte is set to 80. The 43rd byte is for LAG IDs from 1 to 8. But, only one LAG po 1 is set as switch port. Hence, the binary bits will be 10000000. If this converted to Hexadecimal, the value will be 80. Similarly, the first byte for Te 0/1 to Te 0/8 server ports, as the 6th and 7th byte are removed from switch port, the respective bits are set to 0. In binary, the value is 11111001 and the corresponding hex decimal value is F9.

In standalone mode, there are 4000 VLANs, by default. The SNMP output will display for all 4000 VLANs. To view a particular VLAN, issue the `snmp query` with VLAN interface ID.

```
Dell#show interface vlan 1010 | grep "Interface index"
                      Interface index is 1107526642
```

Use the output of the above command in the snmp query.

```
snmpwalk -Os -c public -v 1 10.16.151.151
1.3.6.1.2.1.17.7.1.4.2.1.4.0.1107526642
mib-2.17.7.1.4.2.1.4.0.1107526642 = Hex-STRING: F9 FF FF FF 00 00 00 00 00 00
00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 80 00 00 00 00 00
00 00 00 00 00 00 00 00 00 01 00
```

MIB Support to Display the Available Memory Size on Flash

Dell 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 61. MIB Objects for Displaying the Available Memory Size on Flash via SNMP

MIB Object	OID	Description
chStackUnitFlashUsageUtil	1.3.6.1.4.1.6027.3.19.1.2.8.1.6	Contains flash memory usage in percentage.

The chStackUnitUtilTable MIB table contains the chStackUnitFlashUsageUtil MIB object which contains the flash memory usage percent.

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 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 62. MIB Objects for Displaying the Software Core Files Generated by the System

MIB Object	OID	Description
chSysSwCoresTable	1.3.6.1.4.1.6027.3.19.1.2.9	This is the table that contains the list of software core files generated by the system.
chSysCoresEntry	1.3.6.1.4.1.6027.3.19.1.2.9.1	Entry number.
chSysCoresInstance	1.3.6.1.4.1.6027.3.19.1.2.9.1.1	Stores the indexed information about the available software core files.
chSysCoresFileName	1.3.6.1.4.1.6027.3.19.1.2.9.1.2	Contains the core file names and the file paths.
chSysCoresTimeCreated	1.3.6.1.4.1.6027.3.19.1.2.9.1.3	Contains the time at which core files are created.

MIB Object	OID	Description
chSysCoresStackUnitNumber	1.3.6.1.4.1.6027.3.19.1.2.9.1.4	Contains information that includes which stack unit or processor the core file was originated from.
chSysCoresProcess	1.3.6.1.4.1.6027.3.19.1.2.9.1.5	Contains information that includes the process names that generated each core file.

Viewing the Software Core Files Generated by the System

- To view the viewing 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.

52

Stacking

An Aggregator auto-configures to operate in standalone mode. To use an Aggregator in a stack, you must manually configure it using the CLI to operate in stacking mode.

Stacking is supported on the FN410S and FN410T Aggregators with ports 9 and 10 as the stack ports. The Aggregator supports both ring and daisy-chain topology and stacking of the same type. FN 410S and FN 410T Aggregators support two-unit in-chassis stacking and up to six units stacking across the chassis.

In Stack mode, the lower two external Ethernet ports (ports 9 and 10) operate as stack links. In Programmable MUX (PMUX) mode, you can configure any of the external Ethernet ports to operate as stack links.

Stacking provides a single point of management for high availability and higher throughput. To configure a stack, you must use the CLI.

Supported Modes

Stacking, PMUX

Configuring a Switch Stack

To configure and bring up a switch stack, follow these steps:

1. Connect the ports on the base module of two Aggregators using 10G direct attach or SFP+ fibre cables.
2. Configure each Aggregator to operate in stacking mode.
3. Reload each Aggregator, one after the other in quick succession.

Stacking Prerequisites

Before you cable and configure a stack of the Aggregators, review the following prerequisites.

- All Aggregators in the stack must be powered up with the initial or startup configuration before you attach the cables.
- All stacked Aggregators must run the same Dell Networking OS version. To check the version that a switch is running, use the `show version` command. To download a Dell Networking OS version, go to http://support.dell.com.
- Stacking is supported only with other Aggregators. A maximum of six Aggregators are supported in a single stack.
- A maximum of four stack groups is supported on a stacked Aggregator.
- Interconnect the stack units by following the instructions in Cabling the Switch Stack.
- You cannot stack a Standalone IOA and a PMUX.

Master Selection Criteria

A Master is elected or re-elected based on the following considerations, in order:

1. The switch with the highest priority at boot time.
2. The switch with the highest MAC address at boot time.
3. A unit is selected as Standby by the administrator, and a fail over action is manually initiated or occurs due to a Master unit failure.

No record of previous stack mastership is kept when a stack loses power. As it reboots, the election process will once again determine the Master and Standby switches. As long as the priority has not changed on any members, the stack will retain the same Master and Standby.

NOTE: Each stack members' role (including the Master and Standby) can be defined by the user at any time by setting the priority.

If the entire stack is powered OFF and ON again, the unit that was the Master before the reboot will remain the Master after the stack resumes operation. However, when a stack is powered on, all members are in **sleep** mode for 5 seconds while waiting on the previous Master to join the stack. If the previous Master fails to join within 5 seconds, the remaining members (including the Standby) elect a new Master.

Configuring Priority and stack-group

Perform the following steps to configure the priorities and stack-groups for each of the switches.

1. Set the priorities for the stack-unit.

   CONFIGURATION mode

   ```
   stack-unit unit-number priority 1-14
   ```

   ```
   Dell(conf)# stack-unit 0 priority 12
   ```

   Setting the priority will determine which switch will become the management (Master) switch. The switch with the highest priority number is elected Master. The default priority is 0.

   NOTE: It is best practice to assign priority values to all switches before stacking them in order to acquire and retain complete control over each units role in the stack.

2. Configure the stack-group for each stack-unit.

   CONFIGURATION mode

   ```
   stack-unit unit-no stack-group stack-group-id
   ```

   ```
   Dell(conf)# stack-unit 0 stack-group 0
   Dell (conf)#02:26:05: %STKUNIT0-M:CP %IFMGR-6-STACK_PORTS_ADDED:
   Ports Fo 0/33 have been configured as stacking ports.
   Please save and reload for config to take effect
   ```

   Stack-groups are easier to think of simply as stack ports. For example, using the `stack-group 0` command simply turns the lower port (port 9) into a stacking port. Similarly, `stack-group 1`, `stack-group 2` and `stack-group 3` commands correspond to ports 10, 11 and 12 respectively.

   NOTE: Stack-group is supported only in PMUX mode.

3. Continue to run the `stack-unit 0 stack-group <0-3>` command to add additional stack ports to the switch, using the stack-group mapping.

Cabling the Switch Stack

Dell PowerEdge FN I/O Aggregators are connected to operate as a single stack in a ring topology using the SFP+ or Base-T ports on the front end ports 9 and 10. To create a stack in either a ring or daisy-chain topology, you can use two units on the same chassis or up to six units across multiple chassis.

Prerequisite: Before you attach the stacking cables, all Aggregators in the stack must be powered up with the default or reconfigured settings.

To connect stacking ports, use only SFP+ transceivers, SFP+ cables and BaseT and its supported cables (separately purchased). For example:

1. Insert a cable in port 9 on the first aggregator.
2. Connect the cable to port 10 on the next aggregator.
3. Continue this pattern on up to 6 aggregators.
4. Connect a cable from port 9 on the last aggregator to port 10 on the first aggregator. This creates a ring topology.

NOTE: The resulting topology allows the stack to function as a single switch with resilient failover capabilities.

Accessing the CLI

To configure a stack, you must access the stack master in one of the following ways.

- For remote out-of-band management (OOB), enter the OOB management interface IP address into a Telnet or secure shell (SSH) client and log in to the switch using the user ID and password to access the CLI.
- For local management, use the attached console connection to the master switch to log in to the CLI. Console access to the stack CLI is available on the master only.
- For remote in-band management from a network management station, enter the virtual local area network (VLAN) IP address of the management port and log in to the switch to access the CLI.

Configuring and Bringing Up a Stack

After you attach the cables in a stack of Aggregators, perform the following steps to configure and bring up the stack:

1. Set up a connection to the CLI on an Aggregator as described in Accessing the CLI.
2. Connect the terminal to the console port on an Aggregator. Enter the following commands to access the CLI and configure the base module ports for stacking mode:

   ```
   Login: username
   Password: *****
   Dell> enable
   Dell# configure
   Dell(conf)# stack-unit 0 iom-mode stack
   ```

   Where `stack-unit 0` defines the default stack-unit number in the initial configuration of a switch.
3. Save the stacking configuration on the SFP+ or 10G Base-T ports:

   ```
   Dell# write memory
   ```

4. Reboot the Aggregator by entering the `reload` command in EXEC Privilege mode:
   ```
   Dell# reload
   ```

Repeat the above steps on each Aggregator in the stack by entering the `stackunit 0 iom-mode stack` command and saving the configuration.

If the stacked switches all reboot at approximately the same time, the Aggregator with the highest MAC address is automatically elected as the master switch. The Aggregator with the next highest MAC address is elected as the standby master.

NOTE: You can ensure that a stacked switch becomes the master by rebooting the switch first, and waiting for it to come up before rebooting the second switch in the stack. It is recommended to have the boot image for the stack units via boot from flash.

The Status LED on the Aggregator that is elected as the master switch is blue. Perform VLAN and other software configuration for the stack by connecting to the console port on the master unit. To determine which switch is the stack master, enter the `show system` command at the terminal.

To remove an Aggregator from the stack, enter the `stack-unit 0 iom-mode standalone` command in Global configuration mode, save the configuration and reload the Aggregator for the change to take effect.

In stack mode, all VLAN membership are removed and the port is assigned only to the default VLAN1. You must configure additional VLAN membership, as required.

Adding a Stack Unit

You can add a new unit to an existing stack both when the unit has no stacking ports (stack groups) configured and when the unit already has stacking ports configured. If the units to be added to the stack have been previously used, they are assigned the smallest available unit ID in the stack.

To add a standalone Aggregator to a stack, follow these steps:

1. Power on the switch.
2. Attach SFP+ or direct attach cables to connect 10G ports on the switch to one or more switches in the stack.
3. Log on to the CLI and enter global configuration mode.
   ```
   Login: username

   Password: *****

   Dell> enable

   Dell# configure
   ```
4. Configure the Aggregator to operate in stacking mode.
   CONFIGURATION mode
   ```
   stack-unit 0 iom-mode

   stack
   ```
5. Reload the switch. Dell Operating System 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.

EXEC Privilege mode

```
reload
```

If an Aggregator is already configured to operate in stacking mode, simply attach SFP+ or direct attach cables to connect 10G ports on the base module of each stacked Aggregator. The new unit synchronizes its running and startup configurations with the stack.

Dell Networking OS Behavior: When you add a new Aggregator to a stack:

- If the new unit has been configured 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.

Resetting a Unit on a Stack

Use the following `reset` command to reload any of the member units or the standby in a stack. If you try to reset the stack master, the following error message is displayed:

```
% Error: Reset of master unit is not allowed.
```

To reset a unit on a stack, use the following command:

- Reset any designated stack member, except the management unit.
  EXEC Privilege mode

  ```
  reset stack-unit unit-number {hard}
  ```
- Hard reset any stack unit including master unit.
  EXEC Privilege mode

  ```
  power-cycle stack-unit [unit-number]
  ```

Removing an Aggregator from a Stack

To remove an Aggregator from a stack, follow the below steps:

1. Disconnect the stacking cables from the unit. The unit can be powered on or off and can be online or offline.
2. Log on to the CLI and enter Global Configuration mode.

   ```
   Login: username
   Password: *****
   Dell> enable
   Dell# configure
   ```
3. Configure the Aggregator to operate in standalone mode.
   CONFIGURATION

   ```
   stack-unit 0 iom-mode standalone
   ```
4. Log on to the CLI and reboot each switch, one after another, in as short a time as possible.
   EXEC PRIVILEGE

   reload

The switch functions in standalone mode but retains the running and startup configuration that was last synchronized by the master switch while it operated as a stack unit.

Merging Two Operational Stacks

The recommended procedure for merging two operational stacks is as follows:

1. Always power off all units in one stack before connecting to another stack.
2. Add the units as a group by unplugging one stacking cable in the operational stack and physically connecting all unpowered units.
3. Completely cable the stacking connections, making sure the redundant link is also in place.

Two operational stacks can also be merged by reconnecting stack cables without powering down units in either stack. Connecting a powered-up standalone unit to an existing stack leads to same behavior as when merging two operational stacks. In such cases, Manager re-election is done and the Manager with the higher MAC address wins the election. The losing stack manager resets itself and all its member units. After the reset, all the losing stack members join the winning stack to form a single stack. The winning stack remains functional through the merge process. If the stack merge is performed in this way, then it is strongly recommended that the user set the admin management preference of the desired winner stack manager to a higher value than the stack manager that should lose the election.

NOTE: In case of a stack, when one unit member resets and joins the stack, momentary drops will be observed, which is recovered after few seconds.

Verifying a Stack Configuration

The following lists the status of a stacked switch according to the color of the System Status light emitting diodes (LEDs) on its front panel.

- Blue indicates the switch is operating as the stack master or as a standalone unit.
- Off indicates the switch is a member or standby unit.
- Amber indicates the switch is booting or a failure condition has occurred.

Using Show Commands

To display information on the stack configuration, use the `show` commands on the master switch.

- Displays stacking roles (master, standby, and member units) and the stack MAC address.

  ```
  show system [brief]
  ```
- Displays the stack groups allocated on a stacked switch. The range is from 0 to 5.

  ```
  show system stack-unit unit-number stack-group configured
  ```
- Displays the port numbers that correspond to the stack groups on a switch. The valid stack-unit numbers are from 0 to 5.

  ```
  show system stack-unit unit-number stack-group
  ```
- Displays the type of stack topology (ring or daisy chain) with a list of all stacked ports, port status, link speed, and peer stack-unit connection.

  ```
  show system stack-ports [status | topology]
  ```

Troubleshooting a Switch Stack

To perform troubleshooting operations on a switch stack, use the following commands on the master switch.

1. Displays the status of stacked ports on stack units.

   `show system stack-ports`
2. Displays the master standby unit status, failover configuration, and result of the last master-standby synchronization; allows you to verify the readiness for a stack failover.

   `show redundancy`
3. Displays input and output flow statistics on a stacked port.

   `show hardware stack-unit` *`unit-number`* `stack-port` *`port-number`*
4. Clears statistics on the specified stack unit. The valid stack-unit numbers are from 0 to 5.

   `clear hardware stack-unit` *`unit-number`* `counters`
5. Displays the current operational mode of the Aggregator (standalone or stacking) and the mode in which the Aggregator will operate at the next reload.

   `show system stack-unit` *`unit-number`* `iom-mode`

Failure Scenarios

The following sections describe some of the common fault conditions that can happen in a switch stack and how they are resolved.

Stack Member Fails

- **Problem**: A unit that is not the stack master fails in an operational stack.
- **Resolution**: If a stack member fails in a daisy chain topology, a split stack occurs. If a member unit fails in a ring topology, traffic is re-routed over existing stack links.

The following syslog messages are generated when a member unit fails:

```
Dell#May 31 01:46:17: %STKUNIT3-M:CP %IPC-2-STATUS: target stack unit 4 not
responding

May 31 01:46:17: %STKUNIT3-M:CP %CHMGR-2-STACKUNIT_DOWN: Major alarm: Stack
unit 4 down - IPC
timeout

Dell#May 31 01:46:17: %STKUNIT3-M:CP %IFMGR-1-DEL_PORT: Removed port: Te
4/1-32,41-48, Fo 4/
49,53

Dell#May 31 01:46:18: %STKUNIT5-S:CP %IFMGR-1-DEL_PORT: Removed port: Te
4/1-32,41-48, Fo 4/
49,53
```

Unplugged Stacking Cable

- **Problem**: A stacking cable is unplugged from a member switch. The stack loses half of its bandwidth from the disconnected switch.
- **Resolution**: Intra-stack traffic is re-routed on another link using the redundant stacking port on the switch. A recalculation of control plane and data plane connections is performed.

Master Switch Fails

- **Problem**: The master switch fails due to a hardware fault, software crash, or power loss.
- **Resolution**: A failover procedure begins:

1. Keep-alive messages from the Aggregator master switch time out after 60 seconds and the switch is removed from the stack.
2. The standby switch takes the master role. Data traffic on the new master switch is uninterrupted. Protocol traffic is managed by the control plane.
3. A member switch is elected as the new standby. Data traffic on the new standby is uninterrupted. The control plane prepares for operation in Warm Standby mode.

Stack-Link Flapping Error

Problem/Resolution: Stacked Aggregators monitor their own stack ports and disable any stack port that flaps five times within 10 seconds. If the stacking ports that flap are on the master or standby, KERN-2-INT error messages note the units.

To re-enable a downed stacking port, power cycle the stacked switch on which the port is installed.

The following is an example of the stack-link flapping error message.

```
---------------------------------------MANAGMENT
UNIT-------------------------------------
Error: Stack Port 49 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.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times
within 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.
```

Master Switch Recovers from Failure

- **Problem**: The master switch recovers from a failure after a reboot and rejoins the stack as the standby unit or member unit. Protocol and control plane recovery requires time before the switch is fully online.
- **Resolution**: When the entire stack is reloaded, the recovered master switch becomes the master unit of the stack.

Stack Unit in Card-Problem State Due to Incorrect Dell Networking OS Version

- **Problem**: A stack unit enters a Card-Problem state because the switch has a different Dell Networking OS version than the master unit. The switch does not come online as a stack unit.
- **Resolution**: To restore a stack unit with an incorrect Dell Networking OS version as a member unit, disconnect the stacking cables on the switch and install the correct Dell Networking OS version. Then add the switch to the stack as described in Adding a Stack Unit. To verify that the problem has been resolved and the stacked switch is back online, use the `show system brief` command.

```
Dell#show system brief
Stack MAC : 00:1e:c9:f1:00:9b
```

```
-- Stack Info --
Unit UnitType  Status       ReqTyp          CurTyp          Version    Ports
-----------------------------------------------------------------------------
0   Management online         PE-FN-410S-IOA  PE-FN-410S-IOA  1-0(0-1864) 12
1   Standby    card problem PE-FN-410S-IOA  unknown               12
2   Member    not present
3   Member    not present
4   Member    not present
5   Member    not present
```

Card Problem — Resolved

```
Dell#show system brief
Stack MAC : 00:1e:c9:f1:04:82

-- Stack Info --
Unit UnitType  Status       ReqTyp        CurTyp          Version  Ports
-----------------------------------------------------------------------
0   Management online     PE-FN-410S-IOA  PE-FN-410S-IOA  1-0(0-1864) 12
1   Standby online PE-FN-410S-IOA  PE-FN-410S-IOA  1-0(0-1864) 12
2   Member not present
3   Member not present
4  Member not present
5  Member not present
```

Stack Unit in Card-Problem State Due to Configuration Mismatch

- **Problem**: A stack unit enters a Card-Problem state because there is a configuration mismatch between the logical provisioning stored for the stack-unit number on the master switch and the newly added unit with the same number.
- **Resolution**: From the master switch, reload the stack by entering the `reload` command in EXEC Privilege mode. When the stack comes up, the card problem will be resolved.

Upgrading a Switch Stack

To upgrade all switches in a stack with the same Dell Networking OS version, follow these steps.

1. Copy the new Dell Networking OS image to a network server.
2. Download the Dell Networking OS image by accessing an interactive CLI that requests the server IP address and image filename, and prompts you to upgrade all member stack units.

   EXEC Privilege mode

   upgrade system { *flash:* | *ftp:* | *scp:* | *tftp:* | *usbflash:* } *partition*

   Specify the system partition on the master switch into which you want to copy the Dell Networking OS image. The system then prompts you to upgrade all member units with the new Dell Networking OS version.

   The valid values are a: and b:.
3. Reboot all stack units to load the Dell Networking OS image from the same partition on all switches in the stack.

   CONFIGURATION mode

   `boot system stack-unit all primary system` *partition*
4. Save the configuration.

   EXEC Privilege

```
write memory
```

5. Reload the stack unit to activate the new Dell Networking OS version.

   CONFIGURATION mode

```
reload
```

Example of Upgrading all Stacked Switches

The following example shows how to upgrade all switches in a stack, including the master switch.

```
Dell# upgrade system ftp: A:
Address or name of remote host []: 10.11.200.241
Source file name []: //FTOS-XL-8.3.17.0.bin
User name to login remote host: ftp
Password to login remote host:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Erasing IOM Primary Image, please wait
.!.....................................................................
...
.....................................Writing.........................
...
.....................................................................
...
.....................................................................
...
31972272 bytes successfully copied
System image upgrade completed successfully.
Upgrade system image for all stack-units [yes/no]: yes
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!
!
Image upgraded to all
Dell# configure
Dell(conf)# boot system stack-unit all primary system: A:
Dell(conf)# end
Dell# write memory
Jan 3 14:01:48: %STKUNIT0-M:CP %FILEMGR-5-FILESAVED: Copied running-config to
startup-config
in flash by default
Synchronizing data to peer Stack-unit
!!!!
Dell# reload
Proceed with reload [confirm yes/no]: yes
```

Upgrading a Single Stack Unit

Upgrading a single stacked switch is necessary when the unit was disabled due to an incorrect Dell Networking OS version. This procedure upgrades the image in the boot partition of the member unit from the corresponding partition in the master unit.
To upgrade an individual stack unit with a new Dell Networking OS version, follow the below steps:

1. Download the Dell Networking OS image from the master's boot partition to the member unit, and upgrade the relevant boot partition in the single stack-member unit.

   EXEC Privilege mode

```
upgrade system stack-unit unit-number partition
```

2. Reboot the stack unit from the master switch to load the Dell Networking OS image from the same partition.

CONFIGURATION mode

```
boot system stack-unit unit-number primary system partition
```

3. Save the configuration.
   EXEC Privilege mode

```
write memory
```

4. Reset the stack unit to activate the new Dell Networking OS version.
   EXEC Privilege mode

```
power-cycle stack-unit unit-number
```

Example of Upgrading a Single Stack Unit

The following example shows how to upgrade an individual stack unit.

```
Dell# upgrade system stack-unit 1 A:
!!!
```

53

Storm Control

Storm control is supported on the Dell networking OS.

The storm control feature allows you to control unknown-unicast, muticast, and broadcast control traffic on Layer 2 and Layer 3 physical interfaces.

Dell Networking OS Behavior: The Dell Networking OS supports broadcast control (the `storm-control broadcast` command) for Layer 2 and Layer 3 traffic. The minimum number of packets per second (PPS) that storm control can limit 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 multicast [interface]` command.

EXEC Privilege

Example:

```
Dell#show storm-control multicast Tengigabitethernet 1/1

Multicast storm control configuration

Interface  Direction     Packets/Second
----------------------------------------
Te 1/1     Ingress       5

Dell#
```

To display the storm control unknown-unicast configuration, use the `show storm-control unknown-unicast [interface]` command.

EXEC Privilege

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.
You can only configure storm control for ingress traffic in INTERFACE mode. If you configure storm control from both INTERFACE and CONFIGURATION mode, the INTERFACE mode configurations override the CONFIGURATION mode configurations.

- Configure storm control.
  INTERFACE mode
- Configure the percentage of broadcast traffic allowed on an interface (ingress only).
  INTERFACE mode

  ```
  storm-control broadcast packets_per_second in
  ```
- Configure the percentage 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.
You can configure storm control for ingress traffic in CONFIGURATION mode. Do not apply per-virtual local area network (per-VLAN) quality of service (QoS) on an interface that you have enabled storm-control (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
  ```

54

Broadcast Storm Control

On the Aggregator, the broadcast storm control feature is enabled by default on all ports, and disabled on a port when an iSCSI storage device is detected. Broadcast storm control is re-enabled as soon as the connection with an iSCSI device ends.

Broadcast traffic on Layer 2 interfaces is limited or suppressed during a broadcast storm. You can view the status of a broadcast-storm control operation by using the `show io-aggregator broadcast storm-control status` command. You can disable broadcast storm control by using the `no io-aggregator broadcast storm-control` command.

Dell Networking OS Behavior: If broadcast traffic exceeds 1000 Mbps, the Aggregator limits it to 1000 Mbps per port-pipe.

Supported Modes

Standalone, PMUX, VLT, Stacking

Disabling Broadcast Storm Control

To disable broadcast storm control on an Aggregator, use the `no io-aggregator broadcast storm-control` command from CONFIGURATION mode.
To re-enable broadcast storm control, enter the `io-aggregator broadcast storm-control` command.

Displaying Broadcast-Storm Control Status

To display the status of a current storm control operation, use the `show io-aggregator broadcast storm-control status` command from EXEC Privilege mode.

Configuring Storm Control

The following configurations are available only in PMUX mode.

1. To configure the percentage of broadcast traffic allowed on an interface, use the `storm-control broadcast [`*`packets_per_second`* `in]` command from INTERFACE mode.
2. To configure the percentage of multicast traffic allowed on an interface, use the `storm-control multicast [`*`packets_per_second`* `in]` command from INTERFACE mode.
3. To configure the percentage of unknown-unicast traffic allowed on an interface, use the `storm-control unknown-unicast [`*`packets_per_second`* `in]` command from INTERFACE mode.

55

Spanning Tree Protocol (STP)

Dell Networking OS supports spanning tree protocol (STP).

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.

The Dell Networking OS supports three other variations of spanning tree, as shown in the following table.

Table 63. Dell Networking OS Supported Spanning Tree Protocols

Dell 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
- Removing an Interface from the Spanning Tree Group
- Modifying Global Parameters
- Modifying Interface STP Parameters
- Enabling PortFast
- Prevent Network Disruptions with BPDU Guard
- STP Root Guard
- SNMP Traps for Root Elections and Topology Changes

Important Points to Remember

- STP is disabled by default.
- The Dell Networking operating system (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.

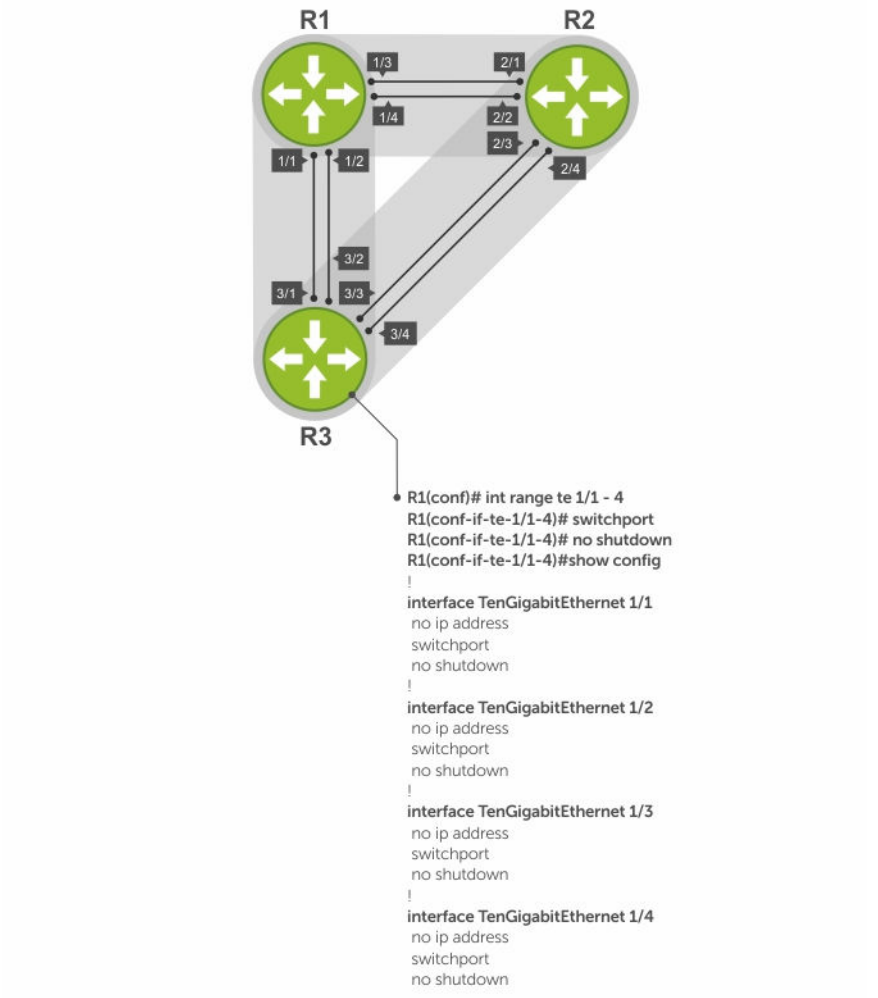


Figure 110. Example of Configuring Interfaces for Layer 2 Mode

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
   ```

Example of the `show config` Command

To verify that an interface is in Layer 2 mode and enabled, use the `show config` command from INTERFACE mode.

```
Dell(conf-if-te-1/1)#show config
!
interface TenGigabitEthernet 1/1
  no ip address
  switchport
  no shutdown
Dell(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.

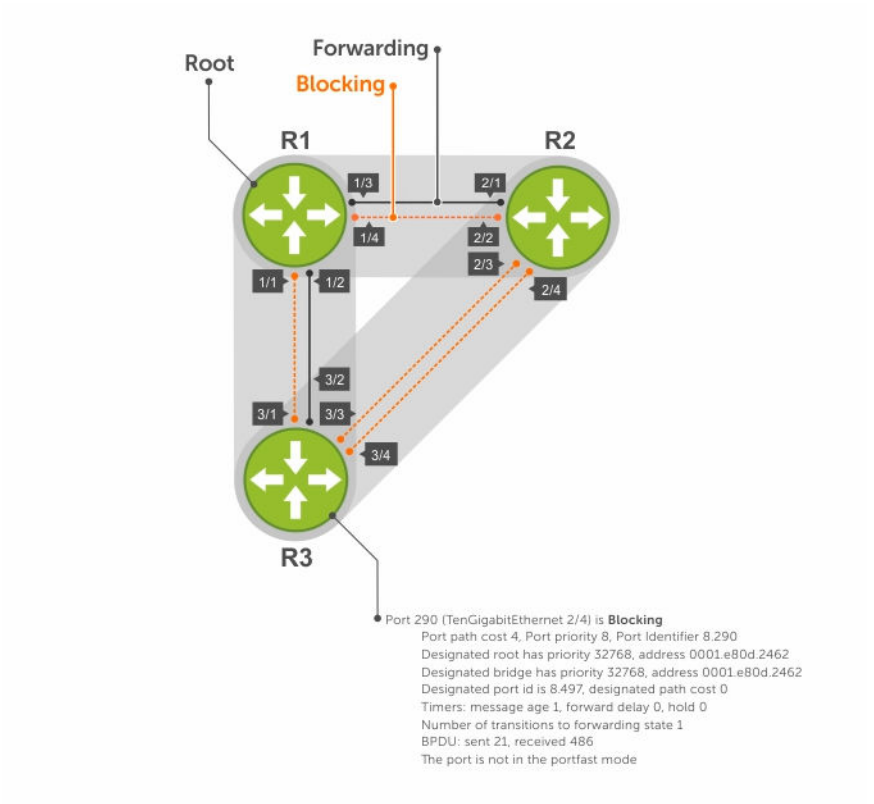


Figure 111. Spanning Tree Enabled Globally

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.

Example of Verifying Spanning Tree is Enabled

```
Dell(conf)#protocol spanning-tree 0
Dell(conf-span)#show config
!
protocol spanning-tree 0
  no disable
Dell#
```

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.

Example of Viewing Spanning Tree Configuration

```
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 (GigabitEthernet 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 GigabitEthernet 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 (GigabitEthernet 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 (GigabitEthernet 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.

Example of Verifying a Port Participates in Spanning Tree

```
Dell#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
-------------- ------ ---- ---- --- ----- --------------------
Tengig 1/1  8.496  8     4 DIS   0     32768 0001.e80d.2462  8.496
Tengig 1/2  8.497  8     4 DIS   0     32768 0001.e80d.2462  8.497
Tengig 1/3  8.513  8     4 FWD   0     32768 0001.e80d.2462  8.513
Tengig 1/4  8.514  8     4 FWD   0     32768 0001.e80d.2462  8.514
Dell#
```

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
  ```

Removing an Interface from the Spanning Tree Group

To remove a Layer 2 interface from the spanning tree topology, use the following command.

- Disable spanning tree on a Layer 2 interface.
  INTERFACE mode

  ```
  no 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 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 64. STP Default Values

STP Parameters	Default Value
Forward Delay	15 seconds
Hello Time	2 seconds
Max Age	20 seconds
Port Cost • 40-Gigabit Ethernet interfaces • 10-Gigabit Ethernet interfaces • Port Channel with 40-Gigabit Ethernet interfaces • Port Channel with 10-Gigabit Ethernet interfaces	• 1 • 2 • 1 • 1
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 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.

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 Port Fast

The Port Fast 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.

PDU Filtering enabled on an interface stops sending and receiving BPDUs on the port fast enabled ports. When you enable BPDU guard and BPDU filter on the port, then BPDU filter takes the highest precedence. By default BPDU filtering on an interface is disabled.

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] |
  bpdufilter]
  ```

Example of Verifying Port Fast is Enabled on an Interface

To verify that Port Fast is enabled on a port, use the `show spanning-tree` command from EXEC Privilege mode or the `show config` command from INTERFACE mode. Dell Networking recommends using the `show config` command.

```
Dell#(conf-if-te-1/1)#show conf
!
interface Tengigabitethernet 1/1
  no ip address
  switchport
  spanning-tree 0 portfast
  no shutdown
Dell#(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 the system displays the following message.

```
3w3d0h: %RPM0-P:RP2 %SPANMGR-5-BPDU_GUARD_RX_ERROR: Received Spanning Tree BPDU
on
BPDU guard port. Disable GigabitEthernet 3/41.
```

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 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 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).

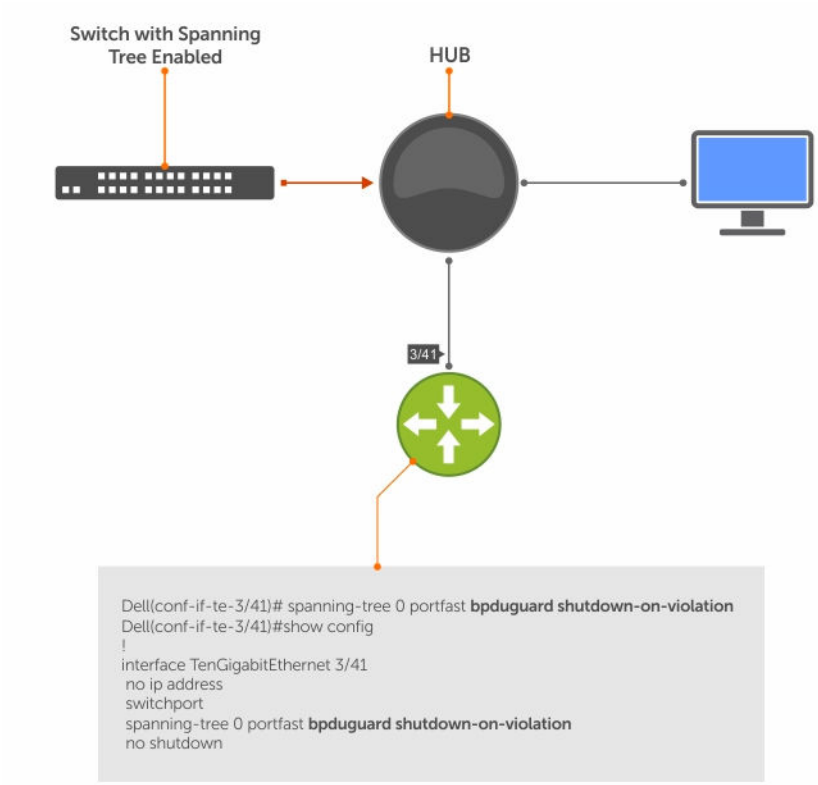


Figure 112. Enabling BPDU Guard

Dell Networking OS Behavior: BPDU guard and BPDU filtering (refer to Removing an Interface from the Spanning Tree Group) 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.

Example of Blocked BPDUs

```
Dell#show spanning-tree 0 brief
  Executing IEEE compatible Spanning Tree Protocol
    Root ID Priority 32768, Address 0001.e88a.fdb3 Cost 1
    Root Port 2 (Port-channel 1)
    Root Bridge hello time 2, max age 20, forward delay 15
    Bridge ID Priority 32768, Address 001e.c9f1.00cf
    Configured hello time 2, max age 20, forward delay 15
    Bpdu filter disabled globally
Interface                      Designated
Name    PortID Prio Cost Sts Cost Bridge ID             PortID
---------- ------- ---- ------ ----------- ------ ------ ------
Po 1    8.2    8    1    FWD  0   32768 0001.e88a.fdb3  8.2
Te 3/20 8.317  8    4    EDS  1   32768 001e.c9f1.00cf  8.317
Te 4/20 8.373  8    4    FWD  1   32768 001e.c9f1.00cf  8.373
Te 4/21 8.374  8    4    FWD  1   32768 001e.c9f1.00cf  8.374
Dell#show ip int br ten 3/20
Interface IP-Address OK Method Status Protocol
TenGigabitEthernet 3/20 unassigned YES None up up
Dell#
```

Global BPDU Filtering

When BPDU Filtering is enabled globally, it stops transmitting BPDUs on the operational port fast enabled ports by default.

When it receives BPDUs, it automatically participates in the spanning tree. By default global bpdu filtering is disabled.

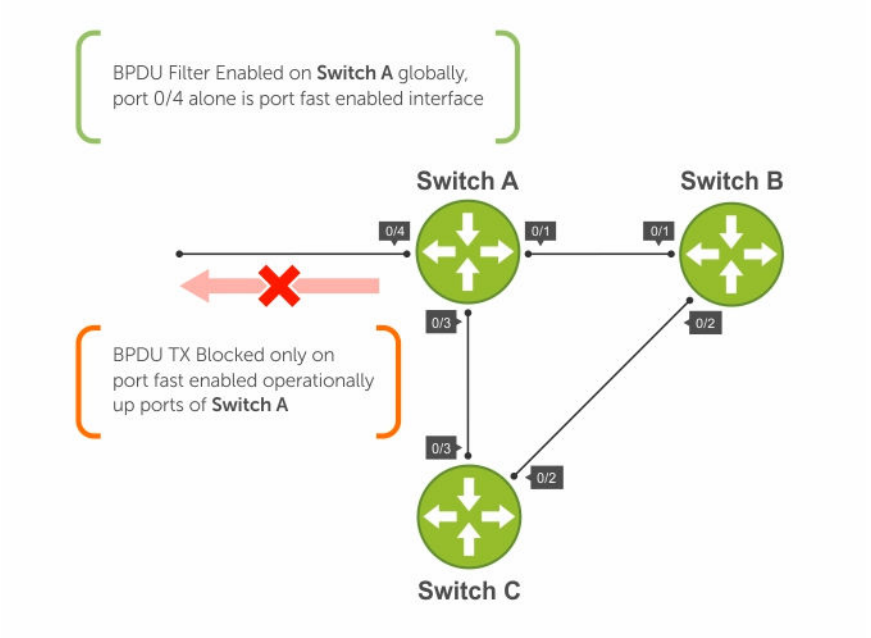


Figure 113. BPDU Filtering Enabled Globally

Interface BPDU Filtering

When BPDU Filtering is enabled on an interface, it should stop sending and receiving BPDUs on the port fast enabled ports.

When BPDU guard and BPDU filter is enabled on the port, then BPDU filter takes the highest precedence. By default bpdu filtering on an interface is disabled.

Add your section content here.

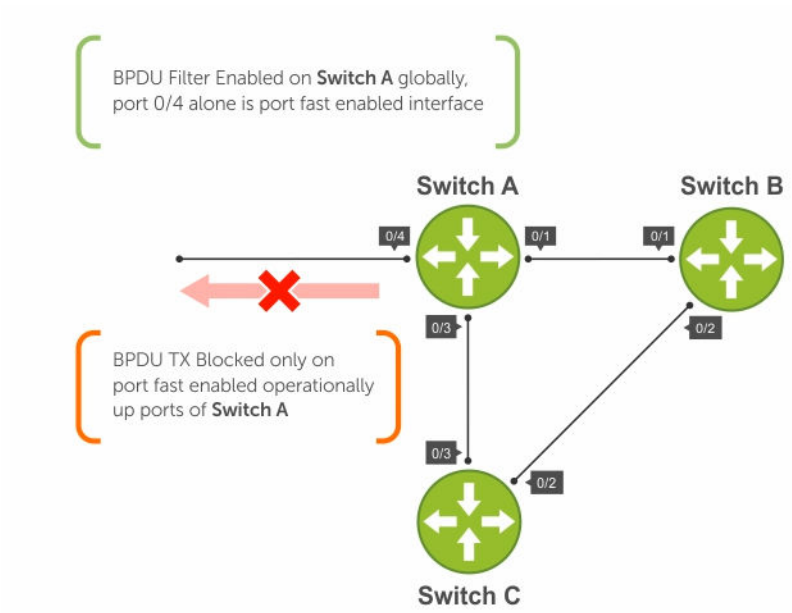


Figure 114. BPDU Filtering Enabled Globally

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**.

Example of Viewing STP Root Information

To view only the root information, use the `show spanning-tree root` command from EXEC privilege mode.

```
Dell#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
Dell#
```

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.

All incoming and outgoing traffic is blocked on an STP port in a Root-Inconsistent state. After the timeout period, the Switch C port automatically transitions to a Forwarding state as soon as device D stops sending BPDUs that advertise a lower priority.

If you enable a root guard on all STP ports on the links where the root bridge should not appear, you can ensure a stable STP network topology and avoid bridging loops.

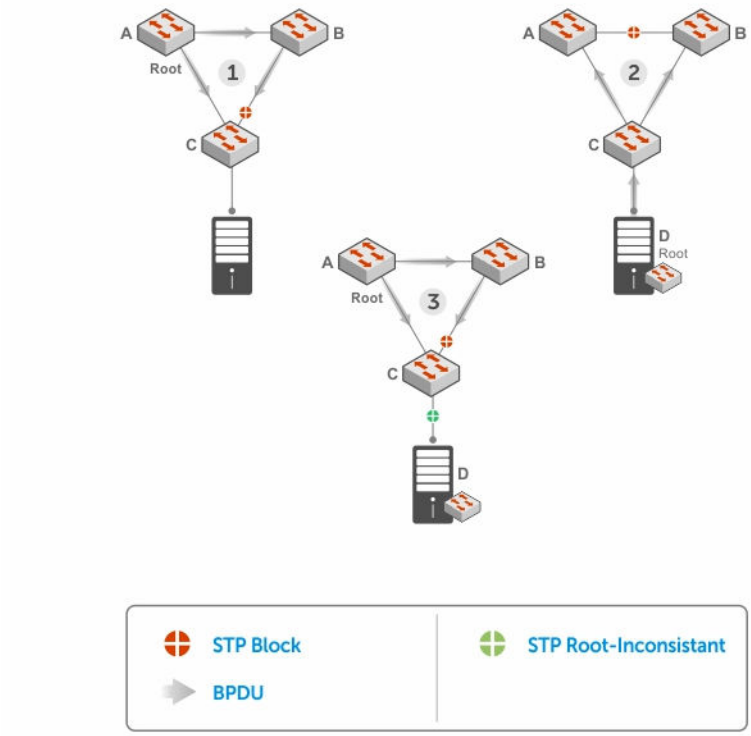


Figure 115. STP Root Guard Prevents Bridging Loops

Configuring Root Guard

Enable STP root guard on a per-port or per-port-channel basis.
Dell 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)
  - Multicast Source Discovery Protocol (MSDP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- When enabled on a port, root guard applies to all VLANs configured on the port.
- 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.

SNMP Traps for Root Elections and Topology Changes

To enable SNMP traps, use the following commands.

- Enable SNMP traps for STP state changes.
  `snmp-server enable traps stp`
- Enable SNMP traps for RSTP, MSTP, and PVST+ collectively.
  `snmp-server enable traps xstp`

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]`

Example of Viewing STP Guard Configuration

```
Dell#show spanning-tree 0 guard
Interface
Name    Instance Sts          Guard type     Bpdu Filter
--------- -------- --------- ----------      ----------
Tengig 0/1  0        INCON(Root)  Rootguard  No
Tengig 0/2  0        LIS          Loopguard  No
Tengig 0/3  0        EDS (Shut)   Bpduguard  No
Dell#
```

56

System Time and Date

The Aggregator auto-configures the hardware and software clocks with the current time and date. If necessary, you can manually set and maintain the system time and date using the CLI commands described in this chapter.

- Setting the Time for the Software Clock
- Setting the Time Zone
- Setting Daylight Savings Time

Supported Modes

Standalone, PMUX, VLT, Stacking

Setting the Time for the 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.

Example of the `clock set` Command

```
Dell#clock set 12:11:00 21 may 2012
Dell#
```

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, you must include the

differentiator between the 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.

Example of the `clock timezone` Command

```
Dell#conf
Dell(conf)#clock timezone Pacific -8
Dell#
```

Setting Daylight Savings Time

Dell Networking OS supports setting the system to daylight savings 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 to 1440. The default is **60 minutes**.

Example of the `clock summer-time` Command

```
Dell(conf)#clock summer-time pacific date Mar 14 2012 00:00 Nov 7 2012 00:00
Dell(conf)#
```

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 savings 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 end daylight saving time.
    * `first`: Enter the keyword `first` to end daylight saving time in the first week of the month.
    * `last`: Enter the keyword `last` to end 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**.

Example of the `clock summer-time recurring` Command

```
Dell(conf)#clock summer-time pacific recurring Mar 14 2012 00:00 Nov 7 2012
00:00
Dell(conf)#
```

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.

Example of Clock Summer-Time Recurring Parameters

```
Dell(conf)#clock summer-time pacific recurring ?
<1-4>     Week number to start
first     Week number to start
last      Week number to start
<cr>
Dell(conf)#clock summer-time pacific recurring
Dell(conf)#
```

57

Tunneling

Tunneling supports RFC 2003, RFC 2473, and 4213. DSCP, hop-limits, flow label values, OSPFv2, and OSPFv3 are also supported. ICMP error relay, PATH MTU transmission, and fragmented packets are not supported.

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).

```
Dell(conf)#interface tunnel 1
Dell(conf-if-tu-1)#tunnel source 30.1.1.1
Dell(conf-if-tu-1)#tunnel destination 50.1.1.1
Dell(conf-if-tu-1)#tunnel mode ipip
Dell(conf-if-tu-1)#ip address 1.1.1.1/24
Dell(conf-if-tu-1)#ipv6 address 1::1/64
Dell(conf-if-tu-1)#no shutdown
Dell(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):

```
Dell(conf)#interface tunnel 2
Dell(conf-if-tu-2)#tunnel source 60.1.1.1
Dell(conf-if-tu-2)#tunnel destination 90.1.1.1
Dell(conf-if-tu-2)#tunnel mode ipv6ip
Dell(conf-if-tu-2)#ipv6 address 2::1/64
Dell(conf-if-tu-2)#no shutdown
Dell(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):

```
Dell(conf)#interface tunnel 3
Dell(conf-if-tu-3)#tunnel source 5::5
Dell(conf-if-tu-3)#tunnel destination 8::9
Dell(conf-if-tu-3)#tunnel mode ipv6
Dell(conf-if-tu-3)#ip address 3.1.1.1/24
Dell(conf-if-tu-3)#ipv6 address 3::1/64
Dell(conf-if-tu-3)#no shutdown
Dell(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

Configure the tunnel keepalive target, interval and attempts.

- By default the tunnel keepalive is disabled.

The following sample configuration shows tunnel keepalive command:

```
Dell(conf-if-te-0/12)#show config
!
interface TenGigabitEthernet 0/12
ip address 40.1.1.1/24
ipv6 address 500:10::1/64
no shutdown
Dell(conf-if-te-0/12)#
Dell(conf)#interface tunnel 1
Dell(conf-if-tu-1)#ipv6 address 1abd::1/64
Dell(conf-if-tu-1)#ip address 1.1.1.1/24
Dell(conf-if-tu-1)#tunnel source 40.1.1.1
Dell(conf-if-tu-1)#tunnel destination 40.1.1.2
Dell(conf-if-tu-1)#tunnel mode ipip
Dell(conf-if-tu-1)#no shutdown
Dell(conf-if-tu-1)#tunnel keepalive 1.1.1.2 attempts 4 interval 6
Dell(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 the ip and ipv6 unnumbered

Dell Networking OS supports configuring the tunnel interface.

You can configure the tunnel in ip unnumbered and ipv6 unnumbered command. To configure the tunnel interface to operate without a unique explicit ip/ ipv6 address, select the interface from which the tunnel will borrow its address.

The following sample configuration shows the IP unnumbered command:

```
Dell(conf-if-te-0/0)#show config
!
interface TenGigabitEthernet 0/0
ip address 20.1.1.1/24
ipv6 address 20:1::1/64
no shutdown
Dell(conf)#interface tunnel 1
Dell(conf-if-tu-1)#ip unnumbered tengigabitethernet 0/0
Dell(conf-if-tu-1)#ipv6 unnumbered tengigabitethernet 0/0
Dell(conf-if-tu-1)#tunnel source 40.1.1.1
Dell(conf-if-tu-1)#tunnel mode ipip decapsulate-any
Dell(conf-if-tu-1)#no shutdown
Dell(conf-if-tu-1)#sho c
!
interface Tunnel 1
ip unnumbered TenGigabitEthernet 0/0
ipv6 unnumbered TenGigabitEthernet 0/0
tunnel source 40.1.1.1
tunnel mode ipip decapsulate-any
no shutdown
Dell(conf-if-tu-1)#
```

Configuring the Tunnel allow-remote

You can configure an IPv4 or IPV6 address or prefix whose tunneled packet will be accepted for decapsulation.

.

- If no allow-remote entries are configured, then tunneled packets from any remote peer address will be accepted.
- Upto eight allow-remote entries can be configured on any particular multipoint receive-only tunnel.

The following sample configuration shows tunnel allow-remote command:

```
Dell(conf)#interface tunnel 1
Dell(conf-if-tu-1)#ipv6 address 1abd::1/64
Dell(conf-if-tu-1)#ip address 1.1.1.1/24
Dell(conf-if-tu-1)#tunnel source 40.1.1.1
Dell(conf-if-tu-1)#tunnel mode ipip decapsulate-any
Dell(conf-if-tu-1)#tunnel allow-remote 40.1.1.2
Dell(conf-if-tu-1)#no shutdown
Dell(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 the Tunnel Source Anylocal

You can use the anylocal argument in place of the ip address or interface, but only with multipoint receive-only mode tunnels. The `tunnel source anylocal` command allows the 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 following sample configuration shows the `tunnel source anylocal` command:

```
Dell(conf)#interface tunnel 1
Dell(conf-if-tu-1)#ipv6 address 1abd::1/64
Dell(conf-if-tu-1)#ip address 1.1.1.1/24
Dell(conf-if-tu-1)#tunnel source anylocal
Dell(conf-if-tu-1)#tunnel mode ipip decapsulate-any
Dell(conf-if-tu-1)#tunnel allow-remote 40.1.1.2
Dell(conf-if-tu-1)#no shutdown
Dell(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
```

58

Uplink Failure Detection (UFD)

Supported Modes

Standalone, PMUX, VLT, Stacking

Feature Description

UFD provides detection of the loss of upstream connectivity and, if used with network interface controller (NIC) teaming, automatic recovery from a failed link.

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.

NOTE: In Standalone and VLT modes, the UFD group number is 1 by default and cannot be changed.

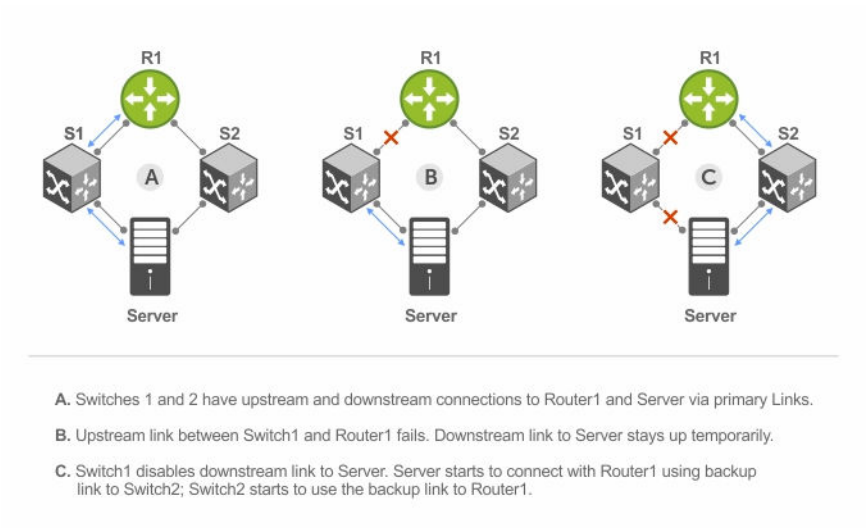


Figure 116. Uplink Failure Detection

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.

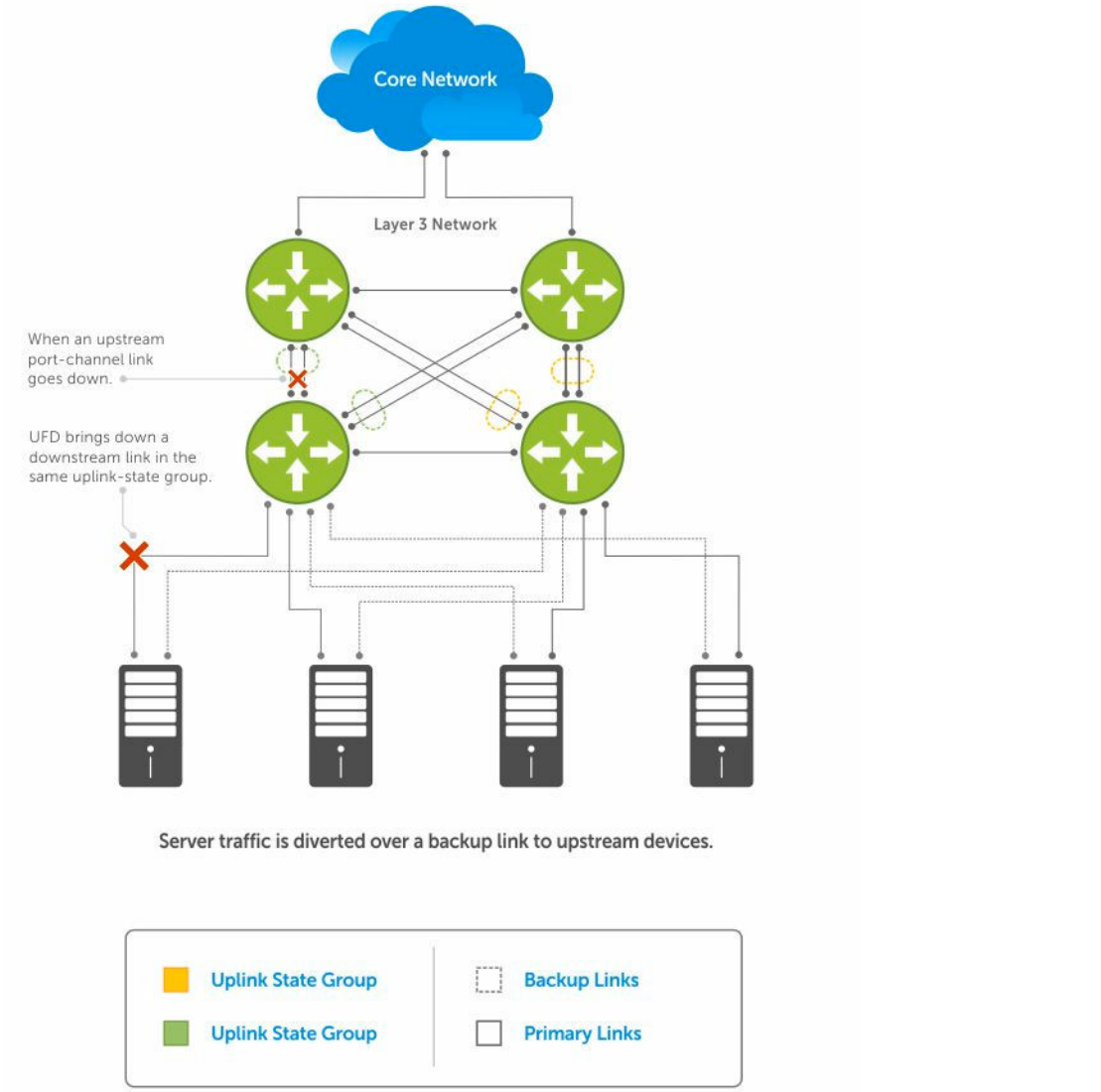


Figure 117. Uplink Failure Detection Example

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 Network Interface Controller (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 to continue to transmit traffic upstream.

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 in PMUX mode and uplink state group 1 is created in Standalone and VLT modes.
  - 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 physical port or port-channel interfaces to an uplink-state group in PMUX mode.
  - 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 .

Uplink Failure Detection (SMUX mode)

In Standalone or VLT modes, by default, all the server-facing ports are tracked by the operational status of the uplink LAG. If the uplink LAG goes down, the aggregator loses its connectivity and is no longer operational. All the server-facing ports are brought down after the specified defer-timer interval, which is 10 seconds by default. If you have configured VLAN, you can reduce the defer time by changing the defer-timer value or remove it by using the `no defer-timer` command.

1. View the Uplink status group.
   EXEC Privilege mode

   ```
   show uplink-state-group
   ```

   ```
   Dell#show uplink-state-group
   Uplink State Group: 1   Status: Enabled, Down
   ```
2. Enable the uplink group tracking.
   UPLINK-STATE-GROUP mode

   ```
   enable
   ```

   ```
   Dell(conf)#uplink-state-group 1
   Dell(conf-uplink-state-group-1)#enable
   ```

   To disable the uplink group tracking, use the `no enable` command.
3. Change the default timer.
   UPLINK-STATE-GROUP mode

   ```
   defer-timer seconds
   ```

   ```
   Dell(conf)#uplink-state-group 1
   Dell(conf-uplink-state-group-1)#defer-timer 20
   Dell(conf-uplink-state-group-1)#show config
   !
   uplink-state-group 1
   downstream TenGigabitEthernet 0/1-12
    upstream Port-channel 128
    defer-timer 20
   ```

Configuring Uplink Failure Detection (PMUX mode)

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:

- TenGigabit Ethernet: enter `tengigabitethernet {slot/port |slot/port-range}`
- Port channel: enter `port-channel {1-128 | 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:

```
upstream tengigabitethernet 0/1-2,5,9,11-12
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. 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}`
   - Port channel: enter `port-channel {1-128 | 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:

   ```
   upstream gigabitethernet 0/1-2,5,9,11-12
   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.

4. (Optional) 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.

   To revert to the default setting, use the `no downstream disable links` command.

5. (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.

6. Specify the time (in seconds) to wait for the upstream port channel (LAG 128) to come back up before server ports are brought down.
   UPLINK-STATE-GROUP mode

   ```
   defer-timer seconds
   ```

   NOTE: This command is available in Standalone and VLT modes only.

   The range is from 1 to 120.
7. (Optional) Enter a text description of the uplink-state group.
   UPLINK-STATE-GROUP mode

   ```
   description text
   ```

   The maximum length is 80 alphanumeric characters.
8. (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 (in PMUX mode)

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:
  - 10 Gigabit Ethernet: enter `tengigabitethernet {slot/port | slot/port-range}`
  - Port channel: enter `port-channel {1-128 | 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 0/1-2,5,9,11-12
      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.

Example of Syslog Messages Before and After Entering the `clear ufd-disable uplink-state-group` Command

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.

```
00:10:12: %STKUNIT0-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to
down: Te 0/1
00:10:12: %STKUNIT0-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to
down: Te 0/2
00:10:12: %STKUNIT0-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to
down: Te 0/3
00:10:12: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down:
Te 0/1
00:10:12: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down:
Te 0/2
00:10:12: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down:
Te 0/3
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed uplink state group state
to down: Group
3
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD
error-disabled:
Te 0/4
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD
error-disabled:
Te 0/5
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD
error-disabled:
Te 0/6
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down:
Te 0/4
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down:
Te 0/5
00:10:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down:
Te 0/6

Dell(conf-if-range-te-0/1-3)#do clear ufd-disable uplink-state-group 3

00:11:50: %STKUNIT0-M:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from
UFD
error-disabled: Te 0/4
00:11:51: %STKUNIT0-M:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from
UFD
error-disabled: Te 0/5
00:11:51: %STKUNIT0-M:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from
UFD
error-disabled: Te 0/6
00:11:51: %STKUNIT0-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te
0/4
00:11:51: %STKUNIT0-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te
0/5
00:11:51: %STKUNIT0-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te
0/6
```

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 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:
  - 10 Gigabit Ethernet: enter `tengigabitethernet slot/port`.
  - Port channel: enter `port-channel {1-128}`.

  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.

Example of Viewing Uplink State Group Status

```
Dell# 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

Dell# show uplink-state-group 16
Uplink State Group: 16 Status: Disabled, Up

Dell#show uplink-state-group detail

(Up): Interface up   (Dwn): Interface down   (Dis): Interface disabled

Uplink State Group    : 1       Status: Enabled, Up
Defer Timer           : 10 sec
Upstream Interfaces   : Po 128(Up)
Downstream Interfaces : Te 0/1(Dwn) Te 0/2(Dwn) Te 0/3(Up) Te 0/4(Dwn) Te
0/5(Up)
                        Te 0/6(Dwn) Te 0/7(Up) Te 0/8(Up)
```

```
Dell#
```

Example of Viewing Interface Status with UFD Information

```
Dell#show interfaces tengigabitethernet 0/7
TenGigabitEthernet 0/7 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
```

Examples of Viewing UFD Output

```
Dell#show running-config uplink-state-group
!
uplink-state-group 1
no enable
downstream TenGigabitEthernet 0/3
upstream TenGigabitEthernet 0/1
Dell#
```

```
Dell(conf-uplink-state-group-16)# show configuration
!
uplink-state-group 16
no enable
description test
downstream disable links all
downstream TengigabitEthernet 0/4
upstream TengigabitEthernet 0/5
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 Gigabitethernet 0/1, 0/2, 0/5, 0/9, 0/11, and 0/12.

- Configure two downstream links to be disabled if an upstream link fails.
- Add upstream links Gigabitethernet 0/3 and 0/4.
- Add a text description for the group.
- Verify the configuration with various `show` commands.

Example of Configuring UFD

```
Dell(conf)#uplink-state-group 3
Dell(conf-uplink-state-group-3)#

00:23:52: %STKUNIT0-M:CP %IFMGR-5-ASTATE_UP: Changed uplink state group Admin
state to up:
Group 3

Dell(conf-uplink-state-group-3)#downstream tengigabitethernet 0/1-2,5,9,11-12
Dell(conf-uplink-state-group-3)#downstream disable links 2
Dell(conf-uplink-state-group-3)#upstream tengigabitethernet 0/3-4
Dell(conf-uplink-state-group-3)#description Testing UFD feature
Dell(conf-uplink-state-group-3)#show config
!
uplink-state-group 3
  description Testing UFD feature
  downstream disable links 2
  downstream TenGigabitEthernet 0/1-2,5,9,11-12
  upstream TenGigabitEthernet 0/3-4

Dell#show running-config uplink-state-group
!
uplink-state-group 3
  description Testing UFD feature
  downstream disable links 2
  downstream TenGigabitEthernet 0/1-2,5,9,11-12
  upstream TenGigabitEthernet 0/3-4

Dell#show uplink-state-group 3

Uplink State Group: 3 Status: Enabled, Up

Dell#show uplink-state-group detail

(Up): Interface up (Dwn): Interface down (Dis): Interface disabled

Uplink State Group    : 3 Status: Enabled, Up
Upstream Interfaces   : Te 0/3(Up) Te 0/4(Up)
Downstream Interfaces : Te 0/1(Up) Te 0/2(Up) Te 0/5(Up) Te 0/9(Up) Te 0/11(Up)
                        Te 0/12(Up)

< After a single uplink port fails >

Dell#show uplink-state-group detail

(Up): Interface up (Dwn): Interface down (Dis): Interface disabled

Uplink State Group    : 3 Status: Enabled, Up
Upstream Interfaces   : Te 0/3(Dwn) Te 0/4(Up)
Downstream Interfaces : Te 0/1(Dis) Te 0/2(Dis) Te 0/5(Up) Te 0/9(Up) Te
0/11(Up)
                        Te 0/12(Up)
```

59

PMUX Mode of the IO Aggregator

This chapter provides an overview of the PMUX mode.

I/O Aggregator (IOA) Programmable MUX (PMUX) Mode

IOA PMUX is a mode that provides flexibility of operation with added configurability. This involves creating multiple LAGs, configuring VLANs on uplinks and the server side, configuring data center bridging (DCB) parameters, and so forth.

By default, IOA starts up in IOA Standalone mode. You can change to PMUX mode by executing the following commands and then reloading the IOA. After the IOA reboots, the IOA operates in PMUX mode. PMUX mode supports both stacking and VLT operations.

Configuring and Changing to PMUX Mode

After the IOA is operational in the default Standalone mode:

1. Connect the terminal to the console port on the IOA to access the CLI and enter the following commands:

```
Login: username
Password: *****
Dell> enable
Dell#
Dell#show system stack-unit 0 iom-mode
Unit Boot-Mode Next-Boot
------------------------------------------------
0 standalone standalone
Dell#
```

2. Change IOA mode to PMUX mode.

```
Dell(conf)# stack-unit 0 iom-mode programmable-mux
```

   Where `stack-unit 0` defines the default stack-unit number.

3. Delete the startup configuration file.

```
Dell# delete startup-config
```

4. Reboot the IOA by entering the `reload` command.

```
Dell# reload
```

5. Repeat the above steps for each member of the IOA in PMUX mode.

   After system is up, you can see the PMUX mode status:

```
Dell#sh system stack-unit 0 iom-mode
Unit          Boot-Mode                  Next-Boot
```

```
-------------------------------------------------------
0           programmable-mux           programmable-mux
Dell#
```

The IOA is now ready for PMUX operations.

Configuring the Commands without a Separate User Account

Starting with Dell Networking OS version 9.3(0.0), you can configure the PMUX mode CLI commands without having to configure a new, separate user profile. The user profile you defined to access and log in to the switch is sufficient to configure the PMUX mode commands.

The IOA PMUX Mode CLI Commands section lists the PMUX mode CLI commands that you can now configure without a separate user account.

Virtual Link Trunking (VLT)

VLT allows physical links between two chassis to appear as a single virtual link to the network core. VLT eliminates the requirement for Spanning Tree protocols by allowing link aggregation group (LAG) terminations on two separate distribution or core switches, and by supporting a loop-free topology. VLT provides Layer 2 multipathing, creating redundancy through increased bandwidth and enabling multiple parallel paths between nodes and load-balancing traffic where alternative paths exist.

NOTE: When you launch the VLT link, the VLT peer-ship is not established if any of the following is **TRUE**:

- The VLT System-MAC configured on both the VLT peers do not match.
- The VLT Unit-Id configured on both the VLT peers are identical.
- The VLT System-MAC or Unit-Id is configured only on one of the VLT peers.
- The VLT domain ID is not the same on both peers.

If the VLT peer-ship is already established, changing the System-MAC or Unit-Id does not cause VLT peer-ship to go down.

Also, if the VLT peer-ship is already established and the VLT Unit-Id or System-MAC are configured on both peers, then changing the CLI configurations on the VLT Unit-Id or System-MAC is rejected if any of the following become **TRUE**:

- After making the CLI configuration change, the VLT Unit-Id becomes identical on both peers.
- After making the CLI configuration change, the VLT System-MAC do not match on both peers.

When the VLT peer-ship is already established, you can remove the VLT Unit-Id or System-MAC configuration from either or both peers. However, removing configuration settings can cause the VLT ports to go down if you configure the Unit-Id or System-MAC on only one of the VLT peers.

Overview

VLT allows physical links between two chassis to appear as a single virtual link to the network core or other switches such as Edge, Access, or top-of-rack (ToR).

VLT reduces the role of spanning tree protocols (STPs) by allowing link aggregation group (LAG) terminations on two separate distribution or core switches, and by supporting a loop-free topology. (To prevent the initial loop that may occur prior to VLT being established, use a spanning tree protocol.

VLT provides Layer 2 multipathing, creating redundancy through increased bandwidth, enabling multiple parallel paths between nodes and load-balancing traffic where alternative paths exist.

Virtual link trunking 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.

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.

Setting up VLT

The following figure shows the sample VLT topology.

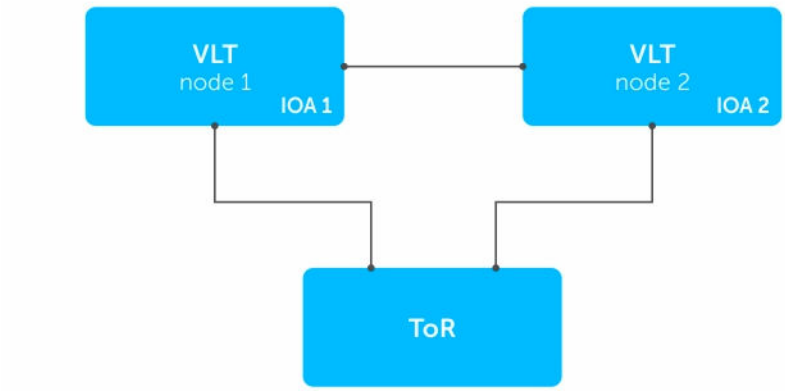


Figure 118. Sample VLT Topology

In PMUX VLT, you can choose any uplink ports for configuring VLT.

NOTE: Ensure the connectivity to ToR from each Aggregator.

To enable VLT and verify the configuration, follow these steps.

1. Enable VLT in node 1 and 2.

   ```
   stack-unit unit iom-mode vlt
   ```

   CONFIGURATION mode

   ```
   Dell(conf)#stack-unit 0 iom-mode vlt
   ```

2. Verify the VLT configurations.

   ```
   show interface port-channel brief
   ```

   EXEC mode

   ```
   Dell# show interfaces port brief
   Codes: L - LACP Port-channel
   O - OpenFlow Controller Port-channel

   LAG      Mode  Status       Uptime      Ports
       127  L2    up           00:18:22    Fo 0/33    (Up)<<<<<<<<ICL LAG
                                           Fo 0/37    (Up)
   L   128  L2    up           00:00:00    Fo 0/41    (Up)<<<<<<<<Uplink LAG
   ```

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 vitality of 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.
- **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).

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.

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

- VLT port channel interfaces must be switch ports.

- Dell Networking strongly recommends that the VLTi (VLT interconnect) be a static LAG and that you disable LACP on the VLTi.
- 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.

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 chassis in the VLT domain is assigned a primary role; the other chassis 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.
  - In a VLT domain, the peer switches must run the same Dell 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 supports a maximum of two 10G ports.
  - In VLT mode, port 33 and 37 are dedicated to VLT interconnect ports.
  - A VLT interconnect over 1G 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 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.

- 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 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 .
  - 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 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.

- Software features supported on VLT port-channels
  - For information about configuring IGMP Snooping in a VLT domain, refer to 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 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.
- 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, port monitoring, and jumbo frames.
- Software features not supported with VLT
  - In a VLT domain, the following software features are supported on non-VLT ports: 802.1x, , DHCP snooping, FRRP, IPv6 dynamic routing, ingress and egress QOS.
- 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

Primary and Secondary VLT Peers are supported on the Aggregator.

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.

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.

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 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 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.

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.

Non-VLT ARP Sync

In the Dell Networking OS version 9.2(0.0), ARP entries (including ND entries) learned on other ports are synced with the VLT peer to support station move scenarios.

Prior to Dell Networking OS version 9.2.(0.0), only ARP entries learned on VLT ports were synced between peers.

Additionally, ARP entries resulting from station movements from VLT to non-VLT ports or to different non-VLT ports are learned on the non-VLT port and synced with the peer node. The peer node is updated to use the new non-VLT port.

NOTE: ARP entries learned on non-VLT, non-spanned VLANs are not synced with VLT peers.

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 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:
    * 10-Gigabit Ethernet: enter `tengigabitethernet slot/port`.
    * Port channel: enter `port-channel {1-128}`.

Example of 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
```

Example of the `show vlt brief` Command

```
Dell_VLTpeer1# show vlt brief
VLT Domain Brief
------------------
  Domain ID:                    1000
  Role:                         Secondary
  Role Priority:                32768
  ICL Link Status:              Up
  HeartBeat Status:             Up
  VLT Peer Status:              Up
  Local Unit Id:                0
  Version:                      5(1)
  Local System MAC address:     00:01:e8:8a:e9:70
  Remote System MAC address:    00:01:e8:8a:e7:e7
  Configured System MAC address: 00:0a:0a:01:01:0a
  Remote system version:        5(1)
  Delay-Restore timer:          90 seconds

Dell_VLTpeer2# show vlt brief
VLT Domain Brief
------------------
Domain ID:                      1000
  Role:                         Primary
  Role Priority:                32768
  ICL Link Status:              Up
  HeartBeat Status:             Up
  VLT Peer Status:              Up
  Local Unit Id:                1
  Version:                      5(1)
  Local System MAC address:     00:01:e8:8a:e7:e7
  Remote System MAC address:    00:01:e8:8a:e9:70
  Configured System MAC address: 00:0a:0a:01:01:0a
  Remote system version:        5(1)
  Delay-Restore timer:          90 seconds
```

Example of 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
```

Example of 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
```

Example of 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
```

Example of 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
```

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)

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)#channel-member TenGigE 0/6,7
Dell_VLTpeer1(conf-if-po-100)#no shutdown
Dell_VLTpeer1(conf-if-po-100)#exit
```

Enable VLT and create a VLT domain with a backup-link VLT 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 port channel to an attached device.

```
Dell_VLTpeer1(conf)#interface port-channel 110
Dell_VLTpeer1(conf-if-po-110)#switchport
Dell_VLTpeer1(conf-if-po-110)#channel-member TenGigE 0/5
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(Te 0/5)
                          T Po100(Te 0/6,7)
```

Configuring Virtual Link Trunking (VLT Peer 2)

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)#channel-member TenGigE 0/3,4
Dell_VLTpeer2(conf-if-po-100)#no shutdown
Dell_VLTpeer2(conf-if-po-100)#exit
```

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 port channel to an attached device.

```
Dell_VLTpeer2(conf)#interface port-channel 110
Dell_VLTpeer2(conf-if-po-110)#switchport
Dell_VLTpeer2(conf-if-po-110)#channel-member TenGigE 0/8
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(Te 0/8)
                          T Po100(Te 0/3,4)
```

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
switchport
channel-member TenGigE 0/1,2
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 Networking representative.

Table 65. 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.
Dell 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 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.
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’.

Description	Behavior at Peer Up	Behavior During Run Time	Action to Take
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 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.

60

NPIV Proxy Gateway

The N-port identifier virtualization (NPIV) Proxy Gateway (NPG) feature provides FCoE-FC bridging capability on the FN 2210S Aggregator, allowing server CNAs to communicate with SAN fabrics over the FN 2210S Aggregator.

NPIV Proxy Gateway Configuration

The Aggregator switches function as a top-of-rack edge switch that supports Converged Enhanced Ethernet (CEE) traffic — Fibre Channel over Ethernet (FCoE) for storage, Interprocess Communication (IPC) for servers, and Ethernet local area network (LAN) (IP cloud) for data — as well as FC links to one or more storage area network (SAN) fabrics.

The NPG provides FCoE-FC bridging capability on the Aggregator.

This chapter describes how to configure and use an NPIV proxy gateway on the Aggregator in a SAN.

NPIV Proxy Gateway Operations and Capabilities

Benefits of an NPIV Proxy Gateway

The Aggregator functions as a top-of-rack edge switch that supports CEE traffic — FCoE for storage, IPC for servers, and Ethernet LAN (IP cloud) for data — as well as FC links to one or more SAN fabrics.

Using an NPG helps resolve the following problems in a storage area network:

- Fibre Channel storage networks typically consist of servers connected to edge switches, which are connected to SAN core switches. As the SAN grows, it is necessary to add more ports and SAN switches. This results in an increase in the required domain IDs, which may surpass the upper limit of 239 domain IDs supported in the SAN network. An NPG avoids the need for additional domain IDs because it is deployed outside the SAN and uses the domain IDs of core switches in its FCoE links.
- With the introduction of 10GbE links, FCoE is being implemented for server connections to optimize performance. However, a SAN traditionally uses Fibre Channel to transmit storage traffic. FCoE servers require an efficient and scalable bridging feature to access FC storage arrays, which an NPG provides.

NPIV Proxy Gateway Operation

Consider a sample scenario of NPG operation. An FX2 server chassis configured as an NPG does not join a SAN fabric, but functions as an FCoE-FC bridge that forwards storage traffic between servers and core SAN switches. The core switches forward SAN traffic to and from FC storage arrays.

An FX2 chassis FC port is configured as an N (node) port that logs in to an F (fabric) port on the upstream FC core switch and creates a channel for N-port identifier virtualization. NPIV allows multiple N-port fabric logins at the same time on a single, physical Fibre Channel link.

Converged Network Adapter (CNA) ports on servers connect to the FX2 chassis Ten-Gigabit Ethernet ports and log in to an upstream FC core switch through the N port. Server fabric login (FLOGI) requests are converted into fabric discovery (FDISC) requests before being forwarded to the FC core switch.

Servers use CNA ports to connect over FCoE to an Ethernet port in ENode mode on the NPIV proxy gateway. FCoE transit with FIP snooping is automatically enabled and configured on the FX2 gateway to prevent unauthorized access and data transmission to the SAN network. FIP is used by server CNAs to discover an FCoE switch operating as an FCoE forwarder (FCF).

The NPIV proxy gateway aggregates multiple locally connected server CNA ports into one or more upstream N port links, conserving the number of ports required on an upstream FC core switch while providing an FCoE-to-FC bridging functionality. The upstream N ports on an FX2 can connect to the same or multiple fabrics.

Using an FCoE map applied to downstream (server-facing) Ethernet ports and upstream (fabric-facing) FC ports, you can configure the association between a SAN fabric and the FCoE VLAN that connects servers over the NPIV proxy gateway to FC switches in the fabric. An FCoE map virtualizes the upstream SAN fabric as an FCF to downstream CNA ports on FCoE-enabled servers as follows:

- As soon as an FC N port comes online (`no shutdown` command), the NPG starts sending FIP multicast advertisements, which contain the fabric name derived from the 64-bit worldwide name (WWN) of the principal SAN switch. (The principal switch in a fabric is the FC switch with the lowest domain ID.)
- When you apply the FCoE map to a server-facing Ethernet port in ENode mode, ACLs are automatically configured to allow only FCoE traffic from servers that perform a successful FLOGI on the FC switch. All other traffic on the VLAN is denied.

You can specify one or more upstream N ports in an FCoE map. The FCoE map also contains the VLAN ID of the dedicated VLAN used to transmit FCoE traffic between the SAN fabric and servers.

NPIV Proxy Gateway: Protocol Services

The Aggregator with the NPG provides the following protocol services:

- Fibre Channel service to create N ports and log in to an upstream FC switch.
- FCoE service to perform:
  - Virtualization of FC N ports on an NPG so that they appear as FCoE FCFs to downstream servers.
  - NPIV service to perform the association and aggregation of FCoE servers to upstream F ports on core switches (through N ports on the NPG). Conversion of server FLOGIs and FDISCs, which are received over the Aggregator with the ENode ports, are converted into FDISCs addressed to the upstream F ports on core switches.

NPIV Proxy Gateway Functionality

The Aggregator with the NPG provides the following functionality in a storage area network:

- FIP Snooping bridge that provides security for FCoE traffic using ACLs.
- FCoE gateway that provides FCoE-to-FC bridging. N-port virtualization using FCoE maps exposes upstream F ports as FCF ports to downstream server-facing ENode ports on the NPG.

NPIV Proxy Gateway: Terms and Definitions

The following table describes the terms used in an NPG configuration on the Aggregator.

Table 66. Aggregator with the NPIV Proxy Gateway: Terms and Definitions

Term	Description
FC port	Fibre Channel port on the Aggregator that operates in autosensing, 2, 4, or 8-Gigabit mode. On an NPIV proxy gateway, an FC port can be used as a downlink for a server connection and an uplink for a fabric connection.
F port	Port mode of an FC port connected to an end node (N) port on an Aggregator with the NPIV proxy gateway.
N port	Port mode of an Aggregator with the FC port that connects to an F port on an FC switch in a SAN fabric. On an Aggregator with the NPIV proxy gateway, an N port also functions as a proxy for multiple server CNA-port connections.
ENode port	Port mode of a server-facing Aggregator with the Ethernet port that provides access to FCF functionality on a fabric.
CNA port	N-port functionality on an FCoE-enabled server port. A converged network adapter (CNA) can use one or more Ethernet ports. CNAs can encapsulate Fibre Channel frames in Ethernet for FCoE transport and de-encapsulate Fibre Channel frames from FCoE to native Fibre Channel.
DCB map	Template used to configure DCB parameters, including priority-based flow control (PFC) and enhanced transmission selection (ETS), on CEE ports.
Fibre Channel fabric	Network of Fibre Channel devices and storage arrays that inter-operate and communicate.
FCF	Fibre Channel forwarder: FCoE-enabled switch that can forward FC traffic to both downstream FCoE and upstream FC devices. An NPIV proxy gateway functions as an FCF to export upstream F port configurations to downstream server CNA ports.
FC-MAP	FCoE MAC-address prefix — The unique 24-bit MAC address prefix in FCoE packets used to generate a fabric-provided MAC address (FPMA). The FPMA is required to send FCoE packets from a server to a SAN fabric.
FCoE map	Template used to configure FCoE and FC parameters on Ethernet and FC ports in a converged fabric.
FCoE VLAN	VLAN dedicated to carrying only FCoE traffic between server CNA ports and a SAN fabric. (FCoE traffic must travel in a VLAN.) When you apply an FCoE map on a port, FCoE is enabled on the port. All non-FCoE traffic is dropped on an FCoE VLAN.
FIP	FCoE Initialization Protocol: Layer 2 protocol for endpoint discovery, fabric login, and fabric association. FIP is used by server CNAs to discover

Term	Description
	an upstream FCoE switch operating as an FCF. FIP keepalive messages maintain the connection between an FCoE initiator and an FCF.
NPIV	N-port identifier virtualization: The capability to map multiple FCoE links from downstream ports to a single upstream FC link.
principal switch	The switch in a fabric with the lowest domain number. The principal switch accesses the master name database and the zone/zone set database.

DCB Maps

A Data Center Bridging (DCB) map is used to configure DCB functionality, such as PFC and ETS, on the Aggregator with the Ethernet ports that support CEE traffic and are DCBx-enabled, by default.

By default, no PFC and ETS settings in a DCB map are applied to the Aggregator with the Ethernet ports when they are enabled. On an Aggregator with the NPG, you must configure PFC and ETS parameters in a DCB map and then apply the map to server-facing Ethernet ports.

FCoE Maps

An FCoE map is used to identify the SAN fabric to which FCoE storage traffic is sent. Using an FCoE map, an Aggregator with the NPG operates as an FCoE-FC bridge between an FC SAN and FCoE network by providing FCoE-enabled servers and switches with the necessary parameters to log in to a SAN fabric.

An FCoE map applies the following parameters on server-facing Ethernet and fabric-facing FC ports on the Aggregator:

- The dedicated FCoE VLAN used to transport FCoE storage traffic.
- The FC-MAP value used to generate a fabric-provided MAC address.
- The association between the FCoE VLAN ID and FC fabric ID where the desired storage arrays are installed. Each Fibre Channel fabric serves as an isolated SAN topology within the same physical network.
- The priority used by a server to select an upstream FCoE forwarder (FCF priority).
- FIP keepalive (FKA) advertisement timeout.

NOTE:
In each FCoE map, the fabric ID, FC-MAP value, and FCoE VLAN must be unique. Use one FCoE map to access one SAN fabric. You cannot use the same FCoE map to access different fabrics.

When you configure an Aggregator with the NPG, FCoE transit with FIP snooping is automatically enabled and configured using the parameters in the FCoE map applied to server-facing Ethernet and fabric-facing FC interfaces.

After you apply an FCoE map on an FC port, when you enable the port (`no shutdown`), the NPG starts sending FIP multicast advertisements on behalf of the FC port to downstream servers in order to advertise the availability of a new FCF port on the FCoE VLAN. The FIP advertisement also contains a keepalive message to maintain connectivity between a SAN fabric and downstream servers.

Configuring an NPIV Proxy Gateway

Prerequisite: Before you configure an NPIV proxy gateway (NPG) on an Aggregator, ensure that the following features are enabled.

- DCB is enabled by default on the Aggregator.
- Autonegotiated DCBx is enabled for converged traffic by default with the Ethernet ports on all Aggregators.
- FCoE transit with FIP snooping is automatically enabled when you configure Fibre Channel on the Aggregator.

To configure an NPG operation on an Aggregator, follow these general configuration steps:

1. Enabling Fibre Channel Capability on the Switch
2. Creating a DCB map
3. Applying a DCB map on server-facing Ethernet ports
4. Creating an FCoE VLAN
5. Creating an FCoE map
6. Applying an FCoE map on server-facing Ethernet ports
7. Applying an FCoE Map on fabric-facing FC ports

NOTE: All these configurations are available only in PMUX mode and you cannot perform these configurations in Standalone mode.

Default Configurations in Standalone mode

By default, the following configurations are set in Standalone mode:

1. All the FC port are applied with the default FCoE map.
2. All the 10G server facing ports with the DCBX frames will have the default FCoE map and default DCB map assigned, and without the DCBX frames, DCB_MAP_PFC_OFF will be applied.

Default DCB map

```
Dell(conf)#do show qos dcb-map SAN_DCB_MAP
-----------------------

State   :Complete
PfcMode:ON
--------------------
PG:0 TSA:ETS  BW:30  PFC:OFF
Priorities:0 1 2 5 6 7


PG:1 TSA:ETS  BW:30  PFC:OFF
Priorities:4


PG:2 TSA:ETS  BW:40  PFC:ON
Priorities:3
```

Default FCoE map

```
Dell(conf)#do show fcoe-map
```

```
Fabric Name         SAN_FABRIC
Fabric Id           1002
Vlan Id             1002
Vlan priority       3
FC-MAP              0efc00
FKA-ADV-Period      8
Fcf Priority        128
Config-State        ACTIVE
Oper-State          UP
Members
Fc 0/9
Te 0/4
```

DCB_MAP_PFC_OFF

```
Dell(conf)#do show qos dcb-map DCB_MAP_PFC_OFF
----------------------

State   :In-Progress
PfcMode:OFF
--------------------
Dell(conf)#
```

Enabling Fibre Channel Capability on the Switch

Enable the Fibre Channel capability on an Aggregator that you want to configure as an NPG for the Fibre Channel protocol. When you enable Fibre Channel capability, FCoE transit with FIP snooping is automatically enabled on all VLANs on the switch, using the default FCoE transit settings.

1. Enable the Fibre Channel capability on an Aggregator for the Fibre Channel protocol.

   CONFIGURATION mode

   ```
   feature fc
   ```

Creating a DCB Map

Configure the priority-based flow control (PFC) and enhanced traffic selection (ETS) settings in a DCB map before you apply them on downstream server-facing ports on an Aggregator.

1. Create a DCB map to specify PFC and ETS settings for groups of dot1p priorities.

   CONFIGURATION mode

   ```
   dcb-map name
   ```

2. Configure the PFC setting (on or off) and the ETS bandwidth percentage allocated to traffic in each priority group. Configure whether the priority group traffic should be handled with strict-priority scheduling. The sum of all allocated bandwidth percentages must be 100 percent. Strict-priority traffic is serviced first. Afterward, 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.

   Restriction: You can enable PFC on a maximum of two priority queues.

Repeat this step to configure PFC and ETS traffic handling for each priority group, for example: `priority-group 0 bandwidth 60 pfc off priority-group 1 bandwidth 20 pfc onpriority-group 2 bandwidth 20 pfc on priority-group 4 strict-priority pfc off`

DCB MAP mode

`priority-group` *group_num* {**bandwidth** *percentage* | **strict-priority**} **pfc** {**on** | **off**}

3. Specify the priority group ID number to handle VLAN traffic for each dot1p class-of-service: 0 through 7. Leave a space between each priority group number. For example, `priority-pgid 0 0 0 1 2 4 4 4` where dot1p priorities 0, 1, and 2 are mapped to priority group 0; dot1p priority 3 is mapped to priority group 1; dot1p priority 4 is mapped to priority group 2; dot1p priorities 5, 6, and 7 are mapped to priority group 4.

   All priorities that map to the same egress queue must be in the same priority group.

   DCB MAP mode

   `priority-pgid` *dot1p0_group_num dot1p1_group_num dot1p2_group_num dot1p3_group_num dot1p4_group_num dot1p5_group_num dot1p6_group_num dot1p7_group_num*

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 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, 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-to-priority group mapping.
- If you delete the dot1p priority-to-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 the network operation.

Applying a DCB Map on Server-facing Ethernet Ports

You can apply a DCB map only on a physical Ethernet interface and can apply only one DCB map per interface.

1. Enter CONFIGURATION mode on a server-facing port or port channel to apply a DCB map.

   You cannot apply a DCB map on a port channel. However, you can apply a DCB map on the ports that are members of the port channel.

   CONFIGURATION mode

   `interface` {**tengigabitEthernet** *slot/port* }

2. Apply the DCB map on an Ethernet port or port channel. The port is configured with the PFC and ETS settings in the DCB map.

Repeat this step to apply a DCB map to more than one port or port channel.

INTERFACE mode

```
dcb-map name
```

```
Dell# interface tengigabitEthernet 0/0
Dell(config-if-te-0/0)# dcb-map SAN_DCB1
```

Creating an FCoE VLAN

Create a dedicated VLAN to send and receive Fibre Channel traffic over FCoE links between servers and a fabric over an NPG. The NPG receives FCoE traffic and forwards decapsulated FC frames over FC links to SAN switches in a specified fabric.

1. Create the dedicated VLAN for FCoE traffic. Range: 2–4094.

   VLAN 1002 is commonly used to transmit FCoE traffic.

   CONFIGURATION mode

   ```
   interface vlan vlan-id
   ```

When you apply an FCoE map to an Ethernet port , the port is automatically configured as a tagged member of the FCoE VLAN.

Creating an FCoE Map

An FCoE map consists of:

- An association between the dedicated VLAN, used to carry FCoE traffic, and the SAN fabric where the storage arrays are installed. Use a separate FCoE VLAN for each fabric to which the FCoE traffic is forwarded. Any non-FCoE traffic sent on a dedicated FCoE VLAN is dropped.
- The FC-MAP value, used to generate the fabric-provided MAC address (FPMA). The FPMA is used by servers to transmit FCoE traffic to the fabric. You can associate an FC-MAP with only one FCoE VLAN and conversely, associate an FCoE VLAN with only one FC-MAP.
- FCF priority, the priority used by a server CNA to select an upstream FCoE forwarder (FCF).
- FIP keepalive (FKA) advertisement timeout.

The values for the FCoE VLAN, fabric ID and FC-MAP must be unique. Apply an FCoE map on downstream server-facing Ethernet ports and upstream fabric-facing Fibre Channel ports.

1. Create an FCoE map that contains parameters used in the communication between servers and a SAN fabric.

   CONFIGURATION mode

   ```
   fcoe-map map-name
   ```

2. Configure the association between the dedicated VLAN and the fabric where the desired storage arrays are installed. The fabric and VLAN ID numbers must be the same. Fabric and VLAN ID range: 2–4094.

   For example: `fabric id 10 vlan 10`

   FCoE MAP mode

```
fabric-id fabric-num vlan vlan-id
```

3. Add a text description of the settings in the FCoE map. Maximum: 32 characters.

   FCoE MAP mode

   ```
   description text
   ```

4. Specify the FC-MAP value used to generate a fabric-provided MAC address, which is required to send FCoE traffic from a server on the FCoE VLAN to the FC fabric specified in Step 2. Enter a unique MAC address prefix as the FC-MAP value for each fabric. Range: 0EFC00–0EFCFF. Default: None.

   FCoE MAP mode

   ```
   fc-map fc-map-value
   ```

5. Configure the priority used by a server CNA to select the FCF for a fabric login (FLOGI). Range: 1–255. Default: 128.

   FCoE MAP mode

   ```
   fcf-priority priority
   ```

6. Enable the monitoring FIP keepalive messages (if it is disabled) to detect if other FCoE devices are reachable. Default: FIP keepalive monitoring is enabled.

   FCoE MAP mode

   ```
   keepalive
   ```

7. Configure the time interval (in seconds) used to transmit FIP keepalive advertisements. Range: 8–90 seconds. Default: 8 seconds.

   FCoE MAP mode

   ```
   fka-adv-period seconds
   ```

Applying an FCoE Map on Server-facing Ethernet Ports

You can apply multiple FCoE maps on an Ethernet port or port channel. When you apply an FCoE map on a server-facing port or port channel:

- The port is configured to operate in hybrid mode (accept both tagged and untagged VLAN frames).
- The associated FCoE VLAN is enabled on the port or port channel.

When you enable a server-facing Ethernet port, the servers respond to the FIP advertisements by performing FLOGIs on upstream virtualized FCF ports. The NPG forwards the FLOGIs as FDISC messages to a SAN switch.

1. Configure a server-facing Ethernet port or port channel with an FCoE map.

   CONFIGURATION mode

   ```
   interface {tengigabitEthernet slot/port | port-channel num}
   ```

2. Apply the FCoE/FC configuration in an FCoE map on the Ethernet port. Repeat this step to apply an FCoE map to more than one port.

INTERFACE or INTERFACE PORT_CHANNEL mode

```
fcoe-map map-name
```

```
Dell# interface tengigabitEthernet 0/0
Dell(config-if-te-0/0)# fcoe-map SAN_FABRIC_A
Dell# interface port-channel 3
Dell(config-if-te-0/0)# dcb-map SAN_DCB1
Dell(config-if-po-3)# fcoe-map SAN_FABRIC_A
```

3. Enable the port for FCoE transmission using the map settings.

   INTERFACE mode

   ```
   no shutdown
   ```

Applying an FCoE Map on Fabric-facing FC Ports

The Aggregator, with the FC ports, are configured by default to operate in N port mode to connect to an F port on an FC switch in a fabric. You can apply only one FCoE map on an FC port.

When you apply an FCoE map on a fabric-facing FC port, the FC port becomes part of the FCoE fabric, whose settings in the FCoE map are configured on the port and exported to downstream server CNA ports.

Each Aggregator, with the FC port, is associated with an Ethernet MAC address (FCF MAC address). When you enable a fabric-facing FC port, the FCoE map applied to the port starts sending FIP multicast advertisements using the parameters in the FCoE map over server-facing Ethernet ports. A server sees the FC port, with its applied FCoE map, as an FCF port.

1. Configure a fabric-facing FC port.

   CONFIGURATION mode

   ```
   interface fibrechannel slot/port
   ```

2. Apply the FCoE and FC fabric configurations in an FCoE map to the port. Repeat this step to apply an FCoE map to more than one FC port.

   INTERFACE FIBRE_CHANNEL mode

   ```
   fabric map-name
   ```

   ```
   Dell# interface fi 0/9
   Dell(config-if-fc-0/9)# fabric SAN_FABRIC_A
   ```

3. Enable the port for FC transmission.

   INTERFACE FIBRE_CHANNEL mode

   ```
   no shutdown
   ```

Important Points to Remember

You can apply a DCB or FCoE map to a range of Ethernet or Fibre Channel interfaces by using the `interface range` command; for example:

```
Dell(config)# interface range tengigabitEthernet 0/1 - 4

Dell(config)# interface range fibrechannel 0/9 - 10
```

Enter the keywords `interface range` followed by an interface type and port range. A port range must contain spaces before and after the dash. Separate each interface type and port range with a space, comma, and space as shown in the preceding examples.

Sample Configuration

1. Configure a DCB map with PFC and ETS settings:

   ```
   Dell(config)# dcb-map SAN_DCB_MAP

   Dell(config-dcbx-name)# priority-group 0 bandwidth 60 pfc off

   Dell(config-dcbx-name)# priority-group 1 bandwidth 20 pfc on

   Dell(config-dcbx-name)# priority-group 2 bandwidth 20 pfc on

   Dell(config-dcbx-name)# priority-group 4 strict-priority pfc off

   Dell(conf-dcbx-name)# priority-pgid 0 0 0 1 2 4 4 4
   ```

2. Apply the DCB map on a downstream (server-facing) Ethernet port:

   ```
   Dell(config)# interface tengigabitethernet 0/1

   Dell(config-if-te-0/0)#dcb-map SAN_DCB_MAP
   ```

3. Create the dedicated VLAN to be used for FCoE traffic:

   ```
   Dell(conf)#interface vlan 1002
   ```

4. Configure an FCoE map to be applied on downstream (server-facing) Ethernet and upstream (core-facing) FC ports:

   ```
   Dell(config)# fcoe-map SAN_FABRIC_A

   Dell(config-fcoe-name)# fabric-id 1002 vlan 1002

   Dell(config-fcoe-name)# description "SAN_FABRIC_A"

   Dell(config-fcoe-name)# fc-map 0efc00

   Dell(config-fcoe-name)# keepalive

   Dell(config-fcoe-name)# fcf-priority 128

   Dell(config-fcoe-name)# fka-adv-period 8
   ```

5. Enable an upstream FC port:

   ```
   Dell(config)# interface fibrechannel 0/9

   Dell(config-if-fc-0)# no shutdown
   ```

6. Enable a downstream Ethernet port:

```
Dell(config)#interface tengigabitEthernet 0/1

Dell(conf-if-te-0)# no shutdown
```

Displaying NPIV Proxy Gateway Information

To display information on the NPG operation, use the show commands in the following table:

Table 67. Displaying NPIV Proxy Gateway Information

Command	Description
`show interfaces status`	Displays the operational status of Ethernet and Fibre Channel interfaces on the Aggregator with the NPG. **NOTE:** Although the show interface status command displays the Fiber Channel (FC) interfaces with the abbreviated label of 'Fc' in the output, if you attempt to specify a FC interface by using the `interface fc` command in the CLI interface, an error message is displayed. You must configure FC interfaces by using the `interface fi` command in CONFIGURATION mode.
`show fcoe-map [`**brief** `|` *map-name*`]`	Displays the Fibre Channel and FCoE configuration parameters in FCoE maps. Enter the **brief** keyword to display an overview of currently configured FCoE maps. Enter the name of an FCoE map to display the FC and FCoE parameters configured in the map to be applied on the Aggregator with the FC ports.
`show qos dcb-map` *map-name*	Displays configuration parameters in a specified DCB map.
`show npiv devices [`**brief**`]`	Displays information on FCoE and FC devices currently logged in to the NPG.
`show fc switch`	Displays the FC mode of operation and worldwide node (WWN) name of an Aggregator.

show interfaces status Command Example

```
Dell# show interfaces status
Port     Description  Status Speed     Duplex Vlan
Te 0/1                Up     10000 Mbit Full   1-4094
Te 0/2                Down   Auto       Auto   1-1001,1003-4094
Te 0/3                Up     10000 Mbit Full   1-1001,1003-4094
Te 0/4                Down   Auto       Auto   1-1001,1003-4094
Te 0/5                Up     10000 Mbit Full   1-4094
Te 0/6                Up     10000 Mbit Full   1-4094
Te 0/7                Up     10000 Mbit Full   1-4094
Te 0/8 toB300         Down   Auto       Auto   1-1001,1003-4094
Fc 0/9                Up     8000 Mbit  Full   --
Fc 0/10               Up     8000 Mbit  Full   --
```

```
Te 0/11               Down   Auto       Auto    --
Te 0/12               Down   Auto       Auto    --
```

Table 68. show interfaces status Field Descriptions

Field	Description
Port	Server-facing 10GbE Ethernet (Te), or fabric-facing Fibre Channel (FC) port with *slot*/*port* information.
Description	Text description of port.
Status	Operational status of port: Ethernet ports - up (transmitting FCoE and LAN storage traffic) or down (not transmitting traffic). Fibre Channel ports - up (link is up and transmitting FC traffic) or down (link is down and not transmitting FC traffic), link-wait (link is up and waiting for FLOGI to complete on peer SW port), or removed (port has been shut down).
Speed	Transmission speed (in Megabits per second) of Ethernet and FC ports, including auto-negotiated speed (Auto).
Duplex	Data transmission mode: Full (allows communication in both directions at the same time), Half (allows communication in both directions but not at the same time), Auto (auto-negotiated transmission).
VLAN	VLAN IDs of the VLANs in which the port is a member.

show fcoe-map Command Examples

```
Dell# show fcoe-map brief
Fabric-Name   Fabric-Id         Vlan-Id  FC-MAP    FCF-Priority  Config-
State      Oper-State
fid_1003        1003            1003     0efc03    128
ACTIVE            UP
fid_1004        1004            1004     0efc04    128
ACTIVE            DOWN


Dell# show fcoe-map fid_1003

Fabric Name         fid_1003
Fabric Id           1003
Vlan Id             1003
Vlan priority       3
FC-MAP              0efc03
FKA-ADV-Period      8
Fcf Priority        128
Config-State        ACTIVE
Oper-State          UP
Members
Fc 0/9
Te 0/11 Te 0/12
```

Table 69. show fcoe-map Field Descriptions

Field	Description
Fabric-Name	Name of a SAN fabric.
Fabric ID	The ID number of the SAN fabric to which FC traffic is forwarded.
VLAN ID	The dedicated VLAN used to transport FCoE storage traffic between servers and a fabric over the NPG. The configured VLAN ID must be the same as the fabric ID.
VLAN priority	FCoE traffic uses VLAN priority 3. This setting is not user-configurable.
FC-MAP	FCoE MAC-address prefix value - The unique 24-bit MAC address prefix that identifies a fabric.
FKA-ADV-period	Time interval (in seconds) used to transmit FIP keepalive advertisements.
FCF Priority	The priority used by a server to select an upstream FCoE forwarder.
Config-State	Indicates whether the configured FCoE and FC parameters in the FCoE map are valid: Active (all mandatory FCoE and FC parameters are correctly configured) or Incomplete (either the FC-MAP value, fabric ID, or VLAN ID are not correctly configured).
Oper-State	Operational status of the link to the fabric: up (link is up and transmitting FC traffic), down (link is down and not transmitting FC traffic), link-wait (link is up and waiting for FLOGI to complete on peer FC port), or removed (port has been shut down).
Members	Aggregator with the FC ports, which are members of the dedicated FCoE VLAN that carries storage traffic to the specified fabric.

show qos dcb-map Command Examples

```
Dell# show qos dcb-map dcbmap2

State   :Complete
PfcMode:ON
--------------------
PG:0 TSA:ETS  BW:50  PFC:OFF
Priorities:0 1 2 4 5 6 7


PG:1 TSA:ETS  BW:50  PFC:ON
Priorities:3
```

Table 70. show qos dcb-map Field Descriptions

Field	Description
State	Complete: All mandatory DCB parameters are correctly configured. In progress: The DCB map configuration is not complete. Some mandatory parameters are not configured.
PFC Mode	PFC configuration in the DCB map: On (enabled) or Off.
PG	Priority group configured in the DCB map.
TSA	Transmission scheduling algorithm used in the DCB map: Enhanced Transmission Selection (ETS).
BW	Percentage of bandwidth allocated to the priority group.
PFC	PFC setting for the priority group: On (enabled) or Off.
Priorities	802.1p priorities configured in the priority group.

show npiv devices brief Command Example

```
Dell# show npiv devices brief

Total NPIV Devices = 2
-------------------------------------------------------------------------------
------------------------

ENode-Intf      ENode-WWPN           FCoE-Vlan  Fabric-Intf     Fabric-Map
LoginMethod     Status
-------------------------------------------------------------------------------
------------------------

Te 0/11      20:01:00:10:18:f1:94:20   1003      Fc 0/9
fid_1003           FLOGI     LOGGED_IN
Te 0/12      10:00:00:00:c9:d9:9c:cb   1003      Fc 0/10
fid_1003           FDISC     LOGGED_IN
```

Table 71. show npiv devices brief Field Descriptions

Field	Description
Total NPIV Devices	Number of downstream ENodes connected to a fabric over the Aggregator with the NPG.
ENode-Intf	Aggregator with the Ethernet interface (*slot*/*port*) to which a server CNA is connected.
ENode-WWPN	Worldwide port name (WWPN) of a server CNA port.
FCoE-Vlan	VLAN ID of the dedicated VLAN used to transmit FCoE traffic to and from the fabric.
Fabric-Intf	Fabric-facing Fibre Channel port (*slot*/*port*) on which FC traffic is transmitted to the specified fabric.

Field	Description
Fabric-Map	Name of the FCoE map containing the FCoE/FC configuration parameters for the server CNA-fabric connection.
Login Method	Method used by the server CNA to log in to the fabric; for example: FLOGI - ENode logged in using a fabric login (FLOGI). FDISC - ENode logged in using a fabric discovery (FDISC).
Status	Operational status of the link between a server CNA port and a SAN fabric: Logged In - Server has logged in to the fabric and is able to transmit FCoE traffic.

show npiv devices Command Example

```
Dell# show npiv devices
ENode[0]:
ENode MAC     :     00:10:18:f1:94:21
ENode Intf    :     Te 0/11
FCF MAC       :     5c:f9:dd:ef:10:c8
Fabric Intf   :     Fc 0/9
FCoE Vlan     :     1003
Fabric Map    :     fid_1003
ENode WWPN    :     20:01:00:10:18:f1:94:20
ENode WWNN    :     20:00:00:10:18:f1:94:21
FCoE MAC      :     0e:fc:03:01:02:01
FC-ID         :     01:02:01
LoginMethod   :     FLOGI
Secs          :     5593
Status        :     LOGGED_IN

ENode[1]:
ENode MAC     :     00:10:18:f1:94:22
ENode Intf    :     Te 0/12
FCF MAC       :     5c:f9:dd:ef:10:c9
Fabric Intf   :     Fc 0/10
FCoE Vlan     :     1003
Fabric Map    :     fid_1003
ENode WWPN    :     10:00:00:00:c9:d9:9c:cb
ENode WWNN    :     10:00:00:00:c9:d9:9c:cd
FCoE MAC      :     0e:fc:03:01:02:02
FC-ID         :     01:02:01
LoginMethod   :     FDISC
Secs          :     5593
Status        :     LOGGED_IN
```

Table 72. show npiv devices Field Descriptions

Field	Description
ENode [*number*]	Server CNA that has successfully logged in to a fabric over an Aggregator with the Ethernet port in ENode mode.
Enode MAC	MAC address of a server CNA port.
Enode Intf	Port number of a server-facing Ethernet port operating in ENode mode.

Field	Description
FCF MAC	Fibre Channel forwarder MAC: MAC address of Aggregator with the FCF interface.
Fabric Intf	Fabric-facing Aggregator with the Fibre Channel port (*slot*/*port*) on which FCoE traffic is transmitted to the specified fabric.
FCoE VLAN	ID of the dedicated VLAN used to transmit FCoE traffic from a server CNA to a fabric and configured on both the server-facing Aggregator with the server CNA port.
Fabric Map	Name of the FCoE map containing the FCoE/FC configuration parameters for the server CNA-fabric connection.
Enode WWPN	Worldwide port name of the server CNA port.
Enode WWNN	Worldwide node name of the server CNA.
FCoE MAC	Fabric-provided MAC address (FPMA). The FPMA consists of the FC-MAP value in the FCoE map and the FC-ID provided by the fabric after a successful FLOGI. In the FPMA, the most significant bytes are the FC-MAP; the least significant bytes are the FC-ID.
FC-ID	FC port ID provided by the fabric.
LoginMethod	Method used by the server CNA to log in to the fabric; for example, FLOGI or FDISC.
Secs	Number of seconds that the fabric connection is up.
State	Status of the fabric connection: logged in.

show fc switch Command Example

```
Dell# show fc switch
Switch Mode : NPG
Switch WWN  : 10:00:5c:f9:dd:ef:10:c0
Dell#
```

Table 73. show fc switch Command Description

Field	Description
Switch Mode	Fibre Channel mode of operation of an Aggregator. Default: NPG (configured as an NPIV proxy gateway).
Switch WWN	Factory-assigned worldwide node (WWN) name of the Aggregator. The Aggregator WWN name is not user-configurable.

61

Upgrade Procedures

To find the upgrade procedures, go to the *Dell Networking OS Release Notes* for your system type to see all the requirements needed to upgrade to the desired Dell Networking OS version. To upgrade your system type, follow the procedures in the *Dell Networking OS Release Notes*.

Get Help with Upgrades

Direct any questions or concerns about the Dell Networking OS upgrade procedures to the Dell Technical Support Center. You can reach Technical Support:

- On the web: http://support.dell.com/
- By email: Dell-Force10_Technical_Support@Dell.com
- By phone: US and Canada: 866.965.5800, International: 408.965.5800.

62

Virtual LANs (VLANs)

Dell Networking OS supports virtual LANs (VLANs).

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 Dell Networking operating system (OS) 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 the Dell Networking OS VLANs, refer to these *Dell 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.

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

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 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

```
Dell(conf)#int tengig 3/2
Dell(conf-if)#no shut
Dell(conf-if)#switchport
Dell(conf-if)#show config
!
interface Tengigabitethernet 3/2
  no ip address
  switchport
  no shutdown
Dell(conf-if)#end
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM Status  Q Ports
* 1 Active    U Tengig 3/2
  2 Active    T Po1(So 0/0-1)
              T Tengig 3/0
Dell#
```

Port-Based VLANs

Port-based VLANs are a broadcast domain defined by different ports or interfaces. In the Dell Networking OS, a port-based VLAN can contain interfaces from different line cards within the chassis. Dell 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.

The Dell Networking OS 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 119. Tagged Frame Format

The tag header contains some key information that the Dell Networking OS 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 (optional)

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.

Example of Verifying a Port-Based VLAN

To view the configured VLANs, use the `show vlan` command in EXEC Privilege mode.

```
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs, R - Remote Port Mirroring 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 - VSN tagged
   i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

  NUM Status Description Q Ports
  1 Inactive a
* 20 Active              U Po32()
                         U Te 0/3,5,13,53-56
  1002 Active            T Te 0/3,13,55-56
Dell#
```

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, refer to 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 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
```

Add an Interface to Another VLAN

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.

```
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM Status Q Ports
* 1 Inactive
  2 Active   T Po1(So 0/0-1)
             T Tengig 3/0
  3 Active   T Po1(So 0/0-1)
             T Tengig 3/1
Dell#config
Dell(conf)#int vlan 4
Dell(conf-if-vlan)#tagged po 1
Dell(conf-if-vlan)#show conf
!
interface Vlan 4
  no ip address
  tagged Port-channel 1
Dell(conf-if-vlan)#end
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs
  NUM Status Q Ports
* 1 Inactive
  2 Active   T Po1(So 0/0-1)
             T Tengig 3/0
  3 Active   T Po1(So 0/0-1)
             T Tengig 3/1
  4 Active T Po1(So 0/0-1)
Dell#
```

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.

Move an Untagged Interface to Another VLAN

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 (gi 3/2) 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 (gi 3/2). Because the Default VLAN no longer contains any interfaces, it is listed as inactive.

```
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs
  NUM Status Q Ports
* 1 Active   U Tengig 3/2
  2 Active   T Po1(So 0/0-1)
             T Tengig 3/0
  3 Active   T Po1(So 0/0-1)
             T Tengig 3/1
  4 Inactive
Dell#conf
Dell(conf)#int vlan 4
Dell(conf-if-vlan)#untagged tengig 3/2
Dell(conf-if-vlan)#show config
!
interface Vlan 4
  no ip address
  untagged Tengigabitethernet 3/2
Dell(conf-if-vlan)#end
Dell#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

  NUM Status Q Ports
* 1 Inactive
  2 Active   T Po1(So 0/0-1)
             T Tengig 3/0
  3 Active   T Po1(So 0/0-1)
             T Tengig 3/1
  4 Active   U Tengig 3/2
Dell#
```

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 the Dell 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 mode 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: 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. The Dell 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`).

63

Virtual Link Trunking (VLT)

Dell Networking OS supports virtual link trunking (VLT).

Overview

VLT allows physical links between two chassis to appear as a single virtual link to the network core.

VLT reduces the role of spanning tree protocols (STPs) by allowing link aggregation group (LAG) terminations on two separate distribution or core switches, and by 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 peer devices have independent management planes. A chassis interconnect trunk between the VLT chassis maintains synchronization of L2/L3 control planes across the two VLT peers. The chassis interconnect trunk uses 10GE or 40GE user ports on the chassis.

VLT provides Layer 2 multipathing, creating redundancy through increased bandwidth, enabling multiple parallel paths between nodes and load-balancing traffic where alternative paths exist.

A separate backup link maintains heartbeat messages across an out-of-band management network. The backup link ensures that node failure conditions are correctly detected and are not confused with failures of the chassis interconnect trunk. VLT ensures that local traffic on a chassis does not traverse the chassis interconnect trunk and takes the shortest path to the destination via directly attached links.

Virtual link trunking 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.

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.

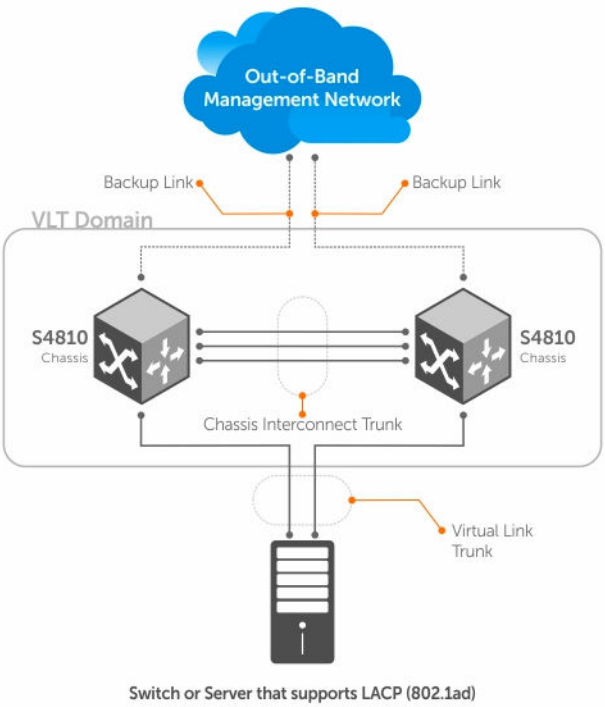


Figure 120. Virtual Link Trunking

Multi-domain VLT

A multi-domain VLT (mVLT) 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 LACP LAG to form a loop-free Layer 2 topology in the aggregation layer.

This configuration supports a maximum of four (4) nodes per mVLT domain, increasing the number of available ports and allowing for dual redundancy of the VLT.

Additionally, a VLT domain that is a member of one mVLT can be used in another mVLT configuration with a different VLT domain. Routing protocols such as OSPF are not compatible with mVLT; however, VLT domains can be used for routing. A separate Layer 3 router is not required for inter-VLAN communication.

The following figure shows how the core/aggregation port density in the Layer 2 topology is increased using mVLT. For inter-VLAN routing and other Layer 3 routing, a separate Layer 3 router is required.

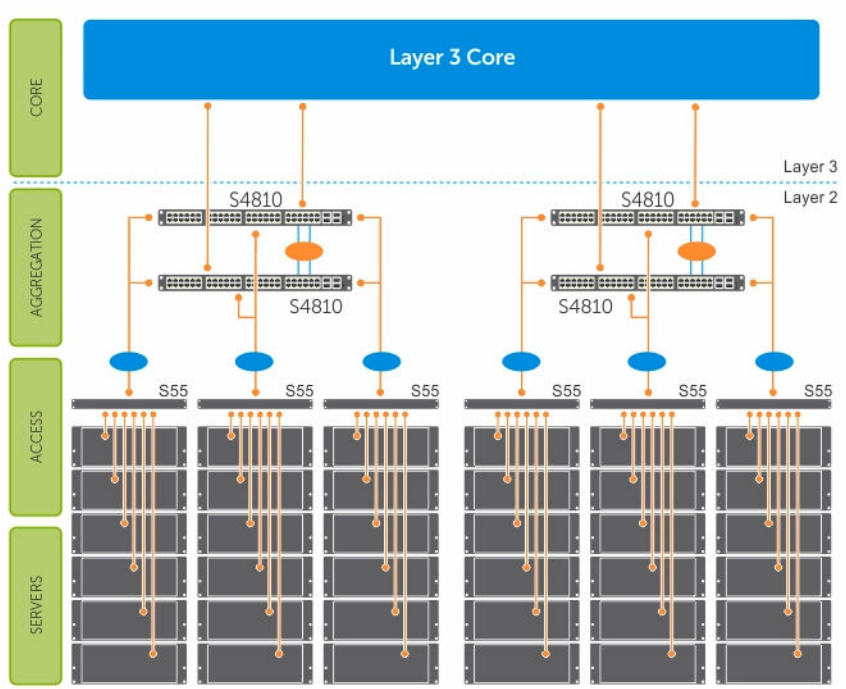


Figure 121. Multi-Domain VLT Example

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 vitality of 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).

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

- VLT port channel interfaces must be switch ports.
- If you include RSTP on the system, configure it before VLT. Refer to RSTP Configuration.
- Ensure that the spanning tree root bridge is at the Aggregation layer. If you enable RSTP on the VLT device, refer to RSTP and VLT for guidelines to avoid traffic loss.

- If you reboot both VLT peers in BMP mode and the VLT LAGs are static, the DHCP server reply to the DHCP discover offer may not be forwarded by the ToR to the correct node. To avoid this scenario, 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 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 Networking recommends configuring both the source and the receiver on a spanned VLT VLAN.
- 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 chassis in the VLT domain is assigned a primary role; the other chassis 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.

- In a VLT domain, the peer switches must run the same Dell Networking operating system (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 must consist of either 10G or 40G ports. A maximum of sixteen 10G or 40G ports are supported. A combination of 10G and 40G ports are not supported.
- A VLT interconnect over 1G 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 Networking strongly recommends configuring a static LAG for VLTi.
- IGMP state information is synchronized between the VLT chassis over the VLT interconnect.
- The traffic transmitted over VLT interconnect is prioritized, and allows you to configure the traffic class-to-queue assignment.
- 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 be synced. Dell Networking recommends retaining the default MTU allocation (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 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 96 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 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 (refer to 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 only).

  NOTE: 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).

- For detailed information about how to use VRRP in a VLT domain, refer to the following *VLT and VRRP Interoperability* section.
- For information about configuring IGMP Snooping in a VLT domain, refer to 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.
- IGMP snooping is supported over VLT ports. The multicast forwarding state is synchronized on both VLT peer switches. The IGMP snooping process on a VLT peer shares the learned group information with the other VLT peer over the chassis interconnect trunk.

- 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, and jumbo frames.
- Software features not supported with VLT
  - In a VLT domain, the following software features are supported on non-VLT ports: 802.1x, DHCP snooping, FRRP, ingress and egress QOS.
- VLT and VRRP interoperability
  - In a VLT domain, VRRP interoperates with virtual link trunks that carry traffic to and from access devices (refer to 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, refer to 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.
  - Also 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.

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 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 Port Delayed Restoration

With the Dell Networking OS version 8.3.12.0, 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.

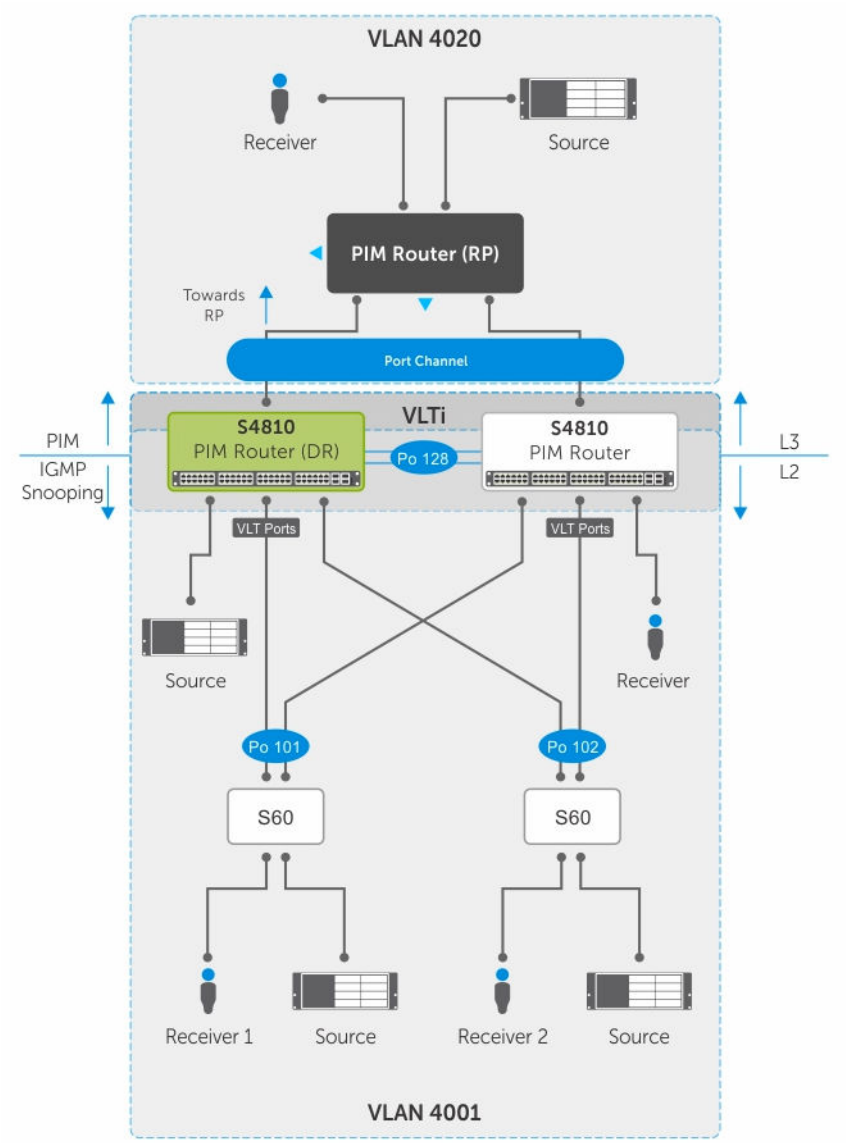


Figure 122. PIM-Sparse Mode Support on VLT

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.

The VLT peer nodes can also act as normal PIM routers on Layer 3 ports and on VLANS that do not have any VLT port members. In addition to being first-hop or last-hop routers, the peer node can also act as an intermediate router.

To route traffic to and from the multicast source and receiver that are connected to VLT ports, enable PIM-Sparse mode on the VLANs to which the VLT ports belong using the `ip pim sparse-mode` command. If IGMP Snooping is configured on these VLANs, the VLTi must be configured as a static multicast router port on both 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, traffic loss occurs until another VLT node is elected the designated router.

VLT Multicast

VLT multicast provides multiple alternate paths for resiliency against link and node failures.

This feature supports inter-server multicast communication between top-of-rack (ToR) switches using an inter-VLAN Layer 3 routing protocol (for example, PIM, IS-IS, or OSPF). It also provides traffic resiliency during multicast routing convergence after failure without disrupting or altering multicast routing behavior.

Layer 2 protocols from the ToR to the server are intra-rack and inter-rack. No spanning tree is required, but interoperability with spanning trees at the aggregation layer is supported. Communication is active-active, with no blocked links. MAC tables are synchronized between VLT nodes for bridging and IGMP snooping can be enabled.

VLT multicast is also scalable, so additional racks can be implemented in an existing system to provide additional resiliency and resources to accommodate an increased need or to anticipate future growth. You can enable VLT unicast and multicast routing across multiple configurations using VLT links. Protocols such as BGP, OSPF, and PIM are compatible with VLT multicast.

Important Points to Remember

- You cannot configure a VLT node as a rendezvous point (RP), but any PIM-SM compatible VLT node can serve as a designated router (DR).
- 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.
- You can enable ECMP on VLT nodes using VLT unicast; however, 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.
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.

VLT Unicast Routing

VLT unicast locally routes 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. In case a VLT node is down, resiliency is provided by a timer that allows you to configure the amount of time needed for peer recovery.

VLT unicast also assists in conflict resolution by updating the routing table when a path for an existing multicast route has been changed. You can enable VLT unicast and multicast routing across multiple configurations using VLT links. ECMP can be enabled on VLT nodes using VLT unicast; however, ECMP is not compatible on VLT nodes using VLT multicast. You must use a single VLAN.

VLT unicast is supported on both IPV6 / IPv4 . To enable VLT unicast, 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, VLAN configuration must be symmetrical on both peers. The same VLAN cannot be configured as Layer 2 on one node and as Layer 3 on the other node. Configuration mismatches are logged in the syslog and displayed in the output of the `show vlt inconsistency` command.

When you enable VLT unicast, VLAN wildcarding is enabled to support up to 4094 VLANs. If you enable VLT unicast, 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.

- For PVLAN, if the IP address is configured for the primary VLAN, L3 routing is enabled.

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.

Non-VLT ARP Sync

In the Dell Networking OS version 9.2(0.0), ARP entries (including ND entries) learned on other ports are synced with the VLT peer to support station move scenarios.

Prior to Dell Networking OS version 9.2.(0.0), only ARP entries learned on VLT ports were synced between peers.

Additionally, ARP entries resulting from station movements from VLT to non-VLT ports or to different non-VLT ports are learned on the non-VLT port and synced with the peer node. The peer node is updated to use the new non-VLT port.

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 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.

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.

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 Protocol 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)

```
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)

```
Dell_VLTpeer2(conf)#protocol spanning-tree rstp
Dell_VLTpeer2(conf-rstp)#no disable
Dell_VLTpeer2(conf-rstp)#bridge-priority 0
```

Configuring VLT

To configure virtual link trunking and create a VLT domain in which two switches are physically connected and treated as a single port channel by access devices, you must configure the following settings on each VLT peer device.
Prerequisites: Before you begin, make sure that both VLT peer switches are running the same Dell Networking OS version and are configured for RSTP as described in RSTP Configuration. 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 Notessection.

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 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 Configuring VLT and Connecting 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:

   - 10-Gigabit Ethernet: Enter `tengigabitethernet slot/port`.
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.

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. Repeat Steps 1 to 3 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. (Optional) After you configure the VLT domain on each peer switch on both sides of the interconnect trunk, by default, the system elects a primary and secondary VLT peer device.
   VLT DOMAIN CONFIGURATION mode

   ```
   primary-priority value
   ```

   To reconfigure the primary role of VLT peer switches, use the `primary-priority` command. To configure the primary role on a VLT peer, enter a lower value than the priority value of the remote peer.

   The priority values are from 1 to 65535. The default is **32768**.
3. (Optional) When you create a VLT domain on a switch, the system 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, the system 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:
   - 10-Gigabit Ethernet: enter `tengigabitethernet slot/port`.

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
   ```

   The valid port-channel ID numbers are from 1 to 128.
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.

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
```

The range is from 1 to 128.

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
   ```

   The range is from 1 to 4094.

Configure Multi-domain VLT (mVLT) (Optional)

To configure a multi-domain VLT between two VLT domains on your network, use the following procedure.
For a sample configuration, refer to theMulti-domain VLT section.

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.
2. Add one or more port interfaces to the port channel.
   INTERFACE PORT-CHANNEL mode

   ```
   channel-member interface
   ```

   interface specifies one of the following interface types:
   - 10-Gigabit Ethernet: Enter `tengigabitethernet slot/port`.
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 will act as the interconnect trunk.
   VLT DOMAIN CONFIGURATION mode

   ```
   peer-link port-channel id-number
   ```

   The range is from 1 to 128.
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, the system automatically creates a VLT-system MAC address used for internal system operations.
   VLT DOMAIN CONFIGURATION mode

   ```
   system-mac mac-address mac-address
   ```

   Use the `system-mac` command to explicitly configure the default MAC address for the domain by entering a new MAC address in the format: aaaa.bbbb.cccc.

   You must 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, the system automatically assigns a unique unit ID (0 or 1) to each peer switch. The unit IDs are used for internal system operations.
   VLT DOMAIN CONFIGURATION mode

   ```
   unit-id {0 | 1}
   ```

   Use the `unit-id` command to explicitly configure the default values on each peer switch.

   You must 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 multi-domain 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.
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
    ```

    Valid port-channel ID numbers are from 1 to 128.
11. Ensure that the port channel is active.
    INTERFACE PORT-CHANNEL mode

    ```
    no shutdown
    ```
12. **Add links to the mVLT 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. Repeat steps 1 through 15 for the VLT peer node in Domain 1.
17. Repeat steps 1 through 15 for the first VLT node in Domain 2.
18. Repeat steps 1 through 15 for the VLT peer node in Domain 2.

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:
    - Fast Ethernet: enter `fastethernet slot/port`.
    - 10-Gigabit Ethernet: enter `tengigabitethernet slot/port`.
    - Port channel: enter `port-channel {1-128}`.

Example of 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
```

Example of the `show vlt brief` Command

```
Dell#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
 Dell#
```

Example of 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
```

Example of 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
```

Example of 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
```

Example of 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
```

Example of 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
```

Connecting a VLT Domain

To connect a VLT domain to an attached access device, use the following commands.
For more information, refer to Verifying a VLT Configuration.

1. Configure the VLT domain with the same ID in VLT peer 1 and VLT peer 2.
   VLT DOMAIN

   ```
   vlt domain domain id
   ```
2. **Configure the VLTi between VLT peer 1 and VLT peer 2.** LACP/Static LAG can be configured between the peer units (not shown).
   CONFIGURATION mode

   ```
   interface port-channel port-channel id
   ```

   NOTE: To benefit from the protocol negotiations, Dell Networking recommends configuring VLTs used as facing hosts/switches with LACP. Both peers must use the same port channel ID.
3. Configure the peer-link port-channel in the VLT domains of each peer unit.
   INTERFACE PORTCHANNEL mode

   ```
   channel-member
   ```
4. **Configure the backup link between the VLT peer units.** Configure the peer 2 management ip/ interface ip for which connectivity is present in VLT peer 1.
   EXEC Privilege mode

   ```
   show running-config vlt
   ```
5. 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
   ```
6. **Configure the VLT links between VLT peer 1 and VLT peer 2 to the top of rack unit.** 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
   ```
7. 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
   ```
8. In the top of rack unit, configure LACP in the physical ports.
   EXEC Privilege mode

   ```
   show running-config entity
   ```
9. Verify VLT is running.
   EXEC mode

   ```
   show vlt brief

   show vlt detail
   ```

10. Verify the VLT LAG is running in both VLT peer units.
    EXEC mode or EXEC Privilege

    ```
    show interfaces interface
    ```

In the following sample VLT configuration steps, VLT peer 1 is FN IOM-2, VLT peer 2 is FN IOM-4, and the ToR is FN IOM-1.

NOTE: If you use a third-party ToR unit, Dell Networking recommends using static LAGs with VLT peers to avoid potential problems if the VLT peers are rebooted.

Configure the VLT domain with the same ID in VLT peer 1 and VLT peer 2

```
fniom-2(conf)#vlt domain 5
fniom-2(conf-vlt-domain)#

fniom-4(conf)#vlt domain 5
fniom-4(conf-vlt-domain)#
```

Configure the VLTi between VLT peer 1 and VLT peer 2

1. LACP/Static LAG can be configured between the peer units (not shown).
2. Configure the peer-link port-channel in the VLT domains of each peer unit.

```
fniom-2(conf)#interface port-channel 1
fniom-2(conf-if-po-1)#channel-member TenGigabitEthernet 0/4-7

fniom-2(conf)#no shutdown
fniom-4(conf)#interface port-channel 1
fniom-4(conf-if-po-1)#channel-member TenGigabitEthernet 0/4-7
fniom-4(conf)#no shutdown
```

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.

```
fniom-2#show running-config vlt
!
vlt domain 5
  peer-link port-channel 1
  back-up destination 10.11.206.58
fniom-2#

fniom-2#show interfaces managementethernet 0/0
Internet address is 10.11.206.43/16

fniom-4#show running-config vlt
!
vlt domain 5
  peer-link port-channel 1
  back-up destination 10.11.206.43
fniom-4#
fniom-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 0/40 in VLT peer 1 is connected to Te 0/48 of TOR and port Te 0/18 in VLT peer 2 is connected to Te 0/50 of TOR.

1. Configure the static LAG/LACP between 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 highlighted vlt-peer-lag port-channel 2 indicates that port-channel 2 is the port-channel id configured in VLT peer 2).

```
fniom-2#show running-config interface tengigabitethernet 0/40
!
interface TenGigabitEthernet 0/40
  no ip address
!
port-channel-protocol LACP
  port-channel 2 mode active
  no shutdown
fniom-2#
configuring VLT peer lag in VLT
fniom-2#show running-config interface port-channel 2
!
  interface Port-channel 2
  no ip address
  switchport
  vlt-peer-lag port-channel 2
  no shutdown
fniom-2#
fniom-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 0/40 (Up)
fniom-2#

fniom-4#show running-config interface tengigabitethernet 0/40
!
interface TenGigabitEthernet 0/40
no ip address
!
port-channel-protocol LACP
  port-channel 2 mode active
  no shutdown
fniom-4#
configuring VLT peer lag in VLT
fniom-4#show running-config interface port-channel 2
!
interface Port-channel 2
no ip address
switchport
vlt-peer-lag port-channel 2
no shutdown
fniom-4#
fniom-4#show interfaces port-channel 2 brief
Codes: L - LACP Port-channel
  LAG Mode Status Uptime    Ports
L 2   L2L3 up     03:33:14  Te 0/40 (Up)
fniom-4#
```

In the ToR unit, configure LACP on the physical ports

```
fniom-1#show running-config interface tengigabitethernet 0/48
!
interface TenGigabitEthernet 0/48
  no ip address
!
```

```
  port-channel-protocol LACP
  port-channel 100 mode active
fniom-1#show running-config interface tengigabitethernet 0/50
!
interface TenGigabitEthernet 0/50
  no ip address
!
port-channel-protocol LACP
port-channel 100 mode active
no shutdown
fniom-1#
fniom-1#show running-config interface port-channel 100
!
interface Port-channel 100
  no ip address
  switchport
  no shutdown
fniom-1#
fniom-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 0/48 (Up)
Te 0/50 (Up)
fniom-1#
```

Verify VLT is up. Verify that the VLTi (ICL) link, backup link connectivity (heartbeat status) and VLT peer link (peer chassis) are all up

```
Dell#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
 Dell#

Dell#FTOS(conf-if-vl-100)#show vlt detail
Local LAG Id Peer LAG Id Local Status Peer Status Active VLANs
------------ ----------- ------------ ------------ -------------
10        10             UP           UP     100, 200, 300, 400,
```

Verify the VLT LAG is up in both VLT peer units

```
fniom-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 0/40 (Up)
fniom-2#
fniom-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 0/18 (Up)
fniom-4#
```

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. For information on PVST+, refer to Per-VLAN Spanning Tree Plus (PVST+).

Run PVST+ on both VLT peer switches. PVST+ instance will be created for every VLAN configured in the system. PVST+ instances running in the Primary Peer will 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. PVST+ instance in Primary peer will send the role/state of VLT-LAGs for all VLANs to the Secondary peer. Secondary peer will use this information to program the hardware. PVST+ instance running in Secondary peer will not control the VLT-LAGs.

Dell 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 (refer to Verifying a VLT Configuration).

```
Dell#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 0/100   128.230  128  2000    FWD         0        0      90b1.1cf4.9b79
128.230
Te 0/103   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       (vltI)P2P No
Po 2       Desg   128.3    128  2000    FWD          0       (vlt) P2P No
Te 0/100   Desg   128.230  128  2000    FWD          0       P2P       Yes
Te 0/103   Desg   128.233  128  2000    FWD          0       P2P       No
Dell#
```

mVLT Configuration Example

The following example demonstrates the steps to configure multi-domain VLT (mVLT) in a network.

In this example there are two domains being configured. Domain 1 consists of Peer 1 and Peer 2; Domain 2 consists of Peer 3 and Peer 4 as shown.

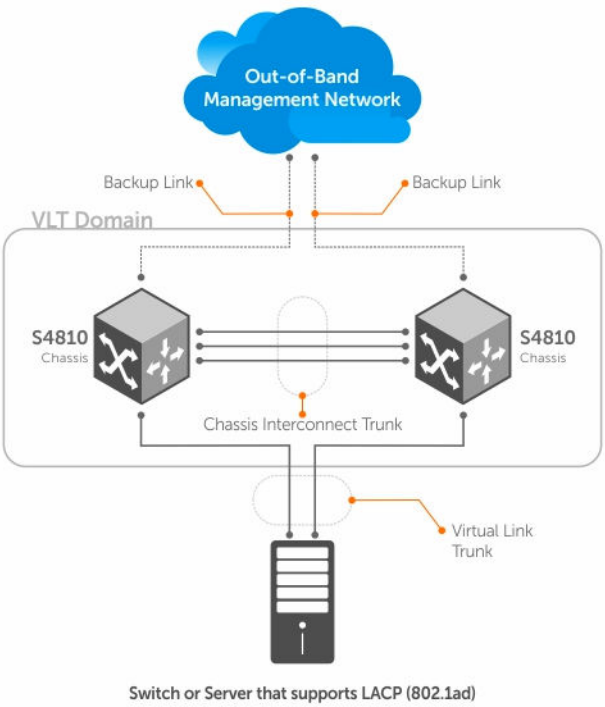


Figure 123. mVLT Configuration Example

In Domain 1, configure Peer 1 first, 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.

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 0/8-9
Domain_1_Peer1#no shutdown

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)#unit-id 0
```

Configure mVLT 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 mVLT port-channel on Peer 1

```
Domain_1_Peer1(conf)#interface range tengigabitethernet 0/16 - 17
Domain_1_Peer1(conf-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_1_Peer1(conf-if-range-te-0/16-17)#port-channel 100 mode active
Domain_1_Peer1(conf-if-range-te-0/16-17)#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 0/8-9
Domain_1_Peer2#no shutdown

Domain_1_Peer2(conf)#vlt domain 200
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)#unit-id 1
```

Configure mVLT 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 mVLT port-channel on Peer 2

```
Domain_1_Peer2(conf)#interface range tengigabitethernet 0/28 - 29
Domain_1_Peer2(conf-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_1_Peer2(conf-if-range-te-0/16-17)#port-channel 100 mode active
Domain_1_Peer2(conf-if-range-te-0/16-17)#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 0/8-9
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)#unit-id 0
```

Configure mVLT 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 mVLT port-channel on Peer 3

```
Domain_2_Peer3(conf)#interface range tengigabitethernet 0/19 - 20
Domain_2_Peer3(conf-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_2_Peer3(conf-if-range-te-0/16-17)#port-channel 100 mode active
Domain_2_Peer3(conf-if-range-te-0/16-17)#no shutdown
```

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 0/8-9
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)#unit-id 1
```

Configure mVLT 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 mVLT port-channel on Peer 4

```
Domain_2_Peer4(conf)#interface range tengigabitethernet 0/31 - 32
Domain_2_Peer4(conf-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_2_Peer4(conf-if-range-te-0/16-17)#port-channel 100 mode active
Domain_2_Peer4(conf-if-range-te-0/16-17)#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.

Example of Configuring PIM-Sparse Mode

Enable PIM Multicast Routing on the VLT node globally.

```
VLT_Peer1(conf)#ip multicast-routing
```

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
```

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
```

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
```

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 tengigabitethernet 0/5,6
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 tengigabitethernet 0/2
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(Te 0/2)
                          T Po100(Te 0/5,6)
```

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 tengigabitethernet 0/3,4
```

```
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 tengigabitethernet 0/8
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(Te 0/8)
                          T Po100(Te 0/3,4)
```

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 tengigabitethernet 1/18,22
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 Networking representative.

Table 74. 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%	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.

Description	Behavior at Peer Up	Behavior During Run Time	Action to Take
	threshold and when it drops below 80%.		
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.
Dell Networking 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 Networking 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 Networking software versions on the VLT

Description	Behavior at Peer Up	Behavior During Run Time	Action to Take
			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.
VLT LAG VLAN mismatch	A syslog error message is generated.	A syslog error message is generated.	Verify that the VLAN configuration is same for the VLT lags on both peers.

Specifying VLT Nodes in a 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 75. 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

VLT LAG Mode		PVLAN Mode of VLT VLAN		ICL VLAN Membership	Mac Synchronization
Peer1	Peer2	Peer1	Peer2		
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
		- 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 contains the following topics that 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.

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:
   - 1-Gigabit Ethernet: Enter `gigabitethernet slot/port`.
   - 10-Gigabit Ethernet: Enter `tengigabitethernet slot/port`.

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
   ```

   The range is from 1 to 128.
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
   ```

   The range is from 1 to 4094.

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

A proxy ARP-enabled device answers the ARP requests that are destined for another host or router. 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 the 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 is displayed in the `show config` command output.

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 enables 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 supported on symmetric VLANs only. Proxy ARP is enabled by default. Routing table must be symmetrically configured to support proxy ARP. For example, consider a sample topology in which VLAN 100 is configured on two VLT nodes, node 1 and node 2. ICL link is not configured 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 will not perform the ARP request for 20.1.1.2. Proxy ARP is supported only for the IP address 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 VLAN 100 and 200 are configured on the VLT peers, and if the VLAN 100 IP address is configured as 10.1.1.0/24 and the VLAN 200 IP address is configured 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 peer routing is enabled on both the VLT peers. If peer routing is disabled 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 the 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, then 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 an IP address is added or deleted, 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 will be stopped 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 will not perform proxy ARP for the peer IP addresses. IP address synchronization occurs again between the VLT peers.

Proxy ARP is enabled only if peer routing is enabled 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 peer routing is disabled 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 the VLT domain is removed on one of the VLT nodes, the peer routing configuration removal will be notified to the peer. In this case VLT peer node disables the proxy ARP. When the ICL link is removed on one of the VLT nodes by 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.

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 VLT domain

```
Dell(conf)#vlt domain 1
Dell(conf-vlt-domain)#peer-link port-channel 1
Dell(conf-vlt-domain)#back-up destination 10.16.151.116
Dell(conf-vlt-domain)#primary-priority 100
Dell(conf-vlt-domain)#system-mac mac-address 00:00:00:11:11:11
Dell(conf-vlt-domain)#unit-id 0
Dell(conf-vlt-domain)#

Dell#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
Dell#
```

Configure VLT LAG as VLAN-Stack Access or Trunk Port

```
Dell(conf)#interface port-channel 10
Dell(conf-if-po-10)#switchport
Dell(conf-if-po-10)#vlt-peer-lag port-channel 10
Dell(conf-if-po-10)#vlan-stack access
Dell(conf-if-po-10)#no shutdown

Dell#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
Dell#

Dell(conf)#interface port-channel 20
Dell(conf-if-po-20)#switchport
Dell(conf-if-po-20)#vlt-peer-lag port-channel 20
Dell(conf-if-po-20)#vlan-stack trunk
Dell(conf-if-po-20)#no shutdown

Dell#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
Dell#
```

Configure VLAN as VLAN-Stack VLAN and add the VLT LAG as Members to the VLAN

```
Dell(conf)#interface vlan 50
Dell(conf-if-vl-50)#vlan-stack compatible
Dell(conf-if-vl-50-stack)#member port-channel 10
Dell(conf-if-vl-50-stack)#member port-channel 20

Dell#show running-config interface vlan 50
!
interface Vlan 50
 vlan-stack compatible
 member Port-channel 10,20
 shutdown
Dell#
```

Verify that the Port Channels used in the VLT Domain are Assigned to the VLAN-Stack VLAN

Sample Configuration of VLAN-Stack Over VLT (Peer 2)

Configure VLT domain

```
Dell(conf)#vlt domain 1
Dell(conf-vlt-domain)#peer-link port-channel 1
Dell(conf-vlt-domain)#back-up destination 10.16.151.115
Dell(conf-vlt-domain)#system-mac mac-address 00:00:00:11:11:11
Dell(conf-vlt-domain)#unit-id 1
Dell(conf-vlt-domain)#
```

```
Dell#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
Dell#
```

Configure VLT LAG as VLAN-Stack Access or Trunk Port

```
Dell(conf)#interface port-channel 10
Dell(conf-if-po-10)#switchport
Dell(conf-if-po-10)#vlt-peer-lag port-channel 10
Dell(conf-if-po-10)#vlan-stack access
Dell(conf-if-po-10)#no shutdown

Dell#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
Dell#

Dell(conf)#interface port-channel 20
Dell(conf-if-po-20)#switchport
Dell(conf-if-po-20)#vlt-peer-lag port-channel 20
Dell(conf-if-po-20)#vlan-stack trunk
Dell(conf-if-po-20)#no shutdown

Dell#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
Dell#
```

Configure the VLAN as VLAN-Stack VLAN and add the VLT LAG as members to the VLAN

```
Dell(conf)#interface vlan 50
Dell(conf-if-vl-50)#vlan-stack compatible
Dell(conf-if-vl-50-stack)#member port-channel 10
Dell(conf-if-vl-50-stack)#member port-channel 20
Dell(conf-if-vl-50-stack)#

Dell#show running-config interface vlan 50
!
interface Vlan 50
 vlan-stack compatible
 member Port-channel 10,20
 shutdown
Dell#
```

Verify that the Port Channels used in the VLT Domain are Assigned to the VLAN-Stack VLAN

64

Virtual Router Redundancy Protocol (VRRP)

Dell Networking OS supports virtual router redundancy protocol (VRRP).

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 GigabitEthernet 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 GigabitEthernet 10/1. 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*.

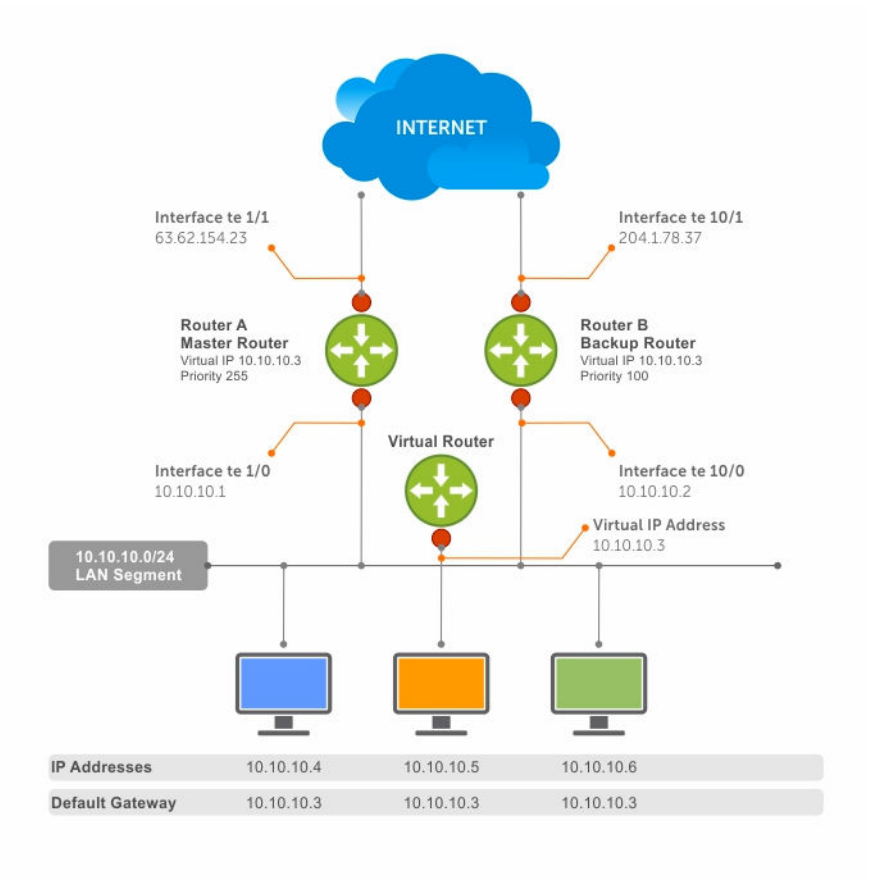


Figure 124. Basic VRRP Configuration

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.

VRRP Implementation

Dell Networking OS supports a total of 2000 VRRP groups on a switch and 255 VRRP groups per interface

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.

Default VRRP settings may affect the maximum number of groups that you can configure and work efficiently, as a result of hardware throttling VRRP advertisement packets reaching the CP on the switch. To avoid throttling VRRP advertisement packets, Dell Networking recommends increasing the VRRP advertisement interval to a value higher than the default value of 1 second.

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.

The recommendations in the following table vary, depending on several factors; for example, address resolution protocol (ARP) broadcasts, IP broadcasts, or spanning tree protocol (STP) before changing the advertisement interval. When the number of packets processed by the CP processor increases or decreases based on the dynamics of the network, the advertisement intervals in may increase or decrease accordingly.

Table 76. Recommended VRRP Advertise Intervals

Recommended Advertise Interval		Groups/Interface
Less than 250	1 second	255
Between 250 and 450	2–3 seconds	255
Between 450 and 600	3–4 seconds	255

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 (optional)
- Setting VRRP Initialization Delay (optional)

For a complete listing of all commands related to VRRP, refer to *Dell Networking OS Command Line Reference Guide*.

Creating a Virtual Router

To enable VRRP, create a virtual router. In Dell, 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
```

Example of Configuring VRRP

```
Dell(conf)#int tengig 1/1
Dell(conf-if-te-1/1)#vrrp-group 111
Dell(conf-if-te-1/1-vrid-111)#
```

Example of Verifying the VRRP Configuration

```
Dell(conf-if-te-1/1)#show conf
!
interface Tengigabitethernet 1/1
  ip address 10.10.10.1/24
!
  vrrp-group 111
  no shutdown
Dell(conf-if-te-1/1)#
```

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.

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.
2. Set the master switch to VRRP protocol version 3.
3. Set the backup switches to 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 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 the system 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.

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.

Example of the `virtual-address` Command

```
Dell(conf-if-te-1/1-vrid-111)#virtual-address 10.10.10.1
Dell(conf-if-te-1/1-vrid-111)#virtual-address 10.10.10.2
```

```
Dell(conf-if-te-1/1-vrid-111)#virtual-address 10.10.10.3
Dell(conf-if-te-1/1-vrid-111)#
```

Example of Verifying the Virtual IP Address Configuration

NOTE: In the following example, the primary IP address and the virtual IP addresses are on the same subnet.

```
Dell(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
Dell(conf-if-te-1/1)#
```

The following example shows the same VRRP group (VRID 111) configured on multiple interfaces on different subnets.

Example of Verifying the VRRP Group Priority

```
Dell#do show vrrp
------------------
Tengigabitethernet 1/1, VRID: 111, Version: 2 Net: 10.10.10.1
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
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: (none)
Dell#
```

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**.

Example of the `priority` Command

```
Dell(conf-if-te-1/2)#vrrp-group 111
Dell(conf-if-te-1/2-vrid-111)#priority 125
```

Example of Verifying the VRRP Group Priority

```
Dell#show vrrp
------------------
Tengigabitethernet 1/1, VRID: 111, Net: 10.10.10.1
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
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)
Dell(conf)#
```

Configuring VRRP Authentication

Simple authentication of VRRP packets ensures that only trusted routers participate in VRRP processes. When you enable authentication, the Dell 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.

To configure simple authentication, use the following command.

NOTE: Authentication for VRRPv3 is not supported.

- 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.

Example of `authentication-type` Command

The bold section shows the encryption type (encrypted) and the password.

```
Dell(conf-if-te-1/1-vrid-111)#authentication-type ?
Dell(conf-if-te-1/1-vrid-111)#authentication-type simple 7 force10
```

Example of Verifying the Configuration of VRRP Authentication

The bold section shows the encrypted password.

```
Dell(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
Dell(conf-if-te-1/1-vrid-111)#
```

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.

Example of Disabling Preempt

```
Dell(conf-if-te-1/1)#vrrp-group 111
Dell(conf-if-te-1/1-vrid-111)#no preempt
Dell(conf-if-te-1/1-vrid-111)#show conf
```

Example of Verifying Preempt is Disabled

```
Dell(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
Dell(conf-if-te-1/1-vrid-111)#
```

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 Networking 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.

1. Change the advertisement interval setting. INTERFACE-VRID mode `advertise-interval` *seconds* The range is from 1 to 255 seconds. The default is **1 second**.
2. 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**.

Example of the `advertise-interval` Command

Example of Verifying the Configured Advertisement Interval

The following example shows how to change the advertise interval using the `advertise-interval` command.

```
Dell(conf-if-te-1/1)#vrrp-group 111
Dell(conf-if-te-1/1-vrid-111)#advertise-interval 10
Dell(conf-if-te-1/1-vrid-111)#
```

The following example shows how to verify the advertise interval change using the `show conf` command

```
Dell(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
Dell(conf-if-te-1/1-vrid-111)#
```

Track an Interface or Object

You can set the Dell 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:

- 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.

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.

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
  ```

Example of the `track` Command

```
Dell(conf-if-te-1/1)#vrrp-group 111
Dell(conf-if-te-1/1-vrid-111)#track tengigabitethernet 1/2
Dell(conf-if-te-1/1-vrid-111)#
```

Example of Verifying the Tracking Configuration

```
Dell(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
Dell(conf-if-te-1/1-vrid-111)#
```

Example of Viewing Tracking Status

```
Dell#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 GigabitEthernet 13/2
  Tracked by:
  VRRP GigabitEthernet 7/30 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 GigabitEthernet 13/2
  Tracked by:
  VRRP GigabitEthernet 7/30 IPv6 VRID 1
```

Example of Viewing VRRP Configuration on an Interface

```
Dell#show running-config interface tengigabitethernet 1/3
!
```

```
  interface TenGigabitEthernet 1/3
  ip address 10.1.1.1/24
!
  vrrp-group 21
  virtual-address 10.1.1.2
  no shutdown
Dell#
```

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.

Set the delay timer on individual interfaces. The delay timer is supported on all physical interfaces, VLANs, and LAGs.

NOTE: When you reload a node that contains VRRP configuration and is enabled for VLT, Dell 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.

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**.

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.

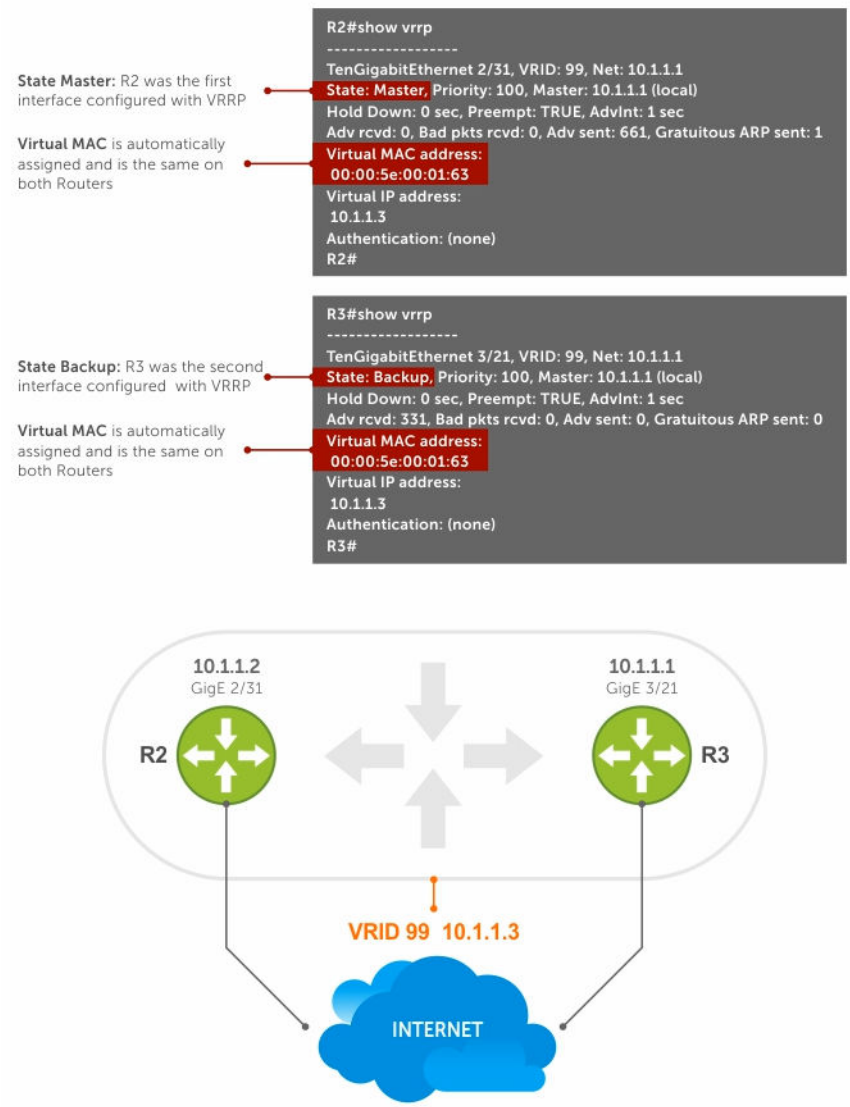


Figure 125. VRRP for IPv4 Topology

Example of Configuring VRRP for IPv4

```
R2(conf)#int tengig 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
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)#int tengig 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 10.1.1.2/24
!
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
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)
```

65

Debugging and Diagnostics

This chapter contains the following sections:.

- Debugging Aggregator Operation
- Software Show Commands
- Offline Diagnostics
- Trace Logs
- Show Hardware Commands

Supported Modes

Standalone, PMUX, VLT

Debugging Aggregator Operation

This section describes common troubleshooting procedures to use for error conditions that may arise during Aggregator operation.

All interfaces on the Aggregator are operationally down

This section describes how you can troubleshoot the scenario in which all the interfaces are down.

Symptom: All Aggregator interfaces are down.

Resolution: Ensure the port channel 128 is up and that the Aggregator-facing port channel on the top-of-rack switch is correctly configured.

Steps to Take:

1. Verify that uplink port-channel 128 is up (show interfaces port-channel 128 brief command) and display the status of member ports (show uplink-state-group 1 detail command).

```
Dell#show interfaces port-channel 128 brief
Codes: L - LACP Port-channel

     LAG  Mode  Status       Uptime      Ports
L    128  L2L3  up           17:36:24    Te 0/1     (Up)
                                         Te 0/2     (Up)
                                         Te 0/3     (Dwn)
                                         Te 0/4     (Dwn)
                                         Te 0/5     (Up)
                                         Te 0/6     (Dwn)
                                         Te 0/7     (Dwn)
                                         Te 0/8     (Up)

Dell#show uplink-state-group 1 detail
```

```
(Up): Interface up   (Dwn): Interface down   (Dis): Interface disabled

Uplink State Group    : 1       Status: Enabled, Up
Defer Timer           : 10 sec
Upstream Interfaces   : Po 128(Up)
Downstream Interfaces : Te 0/1(Up) Te 0/2(Up) Te 0/3(Dwn) Te 0/4(Dwn) Te
0/5(Up)
                        Te 0/6(Dwn) Te 0/7(Dwn) Te 0/8(Up)
```

2. Verify that the downstream port channel in the top-of-rack switch that connect to the Aggregator is configured correctly.

Broadcast, unknown multicast, and DLF packets switched at a very low rate

Symptom: Broadcast, unknown multicast, and DLF packets are switched at a very low rate.

By default, broadcast storm control is enabled on an Aggregator and rate limits the transmission of broadcast, unknown multicast, and DLF packets to 1Gbps. This default behavior is designed to avoid unnecessarily flooding these packets on all (4094) VLANs on all Aggregator interfaces (default configuration).

Resolution: Disable broadcast storm control globally on the Aggregator.

Steps to Take:

1. Display the current status of broadcast storm control on the Aggregator (show io-aggregator broadcast storm-control status command).

```
Dell#show io-aggregator broadcast storm-control status

Storm-Control Enabled

Broadcast Traffic limited to 1000 Mbps
```

2. Disable broadcast storm control (no io-aggregator broadcast storm-control command) and re-display its status.

```
Dell#config terminal
Dell(conf)#no io-aggregator broadcast storm-control
Dell(conf)#end
Dell#show io-aggregator broadcast storm-control status

Storm-Control Disabled
```

Flooded packets on all VLANs are received on a server

Symptom: All packets flooded on all VLANs on an Aggregator are received on a server, even if the server is configured as a member of only a subset of VLANs. This behavior happens because all Aggregator ports are, by default, members of all (4094) VLANs.

Resolution: Configure a port that is connected to the server with restricted VLAN membership.

Steps to Take:

1. Display the current port mode for Aggregator L2 interfaces (show interfaces switchport interface command)..

```
Dell#show interfaces switchport tengigabitethernet 0/1

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 0/1
802.1QTagged: Hybrid
SMUX port mode: Auto VLANs enabled
Vlan membership:
Q       Vlans
U       1
T       2-4094

Native VlanId:    1
```

2. Assign the port to a specified group of VLANs (vlan tagged command) and re-display the port mode status..

```
Dell(conf)#interface tengigabitethernet 0/1
Dell(conf-if-te-0/1)#vlan tagged 2-5,100,4010
Dell(conf-if-te-0/1)#

Dell#show interfaces switchport tengigabitethernet 0/1

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 0/1
802.1QTagged: Hybrid
SMUX port mode: Admin VLANs enabled
Vlan membership:
Q       Vlans
U       1
T       2-5,100,4010

Native VlanId:    1
```

Software show Commands

Use the `show version` and `show system stack-unit 0` commands as a part of troubleshooting an Aggregator's software configuration.

Table 77. Software show Commands

Command	Description
show version	Display the current version of Dell Networking OS software running on an Aggregator.

show version Command Example

```
Dell#show version
Dell Real Time Operating System Software
Dell Operating System Version:  2.0
Dell Application Software Version:  9-4(0-180)
Copyright (c) 1999-2014 by Dell Inc. All Rights Reserved.
Build Time: Sun Mar 30 20:15:19 PDT 2014
Build Path: /work.local/build/toolSpaces/tools05/E9-4-0/SW/SRC
Dell Networking OS uptime is 9 hour(s), 21 minute(s)

System image file is "dv-ci-stomp-tc-1-a1"

System Type: PE-FN-410S-IOA
```

```
Control Processor: MIPS RMI XLP with 2147483648 bytes of memory, core(s) 1.

128M bytes of boot flash memory.

  1 12-port GE/TE (FN)
 12 Ten GigabitEthernet/IEEE 802.3 interface(s)
Dell#
```

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, access, and read-write tests 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 MAC level, Physical level, external 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. These diagnostics also perform snake tests using virtual local area network (VLAN) configurations.

NOTE: Diagnostic is not allowed in Stacking mode, including member stacking. Avoid stacking before executing the diagnostic tests in the chassis.

Important Points to Remember

- You can only perform offline diagnostics on an offline standalone unit. You cannot perform diagnostics if the ports are configured in a stacking group. Remove the port(s) from the stacking group before executing the diagnostic test.
- 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
   ```

   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 - 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/fanout 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]:
```

```
Dell#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/fanout 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
Dell#
```

2. Confirm the offline status.

   EXEC Privilege mode

```
show system brief
```

```
Dell#show system brief

Stack MAC : 00:1e:c9:de:03:7b

--  Stack Info  --
Unit  UnitType     Status         ReqTyp          CurTyp
Version     Ports
--------------------------------------------------------------------------------
--------
  0   Management   offline        PE-FN-410S-IOA  PE-FN-410S-IOA
1-0(0-1862)  12
  1   Member       not present
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present

Dell#
```

Trace Logs

In addition to the syslog buffer, the Dell Networking OS buffers trace messages which are continuously written by various software tasks to report hardware and software events and status information.

Each trace message provides the date, time, and name of the Dell 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.

Example of the `dir flash:` Command

```
Dell#dir flash://TRACE_LOG_DIR
Directory of flash:/TRACE_LOG_DIR
  1 drwx   4096 Jan 17 2011 15:02:16 +00:00 .
  2 drwx   4096 Jan 01 1980 00:00:00 +00:00 ..
  3 -rwx 100583 Feb 11 2011 20:41:36 +00:00 failure_trace0_RPM0_CP

flash: 2143281152 bytes total (2069291008 bytes free)
```

Using the Show Hardware Commands

The `show hardware` command tree consists of commands used with the Aggregator switch. 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.

- View internal interface status of the stack-unit CPU port which connects to the external management interface.

  EXEC Privilege mode

  ```
  show hardware stack-unit {0-5} 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 {0-5} 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 {0-5} buffer total-buffer
  ```
- View the modular packet buffers details per unit and the mode of allocation.

  EXEC Privilege mode

  ```
  show hardware stack-unit {0-5} 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 {0-5} 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 {0-5} 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 {0-5} 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 {0-5} drops unit {0-0} port {33-56}
```

  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 {0-5} stack-port {33-56}
```

- View the counters in the field processors of the stack unit.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} counters
```

- View the details of the FP Devices and Hi gig ports on the stack-unit.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} details
```

- Execute a specified `bShell` command from the CLI without going into the bShell.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} execute-shell-cmd {command}
```

- View the Multicast IPMC replication table from the bShell.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} ipmc-replication
```

- View the internal statistics for each port-pipe (unit) on per port basis.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} port-stats [detail]
```

- View the stack-unit internal registers for each port-pipe.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} register
```

- View the tables from the bShell through the CLI without going into the bShell.

  EXEC Privilege mode

```
show hardware stack-unit {0-5} unit {0-0} table-dump {table name}
```

Environmental Monitoring

Aggregator 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
  ```

Example of the `show interfaces transceiver` Command

```
Dell#show interfaces tengigabitethernet 0/9 transceiver
SFP is present
SFP+ 9 Serial Base ID fields
SFP+ 9 Id                       = 0x03
SFP+ 9 Ext Id                   = 0x04
SFP+ 9 Connector                = 0x21
SFP+ 9 Transceiver Code         = 0x00 0x00 0x00 0x00 0x00 0x04 0x00 0x00
SFP+ 9 Encoding                 = 0x00
SFP+ 9 BR Nominal               = 0x67
SFP+ 9 Length(SFM)     Km       = 0x00
SFP+ 9 Length(OM3)     2m       = 0x00
SFP+ 9 Length(OM2)     1m       = 0x00
SFP+ 9 Length(OM1)     1m       = 0x00
SFP+ 9 Length(Copper)  1m       = 0x01
SFP+ 9 Vendor Rev               = A
SFP+ 9 Laser Wavelength         = 256 nm
SFP+ 9 CheckCodeBase            = 0xf2
SFP+ 9 Serial Extended ID fields
SFP+ 9 Options                  = 0x00 0x00
SFP+ 9 BR max                   = 0
SFP+ 9 BR min                   = 0
SFP+ 9 Vendor SN                = APF11040012888
SFP+ 9 Datecode                 = 110207
SFP+ 9 CheckCodeExt             = 0xb3
SFP+ 9 DOM is not supported

Dell#
```

Recognize an Over-Temperature 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.

NOTE: When the ingress air temperature exceeds 61°C, the Status LED turns Amber and a major alarm is triggered.

Example of the `show alarms threshold` Command

```
Dell#show alarms threshold

--  Temperature Limits (deg C)  --
-------------------------------------------------------------------------------
          Ingress-Air Off    Ingress-Air    Major Off    Major    Shutdown
Unit0        58                 61            84          86        90
Dell#
```

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, Dell Networking requires that you install blanks in all slots without a line card to control airflow for adequate system cooling.

NOTE: Exercise care when removing a card; if it has exceeded the major or shutdown thresholds, the card could be hot to the touch.

Example of the `show enivornment` Command

```
Dell#show environment

-- Unit Environment Status  --
Unit  Status        Temp   Voltage      TempStatus
-------------------------------------------------------------------------------
* 0   online        59C    ok            2

 * Management Unit

--  Thermal Sensor Readings (deg C)  --
Unit  Sensor0  Sensor1  Sensor2
-------------------------------
  0       55       45       58

Dell#
```

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: Line card 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).

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, which provide information about environmental monitoring hardware and hardware components.

Table 78. 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		
.1.3.6.1.4.1.6027.3.16.1.1.4	fpPacketBufferTable	View the modular packet buffers details per stack unit and the mode of allocation.
.1.3.6.1.4.1.6027.3.16.1.1.5	fpStatsPerPortTable	View the forwarding plane statistics containing the packet buffer usage per port per stack unit.
.1.3.6.1.4.1.6027.3.16.1.1.6	fpStatsPerCOSTable	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.

The application-specific integrated circuit (ASICs) implement the key functions of queuing, feature lookups, and forwarding lookups in hardware.

Forwarding processor (FP) ASICs provide Ethernet MAC functions, queueing, and buffering, as well as store feature and forwarding tables for hardware-based lookup and forwarding decisions. 1G and 10G interfaces use different FPs.

You can tune buffers at three locations

1. CSF — Output queues going from the CSF.
2. FP Uplink — Output queues going from the FP to the CSF IDP links.
3. Front-End Link — Output queues going from the FP to the front-end PHY.

All ports support eight queues, four for data traffic and four for control traffic. All eight queues are tunable.

Physical memory is organized into cells of 128 bytes. The cells are organized into two buffer pools — the dedicated buffer and the dynamic buffer.

- **Dedicated buffer** — this pool is reserved memory that other interfaces cannot use on the same ASIC or by other queues on the same interface. This buffer is always allocated, and no dynamic re-carving takes place based on changes in interface status. Dedicated buffers introduce a trade-off. They provide each interface with a guaranteed minimum buffer to prevent an overused and congested interface from starving all other interfaces. However, this minimum guarantee means that the buffer manager does not reallocate the buffer to an adjacent congested interface, which means that in some cases, memory is under-used.
- **Dynamic buffer** — this pool is shared memory that is allocated as needed, up to a configured limit. Using dynamic buffers provides the benefit of statistical buffer sharing. An interface requests dynamic buffers when its dedicated buffer pool is exhausted. The buffer manager grants the request based on three conditions:
  - The number of used and available dynamic buffers.
  - The maximum number of cells that an interface can occupy.
  - Available packet pointers (2k per interface). Each packet is managed in the buffer using a unique packet pointer. Thus, each interface can manage up to 2k packets.

You can configure dynamic buffers per port on both 1G and 10G FPs and per queue on CSFs. By default, the FP dynamic buffer allocation is 10 times oversubscribed. For the 48-port 1G card:

- Dynamic Pool= Total Available Pool(16384 cells) — Total Dedicated Pool = 5904 cells
- Oversubscription ratio = 10
- Dynamic Cell Limit Per port = 59040/29 = 2036 cells

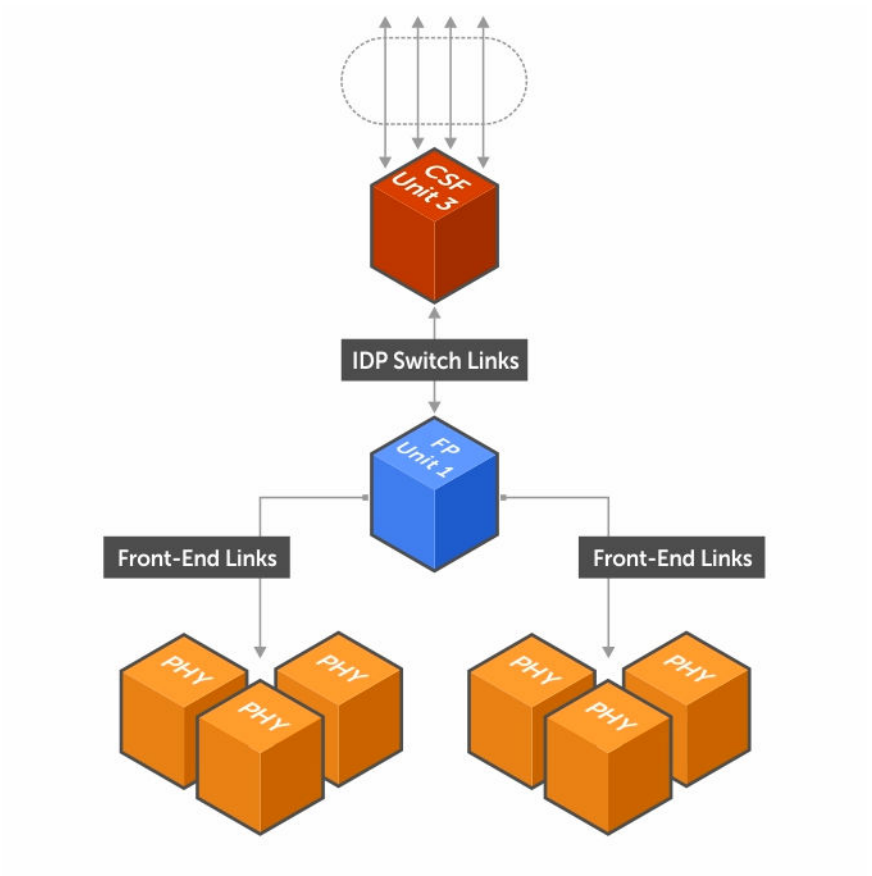


Figure 126. Buffer Tuning Points

Deciding to Tune Buffers

Dell Networking recommends exercising caution when configuring any non-default buffer settings, as tuning can significantly affect system performance. The default values work for most cases.
As a guideline, consider tuning buffers if traffic is bursty (and coming from several interfaces). In this case:

- Reduce the dedicated buffer on all queues/interfaces.
- Increase the dynamic buffer on all interfaces.
- Increase the cell pointers on a queue that you are expecting will receive the largest number of packets.

To define, change, and apply buffers, use the following commands.

- Define a buffer profile for the FP queues.
  CONFIGURATION mode

  ```
  buffer-profile fp fsqueue
  ```
- Define a buffer profile for the CSF queues.
  CONFIGURATION mode

  ```
  buffer-profile csf csqueue
  ```
- Change the dedicated buffers on a physical interface.

BUFFER PROFILE mode

```
buffer dedicated
```

- Change the maximum number of dynamic buffers an interface can request.
  BUFFER PROFILE mode

```
buffer dynamic
```

- Change the number of packet-pointers per queue.
  BUFFER PROFILE mode

```
buffer packet-pointers
```

- Apply the buffer profile to a CSF to FP link.
  CONFIGURATION mode

```
buffer csf linecard
```

Dell Networking OS Behavior: If you attempt to apply a buffer profile to a non-existent port-pipe, the system displays the following message: `%DIFFSERV-2-DSA_BUFF_CARVING_INVALID_PORT_SET: Invalid FP port-set 2 for linecard 2. Valid range of port-set is <0-1>`. However, the configuration still appears in the running-config.

Configuration changes take effect immediately and appear in the running configuration. Because under normal conditions all ports do not require the maximum allocation, the configured dynamic allocations can exceed the actual amount of available memory; this allocation is called oversubscription. If you choose to oversubscribe the dynamic allocation, a burst of traffic on one interface might prevent other interfaces from receiving the configured dynamic allocation, which causes packet loss.

You cannot allocate more than the available memory for the dedicated buffers. If the system determines that the sum of the configured dedicated buffers allocated to the queues is more than the total available memory, the configuration is rejected, returning a syslog message similar to the following: `00:04:20: %S50N:0 %DIFFSERV-2-DSA_DEVICE_BUFFER_UNAVAILABLE: Unable to allocate dedicated buffers for stack-unit 0, port pipe 0, egress port 25 due to unavailability of cells`.

Dell Networking OS Behavior: When you remove a buffer-profile using the `no buffer-profile [fp | csf]` command from CONFIGURATION mode, the buffer-profile name still appears in the output of the `show buffer-profile [detail | summary]` command. After a stack unit reset, the buffer profile correctly returns to the default values, but the profile name remains. Remove it from the `show buffer-profile [detail | summary]` command output by entering `no buffer [fp-uplink | csf] stack-unit port-set buffer-policy` from CONFIGURATION mode and `no buffer-policy` from INTERFACE mode.

To display the allocations for any buffer profile, use the `show` commands.

To display the default buffer profile, use the `show buffer-profile {summary | detail}` command from EXEC Privilege mode.

Example of Viewing the Default Buffer Profile

```
Dell#show buffer-profile detail interface tengigabitethernet 0/1
Interface tengig 0/1
Buffer-profile -
Dynamic buffer 194.88 (Kilobytes)
Queue# Dedicated Buffer  Buffer Packets
```

```
       (Kilobytes)
0      2.50              256
1      2.50              256
2      2.50              256
3      2.50              256
4      9.38              256
5      9.38              256
6      9.38              256
7      9.38              256
```

Example of Viewing the Buffer Profile Allocations

```
Dell#show running-config interface tengigabitethernet 0/6 !
interface TenGigabitEthernet 0/6
mtu 9252
switchport
no shutdown
buffer-policy myfsbufferprofile
```

Example of Viewing the Buffer Profile (Interface)

```
Dell#show buffer-profile detail int te 0/2
Interface Te 0/2
Buffer-profile fsqueue-fp
Dynamic buffer 1256.00 (Kilobytes)
Queue# Dedicated Buffer  Buffer Packets
       Kilobytes)
0      3.00              256
1      3.00              256
2      3.00              256
3      3.00              256
4      3.00              256
5      3.00              256
6      3.00              256
7      3.00              256
```

Example of Viewing the Buffer Profile (Linecard)

```
Dell#show buffer-profile detail fp-uplink stack-unit 0 port-set 0
Linecard 0 Port-set 0
Buffer-profile fsqueue-hig
Dynamic Buffer 1256.00 (Kilobytes)
Queue# Dedicated Buffer Buffer Packets
       (Kilobytes)
0      3.00             256
1      3.00             256
2      3.00             256
3      3.00             256
4      3.00             256
5      3.00             256
6      3.00             256
7      3.00             256
```

Using a Pre-Defined Buffer Profile

The Dell Networking OS provides two pre-defined 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 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 Networking OS Behavior: The buffer profile does not returned to the default, 4Q. 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 Gi 0/1. Please remove global pre-defined buffer profile.`

If the default buffer profile (4Q) is active, the system displays an error message instructing you to remove the default configuration using the `no buffer-profile global` command.

To apply a predefined buffer profile, use the following command.

- Apply one of the pre-defined buffer profiles for all port pipes in the system.
  CONFIGURATION mode

  ```
  buffer-profile global [1Q|4Q]
  ```

Sample Buffer Profile Configuration

The two general types of network environments are sustained data transfers and voice/data.

Dell Networking recommends a single-queue approach for data transfers.

Example of a Single Queue Application with Default Packet Pointers

```
!
buffer-profile fp fsqueue-fp
buffer dedicated queue0 3 queue1 3 queue2 3 queue3 3 queue4 3 queue5 3 queue6 3
queue7 3
buffer dynamic 1256
!
  buffer-profile fp fsqueue-hig
  buffer dedicated queue0 3 queue1 3 queue2 3 queue3 3 queue4 3 queue5 3 queue6
3 queue7 3
  buffer dynamic 1256

!
buffer fp-uplink stack-unit 0 port-set 0 buffer-policy fsqueue-hig
buffer fp-uplink stack-unit 0 port-set 1 buffer-policy fsqueue-hig
!
Interface range gi 0/1 - 48
buffer-policy fsqueue-fp

Dell#show run int Tengig 0/10
```

```
!
interface TenGigabitEthernet 0/10
```

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 0-5 cpu data-plane statistics`
- `show hardware stack-unit 0-5 cpu party-bus statistics`
- `show hardware stack-unit 0-5 drops unit 0-0 port 1-56`
- `show hardware stack-unit 0-5 stack-port 33-56`
- `show hardware stack-unit 0-5 unit 0-0 {counters | details | port-stats [detail] | register | ipmc-replication | table-dump}:`
- `show hardware {layer2| layer3} {eg acl |in acl} stack-unit 0-5 port-set 0-0`
- `show hardware layer3 qos stack-unit 0-5 port-set 0-0`
- `show hardware system-flow layer2 stack-unit 0-5 port-set 0-1 [counters]`
- `show hardware drops interface [range] interface`
- `show hardware stack-unit <id> buffer-stats-snapshot unit <id> resource x`
- `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 0-5 counters`
- `clear hardware stack-unit 0-5 unit 0-0 counters`
- `clear hardware stack-unit 0-5 cpu data-plane statistics`
- `clear hardware stack-unit 0-5 cpu party-bus statistics`
- `clear hardware stack-unit 0-5 stack-port 33-56`

Displaying Drop Counters

To display drop counters, use the following commands.

- Identify which stack unit, port pipe, and port is experiencing internal drops.

  `show hardware stack-unit 0-11 drops [unit 0 [port 0-63]]`
- Display drop counters.

  `show hardware stack-unit drops unit port`
- Identify which interface is experiencing internal drops.

  `show hardware drops interface interface`

Example of the `show hardware stack-unit` Command to View Drop Counters Statistics

```
Dell#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

Dell#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*

```
Dell#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 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

```
Dell#show hardware stack-unit 2 cpu data-plane statistics

bc pci driver statistics for device:
  rxHandle        :0
  noMhdr          :0
  noMbuf          :0
  noClus          :0
  recvd           :0
  dropped         :0
  recvToNet       :0
  rxError         :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(UNIT0)    :0
  rxPkt(UNIT1)    :0
  rxPkt(UNIT2)    :0
  rxPkt(UNIT3)    :0
  transmitted     :0
  txRequested     :0
  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(UNIT0)    :0
```

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

Example of Viewing Party Bus Statistics

```
Dell#show hardware stack-unit 2 cpu party-bus statistics
Input Statistics:
  27550 packets, 2559298 bytes
  0 dropped, 0 errors
Output Statistics:
  1649566 packets, 1935316203 bytes
  0 errors
```

Displaying 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

```
Dell#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
Dell#
```

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

   ```
   Dell(conf)#buffer-stats-snapshot

   Dell(conf)#no disable
   ```

   You must 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

```
Dell(conf)#buffer-stats-snapshot

Dell(conf)#no disable
```

You must 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

```
Dell#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

```
Dell# show hardware buffer-stats-snapshot resource interface
tengigabitethernet 0/0 queue all
Unit 0 unit: 0 port: 1 (interface Te 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
```

Restoring the Factory Default Settings

Restoring factory defaults 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 {0-5 | all} {clear-all | nvram}` 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, all units in the stack are placed into stand-alone mode.
- When you restore a single unit in a stack, only that unit is placed in stand-alone mode. No other units in the stack are affected.

- When you restore the units in stand-alone mode, the units remain in stand-alone mode after the restoration.
- After the restore is complete, the units power cycle immediately.

The following example shows the using the restore factory-defaults command to restore the Factory Default Settings.

Restoring the Factory Default Settings

```
Dell#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).
....
```

66

Standards Compliance

This chapter describes standards compliance for Dell Networking products.

NOTE: Unless noted, when a standard cited here is listed as supported by the Dell Networking Operating System (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.

IEEE Compliance

The following is a list of IEEE compliance.

802.1AB	LLDP
802.1D	Bridging
802.1p	L2 Prioritization
802.1Q	VLAN Tagging, Double VLAN Tagging, GVRP
802.3ac	Frame Extensions for VLAN Tagging
802.3ad	Link Aggregation with LACP
802.3ae	10 Gigabit Ethernet (10GBASE-W, 10GBASE-X)
802.3ak	10 Gigabit Ethernet (10GBASE-CX4)
802.3i	Ethernet (10BASE-T)
802.3u	Fast Ethernet (100BASE-FX, 100BASE-TX)
802.3x	Flow Control
802.1Qaz	Enhanced Transmission Selection
802.1Qbb	Priority-based Flow Control
ANSI/TIA-1057	LLDP-MED
MTU	12,000 bytes

RFC and I-D Compliance

The Dell Networking OS supports the following standards. The standards are grouped by related protocol. The columns showing support by platform indicate which version of Dell Networking OS first supports the standard.

General Internet Protocols

The following table lists the Dell Networking OS support per platform for general internet protocols.

Table 79. General Internet Protocols

RFC#	Full Name
768	User Datagram Protocol
793	Transmission Control Protocol
854	Telnet Protocol Specification
959	File Transfer Protocol (FTP)
1321	The MD5 Message-Digest Algorithm
1350	The TFTP Protocol (Revision 2)
1661	The Point-to-Point Protocol (PPP)
1989	PPP Link Quality Monitoring
1990	The PPP Multilink Protocol (MP)
1994	PPP Challenge Handshake Authentication Protocol (CHAP)
2474	Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers
2698	A Two Rate Three Color Marker
3164	The BSD syslog Protocol

General IPv4 Protocols

The following table lists the Dell Networking OS support per platform for general IPv4 protocols.

Table 80. General IPv4 Protocols

RFC#	Full Name
791	Internet Protocol
792	Internet Control Message Protocol
826	An Ethernet Address Resolution Protocol
1191	Path MTU Discovery
1305	Network Time Protocol (Version 3) Specification, Implementation and Analysis
1542	Clarifications and Extensions for the Bootstrap Protocol
1812	Requirements for IP Version 4 Routers
2131	Dynamic Host Configuration Protocol

Network Management

The following table lists the Dell Networking OS support per platform for network management protocol.

Table 81. Network Management

RFC#	Full Name
1155	Structure and Identification of Management Information for TCP/IP-based Internets
1156	Management Information Base for Network Management of TCP/IP-based internets
1157	A Simple Network Management Protocol (SNMP)
1212	Concise MIB Definitions
1215	A Convention for Defining Traps for use with the SNMP
1493	Definitions of Managed Objects for Bridges [except for the dot1dTpLearnedEntryDiscards object]
1901	Introduction to Community-based SNMPv2
2011	SNMPv2 Management Information Base for the Internet Protocol using SMIv2
2012	SNMPv2 Management Information Base for the Transmission Control Protocol using SMIv2
2013	SNMPv2 Management Information Base for the User Datagram Protocol using SMIv2
2024	Definitions of Managed Objects for Data Link Switching using SMIv2
2096	IP Forwarding Table MIB
2570	Introduction and Applicability Statements for Internet Standard Management Framework
2571	An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks
2572	Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)
2574	User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)
2575	View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)
2576	Coexistence Between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework
2578	Structure of Management Information Version 2 (SMIv2)

RFC#	Full Name
2579	Textual Conventions for SMIv2
2580	Conformance Statements for SMIv2
2618	RADIUS Authentication Client MIB, except the following four counters: radiusAuthClientInvalidServerAddresses radiusAuthClientMalformedAccessResponses radiusAuthClientUnknownTypes radiusAuthClientPacketsDropped
3635	Definitions of Managed Objects for the Ethernet-like Interface Types
2674	Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions
2787	Definitions of Managed Objects for the Virtual Router Redundancy Protocol
2819	Remote Network Monitoring Management Information Base: Ethernet Statistics Table, Ethernet History Control Table, Ethernet History Table, Alarm Table, Event Table, Log Table
2863	The Interfaces Group MIB
2865	Remote Authentication Dial In User Service (RADIUS)
3273	Remote Network Monitoring Management Information Base for High Capacity Networks (64 bits): Ethernet Statistics High-Capacity Table, Ethernet History High-Capacity Table
3416	Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)
3418	Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)
3434	Remote Monitoring MIB Extensions for High Capacity Alarms, High-Capacity Alarm Table (64 bits)
ANSI/TIA-1057	The LLDP Management Information Base extension module for TIA-TR41.4 Media Endpoint Discovery information
draft-grant-tacacs -02	The TACACS+ Protocol
IEEE 802.1AB	Management Information Base module for LLDP configuration, statistics, local system data and remote systems data components.

RFC#	Full Name
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)
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)
sFlow.org	sFlow Version 5
sFlow.org	sFlow Version 5 MIB
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)
FORCE10-LINKAGG-MIB	Force10 Enterprise Link Aggregation MIB
FORCE10-COPY-CONFIG-MIB	Force10 File Copy MIB (supporting SNMP SET operation)
FORCE10-MONMIB	Force10 Monitoring MIB
FORCE10-PRODUCTS-MIB	Force10 Product Object Identifier MIB
FORCE10-SS-CHASSIS-MIB	Force10 S-Series Enterprise Chassis MIB
FORCE10-SMI	Force10 Structure of Management Information
FORCE10-SYSTEM-COMPONENT-MIB	Force10 System Component MIB (enables the user to view CAM usage information)
FORCE10-TC-MIB	Force10 Textual Convention
FORCE10-TRAP-ALARM-MIB	Force10 Trap Alarm MIB
FORCE10-FIPS NOOPING-MI B	Force10 FIP Snooping MIB (Based on T11-FCoE-MIB mentioned in FC-BB-5)
FORCE10-DCB -MIB	Force10 DCB MIB
IEEE 802.1Qaz	Management Information Base extension module for IEEE 802.1 organizationally defined discovery information (LDP-EXT-DOT1-DCBX-MIB)
IEEE 802.1Qbb	Priority-based Flow Control module for managing IEEE 802.1Qbb

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/MIBs/MIB_OIDs.aspx

Some pages of iSupport require a login. To request an iSupport account, go to:

https://www.force10networks.com/CSPortal20/Support/AccountRequest.aspx

If you have forgotten or lost your account information, contact Dell TAC for assistance.

67

FC Flex IO Modules

This part provides a generic, broad-level description of the operations, capabilities, and configuration commands of the Fiber Channel (FC) Flex IO module.

FC Flex IO Module mentioned in this guide refers to FCF Port Combo Card.

FC Flex IO Modules

This part provides a generic, broad-level description of the operations, capabilities, and configuration commands of the Fiber Channel (FC) Flex IO module.

Understanding and Working of the FC Flex IO Modules

This chapter provides a generic, broad-level description of the operations and functionality of the Fiber Channel (FC) Flex IO module, and contains the following sections:

- FC Flex IO Modules Overview
- FC Flex IO Module Capabilities and Operations
- Guidelines for Working with FC Flex IO Modules
- Processing of Data Traffic
- Installing and Configuring the Switch
- Interconnectivity of FC Flex IO Modules with Cisco MDS Switches

FC Flex IO Modules Overview

The Fibre Channel (FC) Flex IO module is supported on FN IOM. The FN IOM switch installed with the FC Flex IO module functions as a top-of-rack edge switch that supports converged enhanced Ethernet (CEE) traffic — Fibre Channel over Ethernet (FCoE) for storage, Interprocess Communication (IPC) for servers, and Ethernet local area network (LAN) (IP cloud) for data — as well as FC links to one or more storage area network (SAN) fabrics. Although the FN IOM can act as a FIP snooping bridge (FSB) to provide FCoE transit switch capabilities, the salient and significant advantage of deploying the FC Flex IO module is to enable more streamlined and cohesive FCoE N-port identifier virtualization (NPIV) proxy gateway functionalities. The NPIV proxy gateway (NPG) provides FCoE-FC bridging behavior.

The FC Flex IO module offers a rich, comprehensive set of FCoE functionalities on the FX2 chassis by splitting the Ethernet and Fibre Channel (FC) traffic at the edge of the chassis. The FC switches that are connected directly to the FC Flex IO module provide Fibre Channel capabilities because the FC Flex IO module does not support full fabric functionalities. With the separation of Ethernet and FC packets performed at the edge of the chassis itself, you can use the FN IOM that contains an FC Flex IO module to connect to a SAN environment without the need for a separate ToR switch to operate as NPIV proxy gateways. The FN IOM can function in NPIV proxy gateway mode when an FC Flex IO module is present or in the FIP snooping bridge (FSB) mode when all the ports are Ethernet ports.

The FC Flex IO module uses the same baseboard hardware of the FN IOM and the FX2 chassis. You can insert the FC Flex IO module into any of the optional module slots of the FN IOM and it provides four FC

ports per module. If you insert only one FC Flex IO module, four ports are supported; if you insert two FC Flex IO modules, eight ports are supported.

By installing an FC Flex IO module, you can enable the FN IOM to directly connect to an existing FC SAN network. The FC Flex IO module uses the existing slots on the FN IOM and provides four or eight FC ports up to speed of 8 GbE per second. You can connect all of the FC ports to the same FC SAN fabric to yield FC bandwidth of up to 64GB. It is possible to connect some of the ports to a different FC SAN fabric to provide access to multiple fabric devices.

In a typical Fibre Channel storage network topology, separate network interface cards (NICs) and host bus adapters (HBAs) on each server (two each for redundancy purposes) are connected to LAN and SAN networks respectively. These deployments typically include a ToR SAN switch in addition to a ToR LAN switch. By employing converged network adapters (CNAs) that the FC Flex IO module supports, CNAs are used to transmit FCoE traffic from the server instead of separate NIC and HBA devices. In such a scenario, you can determine whether the FC or SAN packets and the Ethernet or LAN packets must be split within the chassis or by using a ToR switch to perform this splitting.

If you want to segregate the LAN and SAN traffic within the chassis, you can employ switches such as the Dell M8428-k Converged 10GbE Switch or FC-only switches such as the Dell M5424 switch module. You can also use the S5000 Switch as a ToR switch to separate the LAN and SAN traffic at the ToR. By using the FC Flex IO module, you can optimally and effectively split the LAN and SAN traffic at the edge of the blade chassis itself. You can deploy the FC Flex IO module can be deployed in the enterprise and data center switching networks to leverage and derive the advantages of a converged Ethernet network.

The FC Flex IO module is not an FCF switch, but it offers FCoE capabilities from the server to the FN IOM, and native FC capability in the uplink direction to the SAN switches. Although the FC Flex IO module does not support all of the FCF characteristics, such as full-blown name services or zone parameters, it presents the most flexible solution in interoperating with third-party switches that enable the splitting of LAN and SAN traffic. With the FN IOM being well-established systems in the switch domain, you can install the FC Flex IO module to enhance and increase the converged Ethernet network performance and behavior. With the FC Flex IO module, the FN IOM provide 8 1GbE or 10 GbE server-facing ports and the option to add two FC Flex IO modules that offer up to 8 8Gb Fibre Channel ports for uplink traffic in addition to the fixed two 40GbE ports on the FN IOM.

You can configure one of the following upstream (fabric-facing) FC ports:

- Two 40GbE and eight 8GB FC ports
- Four 40GbE and four 8GB FC ports
- Two 40GbE, four 10GbE, and four 8GB FC ports
- Two 40GbE, four 10GBASE-T, and four 8GB FC ports

FC Flex IO Module Capabilities and Operations

The FC Flex IO module has the following characteristics:

- You can install one or two FC Flex IO modules on the FN IOM. Each module supports four FC ports.
- Each port can operate in 2Gbps, 4Gbps, or 8Gbps of Fibre Channel speed.
- All ports on an FC Flex IO module can function in the NPIV mode that enables connectivity to FC switches or directors, and also to multiple SAN topologies.

- It automatically senses the current speed when the port link is up. Valid link speeds are 2 Gbps, 4 Gbps, and 8 Gbps.
- By default, the FC ports are configured to operate in N-port mode to connect to an F port on an FC switch in a fabric. You can apply only one FCoE map on an FC port. An N-Port is a port on the node of an FC device and is called a node port.
- There should a maximum of 64 server fabric login (FLOGI) requests or fabric discovery (FDISC) requests per server MAC address before being forwarded by the FC Flex IO module to the FC core switch. Without user configuration, only 32 server login sessions are permitted for each server MAC address. To increase the total number of sessions to 64, use the `max sessions` command.
- A distance of up to 300 meters is supported at 8 Gbps for Fibre Channel traffic.
- Multiple domains are supported in an NPIV proxy gateway (NPG).
- You cannot configure the FN IOM switches in Stacking mode if the switches contain the FC Flex IO module. Similarly, FC Flex IO modules do not function when you insert them in to a stack of FN IOM switches.
- If the switch contains FC Flex modules, you cannot create a stack, and a log message states that stacking is not supported unless the switches contain only FC Flex modules.

Guidelines for Working with FC Flex IO Modules

The following guidelines apply to the FC Flex IO module:

- All the ports of FC Flex IO modules operate in FC mode, and do not support Ethernet mode.
- FC Flex IO modules are not supported in the chassis management controller (CMC) GUI.
- The only supported FCoE functionality is NPIV proxy gateway. Configure the other FCoE services, such as name server, zone server, and login server on an external FC switch.
- With the FC Flex IO module, the FN IOM continues to support bare metal provisioning (BMP) on any Ethernet port. BMP is not supported on FC ports. BMP improves accessibility to the FN IOM by automatically loading pre-defined configurations and boot images that are stored in file servers. You can use BMP on a single switch or on multiple switches.
- FC Flex IOM module is a field-replaceable unit (FRU). Its memory type is electrically erasable programmable read-only memory (EEPROM), which enables it to save manufacturing information, such as the serial number. It is hot-swappable, assuming that the module that is removed is replaced by the same type of module in that same slot.
- The FC Flex IO does not have persistent storage for any runtime configuration. All the persistent storage for runtime configuration is on the FN IOM baseboard.
- With both FC Flex IO modules present in the FN IOM switches, the power supply requirement and maximum thermal output are the same as these parameters needed for the M1000 chassis.
- Each port on the FC Flex IO module contains status indicators to denote the link status and transmission activity. For traffic that is being transmitted, the port LED shows a blinking green light. The Link LED displays solid green when a proper link with the peer is established. If there is no connectivity, the LEDs are not lit
- The FN IOM switches continue to operate in FCoE Gateway mode even if connectivity to a ToR switch does not exist.
- Active fabric manager (AFM) is compatible with FC Flex IO modules.

- All SNMP MIBs that are supported for FN IOM switches apply equally for FC Flex IO modules. The interface MIB indicates the FC interface when you install the FC flex IO module. The interface MIB statistical counters compute and display the FC interface metrics.
- When the Dell Networking OS sends FC frames (the initial FLOGI or FLOGO messages), or converts FLOGI to FDISC messages or processes any internally generated FC frames, the software computes and verifies the FC cyclic redundancy check (CRC) value before sending the frame to FC ports.
- Fabric worldwide name (WWN) verification is available for eight FC ports. Single-switching WWN capability is provided when the switch operates in NPIV mode.
- Ensure that the NPIV functionality is enabled on the upstream switches that operate as FC switches or FCoE forwarders (FCF) before you connect the FC port of the FN IOM to these upstream switches.
- While storage traffic traverses through FC Flex IO modules and the Ethernet uplink port-channel status changes (with DCB enabled on an adjacent switch), FCoE traffic is disrupted. This problem does not occur if Ethernet traffic is not involved and only FCoE traffic is transmitted. Also, if DCB on the ToR switch is disabled, traffic disruption does not occur.

Port Numbering for FC Flex IO Modules

Even-numbered ports are at the bottom of the I/O panel and for modules odd-numbered ports are at the top of the I/O panel. When installed in a PowerEdge M1000e Enclosure, the FN IOM ports are numbered 33 to 56 from the bottom to the top of the switch. The following port numbering convention applies to the FC Flex IO module:

- In expansion slot 0, the ports are numbered 41 to 44.
- In expansion slot 1, the ports are numbered 49 to 52.

Installing the Optics

The following optical ports are supported on the FC Flex IO module using one of the supported breakout cables:

- 4G or 8G Fibre Channel small form-factor pluggable plus (SFP+) optics module and LC connectors over a distance of 150 meters.
- 4G or 8G Fibre Channel SFP+ optics module and LC connectors over a distance of 4 km.

CAUTION:
Electrostatic discharge (ESD) damage can occur if the components are mishandled. Always wear an ESD-preventive wrist or heel ground strap when handling the FC Flex IO module and its components.

WARNING: When working with optical fibres, follow all the warning labels and always wear eye protection. Never look directly into the end of a terminated or unterminated fibre or connector as it may cause eye damage.

1.
   - Position the optic so it is in the correct position. The optic has a key that prevents it from being inserted incorrectly.
   - Insert the optic into the port until it gently snaps into place.

NOTE:

1. When you cable the ports, be sure not to interfere with the airflow from the small vent holes above and below the ports.

Processing of Data Traffic

The Dell Networking OS determines the module type that is plugged into the slot. Based on the module type, the software performs the appropriate tasks. The FC Flex IO module encapsulates and decapsulates the FCoE frames. The module directly switches any non-FCoE or non-FIP traffic, and only FCoE frames are processed and transmitted out of the Ethernet network.

When the external device sends FCoE data frames to the switch that contains the FC Flex IO module, the destination MAC address represents one of the Ethernet MAC addresses assigned to FC ports. Based on the destination address, the FCoE header is removed from the incoming packet and the FC frame is transmitted out of the FC port. The flow control mechanism is performed using per-priority flow control to ensure that frame loss does not occur owing to congestion of frames.

Operation of the FIP Application

The NPIV proxy gateway terminates the FIP sessions and responses to FIP messages. The FIP packets are intercepted by the FC Flex IO module and sent to the Dell Networking OS for further analysis. The FIP application responds to the FIP VLAN discovery request from the host based on the configured FCoE VLANs. For every ENode and VN_Port that is logged in, the FIP application responds to keepalive messages for the virtual channel. If the FC link becomes inactive or a logging off of the switch occurs, the FIP engine sends clear virtual link (CVL) messages to the host. The FIP application also responds to solicited advertisements from the end-device. In addition, the FIP application periodically sends advertisement packets to the end-devices for each FCF that is part of the NPIV proxy gateway.

On FN IOM switches, you can configure the switch to operate in FIP Snooping or NPIV mode.

If the FN IOM Switch functions in the NPIV mode and you attempt to set the uplink port to be an FCF or a bridge port, a warning message displays and the settings are not saved.

Operation of the NPIV Proxy Gateway

The NPIV application on the FC Flex IO module manages the FC functionalities configured in Dell Networking OS. After the FC link comes up, the gateway sends the initial FLOGI request to the connected switch using the switch and port WWN methods. After a successful login, the NPIV gateway sends a notification to inform the CNA that the FCF available to log in. The source address of the FIP advertisement and FIP discovery advertisement response contain the MAC address of the FC Flex IO module port. Depending on the number of login sessions on a particular FCF, the NPIV gateway can load-balance the login sessions from ENodes.

The NPIV application performs the FLOGI to FDISC conversion and sends the new FC frame on the associated FC ports. After the external switch responds to the FLOGI request, the NPIV gateway establishes the NPIV session and sends the frame to the FIP application. The FIP application establishes virtual links to convert FCoE FLOGI accept messages into FIP FLOGI accept messages. The corresponding ACL for the accept message is then applied. If a FIP timeout from ENode or VN_PORT occurs, the NPIV application performs the FC fabric logout to the external FC switch. The NPIV application manages the sessions between the FCoE and the FC domain.

Installing and Configuring the Switch

After you unpack the FN IOM, refer to the flow chart in the following figure for an overview of the steps you must follow to install the blade and perform the initial configuration.

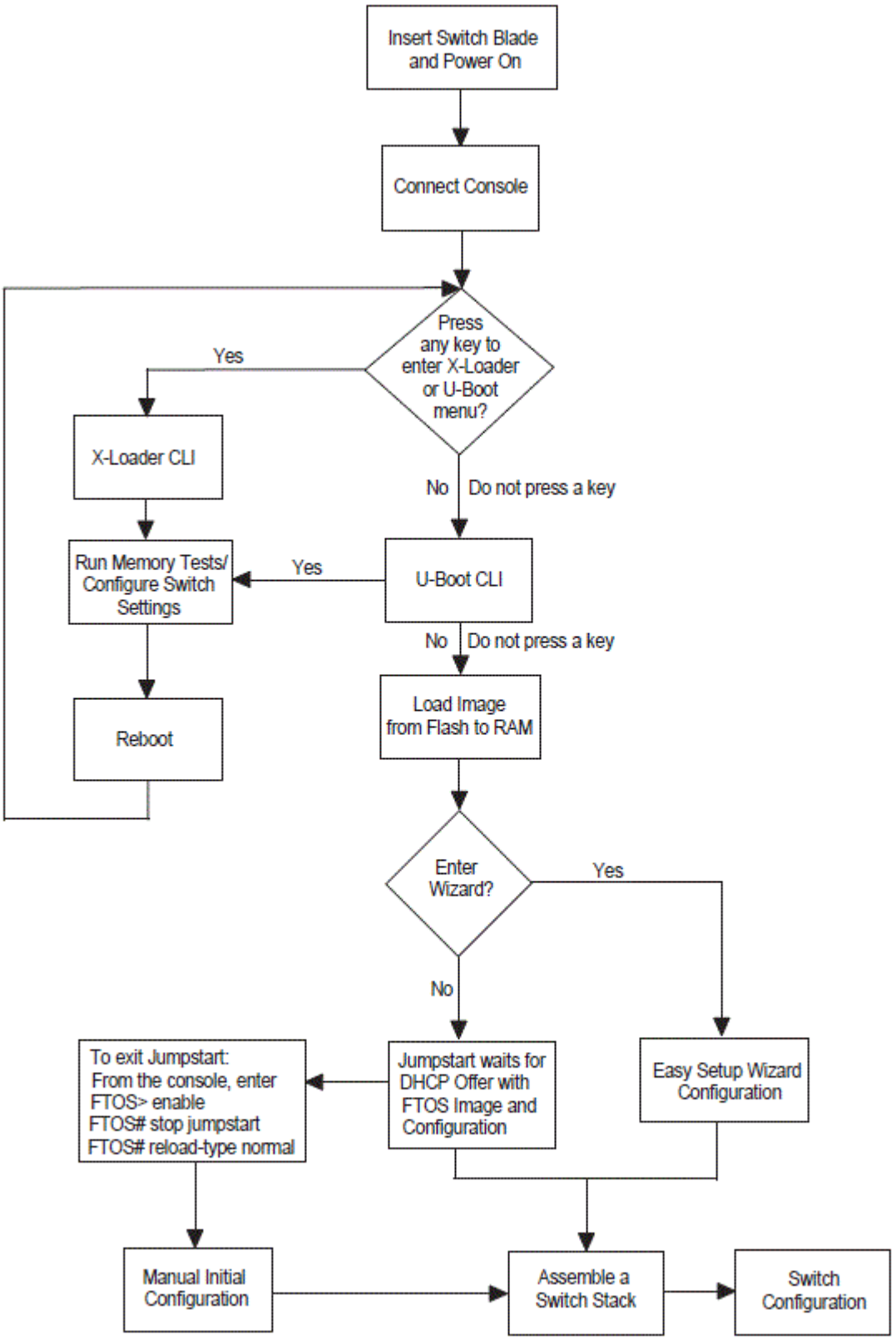


Figure 127. Installing and Configuring Flowchart for FC Flex IO Modules

To see if a switch is running the latest Dell Networking OS version, use the `show version` command. To download a Dell Networking OS version, go to http://support.dell.com.

Installation

Site Preparation

Before installing the switch or switches, make sure that the chosen installation location meets the following site requirements:

- **Clearance** — There is adequate front and rear clearance for operator access. Allow clearance for cabling, power connections, and ventilation.
- **Cabling** — The cabling is routed to avoid sources of electrical noise such as radio transmitters, broadcast amplifiers, power lines, and fluorescent lighting fixtures.
- **Ambient Temperature** — The ambient switch operating temperature range is 10° to 35ºC (50° to 95ºF).

1. Decrease the maximum temperature by 1°C (1.8°F) per 300 m (985 ft.) above 900 m (2955 ft.).
2. **Relative Humidity** — The operating relative humidity is 8 percent to 85 percent (non-condensing) with a maximum humidity gradation of 10 percent per hour.

Unpacking the Switch

Package Contents

When unpacking each switch, make sure that the following items are included:

- One Dell Networking FN IOM module
- One USB type A-to-DB-9 female cable
- Getting Started Guide
- Safety and Regulatory Information
- Warranty and Support Information
- Software License Agreement

Unpacking Steps

1. Before unpacking the switch, inspect the container and immediately report any evidence of damage.
2. Place the container on a clean, flat surface and cut all straps securing the container.
3. Open the container or remove the container top.
4. Carefully remove the switch from the container and place it on a secure and clean surface.
5. Remove all packing material.
6. Inspect the product and accessories for damage.

After you insert a Flex IO module into an empty slot, you must reload the FN IOM for the module. If you remove an installed module and insert a different module type, an error message displays to remind you that the slot is configured for a different type of Flex IO module. You must reload the switch to make the Flex IO module operational.

Interconnectivity of FC Flex IO Modules with Cisco MDS Switches

In a network topology that contains Cisco MDS switches, FC Flex IO modules that are plugged into the FN IOM switches enable interoperation for a robust, effective deployment of the NPIV proxy gateway and FCoE-FC bridging behavior. In an environment that contains FC Flex IO modules and Cisco MDS switches, perform the following steps:

- Insert the FC Flex IO module into any of the optional module slots of the FN IOM Switch and reload the switch.
- When the device is reloaded, NPIV mode is automatically enabled.
- Configure the NPIV-related commands on FN IOM.

After you perform the preceding procedure, the following operations take place:

- A physical link is established between the FC Flex I/O module and the Cisco MDS switch.
- The FC Flex I/O module sends a proxy FLOGI request to the upstream F_Port of the FC switch or the MDS switch. The F_port accepts the proxy FLOGI request for the FC Flex IO virtual N_Port. The converged network adapters (CNAs) are brought online and the FIP application is run.
- Discovery of the VLAN and FCF MAC addresses is completed.
- The CNA sends a FIP fabric login (FLOGI) request to the FC Flex IO module, which converts FLOGI to FDISC messages or processes any internally generated FC frames and sends these messages to the SAN environment.
- When the FC fabric discovery (FDISC) accept message is received from the SAN side, the FC Flex IO module converts the FDISC message again into an FLOGI accept message and transmits it to the CNA.
- Internal tables of the switch are then programmed to enable the gateway device to forward FCoE traffic directly back and forth between the devices.
- The FC Flex IO module sends an FC or FCoE registered state change notification (RSCN) message to the upstream or downstream devices whenever an error occurs in the appropriate direction.
- An F_Port is a port on an FC switch that connects to an N_Port of an FC device and is called a fabric port.

By default, the NPIV functionality is disabled on the Cisco MDS switch; enable this capability before you connect the FC port of the FN IOM to these upstream switches.

Data Center Bridging, Fibre Channel over Ethernet, and NPIV Proxy Gateway features are supported on the FC Flex IO modules. For detailed information about these applications and their working, see the corresponding chapters for these applications in this manual.

The following figures illustrate two deployment scenarios of configuring FC Flex IO modules:

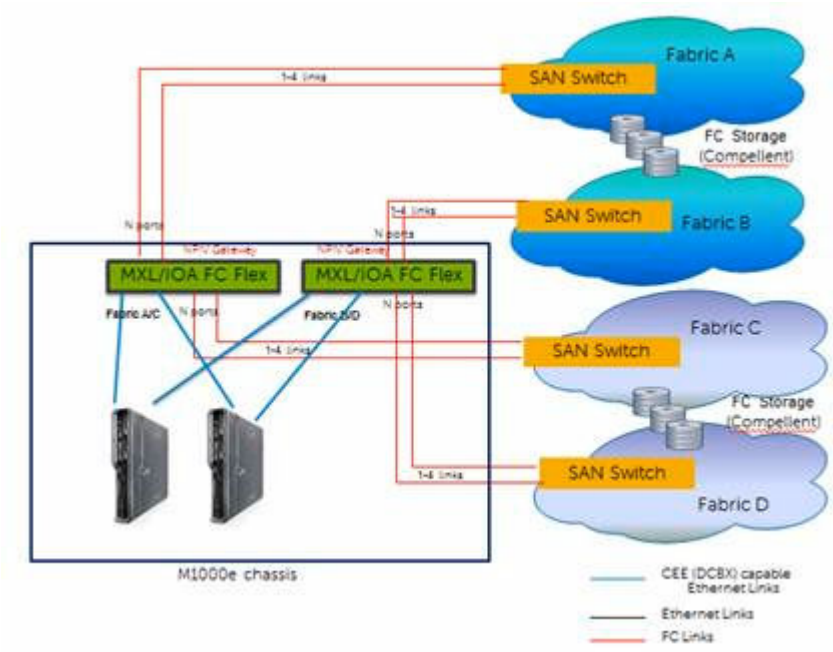


Figure 128. Case 1: Deployment Scenario of Configuring FC Flex IO Modules

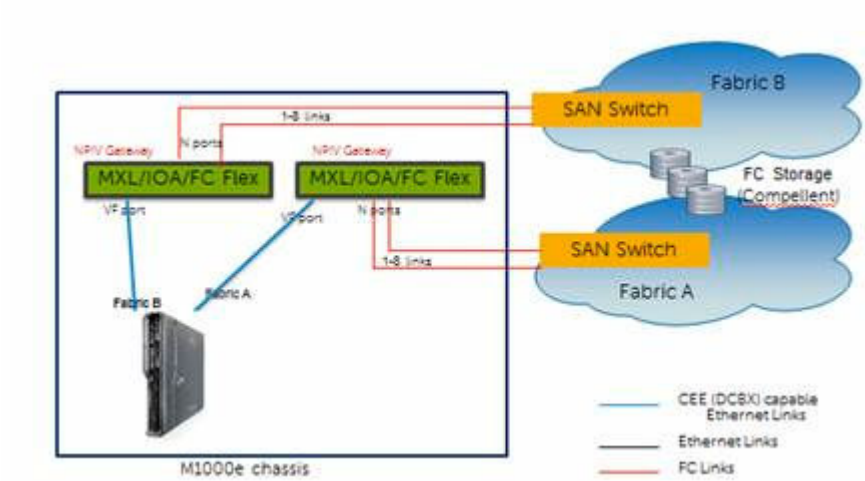


Figure 129. Case 2: Deployment Scenario of Configuring FC Flex IO Modules

Data Center Bridging (DCB)

Data center bridging (DCB) is supported on the FC Flex IO module installed in the FN IOM.

Ethernet Enhancements in Data Center Bridging

The following section describes DCB.

- The device supports the following DCB features:
  - Data center bridging exchange protocol (DCBx)

– Priority-based flow control (PFC)
– Enhanced transmission selection (ETS)

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 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:

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.
Storage traffic	Storage traffic based on Fibre Channel media uses the 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: In the Dell Networking OS version 8.3.12.0, only the PFC, ETS, and DCBx features are supported 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 3.

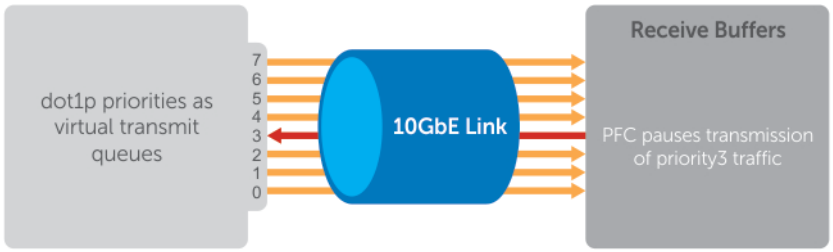


Figure 130. Priority-Based Flow Control

In the system, 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 delay constraints place an upper limit on the transmit time of a queue after receiving a message to pause a specified priority.
- By default, PFC is enabled on an interface with no dot1p priorities configured. You can configure the PFC priorities if the switch negotiates with a remote peer using DCBX.
- During DCBX negotiation with a remote peer:
  - If the negotiation succeeds and the port is in DCBX Willing mode to receive a peer configuration, PFC parameters from the peer are used to configured PFC priorities on the port. If you enable the link-level flow control mechanism on the interface, DCBX negotiation with a peer is not performed.
  - If the negotiation fails and PFC is enabled on the port, any user-configured PFC input policies are applied. If no PFC input policy has been previously applied, the PFC default setting is used (no priorities configured). If you do not enable PFC on an interface, you can enable the 802.3x link-level pause function. By default, the link-level pause is disabled.
- PFC supports buffering to receive data that continues to arrive on an interface while the remote system reacts to the PFC operation.

- PFC uses the DCB MIB IEEE802.1azd2.5 and the PFC MIB IEEE802.1bb-d2.2.

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.

Although you can configure strict-priority queue scheduling for a priority group, ETS introduces flexibility that allows the bandwidth allocated to each priority group to be dynamically managed according to the amount of LAN, storage, and server traffic in a flow. Unused bandwidth is dynamically allocated to prioritized priority groups. Traffic is queued according to its 802.1p priority assignment, while flexible bandwidth allocation and the configured queue-scheduling for a priority group is supported.

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.

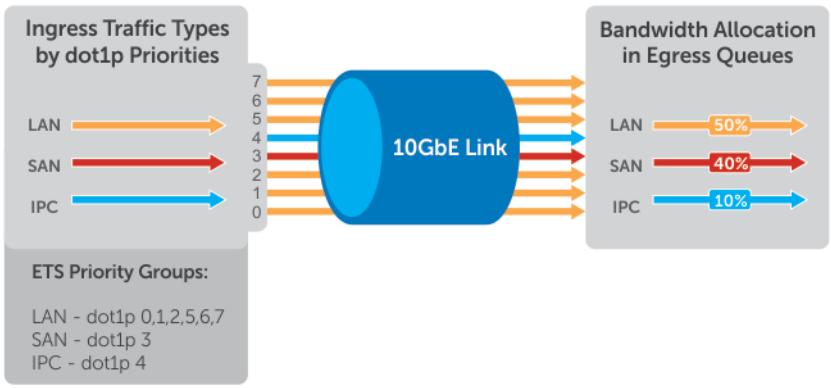


Figure 131. Enhanced Transmission Selection

The following table lists the traffic groupings ETS uses to select multiprotocol traffic for transmission.

Table 82. ETS Traffic Groupings

Traffic Groupings	Description
Priority group	A group of 802.1p priorities used for bandwidth allocation and queue scheduling. All 802.1p priority traffic in a group must have the same traffic handling requirements for latency and frame loss.
Group ID	A 4-bit identifier assigned to each priority group. The range is from 0 to 7.
Group bandwidth	Percentage of available bandwidth allocated to a priority group.

Traffic Groupings	Description
Group transmission selection algorithm (TSA)	Type of queue scheduling a priority group uses.

In the Dell 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
- Bandwidth allocated by the ETS algorithm is made available after strict-priority groups are serviced. If a priority group does not use its allocated bandwidth, the unused bandwidth is made available to other priority groups.
- For ETS traffic selection, an algorithm is applied to priority groups using:
  - Strict priority shaping
  - ETS shaping
- ETS uses the DCB MIB IEEE 802.1azd2.5.

Configuring 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.

1. Enter global configuration mode to create a DCB map or edit PFC and ETS settings.

   CONFIGURATION mode

   ```
   dcb-map name
   ```

2. 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, 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.

   DCB MAP mode

   ```
   priority-group group_num {bandwidth percentage | strict-priority} pfc {on | off}
   ```

   Example:

   ```
   priority-group 0 bandwidth 60 pfc off
   priority-group 1 bandwidth 20 pfc on
   priority-group 2 bandwidth 20 pfc on
   priority-group 4 strict-priority pfc off
   ```

   Repeat this step to configure PFC and ETS traffic handling for each priority group.

3. Specify the dot1p priority-to-priority group mapping for each priority. Priority-group range: 0 to 7.

   DCB MAP mode

   ```
   priority-pgid dot1p0_group_num dot1p1_group_num dot1p2_group_num
   dot1p3_group_num dot1p4_group_num dot1p5_group_num dot1p6_group_num
   dot1p7_group_num
   ```

   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.

Applying a DCB Map on a Port

To apply a DCB map to an Ethernet port, follow these steps:

1. Enter interface configuration mode on an Ethernet port.

   CONFIGURATION mode

   ```
   interfacetengigabitEthernet slot/port
   ```

2. Apply the DCB map on the Ethernet port to configure it with the PFC and ETS settings in the map.

   INTERFACE mode

   ```
   dcb-map name
   ```

   ```
   Dell# interface tengigabitEthernet 0/0
   Dell(config-if-te-0/0)# 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).

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.

1. Enter interface configuration mode on an Ethernet port.

   CONFIGURATION mode

   `interface tengigabitEthernet` *slot*`/port`

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.

   INTERFACE mode

   `pfc priority` *priority-range*

   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).

Configuring Lossless Queues

DCB also supports the manual configuration of lossless queues on an interface after you disable PFC mode in a DCB map and 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.

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 are supported on a port. If the number of lossless queues configured exceeds the maximum supported limit per port (two), an error message is displayed. You must re-configure 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.

1. Enter INTERFACE Configuration mode.

   CONFIGURATION mode

   `interface`**tengigabitEthernet** *slot*`/`*port*

2. Open a DCB map and enter DCB map configuration mode.

   INTERFACE mode

   `dcb-map` *name*

3. Disable PFC.

   DCB MAP mode

   ```
   no pfc mode on
   ```

4. Return to interface configuration mode.

   DCB MAP mode

   ```
   exit
   ```

5. Apply the DCB map, created to disable the PFC operation, on the interface.

   INTERFACE mode

   `dcb-map` {*name* | **default**}

6. Configure the port queues that still function as no-drop queues for lossless traffic.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.The maximum number of lossless queues globally supported on a port is 2.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.

   INTERFACE mode

   `pfc no-drop queues`*queue-range*

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. For more information, refer to Link Layer Discovery Protocol (LLDP). 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.

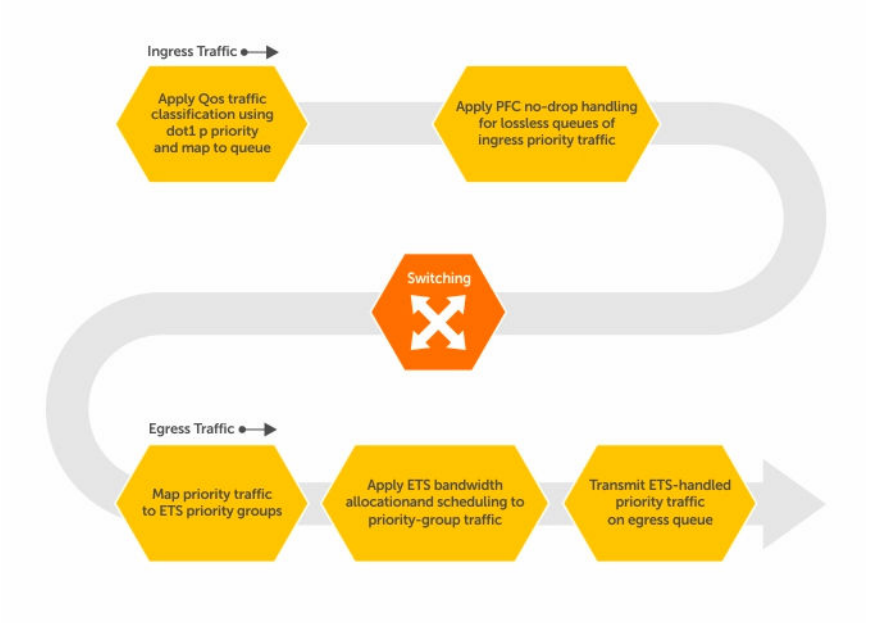


Figure 132. DCB PFC and ETS Traffic Handling

Enabling Data Center Bridging

Data center bridging is enabled by default on an FN IOM to support converged enhanced Ethernet (CEE) in a data center network.
A prerequisite for configuring DCB:

- 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.

On the FN IOM Switch, by default, DCB is enabled and MMU buffers are reserved to achieve no-drop traffic handling for PFC. Disabling DCB does not release the buffers reserved by default. To utilize reserved buffers for non-DCB applications, you have to explicitly release the buffers (Refer to Configuring the PFC Buffer in a Switch Stack).

To disable or re-enable DCB on a switch, enter the following commands.

1. Disable DCB.
   CONFIGURATION mode

```
no dcb enable
```

2. Re-enable DCB.
   CONFIGURATION mode

```
dcb enable
```

NOTE: Dell Networking OS Behavior: DCB is not supported if you enable link-level flow control on one or more interfaces.

After you disable DCB, if link-level flow control is not automatically enabled on an interface, to enable flow control, manually shut down the interface (the `shutdown` command) and re-enable it (the `no shutdown` command).

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 (refer to Honoring dot1p Values on Ingress Packets).
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 (refer to Policy-Based QoS Configurations).

NOTE: Dell Networking does not recommend mapping all ingress traffic to a single queue when using PFC and ETS. However, Dell 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 the system 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. For DCB ETS enabled interfaces, traffic destined to queue that is not mapped to any dot1p priority are dropped.

dot1p Value in the Incoming Frame	Egress Queue Assignment
0	0
1	0
2	0
3	1
4	2
5	3

dot1p Value in the Incoming Frame	Egress Queue Assignment
6	3
7	3

NOTE: If you reconfigure the global dot1p-queue mapping, an automatic re-election of the DCBX configuration source port is performed (refer to Configuration Source Election).

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.

To configure ETS and apply an ETS output policy to an interface, you must:

1. Create a Quality of Service (QoS) output policy with ETS scheduling and bandwidth allocation settings.
2. Create a priority group of 802.1p traffic classes.
3. Configure a DCB output policy in which you associate a priority group with a QoS ETS output policy.
4. Apply the DCB output policy to an interface.

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 previously configured ETS output policy on the port or to the default ETS settings if no ETS output policy was previously applied.
- 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.

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; Enabling Strict-Priority Queueing).

If you configure only the priority group in an ETS output policy or only the dot1p priority for strict-priority scheduling, the flow is handled with group strict priority.

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.

For more information, refer to Allocating Bandwidth to 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. Create a priority group for strict-priority scheduling.
   QoS OUTPUT POLICY mode

   ```
   Dell(conf-qos-policy-out)#scheduler strict
   ```

   NOTE: You can not use `scheduler strict` when bandwidth percentage is configured. It displays an error message.

   ```
   Dell(conf-qos-policy-out)#bandwidth-percentage 100
   Dell(conf-qos-policy-out)#scheduler strict
   % Error: Strict priority scheduler mode is not allowed when bandwidth-
   percentage is configured on qos-policy-output profile.
   Dell(conf-qos-policy-out)#scheduler strict  ?
   ```
5. Exit QoS Output Policy Configuration mode.
   QoS OUTPUT POLICY mode

```
Dell(conf-if-te-0/1)#exit
```

6. Enter INTERFACE Configuration mode.

   CONFIGURATION mode

```
Dell(conf-qos-policy-out)#int te 0/1
```

7. 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
```

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:

- Configuring Priority-Based Flow Control
- Configure Enhanced Transmission Selection
- Configuring FIP Snooping

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 receives its configuration from DCBX peers (ToR or FCF device). 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.

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 DCBX is used on an FN IOM Switch installed in a PowerEdge M1000e chassis in which servers are also installed.

The external 40GbE ports on the base module (ports 33 and 37) of two switches are used for uplinks configured as DCBx auto-upstream ports. The FN IOM switch 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 FN IOM switch, PFC and ETS use DCBx to exchange link-level configuration with DCBx peer devices.

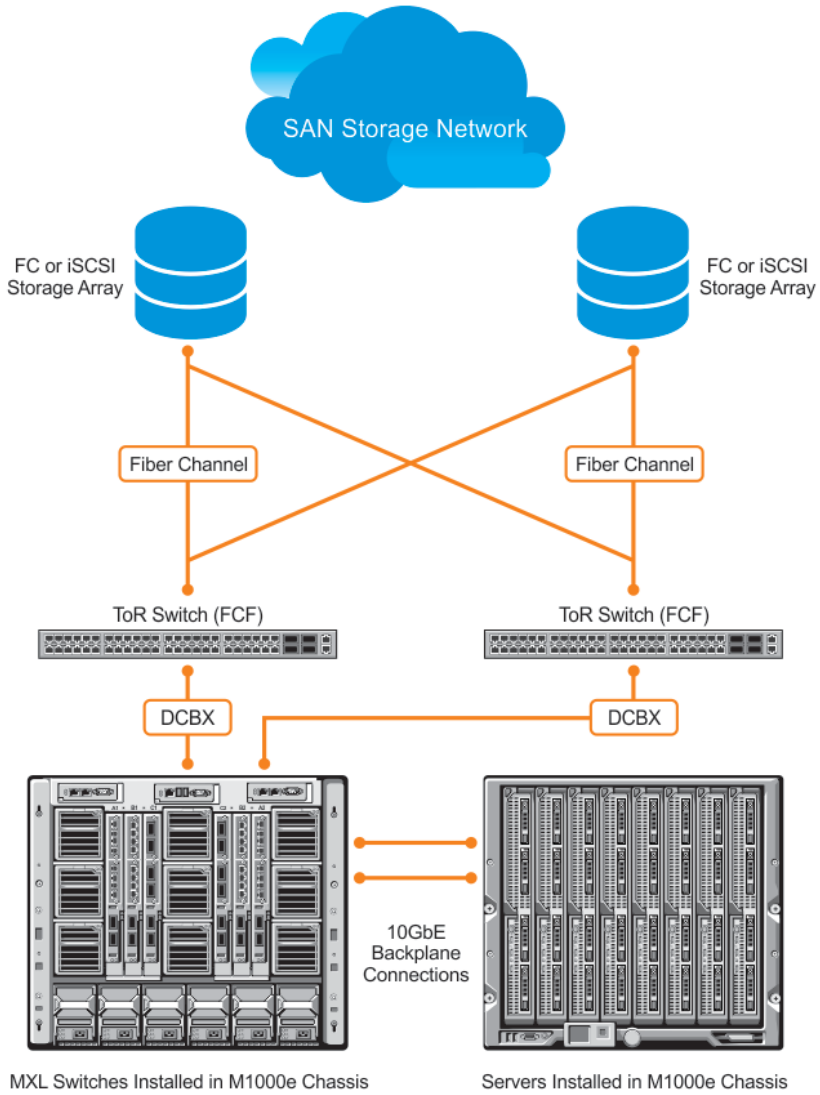


Figure 133. DCBx Sample Topology

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 an FN IOM switch 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 FCoE and iSCSI, refer to Fibre Channel over Ethernet and 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.
   PROTOCL 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`.

   For information about how to use FCoE and iSCSI, refer to Fibre Channel over Ethernet and iSCSI Optimization.
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 83. Displaying DCB Configurations

Command	Output
`show dot1p-queue mapping`	Displays the current 802.1p priority-queue mapping.
`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 slot/port 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 slot/port pfc statistics`	Displays counters for the PFC frames received and transmitted (by dot1p priority class) on an interface.
`show interface port-type slot/port 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.

Example of the `show dot1p-queue mapping` Command

Example of the `show dcb` Command

Example of the `show interfaces pfc summary` Command

Example of the `show interface pfc statistics` Command

Example of the `show interface ets summary` Command

Example of the `show interface ets detail` Command

Example of the `show stack-unit all stack-ports all pfc details` Command

Example of the `show stack-unit all stack-ports all ets details` Command

Example of the `show interface DCBx detail` Command

```
Dell(conf)# show dot1p-queue-mapping
Dot1p Priority: 0 1 2 3 4 5 6 7
Queue         : 0 0 0 1 2 3 3 3

Dell# 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)

Dell# show interfaces tengigabitethernet 0/49 pfc summary
Interface TenGigabitEthernet 0/49
    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

Dell# show interfaces tengigabitethernet 0/49 pfc detail
Interface TenGigabitEthernet 0/49
    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 84. `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

Fields	Description
	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 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.

Fields	Description
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

```
Dell#show interfaces tengigabitethernet 0/3 pfc statistics
Interface TenGigabitEthernet 0/3

Priority Rx XOFF Frames Rx Total Frames Tx Total Frames
-----------------------------------------------------
0        0              0               0
1        0              0               0
2        0              0               0
3        0              0               0
4        0              0               0
5        0              0               0
6        0              0               0
7        0              0               0

Dell(conf)# show interfaces te 0/0 ets summary
Interface TenGigabitEthernet 0/0
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.

```
Dell(conf)# show interfaces tengigabitethernet 0/0 ets detail
Interface TenGigabitEthernet 0/0
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
```

Table 85. `show interface ets detail` Command Description

Field	Description
Interface	Interface type with stack-unit and port number.
Max 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. When on, the scheduling and bandwidth allocation configured in an ETS output policy or received in a DCBx TLV from a peer can take effect on an interface.
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.
Local Parameters	ETS configuration on local port, including Admin mode (enabled when a valid TLV is received from a

Field	Description
	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.

```
Dell(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
```

```
Dell(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                           -           -

Dell(conf)# show interface tengigabitethernet 0/49 dcbx detail
Dell#show interface te 0/49 dcbx detail

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 0/49
  Remote Mac Address 00:00:00:00:00:11
     Port Role is Auto-Upstream
  DCBX Operational Status is Enabled
  Is Configuration Source? TRUE

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: 2
  Acknowledgment Number: 2
  Protocol State: In-Sync

Peer DCBX Status:
----------------
  DCBX Operational Version is 0
  DCBX Max Version Supported is 255
  Sequence Number: 2
```

```
Acknowledgment Number: 2
Total DCBX Frames transmitted 27
Total DCBX Frames received 6
Total DCBX Frame errors 0
Total DCBX Frames unrecognized 0
```

The following table describes the `show interface DCBx detail` command fields.

Table 86. `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.

Field	Description
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.
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.

PFC and ETS Configuration Examples

This section contains examples of how to configure and apply DCB input and output policies on an interface.

Using PFC and ETS to Manage Data Center Traffic

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.

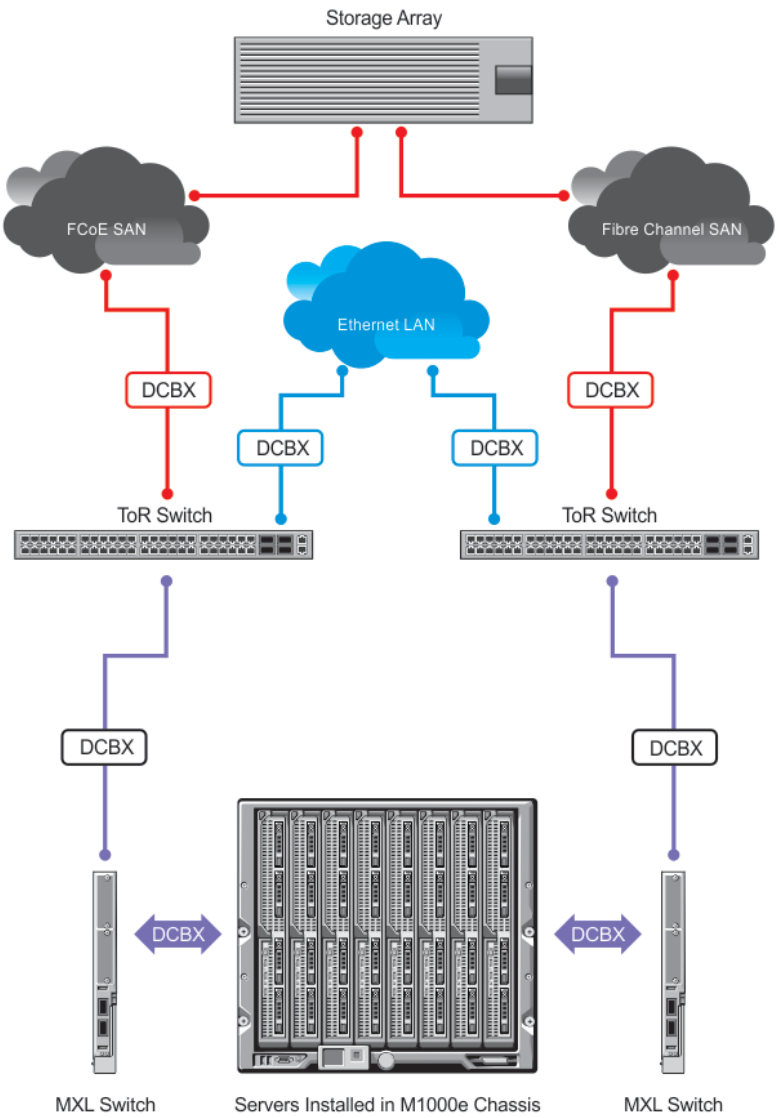


Figure 134. PFC and ETS Applied to LAN, IPC, and SAN Priority Traffic

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.

dot1p Value in Incoming Frame	Queue Assignment
0	0
1	0
2	0
3	1

dot1p Value in Incoming Frame	Queue Assignment
4	2
5	3
6	3
7	3

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.

Example of Configuring QoS Priority-Queue Assignment to Honor Dot1p Priorities

```
Dell(conf)# service-class dynamic dot1p
```

Using PFC and ETS to Manage Converged Ethernet Traffic in a Switch Stack

The following example shows how to apply the DCB PFC input policy (`ipc_san_lan`) and ETS output policy (`ets`) on all FN IOM switches in a switch stack.

This example references the PFC and ETS Configuration Examples section.

Example of Applying DCB PFC Input Policy and ETS Output Policy in a Switch Stack

```
dcb-map stack-unit all stack-ports all <dcb-map-name>
```

Interworking of DCB Map With DCB Buffer Threshold Settings

The `dcb-input` and `dcb-output` configuration commands are deprecated. You must use the `dcp-map` command to create a DCB map to configure priority flow control (PFC) and enhanced transmission selection (ETS) on Ethernet ports that support converged Ethernet traffic.

Configure the `dcb-buffer-threshold` command and its related parameters only on ports with either auto configuration or dcb-map configuration. This command is not supported on existing front-panel interfaces or stack ports that are configured with the `dcb-input` or `dcb-output` commands. Similarly, if the dcb-buffer-threshold configuration is present on a stack port or any interface, the dcb-input or dcb-ouput policies cannot be applied on those interfaces.

Example: When the dcb-buffer-threshold policy is applied on interfaces or stack ports with the dcb-input or dcb-output policies, the following error message is displayed:

```
%Error: dcb-buffer-threshold not supported on interfaces with deprecated
commands
```

Example: When the dcb-input or dcb-output policy is configured on interfaces or stack ports with the dcb-buffer threshold policy, the following error message is displayed:

```
%Error: Deprecated command is not supported on interfaces with dcb-buffer-
threshold configured
```

You must not modify the service-class dot1p mappings when any buffer-threshold-policy is configured on the system.

```
Dell(conf)#service-class dot1p-mapping dot1p0 3

% Error: PFC buffer-threshold policies conflict with dot1p mappings. Please
remove all dcb-buffer-threshold policies to change mappings.
```

The `show dcb` command has been enhanced to display the following additional buffer-related information:

```
 Dell(conf)#do show dcb
 dcb Status :  Enabled
 PFC Queue Count :  2 --Indicate the PFC queue configured.
 Total buffer (lossy + lossless)(in KB): 7787--Total buffer space for lossy and
lossless queues
 PFC total buffer (in KB): 6526 --Indicates the total buffer (configured or
default)
 PFC shared buffer (in KB): 832--Indicates the shared buffer (Configured or
default)
 PFC available buffer ( in KB): 5694--Indicates remaining available buffers for
PFC that are free to be allocated
```

Fibre Channel over Ethernet for FC Flex IO Modules

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.

The Fibre Channel (FC) Flex IO module is supported on Dell Networking Operating System (OS) FN IOM. The FN IOM switch installed with the FC Flex IO module functions as a top-of-rack edge switch that supports converged enhanced Ethernet (CEE) traffic — Fibre channel over Ethernet (FCoE) for storage,

Interprocess Communication (IPC) for servers, and Ethernet local area network (LAN) (IP cloud) for data — as well as FC links to one or more storage area network (SAN) fabrics.

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.

All of the commands that are supported for FCoE on the FN IOM apply to the FC Flex IO modules. Similarly, all of the configuration procedures and the settings that are applicable for FCoE on the FN IOM are valid for the FC Flex IO modules.

NPIV Proxy Gateway for FC Flex IO Modules

The N-port identifier virtualization (NPIV) Proxy Gateway (NPG) feature provides FCoE-FC bridging capability on the FN IOM with the FC Flex IO module switch, allowing server CNAs to communicate with SAN fabrics over the FN IOM with the FC Flex IO module.

To configure the FN IOM with the FC Flex IO module to operate as an NPIV proxy gateway, use the following commands:

NPIV Proxy Gateway Configuration on FC Flex IO Modules

The Fibre Channel (FC) Flex IO module is supported on the FN IOM. The FN IOM switch, installed with the FC Flex IO module, function as a top-of-rack edge switch that supports Converged Enhanced Ethernet (CEE) traffic — Fibre Channel over Ethernet (FCoE) for storage, Interprocess Communication (IPC) for servers, and Ethernet local area network (LAN) (IP cloud) for data — as well as FC links to one or more storage area network (SAN) fabrics.

The N-port identifier virtualization (NPIV) proxy gateway (NPG) provides FCoE-FC bridging capability on the FN IOM with the FC Flex IO module.

This chapter describes how to configure and use an NPIV proxy gateway on the FN IOM with the FC Flex IO module in a (SAN.

NPIV Proxy Gateway Operations and Capabilities

Benefits of an NPIV Proxy Gateway

The FN IOM with the FC Flex IO module functions as a top-of-rack edge switch that supports Converged Enhanced Ethernet (CEE) traffic — FCoE for storage, Interprocess Communication (IPC) for servers, and Ethernet LAN (IP cloud) for data — as well as Fibre Channel (FC) links to one or more SAN fabrics.

Using an NPIV proxy gateway (NPG) helps resolve the following problems in a storage area network:

- Fibre Channel storage networks typically consist of servers connected to edge switches, which are connected to SAN core switches. As the SAN grows, it is necessary to add more ports and SAN switches. This results in an increase in the required domain IDs, which may surpass the upper limit of 239 domain IDs supported in the SAN network. An NPG avoids the need for additional domain IDs because it is deployed outside the SAN and uses the domain IDs of core switches in its FCoE links.
- With the introduction of 10GbE links, FCoE is being implemented for server connections to optimize performance. However, a SAN traditionally uses Fibre Channel to transmit storage traffic. FCoE servers require an efficient and scalable bridging feature to access FC storage arrays, which an NPG provides.

NPIV Proxy Gateway Operation

Consider a sample scenario of NPG operation. An M1000e chassis configured as an NPG does not join a SAN fabric, but functions as an FCoE-FC bridge that forwards storage traffic between servers and core SAN switches. The core switches forward SAN traffic to and from FC storage arrays.

An M1000e chassis FC port is configured as an N (node) port that logs in to an F (fabric) port on the upstream FC core switch and creates a channel for N-port identifier virtualization. NPIV allows multiple N-port fabric logins at the same time on a single, physical Fibre Channel link.

Converged Network Adapter (CNA) ports on servers connect to the M1000e chassis Ten-Gigabit Ethernet ports and log in to an upstream FC core switch through the FC Flex IO module N port. Server fabric login (FLOGI) requests are converted into fabric discovery (FDISC) requests before being forwarded by the FC Flex IO module to the FC core switch.

Servers use CNA ports to connect over FCoE to an Ethernet port in ENode mode on the NPIV proxy gateway. FCoE transit with FIP snooping is automatically enabled and configured on the M1000e gateway to prevent unauthorized access and data transmission to the SAN network (see FCoE Transit). FIP is used by server CNAs to discover an FCoE switch operating as an FCoE forwarder (FCF).

The NPIV proxy gateway aggregates multiple locally connected server CNA ports into one or more upstream N port links, conserving the number of ports required on an upstream FC core switch while providing an FCoE-to-FC bridging functionality. The upstream N ports on an M1000e can connect to the same or multiple fabrics.

Using an FCoE map applied to downstream (server-facing) Ethernet ports and upstream (fabric-facing) FC ports, you can configure the association between a SAN fabric and the FCoE VLAN that connects servers over the NPIV proxy gateway to FC switches in the fabric. An FCoE map virtualizes the upstream SAN fabric as an FCF to downstream CNA ports on FCoE-enabled servers as follows:

- As soon as an FC N port comes online (`no shutdown` command), the NPG starts sending FIP multicast advertisements, which contain the fabric name derived from the 64-bit worldwide name (WWN) of the principal SAN switch. (The principal switch in a fabric is the FC switch with the lowest domain ID.)
- When you apply the FCoE map to a server-facing Ethernet port in ENode mode, ACLs are automatically configured to allow only FCoE traffic from servers that perform a successful FLOGI on the FC switch. All other traffic on the VLAN is denied.

You can specify one or more upstream N ports in an FCoE map. The FCoE map also contains the VLAN ID of the dedicated VLAN used to transmit FCoE traffic between the SAN fabric and servers.

NPIV Proxy Gateway: Protocol Services

The FN IOM with the FC Flex IO module NPG provides the following protocol services:

- Fibre Channel service to create N ports and log in to an upstream FC switch.
- FCoE service to perform:
- Virtualization of FC N ports on an NPG so that they appear as FCoE FCFs to downstream servers.
- NPIV service to perform the association and aggregation of FCoE servers to upstream F ports on core switches (through N ports on the NPG). Conversion of server FLOGIs and FDISCs, which are received

over FN IOM with the FC Flex IO module ENode ports, are converted into FDISCs addressed to the upstream F ports on core switches.

NPIV Proxy Gateway Functionality

The FN IOM with the FC Flex IO module NPG provides the following functionality in a storage area network:

- FIP Snooping bridge that provides security for FCoE traffic using ACLs (see FCoE Transit chapter).
- FCoE gateway that provides FCoE-to-FC bridging. N-port virtualization using FCoE maps exposes upstream F ports as FCF ports to downstream server-facing ENode ports on the NPG (see FCoE Maps).

NPIV Proxy Gateway: Terms and Definitions

The following table describes the terms used in an NPG configuration on the FN IOM with the FC Flex IO module.

Table 87. FN IOM with the FC Flex IO module NPIV Proxy Gateway: Terms and Definitions

Term	Description
FC port	Fibre Channel port on the FN IOM with the FC Flex IO module FC module that operates in autosensing, 2, 4, or 8-Gigabit mode. On an NPIV proxy gateway, an FC port can be used as a downlink for a server connection and an uplink for a fabric connection.
F port	Port mode of an FC port connected to an end node (N) port on the FN IOM with the FC Flex IO module NPIV proxy gateway.
N port	Port mode of the FN IOM with the FC Flex IO module FC port that connects to an F port on an FC switch in a SAN fabric. On the FN IOM with the FC Flex IO module NPIV proxy gateway, an N port also functions as a proxy for multiple server CNA-port connections
ENode port	Port mode of a server-facing FN IOM with the FC Flex IO module Ethernet port that provides access to FCF functionality on a fabric.
CNA port	N-port functionality on an FCoE-enabled server port. A converged network adapter (CNA) can use one or more Ethernet ports. CNAs can encapsulate Fibre Channel frames in Ethernet for FCoE transport and de-encapsulate Fibre Channel frames from FCoE to native Fibre Channel.
DCB map	Template used to configure DCB parameters, including priority-based flow control (PFC) and enhanced transmission selection (ETS), on CEE ports.
Fibre Channel fabric	Network of Fibre Channel devices and storage arrays that interoperate and communicate.
FCF	Fibre Channel forwarder: FCoE-enabled switch that can forward FC traffic to both downstream FCoE and upstream FC devices. An NPIV proxy gateway functions as an FCF to export upstream F port configurations to downstream server CNA ports.

Term	Description
FC-MAP	FCoE MAC-address prefix — The unique 24-bit MAC address prefix in FCoE packets used to generate a fabric-provided MAC address (FPMA). The FPMA is required to send FCoE packets from a server to a SAN fabric.
FCoE map	Template used to configure FCoE and FC parameters on Ethernet and FC ports in a converged fabric.
FCoE VLAN	VLAN dedicated to carrying only FCoE traffic between server CNA ports and a SAN fabric. (FCoE traffic must travel in a VLAN.) When you apply an FCoE map on a port, FCoE is enabled on the port. All non-FCoE traffic is dropped on an FCoE VLAN.
FIP	FCoE Initialization Protocol: Layer 2 protocol for endpoint discovery, fabric login, and fabric association. FIP is used by server CNAs to discover an upstream FCoE switch operating as an FCF. FIP keepalive messages maintain the connection between an FCoE initiator and an FCF.
NPIV	N-port identifier virtualization: The capability to map multiple FCoE links from downstream ports to a single upstream FC link.
principal switch	The switch in a fabric with the lowest domain number. The principal switch accesses the master name database and the zone/zone set database.

DCB Maps

A Data Center Bridging (DCB) map is used to configure DCB functionality, such as PFC and ETS, on FN IOM with the FC Flex IO module Ethernet ports that support CEE traffic and are DCBx-enabled, by default. For more information, on PFC and ETS, see Data Center Bridging (DCB).

By default, no PFC and ETS settings in a DCB map are applied to FN IOM with the FC Flex IO module Ethernet ports when they are enabled. On the FN IOM with the FC Flex IO module NPG, you must configure PFC and ETS parameters in a DCB map and then apply the map to server-facing Ethernet ports (see the “Creating a DCB map" section).

FCoE Maps

An FCoE map is used to identify the SAN fabric to which FCoE storage traffic is sent. Using an FCoE map, the FN IOM with the FC Flex IO module NPG operates as an FCoE-FC bridge between an FC SAN and FCoE network by providing FCoE-enabled servers and switches with the necessary parameters to log in to a SAN fabric.

An FCoE map applies the following parameters on server-facing Ethernet and fabric-facing FC ports on the FN IOM with the FC Flex IO module:

- The dedicated FCoE VLAN used to transport FCoE storage traffic.
- The FC-MAP value used to generate a fabric-provided MAC address.
- The association between the FCoE VLAN ID and FC fabric ID where the desired storage arrays are installed. Each Fibre Channel fabric serves as an isolated SAN topology within the same physical network.
- The priority used by a server to select an upstream FCoE forwarder (FCF priority)

- FIP keepalive (FKA) advertisement timeout

NOTE:
In each FCoE map, the fabric ID, FC-MAP value, and FCoE VLAN must be unique. Use one FCoE map to access one SAN fabric. You cannot use the same FCoE map to access different fabrics.

When you configure the FN IOM with the FC Flex IO module as an NPG, FCoE transit with FIP snooping is automatically enabled and configured using the parameters in the FCoE map applied to server-facing Ethernet and fabric-facing FC interfaces (see FIP Snooping on an NPIV Proxy Gateway).

After you apply an FCoE map on an FC port, when you enable the port (no shutdown), the NPG starts sending FIP multicast advertisements on behalf of the FC port to downstream servers in order to advertise the availability of a new FCF port on the FCoE VLAN. The FIP advertisement also contains a keepalive message to maintain connectivity between a SAN fabric and downstream servers.

Configuring an NPIV Proxy Gateway

Prerequisite: Before you configure an NPIV proxy gateway (NPG) with the FC Flex IO module on the FN IOM, ensure that the following features are enabled.

- DCB is enabled by default with the FC Flex IO module on the FN IOM.
- Autonegotiated DCBx is enabled for converged traffic by default with the FC Flex IO module Ethernet ports on all FN IOM.
- FCoE transit with FIP snooping is automatically enabled when you configure Fibre Channel with the FC Flex IO module on the FN IOM.

To configure an NPG operation with the FC Flex IO module on the FN IOM, follow these general configuration steps:

1. Enabling Fibre Channel Capability on the Switch
2. Creating a DCB map
3. Applying a DCB map on server-facing Ethernet ports
4. Creating an FCoE VLAN
5. Creating an FCoE map
6. Applying an FCoE map on server-facing Ethernet ports
7. Applying an FCoE Map on fabric-facing FC ports

Enabling Fibre Channel Capability on the Switch

Enable the FC Flex IO module on the FN IOM that you want to configure as an NPG for the Fibre Channel protocol. When you enable Fibre Channel capability, FCoE transit with FIP snooping is automatically enabled on all VLANs on the switch, using the default FCoE transit settings.

1. Enable the FN IOM with the FC Flex IO module for the Fibre Channel protocol.

   CONFIGURATION mode

   ```
   feature fc
   ```

Creating a DCB Map

Configure the priority-based flow control (PFC) and enhanced traffic selection (ETS) settings in a DCB map before you apply them on downstream server-facing ports on the FN IOM with the FC Flex IO module.

1. Create a DCB map to specify PFC and ETS settings for groups of dot1p priorities.

   CONFIGURATION mode

   `dcb-map` *name*

2. Configure the PFC setting (on or off) and the ETS bandwidth percentage allocated to traffic in each priority group. Configure whether the priority group traffic should be handled with strict-priority scheduling. The sum of all allocated bandwidth percentages must be 100 percent. Strict-priority traffic is serviced first. Afterward, 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.

   Restriction: You can enable PFC on a maximum of two priority queues.

   Repeat this step to configure PFC and ETS traffic handling for each priority group, for example: `priority-group 0 bandwidth 60 pfc off priority-group 1 bandwidth 20 pfc onpriority-group 2 bandwidth 20 pfc on priority-group 4 strict-priority pfc off`

   DCB MAP mode

   `priority-group` *group_num* {**bandwidth** *percentage* | **strict-priority**} **pfc** {**on** | **off**}

3. Specify the priority group ID number to handle VLAN traffic for each dot1p class-of-service: 0 through 7. Leave a space between each priority group number. For example, `priority-pgid 0 0 0 1 2 4 4 4` where dot1p priorities 0, 1, and 2 are mapped to priority group 0; dot1p priority 3 is mapped to priority group 1; dot1p priority 4 is mapped to priority group 2; dot1p priorities 5, 6, and 7 are mapped to priority group 4.

   All priorities that map to the same egress queue must be in the same priority group.

   DCB MAP mode

   `priority-pgid` *dot1p0_group_num dot1p1_group_num dot1p2_group_num dot1p3_group_num dot1p4_group_num dot1p5_group_num dot1p6_group_num dot1p7_group_num*

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 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, 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-to-priority group mapping.

If you delete the dot1p priority-to-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 the network operation.

Applying a DCB Map on Server-facing Ethernet Ports

You can apply a DCB map only on a physical Ethernet interface and can apply only one DCB map per interface.

1. Enter CONFIGURATION mode on a server-facing port or port channel to apply a DCB map.

   You cannot apply a DCB map on a port channel. However, you can apply a DCB map on the ports that are members of the port channel.

   CONFIGURATION mode

   `interface` **tengigabitEthernet** *slot/port*

2. Apply the DCB map on an Ethernet port or port channel. The port is configured with the PFC and ETS settings in the DCB map.

   Repeat this step to apply a DCB map to more than one port or port channel.

   INTERFACE mode

   `dcb-map` *name*

```
Dell# interface tengigabitEthernet 0/0
Dell(config-if-te-0/0)# dcb-map SAN_DCB1
```

Creating an FCoE VLAN

Create a dedicated VLAN to send and receive Fibre Channel traffic over FCoE links between servers and a fabric over an NPG. The NPG receives FCoE traffic and forwards decapsulated FC frames over FC links to SAN switches in a specified fabric.

1. Create the dedicated VLAN for FCoE traffic. Range: 2–4094.

   VLAN 1002 is commonly used to transmit FCoE traffic.

   CONFIGURATION mode

   `interface vlan` *vlan-id*

When you apply an FCoE map to an Ethernet port (Applying an FCoE map on server-facing Ethernet ports), the port is automatically configured as a tagged member of the FCoE VLAN.

Creating an FCoE Map

An FCoE map consists of:

- An association between the dedicated VLAN, used to carry FCoE traffic, and the SAN fabric where the storage arrays are installed. Use a separate FCoE VLAN for each fabric to which the FCoE traffic is forwarded. Any non-FCoE traffic sent on a dedicated FCoE VLAN is dropped.
- The FC-MAP value, used to generate the fabric-provided MAC address (FPMA). The FPMA is used by servers to transmit FCoE traffic to the fabric. You can associate an FC-MAP with only one FCoE VLAN and conversely, associate an FCoE VLAN with only one FC-MAP.
- FCF priority, the priority used by a server CNA to select an upstream FCoE forwarder (FCF)
- FIP keepalive (FKA) advertisement timeout

The values for the FCoE VLAN, fabric ID and FC-MAP must be unique. Apply an FCoE map on downstream server-facing Ethernet ports and upstream fabric-facing Fibre Channel ports.

1. Create an FCoE map that contains parameters used in the communication between servers and a SAN fabric.

   CONFIGURATION mode

   `fcoe-map` *`map-name`*

2. Configure the association between the dedicated VLAN and the fabric where the desired storage arrays are installed. The fabric and VLAN ID numbers must be the same. Fabric and VLAN ID range: 2–4094.

   For example: `fabric id 10 vlan 10`

   FCoE MAP mode

   `fabric-id` *`fabric-num`* `vlan` *`vlan-id`*

3. Add a text description of the settings in the FCoE map. Maximum: 32 characters.

   FCoE MAP mode

   `description` *`text`*

4. Specify the FC-MAP value used to generate a fabric-provided MAC address, which is required to send FCoE traffic from a server on the FCoE VLAN to the FC fabric specified in Step 2. Enter a unique MAC address prefix as the FC-MAP value for each fabric. Range: 0EFC00–0EFCFF. Default: None.

   FCoE MAP mode

   `fc-map` *`fc-map-value`*

5. Configure the priority used by a server CNA to select the FCF for a fabric login (FLOGI). Range: 1–255. Default: 128.

   FCoE MAP mode

   `fcf-priority` *`priority`*

6. Enable the monitoring FIP keepalive messages (if it is disabled) to detect if other FCoE devices are reachable. Default: FIP keepalive monitoring is enabled.

   FCoE MAP mode

   ```
   keepalive
   ```

7. Configure the time interval (in seconds) used to transmit FIP keepalive advertisements. Range: 8–90 seconds. Default: 8 seconds.

   FCoE MAP mode

   ```
   fka-adv-period seconds
   ```

Applying an FCoE Map on Server-facing Ethernet Ports

You can apply multiple FCoE maps on an Ethernet port or port channel. When you apply an FCoE map on a server-facing port or port channel:

- The port is configured to operate in hybrid mode (accept both tagged and untagged VLAN frames).
- The associated FCoE VLAN is enabled on the port or port channel.

When you enable a server-facing Ethernet port, the servers respond to the FIP advertisements by performing FLOGIs on upstream virtualized FCF ports. The NPG forwards the FLOGIs as fabric discovery (FDISC) messages to a SAN switch.

1. Configure a server-facing Ethernet port or port channel with an FCoE map.

   CONFIGURATION mode

   ```
   interface {tengigabitEthernet slot/port | port-channel num}
   ```

2. Apply the FCoE/FC configuration in an FCoE map on the Ethernet port. Repeat this step to apply an FCoE map to more than one port.

   INTERFACE or INTERFACE PORT_CHANNEL mode

   ```
   fcoe-map map-name
   ```

   ```
   Dell# interface tengigabitEthernet 0/0
   Dell(config-if-te-0/0)# fcoe-map SAN_FABRIC_A
   Dell# interface port-channel 3
   Dell(config-if-te-0/0)# dcb-map SAN_DCB1
   Dell(config-if-po-3)# fcoe-map SAN_FABRIC_A
   ```

3. Enable the port for FCoE transmission using the map settings.

   INTERFACE mode

   ```
   no shutdown
   ```

Applying an FCoE Map on Fabric-facing FC Ports

The FN IOM, with the FC Flex IO module FC ports, are configured by default to operate in N port mode to connect to an F port on an FC switch in a fabric. You can apply only one FCoE map on an FC port.

When you apply an FCoE map on a fabric-facing FC port, the FC port becomes part of the FCoE fabric, whose settings in the FCoE map are configured on the port and exported to downstream server CNA ports.

Each FN IOM, with the FC Flex IO module FC port, is associated with an Ethernet MAC address (FCF MAC address). When you enable a fabric-facing FC port, the FCoE map applied to the port starts sending FIP multicast advertisements using the parameters in the FCoE map over server-facing Ethernet ports. A server sees the FC port, with its applied FCoE map, as an FCF port.

1. Configure a fabric-facing FC port.

   CONFIGURATION mode

   ```
   interface fibrechannel slot/port
   ```

2. Apply the FCoE and FC fabric configurations in an FCoE map to the port. Repeat this step to apply an FCoE map to more than one FC port.

   INTERFACE FIBRE_CHANNEL mode

   ```
   fabric map-name
   ```

   ```
   Dell# interface fi 0/9
   Dell(config-if-fc-0/9)# fabric SAN_FABRIC_A
   ```

3. Enable the port for FC transmission.

   INTERFACE FIBRE_CHANNEL mode

   ```
   no shutdown
   ```

Important Points to Remember

You can apply a DCB or FCoE map to a range of Ethernet or Fibre Channel interfaces by using the`interface range` command; for example:

```
Dell(config)# interface range tengigabitEthernet 1/12 - 23 , tengigabitEthernet
2/24 - 35

Dell(config)# interface range fibrechannel 0/0 - 3 , fibrechannel 0/8 - 11
```

Enter the keywords `interface range` followed by an interface type and port range. A port range must contain spaces before and after the dash. Separate each interface type and port range with a space, comma, and space as shown in the preceding examples.

Sample Configuration

1. Configure a DCB map with PFC and ETS settings:

   ```
   Dell(config)# dcb-map SAN_DCB_MAP
   Dell(config-dcbx-name)# priority-group 0 bandwidth 60 pfc off

   Dell(config-dcbx-name)# priority-group 1 bandwidth 20 pfc on

   Dell(config-dcbx-name)# priority-group 2 bandwidth 20 pfc on
   ```

```
Dell(config-dcbx-name)# priority-group 4 strict-priority pfc off

Dell(conf-dcbx-name)# priority-pgid 0 0 0 1 2 4 4 4
```

2. Apply the DCB map on a downstream (server-facing) Ethernet port:

```
Dell(config)# interface tengigabitethernet 1/0

Dell(config-if-te-0/0)#dcb-map SAN_DCB_MAP
```

3. Create the dedicated VLAN to be used for FCoE traffic:

```
Dell(conf)#interface vlan 1002
```

4. Configure an FCoE map to be applied on downstream (server-facing) Ethernet and upstream (core-facing) FC ports:

```
Dell(config)# fcoe-map SAN_FABRIC_A

Dell(config-fcoe-name)# fabric-id 1002 vlan 1002

Dell(config-fcoe-name)# description "SAN_FABRIC_A"

Dell(config-fcoe-name)# fc-map 0efc00

Dell(config-fcoe-name)# keepalive

Dell(config-fcoe-name)# fcf-priority 128

Dell(config-fcoe-name)# fka-adv-period 8
```

5. Enable an upstream FC port:

```
Dell(config)# interface fibrechannel 0/0

Dell(config-if-fc-0)# no shutdown
```

6. Enable a downstream Ethernet port:

```
Dell(config)#interface tengigabitEthernet 0/0

Dell(conf-if-te-0)# no shutdown
```

Displaying NPIV Proxy Gateway Information

To display information on the NPG operation, use the show commands in the following table

Table 88. Displaying NPIV Proxy Gateway Information

Command	Description
`show interfaces status`	Displays the operational status of Ethernet and Fibre Channel interfaces on the FN IOM with the FC Flex IO module, NPG.

Command	Description
	NOTE: Although the show interface status command displays the Fiber Channel (FC) interfaces with the abbreviated label of 'Fc' in the output, if you attempt to specify a FC interface by using the `interface fc` command in the CLI interface, an error message is displayed. You must configure FC interfaces by using the `interface fi` command in CONFIGURATION mode.
`show fcoe-map [`**brief** `| map-name]`	Displays the Fibre Channel and FCoE configuration parameters in FCoE maps. Enter the **brief** keyword to display an overview of currently configured FCoE maps. Enter the name of an FCoE map to display the FC and FCoE parameters configured in the map to be applied on FN IOM with the FC Flex IO module Ethernet (FCoE) and FC ports.
`show qos dcb-map map-name`	Displays configuration parameters in a specified DCB map.
`show npiv devices [`**brief**`]`	Displays information on FCoE and FC devices currently logged in to the NPG.
`show fc switch`	Displays the FC mode of operation and worldwide node (WWN) name of the FN IOM with the FC Flex IO module.

show interfaces status Command Example

```
Dell# show interfaces status
Port     Description  Status Speed     Duplex Vlan
Fc 0/0                Up     8000 Mbit Auto   --
Fc 0/1                Down   Auto      Auto   --
Fc 0/2                Down   Auto      Auto   --
Fc 0/3                Down   Auto      Auto   --
Fc 0/4                Down   Auto      Auto   --
Fc 0/5                Down   Auto      Auto   --
Fc 0/6                Down   Auto      Auto   --
Fc 0/7                Down   Auto      Auto   --
Fc 0/8                Down   Auto      Auto   --
Fc 0/9                Down   Auto      Auto   --
Fc 0/10               Down   Auto      Auto   --
Fc 0/11               Down   Auto      Auto   --
Te 1/12               Down   Auto      Auto   --
Te 1/13               Down   Auto      Auto   --
Te 1/14               Down   Auto      Auto   --
Te 1/15               Down   Auto      Auto   --
Te 1/16               Down   Auto      Auto   --
Te 1/17               Down   Auto      Auto   --
Te 1/18               Down   Auto      Auto   --
Te 1/19               Up     10000 Mbit Full    --
Te 1/20               Down   Auto      Auto   --
Te 1/21               Down   Auto      Auto   --
```

Table 89. show interfaces status Field Descriptions

Field	Description
Port	Server-facing 10GbE Ethernet (Te), 40GbE Ethernet (Fo), or fabric-facing Fibre Channel (Fc) port with *slot*/*port* information.
Description	Text description of port.
Status	Operational status of port: Ethernet ports - up (transmitting FCoE and LAN storage traffic) or down (not transmitting traffic). Fibre Channel ports - up (link is up and transmitting FC traffic) or down (link is down and not transmitting FC traffic), link-wait (link is up and waiting for FLOGI to complete on peer SW port), or removed (port has been shut down).
Speed	Transmission speed (in Megabits per second) of Ethernet and FC iports, including auto-negotiated speed (Auto).
Duplex	Data transmission mode: Full (allows communication in both directions at the same time), Half (allows communication in both directions but not at the same time), Auto (auto-negotiated transmission).
VLAN	VLAN IDs of the VLANs in which the port is a member.

show fcoe-map Command Examples

```
Dell# show fcoe-map brief
Fabric-Name   Fabric-Id         Vlan-Id  FC-MAP    FCF-Priority  Config-
State      Oper-State
fid_1003        1003            1003     0efc03    128
ACTIVE            UP
fid_1004        1004            1004     0efc04    128
ACTIVE            DOWN


Dell# show fcoe-map fid_1003

Fabric Name         fid_1003
Fabric Id           1003
Vlan Id             1003
Vlan priority       3
FC-MAP              0efc03
FKA-ADV-Period      8
Fcf Priority        128
Config-State        ACTIVE
Oper-State          UP
Members
Fc 0/0
Te 0/14 Te 0/16
```

Table 90. show fcoe-map Field Descriptions

Field	Description
Fabric-Name	Name of a SAN fabric.
Fabric ID	The ID number of the SAN fabric to which FC traffic is forwarded.
VLAN ID	The dedicated VLAN used to transport FCoE storage traffic between servers and a fabric over the NPG. The configured VLAN ID must be the same as the fabric ID.
VLAN priority	FCoE traffic uses VLAN priority 3. (This setting is not user-configurable.)
FC-MAP	FCoE MAC-address prefix value - The unique 24-bit MAC address prefix that identifies a fabric.
FKA-ADV-period	Time interval (in seconds) used to transmit FIP keepalive advertisements.
FCF Priority	The priority used by a server to select an upstream FCoE forwarder.
Config-State	Indicates whether the configured FCoE and FC parameters in the FCoE map are valid: Active (all mandatory FCoE and FC parameters are correctly configured) or Incomplete (either the FC-MAP value, fabric ID, or VLAN ID are not correctly configured).
Oper-State	Operational status of the link to the fabric: up (link is up and transmitting FC traffic), down (link is down and not transmitting FC traffic), link-wait (link is up and waiting for FLOGI to complete on peer FC port), or removed (port has been shut down).
Members	FN IOM with the FC Flex IO module Ethernet and FC ports, which are members of the dedicated FCoE VLAN that carries storage traffic to the specified fabric.

show qos dcb-map Command Examples

```
Dell# show qos dcb-map dcbmap2

State   :Complete
PfcMode:ON
--------------------
PG:0 TSA:ETS  BW:50  PFC:OFF
Priorities:0 1 2 4 5 6 7


PG:1 TSA:ETS  BW:50  PFC:ON
Priorities:3
```

Table 91. show qos dcb-map Field Descriptions

Field	Description
State	Complete: All mandatory DCB parameters are correctly configured. In progress: The DCB map configuration is not complete. Some mandatory parameters are not configured.
PFC Mode	PFC configuration in the DCB map: On (enabled) or Off.
PG	Priority group configured in the DCB map.
TSA	Transmission scheduling algorithm used in the DCB map: Enhanced Transmission Selection (ETS).
BW	Percentage of bandwidth allocated to the priority group.
PFC	PFC setting for the priority group: On (enabled) or Off.
Priorities	802.1p priorities configured in the priority group.

show npiv devices brief Command Example

```
Dell# show npiv devices brief

Total NPIV Devices = 2
-------------------------------------------------------------------------------
------------------------

ENode-Intf      ENode-WWPN           FCoE-Vlan  Fabric-Intf     Fabric-Map
LoginMethod     Status
-------------------------------------------------------------------------------
------------------------

Te 0/12      20:01:00:10:18:f1:94:20   1003      Fc 0/5
fid_1003           FLOGI     LOGGED_IN
Te 0/13      10:00:00:00:c9:d9:9c:cb   1003      Fc 0/0
fid_1003           FDISC     LOGGED_IN
```

Table 92. show npiv devices brief Field Descriptions

Field	Description
Total NPIV Devices	Number of downstream ENodes connected to a fabric over the FN IOM with the FC Flex IO module, NPG.
ENode-Intf	FN IOM with the FC Flex IO module Ethernet interface (*slot*/*port*) to which a server CNA is connected.
ENode-WWPN	Worldwide port name (WWPN) of a server CNA port.
FCoE-Vlan	VLAN ID of the dedicated VLAN used to transmit FCoE traffic to and from the fabric.
Fabric-Intf	Fabric-facing Fibre Channel port (*slot*/*port*) on which FC traffic is transmitted to the specified fabric.

Field	Description
Fabric-Map	Name of the FCoE map containing the FCoE/FC configuration parameters for the server CNA-fabric connection.
Login Method	Method used by the server CNA to log in to the fabric; for example: FLOGI - ENode logged in using a fabric login (FLOGI). FDISC - ENode logged in using a fabric discovery (FDISC).
Status	Operational status of the link between a server CNA port and a SAN fabric: Logged In - Server has logged in to the fabric and is able to transmit FCoE traffic

show npiv devices Command Example

```
Dell# show npiv devices
ENode[0]:
ENode MAC     :     00:10:18:f1:94:21
ENode Intf    :     Te 0/12
FCF MAC       :     5c:f9:dd:ef:10:c8
Fabric Intf   :     Fc 0/5
FCoE Vlan     :     1003
Fabric Map    :     fid_1003
ENode WWPN    :     20:01:00:10:18:f1:94:20
ENode WWNN    :     20:00:00:10:18:f1:94:21
FCoE MAC      :     0e:fc:03:01:02:01
FC-ID         :     01:02:01
LoginMethod   :     FLOGI
Secs          :     5593
Status        :     LOGGED_IN

ENode[1]:
ENode MAC     :     00:10:18:f1:94:22
ENode Intf    :     Te 0/13
FCF MAC       :     5c:f9:dd:ef:10:c9
Fabric Intf   :     Fc 0/0
FCoE Vlan     :     1003
Fabric Map    :     fid_1003
ENode WWPN    :     10:00:00:00:c9:d9:9c:cb
ENode WWNN    :     10:00:00:00:c9:d9:9c:cd
FCoE MAC      :     0e:fc:03:01:02:02
FC-ID         :     01:02:01
LoginMethod   :     FDISC
Secs          :     5593
Status        :     LOGGED_IN
```

Table 93. show npiv devices Field Descriptions

Field	Description
ENode [*number*]	Server CNA that has successfully logged in to a fabric over the FN IOM with the FC Flex IO module Ethernet port in ENode mode.
Enode MAC	MAC address of a server CNA port.
Enode Intf	Port number of a server-facing Ethernet port operating in ENode mode.

Field	Description
FCF MAC	Fibre Channel forwarder MAC: MAC address of FN IOM with the FC Flex IO module FCF interface.
Fabric Intf	Fabric-facing FN IOM with the FC Flex IO module Fibre Channel port (*slot*/*port*) on which FCoE traffic is transmitted to the specified fabric.
FCoE VLAN	ID of the dedicated VLAN used to transmit FCoE traffic from a server CNA to a fabric and configured on both the server-facing FN IOM with the FC Flex IO module port and server CNA port.
Fabric Map	Name of the FCoE map containing the FCoE/FC configuration parameters for the server CNA-fabric connection.
Enode WWPN	Worldwide port name of the server CNA port.
Enode WWNN	Worldwide node name of the server CNA.
FCoE MAC	Fabric-provided MAC address (FPMA). The FPMA consists of the FC-MAP value in the FCoE map and the FC-ID provided by the fabric after a successful FLOGI. In the FPMA, the most significant bytes are the FC-MAP; the least significant bytes are the FC-ID.
FC-ID	FC port ID provided by the fabric.
LoginMethod	Method used by the server CNA to log in to the fabric; for example, FLOGI or FDISC.
Secs	Number of seconds that the fabric connection is up.
State	Status of the fabric connection: logged in.

show fc switch Command Example

```
Dell# show fc switch
Switch Mode : NPG
Switch WWN  : 10:00:5c:f9:dd:ef:10:c0
Dell#
```

Table 94. show fc switch Command Description

Field	Description
Switch Mode	Fibre Channel mode o f operation of the FN IOM with the FC Flex IO module. Default: NPG (configured as an NPIV proxy gateway).
Switch WWN	Factory-assigned worldwide node (WWN) name of the FN IOM with the FC Flex IO module. The FN IOM with the FC Flex IO module WWN name is not user-configurable.