FTOS Configuration Guide for the S55 System FTOS 8.3.5.3

Force10

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.

Information in this publication is subject to change without notice.
© 2012 Dell Force10. All rights reserved.

Reproduction of these materials in any manner whatsoever without the written permission of Dell Inc. is strictly forbidden.

Trademarks used in this text: Dell™, the DELL logo, Dell Precision™, OptiPlex™, Latitude™, PowerEdge™, PowerVault™, PowerConnect™, OpenManage™, EqualLogic™, KACE™, FlexAddress™ and Vostro™ are trademarks of Dell Inc. Intel®, Pentium®, Xeon®, Core™ and Celeron® are registered trademarks of Intel Corporation in the U.S. and other countries. AMD® is a registered trademark and AMD Opteron™, AMD Phenom™, and AMD Sempron™ are trademarks of Advanced Micro Devices, Inc. Microsoft®, Windows®, Windows Server®, MS-DOS® and Windows Vista® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Red Hat Enterprise Linux® and Enterprise Linux® are registered trademarks of Red Hat, Inc. in the United States and/or other countries. Novell® is a registered trademark and SUSE ™ is a trademark of Novell Inc. in the United States and other countries. Oracle® is a registered trademark of Oracle Corporation and/or its affiliates. Citrix®, Xen®, XenServer® and XenMotion® are either registered trademarks or trademarks of Citrix Systems, Inc. in the United States and/or other countries. VMware®, Virtual SMP®, vMotion®, vCenter®, and vSphere® are registered trademarks or trademarks of VMWare, Inc. in the United States or other countries.

Other trademarks and trade names may be used in this publication to refer to either the entities claiming the marks and names or their products. Dell Inc. disclaims any proprietary interest in trademarks and trade names other than its own.

November 2012

1

About this Guide

Objectives

This guide describes the protocols and features supported by the Dell Force10 Operating System (FTOS) and provides configuration instructions and examples for implementing them. It supports the system platforms E-Series, C-Series, and S-Series.

Though this guide contains information on protocols, it is not intended to be a complete reference. This guide is a reference for configuring protocols on Dell Force10 systems. For complete information on protocols, refer to other documentation including IETF Requests for Comment (RFCs). The instructions in this guide cite relevant RFCs, and Appendix 47, Standards Compliance contains a complete list of the supported RFCs and Management Information Base files (MIBs).

Audience

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

Conventions

This document uses the following conventions to describe command syntax:

Convention	Description
keyword	Keywords are in bold and should 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 bar require you to choose one.

Information Symbols

Table 1-1 describes symbols contained in this guide.

Table 1-1. Information Symbols

Symbol	Warning	Description
✎	Note	This symbol informs you of important operational information.
⚙	FTOS Behavior	This symbol informs you of an FTOS behavior. These behaviors are inherent to the Dell Force10 system or FTOS feature and are non-configurable.
C E S	Platform Specific Feature	This symbol informs you of a feature that supported on one or two platforms only: E is for E-Series, C is for C-Series, S is for S-Series.
E T E X	E-Series Specific Feature/Command	If a feature or command applies to only one of the E-Series platforms, a separate symbol calls this to attention: E T for the TeraScale or E X for the ExaScale.
*	Exception	This symbol is a note associated with some other text on the page that is marked with an asterisk.

Related Documents

For more information about the Dell Force10 E-Series, C-Series, and S-Series refer to the following documents:

- *FTOS Command Reference*
- *Dell Force10 Network Operations Guide*
- *Installing and Maintaining the S55 System*
- *FTOS Release Notes*

2

Configuration Fundamentals

The FTOS Command Line Interface (CLI) is a text-based interface through which you can configure interfaces and protocols. The CLI is largely the same for the E-Series, C-Series, and S-Series with the exception of some commands and command outputs. The CLI is structured in modes for security and management purposes. Different sets of commands are available in each mode, and you can limit user access to modes using privilege levels.

In FTOS, after a command is enabled, 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.

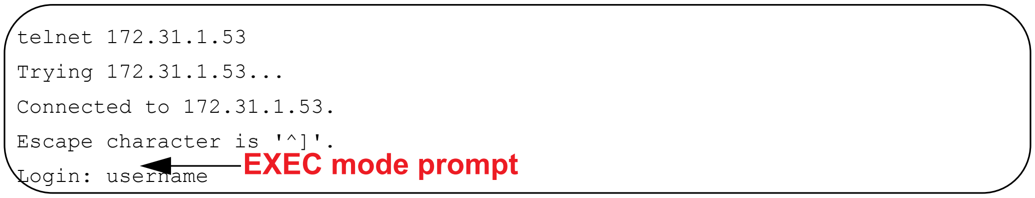
Note: Due to a differences in hardware architecture and the continued system development, features may occasionally differ between the platforms. These differences are identified by the information symbols shown on Table 1-1 on page 24.

Accessing the Command Line

Access the command line through a serial console port or a Telnet session (Figure 2-1). When the system successfully boots, you enter the command line in the EXEC mode.

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

Figure 2-1. 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
```

CLI Modes

Different sets of commands are available in each mode. A command found in one mode cannot be executed from another mode (with the exception of EXEC mode commands preceded by the command **do**; see The do Command on page 30). You can set user access rights to commands and command modes using privilege levels; for more information on privilege levels and security options, refer to Chapter 9, Security, on page 627.

The FTOS 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 **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; see Configure the Enable Password on page 40.
- **CONFIGURATION mode** enables you to configure security features, time settings, set logging and SNMP functions, configure static ARP and MAC addresses, and set line cards on the system.

Beneath CONFIGURATION mode are sub-modes that apply to interfaces, protocols, and features. Figure 2-2 illustrates this sub-mode command structure. Two sub-CONFIGURATION modes are important when configuring the chassis for the first time:

- **INTERFACE sub-mode** is the mode in which you configure Layer 2 and Layer 3 protocols and IP services specific to an interface. An interface can be physical (Management interface, 1-Gigabit Ethernet, 10-Gigabit Ethernet) or logical (Loopback, Null, port channel, or VLAN).
- **LINE sub-mode** is the mode in which you to configure the console and virtual terminal lines.

Note: At any time, entering a question mark (?) will display the available command options. For example, when you are in CONFIGURATION mode, entering the question mark first will list all available commands, including the possible sub-modes.

Figure 2-2. CLI Modes in FTOS

- EXEC
- EXEC Privilege
- CONFIGURATION
  - ARCHIVE
  - AS-PATH ACL
  - INTERFACE
    - GIGABIT ETHERNET
    - 10 GIGABIT ETHERNET
    - INTERFACE RANGE
    - LOOPBACK
    - MANAGEMENT ETHERNET
    - NULL
    - PORT-CHANNEL
    - SONET
    - VLAN
    - VRRP
  - IP
  - IPv6
  - IP COMMUNITY-LIST
  - IP ACCESS-LIST
    - STANDARD ACCESS-LIST
    - EXTENDED ACCESS-LIST
  - LINE
    - AUXILIARY
    - CONSOLE
    - VIRTUAL TERMINAL
  - MAC ACCESS-LIST
  - MONITOR SESSION
  - MULTIPLE SPANNING TREE
  - Per-VLAN SPANNING TREE
  - PREFIX-LIST
  - RAPID SPANNING TREE
  - REDIRECT
  - ROUTE-MAP
  - ROUTER BGP
  - ROUTER ISIS
  - ROUTER OSPF
  - ROUTER RIP
  - SPANNING TREE
  - TRACE-LIST

Navigating CLI Modes

The FTOS prompt changes to indicate the CLI mode. Table 2-1 lists the CLI mode, its prompt, and information on how to access and exit this CLI mode. You must move linearly through the command modes, with the exception of the **end** command which takes you directly to EXEC Privilege mode; the **exit** command moves you up one command mode level.

Note: Sub-CONFIGURATION modes all have the letters “conf” in the prompt with additional modifiers to identify the mode and slot/port information. These are shown in Table 2-1.

www.dell.com | support.dell.com

Table 2-1. FTOS Command Modes

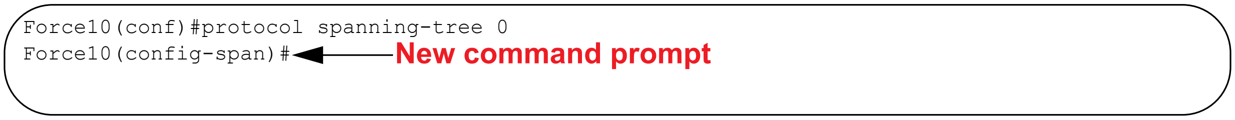
	CLI Command Mode	Prompt	Access Command
	EXEC	`Force10>`	Access the router through the console or Telnet.
	EXEC Privilege	`Force10#`	• From EXEC mode, enter the command **enable**. • From any other mode, use the command **end**.
	CONFIGURATION	`Force10(conf)#`	• From EXEC privilege mode, enter the command **configure**. • From every mode except EXEC and EXEC Privilege, enter the command **exit**.
	Note: Access all of the following modes from CONFIGURATION mode.		
	ARCHIVE	Force10(conf-archive)	**archive**
	AS-PATH ACL	Force10(config-as-path)#	**ip as-path access-list**
INTERFACE modes	Gigabit Ethernet Interface	`Force10(conf-if-gi-0/0)#`	**interface**
	10 Gigabit Ethernet Interface	`Force10(conf-if-te-0/0)#`	
	Interface Range	`Force10(conf-if-range)#`	
	Loopback Interface	`Force10(conf-if-lo-0)#`	
	Management Ethernet Interface	Force10(conf-if-ma-0/0)#	
	Null Interface	Force10(conf-if-nu-0)#	
	Port-channel Interface	Force10(conf-if-po-0)#	
	VLAN Interface	Force10(conf-if-vl-0)#	
IP ACCESS-LIST	STANDARD ACCESS-LIST	Force10(config-std-nacl)#	**ip access-list standard**
	EXTENDED ACCESS-LIST	Force10(config-ext-nacl)#	**ip access-list extended**
	IP COMMUNITY-LIST	Force10(config-community-list)#	**ip community-list**
LINE	AUXILIARY	Force10(config-line-aux)#	**line**
	CONSOLE	Force10(config-line-console)#	
	VIRTUAL TERMINAL	Force10(config-line-vty)#	

Table 2-1. FTOS Command Modes (continued)

CLI Command Mode		Prompt	Access Command
MAC ACCESS-LIST	STANDARD ACCESS-LIST	Force10(config-std-macl)#	**mac access-list standard**
	EXTENDED ACCESS-LIST	Force10(config-ext-macl)#	**mac access-list extended**
	MULTIPLE SPANNING TREE	Force10(config-mstp)#	**protocol spanning-tree mstp**
	Per-VLAN SPANNING TREE Plus	Force10(config-pvst)#	**protocol spanning-tree pvst**
	PREFIX-LIST	Force10(conf-nprefixl)#	**ip prefix-list**
	RAPID SPANNING TREE	Force10(config-rstp)#	**protocol spanning-tree rstp**
	REDIRECT	Force10(conf-redirect-list)#	**ip redirect-list**
	ROUTE-MAP	Force10(config-route-map)#	**route-map**
	ROUTER BGP	Force10(conf-router_bgp)#	**router bgp**
	ROUTER ISIS	Force10(conf-router_isis)#	**router isis**
	ROUTER OSPF	Force10(conf-router_ospf)#	**router ospf**
	ROUTER RIP	Force10(conf-router_rip)#	**router rip**
	SPANNING TREE	Force10(config-span)#	**protocol spanning-tree 0**
	TRACE-LIST	Force10(conf-trace-acl)#	**ip trace-list**

Figure 2-3 illustrates how to change the command mode from CONFIGURATION mode to PROTOCOL SPANNING TREE.

Figure 2-3. Changing CLI Modes



```
Force10(conf)#protocol spanning-tree 0
Force10(config-span)#
```

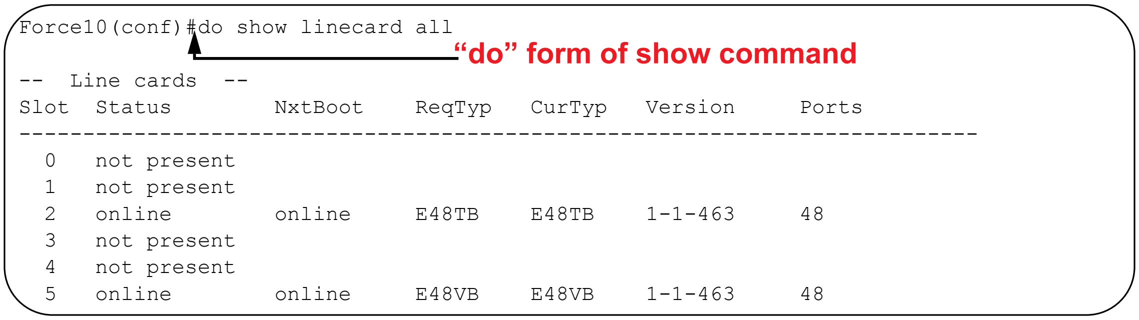
New command prompt

The do Command

Enter an EXEC mode command from any CONFIGURATION mode (CONFIGURATION, INTERFACE, SPANNING TREE, etc.) without returning to EXEC mode by preceding the EXEC mode command with the command **do**. Figure 2-4 illustrates the **do** command.

Note: The following commands cannot be modified by the **do** command: **enable, disable, exit**, and **configure**.

Figure 2-4. Using the do Command



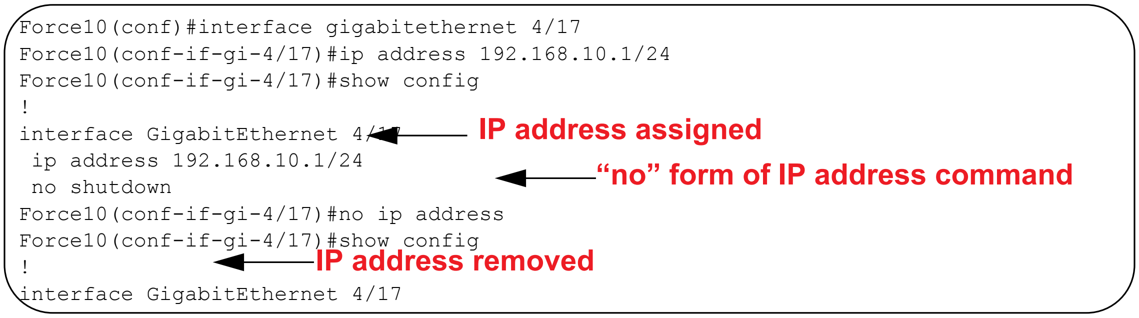
```
Force10(conf)#do show linecard all

--  Line cards  --
Slot  Status         NxtBoot    ReqTyp   CurTyp   Version     Ports
---------------------------------------------------------------------------
  0   not present
  1   not present
  2   online         online     E48TB    E48TB    1-1-463     48
  3   not present
  4   not present
  5   online         online     E48VB    E48VB    1-1-463     48
```

Undoing Commands

When you enter a command, the command line is added to the running configuration file. Disable a command and remove it from the running-config by entering the original command preceded by the command **no**. For example, to delete an ip address configured on an interface, use the **no ip address** *ip-address* command, as shown in Figure 2-5.

Note: Use the **help** or **?** command as discussed in Obtaining Help command to help you construct the “no” form of a command.

Figure 2-5. Undoing a command with the no Command



```
Force10(conf)#interface gigabitethernet 4/17
Force10(conf-if-gi-4/17)#ip address 192.168.10.1/24
Force10(conf-if-gi-4/17)#show config
!
interface GigabitEthernet 4/17
 ip address 192.168.10.1/24
 no shutdown
Force10(conf-if-gi-4/17)#no ip address
Force10(conf-if-gi-4/17)#show config
!
interface GigabitEthernet 4/17
```

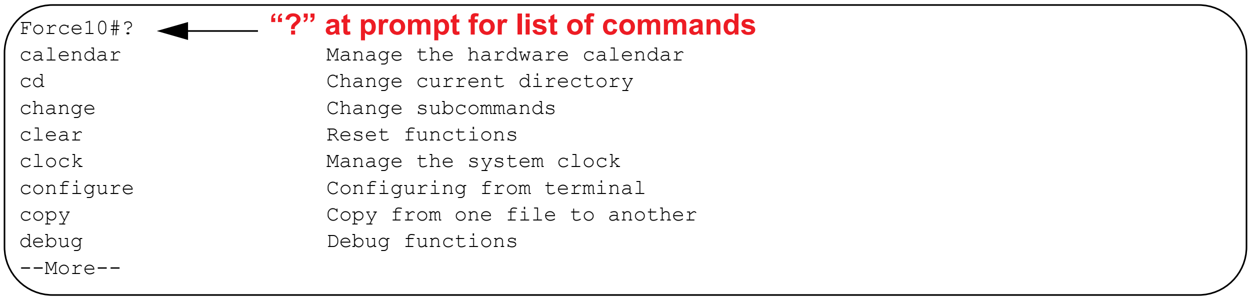
Layer 2 protocols are disabled by default. Enable them using the **no disable** command. For example, in PROTOCOL SPANNING TREE mode, enter **no disable** to enable Spanning Tree.

Obtaining Help

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

- Enter **?** at the prompt or after a keyword to list the keywords available in the current mode.
  - **?** after a prompt lists all of the available keywords. The output of this command is the same for the **help** command.

Figure 2-6. ? Command Example

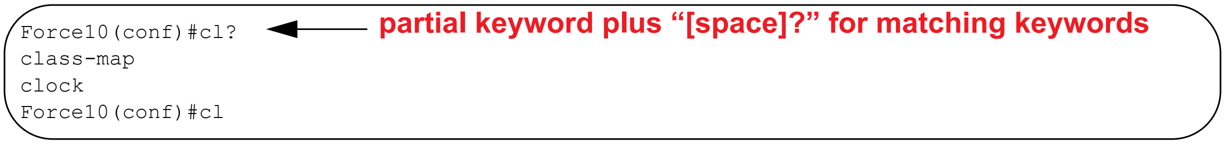


```
Force10#?
calendar                Manage the hardware calendar
cd                      Change current directory
change                  Change subcommands
clear                   Reset functions
clock                   Manage the system clock
configure               Configuring from terminal
copy                    Copy from one file to another
debug                   Debug functions
--More--
```

"?" at prompt for list of commands

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

Figure 2-7. Keyword? Command Example

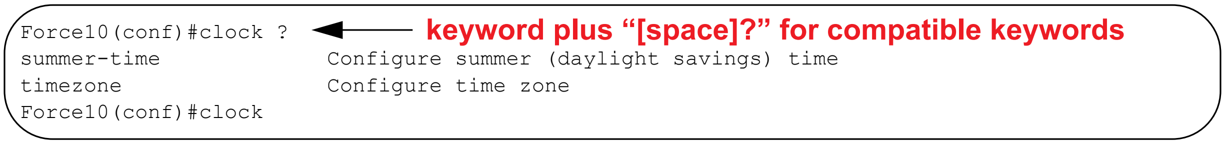


```
Force10(conf)#cl?
class-map
clock
Force10(conf)#cl
```

partial keyword plus "[space]?" for matching keywords

-
  - A keyword followed by [space]**?** lists all of the keywords that can follow the specified keyword.

Figure 2-8. Keyword ? Command Example



```
Force10(conf)#clock ?
summer-time             Configure summer (daylight savings) time
timezone                Configure time zone
Force10(conf)#clock
```

keyword plus "[space]?" for compatible keywords

Entering and Editing Commands

When entering commands:

- The CLI is not case sensitive.
- You can enter partial CLI keywords.
  - You must enter the minimum number of letters to uniquely identify a command. For example, **cl** cannot be entered as a partial keyword because both the **clock** and **class-map** commands begin with the letters "cl." **clo**, however, can be entered as a partial keyword because only one command begins with those three letters.
- The TAB key auto-completes keywords in commands. You must enter the minimum number of letters to uniquely identify a command.

- The UP and DOWN arrow keys display previously entered commands (see Command History).
- The BACKSPACE and DELETE keys erase the previous letter.
- Key combinations are available to move quickly across the command line, as described in Table 2-2.

Table 2-2. Short-Cut Keys and their Actions

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-R	Re-enters the previous command.
CNTL-U	Deletes the line.
CNTL-W	Deletes the previous word.
CNTL-X	Deletes the line.
CNTL-Z	Ends continuous scrolling of command outputs.
Esc B	Moves the cursor back one word.
Esc F	Moves the cursor forward one word.
Esc D	Deletes all characters from the cursor to the end of the word.

Command History

FTOS 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 the **ignore-case** sub-option is implemented.

Starting with FTOS 7.8.1.0, the **grep** command accepts an **ignore-case** sub-option that forces the search to case-*in*sensitive. For example, the commands:

- **show run | grep Ethernet** returns a search result with instances containing a capitalized “Ethernet,” such as `interface GigabitEthernet 0/0`.
- **show run | grep ethernet** would not return that search result because it only searches for instances containing a non-capitalized “ethernet.”

Executing the command **show run | grep Ethernet ignore-case** would return instances containing both “Ethernet” and “ethernet.”

- **grep** displays only the lines containing specified text. Figure 2-9 shows this command used in combination with the command **show linecard all**.

Figure 2-9. Filtering Command Outputs with the grep Command

```
Force10(conf)#do show linecard all | grep 0
  0   not present
```

Note: FTOS accepts a space or no space before and after the pipe. To filter on a phrase with spaces, underscores, or ranges, enclose the phrase with double quotation marks.

- **except** displays text that does not match the specified text. Figure 2-10 shows this command used in combination with the command **show linecard all**.

Figure 2-10. Filtering Command Outputs with the except Command

```
Force10#show linecard all | except 0

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
---------------------------------------------------------------------------
  2   not present
  3   not present
  4   not present
  5   not present
  6   not present
```

- **find** displays the output of the show command beginning from the first occurrence of specified text Figure 2-11 shows this command used in combination with the command **show linecard all**.

Figure 2-11. Filtering Command Outputs with the find Command

```
Force10(conf)#do show linecard all | find 0
  0   not present
  1   not present
  2   online        online     E48TB    E48TB    1-1-463     48
  3   not present
  4   not present
  5   online        online     E48VB    E48VB    1-1-463     48
  6   not present
  7   not present
```

- **display** displays additional configuration information.
- **no-more** displays the output all at once rather than one screen at a time. This is similar to the command **terminal length** except that the **no-more** option affects the output of the specified command only.
- **save** copies the output to a file for future reference.

Note: You can filter a single command output multiple times. The save option should be the last option entered. For example:
Force10# *command* **| grep** *regular-expression* **| except** *regular-expression* **| grep** *other-regular-expression* **| find** *regular-expression* **| save**

Multiple Users in Configuration mode

FTOS notifies all users in the event that there are multiple users logged into 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, Message 1 appears:

Message 1 Multiple Users in Configuration mode Telnet Message

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

- On the system that is connected over the console, Message 2 appears:

Message 2 Multiple Users in Configuration mode Telnet Message

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

If either of these messages appears, Dell Force10 recommends that you coordinate with the users listed in the message so that you do not unintentionally overwrite each other's configuration changes.

3

Getting Started

This chapter contains the following major sections:

- Default Configuration
- Configure a Host Name
- Access the System Remotely
- Configure the Enable Password
- Configuration File Management
- File System Management

When you power up the chassis, the system performs\ a Power-On Self Test (POST) during which Route Processor Module (RPM), Switch Fabric Module (SFM), and line card status LEDs blink green.The system then loads FTOS and 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 is complete, the RPM and line card status LEDs remain online (green), and the console monitor displays the EXEC mode prompt.

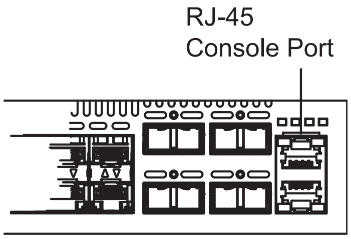
For details on using the Command Line Interface (CLI), refer to Accessing the Command Line in the Configuration Fundamentals chapter.

Console access

The S4810 has 2 management ports available for system access: a serial console port and an Out-of-Bounds (OOB) port.

Serial console

The RJ-45/RS-232 console port is labeled on the S4810 chassis. It is in the upper right-hand side, as you face the I/O side of the chassis.



To access the console port, follow the procedures below. Refer to Table 3-1, "Pin Assignments Between the Console and a DTE Terminal Server," in Getting Started for the console port pinout.

Step	Task
1	Install an RJ-45 copper cable into the console port.Use a rollover (crossover) cable to connect the S4810 console port to a terminal server.
2	Connect the other end of the cable to the DTE terminal server.
3	Terminal settings on the console port cannot be changed in the software and are set as follows: 9600 baud rate No parity 8 data bits 1 stop bit No flow control

Accessing the RJ-45 console port with a DB-9 adapter

You can connect to the console using a RJ-45 to RJ-45 rollover cable and a RJ-45 to DB-9 female DTE adapter to a terminal server (for example, PC). Table 3-1, "Pin Assignments Between the Console and a DTE Terminal Server," in Getting Started lists the pin assignments.

Table 3-1. Pin Assignments Between the Console and a DTE Terminal Server

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

Default Configuration

A version of FTOS 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 FTOS). You must configure the system using the CLI.

Configure a Host Name

The host name appears in the prompt. The default host name is FTOS.

- 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 configure a host name:

Step	Task	Command Syntax	Command Mode
1	Create a new host name.	hostname *name*	CONFIGURATION

The example below illustrates the hostname command.

```
FTOS(conf)#hostname R1
R1(conf)#
```

Access the System Remotely

You can configure the system to access it remotely by Telnet. The method for configuring the C-Series and E-Series for Telnet access is different from S-Series.

- The C-Series, E-Series and the S4810 have a dedicated management port and a management routing table that is separate from the IP routing table.
- The S-Series (except the S4810) does not have a dedicated management port, but is managed from any port. It does not have a separate management routing table.

Access the C-Series and E-Series and the S4810 Remotely

Configuring the system for Telnet is a three-step process:

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

Configure the Management Port IP Address

Assign IP addresses to the management ports in order to access the system remotely.

Note: Assign different IP addresses to each RPM's management port.

To configure the management port IP address:

Step	Task	Command Syntax	Command Mode
1	Enter INTERFACE mode for the Management port.	interface ManagementEthernet *slot/port* • *slot* range: 0 to 1 • *port* range: 0	CONFIGURATION
2	Assign an IP address to the interface.	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).	INTERFACE
3	Enable the interface.	no shutdown	INTERFACE

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:

Step	Task	Command Syntax	Command Mode
1	Configure a management route to the network from which you are accessing the system.	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.	CONFIGURATION

Configure a Username and Password

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

To configure a username and password:

Step	Task	Command Syntax	Command Mode
1	Configure a username and password to access the system remotely.	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 Force10 system.	CONFIGURATION

Access the S-Series Remotely

The S-Series does not have a dedicated management port nor a separate management routing table. Configure any port on the S-Series to be the port through which you manage the system and configure an IP route to that gateway.

Note: The S4810 system uses management ports and should be configured similar to the C-Series and E-Series systems. Refer to Access the C-Series and E-Series and the S4810 Remotely

Configuring the system for Telnet access is a three-step process:

1. Configure an IP address for the port through which you will manage the system using the command ip address from INTERFACE mode, as shown in the example below.
2. Configure a IP route with a default gateway using the command ip route from CONFIGURATION mode, as shown in the example below.
3. Configure a username and password using the command username from CONFIGURATION mode, as shown in the example below.

```
R5(conf)#int gig 0/48
R5(conf-if-gi-0/48)#ip address 10.11.131.240
R5(conf-if-gi-0/48)#show config
!
interface GigabitEthernet 0/48
 ip address 10.11.131.240/24
 no shutdown
R5(conf-if-gi-0/48)#exit
R5(conf)#ip route 10.11.32.0/23 10.11.131.254
R5(conf)#username admin pass FTOS
```

Configure the Enable Password

Access the EXEC Privilege mode using the enable command. The 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 Force10 recommends using the enable secret password.

To configure an enable password:

Task	Command Syntax	Command Mode
Create a password to access EXEC Privilege 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 Force10 system. • 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 Force10 system.	CONFIGURATION

Configuration File Management

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

The E-Series EtherScale platform architecture uses MMC cards for both the internal and external Flash memory. MMC cards support a maximum of 100 files. The E-Series TeraScale and ExaScale platforms architecture use Compact Flash for the internal and external Flash memory. It has a space limitation but does not limit the number of files it can contain.

Note: Using flash memory cards in the system that have not been approved by Dell Force10 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: See the *FTOS Command Reference* for a detailed description of the copy command.

- 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 shown in Table 3-2, "Forming a copy Command," in Getting Started.
- To copy a remote file to Dell Force10 system, combine the *file-origin* syntax for a remote file location with the *file-destination* syntax for a local file location shown in Table 3-2, "Forming a copy Command," in Getting Started.

Table 3-2. Forming a copy Command

	source-file-url **Syntax**	*destination-file-url* **Syntax**
Local File Location		
Internal flash:		
primary RPM	copy flash://*filename*	flash://*filename*
standby RPM	copy rpm{0|1}flash://*filename*	rpm{0|1}flash://*filename*
External flash:		
primary RPM	copy rpm{0|1}slot0://*filename*	rpm{0|1}slot0://*filename*
standby RPM	copy rpm{0|1}slot0://*filename*	rpm{0|1}slot0://*filename*
USB Drive (E-Series ExaScale)		
USB drive on RPM0	copy rpm0usbflash://*filepath*	rpm0usbflash://*filename*
External USB drive	copy usbflash://*filepath*	usbflash://*filename*
Remote File Location		
FTP server	copy ftp://*username*:*password*@{*hostip* | *hostname*}/*filepath*/*filename*	ftp://*username*:*password*@{*hostip* | *hostname*}/*filepath*/*filename*
TFTP server	copy tftp://{*hostip* | *hostname*}/*filepath*/*filename*	tftp://{*hostip* | *hostname*}/*filepath*/*filename*
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.
- The internal flash memories on the RPMs are synchronized whenever there is a change, but only if both RPMs are running the same version of FTOS.
- When copying to a server, a hostname can only be used if a DNS server is configured.

- The usbflash and rpm0usbflash commands are supported on E-Series ExaScale systems. Refer to your system’s Release Notes for a list of approved USB vendors.

The following text is an example of using the copy command to save a file to an FTP server.

```
FTOS#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
```

The following text is an example of using the copy command to import a file to the Dell Force10 system from an FTP server.

```
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 Force10 recommends that you copy 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 primary RPM by default, but it can be saved onto an external flash (on an RPM) or a remote server.

To save the running-configuration:

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

Task	Command Syntax	Command Mode
Save the running-configuration to:		EXEC Privilege
the startup-configuration on the internal flash of the primary RPM	copy running-config startup-config	
the internal flash on an RPM	copy running-config rpm{0|1}flash://*filename*	
Note: The internal flash memories on the RPMs are synchronized whenever there is a change, but only if the RPMs are running the same version of FTOS.		
the external flash of an RPM	copy running-config rpm{0|1}slot0://*filename*	
an FTP server	copy running-config ftp://*username*:*password*@{*hostip* | *hostname*}/*filepath*/*filename*	
a TFTP server	copy running-config tftp://{*hostip* | *hostname*}/*filepath*/*filename*	
an SCP server	copy running-config scp://{*hostip* | *hostname*}/*filepath*/*filename*	
Note: When copying to a server, a hostname can only be used if a DNS server is configured.		
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.	copy running-config startup-config duplicate	EXEC Privilege

FTOS 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 command copy *startup-config.bak startup-config*.

Configure the Overload bit for Startup Scenario

For information on setting the router overload bit for a specific period of time after a switch reload is implemented, see the *FTOS Command Line Reference Guide*, Chapter 18 - Intermediate System to Intermediate System (IS-IS).

View Files

File information and content can only be viewed on local file systems. To view a list of files on the internal or external Flash:

Step	Task	Command Syntax	Command Mode
1	View a list of files on:		
	the internal flash of an RPM	dir flash:	EXEC Privilege
	the external flash of an RPM	dir slot:	

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

```
FTOS#dir
Directory of flash:

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

To view the contents of a file:

Step	Task	Command Syntax	Command Mode
1	View the:		
	contents of a file in the internal flash of an RPM	show file rpm{0|1}flash://*filename*	
	contents of a file in the external flash of an RPM	show file rpm{0|1}slot0://*filename*	EXEC Privilege
	running-configuration	show running-config	
	startup-configuration	show startup-config	

View Configuration Files

Configuration files have three commented lines at the beginning of the file, as shown in the example below, 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,” then you have made changes that have not been saved and will not be preserved upon a system reboot.

```
FTOS#show running-config
Current Configuration ...
! Version 8.2.1.0
! Last configuration change at Thu Apr 3 23:06:28 2008 by admin
! Startup-config last updated at Thu Apr 3 23:06:55 2008 by admin
!
boot system rpm0 primary flash://FTOS-EF-8.2.1.0.bin
boot system rpm0 secondary flash://FTOS-EF-7.8.1.0.bin
boot system rpm0 default flash://FTOS-EF-7.7.1.1.bin
boot system rpm1 primary flash://FTOS-EF-7.8.1.0.bin
boot system gateway 10.10.10.100
--More--
```

File System Management

The Dell Force10 system can use the internal Flash, external Flash, or remote devices to store files. It stores files on the internal Flash by default but can be configured to store files elsewhere.

To view file system information:

Task	Command Syntax	Command Mode
View information about each file system.	show file-systems	EXEC Privilege

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

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

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

To change the default storage location:

Task	Command Syntax	Command Mode
Change the default directory.	cd *directory*	EXEC Privilege

In the example below, the default storage location is changed to the external Flash of the primary RPM. File management commands then apply to the external Flash rather than the internal Flash.

```
FTOS#cd slot0:
FTOS#copy running-config test
FTOS#copy run test
!
7419 bytes successfully copied
FTOS#dir
Directory of slot0:

  1  drw-      32768   Jan 01 1980 00:00:00  .
  2  drwx        512   Jul 23 2007 00:38:44  ..
  3  ----          0   Jan 01 1970 00:00:00  DCIM
  4  -rw-       7419   Jul 23 2007 20:44:40  test
  5  ----          0   Jan 01 1970 00:00:00  BT
  6  ----          0   Jan 01 1970 00:00:00  200702~1VSN
  7  ----          0   Jan 01 1970 00:00:00  G
  8  ----          0   Jan 01 1970 00:00:00  F
  9  ----          0   Jan 01 1970 00:00:00  F

slot0: 127772672 bytes total (21927936 bytes free)
```

View command history

The command-history trace feature captures all commands entered by all users of the system with a time stamp and writes these messages to a dedicated trace log buffer. The system generates a trace message for each executed command. No password information is saved to the file.

To view the command-history trace, use the show command-history command, as shown in the example below.

```
FTOS#show command-history
[12/5 10:57:8]: CMD-(CLI):service password-encryption
[12/5 10:57:12]: CMD-(CLI):hostname Force10
[12/5 10:57:12]: CMD-(CLI):ip telnet server enable
[12/5 10:57:12]: CMD-(CLI):line console 0
[12/5 10:57:12]: CMD-(CLI):line vty 0 9
[12/5 10:57:13]: CMD-(CLI):boot system rpm0 primary flash://FTOS-CB-1.1.1.2E2.bin
```

Upgrading FTOS

Note: To upgrade FTOS, see the release notes for the version you want to load on the system.

4

Management

Management is supported on platforms: C E S

This chapter explains the different protocols or services used to manage the Dell Force10 system including:

Configure Privilege Levels

Privilege levels restrict access to commands based on user or terminal line. There are 16 privilege levels, of which three are pre-defined. The default privilege level is 1.

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

Create a Custom Privilege Level

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

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

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

Removing a command from EXEC mode

Remove a command from the list of available commands in EXEC mode for a specific privilege level using the command **privilege exec** from CONFIGURATION mode. In the command, specify a level *greater* than the level given to a user or terminal line, followed by the first keyword of each command to be restricted.

Move a command from EXEC privilege mode to EXEC mode

Move a command from EXEC Privilege to EXEC mode for a privilege level using the command **privilege exec** 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.

Allow Access to CONFIGURATION mode commands

Allow access to CONFIGURATION mode using the command **privilege exec level** *level* **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**. You must individually specify each CONFIGURATION mode command to which you want to allow access using the command **privilege configure level** *level*. 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.

Allow 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, you must first allow access to the command that enters you into the mode. For example, allow a user to enter INTERFACE mode using the command **privilege configure level** *level* **interface gigabitethernet**
2. Then, individually identify the INTERFACE, LINE, ROUTE-MAP or ROUTER commands to which you want to allow access using the command **privilege** {**interface** | **line** | **route-map** | **router**} **level** *level*. 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.

The following table lists the configuration tasks you can use to customize a privilege level:

Task	Command Syntax	Command Mode
Remove a command from the list of available commands in EXEC mode.	**privilege exec level** *level* {*command* ||...|| *command*}	CONFIGURATION
Move a command from EXEC Privilege to EXEC mode.	**privilege exec level** *level* {*command* ||...|| *command*}	CONFIGURATION
Allow access to CONFIGURATION mode.	**privilege exec level** *level* **configure**	CONFIGURATION

Task	Command Syntax	Command Mode
Allow access to INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode. Specify *all* keywords in the command.	**privilege configure level** *level* {**interface** | **line** | **route-map** | **router**} {*command-keyword* ||...|| *command-keyword*}	CONFIGURATION
Allow access to a CONFIGURATION, INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode command.	**privilege** {**configure** |**interface** | **line** | **route-map** | **router**} **level** *level* {*command* ||...|| *command*}	CONFIGURATION

The configuration in Figure 4-1 creates privilege level 3. This level:

- removes the **resequence** command from EXEC mode by requiring a minimum of privilege level 4,
- moves the command **capture bgp-pdu max-buffer-size** 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, and
- allows access to INTERFACE and LINE modes are allowed with no commands.

Figure 4-1. Create a Custom Privilege Level

```
Force10(conf)#do show run priv
!
privilege exec level 3 capture
privilege exec level 3 configure
privilege exec level 4 resequence
privilege exec level 3 capture bgp-pdu
privilege exec level 3 capture bgp-pdu max-buffer-size
privilege configure level 3 line
privilege configure level 3 interface
Force10(conf)#do telnet 10.11.80.201
[telnet output omitted]
Force10#show priv
Current privilege level is 3.
Force10#?
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]
Force10#config
[output omitted]
Force10(conf)#do show priv
Current privilege level is 3.
Force10(conf)#?
end                     Exit from configuration mode
exit                    Exit from configuration mode
interface               Select an interface to configure
line                    Configure a terminal line
linecard                Set line card type
Force10(conf)#interface ?
fastethernet            Fast Ethernet interface
gigabitethernet         Gigabit Ethernet interface
loopback                Loopback interface
managementethernet      Management Ethernet interface
null                    Null interface
port-channel            Port-channel interface
range                   Configure interface range
sonet                   SONET interface
tengigabitethernet      TenGigabit Ethernet interface
vlan                    VLAN interface
Force10(conf)#interface gigabitethernet 1/1
Force10(conf-if-gi-1/1)#?
end                             Exit from configuration mode
exit                            Exit from interface configuration mode
Force10(conf-if-gi-1/1)#exit
Force10(conf)#line ?
aux                             Auxiliary line
console                         Primary terminal line
vty                             Virtual terminal
Force10(conf)#line vty 0
Force10(config-line-vty)#?
exit                            Exit from line configuration mode
Force10(config-line-vty)#
```

Apply a Privilege Level to a Username

To set a privilege level for a user:

Task	Command Syntax	Command Mode
Configure a privilege level for a user.	**username** *username* **privilege** *level*	CONFIGURATION

Apply a Privilege Level to a Terminal Line

To set a privilege level for a terminal line:

Task	Command Syntax	Command Mode
Configure a privilege level for a terminal line.	**privilege level** *level*	LINE

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*>.

Configure Logging

FTOS tracks changes in the system using event and error messages. By default, FTOS logs these messages on:

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

Disable Logging

To disable logging:

Task	Command Syntax	Command Mode
Disable all logging except on the console.	**no logging on**	CONFIGURATION
Disable logging to the logging buffer.	no logging buffer	CONFIGURATION
Disable logging to terminal lines.	no logging monitor	CONFIGURATION
Disable console logging.	no logging console	CONFIGURATION

Log Messages in the Internal Buffer

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

Message 1 BootUp Events

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

Configuration Task List for System Log Management

The following list includes the configuration tasks for system log management:

- Disable System Logging on page 52
- Send System Messages to a Syslog Server on page 53

Disable System Logging

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

Enable and disable system logging using the following commands:

Task	Command Syntax	Command Mode
Disable all logging except on the console.	**no logging on**	CONFIGURATION
Disable logging to the logging buffer.	no logging buffer	CONFIGURATION
Disable logging to terminal lines.	no logging monitor	CONFIGURATION
Disable console logging.	no logging console	CONFIGURATION

Send System Messages to a Syslog Server

Send system messages to a syslog server by specifying the server with the following command:

Task	Command Syntax	Command Mode
Specify the server to which you want to send system messages. You can configure up to eight syslog servers.	**logging** {*ip-address* | *hostname*}	CONFIGURATION

Configure a Unix System as a Syslog Server

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.

- on a 4.1 BSD UNIX system, add the line: **local7.debugging /var/log/force10.log**
- on a 5.7 SunOS UNIX system, add the line: **local7.debugging /var/adm/force10.log**

In the lines above, **local7** is the logging facility level and debugging is the severity level.

Change 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.

Task	Command Syntax	Command Mode
Specify the minimum severity level for logging to the logging buffer.	**logging buffered** *level*	CONFIGURATION
Specify the minimum severity level for logging to the console.	**logging console** *level*	CONFIGURATION
Specify the minimum severity level for logging to terminal lines.	**logging monitor** *level*	CONFIGURATION
Specifying the minimum severity level for logging to a syslog server.	**logging trap** *level*	CONFIGURATION
Specify the minimum severity level for logging to the syslog history table.	**logging history** *level*	CONFIGURATION

Task	Command Syntax	Command Mode
Specify the size of the logging buffer. **Note**: When you decrease the buffer size, FTOS deletes all messages stored in the buffer. Increasing the buffer size does not affect messages in the buffer.	**logging buffered** *size*	CONFIGURATION
Specify the number of messages that FTOS saves to its logging history table.	**logging history size** *size*	CONFIGURATION

To change one of the settings for logging system messages, use any or all of the following commands in the CONFIGURATION mode:

To view the logging buffer and configuration, use the **show logging** command (Figure 35) in the EXEC privilege mode.

To change the severity level of messages logged to a syslog server, use the following command in the CONFIGURATION mode:

To view the logging configuration, use the **show running-config logging** command (Figure 37) in the EXEC privilege mode.

Display the Logging Buffer and the Logging Configuration

Display the current contents of the logging buffer and the logging settings for the system, use the **show logging** command (Figure 35) in the EXEC privilege mode.

Figure 4-2. show logging Command Example

```
Force10#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 (Figure 37) in the EXEC privilege mode.

Configure 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 in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
logging facility [*facility-type*]	CONFIGURATION	Specify one of the following parameters. • 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). This is the default. • 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) The default is local7.

To view nondefault settings, use the **show running-config logging** command (Figure 37) in the EXEC mode.

Figure 4-3. show running-config logging Command Example

```
Force10#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
Force10#
```

Synchronize log messages

You can configure FTOS 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.

To synchronize log messages, use these commands in the following sequence starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**line** {**console 0** | **vty** *number* [*end-number*] | **aux 0**}	CONFIGURATION	Enter the LINE mode. Configure the following parameters for the virtual terminal lines: • *number* range: zero (0) to 8. • *end-number* range: 1 to 8. You can configure multiple virtual terminals at one time by entering a *number* and an *end-number.*
2	**logging synchronous** [**level** *severity-level* | **all**] [*limit*]	LINE	Configure a level and set the maximum number of messages to be printed. Configure the following optional parameters: • **level** *severity-level* range: 0 to 7. Default is 2. Use the **all** keyword to include all messages. • *limit* range: 20 to 300. Default is 20.

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

Enable timestamp on syslog messages

syslog messages, by default, do not include a time/date stamp stating when the error or message was created.

To have FTOS include a timestamp with the syslog message, use the following command syntax in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
service timestamps [**log** | **debug**] [**datetime** [**localtime**] [**msec**] [**show-timezone**] | **uptime**]	CONFIGURATION	Add timestamp to syslog messages. 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 neither parameter is specified, FTOS configures **uptime**.

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

To disable time stamping on syslog messages, enter **no service timestamps** [**log** | **debug**].

File Transfer Services

With FTOS, you can configure the system to transfer files over the network using 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 VLAN interfaces.

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

Configuration Task List for File Transfer Services

The following list includes the configuration tasks for file transfer services:

- Enable FTP server on page 59 (mandatory)
- Configure FTP server parameters on page 59 (optional)
- Configure FTP client parameters on page 60 (optional)

For a complete listing of FTP related commands, refer to .

Enable FTP server

To enable the system as an FTP server, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ftp-server enable	CONFIGURATION	Enable FTP on the system.

To view FTP configuration, use the **show running-config ftp** command (Figure 41) in the EXEC privilege mode.

Figure 4-4. show running-config ftp Command Output

```
Force10#show running ftp
!
ftp-server enable
ftp-server username nairobi password 0 zanzibar
Force10#
```

Configure FTP server parameters

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

To configure FTP server parameters, use any or all of the following commands in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ftp-server topdir *dir*	CONFIGURATION	Specify the directory for users using FTP to reach the system. The default is the internal flash directory.
ftp-server username *username* **password** [*encryption-type*] *password*	CONFIGURATION	Specify a user name for all FTP users and configure either a plain text or encrypted 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 **ftp-server topdir** has been configured.

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

Configure FTP client parameters

To configure FTP client parameters, use the following commands in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip ftp source-interface *interface*	CONFIGURATION	Enter the following keywords and slot/port or number information: • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by a number between 0 and 16383. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.
ip ftp password *password*	CONFIGURATION	Configure a password.
ip ftp username *name*	CONFIGURATION	Enter username to use on FTP client.

To view FTP configuration, use the **show running-config ftp** command (Figure 41) in the EXEC privilege mode.

Terminal Lines

You can access the system remotely and restrict access to the system by creating user profiles. The terminal lines on the system provide different means of accessing the system. The console line (console) connects you through the Console port in the RPMs. The virtual terminal lines (VTY) connect you through Telnet to the system. The auxiliary line (aux) connects secondary devices such as modems.

Deny and Permit Access to a Terminal Line

Dell Force10 recommends applying only standard ACLs to deny and permit access to VTY lines.

- Layer 3 ACL deny all traffic that is not explicitly permitted, but in the case of VTY lines, an ACL with no rules does not deny any traffic.
- You cannot use **show ip accounting access-list** to display the contents of an ACL that is applied only to a VTY line.

To apply an IP ACL to a line:

Task	Command Syntax	Command Mode
Apply an ACL to a VTY line.	**ip access-class** *access-list*	LINE

To view the configuration, enter the **show config** command in the LINE mode, as shown in Figure 4-5.

Figure 4-5. Applying an Access List to a VTY Line

```
Force10(config-std-nacl)#show config
!
ip access-list standard myvtyacl
 seq 5 permit host 10.11.0.1
Force10(config-std-nacl)#line vty 0
Force10(config-line-vty)#show config
line vty 0
 access-class myvtyacl
```

FTOS Behavior: Prior to FTOS version 7.4.2.0, in order to deny access on a VTY line, you must apply an ACL and AAA authentication to the line. Then users are denied access only *after* they enter a username and password. Beginning in FTOS version 7.4.2.0, only an ACL is required, and users are denied access *before* they are prompted for a username and password.

Configure Login Authentication for Terminal Lines

You can use any combination of up to 6 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, FTOS 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 e password you assigned to the terminal line. You must 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 command password from LINE mode.
- **local**—Prompt for the 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.

To configure authentication for a terminal line:

Step	Task	Command Syntax	Command Mode
1	Create an authentication method list. You may use a mnemonic name or use the keyword **default**. The default authentication method for terminal lines is **local**, and the **default** method list is empty.	**aaa authentication login** {*method-list-name* | **default**} [*method-1*] [*method-2*] [*method-3*] [*method-4*] [*method-5*] [*method-6*]	CONFIGURATION
2	Apply the method list from Step 1 to a terminal line.	**login authentication** {*method-list-name* | **default**}	CONFIGURATION
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.	**password**	LINE

In Figure 4-6 VTY lines 0-2 use a single authentication method, **line**.

Figure 4-6. Configuring Login Authentication on a Terminal Line

```
Force10(conf)#aaa authentication login myvtymethodlist line
Force10(conf)#line vty 0 2
Force10(config-line-vty)#login authentication myvtymethodlist
Force10(config-line-vty)#password myvtypassword
Force10(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
Force10(config-line-vty)#
```

Time out of EXEC Privilege Mode

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

To change the timeout period or disable EXEC timeout.

Task	Command Syntax	Command Mode
Set the number of minutes and seconds. Default: 10 minutes on console, 30 minutes on VTY. Disable EXEC timeout by setting the timeout period to 0.	**exec-timeout** *minutes* [*seconds*]	LINE
Return to the default timeout values.	**no exec-timeout**	LINE

View the configuration using the command **show config** from LINE mode.

Figure 4-7. Configuring EXEC Timeout

```
Force10(conf)#line con 0
Force10(config-line-console)#exec-timeout 0
Force10(config-line-console)#show config
line console 0
 exec-timeout 0 0
Force10(config-line-console)#
```

Telnet to Another Network Device

To telnet to another device:

Task	Command Syntax	Command Mode
Telnet to the peer RPM. You do not need to configure the management port on the peer RPM to be able to telnet to it.	**telnet-peer-rpm**	EXEC Privilege
Telnet to a device with an IPv4 or IPv6 address. If you do not enter an IP address, FTOS enters a Telnet dialog that prompts you for one. • Enter an IPv4 address in dotted decimal format (A.B.C.D). • Enter an IPv6 address in the format 0000:0000:0000:0000:0000:0000:0000:0000. Elision of zeros is supported.	**telnet** [*ip-address*]	EXEC Privilege

Figure 4-8. Telnet to Another Network Device

```
Force10# telnet 10.11.80.203
Trying 10.11.80.203...
Connected to 10.11.80.203.
Exit character is '^]'.
Login:
Login: admin
Password:
Force10>exit
Force10#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
Force10#
```

Lock CONFIGURATION mode

FTOS 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).

A two types of locks can be set: auto and manual.

- Set an auto-lock using the command **configuration mode exclusive auto** from CONFIGURATION mode. When you set an 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 a manual lock using the command **configure terminal lock** from CONFIGURATION mode. When you configure a manual lock, which is the default, you must enter this command time you want to enter CONFIGURATION mode and deny access to others.

Figure 4-9. Locking CONFIGURATION mode

```
Force10(conf)#configuration mode exclusive auto
Force10(conf)#exit
3d23h35m: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by  console

Force10#config
! Locks configuration mode exclusively.
Force10(conf)#
```

If another user attempts to enter CONFIGURATION mode while a lock is in place, Message 1 appears on their terminal.

Message 1 CONFIGURATION mode Locked Error

```
% 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, Message 2 appears on their terminal.

Message 2 Cannot Lock CONFIGURATION mode Error

```
% 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.

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 command **show configuration lock** 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 command **clear** from EXEC Privilege mode. If you clear a console session, the user is returned to EXEC mode.

Recovering from a Forgotten Password on the S55

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.

If you forget your password:

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

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

Recovering from a Forgotten Enable Password on the S55

If you forget the enable password:

Step	Task	Command Syntax	Command Mode
1	Log onto the system via console.		
2	Power-cycle the chassis by switching off all of the power modules and then switching them back on.		
3	Press any key to abort the boot process. You enter uBoot immediately, as indicated by the => prompt.	press any key	(during bootup)
4	Set the system parameters to ignore the enable password when the system reloads.	**setenv enablepwdignore true**	uBoot
5	Reload the system.	**reset**	uBoot
6	Configure a new enable password.	**enable {secret | password}**	CONFIGURATION
7	Save the running-config to the startup-config.	**copy running-config startup-config**	EXEC Privilege

Recovering from a Failed Start on the S55

A system that does not start correctly might be attempting to boot from a corrupted FTOS image or from a mis-specified location. In that 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 below. For details on the **setenv** command, its supporting commands, and other commands that can help recover from a failed start, see the BuBoot chapter in the *FTOS Command Line Reference for the S55*.

Step	Task	Command Syntax	Command Mode
1	Power-cycle the chassis (pull the power cord and reinsert it).		
2	Press any key to abort the boot process. You enter uBoot immediately, as indicated by the => prompt.	press any key	(during bootup)
3	Assign the new location to the FTOS image to be used when the system reloads.	**setenv [primary_image f10boot** *location* **| secondary_image f10boot** *location* **| default_image f10boot** *location***]**	uBoot
4	Assign an IP address to the Management Ethernet interface.	**setenv ipaddr** *address*	uBoot
5			
6	Assign an IP address as the default gateway for the system.	**setenv gatewayip** *address*	uBoot
7	Reload the system.	**reset**	uBoot

5

802.1ag

802.1ag is available only on platform: S

Ethernet Operations, Administration, and Maintenance (OAM) is a set of tools used to install, monitor, troubleshoot and manage Ethernet infrastructure deployments. Ethernet OAM consists of three main areas:

1. Service Layer OAM: IEEE 802.1ag Connectivity Fault Management (CFM)
2. Link Layer OAM: IEEE 802.3ah OAM
3. Ethernet Local management Interface (MEF-16 E-LMI)

Ethernet CFM

Ethernet CFM is an end-to-end per-service-instance Ethernet OAM scheme which enables: proactive connectivity monitoring, fault verification, and fault isolation.

The service-instance with regard to OAM for Metro/Carrier Ethernet is a VLAN. This service is sold to an end-customer by a network service provider. Typically the service provider contracts with multiple network operators to provide end-to-end service between customers. For end-to-end service between customer switches, connectivity must be present across the service provider through multiple network operators.

Layer 2 Ethernet networks usually cannot be managed with IP tools such as ICMP Ping and IP Traceroute. Traditional IP tools often fail because:

- there are complex interactions between various Layer 2 and Layer 3 protocols such as STP, LAG, VRRP and ECMP configurations.
- Ping and traceroute are not designed to verify data connectivity in the network and within each node in the network (such as in the switching fabric and hardware forwarding tables).
- when networks are built from different operational domains, access controls impose restrictions that cannot be overcome at the IP level, resulting in poor fault visibility. There is a need for hierarchical domains that can be monitored and maintained independently by each provider or operator.
- routing protocols choose a subset of the total network topology for forwarding, making it hard to detect faults in links and nodes that are not included in the active routing topology. This is made more complex when using some form of Traffic Engineering (TE) based routing.
- network and element discovery and cataloging is not clearly defined using IP troubleshooting tools.

There is a need for Layer 2 equivalents to manage and troubleshoot native Layer 2 Ethernet networks. With these tools, you can identify, isolate, and repair faults quickly and easily, which reduces operational cost of running the network. OAM also increases availability and reduces mean time to recovery, which allows for tighter service level agreements, resulting in increased revenue for the service provider.

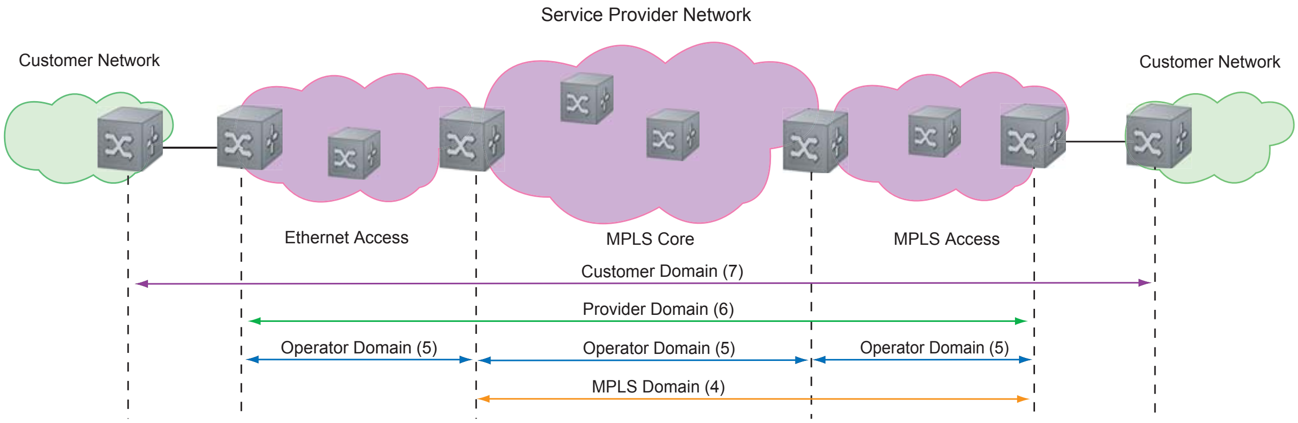
In addition to providing end-to-end OAM in native Layer 2 Ethernet Service Provider/Metro networks, you can also use CFM to manage and troubleshoot any Layer 2 network including enterprise, datacenter, and cluster networks.

Maintenance Domains

Connectivity Fault Management (CFM) divides a network into hierarchical maintenance domains, as shown in Figure 5-1.

A CFM maintenance domain is a management space on a network that is owned and operated by a single management entity. The network administrator assigns a unique maintenance level (0 to 7) to each domain to define the hierarchical relationship between domains. Domains can touch or nest but cannot overlap or intersect as that would require management by multiple entities.

Figure 5-1. OAM Domains



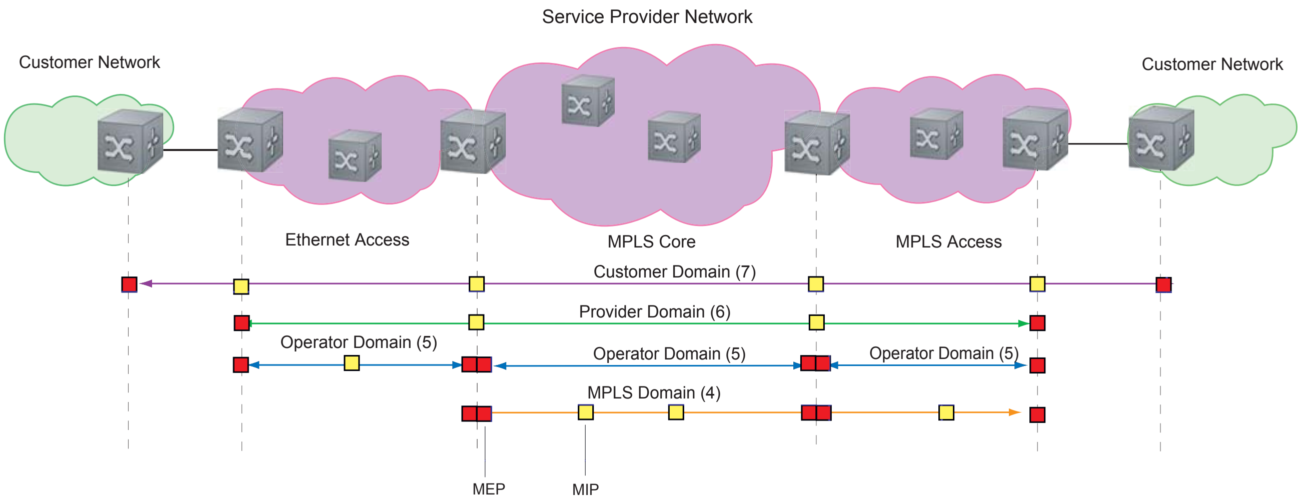
Maintenance Points

Domains are comprised of logical entities called Maintenance Points. A maintenance point is an interface demarcation that confines CFM frames to a domain. There are two types of maintenance points:

- **Maintenance End Points (MEPs)**: a logical entity that marks the end-point of a domain
- **Maintenance Intermediate Points (MIPs)**: a logical entity configured at a port of a switch that is an intermediate point of a Maintenance Entity (ME). An ME is a point-to-point relationship between two MEPs within a single domain. MIPs are internal to a domain, not at the boundary, and respond to CFM only when triggered by linktrace and loopback messages. MIPs can be configured to snoop Continuity Check Messages (CCMs) to build a MIP CCM database.

These roles define the relationships between all devices so that each device can monitor the layers under its responsibility. Maintenance points drop all lower-level frames and forward all higher-level frames.

Figure 5-2. Maintenance Points



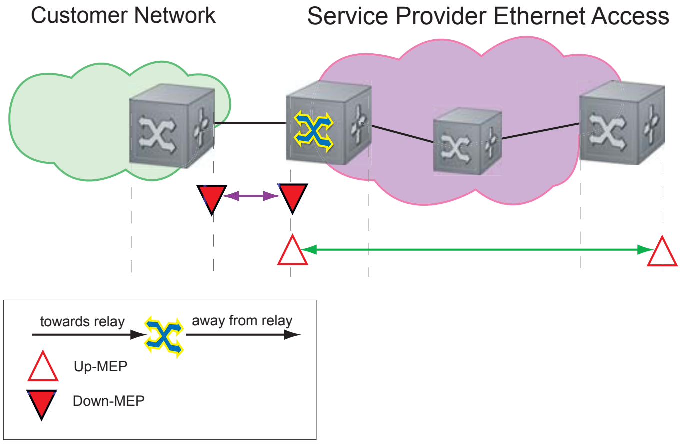
Maintenance End Points

A Maintenance End Point (MEP) is a logical entity that marks the end-point of a domain. There are two types of MEPs defined in 802.1ag for an 802.1 bridge:

- **Up-MEP**: monitors the forwarding path internal to an bridge on the customer or provider edge; on Dell Force10 systems the internal forwarding path is effectively the switch fabric and forwarding engine.
- **Down-MEP**: monitors the forwarding path external another bridge.

Configure Up- MEPs on ingress ports, ports that send traffic towards the bridge relay. Configure Down-MEPs on egress ports, ports that send traffic away from the bridge relay.

Figure 5-3. Up-MEP versus Down-MEP



Implementation Information

- Since the S-Series has a single MAC address for all physical/LAG interfaces, only one MEP is allowed per MA (per VLAN or per MD level).

Configure CFM

Configuring CFM is a five-step process:

1. Configure the ecfmacl CAM region using the **cam-acl** command.
2. Enable Ethernet CFM. See page 73.
3. Create a Maintenance Domain. See page 73.
4. Create a Maintenance Association. See page 74.
5. Create Maintenance Points. See page 74.
6. Use CFM tools:
   a Continuity Check Messages on page 77
   b Loopback Message and Response on page 78
   c Linktrace Message and Response on page 78

Related Configuration Tasks

- Enable CFM SNMP Traps. on page 80
- Display Ethernet CFM Statistics on page 81

Enable Ethernet CFM

Task	Command Syntax	Command Mode
Spawn the CFM process. No CFM configuration is allowed until the CFM process is spawned.	**ethernet cfm**	CONFIGURATION
Disable Ethernet CFM without stopping the CFM process.	disable	ETHERNET CFM

Create a Maintenance Domain

Connectivity Fault Management (CFM) divides a network into hierarchical maintenance domains, as shown in Figure 5-1.

Step	Task	Command Syntax	Command Mode
1	Create maintenance domain.	**domain** *name* **md-level** *number* Range: 0-7	ETHERNET CFM
2	Display maintenance domain information.	**show ethernet cfm domain** [*name* | **brief**]	EXEC Privilege

```
Force10# show ethernet cfm domain

Domain Name: customer
Level: 7
Total Service: 1
    Services
              MA-Name           VLAN         CC-Int        X-CHK Status

              My_MA             200            10s            enabled

Domain Name: praveen
Level: 6
Total Service: 1
    Services
              MA-Name           VLAN         CC-Int        X-CHK Status

              Your_MA           100            10s            enabled
```

Create a Maintenance Association

A Maintenance Association MA is a subdivision of an MD that contains all managed entities corresponding to a single end-to-end service, typically a VLAN. An MA is associated with a VLAN ID.

Task	Command Syntax	Command Mode
Create maintenance association.	**service** *name* **vlan** *vlan-id*	ECFM DOMAIN

Create Maintenance Points

Domains are comprised of logical entities called Maintenance Points. A maintenance point is a interface demarcation that confines CFM frames to a domain. There are two types of maintenance points:

- **Maintenance End Points (MEPs)**: a logical entity that marks the end-point of a domain
- **Maintenance Intermediate Points (MIPs)**: a logical entity configured at a port of a switch that constitutes intermediate points of an Maintenance Entity (ME). An ME is a point-to-point relationship between two MEPs within a single domain.

These roles define the relationships between all devices so that each device can monitor the layers under its responsibility.

Create a Maintenance End Point

A Maintenance End Point (MEP) is a logical entity that marks the end-point of a domain. There are two types of MEPs defined in 802.1ag for an 802.1 bridge:

- **Up-MEP**: monitors the forwarding path internal to an bridge on the customer or provider edge; on Dell Force10 systems the internal forwarding path is effectively the switch fabric and forwarding engine.
- **Down-MEP**: monitors the forwarding path external another bridge.

Configure Up- MEPs on ingress ports, ports that send traffic towards the bridge relay. Configure Down-MEPs on egress ports, ports that send traffic away from the bridge relay.

Task	Command Syntax	Command Mode
Create an MEP.	**ethernet cfm mep** {**up-mep | down-mep**} **domain** {*name* | *level* } **ma-name** *name* **mepid** *mep-id* Range: 1-8191	INTERFACE
Display configured MEPs and MIPs.	**show ethernet cfm maintenance-points local** [**mep** | **mip**]	EXEC Privilege

Task	Command Syntax	Command Mode

```
Force10#show ethernet cfm maintenance-points local mep
 -------------------------------------------------------------------------------
MPID          Domain Name      Level   Type          Port           CCM-Status
                MA Name        VLAN    Dir            MAC
-------------------------------------------------------------------------------

 100                     cfm0     7      MEP         Gi 4/10            Enabled
                      test0      10    DOWN     00:01:e8:59:23:45

 200                     cfm1     6      MEP         Gi 4/10            Enabled
                      test1      20    DOWN     00:01:e8:59:23:45

 300                     cfm2     5      MEP         Gi 4/10            Enabled
                      test2      30    DOWN     00:01:e8:59:23:45
```

Create a Maintenance Intermediate Point

Maintenance Intermediate Point (MIP) is a logical entity configured at a port of a switch that constitutes intermediate points of an Maintenance Entity (ME). An ME is a point-to-point relationship between two MEPs within a single domain. An MIP is not associated with any MA or service instance, and it belongs to the entire MD.

Task	Command Syntax	Command Mode
Create an MIP.	**ethernet cfm mip domain** {*name* | *level* } **ma-name** *name*	INTERFACE
Display configured MEPs and MIPs.	**show ethernet cfm maintenance-points local** [**mep** | **mip**]	EXEC Privilege

```
Force10#show ethernet cfm maintenance-points local mip
 -------------------------------------------------------------------------------
MPID          Domain Name      Level   Type          Port           CCM-Status
                MA Name        VLAN    Dir            MAC
-------------------------------------------------------------------------------

   0                 service1     4      MIP          Gi 0/5           Disabled
                     My_MA       3333    DOWN     00:01:e8:0b:c6:36

   0                 service1     4      MIP          Gi 0/5           Disabled
                     Your_MA     3333      UP     00:01:e8:0b:c6:36
```

MP Databases

CFM maintains two MP databases:

- **MEP Database (MEP-DB)**: Every MEP must maintain a database of all other MEPs in the MA that have announced their presence via CCM.

- **MIP Database (MIP-DB)**: Every MIP must maintain a database of all other MEPs in the MA that have announced their presence via CCM

<table>
<tr><th>Task</th><th>Command Syntax</th><th>Command Mode</th></tr>
<tr><td>Display the MEP Database.</td><td><b>show ethernet cfm maintenance-points remote detail</b> [<b>active</b> | <b>domain</b> {<i>level</i> | <i>name</i>} | <b>expired</b> | <b>waiting</b>]</td><td>EXEC Privilege</td></tr>
<tr><td colspan="3"><pre>
Force10#show ethernet cfm maintenance-points remote detail

MAC Address: 00:01:e8:58:68:78
Domain Name: cfm0
MA Name: test0
Level: 7
VLAN: 10
MP ID: 900
Sender Chassis ID: Force10
MEP Interface status: Up
MEP Port status: Forwarding
Receive RDI: FALSE
MP Status: Active
</pre></td></tr>
<tr><td>Display the MIP Database.</td><td><b>show ethernet cfm mipdb</b></td><td>EXEC Privilege</td></tr>
</table>

MP Database Persistence

Task	Command Syntax	Command Mode
Set the amount of time that data from a missing MEP is kept in the Continuity Check Database.	**database hold-time** *minutes* Default: 100 minutes Range: 100-65535 minutes	ECFM DOMAIN

Continuity Check Messages

Continuity Check Messages (CCM) are periodic hellos used to:

- discover MEPs and MIPs within a maintenance domain
- detect loss of connectivity between MEPs
- detect mis-configuration, such as VLAN ID mismatch between MEPs
- to detect unauthorized MEPs in a maintenance domain

Continuity Check Messages (CCM) are multicast Ethernet frames sent at regular intervals from each MEP. They have a destination address based on the MD level (01:80:C2:00:00:3X where X is the MD level of the transmitting MEP from 0 to 7). All MEPs must listen to these multicast MAC addresses and process these messages. MIPs may optionally processes the CCM messages originated by MEPs and construct a MIP CCM database.

MEPs and MIPs filter CCMs from higher and lower domain levels as described in Table 5-1.

Table 5-1. Continuity Check Message Processing

Frames at	Frames from	UP-MEP Action	Down-MEP Action	MIP Action
Less than my level	Bridge-relay side or Wire side	Drop	Drop	Drop
My level	Bridge-relay side	Consume	Drop	Add to MIP-DB and forward
	Wire side	Drop	Consume	
Greater than my level	Bridge-relay side or Wire side	Forward	Forward	Forward

All the remote MEPs in the maintenance domain are defined on each MEP. Each MEP then expects a periodic CCM from the configured list of MEPs. A connectivity failure is then defined as:

1. Loss of 3 consecutive CCMs from any of the remote MEP, which indicates a network failure
2. Reception of a CCM with an incorrect CCM transmission interval, which indicates a configuration error.
3. Reception of CCM with an incorrect MEP ID or MAID, which indicates a configuration or cross-connect error. This could happen when different VLANs are cross-connected due to a configuration error.
4. Reception of a CCM with an MD level lower than that of the receiving MEP, which indicates a configuration or cross-connect error.
5. Reception of a CCM containing a port status/interface status TLV, which indicates a failed bridge or aggregated port.

The Continuity Check protocol sends fault notifications (Syslogs, and SNMP traps if enabled) whenever any of the above errors are encountered.

Enable CCM

Step	Task	Command Syntax	Command Mode
1	Enable CCM.	**no ccm disable** Default: Disabled	ECFM DOMAIN
2	Configure the transmit interval (mandatory). The interval specified applies to all MEPs in the domain.	ccm transmit-interval *seconds* Default: 10 seconds	ECFM DOMAIN

Enable Cross-checking

Task	Command Syntax	Command Mode
Enable cross-checking.	**mep cross-check enable** Default: Disabled	ETHERNET CFM
Start the cross-check operation for an MEP.	mep cross-check *mep-id*	ETHERNET CFM
Configure the amount of time the system waits for a remote MEP to come up before the cross-check operation is started.	**mep cross-check start-delay** *number*	ETHERNET CFM

Loopback Message and Response

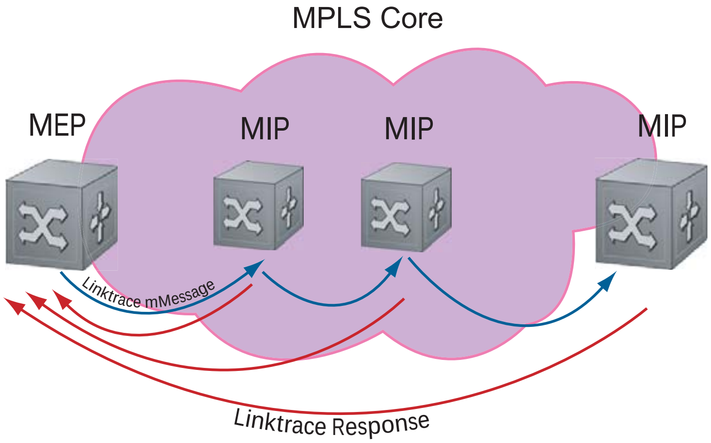
Loopback Message and Response (LBM, LBR), also called Layer 2 Ping, is an administrative echo transmitted by MEPs to verify reachability to another MEP or MIP within the maintenance domain. LBM and LBR are unicast frames.

Task	Command Syntax	Command Mode
Send a Loopback message.	**ping ethernet domain** *name* **ma-name** *ma-name* **remote** {*mep-id* | **mac-addr** *mac-address*} **source** {*mep-id* | **port** *interface*}	EXEC Privilege

Linktrace Message and Response

Linktrace Message and Response (LTM, LTR), also called Layer 2 Traceroute, is an administratively sent multicast frames transmitted by MEPs to track, hop-by-hop, the path to another MEP or MIP within the maintenance domain. All MEPs and MIPs in the same domain respond to an LTM with a unicast LTR. Intermediate MIPs forward the LTM toward the target MEP.

Figure 5-4. Linktrace Message and Response



Link trace messages carry a unicast target address (the MAC address of an MIP or MEP) inside a multicast frame. The destination group address is based on the MD level of the transmitting MEP (01:80:C2:00:00:3[8 to F]). The MPs on the path to the target MAC address reply to the LTM with an LTR, and relays the LTM towards the target MAC until the target MAC is reached or TTL equals 0.

Task	Command Syntax	Command Mode
Send a Linktrace message. Since the LTM is a Multicast message sent to the entire ME, there is no need to specify a destination.	**traceroute ethernet domain**	EXEC Privilege

Link Trace Cache

After a Link Trace command is executed, the trace information can be cached so that you can view it later without retracing.

Task	Command Syntax	Command Mode
Enable Link Trace caching.	**traceroute cache**	CONFIGURATION
Set the amount of time a trace result is cached.	**traceroute cache hold-time** *minutes* Default: 100 minutes Range: 10-65535 minutes	ETHERNET CFM
Set the size of the Link Trace Cache.	**traceroute cache size** *entries* Default: 100 Range: 1 - 4095 entries	ETHERNET CFM
Display the Link Trace Cache.	**show ethernet cfm traceroute-cache**	EXEC Privilege

Task	Command Syntax	Command Mode

```
Force10#show ethernet cfm traceroute-cache

Traceroute to 00:01:e8:52:4a:f8 on Domain Customer2, Level 7, MA name Test2 with
VLAN 2

 ------------------------------------------------------------------------------
  Hops              Host               IngressMAC     Ingr Action   Relay Action
                 Next Host             Egress MAC     Egress Action  FWD Status
------------------------------------------------------------------------------

   4      00:00:00:01:e8:53:4a:f8   00:01:e8:52:4a:f8  IngOK            RlyHit
         00:00:00:01:e8:52:4a:f8                                   Terminal MEP
```

Delete all Link Trace Cache entries.	**clear ethernet cfm traceroute-cache**	EXEC Privilege

Enable CFM SNMP Traps.

Task	Command Syntax	Command Mode
Enable SNMP trap messages for Ethernet CFM.	**snmp-server enable traps ecfm**	CONFIGURATION

A Trap is sent only when one of the five highest priority defects occur, as shown in Table 5-2.

Table 5-2. ECFM SNMP Traps

Cross-connect defect	`%ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000`
Error-CCM defect	`%ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000`
MAC Status defect	`%ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain provider at Level 4 VLAN 3000`
Remote CCM defect	`%ECFM-5-ECFM_REMOTE_ALARM: Remote CCM Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000`
RDI defect	`%ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000`

Three values are giving within the trap messages: MD Index, MA Index, and MPID. You can reference these values against the output of **show ethernet cfm domain** and **show ethernet cfm maintenance-points local mep**.

```
Force10#show ethernet cfm maintenance-points local mep
 -------------------------------------------------------------------------------
MPID          Domain Name      Level   Type          Port           CCM-Status
                MA Name        VLAN    Dir            MAC
-------------------------------------------------------------------------------

 100                     cfm0      7      MEP         Gi 4/10            Enabled
                      test0      10     DOWN     00:01:e8:59:23:45

Force10(conf-if-gi-0/6)#do show ethernet cfm domain

Domain Name: My_Name
MD Index: 1
Level: 0
Total Service: 1
    Services
MA-Index        MA-Name           VLAN         CC-Int        X-CHK Status

      1         test               0            1s             enabled

Domain Name: Your_Name
MD Index: 2
Level: 2
Total Service: 1
    Services
MA-Index        MA-Name           VLAN         CC-Int        X-CHK Status

      1         test             100            1s             enabled
```

Display Ethernet CFM Statistics

Task	Command Syntax	Command Mode
Display MEP CCM statistics.	**show ethernet cfm statistics** [**domain** {*name* | *level*} **vlan-id** *vlan-id* **mpid** *mpid*	EXEC Privilege

```
Force10#  show ethernet cfm statistics

Domain Name: Customer
Domain Level: 7
MA Name: My_MA
MPID: 300

     CCMs:
       Transmitted:                  1503     RcvdSeqErrors:                0
     LTRs:
       Unexpected Rcvd:                 0
     LBRs:
       Received:                        0     Rcvd Out Of Order:           0
       Received Bad MSDU:               0
       Transmitted:                     0
```

Task	Command Syntax	Command Mode
Display CFM statistics by port.	**show ethernet cfm port-statistics** [**interface**]	EXEC Privilege

```
Force10#show ethernet cfm port-statistics interface gigabitethernet 0/5
Port statistics for port: Gi 0/5
==================================

RX Statistics
=============
Total CFM Pkts 75394 CCM Pkts 75394
LBM Pkts 0 LTM Pkts 0
LBR Pkts 0 LTR Pkts 0
Bad CFM Pkts 0 CFM Pkts Discarded 0
CFM Pkts forwarded 102417

TX Statistics
=============
Total CFM Pkts 10303 CCM Pkts 0
LBM Pkts 0 LTM Pkts 3
LBR Pkts 0 LTR Pkts 0
```

6

802.1X

802.1X is supported on platforms: C E S

Protocol Overview

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) via a mandatory intermediary network access device, in this case, a Dell Force10 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.

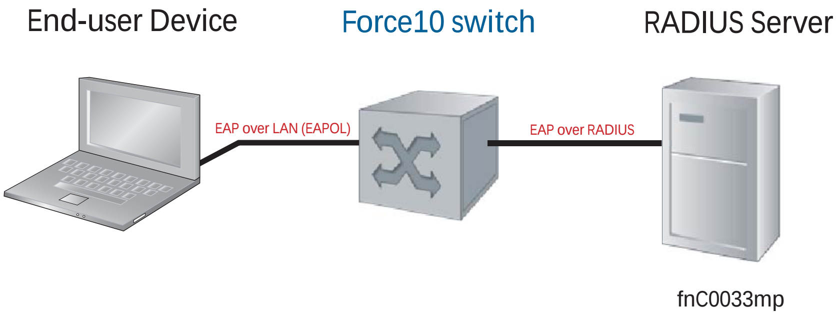
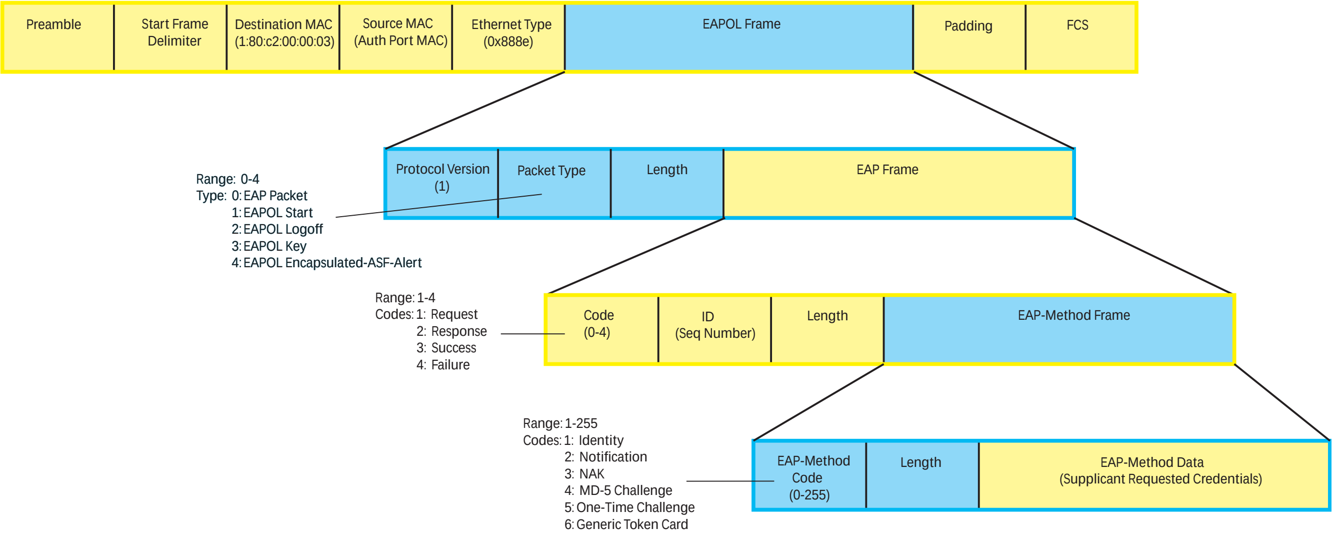


Figure 6-1 and Figure show how EAP frames are encapsulated in Ethernet and Radius frames.

Note: FTOS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.

Figure 6-1. EAPOL Frame Format



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 port is authorized by the authenticator. It can only communicate with the authenticator in response to 802.1X requests.
- The device with which the supplicant communicates is the **authenticator**. The authenicator 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 Force10 switch is the authenticator.
- The **authentication-server** selects the authentication method, verifies the information provided by the supplicant, 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 Force10 switches place 802.1X-enabled ports in the unauthorized state by default.

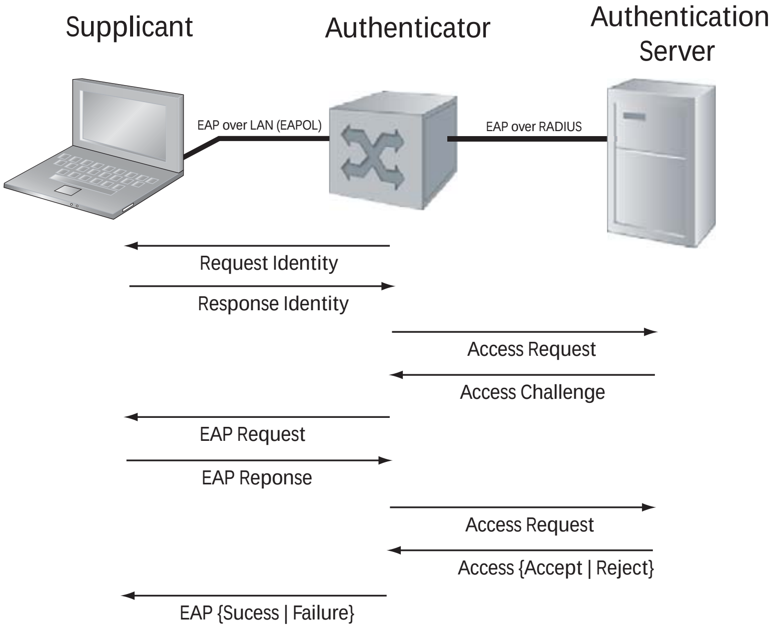
The Port-authentication Process

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. The Access-Challenge is request 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.
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 and Access-Reject frame. The port state remains unauthorized, and the authenticator forwards EAP Failure frame.

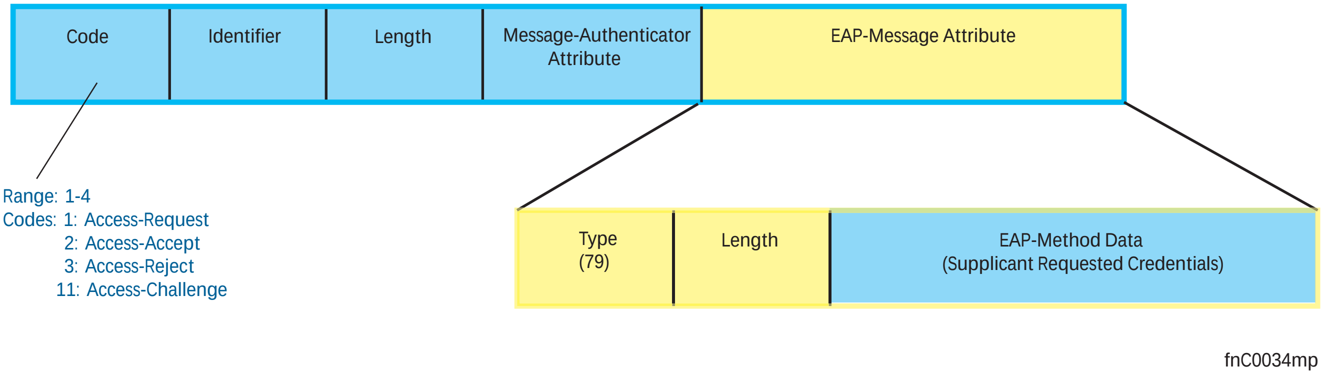
Figure 6-2. 802.1X Authentication Process



EAP over RADIUS

802.1X uses RADIUS to shuttle EAP packets between the authenticator and the authentication server, as defined in RFC 3579. EAP messages are encapsulated in RADIUS packets as a type of *attribute* in Type, Length, Value (TLV) format. The Type value for EAP messages is 79.

Figure 6-3. RADIUS Frame Format



RADIUS Attributes for 802.1 Support

Dell Force10 systems includes 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. See page 87.
2. Enable 802.1X on an interface. See page 87.

Related Configuration Tasks

- Configuring Request Identity Re-transmissions on page 90
- Configuring Port-control on page 92
- Re-authenticating a Port on page 93
- Configuring Timeouts on page 94
- Configuring a Guest VLAN on page 97
- Configuring an Authentication-fail VLAN on page 97

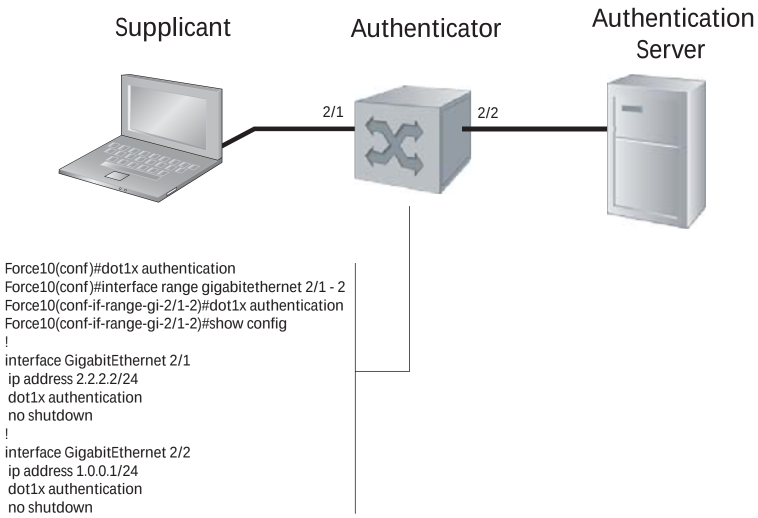
Important Points to Remember

- FTOS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.
- E-Series and C-Series support only RADIUS as the authentication server.
- 802.1X is not supported on port-channels or port-channel members.

Enabling 802.1X

802.1X must be enabled globally and at interface level.

Figure 6-4. Enabling 802.1X



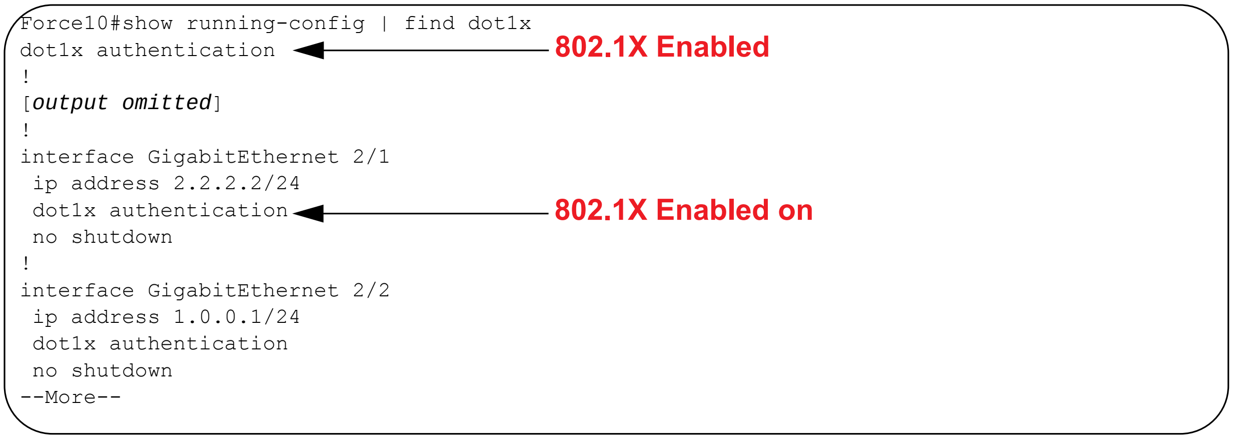
```
Force10(conf)#dot1x authentication
Force10(conf)#interface range gigabitethernet 2/1 - 2
Force10(conf-if-range-gi-2/1-2)#dot1x authentication
Force10(conf-if-range-gi-2/1-2)#show config
!
interface GigabitEthernet 2/1
 ip address 2.2.2.2/24
 dot1x authentication
 no shutdown
!
interface GigabitEthernet 2/2
 ip address 1.0.0.1/24
 dot1x authentication
 no shutdown
```

To enable 802.1X:

Step	Task	Command Syntax	Command Mode
1	Enable 802.1X globally.	**dot1x authentication**	CONFIGURATION
2	Enter INTERFACE mode on an interface or a range of interfaces.	**interface** [**range**]	INTERFACE
3	Enable 802.1X on an interface or a range of interfaces.	dot1x authentication	INTERFACE

Verify that 802.1X is enabled globally and at interface level using the command **show running-config | find dot1x** from EXEC Privilege mode, as shown in Figure 6-5.

Figure 6-5. Verifying 802.1X Global Configuration



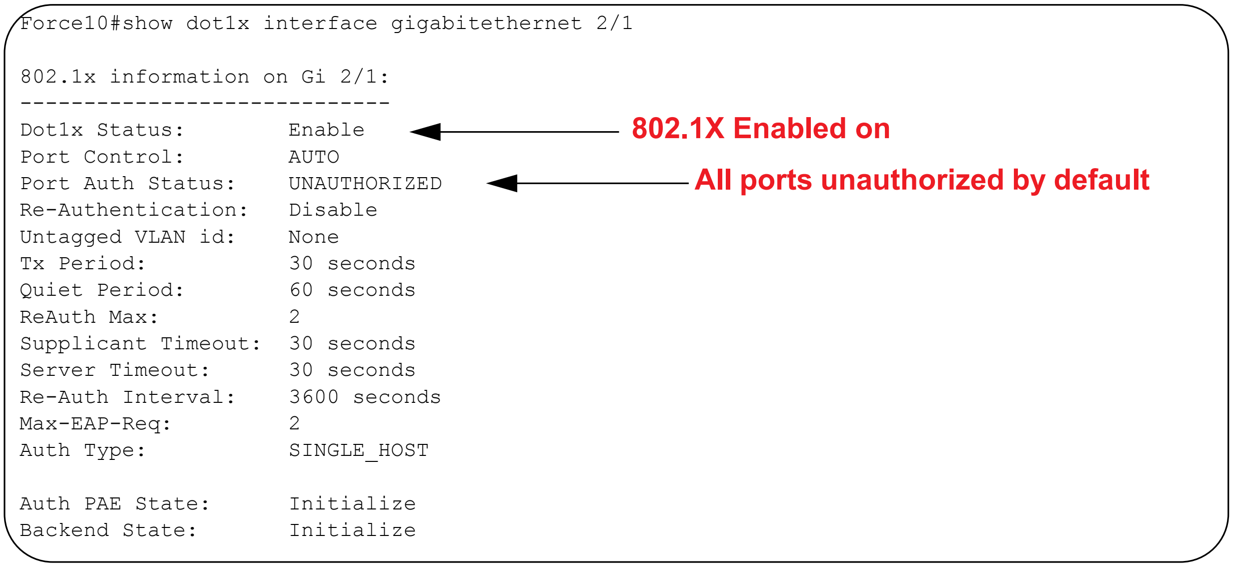
```
Force10#show running-config | find dot1x
dot1x authentication
!
[output omitted]
!
interface GigabitEthernet 2/1
 ip address 2.2.2.2/24
 dot1x authentication
 no shutdown
!
interface GigabitEthernet 2/2
 ip address 1.0.0.1/24
 dot1x authentication
 no shutdown
--More--
```

802.1X Enabled

802.1X Enabled on

View 802.1X configuration information for an interface using the command **show dot1x interface**, as shown in Figure 6-6.

Figure 6-6. Verifying 802.1X Interface Configuration



```
Force10#show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:          Enable
Port Control:          AUTO
Port Auth Status:      UNAUTHORIZED
Re-Authentication:     Disable
Untagged VLAN id:      None
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
Auth Type:             SINGLE_HOST

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

802.1X Enabled on

All ports unauthorized by default

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; the supplicant might have been booting when the request arrived, or there might be a physical layer problem.

To configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame:

Step	Task	Command Syntax	Command Mode
1	Configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame.	**dot1x tx-period** *number* Range: 1-31536000 (1 year) Default: 30	INTERFACE

To configure a maximum number of Request Identity re-transmissions:

Step	Task	Command Syntax	Command Mode
1	Configure a maximum number of times that a Request Identity frame can be re-transmitted by the authenticator.	**dot1x max-eap-req** *number* Range: 1-10 Default: 2	INTERFACE

Figure 6-7 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 this period can be configured.

Note: The quiet period (**dot1x quiet-period**) is an transmit interval for after a failed authentication where as the Request Identity Re-transmit interval (**dot1x tx-period**) is for an unresponsive supplicant.

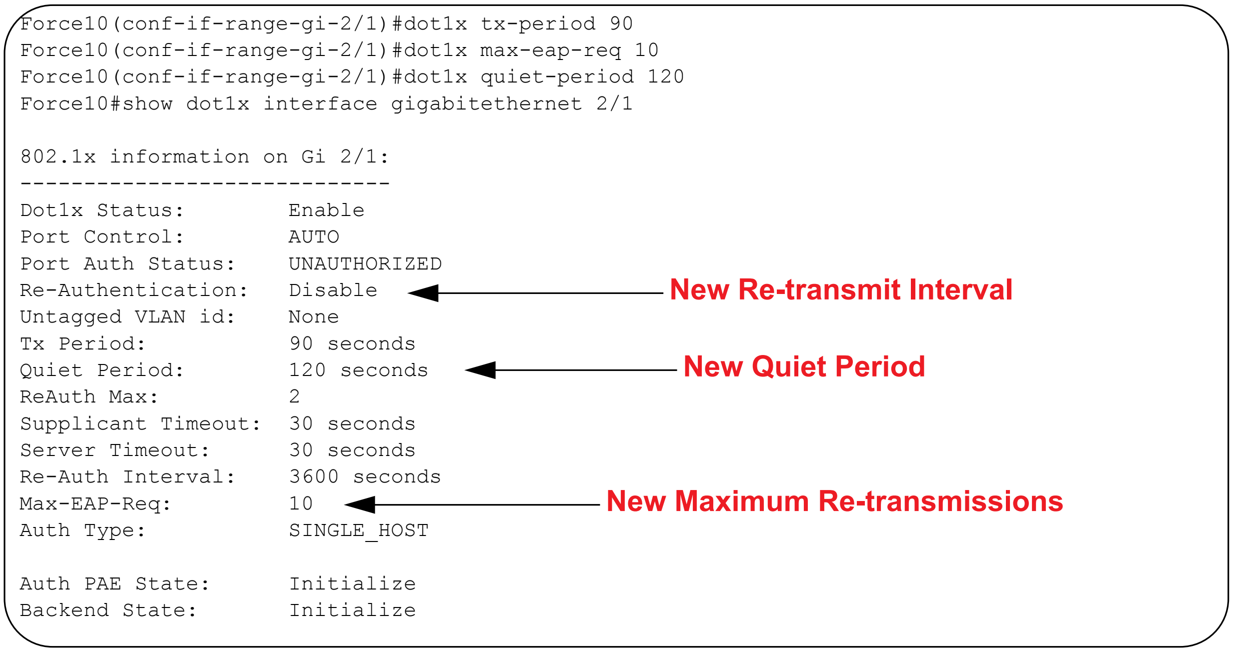
To configure the quiet period after a failed authentication:

Step	Task	Command Syntax	Command Mode
1	Configure the amount of time that the authenticator waits to re-transmit a Request Identity frame after a failed authentication.	**dot1x quiet-period** *seconds* Range: 1-65535 Default: 60	INTERFACE

Figure 6-7 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

Figure 6-7. Configuring a Request Identity Re-transmissions



```
Force10(conf-if-range-gi-2/1)#dot1x tx-period 90
Force10(conf-if-range-gi-2/1)#dot1x max-eap-req 10
Force10(conf-if-range-gi-2/1)#dot1x quiet-period 120
Force10#show dot1x interface gigabitethernet 2/1

802.1x information on Gi 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** is 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.

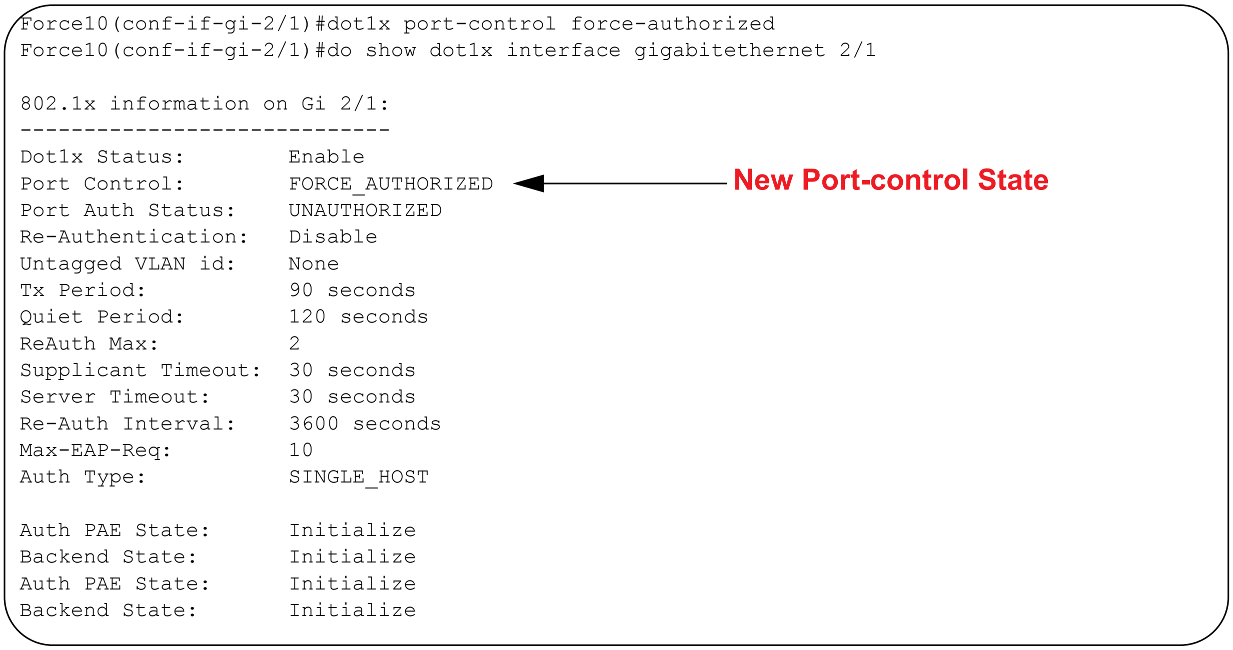
- **ForceUnauthorized** an unauthorized state. A device connected to a port in this state is never subjected to the authentication process and is not allowed to communicate on the network. Placing the port in this state is the same as shutting down the port. Any attempt by the supplicant to initiate authentication is ignored.
- **Auto** is an unauthorized state by default. A device connected to this port is 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 place a port in one of these three states:

Step	Task	Command Syntax	Command Mode
1	Place a port in the ForceAuthorized, ForceUnauthorized, or Auto state.	**dot1x port-control** {**force-authorized** | **force-unauthorized** | **auto**} Default: auto	INTERFACE

Figure 6-8 shows configuration information for a port that has been force-authorized.

Figure 6-8. Configuring Port-control



```
Force10(conf-if-gi-2/1)#dot1x port-control force-authorized
Force10(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
```

New Port-control State

Re-authenticating a Port

Periodic Re-authentication

After the supplicant has been authenticated, and the port has been authorized, the authenticator can be configured to re-authenticates the supplicant periodically. If re-authentication is enabled, the supplicant is required to re-authenticate every 3600 seconds, but this interval can be configured. A maximum number of re-authentications can be configured as well.

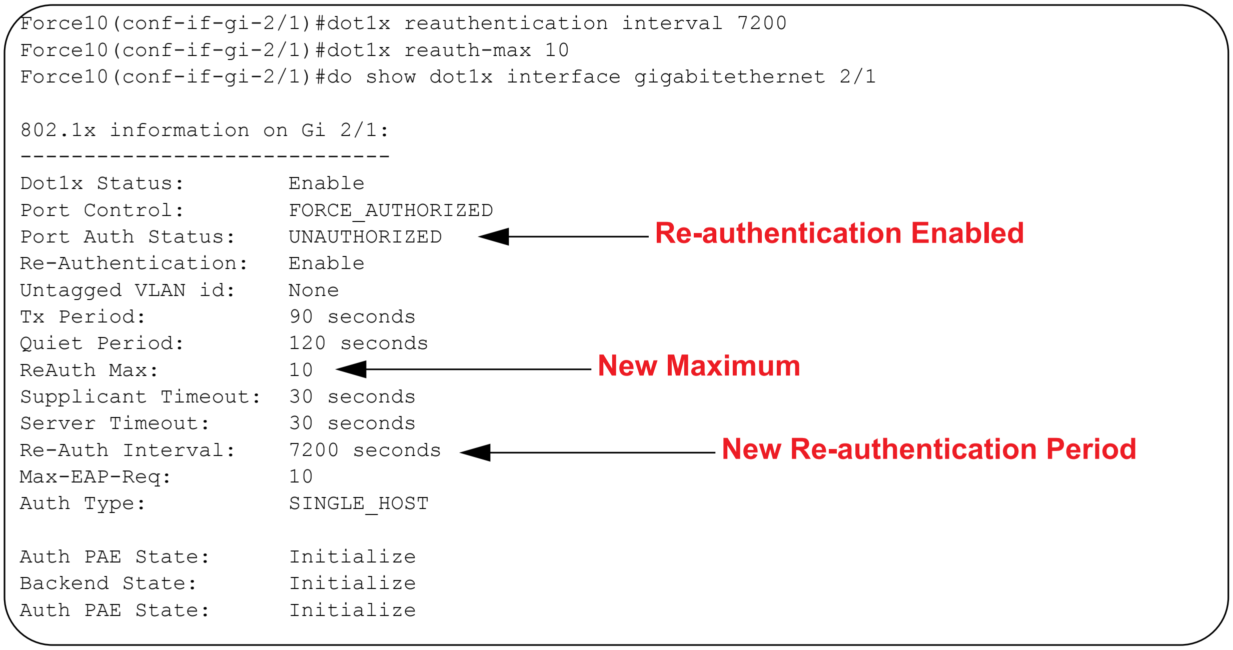
To configure a re-authentication or a re-authentication period:

Step	Task	Command Syntax	Command Mode
1	Configure the authenticator to periodically re-authenticate the supplicant.	**dot1x reauthentication** [**interval**] *seconds* Range: 1-65535 Default: 60	INTERFACE

To configure a maximum number of re-authentications:

Step	Task	Command Syntax	Command Mode
1	Configure the maximum number of times that the supplicant can be reauthenticated.	**dot1x reauth-max** *number* Range: 1-10 Default: 2	INTERFACE

Figure 6-9. Configuring a Reauthentiction Period



```
Force10(conf-if-gi-2/1)#dot1x reauthentication interval 7200
Force10(conf-if-gi-2/1)#dot1x reauth-max 10
Force10(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. This amount of time that the authenticator waits for a response can be configured.

To terminate the authentication process due to an unresponsive supplicant:

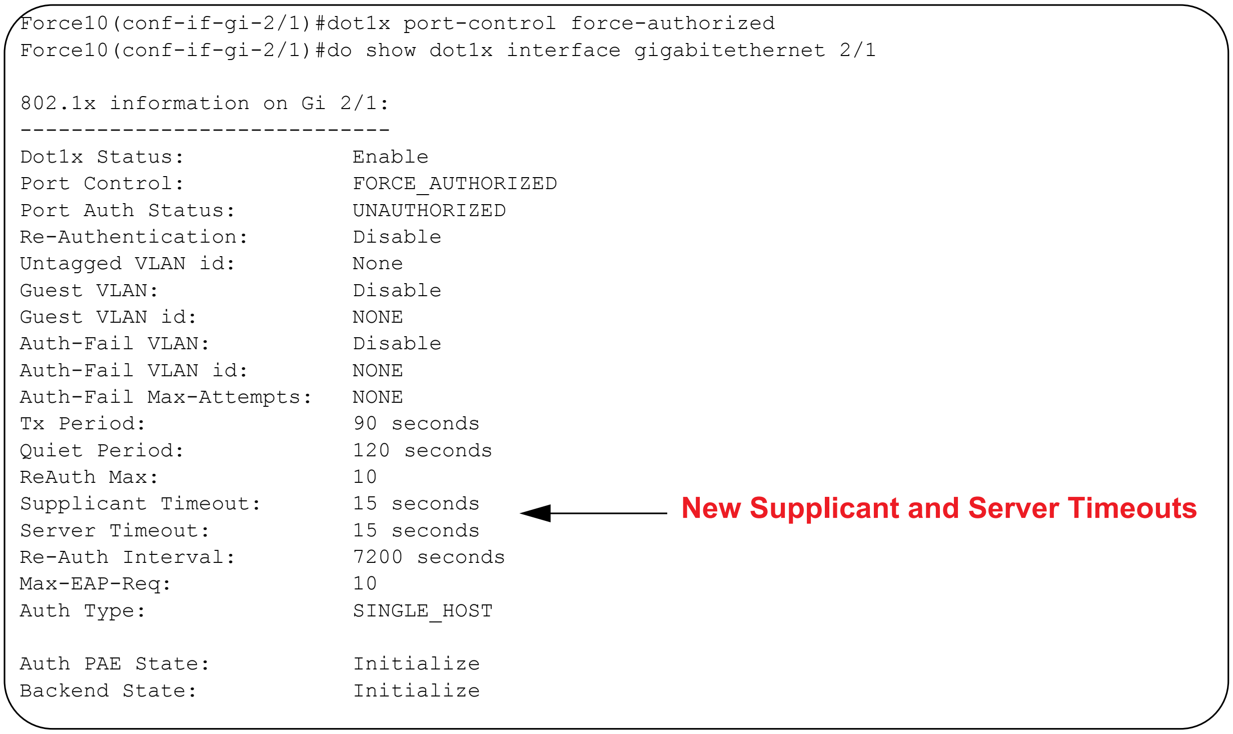
Step	Task	Command Syntax	Command Mode
1	Terminate the authentication process due to an unresponsive supplicant.	**dot1x supplicant-timeout** *seconds* Range: 1-300 Default: 30	INTERFACE

To terminate the authentication process due to an unresponsive authentication server:

Step	Task	Command Syntax	Command Mode
1	Terminate the authentication process due to an unresponsive authentication server.	**dot1x server-timeout** *seconds* Range: 1-300 Default: 30	INTERFACE

Figure 6-10 shows configuration information for a port for which the authenticator terminates the authentication process for an unresponsive supplicant or server after 15 seconds.

Figure 6-10. Configuring a Timeout



```
Force10(conf-if-gi-2/1)#dot1x port-control force-authorized
Force10(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
```

New Supplicant and Server Timeouts

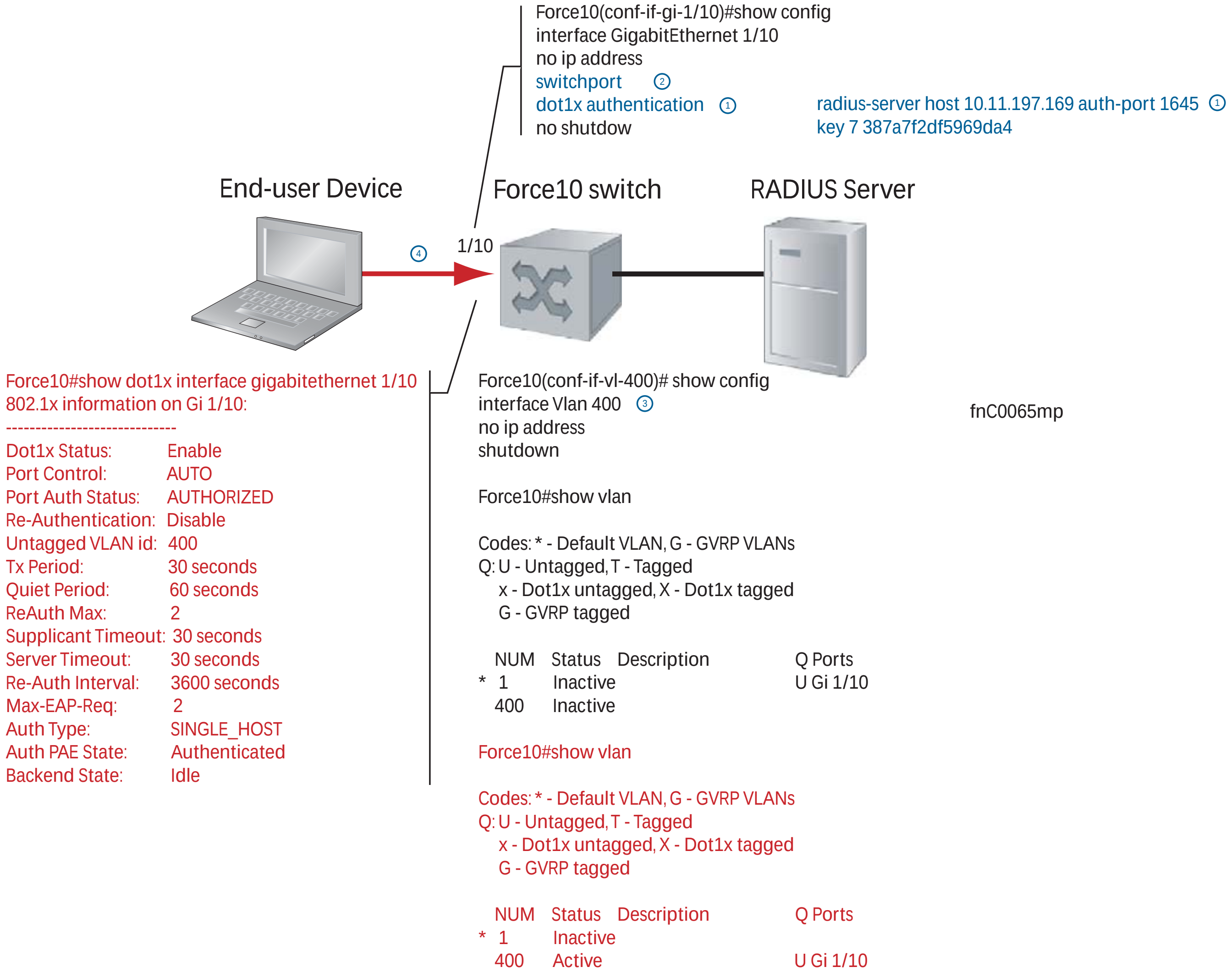
Dynamic VLAN Assignment with Port Authentication

FTOS 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 Force10 system, 2) the system forwards a RADIUS REQEST packet containing the host MAC address and ingress port number, and 3) the RADIUS server authenticates the request and returns a RADIUS ACCEPT message with the VLAN assignment using Tunnel-Private-Group-ID.

Step	Task
1	Configure 8021.x globally and at interface level (see Enabling 802.1X on page 87) along with relevant RADIUS server configurations (Figure 6-11)
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 (Figure 6-11, red text).

In Figure 6-11 shows the configuration on the Dell Force10 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 6-11. Dynamic VLAN Assignment with 802.1X



```
Force10(conf-if-gi-1/10)#show config
interface GigabitEthernet 1/10
no ip address
switchport  ②
dot1x authentication  ①
no shutdow
```

```
radius-server host 10.11.197.169 auth-port 1645  ①
key 7 387a7f2df5969da4
```

```
Force10#show dot1x interface gigabitethernet 1/10
802.1x information on Gi 1/10:
-----------------------------
Dot1x Status:          Enable
Port Control:          AUTO
Port Auth Status:      AUTHORIZED
Re-Authentication:     Disable
Untagged VLAN id:      400
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
Auth Type:             SINGLE_HOST
Auth PAE State:        Authenticated
Backend State:         Idle
```

```
Force10(conf-if-vl-400)# show config
interface Vlan 400  ③
no ip address
shutdown

Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged

  NUM   Status    Description          Q Ports
*  1    Inactive                       U Gi 1/10
  400   Inactive

Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged

  NUM   Status    Description          Q Ports
*  1    Inactive
  400   Active                         U Gi 1/10
```

Guest and Authentication-fail VLANs

Typically, the authenticator (Dell Force10 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, then 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*, see Configuring Timeouts on page 94) the system assumes that the host does not have 802.1X capability, and and the port is placed in the Guest VLAN.

Configure a port to be placed in the Guest VLAN after failing to respond within the timeout period using the command **dot1x guest-vlan** from INTERFACE mode, as shown in Figure 6-12.

Figure 6-12. Configuring a Guest VLAN

```
Force10(conf-if-gi-1/2)#dot1x guest-vlan 200
Force10(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
 switchport
 dot1x guest-vlan 200
 no shutdown
Force10(conf-if-gi-1/2)#
```

View your configuration using the command **show config** from INTERFACE mode, as shown in Figure 6-12, or using the command **show dot1x interface** command from EXEC Privilege mode as shown in Figure 6-14.

Configuring an Authentication-fail VLAN

If the supplicant fails authentication, the authenticator re-attempts to authenticate after a specified amount of time (30 seconds by default, see Configuring a Quiet Period after a Failed Authentication on page 90). 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 command **dot1x auth-fail-vlan** from INTERFACE mode, as shown in Figure 6-13. Configure the maximum number of authentication attempts by the authenticator using the keyword **max-attempts** with this command.

Figure 6-13. Configuring an Authentication-fail VLAN

```
Force10(conf-if-gi-1/2)#dot1x auth-fail-vlan 100 max-attempts 5
Force10(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
Force10(conf-if-gi-1/2)#
```

View your configuration using the command **show config** from INTERFACE mode, as shown in Figure 6-12, or using the command **show dot1x interface** command from EXEC Privilege mode as shown in Figure 6-14.

Figure 6-14. Viewing Guest and Authentication-fail VLAN Configurations

```
Force10(conf-if-gi-2/1)#dot1x port-control force-authorized
Force10(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 Lists (ACL), Prefix Lists, and Route-maps

Access Control Lists, Prefix Lists, and Route-maps are supported on platforms: C E S

Ingress IP and MAC ACLs are supported on platforms: C E S

Egress IP and MAC ACLs are supported on platforms: E S55 S60

Overview

At their simplest, Access Control Lists (ACLs), Prefix lists, and Route-maps permit or deny traffic based on MAC and/or IP addresses. This chapter discusses implementing IP ACLs, IP Prefix lists and Route-maps. For MAC ACLS, refer to Chapter 10, Layer 2, on page 47.

An ACL is essentially a filter containing some criteria to match (examine IP, TCP, or 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 CAM size. See CAM Profiling, CAM Allocation, and CAM Optimization in this chapter for more information. Refer to Chapter 10, Content Addressable Memory, on page 219 for complete CAM profiling information.

This chapter covers the following topics:

- IP Access Control Lists (ACLs) on page 100
  - CAM Profiling, CAM Allocation, and CAM Optimization on page 100
  - Implementing ACLs on FTOS on page 103
- IP Fragment Handling on page 104
- Configure a standard IP ACL on page 106
- Configure an extended IP ACL on page 109
- Configuring Layer 2 and Layer 3 ACLs on an Interface on page 112
- Assign an IP ACL to an Interface on page 112
- Configuring Ingress ACLs on page 114
- Configuring Egress ACLs on page 115

IP Access Control Lists (ACLs)

In the Dell Force10 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 (for more information on ACL supported options see the *FTOS Command Reference*):

- 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 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 FTOS will assign numbers in the order the filters are created. The sequence numbers, whether configured or assigned by FTOS, are listed in the **show config** and **show ip accounting access-list** command display output.

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 CAM are shuffled to accommodate the new entries. Hot Lock ACLs are enabled by default and support both standard and extended ACLs and on all platforms.

Note: Hot Lock ACLs are supported for Ingress ACLs only.

CAM Profiling, CAM Allocation, and CAM Optimization

CAM Profiling is supported on platform E

User Configurable CAM Allocations are supported on platform C and S55

CAM optimization is supported on platforms C S

CAM Profiling

CAM optimization is supported on platforms E T

The default CAM profile has 1K Layer 2 ingress ACL entries. If you need more memory for Layer 2 ingress ACLs, select the profile *l2-ipv4-inacl*.

When budgeting your CAM allocations for ACLs and QoS configurations, remember that ACL and QoS rules might consume more than one CAM entry depending on complexity. For example, TCP and UDP rules with port range options might require more than one CAM entry.

The Layer 2 ACL CAM partition has sub-partitions for several types of information. Table 7-1 lists the sub-partition and the percentage of the Layer 2 ACL CAM partition that FTOS allocates to each by default.

Table 7-1. Layer 2 ACL CAM Sub-partition Sizes

Partition	% Allocated
Sysflow	6
L2ACL	14
*PVST	50
QoS	12
L2PT	13
FRRP	5

You can re-configure the amount of space, in percentage, allocated to each sub-partition. As with the IPv4Flow partition, you can configure the Layer 2 ACL partition from EXEC Privilege mode or CONFIGURATION mode.

The amount of space that you can distribute to the sub-partitions is equal to the amount of CAM space that the selected CAM profile allocates to the Layer 2 ACL partition. FTOS requires that you specify the amount of CAM space for all sub-partitions and that the sum of all sub-partitions is 100%. FTOS displays the following message if the total allocated space is not correct:

```
% Error: Sum of all regions does not total to 100%.
```

User Configurable CAM Allocation

User Configurable CAM Allocations are supported on platform C and S55

Allocate space for IPV6 ACLs on the by using the **cam-acl** command in CONFIGURATION mode.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. Note that there are 16 FP blocks, but the System Flow requires 3 blocks that cannot be reallocated. The default CAM Allocation settings on a C-Series matching are:

- L3 ACL (ipv4acl): 6
- L2 ACL(l2acl) : 5
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1

The **ipv6acl** allocation 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.

You must save the new CAM settings to the startup-config (**write-mem or copy run start**) then reload the system for the new settings to take effect.

CAM optimization

CAM optimization is supported on platforms C S

When this command is enabled, 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 will be used). When the command is disabled, the system behaves as described in this chapter.

Test CAM Usage

The test cam-usage command is supported on platforms C E S

This command applies to both IPv4 and IPv6 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. Figure 7-1 gives a sample of the output shown when executing the command. The status column indicates whether or not the policy can be enabled.

Figure 7-1. Command Example: test cam-usage (C-Series)

```
Force10#test cam-usage service-policy input TestPolicy linecard all

Linecard | Portpipe | CAM Partition | Available CAM | Estimated CAM per Port | Status
------------------------------------------------------------------------------------------
       2 |        1 | IPv4Flow      |           232 |                      0 | Allowed
       2 |        1 | IPv6Flow      |             0 |                      0 | Allowed
       4 |        0 | IPv4Flow      |           232 |                      0 | Allowed
       4 |        0 | IPv6Flow      |             0 |                      0 | Allowed
Force10#
```

Implementing ACLs on FTOS

One IP ACL can be assigned per interface with FTOS. If an IP ACL is not assigned to an interface, it is not used by the software in any other capacity.

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

If counters are enabled 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 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 would get installed in the ACL CAM on the port-pipe. The entry would look 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 would be installed for each port belonging to a port-pipe.

When you use the **log** keyword, CP processor will have to log details about the packets that match. Depending on how many packets match the log entry and at what rate, CP might become busy as it has to log these packets' details. However the other processors (RP1 and RP2) should be unaffected. This option is typically useful when debugging some problem related to control traffic. We have used this option numerous times in the field and have not encountered any problems in such usage so far.

ACL Optimization

If an access list contains duplicate entries, FTOS deletes one entry to conserve CAM space.

Standard and Extended ACLs take up the same amount of CAM space. A single ACL rule uses 2 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 command **service-queue**, FTOS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities). For example, in Figure 7-2, 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, as shown in Figure 7-2. The order can range from 0 to 254. FTOS 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 254.

Figure 7-2. Using the Order Keyword in ACLs

```
Force10(conf)#ip access-list standard acl1
Force10(config-std-nacl)#permit 20.0.0.0/8
Force10(config-std-nacl)#exit
Force10(conf)#ip access-list standard acl2
Force10(config-std-nacl)#permit 20.1.1.0/24 order 0
Force10(config-std-nacl)#exit
Force10(conf)#class-map match-all cmap1
Force10(conf-class-map)#match ip access-group acl1
Force10(conf-class-map)#exit
Force10(conf)#class-map match-all cmap2
Force10(conf-class-map)#match ip access-group acl2
Force10(conf-class-map)#exit
Force10(conf)#policy-map-input pmap
Force10(conf-policy-map-in)#service-queue 7 class-map cmap1
Force10(conf-policy-map-in)#service-queue 4 class-map cmap2
Force10(conf-policy-map-in)#exit
Force10(conf)#interface gig 1/0
Force10(conf-if-gi-1/0)#service-policy input pmap
```

IP Fragment Handling

FTOS supports a configurable option to explicitly deny IP fragmented packets, particularly second and subsequent packets. It extends the existing ACL command syntax with the **fragments** keyword for all Layer 3 rules applicable to all Layer protocols (permit/deny ip/tcp/udp/icmp).

- Both standard and extended ACLs support IP fragments.
- Second and subsequent fragments are allowed because a Layer 4 rule cannot be applied to these fragments. If the packet is to be denied eventually, the first fragment would be denied and hence the packet as a whole cannot be reassembled.
- Implementing the required rules will use a significant number of CAM entries per TCP/UDP entry.
- For IP ACL, FTOS always applies implicit deny. You do not have to configure it.
- For IP ACL, FTOS applies implicit permit for second and subsequent fragment just prior to the implicit deny.
- If an *explicit* deny is configured, the second and subsequent fragments will not hit the implicit permit rule for fragments.

- Loopback interfaces do not support ACLs using the IP fragment option. If you configure an ACL with the fragments option and apply it to a loopback interface, the command is accepted, but the ACL entries are not actually installed the offending rule in CAM.

IP fragments ACL examples

The following configuration permits all packets (both fragmented & non-fragmented) with destination IP 10.1.1.1. The second rule does not get hit at all.

```
Force10(conf)#ip access-list extended ABC
Force10(conf-ext-nacl)#permit ip any 10.1.1.1/32
Force10(conf-ext-nacl)#deny ip any 10.1.1.1./32 fragments
Force10(conf-ext-nacl)
```

To deny second/subsequent fragments, use the same rules in a different order. These ACLs deny all second & subsequent fragments with destination IP 10.1.1.1 but permit the first fragment & non fragmented packets with destination IP 10.1.1.1 .

```
Force10(conf)#ip access-list extended ABC
Force10(conf-ext-nacl)#deny ip any 10.1.1.1/32 fragments
Force10(conf-ext-nacl)#permit ip any 10.1.1.1/32
Force10(conf-ext-nacl)
```

Layer 4 ACL rules examples

In the below scenario, first fragments non-fragmented TCP packets from 10.1.1.1 with TCP destination port equal to 24 are permitted. All other fragments are denied.

```
Force10(conf)#ip access-list extended ABC
Force10(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
Force10(conf-ext-nacl)#deny ip any any fragment
Force10(conf-ext-nacl)
```

In the following, 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.

```
Force10(conf)#ip access-list extended ABC
Force10(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
Force10(conf-ext-nacl)#permit tcp host 10.1.1.1 any fragment
Force10(conf-ext-nacl)#deny ip any any fragment
Force10(conf-ext-nacl)
```

To log all the packets denied and to override the implicit deny rule and the implicit permit rule for TCP/UDP fragments, use a configuration similar to the following.

```
Force10(conf)#ip access-list extended ABC
Force10(conf-ext-nacl)#permit tcp any any fragment
Force10(conf-ext-nacl)#permit udp any any fragment
Force10(conf-ext-nacl)#deny ip any any log
Force10(conf-ext-nacl)
```

Note the following when configuring ACLs with the **fragments** keyword.

When an ACL filters packets it looks at the Fragment Offset (FO) to determine whether or not 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 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 fragment offset (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 fragment offset (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.

Configure a standard IP ACL

To configure an ACL, use commands in the IP ACCESS LIST mode and the INTERFACE mode. The following list includes the configuration tasks for IP ACLs:

For a complete listing of all commands related to IP ACLs, refer to the *FTOS Command Line Interface Reference* document.

Refer to Configure an extended IP ACL on page 109 to set up extended ACLs.

A standard IP ACL uses the source IP address as its match criterion.

To configure a standard IP ACL, use these commands in the following sequence:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list standard** *access-listname*	CONFIGURATION	Enter IP ACCESS LIST mode by naming a standard IP access list.

Step	Command Syntax	Command Mode	Purpose
2	**seq** *sequence-number* {**deny** | **permit**} {*source* [*mask*] | **any** | **host** *ip-address*} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-STD-NACL	Configure a drop or forward filter. The parameters are: • **log** and **monitor** options are supported on E-Series only.

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 or another number.

When you use the **log** keyword, CP processor 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 (Figure 226) in EXEC Privilege mode.

Figure 7-3. Command Example: show ip accounting access-list

```
Force10#show ip accounting access 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
Force10#
```

Figure 7-4 illustrates 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.

Figure 7-4. Command example: seq

```
Force10(config-std-nacl)#seq 25 deny ip host 10.5.0.0 any log
Force10(config-std-nacl)#seq 15 permit tcp 10.3.0.0 /16 any
Force10(config-std-nacl)#show config
!
ip access-list standard dilling
 seq 15 permit tcp 10.3.0.0/16 any
 seq 25 deny ip host 10.5.0.0 any log
Force10(config-std-nacl)#
```

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

If you are creating a standard ACL with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. The software assigns filters in multiples of 5.

To configure a filter without a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list standard** *access-list-name*	CONFIGURATION	Create a standard IP ACL and assign it a unique name.
2	{**deny** | **permit**} {*source* [*mask*] | **any** | **host** *ip-address*} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-STD-NACL	Configure a drop or forward IP ACL filter. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor 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.

Figure 7-5 illustrates a standard IP ACL in which the sequence numbers were assigned by the FTOS. 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 the IP ACCESS LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 7-5. Standard IP ACL

```
Force10(config-route-map)#ip access standard kigali
Force10(config-std-nacl)#permit 10.1.0.0/16
Force10(config-std-nacl)#show config
!
ip access-list standard kigali
 seq 5 permit 10.1.0.0/16
Force10(config-std-nacl)#
```

To view all configured IP ACLs, use the **show ip accounting access-list** command (Figure 229) in the EXEC Privilege mode.

Figure 7-6. Command Example: show ip accounting access-list

```
Force10#show ip accounting access example interface gig 4/12
Extended IP access list example
seq 10 deny tcp any any eq 111
 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 the 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 the 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.

Since traffic passes through the filter in the order of the filter's sequence, you can configure the extended IP ACL by first entering the IP ACCESS LIST mode and then assigning a sequence number to the filter.

Note: On E-Series ExaScale systems, TCP ACL flags are not supported in an extended ACL with IPv6 microcode. An error message is shown if IPv6 microcode is configured and an ACL is entered with a TCP filter included.

```
Force10(conf-ipv6-acl)#seq 8 permit tcp any any urg
May 5 08:32:34: %E90MJ:0 %ACL_AGENT-2-ACL_AGENT_ENTRY_ERROR: Unable to write seq 8 of
list test as individual TCP flags are not supported on linecard 0
```

Configure filters with sequence number

To create a filter for packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list extended** *access-list-name*	CONFIGURATION	Enter the IP ACCESS LIST mode by creating an extended IP ACL.
2	**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**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure a drop or forward filter. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor 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.

TCP packets: To create a filter for TCP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list extended** *access-list-name*	CONFIGURATION	Create an extended IP ACL and assign it a unique name.

Step	Command Syntax	Command Mode	Purpose
2	**seq** *sequence-number* {**deny** | **permit**} **tcp** {*source mask* | **any** | **host** *ip-address*}} [**count** [**byte**] | **log**] [**order**] [**monitor]** [**fragments**]	CONFIG-EXT-NACL	Configure an extended IP ACL filter for TCP packets. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor 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.

UDP packets: To create a filter for UDP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list extended** *access-list-name*	CONFIGURATION	Create a extended IP ACL and assign it a unique name.
2	**seq** *sequence-number* {**deny** | **permit**} {*ip-protocol-number* **udp**} {*source mask* | **any** | **host** *ip-address*} {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure an extended IP ACL filter for UDP packets. • **log** and **monitor** options are supported on E-Series only.

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, keep in mind that you might need to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five or another number.

Figure 7-7 illustrates 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.

Figure 7-7. Command Example: seq

```
Force10(config-ext-nacl)#seq 15 deny ip host 112.45.0.0 any log
Force10(config-ext-nacl)#seq 5 permit tcp 12.1.3.45 0.0.255.255 any
Force10(config-ext-nacl)#show confi
!
ip access-list extended dilling
 seq 5 permit tcp 12.1.0.0 0.0.255.255 any
 seq 15 deny ip host 112.45.0.0 any log
Force10(config-ext-nacl)#
```

Configure filters without sequence number

If you are creating an extended ACL with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. FTOS assigns filters in multiples of 5.

To configure a filter for an extended IP ACL without a specified sequence number, use any or all of the following commands in the IP ACCESS LIST mode:

Command Syntax	Command Mode	Purpose
{**deny** | **permit**} {*source mask* | **any** | **host** *ip-address*} [**count** [**byte**] | **log**] [**order**] [**monitor]** [**fragments**]	CONFIG-EXT-NACL	Configure a deny or permit filter to examine IP packets. • **log** and **monitor** options are supported on E-Series only.
{**deny** | **permit**} **tcp** {*source mask*] | **any** | **host** *ip-address*}} [**count** [**byte**] | **log**] [**order**] [**monitor]** [**fragments**]	CONFIG-EXT-NACL	Configure a deny or permit filter to examine TCP packets. • **log** and **monitor** options are supported on E-Series only.
{**deny** | **permit**} **udp** {*source mask* | **any** | **host** *ip-address*}} [**count** [**byte**] | **log**] [**order**] [**monitor]** [**fragments**]	CONFIG-EXT-NACL	Configure a deny or permit filter to examine UDP packets. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor 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.

Figure 7-8 illustrates 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 the IP ACCESS LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 7-8. Extended IP ACL

```
Force10(config-ext-nacl)#deny tcp host 123.55.34.0 any
Force10(config-ext-nacl)#permit udp 154.44.123.34 0.0.255.255 host 34.6.0.0
Force10(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
Force10(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 (Figure 232) in the EXEC Privilege mode.

Configuring Layer 2 and Layer 3 ACLs on an Interface

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:

- The packets routed by FTOS are governed by the L3 ACL only, since they are not filtered against an L2 ACL.
- The packets switched by FTOS are first filtered by the L3 ACL, then by the L2 ACL.
- When packets are switched by FTOS, the egress L3 ACL does not filter the packet.

For the following features, if counters are enabled on rules that have already been configured and a new rule is either inserted or prepended, all the existing counters will be reset:

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

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

Table 7-2. L2 and L3 ACL Filtering on Switched Packets

L2 ACL Behavior	L3 ACL Behavior	Decision on Targeted Traffic
Deny	Deny	Denied by L3 ACL
Deny	Permit	Permitted by L3 ACL
Permit	Deny	Denied by L3 ACL
Permit	Permit	Permitted by L3 ACL

Note: If an interface is configured as a "**vlan-stack access**" port, the packets are filtered by an L2 ACL only. The L3 ACL applied to such a port does not affect traffic. That is, existing rules for other features (such as trace-list, PBR, and QoS) are applied accordingly to the permitted traffic.

For information on MAC ACLs, refer to Chapter 20, "Layer 2," on page 387.

Assign an IP ACL to an Interface

Ingress IP ACLs are supported on platforms: C and S

Ingress and Egress IP ACL are supported on platform: E S55 S60 and S55

To pass traffic through a configured IP ACL, you must assign that ACL to a physical interface, a port channel interface, or a VLAN. The IP ACL is applied to all traffic entering a physical or port channel interface and the traffic is either forwarded or dropped depending on the criteria and actions specified in the ACL.

The same ACL may be applied to different interfaces and that changes its functionality. For example, you can take ACL "ABCD", and apply it using the **in** keyword and it becomes an ingress access list. If you apply the same ACL using the **out** keyword, it becomes an egress access list. If you apply the same ACL to the loopback interface, it becomes a loopback access list.

This chapter covers the following topics:

- Configuring Ingress ACLs on page 114
- Configuring Egress ACLs on page 115
- Configuring ACLs to Loopback on page 116

For more information on Layer-3 interfaces, refer to Chapter 13, Interfaces, on page 47.

To apply an IP ACL (standard or extended) to a physical or port channel interface, use these commands in the following sequence in the INTERFACE mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface interface slot/port**	CONFIGURATION	Enter the interface number.
2	**ip address** *ip-address*	INTERFACE	Configure an IP address for the interface, placing it in Layer-3 mode.
3	**ip access-group** *access-list-name* {**in | out**} [**implicit-permit**] [**vlan** *vlan-range*]	INTERFACE	Apply an IP ACL to traffic entering or exiting an interface. • **out:** configure the ACL to filter outgoing traffic. This keyword is supported only on E-Series. **Note:** The number of entries allowed per ACL is hardware-dependent. Refer to your line card documentation for detailed specification on entries allowed per ACL.
4	**ip access-list [standard | extended]** *name*	INTERFACE	Apply rules to the new ACL.

To view which IP ACL is applied to an interface, use the **show config** command (Figure 232) in the INTERFACE mode or the **show running-config** command in the EXEC mode.

Figure 7-9. Command example: show config in the INTERFACE Mode

```
Force10(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
Force10(conf-if)#
```

Use only Standard ACLs in the **access-class** command to filter traffic on Telnet sessions.

Counting ACL Hits

You can view the number of packets matching the ACL by using the **count** option when creating ACL entries. E-Series supports packet and byte counts simultaneously. C-Series and S-Series support only one at any given time.

To view the number of packets matching an ACL that is applied to an interface:

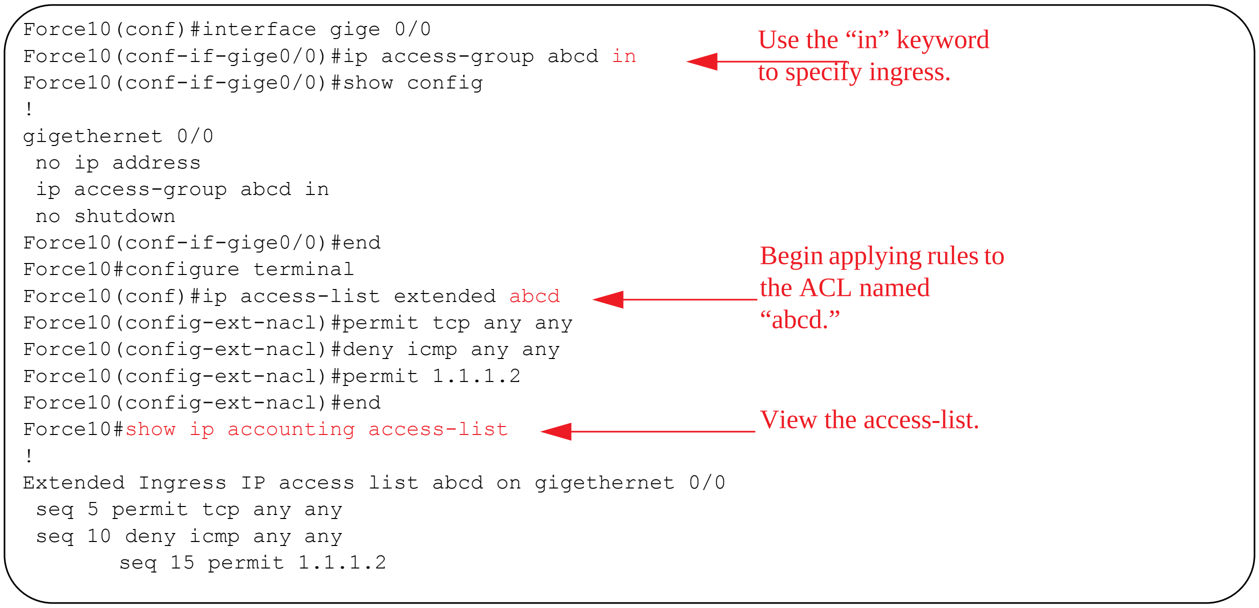
Step	Task
1	Create an ACL that uses rules with the count option. See Configure a standard IP ACL on page 106
2	Apply the ACL as an inbound or outbound ACL on an interface. See Assign an IP ACL to an Interface on page 112
3	View the number of packets matching the ACL using the **show ip accounting access-list** from EXEC Privilege mode.

Configuring 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 ACLs, use the **ip access-group** command (Figure 233) in the EXEC Privilege mode. This example also shows applying the ACL, applying rules to the newly created access group, and viewing the access list:

Figure 7-10. Creating an Ingress ACL



```
Force10(conf)#interface gige 0/0
Force10(conf-if-gige0/0)#ip access-group abcd in
Force10(conf-if-gige0/0)#show config
!
gigethernet 0/0
 no ip address
 ip access-group abcd in
 no shutdown
Force10(conf-if-gige0/0)#end
Force10#configure terminal
Force10(conf)#ip access-list extended abcd
Force10(config-ext-nacl)#permit tcp any any
Force10(config-ext-nacl)#deny icmp any any
Force10(config-ext-nacl)#permit 1.1.1.2
Force10(config-ext-nacl)#end
Force10#show ip accounting access-list
!
Extended Ingress IP access list abcd on gigethernet 0/0
 seq 5 permit tcp any any
 seq 10 deny icmp any any
        seq 15 permit 1.1.1.2
```

Use the “in” keyword to specify ingress.

Begin applying rules to the ACL named “abcd.”

View the access-list.

Configuring Egress ACLs

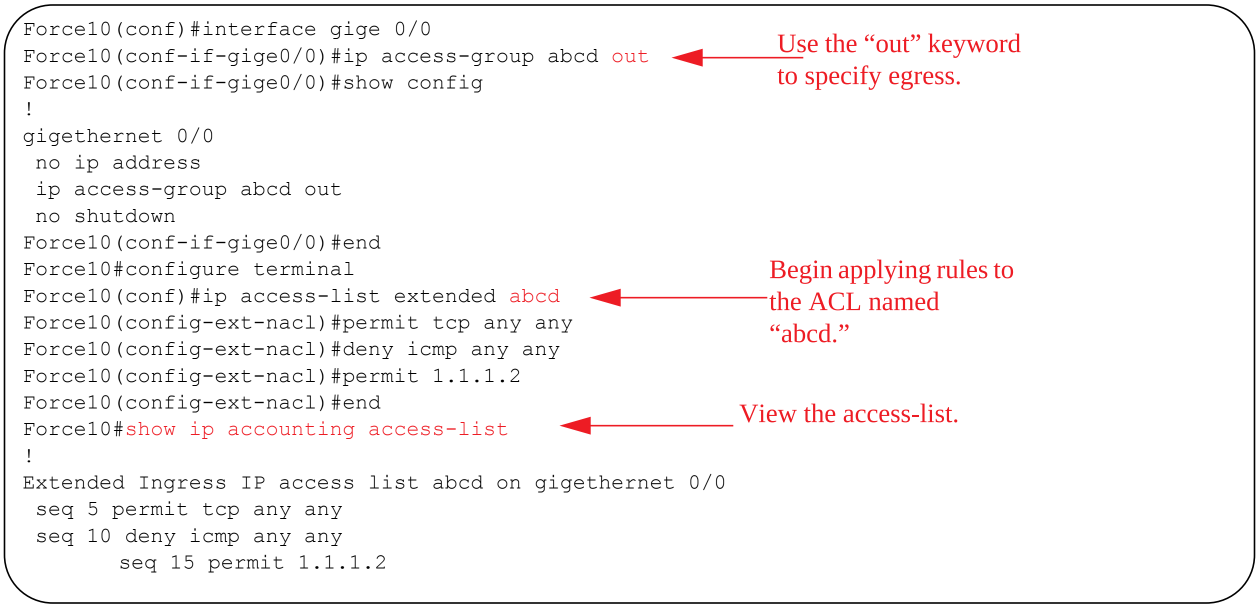
Egress ACLs are supported on platforms E and the S55

Egress ACLs are applied to line cards and affect the traffic leaving the system. Configuring egress ACLs onto physical interfaces protects the system infrastructure from attack—malicious and incidental—by explicitly allowing only authorized traffic.These system-wide ACLs eliminate the need to apply ACLs onto each interface and achieves the same results. By localizing target traffic, it is a simpler implementation.

An egress ACL is used when users would like to restrict egress traffic. For example, when a DOS attack traffic is isolated to one particular interface, you can apply an egress ACL to block that particular flow from exiting the box, thereby protecting downstream devices.

To create an egress ACLs, use the **ip access-group** command (Figure 234) in the EXEC Privilege mode. This example also shows viewing the configuration, applying rules to the newly created access group, and viewing the access list:

Figure 7-11. Creating an Egress ACL



```
Force10(conf)#interface gige 0/0
Force10(conf-if-gige0/0)#ip access-group abcd out
Force10(conf-if-gige0/0)#show config
!
gigethernet 0/0
 no ip address
 ip access-group abcd out
 no shutdown
Force10(conf-if-gige0/0)#end
Force10#configure terminal
Force10(conf)#ip access-list extended abcd
Force10(config-ext-nacl)#permit tcp any any
Force10(config-ext-nacl)#deny icmp any any
Force10(config-ext-nacl)#permit 1.1.1.2
Force10(config-ext-nacl)#end
Force10#show ip accounting access-list
!
Extended Ingress IP access list abcd on gigethernet 0/0
 seq 5 permit tcp any any
 seq 10 deny icmp any any
        seq 15 permit 1.1.1.2
```

Use the “out” keyword to specify egress.

Begin applying rules to the ACL named “abcd.”

View the access-list.

Egress Layer 3 ACL Lookup for Control-plane IP Traffic

By default, packets originated from the system are not filtered by egress ACLs. If you initiate a ping session from the system, for example, 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..

Task	Command Syntax	Command Mode
Apply Egress ACLs to IPv4 system traffic.	**ip control-plane** [**egress filter**]	CONFIGURATION
Apply Egress ACLs to IPv6 system traffic.	**ipv6 control-plane** [**egress filter**]	CONFIGURATION
Create a Layer 3 ACL using **permit** rules with the **count** option to describe the desired CPU traffic	**permit ip** {*source mask* | **any** | **host** *ip-address} {destination mask* | **any** | **host** *ip-address*} **count**	CONFIG-NACL

Note: The **ip control-plane** [**egress filter**] and the **ipv6 control-plane** [**egress filter**] commands are not supported on S55 systems.

FTOS Behavior: VRRP hellos and IGMP packets are not affected when egress ACL filtering for CPU traffic is enabled. 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.

Configuring ACLs to Loopback

ACLs can be supplied on Loopback interfaces supported on platform E

Configuring ACLs onto the CPU in a loopback interface protects the system infrastructure from attack—malicious and incidental—by explicate allowing only authorized traffic.

The ACLs on loopback interfaces are applied only to the CPU on the RPM—this eliminates the need to apply specific ACLs onto all ingress interfaces and achieves the same results. By localizing target traffic, it is a simpler implementation.

The ACLs target and handle Layer 3 traffic destined to terminate on the system including routing protocols, remote access, SNMP, ICMP, and etc. Effective filtering of Layer 3 traffic from Layer 3 routers reduces the risk of attack.

Note: Loopback ACLs are supported only on ingress traffic.

Loopback interfaces do not support ACLs using the IP fragment option. If you configure an ACL with the fragments option and apply it to a loopback interface, the command is accepted, but the ACL entries are not actually installed the offending rule in CAM.

See also Loopback Interfaces in the Interfaces chapter.

Applying an ACL on Loopback Interfaces

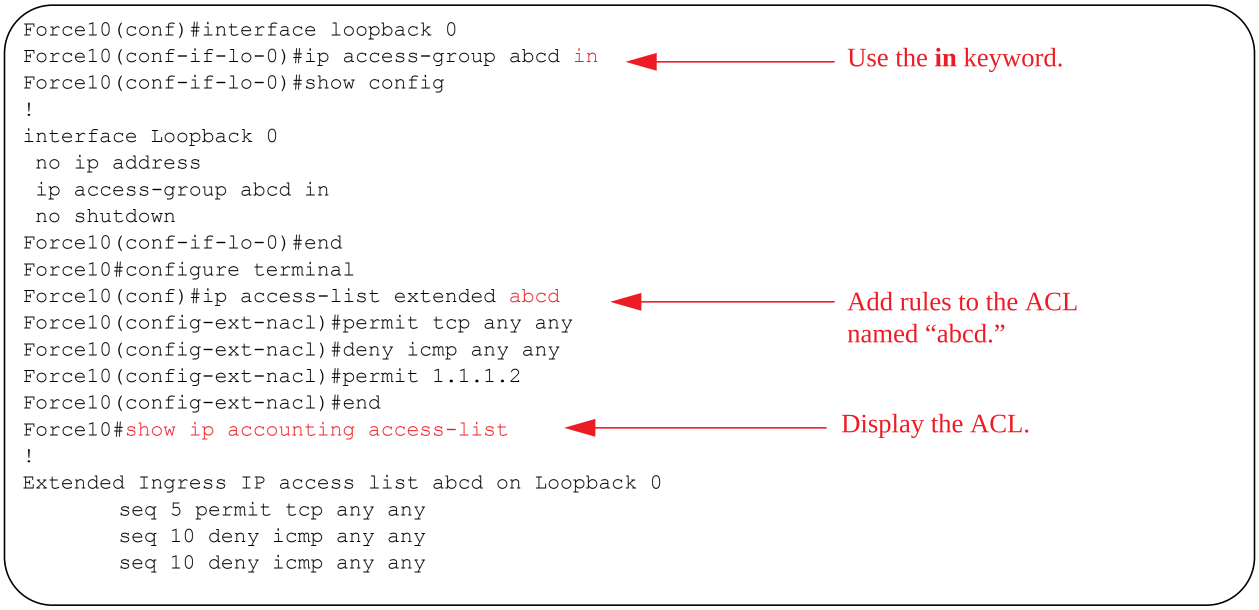
ACLs can be applied on Loopback interfaces supported on platform E

To apply an ACL (standard or extended) for loopback, use these commands in the following sequence:

Step	Command Syntax	Command Mode	Purpose
1	**interface loopback 0**	CONFIGURATION	Only loopback 0 is supported for the loopback ACL.
2	**ip access-list** [**standard** | **extended**] *name*	CONFIGURATION	Apply rules to the new ACL.
3	**ip access-group** *name* **in**	INTERFACE	Apply an ACL to traffic entering loopback. • **in:** configure the ACL to filter incoming traffic **Note:** ACLs for loopback can only be applied to incoming traffic.

To apply ACLs on loopback, use the **ip access-group** command (Figure 235) in the INTERFACE mode. This example also shows the interface configuration status, adding rules to the access group, and displaying the list of rules in the ACL:

Figure 7-12. Applying an ACL to the Loopback Interface



```
Force10(conf)#interface loopback 0
Force10(conf-if-lo-0)#ip access-group abcd in
Force10(conf-if-lo-0)#show config
!
interface Loopback 0
 no ip address
 ip access-group abcd in
 no shutdown
Force10(conf-if-lo-0)#end
Force10#configure terminal
Force10(conf)#ip access-list extended abcd
Force10(config-ext-nacl)#permit tcp any any
Force10(config-ext-nacl)#deny icmp any any
Force10(config-ext-nacl)#permit 1.1.1.2
Force10(config-ext-nacl)#end
Force10#show ip accounting access-list
!
Extended Ingress IP access list abcd on Loopback 0
        seq 5 permit tcp any any
        seq 10 deny icmp any any
        seq 10 deny icmp any any
```

Use the **in** keyword.

Add rules to the ACL named “abcd.”

Display the ACL.

Note: See also the section VTY Line Local Authentication and Authorization on page 644.

IP Prefix Lists

Prefix Lists are supported on platforms: C E S

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, FTOS 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.

Below are some examples that 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 le 12`
- 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.
- Once a route matches a filter, the filter’s action is applied. No additional filters are applied to the route.

Implementation Information

In FTOS, prefix lists are used in processing routes for routing protocols (for example, RIP, OSPF, and BGP).

Note: The S-Series platform 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, you must use commands in the PREFIX LIST, the ROUTER RIP, ROUTER OSPF, and ROUTER BGP modes. Basically, you create the prefix list in the PREFIX LIST mode, and assign that list to commands in the ROUTER RIP, ROUTER OSPF and ROUTER BGP modes.

The following list includes the configuration tasks for prefix lists:

- Configure a prefix list on page 119
- Use a prefix list for route redistribution on page 121

For a complete listing of all commands related to prefix lists, refer to the *FTOS Command Line Interface Reference* document.

Configure a prefix list

To configure a prefix list, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip prefix-list** *prefix-name*	CONFIGURATION	Create a prefix list and assign it a unique name. You are in the PREFIX LIST mode.
2	**seq** *sequence-number* {**deny** | **permit**} *ip-prefix* [**ge** *min-prefix-length*] [**le** *max-prefix-length*]	CONFIG-NPREFIXL	Create a prefix list with a sequence number and a deny or permit action. The optional parameters are: • **ge** *min-prefix-length:* is the minimum prefix length to be matched (0 to 32). • **le** *max-prefix-length:* is the maximum prefix length to be matched (0 to 32).

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 (**permit 0.0.0.0/0 le 32**). The “permit all” filter should be the last filter in your prefix list. To permit the default route only, enter **permit 0.0.0.0/0**.

Figure 7-13 illustrates 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.

Figure 7-13. Command Example: seq

```
Force10(conf-nprefixl)#seq 20 permit 0.0.0.0/0 le 32
Force10(conf-nprefixl)#seq 12 deny 134.23.0.0 /16
Force10(conf-nprefixl)#seq 15 deny 120.23.14.0 /8 le 16
Force10(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
Force10(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 the PREFIX LIST mode.

If you are creating a standard prefix list with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. The FTOS assigns filters in multiples of five.

To configure a filter without a specified sequence number, use these commands in the following sequence starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip prefix-list** *prefix-name*	CONFIGURATION	Create a prefix list and assign it a unique name.
2	{**deny** | **permit**} *ip-prefix* [**ge** *min-prefix-length*] [**le** *max-prefix-length*]	CONFIG-NPREFIXL	Create a prefix list filter with a deny or permit action. The optional parameters are: • **ge** *min-prefix-length:* is the minimum prefix length to be matched (0 to 32). • **le** *max-prefix-length:* is the maximum prefix length to be matched (0 to 32).

Figure 7-14 illustrates 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 the PREFIX LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 7-14. Prefix List

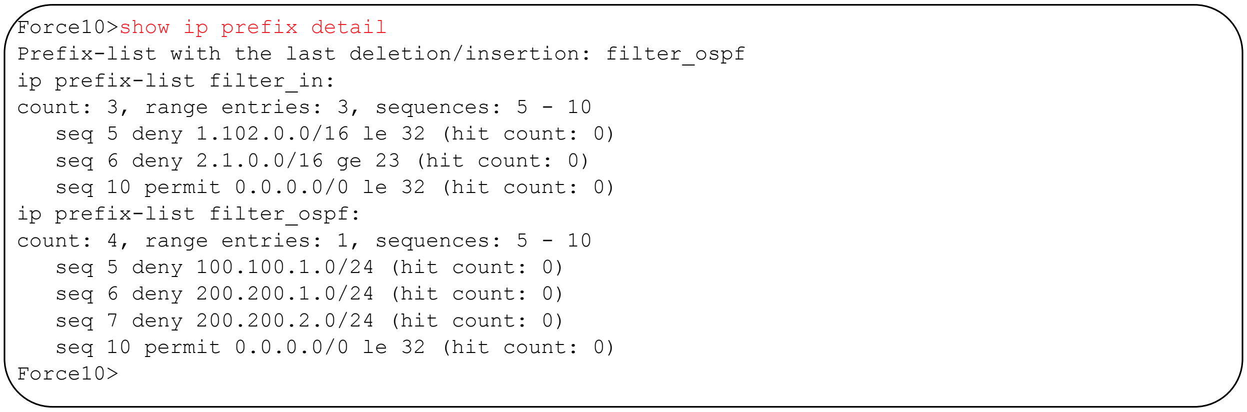
```
Force10(conf-nprefixl)#permit 123.23.0.0 /16
Force10(conf-nprefixl)#deny 133.24.56.0 /8
Force10(conf-nprefixl)#show conf
!
ip prefix-list awe
 seq 5 permit 123.23.0.0/16
 seq 10 deny 133.0.0.0/8
Force10(conf-nprefixl)#
```

To delete a filter, enter the **show config** command in the PREFIX LIST mode and locate the sequence number of the filter you want to delete; then use the **no seq** *sequence-number* command in the PREFIX LIST mode.

To view all configured prefix lists, use either of the following commands in the EXEC mode:

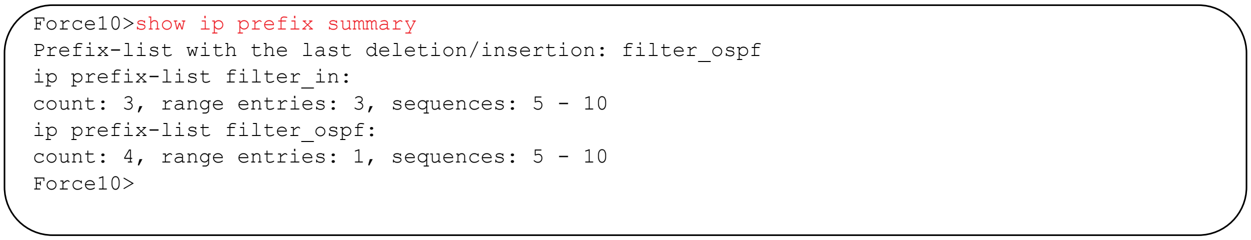
Command Syntax	Command Mode	Purpose
show ip prefix-list detail [*prefix-name*]	EXEC Privilege	Show detailed information about configured Prefix lists.
show ip prefix-list summary [*prefix-name*]	EXEC Privilege	Show a table of summarized information about configured Prefix lists.

Figure 7-15. Command example: show ip prefix-list detail



```
Force10>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)
Force10>
```

Figure 7-16. Command Example: show ip prefix-list summary



```
Force10>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
Force10>
```

Use a prefix list for route redistribution

To pass traffic through a configured prefix list, you must use the prefix list in a route redistribution command. The prefix list is applied to all traffic redistributed into the routing process and 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 (RIP is supported on C and E-Series.), use either of the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
router rip	CONFIGURATION	Enter RIP mode
distribute-list *prefix-list-name* **in** [*interface*]	CONFIG-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.
distribute-list *prefix-list-name* **out** [*interface* | **connected** | **static** | **ospf**]	CONFIG-ROUTER-RIP	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.

To view the configuration, use the **show config** command in the ROUTER RIP mode (Figure 240) or the **show running-config rip** command in the EXEC mode.

Figure 7-17. Command Example: show config in the ROUTER RIP Mode

```
Force10(conf-router_rip)#show config
!
router rip
 distribute-list prefix juba out
 network 10.0.0.0
Force10(conf-router_rip)#router ospf 34
```

To apply a filter to routes in OSPF, use either of the following commands in the ROUTER OSPF mode:

Command Syntax	Command Mode	Purpose
router ospf	CONFIGURATION	Enter OSPF mode
distribute-list *prefix-list-name* **in** [*interface*]	CONFIG-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.
distribute-list *prefix-list-name* **out** [**connected** | **rip** | **static**]	CONFIG-ROUTER-OSPF	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.

To view the configuration, use the **show config** command in the ROUTER OSPF mode (Figure 241) or the **show running-config ospf** command in the EXEC mode.

Figure 7-18. Command Example: show config in ROUTER OSPF Mode

```
Force10(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
Force10(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. Use Resequencing whenever there is no longer an opportunity to order new rules as desired using current numbering scheme.

For example, Table 7-3 contains some rules that are numbered in increments of 1. No new rules can be placed between these, so apply resequencing to create numbering space, as shown in Table 7-4. In the same example, apply resequencing if more than two rules must be placed between rules 7 and 10.

IPv4 and IPv6 ACLs and prefixes and MAC ACLs can be resequenced. No CAM writes happen as a result of resequencing, so there is no packet loss; the behavior is like Hot-lock ACLs.

Note: ACL Resequencing does not affect the rules or remarks or the order in which they are applied. It merely renumbers them so that new rules can be placed within the list as desired.

Table 7-3. ACL Resequencing Example (Insert New Rules)

seq 5 permit any host 1.1.1.1
seq 6 permit any host 1.1.1.2
seq 7 permit any host 1.1.1.3
seq 10 permit any host 1.1.1.4

Table 7-4. ACL Resequencing Example (Resequenced)

seq 5 permit any host 1.1.1.1
seq 10 permit any host 1.1.1.2
seq 15 permit any host 1.1.1.3
seq 20 permit any host 1.1.1.4

Resequencing an ACL or Prefix List

Resequencing is available for IPv4 and IPv6 ACLs and prefix lists and MAC ACLs. To resequence an ACL or prefix list use the appropriate command in Table 7-5. You must specify the list name, starting number, and increment when using these commands.

Table 7-5. Resequencing ACLs and Prefix Lists

List	Command	Command Mode
IPv4, IPv6, or MAC ACL	**resequence access-list** {**ipv4** | **ipv6 | mac**} {*access-list-name StartingSeqNum Step-to-Increment*}	Exec
IPv4 or IPv6 prefix-list	**resequence prefix-list** {**ipv4** | **ipv6**} {*prefix-list-name StartingSeqNum Step-to-Increment*}	Exec

Figure 7-19 shows the resequencing of an IPv4 access-list beginning with the number 2 and incrementing by 2.

Figure 7-19. Resequencing ACLs

```
Force10(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
Force10# end
Force10# resequence access-list ipv4 test 2 2
Force10# 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 and rules that originally have the same sequence number have the same sequence number after the **resequence** command is applied. Remarks that do not have a corresponding rule will be incremented as as a rule. These two mechanisms allow remarks to retain their original position in the list.

For example, in Figure 7-20, remark 10 corresponds 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.

Figure 7-20. Resequencing Remarks

```
Force10(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
Force10# end
Force10# resequence access-list ipv4 test 2 2
Force10# 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

Route-maps are supported on platforms: C E S

Like ACLs and prefix lists, route maps are composed of a series of commands that contain a matching criterion and an action, yet 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 FTOS implementation of route maps allows route maps with no match command or no set command. 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; 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

You configure route maps in the ROUTE-MAP mode and apply them in various commands in the ROUTER RIP and ROUTER OSPF modes.

The following list includes the configuration tasks for route maps:

- Create a route map on page 126 (mandatory)
- Configure route map filters on page 128 (optional)
- Configure a route map for route redistribution on page 131 (optional)
- Configure a route map for route tagging on page 132 (optional)

Create a route map

Route maps, ACLs, and prefix lists are similar in composition because all three contain filters, but route map filters are 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 and enter the ROUTE-MAP mode, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
route-map *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Create a route map and assign it a unique name. The optional **permit** and **deny** keywords are the action of the route map. The default is **permit**. The optional parameter **seq** allows you to assign a sequence number to the route map instance.

The default action is permit and the default sequence number starts at 10. When the keyword **deny** is used in configuring a route map, routes that meet the match filters are not redistributed.

To view the configuration, use the **show config** command in the ROUTE-MAP mode (Figure 244).

Figure 7-21. Command Example: show config in the ROUTE-MAP Mode

```
Force10(config-route-map)#show config
!
route-map dilling permit 10
Force10(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. FTOS processes the route maps with the lowest sequence number first. When a configured route map is applied to a command, like **redistribute**, traffic passes through all instances of that route map until a match is found. Figure 7-22 shows an example with two instances of a route map.

Figure 7-22. Command Example: show route-map with Multiple Instances of a Route Map

```
Force10#show route-map
route-map zakho, permit, sequence 10
 Match clauses:
 Set clauses:
route-map zakho, permit, sequence 20
 Match clauses:
  interface  GigabitEthernet 0/1
 Set clauses:
  tag  35
  level  stub-area
Force10#
```

Route map zakho has two instances

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 (Figure 246).

Figure 7-23. Deleting One Instance of a Route Map

```
Force10(conf)#no route-map zakho 10
Force10(conf)#end
Force10#show route-map
route-map zakho, permit, sequence 20
 Match clauses:
  interface  GigabitEthernet 0/1
 Set clauses:
  tag  35
  level  stub-area
Force10#
```

Figure 7-24 shows an example of 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.

Figure 7-24. Command Example: show route-map

```
Force10#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
Force10#
```

To delete a route map, use the **no route-map** *map-name* command in the CONFIGURATION mode.

Configure route map filters

Within the ROUTE-MAP mode, there are **match** and **set** commands. Basically, **match** commands search for a certain criterion in the routes and the **set** commands change the characteristics of those routes, either adding something or specifying a level.

When there are multiple match commands of the same parameter under one instance of route-map, then FTOS does a match between either of those match commands. If there are multiple match commands of different parameter, then FTOS does a match ONLY if there is a match among ALL match commands. The following example explains better:

Example 1

```
Force10(conf)#route-map force permit 10
Force10(config-route-map)#match tag 1000
Force10(config-route-map)#match tag 2000
Force10(config-route-map)#match tag 3000
```

In the above route-map, if a route has any of the tag value specified in the match commands, then there is a match.

Example 2

```
Force10(conf)#route-map force permit 10
Force10(config-route-map)#match tag 1000
Force10(config-route-map)#match metric 2000
```

In the above route-map, *only* if a route has *both* the characteristics mentioned in the route-map, it is matched. Explaining further, the route *must* have a tag value of 1000 *and* a metric value of 2000. Only then is there a match.

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. As an example:

```
Force10(conf)#route-map force permit 10
Force10(config-route-map)#match tag 1000

Force10(conf)#route-map force deny 20
Force10(config-route-map)#match tag 1000

Force10(conf)#route-map force deny 30
Force10(config-route-map)#match tag 1000
```

In the above route-map, instance 10 permits the route having a tag value of 1000 and instances 20 & 30 denies the route having a tag value of 1000. In the above scenario, FTOS scans all the instances of the route-map for any permit statement. If there is a match anywhere, the route is permitted, though other instances of the route-map denies it.

To configure match criterion for a route map, use any or all of the following commands in the ROUTE-MAP mode:

Command Syntax	**Command Mode**	**Purpose**
match as-path *as-path-name*	CONFIG-ROUTE-MAP	Match routes with the same AS-PATH numbers.
match community *community-list-name* [**exact**]	CONFIG-ROUTE-MAP	Match routes with COMMUNITY list attributes in their path.
match interface *interface*	CONFIG-ROUTE-MAP	Match routes whose next hop is a specific interface. The parameters are: • For a Fast Ethernet interface, enter the keyword **FastEthernet** followed by the slot/port information. • For a 1-Gigabit Ethernet interface, enter the keyword **gigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by a number between zero (0) and 16383. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **tengigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword vlan followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Command Syntax	Command Mode	Purpose
match ip address *prefix-list-name*	CONFIG-ROUTE-MAP	Match destination routes specified in a prefix list (IPv4).
match ipv6 address *prefix-list-name*	CONFIG-ROUTE-MAP	Match destination routes specified in a prefix list (IPv6).
match ip next-hop {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match next-hop routes specified in a prefix list (IPv4).
match ipv6 next-hop {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match next-hop routes specified in a prefix list (IPv6).
match ip route-source {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match source routes specified in a prefix list (IPv4).
match ipv6 route-source {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match source routes specified in a prefix list (IPv6).
match metric *metric-value*	CONFIG-ROUTE-MAP	Match routes with a specific value.
match origin {**egp** | **igp** | **incomplete**}	CONFIG-ROUTE-MAP	Match BGP routes based on the ORIGIN attribute.
match route-type {**external** [**type-1** | **type-2**] **| internal | level-1 | level-2** | **local** }	CONFIG-ROUTE-MAP	Match routes specified as internal or external to OSPF, ISIS level-1, ISIS level-2, or locally generated.
match tag *tag-value*	CONFIG-ROUTE-MAP	Match routes with a specific tag.

To configure a set condition, use any or all of the following commands in the ROUTE-MAP mode:

Command Syntax	Command Mode	Purpose
set as-path prepend *as-number* [... *as-number*]	CONFIG-ROUTE-MAP	Add an AS-PATH number to the beginning of the AS-PATH
set automatic-tag	CONFIG-ROUTE-MAP	Generate a tag to be added to redistributed routes.
set level {**backbone** | **level-1** | **level-1-2** | **level-2** | **stub-area** }	CONFIG-ROUTE-MAP	Specify an OSPF area or ISIS level for redistributed routes.
set local-preference *value*	CONFIG-ROUTE-MAP	Specify a value for the BGP route’s LOCAL_PREF attribute.
set metric {+ | - | *metric-value*}	CONFIG-ROUTE-MAP	Specify a value for redistributed routes.
set metric-type {**external** | **internal** | **type-1** | **type-2**}	CONFIG-ROUTE-MAP	Specify an OSPF or ISIS type for redistributed routes.
set next-hop *ip-address*	CONFIG-ROUTE-MAP	Assign an IP address as the route’s next hop.

Command Syntax	Command Mode	Purpose
set ipv6 next-hop *ip-address*	CONFIG-ROUTE-MAP	Assign an IPv6 address as the route's next hop.
set origin {**egp** | **igp** | **incomplete**}	CONFIG-ROUTE-MAP	Assign an ORIGIN attribute.
set tag *tag-value*	CONFIG-ROUTE-MAP	Specify a tag for the redistributed routes.
set weight *value*	CONFIG-ROUTE-MAP	Specify a value as the route's weight.

Use these commands to create route map instances. There is no limit to the number of set and match commands per route map, but the convention is to keep the number of match and 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 on the E-Series, you must call or include that route map in a command such as the **redistribute** or **default-information originate** commands in OSPF, ISIS, and BGP.

Route redistribution occurs when FTOS 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 Figure 7-25, 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 will be redistributed into the OSPF backbone area.

Note: When re-distributing routes using route-maps, the user must take care to create the route-map defined in the **redistribute** command under the routing protocol. If no route-map is created, then NO routes are redistributed.

Figure 7-25. Route Redistribution into OSPF

```
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 and that tag is passed along with the route as it passes through different routing protocols. This tag can then be used when the route leaves a routing domain to redistribute those routes again.

In Figure 7-26, the **redistribute ospf** command with a route map is used in the ROUTER RIP mode to apply a tag of 34 to all internal OSPF routes that are redistributed into RIP.

Figure 7-26. Tagging OSPF Routes Entering a RIP Routing Domain

```
!
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 the **continue** command is configured at the end of a module, the next module (or a specified module) is processed even after a match is found. Figure 7-27 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 will be processed.

Note: If the continue clause is configured without specifying a module, the next sequential module is processed.

Figure 7-27. Command Example: continue

```
!
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!
```

8

Border Gateway Protocol IPv4 (BGPv4)

Border Gateway Protocol IPv4 (BGPv4) version 4 (BGPv4) is supported on platforms: C E S

Platforms support BGP according to the following table:

FTOS version	Platform support	
8.1.1.0	E-Series ExaScale	E X
7.8.1.0	S-Series	S
7.7.1.0.	C-Series	C
pre-7.7.1.0	E-Series TeraScale	E T

This chapter is intended to provide a general description of Border Gateway Protocol version 4 (BGPv4) as it is supported in the Dell Force10 Operating System (FTOS).

This chapter includes the following topics:

- Protocol Overview
  - Autonomous Systems (AS)
  - Sessions and Peers
  - Route Reflectors
  - Confederations
- BGP Attributes
  - Best Path Selection Criteria
  - Weight
  - Local Preference
  - Multi-Exit Discriminators (MEDs)
  - AS Path
  - Next Hop
- Multiprotocol BGP
- Implementing BGP with FTOS
  - Advertise IGP cost as MED for redistributed routes

BGP protocol standards are listed in the Appendix 47, Standards Compliance chapter.

Protocol Overview

Border Gateway Protocol (BGP) is an external gateway protocol that transmits interdomain routing information within and between Autonomous Systems (AS). Its primary function is to exchange network reachability information with other BGP systems. BGP generally operates with an Internal Gateway Protocol (IGP) such as OSPF or RIP, allowing you to communicate to external ASs smoothly. BGP adds reliability to network connections be 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, already assigned by an internet authority. You do not assign the BGP number.

AS Numbers (ASNs) are important because the ASN uniquely identifies each network on the Internet. The IANA has reserved AS numbers 64512 through 65534 to be used for private purposes. The ASNs 0 and 65535 are reserved by the IANA and should not be used in a live environment.

Autonomous Systems can be grouped into three categories, defined by their connections and operation.

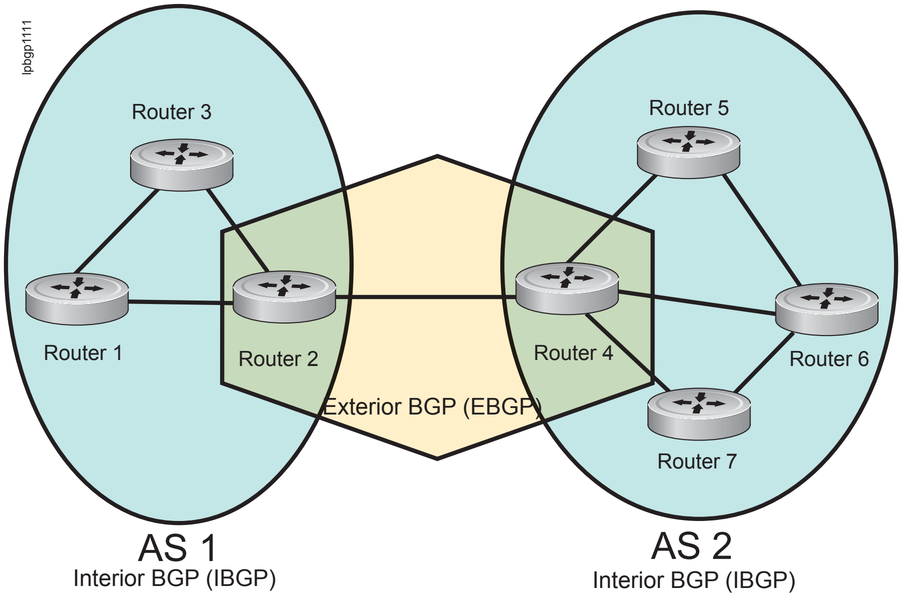
A **multihomed** AS is one that maintains connections to more than one other AS. This allows the AS to remain connected to the internet in the event of a complete failure of one of their connections. However, this type of AS does not allow traffic from one AS to pass through on its way to another AS. A simple example of this is seen in Figure 8-1.

A **stub** AS is one that is connected to only one other AS.

A **transit** AS is one that provides connections through itself to separate networks. For example as seen in Figure 8-1, Router 1 can use Router 2 (the transit AS) to connect to Router 4. 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 Autonomous System (AS1 or AS2 as seen in Figure 8-1), it is referred to as Internal BGP (IBGP *Interior Border Gateway Protocol*). When BGP operates between Autonomous Systems (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 8-1. BGP Autonomous Zones

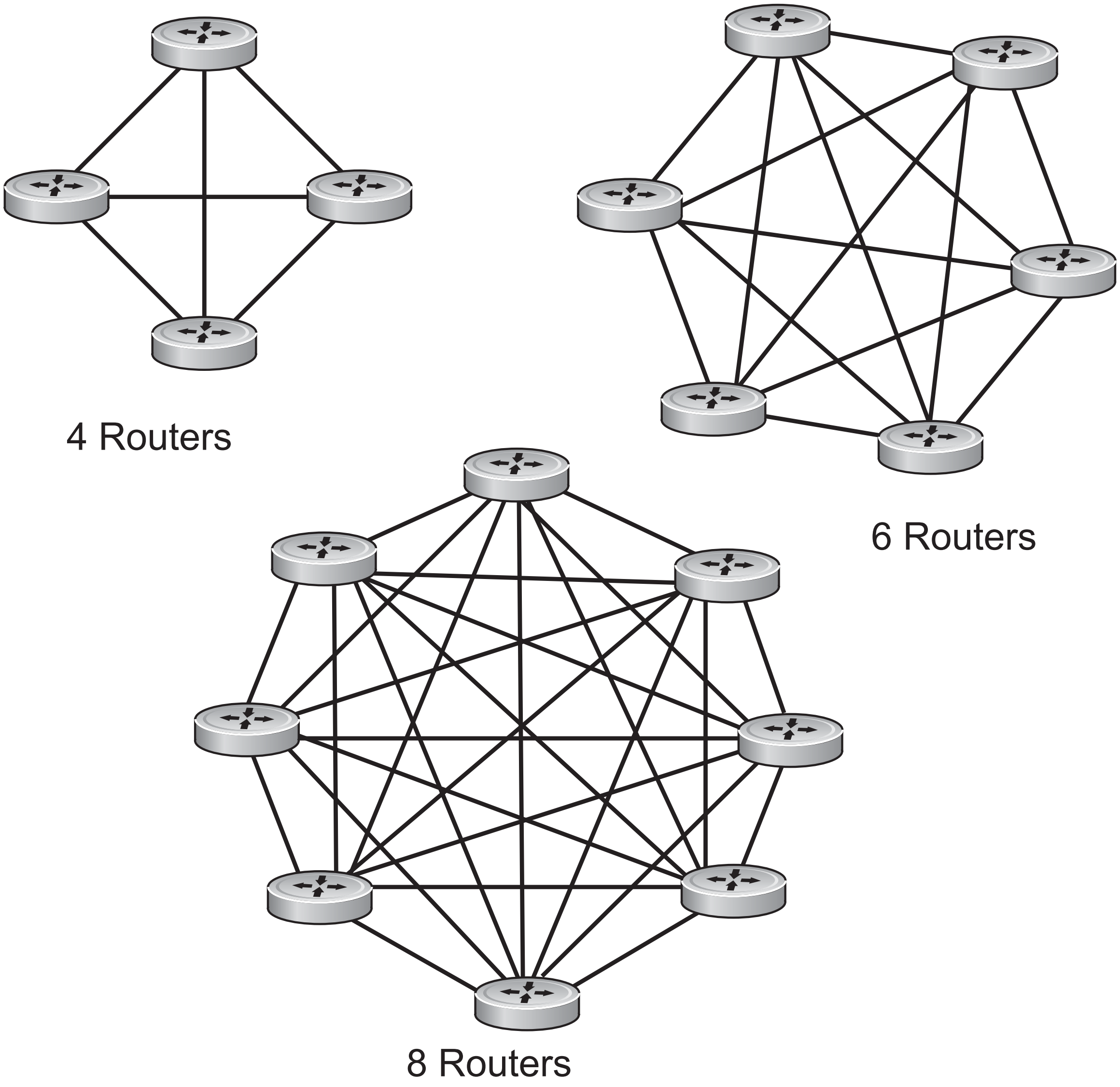


BGP version 4 (BGPv4) supports classless interdomain routing and aggregate routes and AS paths. BGP is a path vector protocol - a computer network in which BGP maintains the path that update information takes as it diffuses through the network. Updates traveling through the network and returning to the same node are easily detected and discarded.

BGP does not use traditional Interior Gateway Protocol (IGP) matrix, but makes routing decisions based on path, network policies and/or rulesets. Unlike most protocols, BGP uses TCP as its transport protocol.

Since each BGP routers 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. For example, as seen in Figure 8-2, four routers connected in a full mesh have three peers each, six routers have 5 peers each, and eight routers in full mesh will have seven peers each.

Figure 8-2. Full Mesh Examples



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.

Establishing 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 peer 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.

The first state is the **Idle** mode. BGP initializes all resources, refuses all inbound BGP connection attempts, and initiates a TCP connection to the peer.

The next state is **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.

In the **Active** state, the router resets the ConnectRetry timer to zero, and returns to the **Connect** state.

Upon successful **OpenSent** transition, the router sends an Open message and waits for one in return.

Once the Open message parameters are agreed between peers then the neighbor relation is established and is in **Open confirm** state. This is when the router receives and checks for agreement on the parameters of open messages to establish a session.

Keepalive messages are exchanged next, and upon 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.

Once established, the router can now send/receive Keepalive, Update, and Notification messages to/from its peer.

Peer Groups

Peer Groups are neighbors grouped according to common routing policies. They enable easier system configuration and management by allowing groups of routers to share and inherit policies.

Peer groups also aid in convergence speed. When a BGP process needs to send the same information to a large number of peers, it needs to set up a long output queue to get that information to all the proper peers. If they are members of a peer group, however, the information can be sent to one place then passed onto the peers within the group.

Route Reflectors

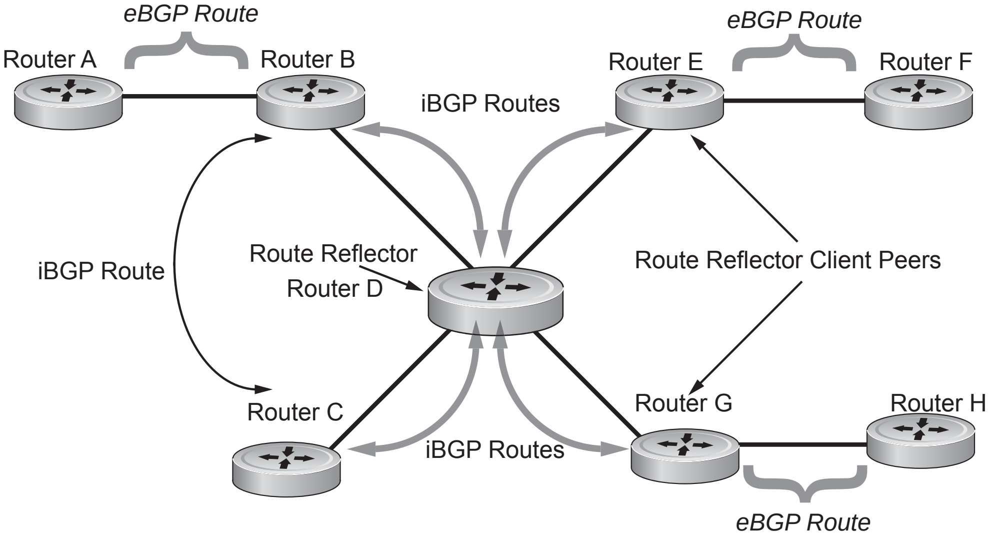
Route Reflectors reorganize the iBGP core into a hierarchy and allows some route advertisement rules.

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. Since BGP speakers announce only the best route for a given prefix, route reflector rules are applied after the router makes its best path decision.

- If a route was received from a nonclient peer, reflect the route to all client peers.
- If the route was received from a client peer, reflect the route to all nonclient and all client peers.

To illustrate how these rules affect routing, see Figure 8-3 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 8-3. Route Reflection Example



1. Router B receives an advertisement from Router A through eBGP. Since 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 non-client 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.

Confederations

Communities

BGP communities are sets of routes with one or more common attributes. This is a way to assign common attributes to multiple routes at the same time.

BGP Attributes

Routes learned via 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

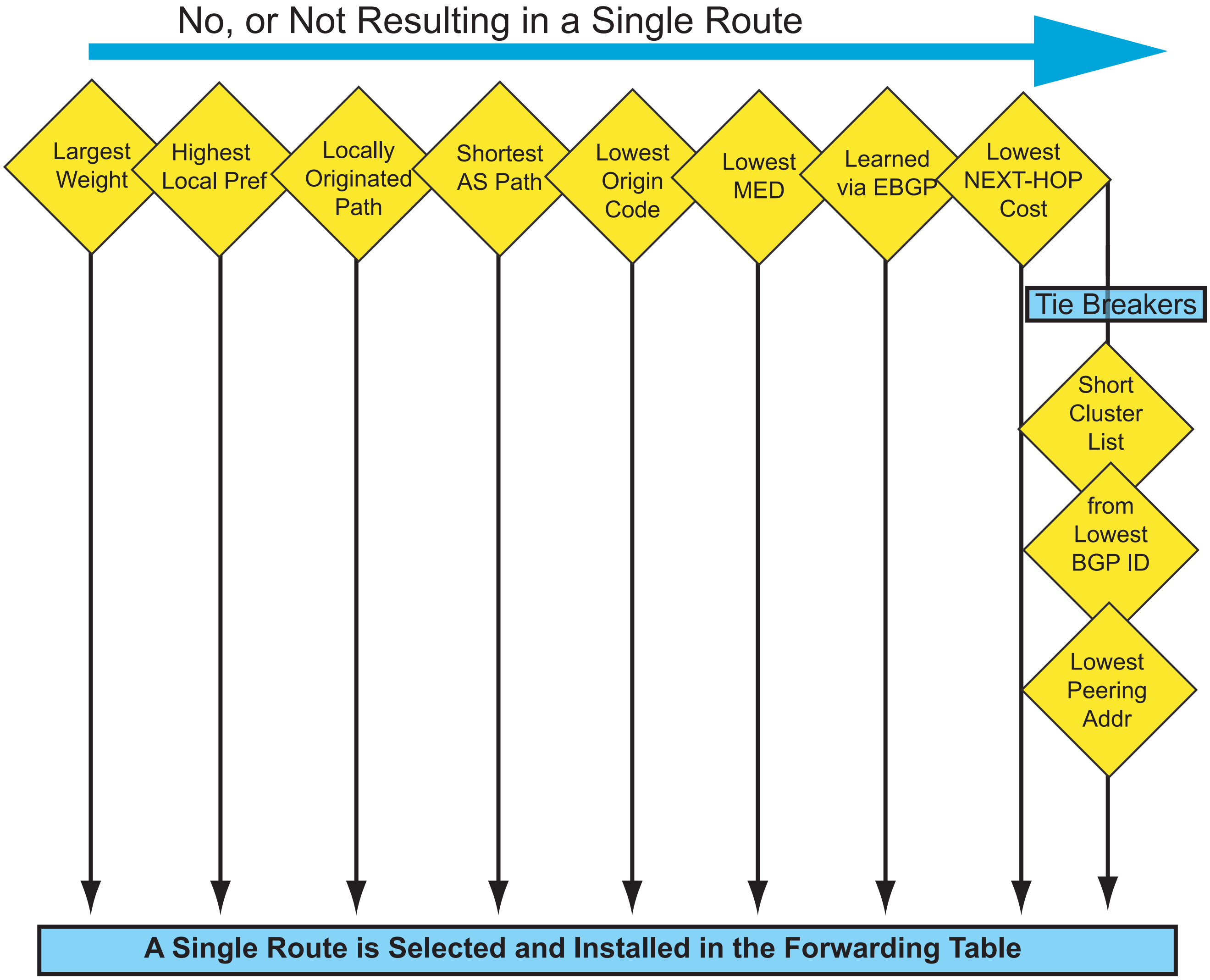
Paths for active routes are grouped in ascending order according to their neighboring external AS number (BGP best path selection is deterministic by default, which means the **bgp non-deterministic-med** command is NOT applied).

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 is determined, this selection criteria is applied to group’s best to determine the ultimate best path.

In non-deterministic mode (the **bgp non-deterministic-med** command is applied), paths are compared in the order in which they arrive. This method can lead to FTOS choosing different best paths from a set of paths, depending on the order in which they were received from the neighbors, since MED may or may not get compared between adjacent paths. In deterministic mode, FTOS compares MED between adjacent paths within an AS group since all paths in the AS group are from the same AS.

Figure 8-4 illustrates the decisions BGP goes through to select the best path. The list following the illustration details the path selection criteria.

Figure 8-4. 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.
    - Routes originated with the **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:
    - An AS_SET has a path length of 1, no matter how many ASs are in the set.
    - A path with no AS_PATH configured has a path length of 0.
    - AS_CONFED_SET is not included in the AS_PATH length.

- 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:
    - 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.
    - If the **bgp always-compare-med** command is entered, MEDs are compared for all paths.
    - 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 next-hop is selected when synchronization is disabled and only an internal path remains.
9. FTOS deems the paths as equal and does not perform steps 9 through 11 listed below, if the following criteria is met:
    - the IBGP multipath or EBGP multipath are configured (**maximum-path** command)
    - the paths being compared were received from the same AS with the same number of ASs in the AS Path but with different NextHops
    - the paths were received from IBGP or EBGP neighbor respectively
10. If the **bgp bestpath router-id ignore** command is enabled and:
    - If the Router-ID is the same for multiple paths (because the routes were received from the same route) skip this step.
    - 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 should return without performing any of the checks outlined below.
11. Prefer the path originated from the BGP router with the lowest router ID. For paths containing a Route Reflector (RR) attribute, the originator ID is substituted for the router ID.
12. If two paths have the same router ID, prefer the path with the lowest cluster ID length. Paths without a cluster ID length are set to a 0 cluster ID length.
13. Prefer the path originated from the neighbor with the lowest address. (The neighbor address is used in the BGP neighbor configuration, and corresponds to the remote peer used in the TCP connection with the local router.)

After a number of best paths is determined, this selection criteria is applied to group’s best to determine the ultimate best path.

In non-deterministic mode (the **bgp non-deterministic-med** command is applied), paths are compared in the order in which they arrive. This method can lead to FTOS choosing different best paths from a set of paths, depending on the order in which they were received from the neighbors since MED may or may not get compared between adjacent paths. In deterministic mode, FTOS compares MED between adjacent paths within an AS group since all paths in the AS group are from the same AS.

Weight

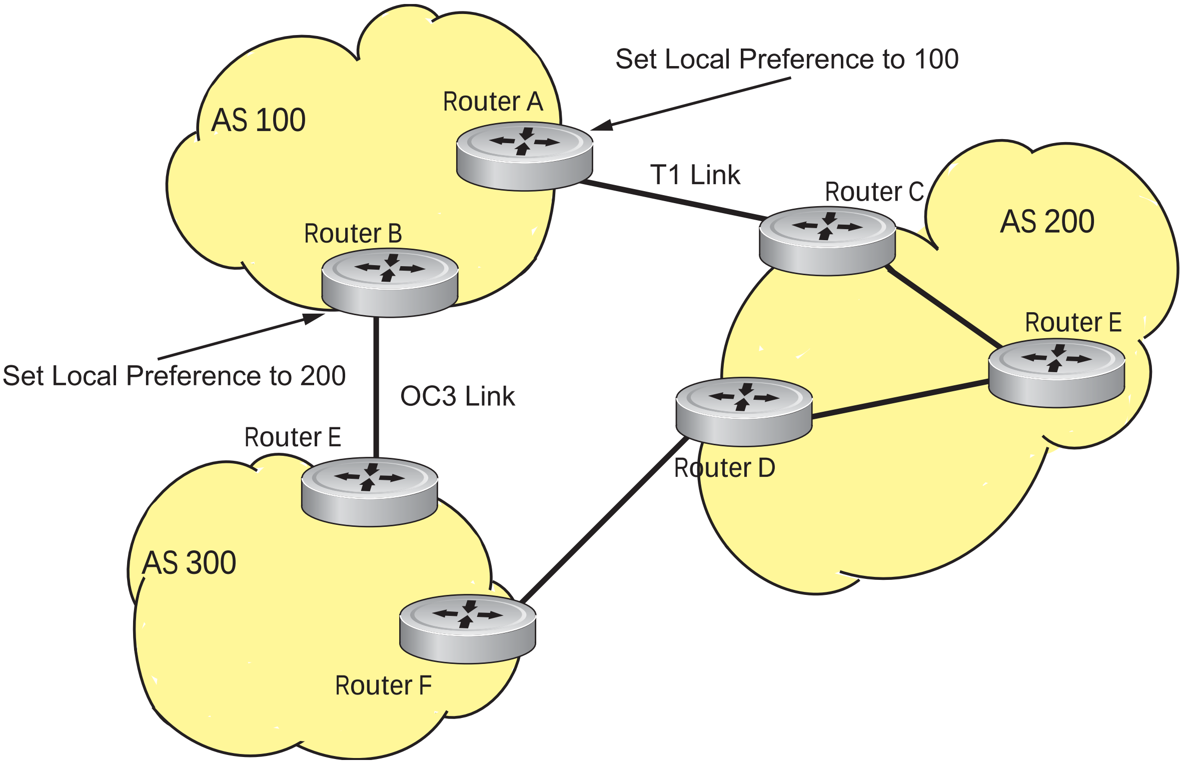
The Weight attribute is local to the router and is not advertised to neighboring routers. If the router learns about more than one route to the same destination, the route with the highest weight will be preferred. The route with the highest weight is installed in the IP routing table.

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.

The 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 Figure 8-4. For this example, assume that LOCAL_PREF is the only attribute applied. In Figure 8-5, AS100 has two possible paths to AS 200. Although the path through the 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 8-5. LOCAL_PREF Example



Multi-Exit Discriminators (MEDs)

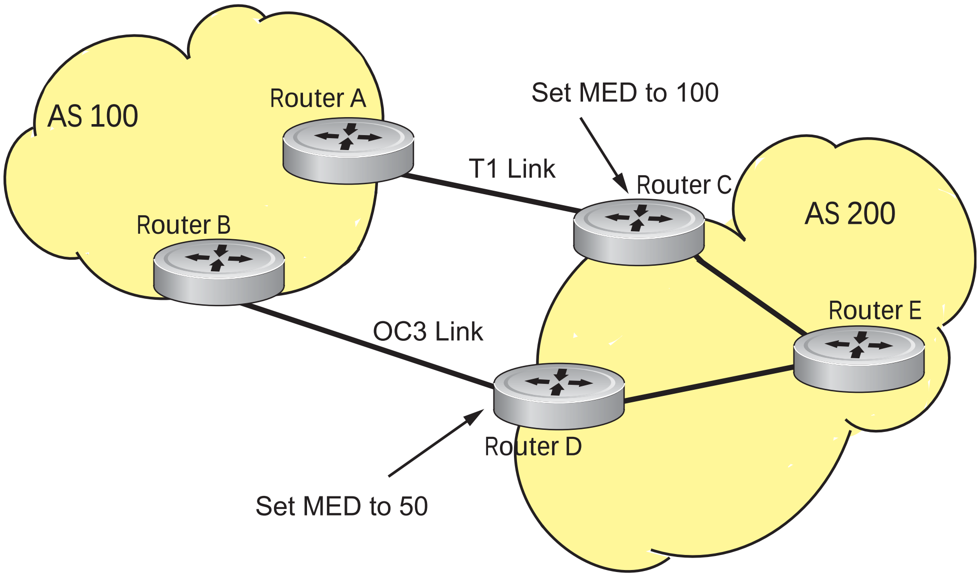
If two Autonomous Systems (AS) connect in more than one place, a Multi-Exit Discriminator (MED) can be used to assign a preference to a preferred path. The MED is one of the criteria used to determine the best path, so keep in mind that other criteria may impact selection, as shown in Figure 8-4.

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 Figure 8-6, 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.

An MED is a non-transitive attribute. 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 Autonomous Systems (AS100 and AS200).

Note that the MEDs are advertised across both links, so that if a link goes down AS 1 still has connectivity to AS300 and AS400.

Figure 8-6. MED Route Example



Note: With FTOS Release 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

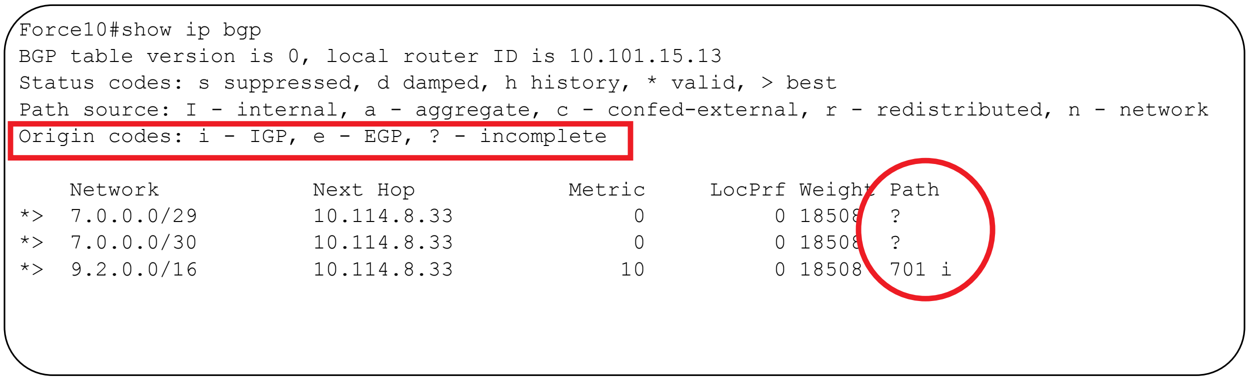
The Origin indicates the origin of the prefix, or how the prefix came into BGP. There are three Origin codes: IGP, EGP, INCOMPLETE.

- IGP indicated the prefix originated from information learned through an interior gateway protocol.
- EGP indicated 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 FTOS, these origin codes appear as shown in Figure 8-7. The question mark (?) indicates an Origin code of INCOMPLETE. The lower case letter (i) indicates an Origin code of IGP.

Figure 8-7. Origin attribute reported



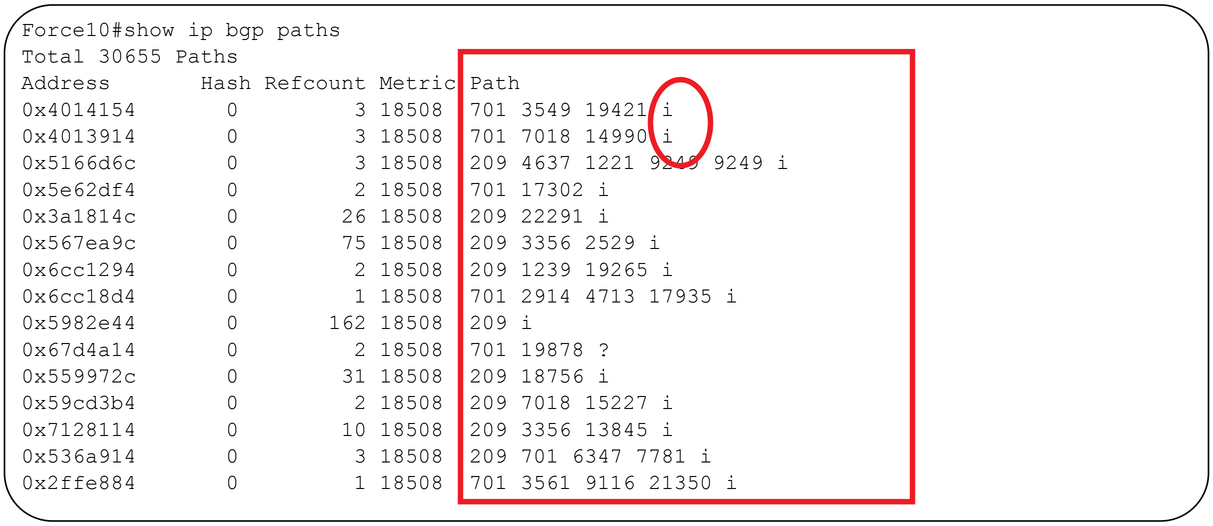
```
Force10#show ip bgp
BGP table version is 0, local router ID is 10.101.15.13
Status codes: s suppressed, d damped, h history, * valid, > best
Path source: I - internal, a - aggregate, c - confed-external, r - redistributed, n - network
Origin codes: i - IGP, e - EGP, ? - incomplete

    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 Autonomous Systems 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.

In FTOS the AS Path is shown in Figure 8-8. Note that the Origin attribute is shown following the AS Path information.

Figure 8-8. AS Path attribute reported



```
Force10#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. It can also be set 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.

FTOS 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

MBGP for IPv6 unicast is supported on platforms E C

MBGP for IPv4 Multicast is supported on platform C E S

Multiprotocol Extensions for BGP (MBGP) is defined in IETF RFC 2858. MBGP allows different types of address families to be distributed in parallel. This allows information about the topology of 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. Therefor, You cannot redistribute Multiprotocol BGP routes into BGP.

Implementing BGP with FTOS

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.

FTOS 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 in conjunction with the **route-map** command, you can specify whether a peer advertises the standard MED or uses the IGP cost as the MED.

Note the following when configuring this functionality:

- If the **redistribute** command does not have any **metric** configured and BGP Peer out-bound 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 out-bound route-map has **metric-type internal** configured, BGP advertises the metric configured in the redistribute command as MED.
- If BGP peer out-bound route-map has **metric** configured, then all other metrics are overwritten by this.

Note: When redistributing static, connected or OSPF routes, there is no metric option. Simply assign the appropriate route-map to the redistributed route.

Table 8-1gives some examples of these rules.

Table 8-1. Example MED advertisement

Command Settings	BGP Local Routing Information Base	MED Advertised to Peer	
		WITH route-map metric-type internal	**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
redistribute *isis* metric 100	MED: IGP cost 100	MED: 100	MED: 100

Ignore Router-ID for some best-path calculations

FTOS 8.3.1.0 and later allow 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.

4-Byte AS Numbers

FTOS Version 7.7.1 and later support 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 will be different with the peer depending on whether the peer is 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 will appear 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.**

Table 8-2. 4-Byte ASN Dot Format Examples

Traditional Format		Dot Format
65001	**Is**	0.65501
65536	**The**	1.0
100000	**Same As**	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 the 4-byte AS numbers with the **four-octet-support** command.

AS4 Number Representation

FTOS version 8.2.1.0 supports multiple representations of an 4-byte AS Numbers: **asplain**, **asdot+**, and **asdot**.

Note: The ASDOT and ASDOT+ representations are supported only in conjunction with the 4-Byte AS Numbers feature. If 4-Byte AS Numbers are not implemented, only ASPLAIN representation is supported.

ASPLAIN is the method FTOS has used for all previous FTOS versions.It remains the default method with FTOS 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 0-65535 are represented as a decimal number when entered in the CLI as well as when displayed in the show command outputs.
- AS Numbers larger than 65535 are represented using ASPLAIN notation as well. 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 Table 8-2.

- All AS Numbers between 0-65535 are represented as a decimal number, when entered in the CLI as well as when displayed in the show command 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 using the decimal method (asdot+). For example, the AS Number 65526 appears as 65526, and the AS Number 65546 appears as 1.10.

Dynamic AS Number Notation application

FTOS 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 (Figure 8-9 and Figure 8-10).

Figure 8-9. Dynamic changes of the bgp asnotation command in the show running config

ASDOT

```
Force10(conf-router_bgp)#bgp asnotation asdot
Force10(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>

Force10(conf-router_bgp)#do show ip bgp
BGP table version is 24901, local router ID is 172.30.1.57
<output truncated>
```

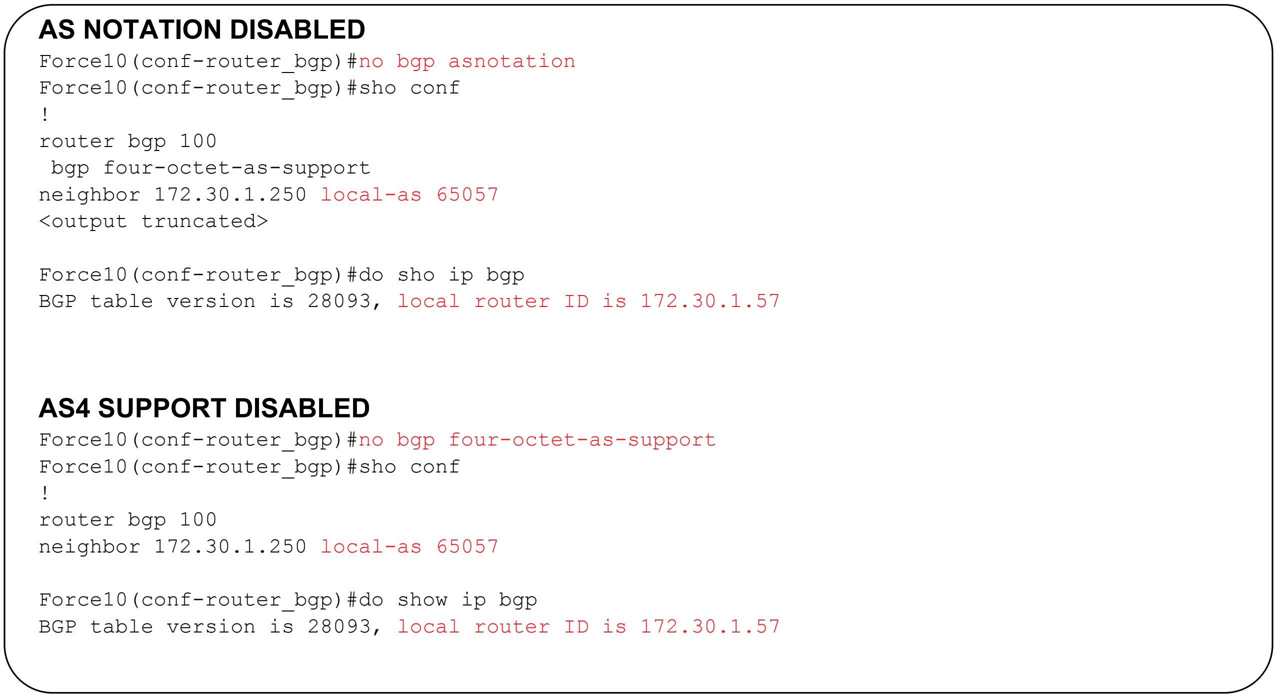
ASDOT+

```
Force10(conf-router_bgp)#bgp asnotation asdot+
Force10(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>

Force10(conf-router_bgp)#do show ip bgp
BGP table version is 31571, local router ID is 172.30.1.57
<output truncated>
```

AS-PLAIN

```
Force10(conf-router_bgp)#bgp asnotation asplain
Force10(conf-router_bgp)#sho conf
!
router bgp 100
 bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

Force10(conf-router_bgp)#do sho ip bgp
BGP table version is 34558, local router ID is 172.30.1.57
<output truncated>
```

Figure 8-10. Dynamic changes when bgp asnotation command is disabled in the show running config



AS NOTATION DISABLED

```
Force10(conf-router_bgp)#no bgp asnotation
Force10(conf-router_bgp)#sho conf
!
router bgp 100
 bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

Force10(conf-router_bgp)#do sho ip bgp
BGP table version is 28093, local router ID is 172.30.1.57
```

AS4 SUPPORT DISABLED

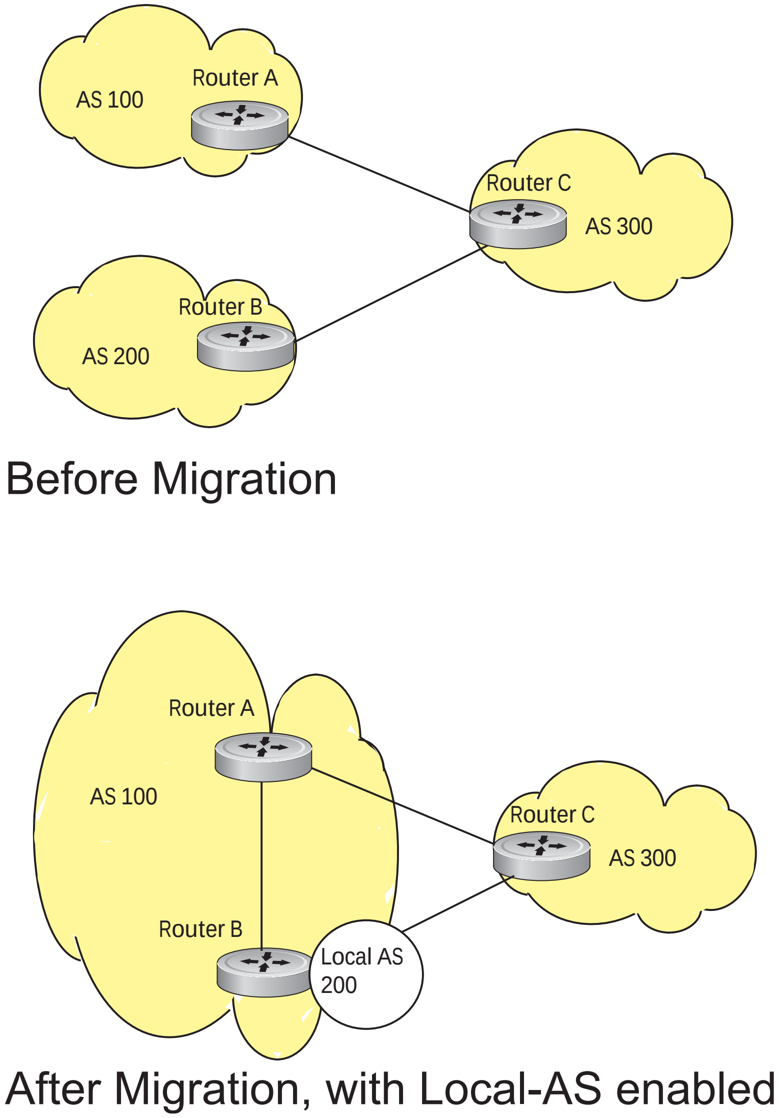
```
Force10(conf-router_bgp)#no bgp four-octet-as-support
Force10(conf-router_bgp)#sho conf
!
router bgp 100
neighbor 172.30.1.250 local-as 65057

Force10(conf-router_bgp)#do show ip bgp
BGP table version is 28093, local router ID is 172.30.1.57
```

AS Number Migration

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 need to be updated to maintain network reachability. With this feature you can transparently change the AS number of entire BGP network and ensure that the routes are propagated throughout the network while the migration is in progress. 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.

Figure 8-11 shows a scenario where Router A, Router B and Router C belong to AS 100, 200, 300 respectively. Router A acquired Router B; Router B has Router C as its customer. When Router B is migrating to Router A, it needs to maintain the connection with Router C without immediately updating Router C's configuration. **Local-AS** allows this 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 8-11. Local-AS Scenario



Before Migration

After Migration, with Local-AS enabled

When you complete your migration, and you have reconfigured your network with the new information you must disable this feature.

If the “no prepend” option is used, the local-as will not be prepended to the updates received from the eBGP peer. If “no prepend” is not selected (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 Figure 8-11. If Router B has an inbound route-map applied on Router C to prepend "65001 65002" to the as-path, the following events will 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 thru a router in AS 200 before it reached Router B.

BGP4 Management Information Base (MIB)

The FORCE10-BGP4-V2-MIB enhances FTOS BGP Management Information Base (MIB) support with many new SNMP objects and notifications (traps) defined in the *draft-ietf-idr-bgp4-mibv2-05*. To see these enhancements, download the MIB from the Dell Force10 website, www.force10networks.com.

Note: See the Dell Force10 iSupport webpage for the *Force10-BGP4-V2-MIB* and other MIB documentation.

Important Points to Remember

- In 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 be 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 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.
- Inbound BGP soft-reconfiguration must be configured on a peer for f10BgpM2PrefixInPrefixesRejected to display the number of prefixes filtered due to a policy. If BGP soft-reconfig is not enabled, 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, Originator ID attributes are not stored in the PA Table and cannot be retrieved using the index passed in. These fields are not populated in f10BgpM2PathAttrEntry, f10BgpM2PathAttrClusterEntry, f10BgpM2PathAttrOriginatorIdEntry.
- F10BgpM2PathAttrUnknownEntry contains the optional-transitive attribute details.
- Query for f10BgpM2LinkLocalNextHopEntry returns default value for Link-local Next-hop.
- RFC 2545 and the f10BgpM2Rfc2545Group are not supported.
- An SNMP query will display up to 89 AS paths. A query for a larger AS path count will display 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 will return 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 BGP client-2-client reflection is enabled and that the BGP speaker acting as reflector will advertise 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 Force10 recommends setting the timeout and retry count values to a relatively higher number. e.g. 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 Force10 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.
- F10BgpM2NlriIndex and f10BgpM2AdjRibsOutIndex fields are not used.
- Carrying MPLS labels in BGP is not supported. F10BgpM2NlriOpaqueType and f10BgpM2NlriOpaquePointer fields are set to zero.
- 4-byte ASN is supported. 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, you must assign an AS number and use commands in the ROUTER BGP mode to configure a BGP neighbor.

Defaults

By default, BGP is disabled.

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

Note: In FTOS, all newly configured neighbors and peer groups are disabled. You must enter the **neighbor** {*ip-address* | *peer-group-name*} **no shutdown** command to enable a neighbor or peer group.

Table 8-3 displays the default values for BGP on FTOS.

Table 8-3. FTOS BGP Defaults

Item	Default
BGP Neighbor Adjacency changes	All BGP neighbor changes are logged.
Fast External Fallover feature	Enabled
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

Configuration Task List for BGP

The following list includes the configuration tasks for BGP:

- Enable BGP
- Configure AS4 Number Representations
- Configure Peer Groups
- BGP fast fall-over

- Configure passive peering
- Maintain existing AS numbers during an AS migration
- Allow an AS number to appear in its own AS path
- Enable graceful restart
- Filter on an AS-Path attribute
- Configure IP community lists
- Manipulate the COMMUNITY attribute
- Change MED attribute
- Change LOCAL_PREFERENCE attribute
- Change NEXT_HOP attribute
- Change WEIGHT attribute
- Enable multipath
- Filter BGP routes
- Redistribute routes on page 176
- Configure BGP route reflectors
- Aggregate routes
- Configure BGP confederations
- Enable route flap dampening
- Change BGP timers
- BGP neighbor soft-reconfiguration
- Route map continue

Enable BGP

By default, BGP is not enabled on the system. FTOS supports one Autonomous System (AS) and you must assign the AS Number (ASN). To establish BGP sessions and route traffic, you must configure at least one BGP neighbor or peer.

In BGP, routers with an established TCP connection are called neighbors or peers. Once a connection is established, the neighbors exchange full BGP routing tables with incremental updates afterwards. In addition, neighbors exchange KEEPALIVE messages to maintain the connection.

In BGP, neighbor routers or peers can be classified as internal or external. External BGP peers must be connected physically to one another (unless you enable the EBGP multihop feature), while internal BGP peers do not need to be directly connected. The IP address of an EBGP neighbor is usually the IP address of the interface directly connected to the router. First, the BGP process determines if all internal BGP peers are reachable, and then it determines which peers outside the AS are reachable.

Note: Sample Configurations for enabling BGP routers are found at the end of this chapter.

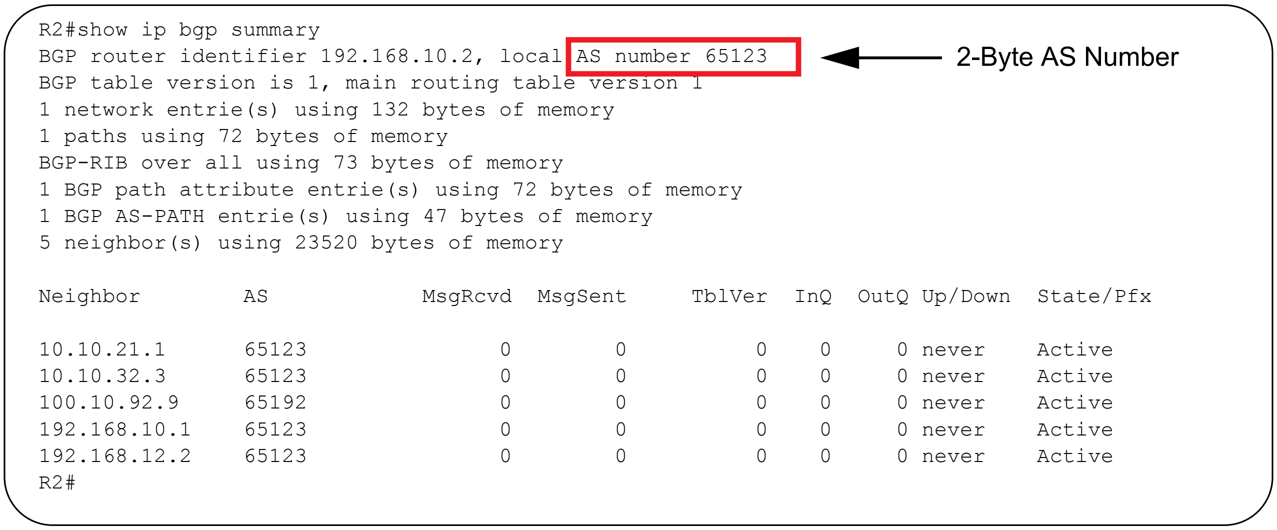
Use these commands in the following sequence, starting in the CONFIGURATION mode to establish BGP sessions on the router.

Step	Command Syntax	Command Mode	Purpose
1	**router bgp** *as-number*	CONFIGURATION	Assign an AS number and enter the ROUTER BGP mode. AS Number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format) Only one AS is supported per system If you enter a 4-Byte AS Number, 4-Byte AS Support is enabled automatically.
1a	bgp four-octet-as-support	CONFIG-ROUTER-BGP	Enable 4-Byte support for the BGP process. **Note:** This is an OPTIONAL command. Enable if you want to use 4-Byte AS numbers or if you support AS4 Number Representation. Use it only if you support 4-Byte AS Numbers or if you support AS4 Number Representation. If you are supporting 4-Byte ASNs, this command must be enabled first. 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 will be displayed in ASPLAIN format.
1b	address-family [ipv4 | ipv6}	CONFIG-ROUTER-BGP	Enable IPv4 multicast or IPv6 mode. Use this command to enter BGP for IPv6 mode (CONF-ROUTER_BGPv6_AF).
2	**neighbor** {*ip-address* | *peer-group name*} **remote-as** *as-number*	CONFIG-ROUTER-BGP	Add a neighbor as a remote AS. Formats: IP Address A.B.C.D Peer-Group Name: 16 characters AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format) You must Configure Peer Groups *before* assigning it a remote AS.
3	**neighbor** {*ip-address* | *peer-group-name*} **no shutdown**	CONFIG-ROUTER-BGP	Enable the BGP neighbor.

Note: When you change the configuration of a BGP neighbor, always reset it by entering the **clear ip bgp** command in EXEC Privilege mode.

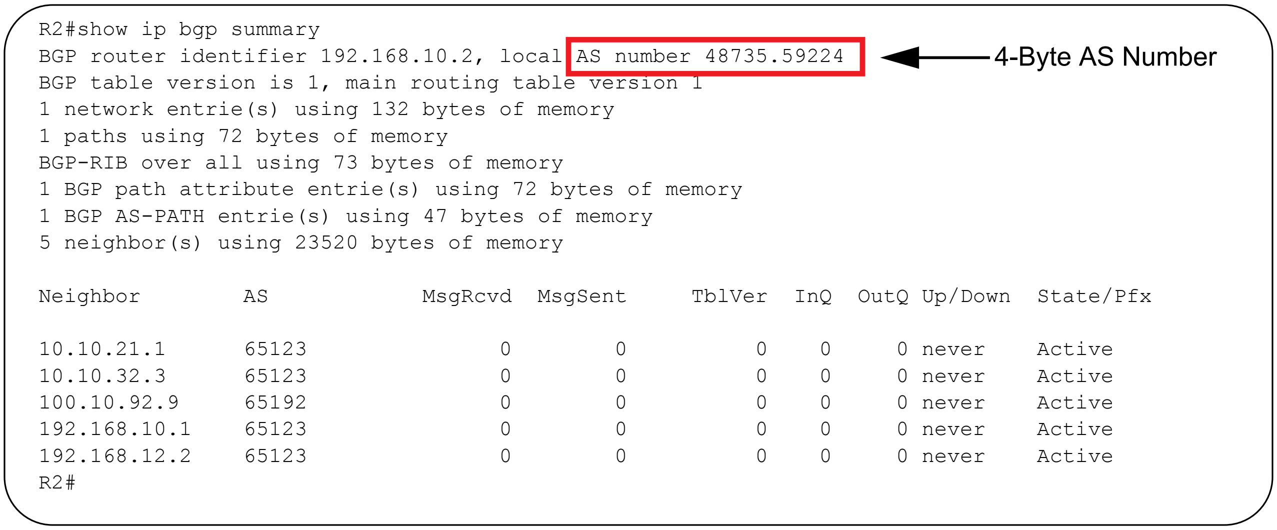
Enter **show config** in CONFIGURATION ROUTER BGP mode to view the BGP configuration. Use the **show ip bgp summary** command in EXEC Privilege mode to view the BGP status. Figure 8-12 shows the summary with a 2-Byte AS Number displayed; Figure 8-13 shows the summary with a 4-Byte AS Number displayed.

Figure 8-12. Command example: show ip bgp summary (2-Byte AS Number displayed)



```
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
100.10.92.9     65192                0        0          0    0      0 never     Active
192.168.10.1    65123                0        0          0    0      0 never     Active
192.168.12.2    65123                0        0          0    0      0 never     Active
R2#
```

Figure 8-13. Command example: show ip bgp summary (4-Byte AS Number displayed)



```
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
100.10.92.9     65192                0        0          0    0      0 never     Active
192.168.10.1    65123                0        0          0    0      0 never     Active
192.168.12.2    65123                0        0          0    0      0 never     Active
R2#
```

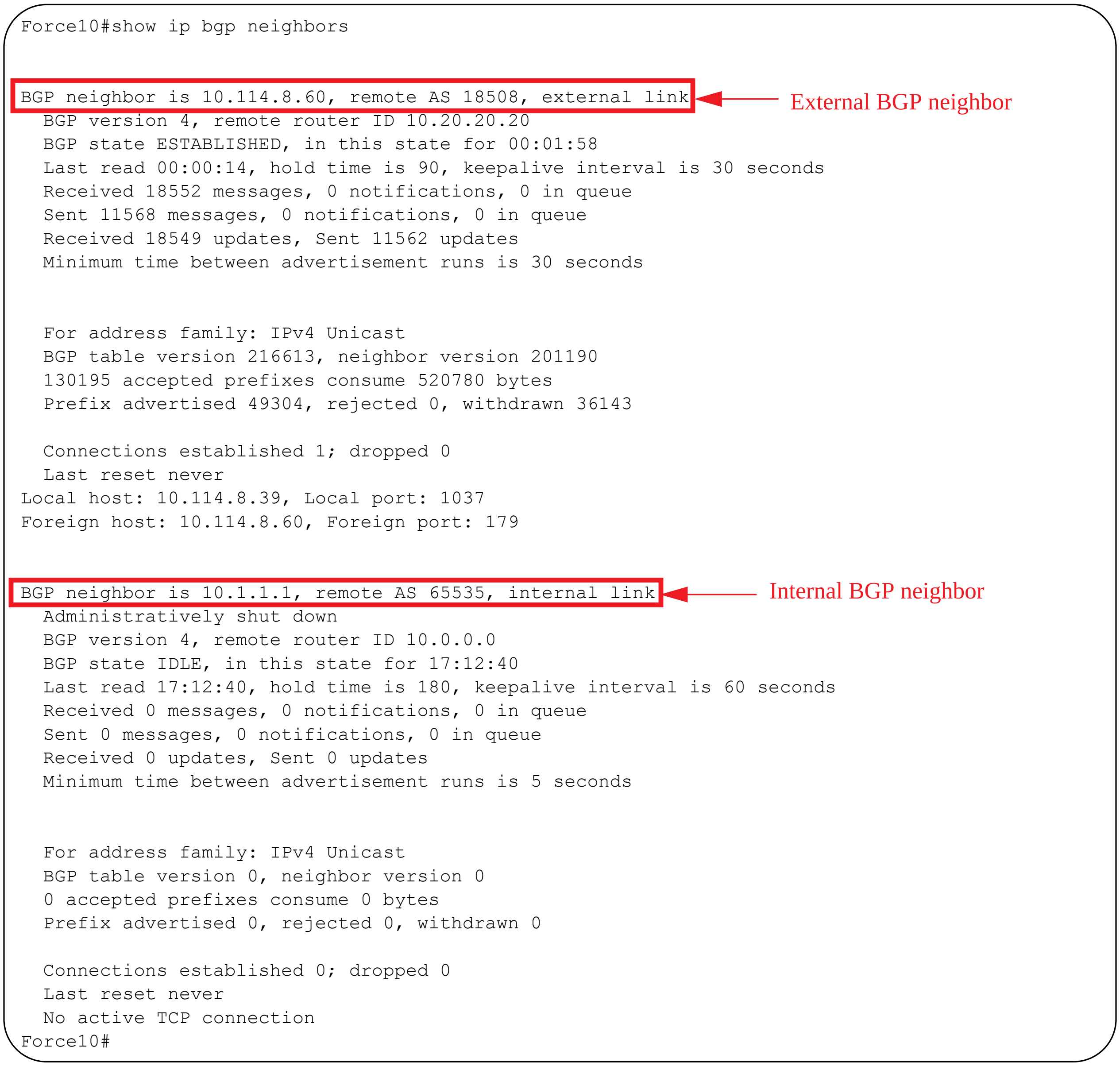
For the router’s identifier, FTOS uses the highest IP address of the Loopback interfaces configured. Since Loopback interfaces are virtual, they cannot go down, thus preventing changes in the router ID. If no Loopback interfaces are configured, the highest IP address of any interface is used as the router ID.

To view the status of BGP neighbors, use the **show ip bgp neighbors** (Figure 8-14) command in EXEC Privilege mode. For BGP neighbor configuration information, use the **show running-config bgp** command in EXEC Privilege mode (Figure 8-15). Note that the **showconfig** command in CONFIGURATION ROUTER BGP mode gives the same information as thew **show running-config bgp**.

Figure 8-14 displays two neighbors, one is an external 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.

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 details on using the **show ip bgp neighbors** command, refer to the *FTOS Command Line Interface Reference*.

Figure 8-14. Command example: show ip bgp neighbors



```
Force10#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
Force10#
```

Figure 8-15. Command example: show running-config bgp

```
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
 neighbor 192.168.10.1 no shutdown
 neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2#
```

Configure AS4 Number Representations

Enable one type of AS Number Representation: ASPLAIN, ASDOT+, or ASDOT.

- ASPLAIN is the method FTOS has used for all previous FTOS versions. It remains the default method with FTOS 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 in conjunction with the 4-Byte AS Numbers feature. If 4-Byte AS Numbers are not implemented, 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.

Task	Command Syntax	Command Mode
Enable ASPLAIN AS Number representation. Figure 8-16	**bgp asnotation asplain**	CONFIG-ROUTER-BGP
	Note: ASPLAIN is the default method FTOS uses and does not appear in the configuration display.	

Task	Command Syntax	Command Mode
Enable ASDOT AS Number representation. Figure 8-17	**bgp asnotation asdot**	CONFIG-ROUTER-BGP
Enable ASDOT+ AS Number representation.Figure 8-18	**bgp asnotation asdot+**	CONFIG-ROUTER-BGP

Figure 8-16. Command example and output: bgp asnotation asplain

```
Force10(conf-router_bgp)#bgp asnotation asplain
Force10(conf-router_bgp)#sho conf
!
router bgp 100
 bgp four-octet-as-support
 neighbor 172.30.1.250 remote-as 18508
 neighbor 172.30.1.250 local-as 65057
 neighbor 172.30.1.250 route-map rmap1 in
 neighbor 172.30.1.250 password 7
5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
 neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Figure 8-17. Command example and output: bgp asnotation asdot

```
Force10(conf-router_bgp)#bgp asnotation asdot
Force10(conf-router_bgp)#sho conf
!
router bgp 100
 bgp asnotation asdot
bgp four-octet-as-support
 neighbor 172.30.1.250 remote-as 18508
 neighbor 172.30.1.250 local-as 65057
 neighbor 172.30.1.250 route-map rmap1 in
 neighbor 172.30.1.250 password 7
5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
 neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Figure 8-18. Command example and output: bgp asnotation asdot+

```
Force10(conf-router_bgp)#bgp asnotation asdot+
Force10(conf-router_bgp)#sho conf
!
router bgp 100
 bgp asnotation asdot+
bgp four-octet-as-support
 neighbor 172.30.1.250 remote-as 18508
 neighbor 172.30.1.250 local-as 65057
 neighbor 172.30.1.250 route-map rmap1 in
 neighbor 172.30.1.250 password 7
5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
 neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Configure Peer Groups

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

A *maximum* of 256 Peer Groups are allowed on the system.

You create a peer group by assigning it a name, then adding members to the peer group. Once a peer group is created, you can configure route policies for it. Refer to Filter BGP routes for information on configuring route policies for a peer group.

Note: Sample Configurations for enabling Peer Groups are found at the end of this chapter.

Use these commands in the following sequence starting in the CONFIGURATION ROUTER BGP mode to create a peer group

Step	Command Syntax	Command Mode	Purpose
1	**neighbor** *peer-group-name* **peer-group**	CONFIG-ROUTER-BGP	Create a peer group by assigning a name to it.
2	**neighbor** *peer-group-name* **no shutdown**	CONFIG-ROUTER-BGP	Enable the peer group. By default, all peer groups are disabled
3	**neighbor** *ip-address* **remote-as** *as-number*	CONFIG-ROUTER-BGP	Create a BGP neighbor.
4	**neighbor** *ip-address* **no shutdown**	CONFIG-ROUTER-BGP	Enable the neighbor.
5	**neighbor** *ip-address* **peer-group** *peer-group-name*	CONFIG-ROUTER-BGP	Add an enabled neighbor to the peer group.
6	**neighbor** {*ip-address* | *peer-group name*} **remote-as** *as-number*	CONFIG-ROUTER-BGP	Add a neighbor as a remote AS. Formats: IP Address A.B.C.D Peer-Group Name16 characters AS-number: 0-65535 (2-Byte) or 1-4294967295 | 0.1- 65535.65535 (4-Byte) or 0.1-65535.65535 (Dotted format)
		To add an external BGP (EBGP) neighbor, configure the *as-number* parameter with a number *different* from the BGP *as-number* configured in the **router bgp** *as-number* command. To add an internal BGP (IBGP neighbor, configure the *as-number* parameter with the *same* BGP *as-number* configured in the **router bgp** *as-number* command.	

After you create a peer group, you can use any of the commands beginning with the keyword **neighbor** to configure that peer group.

When you add a peer to a peer group, it inherits all the peer group's configured parameters.

A neighbor *cannot* become part of a peer group if it has any of the following commands are configured:

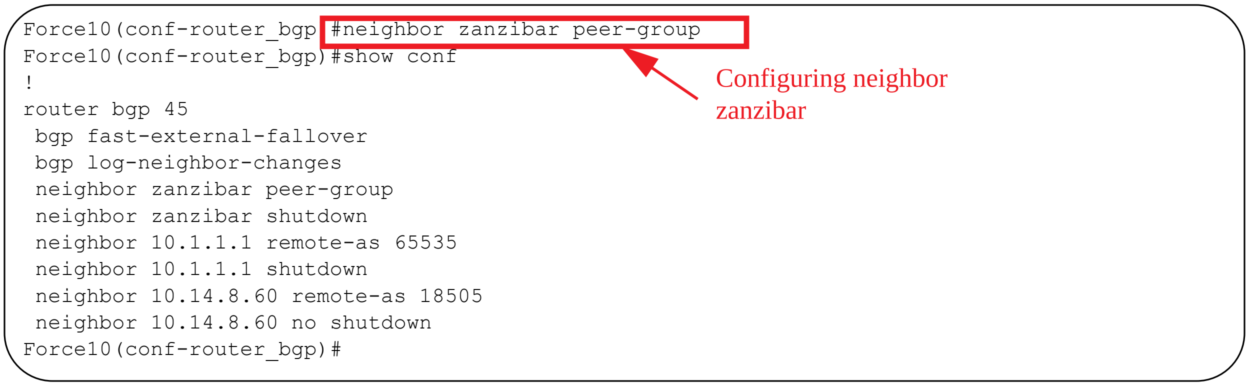
- neighbor advertisement-interval
- neighbor distribute-list out
- neighbor filter-list out
- neighbor next-hop-self
- neighbor route-map out
- neighbor route-reflector-client
- neighbor send-community

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 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.

Use the **show config** command in the CONFIGURATION ROUTER BGP mode to view the configuration. When you create a peer group, it is disabled (**shutdown**). Figure 8-19 shows the creation of a peer group (zanzibar).

Figure 8-19. Command example: show config (creating peer-group)

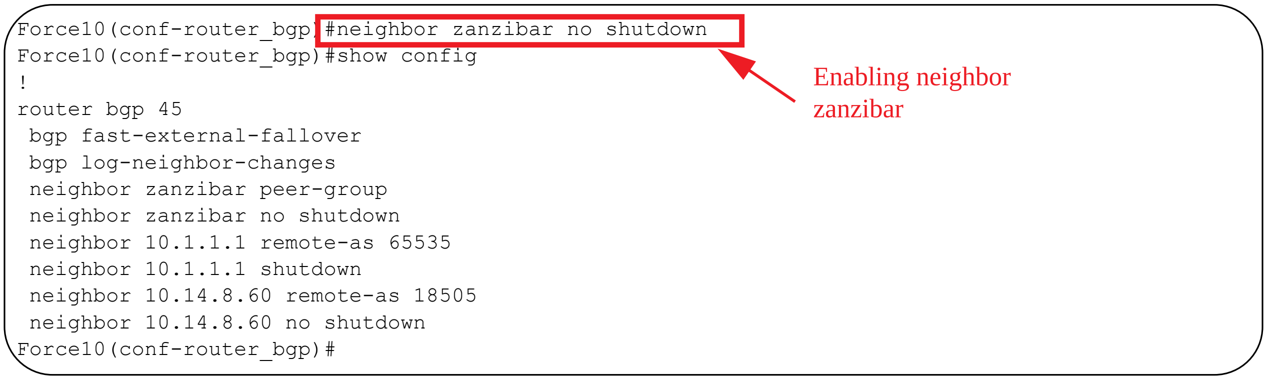


```
Force10(conf-router_bgp)#neighbor zanzibar peer-group
Force10(conf-router_bgp)#show conf
!
router bgp 45
 bgp fast-external-fallover
 bgp log-neighbor-changes
 neighbor zanzibar peer-group
 neighbor zanzibar shutdown
 neighbor 10.1.1.1 remote-as 65535
 neighbor 10.1.1.1 shutdown
 neighbor 10.14.8.60 remote-as 18505
 neighbor 10.14.8.60 no shutdown
Force10(conf-router_bgp)#
```

Configuring neighbor zanzibar

Use the **neighbor peer-group-name no shutdown** command in the CONFIGURATION ROUTER BGP mode to enable a peer group.

Figure 8-20. Command example: show config (peer-group enabled



```
Force10(conf-router_bgp)#neighbor zanzibar no shutdown
Force10(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
 neighbor 10.14.8.60 remote-as 18505
 neighbor 10.14.8.60 no shutdown
Force10(conf-router_bgp)#
```

Enabling neighbor zanzibar

To disable a peer group, use the **neighbor** *peer-group-name* **shutdown** command in the 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 ESTABLISHED state are moved to IDLE state.

Use the show **ip bgp peer-group** command in EXEC Privilege mode (Figure 8-21) to view the status of peer groups.

Figure 8-21. Command example: show ip bgp peer-group

```
Force10>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
  10.68.179.1
  10.68.180.1
  10.68.181.1
  10.68.182.1
  10.68.183.1
  10.68.184.1
  10.68.185.1
Force10>
```

BGP fast fall-over

By default, a BGP session is governed by the hold time. BGP routers typically carry large routing tables, so frequent session resets are not desirable. The BGP fast fall-over feature reduces the convergence time while maintaining stability. The connection to a BGP peer is immediately reset if a link to a directly connected external peer fails.

When fall-over is enabled, BGP tracks IP reachability to the peer remote address and the peer local address. Whenever either address becomes unreachable (for example, no active route exists in the routing table for peer IPv6 destinations/local address), BGP brings down the session with the peer.

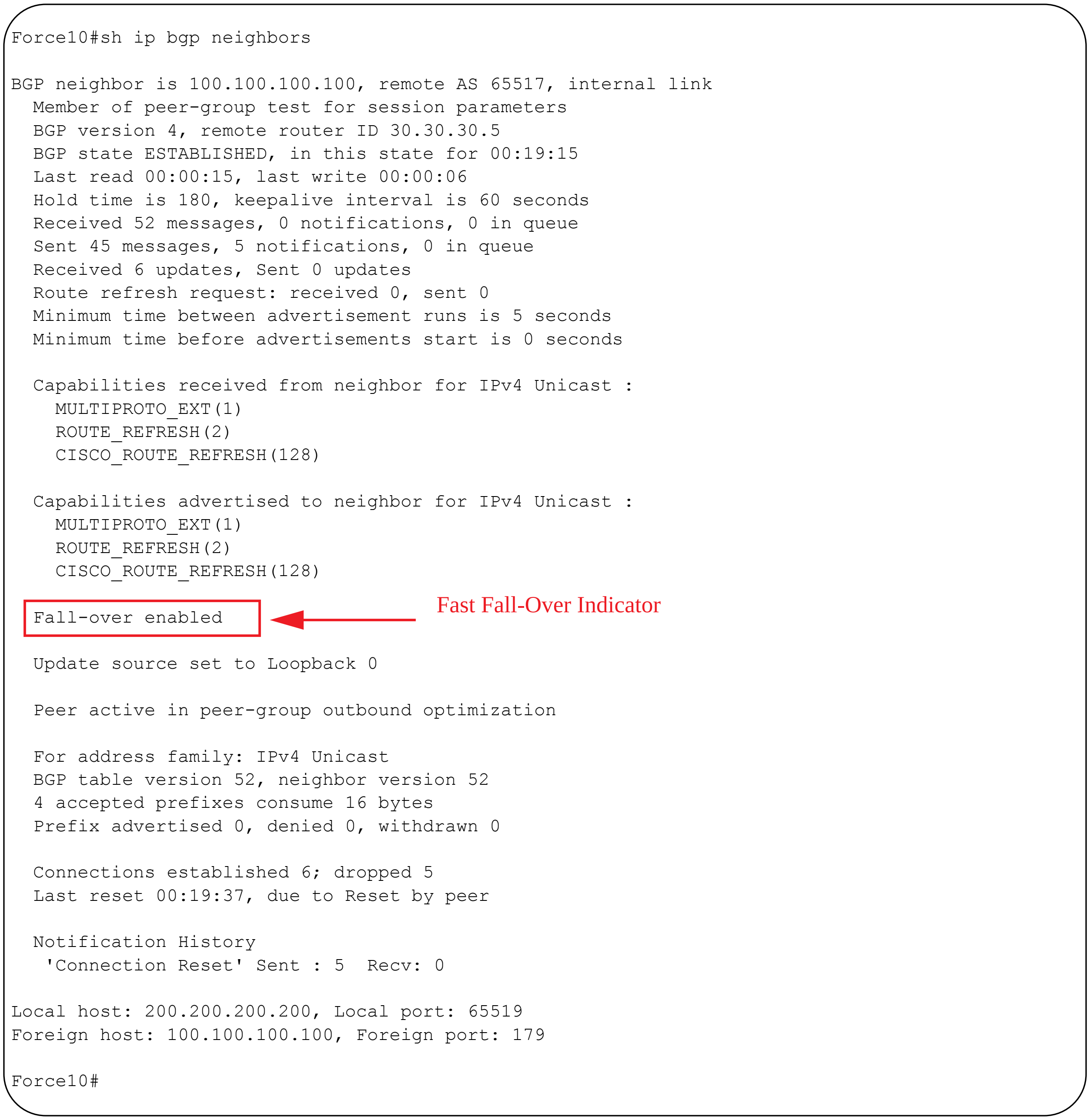
The BGP fast fall-over feature is configured on a per-neighbor or peer-group basis and is disabled by default.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **fall-over**	CONFIG-ROUTER-BGP	Enable BGP Fast Fall-Over

To disable Fast Fall-Over, use the **[no] neighbor [neighbor | peer-group] fall-over** command in CONFIGURATION ROUTER BGP mode

Use the **show ip bgp neighbors** command as shown in Figure 8-22 to verify that fast fall-over is enabled on a particular BGP neighbor. Note that since Fast Fall-Over is disabled by default, it will appear only if it has been enabled

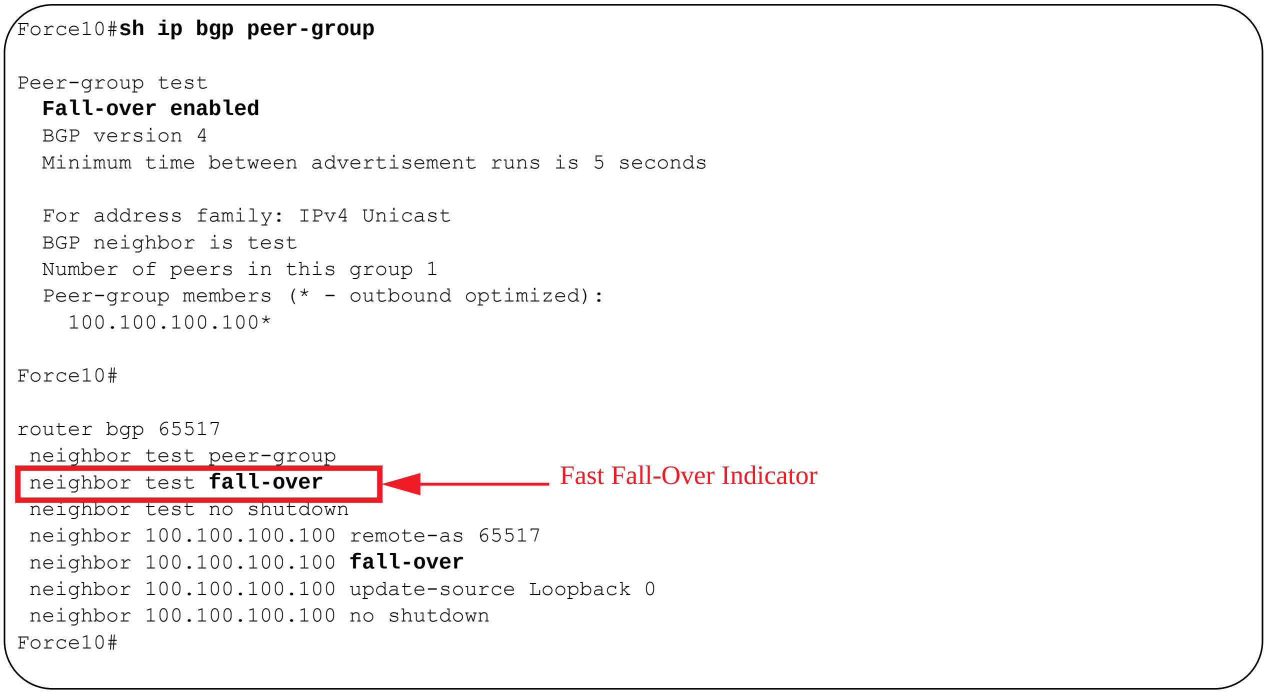
Figure 8-22. Command example: show ip bgp neighbors



```
Force10#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)

  Fall-over enabled

  Update source set to Loopback 0

  Peer active in peer-group outbound optimization

  For address family: IPv4 Unicast
  BGP table version 52, neighbor version 52
  4 accepted prefixes consume 16 bytes
  Prefix advertised 0, denied 0, withdrawn 0

  Connections established 6; dropped 5
  Last reset 00:19:37, due to Reset by peer

  Notification History
   'Connection Reset' Sent : 5  Recv: 0

Local host: 200.200.200.200, Local port: 65519
Foreign host: 100.100.100.100, Foreign port: 179

Force10#
```

Use the **show ip bgp peer-group** command to verify that fast fall-over is enabled on a peer-group.

Figure 8-23. Command example: show ip bgp peer-group



```
Force10#sh ip bgp peer-group

Peer-group test
  Fall-over enabled
  BGP version 4
  Minimum time between advertisement runs is 5 seconds

  For address family: IPv4 Unicast
  BGP neighbor is test
  Number of peers in this group 1
  Peer-group members (* - outbound optimized):
    100.100.100.100*

Force10#

router bgp 65517
 neighbor test peer-group
 neighbor test fall-over
 neighbor test no shutdown
 neighbor 100.100.100.100 remote-as 65517
 neighbor 100.100.100.100 fall-over
 neighbor 100.100.100.100 update-source Loopback 0
 neighbor 100.100.100.100 no shutdown
Force10#
```

Configure 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 will respond to an OPEN message.

When a BGP neighbor connection with authentication configured is rejected by a passive peer-group, FTOS 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.

Use these commands in the following sequence, starting in the CONFIGURATION ROUTER BGP mode to configure passive peering.

Step	Command Syntax	Command Mode	Purpose
1	**neighbor** *peer-group-name* **peer-group passive**	CONFIG-ROUTER-BGP	Configure a peer group that does not initiate TCP connections with other peers.
2	**neighbor** *peer-group-name* **subnet** *subnet-number mask*	CONFIG-ROUTER-BGP	Assign a subnet to the peer group. The peer group will respond to OPEN messages sent on this subnet.
3	**neighbor** *peer-group-name* **no shutdown**	CONFIG-ROUTER-BGP	Enable the peer group.
4	**neighbor** *peer-group-name* **remote-as** *as-number*	CONFIG-ROUTER-BGP	Create and specify a remote peer for BGP neighbor.

Only after the peer group responds to an OPEN message sent on the subnet does its BGP state change to ESTABLISHED. Once the peer group is ESTABLISHED, the peer group is the same as any other peer group.

For more information on peer groups, refer to Configure Peer Groups on page 162.

Maintain existing AS numbers during an AS migration

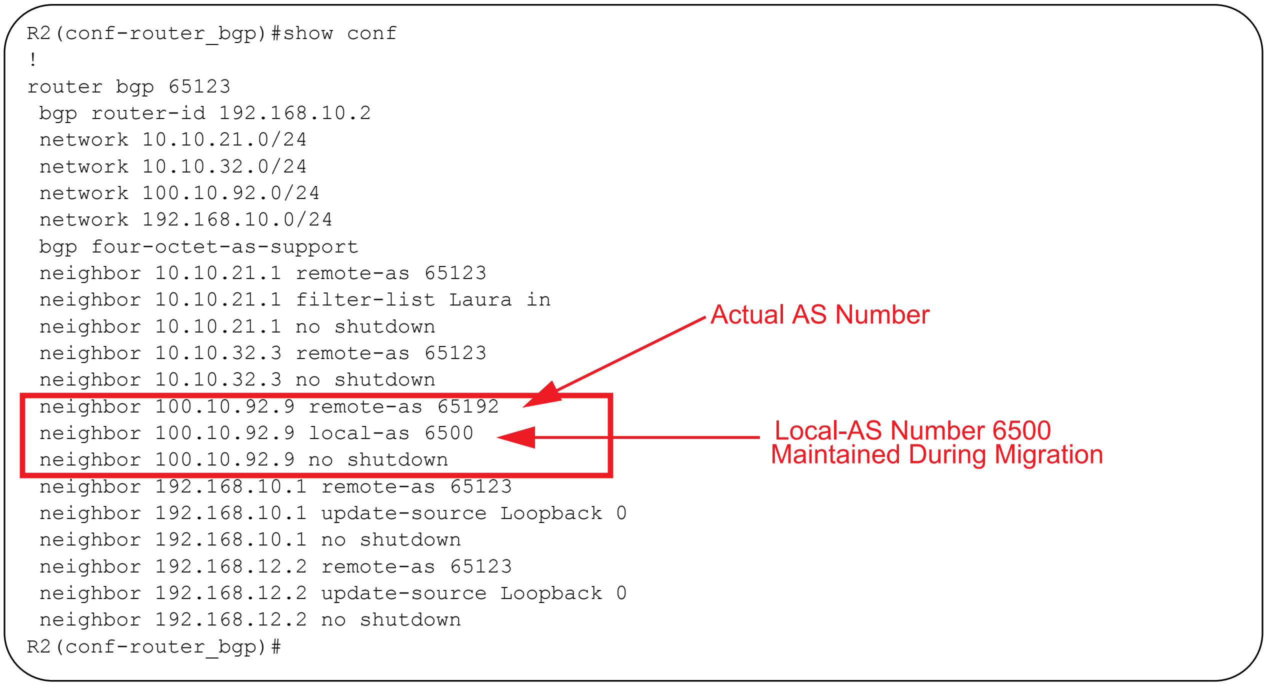
The **local-as** feature smooths out the BGP network migration operation and allows you to maintain existing ASNs during a BGP network migration.

When you complete your migration, be sure to reconfigure your routers with the new information and disable this feature.

Command Syntax	Command Mode	Purpose
neighbor {*IP address* | *peer-group-name* **local-as** *as number* [no prepend]	CONFIG-ROUTER-BGP	Allow external routes from this neighbor. Format: IP Address: A.B.C.D Peer Group Name: 16 characters AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format) No Prepend specifies that local AS values are not prepended to announcements from the neighbor.
	You must Configure Peer Groups *before* assigning it to an AS. This feature is not supported on passive peer groups.	

Disable this feature, using the **no neighbor local-as** command in CONFIGURATION ROUTER BGP mode.

Figure 8-24. Local-as information shown



```
R2(conf-router_bgp)#show conf
!
router bgp 65123
 bgp router-id 192.168.10.2
 network 10.10.21.0/24
 network 10.10.32.0/24
 network 100.10.92.0/24
 network 192.168.10.0/24
 bgp four-octet-as-support
 neighbor 10.10.21.1 remote-as 65123
 neighbor 10.10.21.1 filter-list Laura in
 neighbor 10.10.21.1 no shutdown
 neighbor 10.10.32.3 remote-as 65123
 neighbor 10.10.32.3 no shutdown
 neighbor 100.10.92.9 remote-as 65192
 neighbor 100.10.92.9 local-as 6500
 neighbor 100.10.92.9 no shutdown
 neighbor 192.168.10.1 remote-as 65123
 neighbor 192.168.10.1 update-source Loopback 0
 neighbor 192.168.10.1 no shutdown
 neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2(conf-router_bgp)#
```

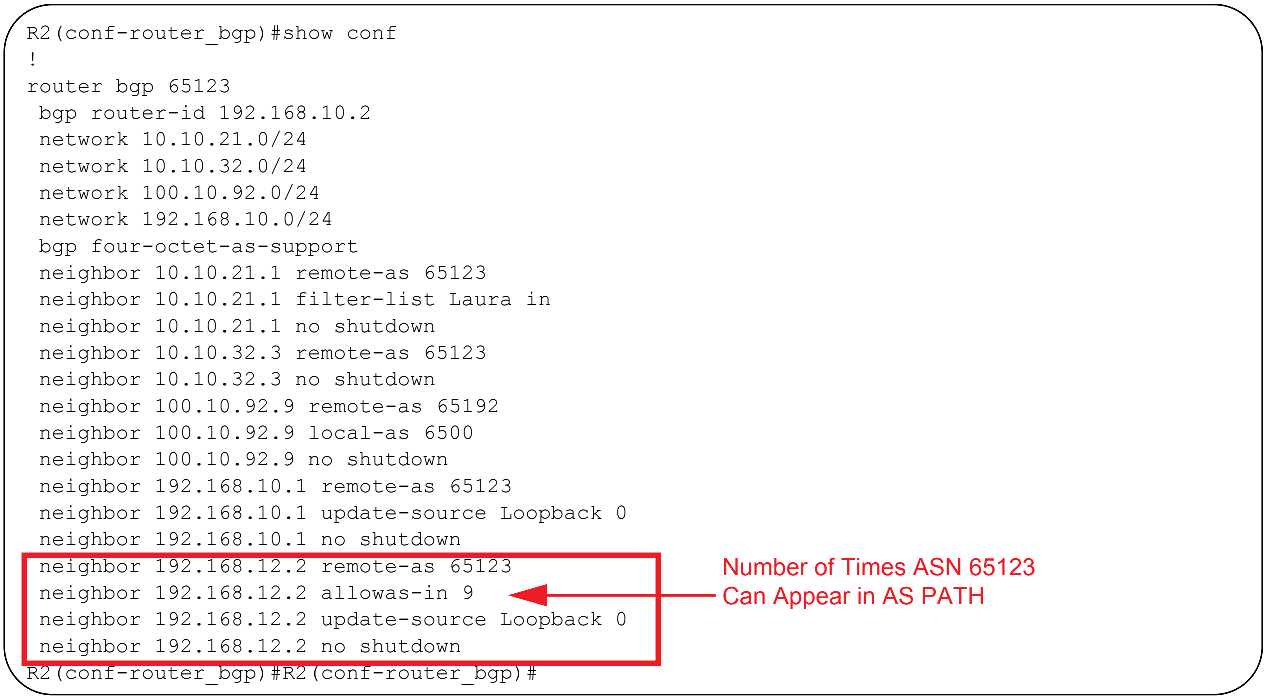
Allow an AS number to appear in its own AS path

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 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.

Command Syntax	Command Mode	Purpose
neighbor {*IP address* | *peer-group-name*} **allowas-in** *number*	CONFIG-ROUTER-BGP	Allow this neighbor ID to use the AS path the specified number of times. Format: IP Address: A.B.C.D Peer Group Name: 16 characters Number: 1-10
	You must Configure Peer Groups *before* assigning it to an AS.	

To disable this feature, use the **no neighbor allow-as in** *number* command in the CONFIGURATION ROUTER BGP mode.

Figure 8-25. Allowas-in information shown



```
R2(conf-router_bgp)#show conf
!
router bgp 65123
 bgp router-id 192.168.10.2
 network 10.10.21.0/24
 network 10.10.32.0/24
 network 100.10.92.0/24
 network 192.168.10.0/24
 bgp four-octet-as-support
 neighbor 10.10.21.1 remote-as 65123
 neighbor 10.10.21.1 filter-list Laura in
 neighbor 10.10.21.1 no shutdown
 neighbor 10.10.32.3 remote-as 65123
 neighbor 10.10.32.3 no shutdown
 neighbor 100.10.92.9 remote-as 65192
 neighbor 100.10.92.9 local-as 6500
 neighbor 100.10.92.9 no shutdown
 neighbor 192.168.10.1 remote-as 65123
 neighbor 192.168.10.1 update-source Loopback 0
 neighbor 192.168.10.1 no shutdown
 neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 allowas-in 9
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2(conf-router_bgp)#R2(conf-router_bgp)#
```

Number of Times ASN 65123
Can Appear in AS PATH

Enable graceful restart

Use this feature to lessen the negative effects of a BGP restart. FTOS 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 on is as a receiving or restarting peer. If you enable BGP, when a peer that supports graceful restart resumes operating, FTOS 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, FTOS can perform the following actions during a hot failover:

- Save all FIB and CAM entries on the line card and continue forwarding traffic while the secondary 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 from your E-Series 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.

Enable graceful restart using the **configure router bgp graceful-restart** command. The table below shows the command and its available options:

Command Syntax	Command Mode	Usage
bgp graceful-restart	CONFIG-ROUTER-BGP	Enable graceful restart for the BGP node.
bgp graceful-restart [**restart-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum restart time for all peers. Default is 120 seconds.
bgp graceful-restart [**stale-path-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum time to retain the restarting peer’s stale paths. Default is 360 seconds.
bgp graceful-restart [**role receiver-only**]	CONFIG-ROUTER-BGP	Local router supports graceful restart as a receiver only.

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, FTOS 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. However, the E-Series does not advertise that it saves these forwarding states when it restarts. 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, please see the following table or the *FTOS Command Line Interface Reference*.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **graceful-restart**	CONFIG-ROUTER-BGP	Add graceful restart to a BGP neighbor or peer-group.
neighbor {*ip-address* | *peer-group-name*} **graceful-restart** [**restart-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum restart time for the neighbor or peer-group. Default is 120 seconds.
neighbor {*ip-address* | *peer-group-name*} **graceful-restart** [**role receiver-only**]	CONFIG-ROUTER-BGP	Local router supports graceful restart for this neighbor or peer-group as a receiver only.
neighbor {*ip-address* | *peer-group-name*} **graceful-restart** [**stale-path-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum time to retain the restarting neighbor’s or peer-group’s stale paths. Default is 360 seconds.

Filter on an AS-Path attribute

The BGP attribute, AS_PATH, can be used to manipulate routing policies. The AS_PATH attribute contains a sequence of AS numbers representing the route's path. As the route traverses an Autonomous System, the AS number is prepended to the route. You can manipulate routes based on their AS_PATH to affect interdomain routing. By identifying certain AS numbers in the AS_PATH, you can permit or deny routes based on the number in its AS_PATH.

To view all BGP path attributes in the BGP database, use the **show ip bgp paths** command in EXEC Privilege mode (Figure 8-26).

Figure 8-26. Command example: show ip bgp paths

```
Force10#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--
```

AS-PATH ACLs use regular expressions to search AS_PATH values. AS-PATH ACLs have an "implicit deny." This means that routes that do not meet a deny or match filter are dropped.

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an AS-PATH ACL to filter a specific AS_PATH value.

Step	Command Syntax	Command Mode	Purpose
1	**ip as-path access-list** *as-path-name*	CONFIGURATION	Assign a name to a AS-PATH ACL and enter AS-PATH ACL mode.

Step	Command Syntax	Command Mode	Purpose
2	{**deny** | **permit**} *filter parameter*	CONFIG-AS-PATH	Enter the parameter to match BGP AS-PATH for filtering. This is the filter that will be used to match the AS-path. The entries can be any format, letters, numbers, or regular expressions. This command can be entered multiple times if multiple filters are desired. See Table 8-4 for accepted expressions.
3	**exit**	AS-PATH ACL	Return to CONFIGURATION mode
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **filter-list** *as-path-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Use a configured AS-PATH ACL for route filtering and manipulation. If you assign an non-existent or empty AS-PATH ACL, the software allows all routes.

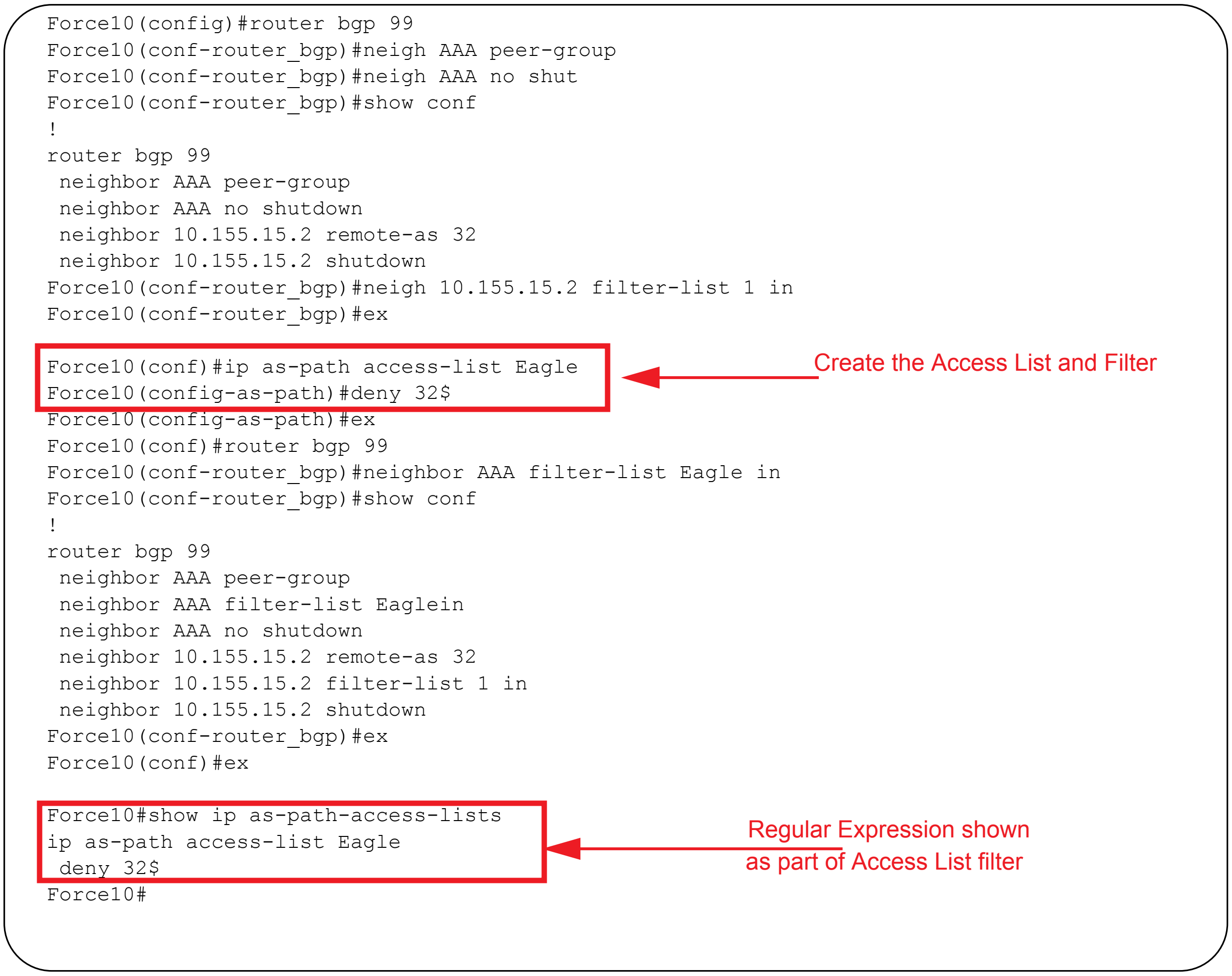
Regular Expressions as filters

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 commands above, if the AS path matches the regular expression in the access list, then the route matches the access list.

Figure 8-27 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.

Figure 8-27. Filtering with Regular Expression



```
Force10(config)#router bgp 99
Force10(conf-router_bgp)#neigh AAA peer-group
Force10(conf-router_bgp)#neigh AAA no shut
Force10(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
Force10(conf-router_bgp)#neigh 10.155.15.2 filter-list 1 in
Force10(conf-router_bgp)#ex

Force10(conf)#ip as-path access-list Eagle
Force10(config-as-path)#deny 32$
Force10(config-as-path)#ex
Force10(conf)#router bgp 99
Force10(conf-router_bgp)#neighbor AAA filter-list Eagle in
Force10(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
Force10(conf-router_bgp)#ex
Force10(conf)#ex

Force10#show ip as-path-access-lists
ip as-path access-list Eagle
 deny 32$
Force10#
```

Create the Access List and Filter

Regular Expression shown as part of Access List filter

Table 8-4 lists the Regular Expressions accepted in FTOS.

Table 8-4. Regular Expressions

Regular Expression	Definition
^ (carrot)	Matches the beginning of the input string. Alternatively, when used as the first character within brackets [^] 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.

Table 8-4. Regular Expressions

Regular Expression	Definition
( ) (parenthesis)	Specifies patterns for multiple use when followed by one of the multiplier metacharacters: asterisk *, plus sign +, or question mark ?
[ ] (brackets)	Matches any enclosed character; 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, a {, or a }. Placed on either side of a string to specify a literal and disallow substring matching. Numerals enclosed by underscores can be preceded or followed by any of the characters listed above.
| (pipe)	Matches characters on either side of the metacharacter; logical OR.

As seen in Figure 8-27, the expressions are displayed when using the **show** commands. Use the **show config** command in the CONFIGURATION AS-PATH ACL mode and the **show ip as-path-access-list** command in EXEC Privilege mode to view the AS-PATH ACL configuration.

For more information on this command and route filtering, refer to Filter BGP routes.

Redistribute routes

In addition to filtering routes, you can add routes from other routing instances or protocols to the BGP process. With the **redistribute** command syntax, you can include ISIS, OSPF, static, or directly connected routes in the BGP process.

Use any of the following commands in ROUTER BGP mode to add routes from other routing instances or protocols.

Command Syntax	Command Mode	Purpose
redistribute {**connected** | **static**} [**route-map** *map-name*]	ROUTER BGP or CONF-ROUTER_BGPv6_AF	Include, directly connected or user-configured (static) routes in BGP. Configure the following parameters: • *map-name*: name of a configured route map.
redistribute isis [**level-1** | **level-1-2** | **level-2**] [**metric** *value*] [**route-map** *map-name*]	ROUTER BGP or CONF-ROUTER_BGPv6_AF	Include specific ISIS routes in BGP. Configure the following parameters: • **level-1**, l**evel-1-2**, or **level-2**: Assign all redistributed routes to a level. Default is **level-2**. • *metric* range: 0 to 16777215. Default is 0. • *map-name*: name of a configured route map.

Command Syntax	Command Mode	Purpose
redistribute ospf *process-id* [**match external** {**1 | 2**} | **match internal**] [**metric-type** {**external** | **internal**}] [**route-map** *map-name*]	ROUTER BGP or CONF-ROUTER_BGPv6_ AF	Include specific OSPF routes in IS-IS. Configure the following parameters: • *process-id* range: 1 to 65535 • **match external** range: 1 or 2 • **match internal** • **metric-type**: external or internal. • *map-name*: name of a configured route map.

Configure IP community lists

Within FTOS, you have multiple methods of manipulating routing attributes. One attribute you can manipulate is the COMMUNITY attribute. This attribute is an optional attribute that is defined for a group of destinations. In FTOS, 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 pre-defined communities of INTERNET, NO_EXPORT_SUBCONFED, NO_ADVERTISE, and NO_EXPORT. All BGP routes belong to the INTERNET community. In the RFC, the other communities are defined as follows:

- All routes with the NO_EXPORT_SUBCONFED (0xFFFFFF03) community attribute are not sent to CONFED-EBGP or EBGP peers, but are sent to IBGP peers within CONFED-SUB-AS.
- All routes with the NO_ADVERTISE (0xFFFFFF02) community attribute must not be advertised.
- All routes with the NO_EXPORT (0xFFFFFF01) community attribute must not be advertised outside a BGP confederation boundary, but are sent to CONFED-EBGP and IBGP peers.

FTOS also supports BGP Extended Communities as described in RFC 4360—BGP Extended Communities Attribute.

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an IP community list.

Step	Command Syntax	Command Mode	Purpose
1	**ip community-list** *community-list-name*	CONFIGURATION	Create a Community list and enter the COMMUNITY-LIST mode.
2	{**deny** | **permit**} {*community-number* | **local-AS** | **no-advertise** | **no-export** | **quote-regexp** *regular-expression-list* | **regexp** *regular-expression*}	CONFIG-COMMUNITY-LIST	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. • **no-advertise:** routes with the COMMUNITY attribute of NO_ADVERTISE. • **no-export:** routes with the COMMUNITY attribute of NO_EXPORT. • **quote-regexp:** followed by any number of regular expressions. The software applies all regular expressions in the list. • **regexp:** followed by a regular expression.

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an IP extended community list.

Step	Command Syntax	Command Mode	Purpose
1	**ip extcommunity-list** *extcommunity-list-name*	CONFIGURATION	Create a extended community list and enter the EXTCOMMUNITY-LIST mode.
2	{**permit** | **deny**} {{**rt** | **soo**} {*ASN:NN* | *IPADDR:N*} | **regex** *REGEX-LINE*}	CONFIG-COMMUNITY-LIST	Two types of extended communities are supported. Filter routes based on the type of extended communities they carry using one of the following keywords: • **rt**: Route Target • **soo**: Route Origin or Site-of-Origin. Support for matching extended communities against regular expression is also supported. Match against a regular expression using the following keyword: • **regexp**: regular expression

To set or modify an extended community attribute, use the **set extcommunity** {**rt** | **soo**} {*ASN:NN* | *IPADDR:NN*} command.

To view the configuration, use the **show config** command in the CONFIGURATION COMMUNITY-LIST or CONFIGURATION EXTCOMMUNITY LIST mode or the **show ip {community-lists | extcommunity-list} command** in EXEC Privilege mode (Figure 8-28).

Figure 8-28. Command example: show ip community-lists

```
Force10#show ip community-lists
ip community-list standard 1
 deny 701:20
 deny 702:20
 deny 703:20
 deny 704:20
 deny 705:20
 deny 14551:20
 deny 701:112
 deny 702:112
 deny 703:112
 deny 704:112
 deny 705:112
 deny 14551:112
 deny 701:667
 deny 702:667
 deny 703:667
```

Use these commands in the following sequence, starting in the CONFIGURATION mode, To use an IP Community list or 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.

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Enter the ROUTE-MAP mode and assign a name to a route map.
2	**match** {**community** *community-list-name* [**exact**] **|** **extcommunity** *extcommunity-list-name* [**exact**]}	CONFIG-ROUTE-MAP	Configure a match filter for all routes meeting the criteria in the IP Community or Extended Community list.
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter the ROUTER BGP mode. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format)
5	**neighbor** {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Apply the route map to the neighbor or peer group's incoming or outgoing 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.

To view which BGP routes meet an IP Community or Extended Community list's criteria, use the **show ip bgp** {**community-list | extcommunity-list} command** in EXEC Privilege mode.

Manipulate 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, FTOS does not send the COMMUNITY attribute.

Use the following command in the CONFIGURATION ROUTER BGP mode to send the COMMUNITY attribute to BGP neighbors.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **send-community**	CONFIG-ROUTER-BGP	Enable the software to send the router's COMMUNITY attribute to the BGP neighbor or peer group specified.

To view the BGP configuration, use the **show config** command in the 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. Use these commands in the following sequence, starting in the CONFIGURATION mode:

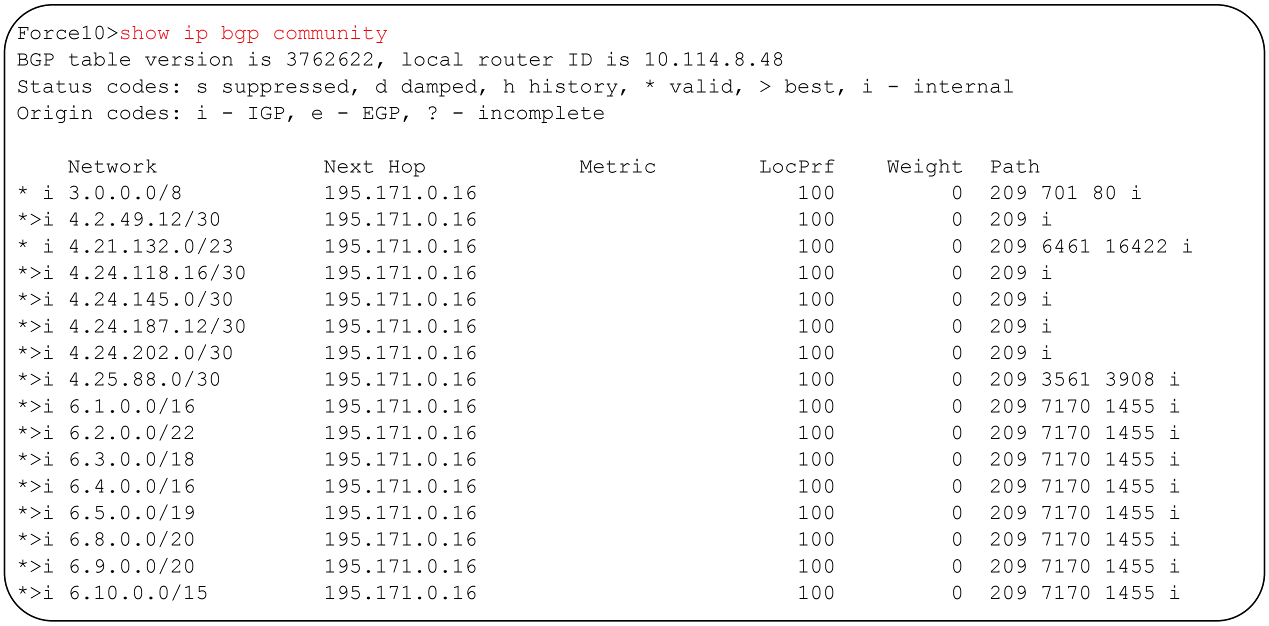
Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Enter the ROUTE-MAP mode and assign a name to a route map.
2	**set comm-list** *community-list-name* **delete**	CONFIG-ROUTE-MAP	Configure a set filter to delete all COMMUNITY numbers in the IP Community list.
	set community {*community-number* | **local-as** | **no-advertise** | **no-export** | **none**}	CONFIG-ROUTE-MAP	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.

Step	Command Syntax	Command Mode	Purpose
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter the ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Apply the route map to the neighbor or peer group's incoming or outgoing 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.

Use the **show ip bgp community** command in EXEC Privilege mode (Figure 8-29) to view BGP routes matching a certain community number or pre-defined BGP community.

Figure 8-29. Command example: show ip bgp community (Partial)



```
Force10>show ip bgp community
BGP table version is 3762622, local router ID is 10.114.8.48
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

    Network              Next Hop            Metric        LocPrf    Weight  Path
* i 3.0.0.0/8            195.171.0.16                         100         0  209 701 80 i
*>i 4.2.49.12/30         195.171.0.16                         100         0  209 i
* i 4.21.132.0/23        195.171.0.16                         100         0  209 6461 16422 i
*>i 4.24.118.16/30       195.171.0.16                         100         0  209 i
*>i 4.24.145.0/30        195.171.0.16                         100         0  209 i
*>i 4.24.187.12/30       195.171.0.16                         100         0  209 i
*>i 4.24.202.0/30        195.171.0.16                         100         0  209 i
*>i 4.25.88.0/30         195.171.0.16                         100         0  209 3561 3908 i
*>i 6.1.0.0/16           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.2.0.0/22           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.3.0.0/18           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.4.0.0/16           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.5.0.0/19           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.8.0.0/20           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.9.0.0/20           195.171.0.16                         100         0  209 7170 1455 i
*>i 6.10.0.0/15          195.171.0.16                         100         0  209 7170 1455 i
```

Change MED attribute

By default, FTOS uses the MULTI_EXIT_DISC or MED attribute when comparing EBGP paths from the same AS.

Use any or all of the following commands in the CONFIGURATION ROUTER BGP mode to change how the MED attribute is used.

Command Syntax	Command Mode	Purpose
bgp always-compare-med	CONFIG-ROUTER-BGP	Enable MED comparison in the paths from neighbors with different ASs. By default, this comparison is not performed.
bgp bestpath med {confed | missing-as-best}	CONFIG-ROUTER-BGP	Change the bestpath MED selection to one of the following: **confed**: Chooses the bestpath MED comparison of paths learned from BGP confederations. **missing-as-bes**t: Treat a path missing an MED as the most preferred one

Use the **show config** command in the CONFIGURATION ROUTER BGP mode to view the nondefault values.

Change LOCAL_PREFERENCE attribute

In FTOS, you can change the value of the LOCAL_PREFERENCE attribute.

Use the following command in the CONFIGURATION ROUTER BGP mode to change the default values of this attribute for all routes received by the router.

Command Syntax	Command Mode	Purpose
bgp default local-preference *value*	CONFIG-ROUTER-BGP	Change the LOCAL_PREF value. • *value* range: 0 to 4294967295 • Default is 100.

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view BGP configuration.

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

Use these commands in the following sequence, starting CONFIGURATION mode to change the default value of the LOCAL_PREF attribute for specific routes.

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Enter the ROUTE-MAP mode and assign a name to a route map.
2	**set local-preference** *value*	CONFIG-ROUTE-MAP	Change LOCAL_PREF value for routes meeting the criteria of this route map.
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.

Step	Command Syntax	Command Mode	Purpose
4	**router bgp** *as-number*	CONFIGURATION	Enter the ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Apply the route map to the neighbor or peer group's incoming or outgoing routes.

To view the BGP configuration, use the **show config** command in the CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the **show route-map** command in EXEC Privilege mode.

Change NEXT_HOP attribute

You can change how the NEXT_HOP attribute is used.

Use the following command in the CONFIGURATION ROUTER BGP mode to change the how the NEXT_HOP attribute is used.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **next-hop-self**	CONFIG-ROUTER-BGP	Disable next hop processing and configure the router as the next hop for a BGP neighbor.

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view BGP configuration.

You can also use route maps to change this and other BGP attributes. For example, you can include the following command in a route map to specify the next hop address:

Command Syntax	Command Mode	Purpose
set next-hop *ip-address*	CONFIG-ROUTE-MAP	Sets the next hop address.

Change WEIGHT attribute

Use the following command in CONFIGURATION ROUTER BGP mode to change the how the WEIGHT attribute is used.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **weight** *weight*	CONFIG-ROUTER-BGP	Assign a weight to the neighbor connection. • *weight* range: 0 to 65535 • Default is 0

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view BGP configuration.

You can also use route maps to change this and other BGP attributes. For example, you can include the following command in a route map to specify the next hop address:

Command Syntax	Command Mode	Purpose
set weight *weight*	CONFIG-ROUTE-MAP	Sets weight for the route. • *weight* range: 0 to 65535

Enable 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.

Use the following command in the CONFIGURATION ROUTER BGP mode to allow more than one path.

Command Syntax	Command Mode	Purpose
maximum-paths {**ebgp** | **ibgp**} *number*	CONFIG-ROUTER-BGP	Enable multiple parallel paths. • *number* range: 1 to 16 • Default is 1

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

Filter 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 (via a route map) to control which routes are accepted and advertised by the BGP neighbor or peer group. Prefix lists filter routes based on route and prefix length, while AS-Path ACLs filter routes based on the Autonomous System number. Route maps can filter and set conditions, change attributes, and assign update policies.

Note: FTOS supports up to 255 characters in a set community statement inside a route map.

Note: With FTOS, you can create inbound and outbound policies. Each of the commands used for filtering, has **in** and **out** parameters that must be applied. In FTOS, 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 **neighbor distribute-list** command)
- AS-PATH ACLs (using **neighbor filter-list** command)
- route maps (using **neighbor route-map** command)

Prior to filtering BGP routes, you must create the prefix list, AS-PATH ACL, or route map to be used.

Refer to Chapter 7, “Access Control Lists (ACL), Prefix Lists, and Route-maps,” on page 99 for configuration information on prefix lists, AS-PATH ACLs, and route maps.

Note: When you configure a new set of BGP policies, always reset the neighbor or peer group by entering the **clear ip bgp** command in EXEC Privilege mode.

Use these commands in the following sequence, starting in the CONFIGURATION mode to filter routes using prefix lists.

Step	Command Syntax	Command Mode	Purpose
1	**ip prefix-list** *prefix-name*	CONFIGURATION	Create a prefix list and assign it a name.
2	**seq** *sequence-number* {**deny** | **permit**} {*any* | *ip-prefix* [**ge | le**] }	CONFIG-PREFIX LIST	Create multiple prefix list filters with a deny or permit action. **ge**: Minimum prefix length to be matched **le**: maximum prefix length to me matched Refer to Chapter 7, “Access Control Lists (ACL), Prefix Lists, and Route-maps,” on page 99 for information on configuring prefix lists.
3	**exit**	CONFIG-PREFIX LIST	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **distribute-list** *prefix-list-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Filter routes based on the criteria in the configured prefix list. 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, below are some 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**).
- Once 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 the 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.

Use these commands in the following sequence, starting in the CONFIGURATION mode to filter routes using a route map.

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Create a route map and assign it a name.
2	{**match** | **set**}	CONFIG-ROUTE-MAP	Create multiple route map filters with a match or set action. Refer to Chapter 7, “Access Control Lists (ACL), Prefix Lists, and Route-maps,” on page 99 for information on configuring route maps.
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.
	neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Filter routes based on the criteria in the configured route map. 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.

Use the **show config** command in CONFIGURATION ROUTER BGP mode to view the BGP configuration. Use the **show route-map** command in EXEC Privilege mode to view a route map configuration.

Use these commands in the following sequence, beginning in the CONFIGURATION mode to filter routes based on AS-PATH information.

Step	Command Syntax	Command Mode	Purpose
1	**ip as-path access-list** *as-path-name*	CONFIGURATION	Create a AS-PATH ACL and assign it a name.
2	{**deny** | **permit**} *as-regular-expression*	AS-PATH ACL	Create a AS-PATH ACL filter with a deny or permit action.
3	**exit**	AS-PATH ACL	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.

Step	Command Syntax	Command Mode	Purpose
5	**neighbor** {*ip-address* | *peer-group-name*} **filter-list** *as-path-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Filter routes based on the criteria in the configured route map. 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.

Use the **show config** command in CONFIGURATION ROUTER BGP mode and **show ip as-path-access-list** command in EXEC Privilege mode to view which commands are configured.

Include this filter **permit .*** in your AS-PATH ACL to forward all routes not meeting the AS-PATH ACL criteria.

Configure BGP route reflectors

BGP route reflectors are intended for Autonomous Systems with a large mesh and 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 others are clients who receive their updates from the concentration router.

Use the following commands in the CONFIGURATION ROUTER BGP mode to configure a route reflector.

Command Syntax	Command Mode	Purpose
bgp cluster-id *cluster-id*	CONFIG-ROUTER-BGP	Assign an ID to a router reflector cluster. You can have multiple clusters in an AS.
neighbor {*ip-address* | *peer-group-name*} **route-reflector-client**	CONFIG-ROUTER-BGP	Configure the local router as a route reflector and the neighbor or peer group identified is the route reflector client.

To view a route reflector configuration, use the **show config** command in the CONFIGURATION ROUTER BGP mode or **show running-config bgp** in EXEC Privilege mode.

When you enable a route reflector, FTOS 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 should be fully meshed before you disable route reflection.

Aggregate routes

FTOS 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.

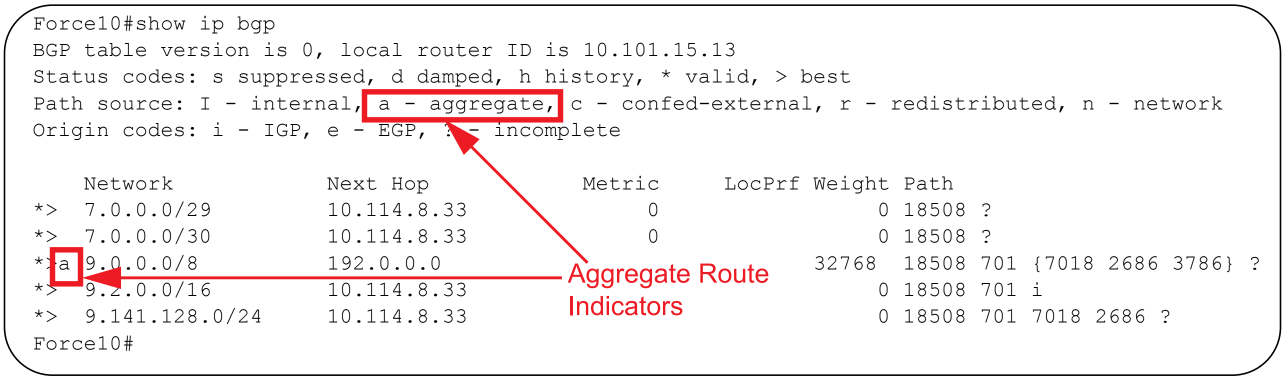
Use the following command in the CONFIGURATION ROUTER BGP mode to aggregate routes.

Command Syntax	Command Mode	Purpose
aggregate-address *ip-address mask* [**advertise-map** *map-name*] [**as-set**] [**attribute-map** *map-name*] [**summary-only**] [**suppress-map** *map-name*]	CONFIG-ROUTER-BGP	Assign the IP address and mask of the prefix to be aggregated. Optional parameters are: • **advertise-map** *map-name*: set filters for advertising an aggregate route • **as-set**: generate path attribute information and include it in the aggregate. • **attribute-map** *map-name: modify* attributes of the aggregate, except for the AS_PATH and NEXT_HOP attributes • **summary-only**: advertise only the aggregate address. Specific routes will not be advertised • **suppress-map** *map-name*: identify which more-specific routes in the aggregate are suppressed

AS_SET includes AS_PATH and community information from the routes included in the aggregated route.

In the **show ip bgp** command, aggregates contain an ‘a’ in the first column and routes suppressed by the aggregate contain an ‘s’ in the first column.

Figure 8-30. Command Example: show ip bgp



```
Force10#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} ?
*>  9.2.0.0/16         10.114.8.33                                0 18508 701 i
*>  9.141.128.0/24     10.114.8.33                                0 18508 701 7018 2686 ?
Force10#
```

Configure 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 a large number of 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.

Use the following commands in the CONFIGURATION ROUTER BGP mode to configure BGP confederations.

Command Syntax	Command Mode	Purpose
bgp confederation identifier *as-number*	CONFIG-ROUTER-BGP	Specifies the confederation ID. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte)
bgp confederation peers *as-number* [... *as-number*]	CONFIG-ROUTER-BGP	Specifies which confederation sub-AS are peers. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte)
	All Confederation routers must be either 4-Byte or 2-Byte. You cannot have a mix of router ASN support,	

Use the **show config** command in the CONFIGURATION ROUTER BGP mode to view the configuration.

Enable 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 FTOS, 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:

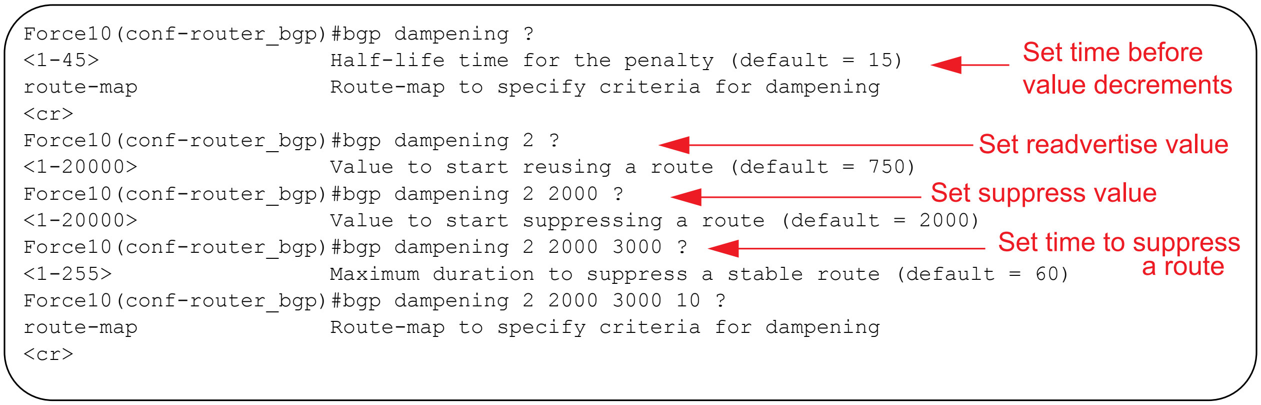
- Withdraw
- Readvertise
- Attribute change

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

The CLI example below shows configuring values to start reusing or restarting a route, as well as their default values.

Figure 8-31. Setting Reuse and Restart Route Values



```
Force10(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>
Force10(conf-router_bgp)#bgp dampening 2 ?
<1-20000>               Value to start reusing a route (default = 750)
Force10(conf-router_bgp)#bgp dampening 2 2000 ?
<1-20000>               Value to start suppressing a route (default = 2000)
Force10(conf-router_bgp)#bgp dampening 2 2000 3000 ?
<1-255>                 Maximum duration to suppress a stable route (default = 60)
Force10(conf-router_bgp)#bgp dampening 2 2000 3000 10 ?
route-map               Route-map to specify criteria for dampening
<cr>
```

Set time before value decrements

Set readvertise value

Set suppress value

Set time to suppress a route

Use the following command in the CONFIGURATION ROUTER BGP mode to configure route flap dampening parameters.

Command Syntax	Command Mode	Purpose
bgp dampening [*half-life* | *reuse* | *suppress* *max-suppress-time*] [**route-map** *map-name*]	CONFIG-ROUTER-BGP	Enable route dampening. Enter the following optional parameters to configure route dampening parameters: • *half-life* range: 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. (Default: 15 minutes) • *reuse* range: 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. (Default: 750) • *suppress* range: 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). (Default: 2000.) • *max-suppress-time* range: 1 to 255. The maximum number of minutes a route can be suppressed. The default is four times the half-life value. (Default: 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.

To view the BGP configuration, use **show config** in the CONFIGURATION ROUTER BGP mode or **show running-config bgp** in EXEC Privilege mode.

To set dampening parameters via a route map, use the following command in CONFIGURATION ROUTE-MAP mode:

Command Syntax	Command Mode	Purpose
set dampening *half-life reuse suppress max-suppress-time*	CONFIG-ROUTE-MAP	Enter the following optional parameters to configure route dampening parameters: • *half-life* range: 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. (Default: 15 minutes) • *reuse* range: 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). (Default: 750) • *suppress* range: 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). (Default: 2000.) • *max-suppress-time* range: 1 to 255. The maximum number of minutes a route can be suppressed. The default is four times the half-life value. (Default: 60 minutes.)

To view a count of dampened routes, history routes and penalized routes when route dampening is enabled, look at the seventh line of the **show ip bgp** summary command output (Figure 8-32).

Figure 8-32. Command example: show ip bgp summary

```
Force10>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
Force10>
```

Dampening Information

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

Use the following command in EXEC Privilege mode to clear information on route dampening and return suppressed routes to active state.

Command Syntax	Command Mode	Purpose
clear ip bgp dampening [*ip-address mask*]	EXEC Privilege	Clear all information or only information on a specific route.

Use the following command in EXEC and EXEC Privilege mode to view statistics on route flapping.

Command Syntax	Command Mode	Purpose
show ip bgp flap-statistics [*ip-address* [*mask*]] [**filter-list** *as-path-name*] [**regexp** *regular-expression*]	EXEC EXEC Privilege	View all flap statistics or for specific routes meeting the following criteria: • *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 FTOS 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.

Use the following command in CONFIGURATION ROUTER BGP mode to change the path selection from the default mode (deterministic) to non-deterministic.

Command Syntax	Command Mode	Purpose
bgp non-deterministic-med	CONFIG-ROUTER-BGP	Change the best path selection method to non-deterministic.

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.

Change BGP timers

Use either or both of the following commands in the CONFIGURATION ROUTER BGP mode to configure BGP timers.

Command Syntax	Command Mode	Purpose
neighbors {*ip-address* | *peer-group-name*} **timers** *keepalive holdtime*	CONFIG-ROUTER-BGP	Configure timer values for a BGP neighbor or peer group. • *keepalive* range: 1 to 65535. Time interval, in seconds, between keepalive messages sent to the neighbor routers. (Default: 60 seconds) • *holdtime* range: 3 to 65536. Time interval, in seconds, between the last keepalive message and declaring the router dead. (Default: 180 seconds)
timers bgp *keepalive holdtime*	CONFIG-ROUTER-BGP	Configure timer values for all neighbors. • *keepalive* range: 1 to 65535. Time interval, in seconds, between keepalive messages sent to the neighbor routers. (Default: 60 seconds) • *holdtime* range: 3 to 65536. Time interval, in seconds, between the last keepalive message and declaring the router dead. (Default: 180 seconds)

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view non-default values.

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 will be 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.

BGP neighbor soft-reconfiguration

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.

Use the **clear ip bgp** command in EXEC Privilege mode at the system prompt to reset a BGP connection using BGP soft reconfiguration.

Command Syntax	Command Mode	Purpose
clear ip bgp {* | *neighbor-address* **|** *AS Numbers* | *ipv4* | *peer-group-name***}** **[soft [in | out]]**	EXEC Privilege	Clear all information or only specific details. *: Clear all peers *neighbor-address*: Clear the neighbor with this IP address *AS Numbers*: Peers' AS numbers to be cleared *ipv4:* Clear information for IPv4 Address family *peer-group-name:* Clear all members of the specified peer group
neighbor {*ip-address* **|** *peer-group-name***}** **soft-reconfiguration inbound**	CONFIG-ROUTER-BGP	Enable soft-reconfiguration for the BGP neighbor specified. 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.	

When soft-reconfiguration is enabled for a neighbor and the **clear ip bgp soft in** command is executed, 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 route-refresh request is not negotiated with the peer. If the request is indeed negotiated (upon execution of **clear ip bgp soft in**), then 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 should be displayed:

```
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.

The following (Figure 8-33) 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.

Figure 8-33. Command example: router bgp

```
Force10>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 (in ROUTE-MAP mode) allows movement from one route-map entry to a specific route-map entry (the *sequence number*). If the sequence number is not specified, 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.

continue [*sequence-number***]**

Match 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 upon 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 Clause with a Continue Clause

If the route-map entry contains sets with the continue clause, then 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.

MBGP Configuration

MBGP for IPv6 unicast is supported on platforms ET C

MBGP for IPv4 Multicast is supported on platform C ET S

MBGP is *not* supported on the E-Series ExaScale EX platform.

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.

FTOS MBGP is implemented as per RFC 1858. The MBGP feature can be enabled per router and/or per peer/peer-group.

Default is IPv4 Unicast routes.

Command Syntax	Command Mode	Purpose
address family ipv4 multicast	CONFIG-ROUTER-BGP	Enables support for the IPv4 Multicast family on the BGP node
neighbor [*ip-address* | *peer-group-name*] **activate**	CONFIG-ROUTER-BGP-AF (Address Family)	Enable IPv4 Multicast support on a BGP neighbor/peer group

When a peer is configured to support IPv4 Multicast, FTOS takes the following actions:

- Send a capacity 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 will mark 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 FTOS BGP IPv4 Unicast commands are extended to support the IPv4 Multicast RIB using extra options to the command. See the *FTOS Command Line Interface Reference* for a detailed description of the MBGP commands.

BGP Regular Expression Optimization

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. FTOS 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. This feature is turned on by default. Use the command **bgp regex-eval-optz-disable** in CONFIGURATION ROUTER BGP mode to disable it if necessary.

Debugging BGP

Use any of the commands in EXEC Privilege mode to enable BGP debugging.

Command Syntax	Command Mode	Purpose
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] [**in** | **out**]	EXEC Privilege	View all information on BGP, including BGP events, keepalives, notifications, and updates.
debug ip bgp dampening [**in | out**]	EXEC Privilege	View information on BGP route being dampened.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **events** [**in** | **out**]	EXEC Privilege	View information on local BGP state changes and other BGP events.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **keepalive** [**in** | **out**]	EXEC Privilege	View information about BGP KEEPALIVE messages.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **notifications** [**in** | **out**]	EXEC Privilege	View information about BGP notifications received from or sent to neighbors.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **updates** [**in** | **out**] [**prefix-list name**]	EXEC Privilege	View information about BGP updates and filter by prefix name
debug ip bgp {*ip-address* | *peer-group-name*} **soft-reconfiguration**	EXEC Privilege	Enable soft-reconfiguration debug. Enable soft-reconfiguration debug. To enhance debugging of soft reconfig, use the following command only when route-refresh is not negotiated to avoid the peer from resending messages: **bgp soft-reconfig-backup** In-BGP is shown via the **show ip protocols** command.

FTOS displays debug messages on the console. To view which debugging commands are enabled, use the **show debugging** command in EXEC Privilege mode.

Use the keyword no followed by the debug command To disable a specific debug command. For example, to disable debugging of BGP updates, enter **no debug ip bgp updates** command.

Use **no debug ip bgp** to disable all BGP debugging.

Use **undebug all** to disable all debugging.

Storing Last and Bad PDUs

FTOS stores the last notification sent/received, and the last bad PDU received on per peer basis. The last bad PDU is the one that causes a notification to be issued. These PDUs are shown in the output of the command **show ip bgp neighbor**, as shown in Figure 8-34.

Figure 8-34. Viewing the Last Bad PDU from BGP Peers

```
Force10(conf-router_bgp)#do show ip bgp neighbors 1.1.1.2

BGP neighbor is 1.1.1.2, remote AS 2, external link
  BGP version 4, remote router ID 2.4.0.1
  BGP state ESTABLISHED, in this state for 00:00:01
  Last read 00:00:00, last write 00:00:01
  Hold time is 90, keepalive interval is 30 seconds
  Received 1404 messages, 0 in queue
    3 opens, 1 notifications, 1394 updates
    6 keepalives, 0 route refresh requests
  Sent 48 messages, 0 in queue
    3 opens, 2 notifications, 0 updates
    43 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  For address family: IPv4 Unicast
  BGP table version 1395, neighbor version 1394
  Prefixes accepted 1 (consume 4 bytes), 0 withdrawn by peer
  Prefixes advertised 0, rejected 0, 0 withdrawn from peer

  Connections established 3; dropped 2
  Last reset 00:00:12, due to Missing well known attribute

  Notification History
   'UPDATE error/Missing well-known attr' Sent : 1  Recv: 0
   'Connection Reset' Sent : 1  Recv: 0

   Last notification (len 21) sent 00:26:02 ago
    ffffffff ffffffff ffffffff ffffffff 00160303 03010000
   Last notification (len 21) received 00:26:20 ago
    ffffffff ffffffff ffffffff ffffffff 00150306 00000000
   Last PDU (len 41) received 00:26:02 ago that caused notification to be issued
    ffffffff ffffffff ffffffff ffffffff 00290200 00000e01 02040201 00024003 04141414
0218c0a8
    01000000
Local host: 1.1.1.1, Local port: 179
Foreign host: 1.1.1.2, Foreign port: 41758
```

Last PDUs

Capturing PDUs

Capture incoming and outgoing PDUs on a per-peer basis using the command **capture bgp-pdu neighbor direction.** Disable capturing using the no form of this command.

The buffer size supports a maximum value between 40 MB (the default) and 100 MB. The capture buffers are cyclic and reaching the limit prompts the system to overwrite the oldest PDUs when new ones are received for a given neighbor or direction. Setting the buffer size to a value lower than the current max, might cause captured PDUs to be freed to set the new limit.

Note: Memory on RP1 is not pre-allocated, and is allocated only when a PDU needs to be captured.

Use the command **capture bgp-pdu max-buffer-size** (Figure 8-35) to change the maximum buffer size. View the captured PDUs using the command **show capture bgp-pdu neighbor**.

Figure 8-35. Viewing Captured PDUs

```
Force10#show capture bgp-pdu neighbor 20.20.20.2

Incoming packet capture enabled for BGP neighbor 20.20.20.2
Available buffer size 40958758, 26 packet(s) captured using 680 bytes
  PDU[1] : len 101, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00650100 00000013 00000000 00000000 419ef06c
00000000
    00000000 00000000 00000000 00000000 0181a1e4 0181a25c 41af92c0 00000000 00000000
00000000
    00000000 00000001 0181a1e4 0181a25c 41af9400 00000000
  PDU[2] : len 19, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[3] : len 19, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[4] : len 19, captured 00:34:22 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
[. . .]

Outgoing packet capture enabled for BGP neighbor 20.20.20.2
Available buffer size 40958758, 27 packet(s) captured using 562 bytes
  PDU[1] : len 41, captured 00:34:52 ago
    ffffffff ffffffff ffffffff ffffffff 00290104 000100b4 14141401 0c020a01 04000100
01020080
    00000000
  PDU[2] : len 19, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[3] : len 19, captured 00:34:50 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[4] : len 19, captured 00:34:20 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
[. . .]
```

The buffers storing the PDU free memory when:

- BGP is disabled
- A neighbor is unconfigured
- **clear ip bgp** is issued
- New PDU are captured and there is no more space to store them
- The max buffer size is reduced. (This may cause PDUs to be cleared depending upon the buffer space consumed and the new limit.)

With full internet feed (205K) captured, approximately 11.8MB is required to store all of the PDUs, as shown in Figure 8-36.

Figure 8-36. Required Memory for Captured PDUs

```
Force10(conf-router_bgp)#do show capture bgp-pdu neighbor 172.30.1.250

Incoming packet capture enabled for BGP neighbor 172.30.1.250
Available buffer size 29165743, 192991 packet(s) captured using 11794257 bytes
 [. . .]

Force10(conf-router_bgp)#do sho ip bg s
BGP router identifier 172.30.1.56, local AS number 65056
BGP table version is 313511, main routing table version 313511
207896 network entrie(s) and 207896 paths using 42364576 bytes of memory
59913 BGP path attribute entrie(s) using 2875872 bytes of memory
59910 BGP AS-PATH entrie(s) using 2679698 bytes of memory
3 BGP community entrie(s) using 81 bytes of memory

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

1.1.1.2         2           17    18966          0    0     0 00:08:19 Active
172.30.1.250    18508   243295       25     313511    0     0 00:12:46    207896
```

PDU Counters

FTOS 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 command **show ip bgp neighbor**.

Sample Configurations

The following configurations are examples for enabling BGP and setting up some peer groups. 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.

Figure 8-37 is a graphic illustration of the configurations shown on the following pages. 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 8-37. Sample Configuration Illustration

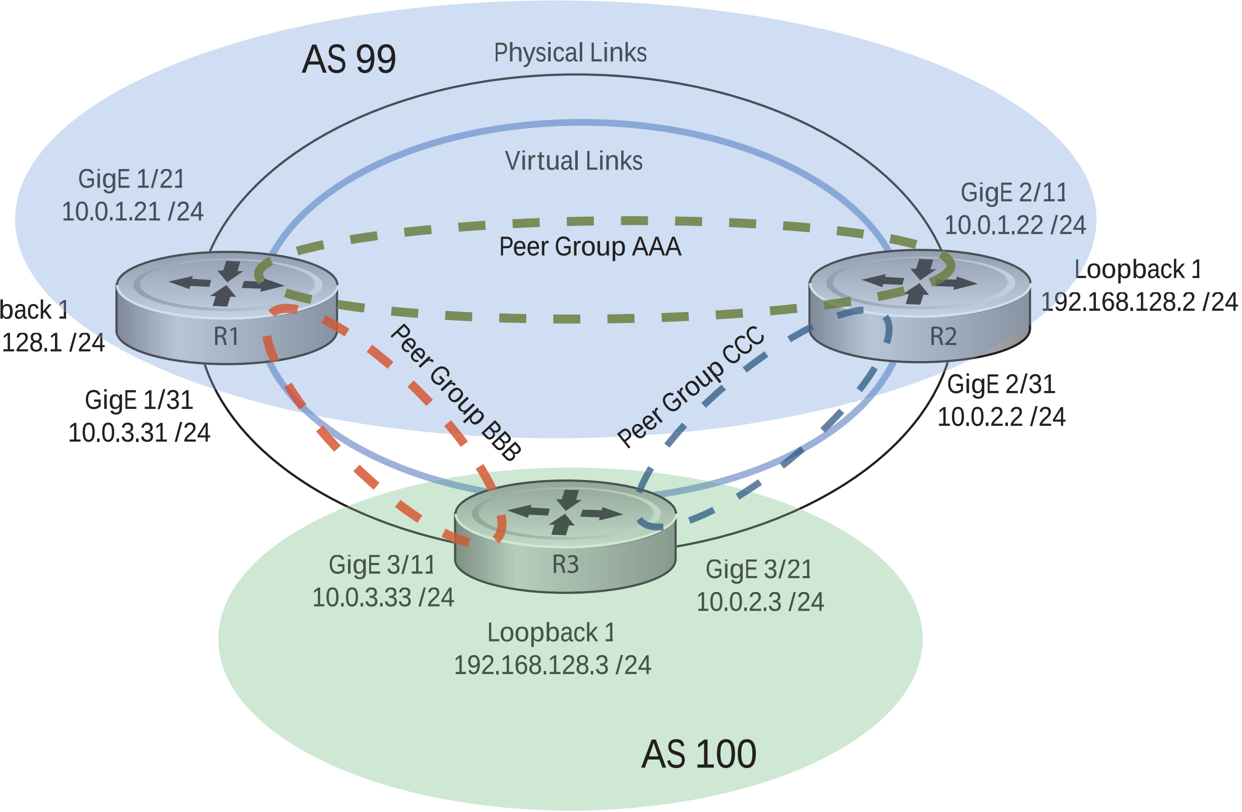


Figure 8-38.   Enable BGP - Router 1

```
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#
```

Figure 8-39. Enable BGP - Router 2

```
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#
```

Figure 8-40. Enable BGP - Router 3

```
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
```

Figure 8-41. Enable Peer Group - Router 1

```
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
```

Figure 8-42. Enable Peer Groups - Router 1 continued

```
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#
```

Figure 8-43. Enable Peer Groups - Router 2

```
R2#conf
R2(conf)#router bgp 99
R2(conf-router_bgp)# neighbor CCC peer-group
R2(conf-router_bgp)# neighbor CC no shutdown
R2(conf-router_bgp)# neighbor BBB peer-group
R2(conf-router_bgp)# neighbor BBB no shutdown
R2(conf-router_bgp)# neighbor 192.168.128.1 peer AAA
R2(conf-router_bgp)# neighbor 192.168.128.1 no shut
R2(conf-router_bgp)# neighbor 192.168.128.3 peer BBB
R2(conf-router_bgp)# neighbor 192.168.128.3 no shut
R2(conf-router_bgp)#show conf
!
router bgp 99
 network 192.168.128.0/24
 neighbor AAA peer-group
 neighbor AAA no shutdown
 neighbor BBB peer-group
 neighbor BBB no shutdown
 neighbor 192.168.128.1 remote-as 99
 neighbor 192.168.128.1 peer-group CCC
 neighbor 192.168.128.1 update-source Loopback 0
 neighbor 192.168.128.1 no shutdown
 neighbor 192.168.128.3 remote-as 100
 neighbor 192.168.128.3 peer-group BBB
 neighbor 192.168.128.3 update-source Loopback 0
 neighbor 192.168.128.3 no shutdown
R2(conf-router_bgp)#end

R2#
R2#show ip bgp summary
BGP router identifier 192.168.128.2, local AS number 99
BGP table version is 2, main routing table version 2
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.1   99         140      136          2    0   (0) 00:11:24         1
192.168.128.3   100        138      140          2    0   (0) 00:18:31         1

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
```

Figure 8-44. Enable Peer Group - Router 3

```
R3#conf
R3(conf)#router bgp 100
R3(conf-router_bgp)# neighbor AAA peer-group
R3(conf-router_bgp)# neighbor AAA no shutdown
R3(conf-router_bgp)# neighbor CCC peer-group
R3(conf-router_bgp)# neighbor CCC no shutdown
R3(conf-router_bgp)# neighbor 192.168.128.2 peer-group BBB
R3(conf-router_bgp)# neighbor 192.168.128.2 no shutdown
R3(conf-router_bgp)# neighbor 192.168.128.1 peer-group BBB
R3(conf-router_bgp)# neighbor 192.168.128.1 no shutdown
R3(conf-router_bgp)#

R3(conf-router_bgp)#end

R3#show ip bgp summary
BGP router identifier 192.168.128.3, local AS number 100
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory

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

192.168.128.1   99          93       99          1    0   (0) 00:00:15         1
192.168.128.2   99         122      120          1    0   (0) 00:00:11         1
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
```

Figure 8-45. Enable Peer Groups - Router 3 continued

```
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
```

9

Bare Metal Provisioning 2.0

Bare Metal Provisioning 2.0 is included as part of the FTOS image. It is supported on the following platforms: S55 S60 S4810 □

Bare Metal Provisioning (BMP) improves accessibility to the switch by automatically loading pre-defined configurations and boot images that are stored in file servers. BMP can be used on a single switch or on multiple switches.

For more information on using BMP and the different types of modes, refer to the *Open Automation Guide*.

BMP eases configuration by automating the following steps:

- Boot images and running configurations are specified in a DHCP server.
- Switch boots up in Layer 3 mode with interfaces already in no shutdown mode and only enabling some basic protocols to protect the switch and the network.
- The first port that receives the DHCP server response retains the IP address provided by the DHCP server during the BMP process. All other management and user ports are shut down.
- Files are automatically downloaded from a file server.
- After the BMP process is complete, the IP address is released and the configuration is applied by the switch.

BMP is enabled on a brand new, factory-loaded switch. You can enable and disable BMP using the following steps:

1. Configure a reload mode using the **reload-type** command.
2. Reload the switch in the configured mode using the **reload** command.

Prerequisites

Before you use BMP 2.0 to auto-configure a supported Dell Force10 switch, you must first configure a Dynamic Host Configuration Protocol (DHCP) server and a file server in the network. Optionally, you can also configure a Domain Name Server (DNS). For more information, refer to DHCP Server, Domain Name Server, and File Server.

Note: If the switch is connected to upstream aggregation switches that have VLT enabled, and the DHCP and file servers are reachable through the VLT LAG interface, you must configure the VLT members with the **lacp ungroup member-independent vlt** command. This allows the bottom switch to establish communication with the VLT switches.

Restrictions

BMP 2.0 is supported on the user ports and management ports of a switch.

BMP 2.0 is not supported in a stacking environment.

Overview

On a new factory-loaded switch, the switch boots up in **Jumpstart** mode. You can reconfigure a switch to reload between **Normal** and **Jumpstart** mode.

- **Jumpstart (BMP) mode**: The switch automatically configures all ports (management and user ports) as Layer 3 physical ports and acts as a DHCP client on the ports for a user-configured time (DHCP timeout). This is the default startup mode. It is set with the **reload-type jump-start** command.
- **Normal mode**: The switch loads the FTOS image and startup configuration file stored in the local flash. New configurations require that the Management IP and Management Interface be configured manually. This mode is set with the **reload-type normal-reload** command.

To reconfigure a switch to reload between **Normal** and **Jumpstart** mode, use the **reload-type** command.

Command Syntax	Command Mode	Purpose
reload-type {**normal-reload** | **jump-start** [**config-download** {*enable* | *disable*}] [**dhcp-timeout** *minutes*]}	EXEC Privilege	Reload a switch running BMP version 2.0 in either **Normal** or **Jumpstart** mode. If you reload in **Jumpstart** mode, you can configure: • **config-download**: Whether the switch boots up using the configuration file downloaded from the DHCP/file servers (*enable*) OR if the downloaded file will be discarded and the startup configuration file stored in the local flash will be used (*disable*). • **dhcp-timeout**: The amount of time the switch waits for a DHCP server response before reverting to **Normal** mode and loading the startup configuration from the flash. The default time is infinity, which makes the switch continue to wait forever unless the **stop jump-start** command is given. Range: 1 to 50 minutes. Default: The switch tries to contact a DHCP server an infinite number of times.
stop jump-start	EXEC Privilege	This command stops the jumpstart reload process while it is in progress and changes the reload type to **Normal** mode. If the command is initiated while the switch is downloading an image or configuration file, the command takes effect when the DHCP release is sent.

The reload settings that you configure with the **reload-type** command are stored in non-volatile memory and retained for future reboots. Enter the **reload** command to reload the switch in the current configured mode: Normal or Jumpstart mode.

To display the currently configured reload mode for a switch running BMP version 2.0, enter the **show reload-type** or **show bootvar** command.

```
FTOS#show reload type
Reload-Type         :   jump-start [Next boot :jump-start]
config-download     :   enable
dhcp-timeout        :   10

FTOS#show bootvar
. . content truncated..
Reload Mode =  jump-start
File URL =  tftp:/30.0.0.1/FTOS-SE-8-3-8-17.bin
```

Note: If a switch enters a loop while reloading in Jumpstart mode because it continuously tries to contact a DHCP server and a DHCP server is not found, connect to the switch using the console terminal and enter the **stop jump-start** command to interrupt the repeated discovery attempts. The startup configuration file stored in the local flash on the switch is loaded and the auto-configuration mode is automatically changed to **Normal** reload, i.e., BMP is disabled.

Jumpstart mode

Jumpstart (BMP) mode is the default boot mode configured for a new switch arriving from Dell Force10. This mode obtains the FTOS image and configuration file from a network source (DHCP server and file server).

DHCP Server

DHCP Configuration

You must first configure an external DHCP server before you can use the Jumpstart mode on a switch. Configure the DHCP server with the set of parameters described below for each client switch. Refer to the *FTOS Configuration Guide: Dynamic Host Configuration Protocol* chapter for detailed information. The DHCP server is configured to assign an IP address to the switch and specify the files to download.

One or more of the following parameters must be configured on the DHCP server.

- Boot File Name: The FTOS image to be loaded on the switch. The boot file name is expected to use Option 67 or the boot filename in the boot payload of the DHCP offer. If both are specified, Option 67 will be used.
- Configuration File Name: The configurations to be applied to the switch. The configuration file name is expected to use Option 209.
- File Server Address: The server where the Image and Configurations file are placed. The address is assumed to be a TFTP address unless it is given as a URL. The switch supports TFTP, HTTP, and FTP protocols, as well as files stored in Flash and external USB memory. If TFTP is used, you can add Option 66 or Option 150.
- Domain Name Server: (Optional.) The DNS server to be contacted to resolve the hostname through Option 6.
- IP Address: Dynamic IP address for the switch. This is used only for file transfers.

The DHCP option codes used are:

- 6 Domain Name Server IP
- 66 TFTP Server name
- 67 Boot filename
- 150 TFTP server IP address
- 209 Configuration File

Note: The boot file name and configuration file name must be in the correct format. If it is not, the switch will be unable to download the file from the DHCP server, and will behave as if the server could not be reached. The discovery process will continue, despite configured time-out, until the **stop jump-start** command is given.

URL Examples	Description
`##### FTOS image`	
`option bootfile-name "ftp://user:passwd@myserver/FTOS-SE-8.3.10.1.bin";`	FTP URL with hostname (requires DNS)
`option bootfile-name "http://10.20.4.1/FTOS-SE-8.3.10.1.bin";`	HTTP URL with IP address
`option bootfile-name "tftp://10.20.4.1/FTOS-SE-8.3.10.1.bin";`	TFTP URL with IP address
`option bootfile-name "flash://FTOS-SE-8.3.10.1.bin";`	Flash path relative to /f10/flash directory
`##### Configuration file could be given in the following way`	
`option config-file "ftp://user:passwd@10.20.4.1//home/user/S4810-1.conf";`	FTP URL with IP address
`option config-file "http://myserver/S4810-1.conf";`	HTTP URL with hostname (requires DNS)
`option config-file "tftp://10.10.4.1/S4810-1.conf";`	TFTP URL with IP address
`option config-file "flash://S4810-1.conf";`	Flash path relative to /f10/flash directory
`option config-file "usbflash://S55-1.conf";`	External USB memory

MAC-Based IP assignment

One way to use the BMP mode most efficiently is to configure the DHCP server to assign a fixed IP address, FTOS image, and configuration file based on the switch's MAC address. When this is done, the same IP address is assigned to the switch even on repetitive reloads and the same configuration file will be retrieved when using the DNS server or the **network-config** file to determine the hostname.

The assigned IP address is only used to retrieve the files from the file server. It is discarded after the files are retrieved.

Following is an example of a configuration of the DHCP server included on the most popular Linux distributions. The **dhcpd.conf** file shows assignment of a fixed IP address and configuration file based on the MAC address of the switch.

Parameter Example	Description

`option boot-filename code 67 = text;`
`option tftp-server-address code 150 = ip-address;`
`option config-file code 209 = text;`
 	
`subnet 10.20.30.0 netmask 255.255.255.0 {`
`    option domain-name-servers 20.30.40.1, 20.30.40.2;`	

` host S4810-1 {`	`BMP 2.0 Syntax`

`        hardware ethernet 00:01:e8:8c:4d:0e;`
`        fixed-address 10.20.30.41;`
	MAC to IP mapping

`       option boot-filename "tftp://10.20.4.1/FTOS-SE-8.3.10.1.bin";`	FTOS image
`        option config-file "http://10.20.4.1/S4810-1.conf";`	Config file
`}`	
`    host S4810-2 {`	**`BMP1.0 syntax`**

`        hardware ethernet 00:01:e8:8c:4c:04;`
`        fixed-address 10.20.30.42;`
	MAC to IP mapping

`        option tftp-server-address 10.20.4.1;`
`        filename "FTOS-SE-8.3.10.1.bin";`
	FTOS image

`        option config-file "S4810-2.conf";`	Config file
`    }`	

DHCP Retry Mechanism

BMP will request a different DHCP offer in the following scenarios:

- If the command **reload-type config-download** is enabled, the DHCP offer specifies both the boot image and the configuration file, and either download is successful, BMP will not request another DHCP offer.
- If the offer contains only a boot image that cannot be downloaded, BMP will request another DHCP offer.
- If the command **reload-type config-download** is enabled and the configuration file in the offer cannot be downloaded, BMP will request another DHCP offer.

DHCP Server IP Blacklist

If the process does not complete successfully, the DHCP server IP will be blacklisted and the DHCP process will be re-initiated. A DHCP server is maintained in the blacklist for ten minutes. If a DHCP offer is received from the blacklisted DHCP server, the offer will be rejected.

File Server

Set up a file server and ensure connectivity.

The server that holds the boot and configuration files must be configured as the network source for the switch. The switch recognizes HTTP, TFTP, FTP, external USB memory and Flash URLs.

For example:

- `tftp://serverip/filename`
- `tftp://hostname/filename`
- `ftp://user:passwd@serverip//mypath/filename`
- `ftp://user:passwd@hostname//mypath/filename`
- `http://serverip/filename`
- `http://hostname/filename`
- `flash://filename`
- `usbflash://filename`
- `filename (`Assumes TFTP`)`

When loading the FTOS image, if the FTOS image on the server is different from the image on the local flash, the switch downloads the image from the server onto the local flash and reloads using that image.

Next, the switch tries to load the configuration file. If the configuration file is not specified or if the **config-download** parameter is disabled, the switch loads the startup-config from the local flash.

Domain Name Server

Set up a Domain Name Server (DNS) to determine the host name applied in the switch startup configuration when no configuration file is retrieved from the DHCP server. The DNS server is contacted only when no configuration file is contained in a DHCP server response and the host name is not resolved from the network-config file on the switch. Refer to the *FTOS Configuration Guide IPv4 Addressing* chapter, *Resolution of Host Names* for information.

Switch boot and set-up behavior in Jumpstart Mode

When the switch boots up in jumpstart mode all ports, including management ports, are placed in **L3** mode in a **no shut** state. The switch acts as a DHCP client on these ports for a period of time (dhcp-timeout). This allows the switch time to send out a DHCP DISCOVER on all the **interface up** ports to the DHCP Server in order to obtain its IP address, boot image filename and configuration file from the DHCP server.

1. The switch begins boot up process in jumpstart mode (default mode)

2. The switch sends DHCP Discover on all the interface up ports.

```
00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Ma 0/0.
00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/0.
00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/5.
00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/6.
00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/8.
00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/35.
00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Fo 0/56.
00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Fo 0/60.
00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Ma 0/0.
```

3. The IP address, boot image filename and the configuration filename are reserved for the switch and provided in the DHCP reply (one-file read method). The switch receives its IP address, subnet mask, DHCP server IP, TFTP server address, DNS server IP, bootfile name and the configuration filename from the DHCP server.

   If a DHCP offer has no image path or configuration file path it is considered to be an invalid BMP DHCP offer, the offer is ignored. The first DHCP offer with IP address, FTOS image and configuration file, *or* the IP address and FTOS image, *or* the IP address and configuration file is chosen.

4. The DHCP OFFER is selected.

   a All other ports are set to shutdown mode.

```
 00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP acquired IP 30.0.0.20 mask 255.255.0.0 server IP 30.1.1.1.
 00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP tftp IP 30.0.0.1 dns IP 30.0.0.1 router IP 30.0.0.14.
 00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP image file FTOS-SE-8.3.10.1.bin.
 00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP config file pt-s4810-12.
```

5. The switch sends a unicast message to the file server to retrieve the named FTOS file and/or the configuration file from the base directory of the server.

   a If an option **bootfile-name** is used, the file name can be 256 bytes. If a **filename** field is specified in the DHCP Offer, the filename can be 128 bytes. The name can be a fully qualified URL or it can be a file name only.

   b When an FTOS image is found, the switch compares that image to the version the chassis currently has loaded.

   - If there is a mismatch, the switch applies the downloaded version and reloads.

```
*********VALID IMAGE***********

 DOWNLOADED RELEASE HEADER :
Release Image Major Version  : 8
Release Image Minor Version  : 3
Release Image Main Version   : 8
Release Image Patch Version  : 33

 FLASH RELEASE HEADER B :
Release Image Major Version  : 8
Release Image Minor Version  : 3
Release Image Main  Version  : 10
Release Image Patch Version  : 1
00:04:05: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DOWNLOAD: The FTOS image download is
successful.


Erasing Sseries Primary Image, please wait
```

```
..........................................................................................
..........................................................................................
..........................................................................................
..........................................................................................
..........................................................................................
...............................00:09:50: %STKUNsyncing disks... IT0-M:CP %CHMGR-1
5-RELOAD: User done
request to reload the chassis
rebooting
```

- If there is no version mismatch the switch downloads the configuration file.

```
00:27:12: %STKUNIT0-M:CP %JUMPSTART-2-JUMPSTART_DOWNLOAD_START: The config file download has started.
00:27:12: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE: DHCP RELEASE sent on Gi 0/1.
00:27:12: %STKUNIT0-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 0/1
00:27:12: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 0/1
00:27:37: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DOWNLOAD: The config file download is successful.
00:27:37: %STKUNIT0-M:CP %JUMPSTART-5-CFG_APPLY: The downloaded config from dhcp server is being applied
00:27:37: %STKUNIT0-M:CP %SYS-5-CONFIG_LOAD: Loading configuration file,
```

c If the configuration file is downloaded from the server, any saved startup-configuration on the flash is ignored. If no configuration file is downloaded from the server or the **config-download** parameter is disable, the startup-configuration file on the flash is loaded as in normal reload.

6. When the FTOS image and the configuration file have been downloaded, the IP address is released.

```
00:04:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE: DHCP RELEASE sent on Fo 0/56.
00:04:06: %STKUNIT0-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Fo 0/56
```

7. The switch applies the configuration. The switch is now up and running. It can be managed as usual.

10

Content Addressable Memory

Content Addressable Memory is supported on platforms C E T S

Content Addressable Memory

Content Addressable Memory (CAM) is a type of memory that stores information in the form of a lookup table. On Dell Force10 systems, the CAM stores Layer 2 and Layer 3 forwarding information, access-lists (ACL), flows, and routing policies. On Dell Force10 systems, there are one or two CAM (Dual-CAM) modules per port-pipe depending on the type of line card.

- The ExaScale EH and EJ series line cards are single-CAM line cards that support 10M and 40M CAM for storing the lookup information.
- The TeraScale EG-series line cards are dual-CAM and use two 18 Megabit CAM modules with a dedicated 512 IPv4 Forwarding Information Base (FIB), and flexible CAM allocations for Layer2, FIB, and ACLs.
- Either ExaScale 10G or 40G CAM line cards can be used in a system.

CAM Profiles

Dell Force10 systems partition each CAM module so that it can store the different types of information. The size of each partition is specified in the CAM profile. A CAM profile is stored on every card, including each RPM. The same profile must be on every line card and RPM in the chassis.

There is a default CAM profile and several other CAM profiles available so that you can partition the CAM according to your performance requirements. For example, the default profile has 1K Layer 2 ingress ACL entries. If you need more memory for Layer 2 ingress ACLs, select the profile *l2-ipv4-inacl*.

Table 10-1 describes the available profiles. The default profile is an all-purpose profile that allocates CAM space according to the way Dell Force10 systems are most commonly used. In general, non-default profiles allocate more space to particular regions to accommodate specific applications. The size of CAM partitions is measured in entries. The total CAM space is finite, therefor adding entries to one region necessarily decreases the number available to other regions.

Note: Not all CAM profiles and microcodes are available for all systems. Refer to the Command Line Interface Reference Guide for details regarding available profiles for each system.

Table 10-1. CAM Profile Descriptions

CAM Profile	Description
Default	An all-purpose profile that allocates CAM space according to the way Dell Force10 systems are most commonly used. Available Microcodes: default, lag-hash-align, lag-hash-mpls
eg-default	For EG-series line cards only. EG series line cards have two CAM modules per Port-pipe. Available Microcodes: default, ipv6-extacl
ipv4-320k	Provides 320K entries for the IPv4 Forwarding Information Base (FIB) and reduces the IPv4 Flow partition to 12K. Available Microcodes: default, lag-hash-mpls
ipv4-egacl-16k	Provides 16K entries for egress ACLs Available Microcodes: acl-group
ipv6-extacl	Provides IPv6 functionality. Available Microcodes: ipv6-extacl
l2-ipv4-inacl	Provides 32K entries for Layer 2 ingress ACLs and 28K entries for Layer 3 IPv4 ingress ACLs. Available Microcodes: default
unified-default	Maintains the CAM allocations for the and IPv4 FIB while allocating more CAM space for the Ingress and Egress Layer 2 ACL, and IPv4 ACL regions. Available Microcodes: ipv6-extacl
ipv4-VRF	Provides VRF functionality for IPv4. Available Microcodes:ipv4-vrf
ipv4-v6-VRF	Provides VRF functionality for both IPv4 and I.Pv6 Available Microcodes: ipv4-v6-vrf

Table 10-1. CAM Profile Descriptions (continued)

CAM Profile	Description
ipv4-64k-ipv6	Provides IPv6 functionality; an alternate to ipv6-extacl that redistributes CAM space from the IPv4FIB to IPv4Flow and IPv6FIB. Available Microcodes: ipv6-extacl

The size of CAM partitions is measured in entries. Table 10-1 shows the number of entries available in each partition for all CAM profiles. The total CAM space is finite, therefor adding entries to one region necessarily decreases the number available to other regions.

Table 10-2. CAM entries per partition

Profile Partition	L2FIB	L2ACL	IPv4FIB	IPv4ACL	IPv4Flow	EgL2ACL	EgIPv4ACL	Reserved	IPv6FIB	IPv6ACL	IPv6Flow	EgIPv6ACL
Default	32K	2K	256K	12K	24K	1K	1K	8K	0	0	0	0
eg-default	32K	2K	512K	12K	24K	1K	1K	8K	32K	3K	4K	1K
ipv4-320k	32K	2K	320K	12K	12K	1K	1K	4K	0	0	0	0
pv4-egacl-16k	32K	2K	192K	8K	24K	0	16K	8K	0	0	0	0
ipv6-extacl	32K	2K	192K	12K	8K	1K	1K	2K	6K	3K	4K	2K
l2-ipv4-inacl	32K	33K	64K	27K	8K	2K	2K	2K	0	0	0	0
unified-default	32K	3K	192K	9K	8K	2K	2K	2K	6K	2K	4K	2K
IPv4-VRF	32K	3K	160K	2K	12K	1K	12K	2K	0	0	0	0
IPv4-v6-VRF	32K	3K	64K	1K	12K	1K	11K	2K	18K	4K	3K	1K
ipv4-64k-ipv6	32K	2K	64K	12K	24K	1K	1K	8K	16K	3K	4K	1K

Microcode

Microcode is a compiled set of instructions for a CPU. On Dell Force10 systems, the microcode controls how packets are handled.

There is a default microcode, and several other microcodes are available, so that you can adjust packet handling according to your application. Specifying a microcode is mandatory when selecting a CAM profile (though you are not required to change it).

Note: Not all CAM profiles and microcodes are available for all systems. Refer to the Command Line Interface Reference Guide for details regarding available profiles for each system.

Table 10-3. Microcode Descriptions

Microcode	Description
default	Distributes CAM space for a typical deployment
lag-hash-align	For applications that require the same hashing for bi-directional traffic (for example, VoIP call or P2P file sharing). For port-channels, this microcode maps both directions of a bi-directional flow to the same output link.
lag-hash-mpls	For hashing based on MPLS labels (up to five labels deep). With the default microcode, MPLS packets are distributed over a port-channel based on the MAC source and destination address. With the lag-hash-mpls microcode, MPLS packets are distributed across the port-channel based on IP source and destination address and IP protocol. This is applicable for MPLS packets with up to five labels. When the IP header is not available after the 5th label, hashing for default load-balance is based on MPLS labels. For packets with more than 5 labels, hashing is always based on the MAC source and destination address.
ipv6-extacl	Use this microcode when IPv6 is enabled.
acl-group	For applications that need 16k egress IPv4 ACLs (for example, the VLAN ACL Group feature, which permits group VLANs IP egress ACLs.
ipv4-vrf	Apply to IPv4 VRF CAM profile.
ipv4-v6-vrf	Enable IPv4 and IPv6 CAM profiles for VRF.

CAM Profiling for ACLs

CAM Profiling for ACLs is supported on platform E T only.

The default CAM profile has 1K Layer 2 ingress ACL entries. If you need more memory for Layer 2 ingress ACLs, select the profile *l2-ipv4-inacl*.

When budgeting your CAM allocations for ACLs and QoS configurations, remember that ACL and QoS rules might consume more than one CAM entry depending on complexity. For example, TCP and UDP rules with port range options might require more than one CAM entry.

The Layer 2 ACL CAM partition has sub-partitions for several types of information. Table 10-4 lists the sub-partition and the percentage of the Layer 2 ACL CAM partition that FTOS allocates to each by default.

Table 10-4. Layer 2 ACL CAM Sub-partition Sizes

Partition	% Allocated
Sysflow	6
L2ACL	14
*PVST	50
QoS	12
L2PT	13
FRRP	5

You can re-configure the amount of space, in percentage, allocated to each sub-partition. As with the IPv4Flow partition, you can configure the Layer 2 ACL partition from EXEC Privilege mode or CONFIGURATION mode.

The amount of space that you can distribute to the sub-partitions is equal to the amount of CAM space that the selected CAM profile allocates to the Layer 2 ACL partition. FTOS requires that you specify the amount of CAM space for all sub-partitions and that the sum of all sub-partitions is 100%. FTOS displays the following message if the total allocated space is not correct:

```
% Error: Sum of all regions does not total to 100%.
```

Boot Behavior

The profile and microcode loaded on the primary RPM determines the profile and microcode that is required on all other chassis components and is called the “chassis profile.” A profile mismatch condition exists if either the CAM profile or the microcode does not match. The following points describe line card boot behavior when the line card profile does not match the chassis profile.

- A microcode mismatch constitutes a profile mismatch.
- When the line card profile and chassis profile are of the same type (single-CAM or dual-CAM), but their CAM profiles do not match, the line card must load a new profile and therefore takes longer to come online.
- If you insert a single-CAM line card into a chassis with a dual-CAM profile, the system displays Message 1. The line card boots with the default (single-CAM) profile and remains in a problem state (Figure 10-1). The line card cannot forward traffic in a problem state.

- If you insert a dual-CAM line card into a chassis with a single-CAM profile, the line card boots with a matching profile, but operates with a lower capability.

Message 1 EF Line Card with EG Chassis Profile Error

```
# Before reload:
01:09:56: %RPM0-P:CP %CHMGR-4-EG_PROFILE_WARN: If EG CAM profile is selected, non-EG cards
will be in problem state after reload
# After reload:
00:04:46: %RPM0-P:CP %CHMGR-3-PROFILE_MISMATCH: Mismatch: line card 1 has mismatch CAM
profile or microcode
```

Message 2 EH Line Card with EG Chassis Profile Error

```
# Before reload:
01:09:56: %RPM0-P:CP %CHMGR-4-EH_PROFILE_WARN: If EH CAM profile is selected, non-EJ cards
will be in problem state after reload
# After reload:
00:04:46: %RPM0-P:CP %CHMGR-3-PROFILE_MISMATCH: Mismatch: line card 1 has mismatch CAM
profile or microcode
```

Figure 10-1. EF Line Card with EG Chassis Profile—Card Problem

```
R1#show linecard 1 brief

--  Line card 1 --
Status         : card problem - mismatch cam profile
Next Boot      : online
Required Type : E48TF - 48-port 10/100/1000Base-T line card with RJ-45 interfaces (EF)
Current Type   : E48TF - 48-port 10/100/1000Base-T line card with RJ-45 interfaces (EF)
Hardware Rev   : Base  - 1.1  PP0 - 1.1  PP1 - 1.1
Num Ports      : 48
Up Time        : 0 sec
FTOS Version   : 7.6.1.0
Jumbo Capable : yes
```

Figure 10-2. EH Line Card with EG Chassis Profile—Card Problem

```
R1#show linecard 1 brief

--  Line card 1 --
Status         : card problem - mismatch cam profile
Next Boot      : online
Required Type : E90MH - 90-port 10/100/1000Base-T line card with mini RJ-21 interfaces (EH)
Current Type   : E90MH - 90-port 10/100/1000Base-T line card with mini RJ-21 interfaces (EH)
Hardware Rev   : Base  - 0.3 PP0 - 1.1  PP0 - PP1 -
Num Ports      : 90
Up Time        : 0 sec
FTOS Version   : 8.1.1.0
Jumbo Capable : yes
```

When to Use CAM Profiling

The CAM profiling feature enables you to partition the CAM to best suit your application. For example:

- Configure more Layer 2 FIB entries when the system is deployed as a switch.
- Configure more Layer 3 FIB entries when the system is deployed as a router.
- Configure more ACLs (when IPv6 is not employed).
- Hash MPLS packets based on source and destination IP addresses for LAGs. See LAG Hashing on page 230.
- Hash based on bidirectional flow for LAGs. See LAG Hashing based on Bidirectional Flow on page 231.
- Optimize the VLAN ACL Group feature, which permits group VLANs for IP egress ACLs. See CAM profile for the VLAN ACL group feature on page 231.

Important Points to Remember

- CAM Profiling is available on the E-Series TeraScale with FTOS versions 6.3.1.1 and later.
- All line cards within a single system must have the same CAM profile; this profile must match the system CAM profile (the profile on the primary RPM).
  - FTOS automatically reconfigures the CAM profile on line cards and the secondary RPM to match the system CAM profile by saving the correct profile on the card and then rebooting it.
- The CAM configuration is applied to entire system when you use CONFIGURATION mode commands. You must save the running-configuration to affect the change.
- All CAM configuration commands require you to reboot the system.
- When budgeting your CAM allocations for ACLs and QoS configurations, remember that ACL and QoS rules might consume more than one CAM entry depending on complexity. For example, TCP and UDP rules with port range options might require more than one CAM entry. See Pre-calculating Available QoS CAM Space on page 582.
- After you install a secondary RPM, copy the running-configuration to the startup-configuration so that the new RPM has the correct CAM profile.

Select CAM Profiles

A CAM profile is selected in CONFIGURATION mode. The CAM profile is applied to entire system, however, you must save the running-configuration to affect the change.

All components in the chassis must have the same CAM profile and microcode. The profile and microcode loaded on the primary RPM determines the profile that is required on all other chassis components.

- If a newly installed line card has a profile different from the primary RPM, the card reboots so that it can load the proper profile.

- If a the standby RPM has a profile different from the primary RPM, the card reboots so that it can load the proper profile.

To change the CAM profile on the entire system:

Step	Task	Command Syntax	Command Mode
1	Select a CAM profile.	**cam-profile** *profile* **microcode** *microcode*	CONFIGURATION
✎	**Note:** If selecting a cam-profile for VRF (**cam-profile ipv4-vrf or ipv4-v6-vrf**), implement the command in the CONFIGURATION mode only. If you use EXEC Privilege mode, the linecards may go into an error state.		
2	Save the running-configuration.	**copy running-config startup-config**	EXEC Privilege
3	Verify that the new CAM profile will be written to the CAM on the next boot.	**show cam-profile summary**	EXEC Privilege
4	Reload the system.	**reload**	EXEC Privilege

CAM Allocation

User Configurable CAM Allocations is available on platforms: C S

Allocate space for IPV4 ACLs and QoS regions, and IPv6 6 ACLs and QoS regions on the C-Series and S-Series by using the **cam-acl** command in CONFIGURATION mode.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. Note that there are 16 FP blocks, but the System Flow requires 3 blocks that cannot be reallocated. The default CAM Allocation settings on a C-Series system are:

- L3 ACL (ipv4acl): 6
- L2 ACL(l2acl) : 5
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1
- L2PT (l2pt): 1
- MAC ACLs (ipmacacl): 2
- ECFMACL (ecfmacl): 0
- 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.

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.

To configure the IPv4 and IPv6 ACLs and Qos regions on the entire system:

Step	Task	Command Syntax	Command Mode
1	Select a cam-acl action	**cam-acl** [**default** | **l2acl**]	CONFIGURATION
	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.	**l2acl** *number* **ipv4acl** *number* **ipv6acl** *number*, **ipv4qos** *number* **l2qos** *number*, **l2pt** *number* **ipmacacl** *number* **ecfmacl** *number* [**vman-qos** | **vman-dual-qos** *number*	EXEC Privilege
3	Verify that the new settings will be written to the CAM on the next boot.	**show cam-acl**	EXEC Privilege
4	Reload the system.	**reload**	EXEC Privilege

Test CAM Usage

The **test cam-usage** command is supported on platforms C E S

This command applies to both IPv4 and IPv6 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. Figure 10-3 gives a sample of the output shown when executing the command. The status column indicates whether or not the policy can be enabled.

Figure 10-3. Command Example: test cam-usage (C-Series)

```
Force10#test cam-usage service-policy input TestPolicy linecard all

Linecard | Portpipe | CAM Partition | Available CAM | Estimated CAM per Port | Status
------------------------------------------------------------------------------------------
       2 |        1 | IPv4Flow      |           232 |                      0 | Allowed
       2 |        1 | IPv6Flow      |             0 |                      0 | Allowed
       4 |        0 | IPv4Flow      |           232 |                      0 | Allowed
       4 |        0 | IPv6Flow      |             0 |                      0 | Allowed
Force10#
```

View CAM Profiles

View the current CAM profile for the chassis and each component using the command **show cam-profile**, as shown in Figure 10-4. This command also shows the profile that will be loaded upon the next chassis or component reload.

Figure 10-4. Viewing CAM Profiles on E-Series TeraScale

```
Force10#show cam-profile

-- Chassis CAM Profile --

CamSize           : 18-Meg
                  : Current Settings : Next Boot
Profile Name      : Default          : Default
L2FIB             : 32K entries      : 32K entries
L2ACL             : 1K entries       : 1K entries
IPv4FIB           : 256K entries     : 256K entries
IPv4ACL           : 12K entries      : 12K entries
IPv4Flow          : 24K entries      : 24K entries
EgL2ACL           : 1K entries       : 1K entries
EgIPv4ACL         : 1K entries       : 1K entries
Reserved          : 8K entries       : 8K entries
FIB           : 0  entries       : 0  entries
ACL           : 0  entries       : 0  entries
Flow          : 0  entries       : 0  entries
EgACL         : 0  entries       : 0  entries
MicroCode Name    : Default          : Default
--More--
```

View a brief output of the command **show cam-profile** using the **summary** option.

The command **show running-config cam-profile** shows the current profile and microcode (Figure 10-5).

Note: If you select the CAM profile from CONFIGURATION mode, the output of this command does not reflect any changes until you save the running-configuration and reload the chassis.

Figure 10-5. Viewing CAM Profile Information in the Running-configuration

```
Force10#show running-config cam-profile
!
cam-profile default microcode default

Force10#
```

View CAM-ACL settings

View the current cam-acl settings for the C-Series and S-Series systems chassis and each component using the command **show cam-acl**, as shown in Figure 10-6.

Figure 10-6. View CAM-ACl settings on C-Series and S-Series

```
Force10# show cam-acl

-- Chassis Cam ACL --
             Current Settings(in block sizes)
L2Acl        :         2
Ipv4Acl      :         2
Ipv6Acl      :         2
Ipv4Qos      :         2
L2Qos        :         2
L2PT         :         1
IpMacAcl     :         2
VmanQos      :         0
VmanDualQos  :         0

-- Line card 0 --
           Current Settings(in block sizes)
L2Acl        :         2
Ipv4Acl      :         2
Ipv6Acl      :         2
Ipv4Qos      :         2
L2Qos        :         2
L2PT         :         1
IpMacAcl     :         2
VmanQos      :         0
VmanDualQos  :         0

-- Line card 6 --
           Current Settings(in block sizes)
L2Acl        :         2
Ipv4Acl      :         2
Ipv6Acl      :         2
Ipv4Qos      :         2
L2Qos        :         2
L2PT         :         1
IpMacAcl     :         2
VmanQos      :         0
VmanDualQos  :         0
```

View 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 command **show cam-usage** from EXEC Privilege mode, as shown in Figure 10-7.

Figure 10-7. Viewing CAM Usage Information

```
R1#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--
```

CAM Optimization

CAM optimization is supported on platforms C S

When this command is enabled, 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 will be used). When the command is disabled, the system behaves as described in this chapter.

Applications for CAM Profiling

LAG Hashing

FTOS includes a CAM profile and microcode that treats MPLS packets as non-IP packets. Normally, switching and LAG hashing is based on source and destination MAC addresses. Alternatively, you can base LAG hashing for MPLS packets on source and destination IP addresses. This type of hashing is allowed for MPLS packets with 5 labels or less.

MPLS packets are treated as follows:

- When MPLS IP packets are received, FTOS looks up to 5 labels deep for the IP header.

- When an IP header is present, hashing is based on IP 3 tuple (source IP address, destination IP address, and IP protocol).
- If an IP header is not found after the 5th label, hashing is based on the MPLS labels.
- If the packet has more than 5 MPLS labels, hashing is based on the source and destination MAC address.

To enable this type of hashing, use the default CAM profile with the microcode *lag-hash-mpls*.

LAG Hashing based on Bidirectional Flow

To hash LAG packets such that both directions of a bidirectional flow (for example, VoIP or P2P file sharing) are mapped to the same output link in the LAG bundle, use the default CAM profile with the microcode *lag-hash-align*.

CAM profile for the VLAN ACL group feature

IPv4Flow sub-partitions are supported on platform E T only.

To optimize for the VLAN ACL Group feature, which permits group VLANs for the IP egress ACL, use the CAM profile *ipv4-egacl-16k* with the default microcode.

Note: Do not use this CAM profile for Layer 2 egress ACLs.

Troubleshoot CAM Profiling

CAM Profile Mismatches

The CAM profile on all cards must match the system profile. In most cases, the system corrects mismatches by copying the correct profile to the card, and rebooting the card. If three resets do not bring up the card, or if the system is running an FTOS version prior to 6.3.1.1, the system presents an error message. In this case, manually adjust the CAM configuration on the card to match the system configuration.

Dell Force10 recommends the following to prevent mismatches:

- Use the eg-default CAM profile in a chassis that has only EG Series line cards. If this profile is used in a chassis with non-EG line cards, the non-EG line cards enter a problem state.
- Before moving a card to a new chassis, change the CAM profile on a card to match the new system profile.
- After installing a secondary RPM into a chassis, copy the running-configuration to the startup-configuration.
- Change to the default profile if downgrading to and FTOS version earlier than 6.3.1.1.

- Use the CONFIGURATION mode commands so that the profile is change throughout the system.
- Use the EXEC Privilege mode commands to match the profile of a component to the profile of the target system.

QoS CAM Region Limitation

The default CAM profile allocates a partition within the IPv4Flow region to store QoS service policies. If the QoS CAM space is exceeded, messages similar to the ones in Message 3 are displayed.

Message 3 QoS CAM Region Exceeded

```
%EX2YD:12 %DIFFSERV-2-DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for
class 2 (Gi 12/20) entries on portpipe 1 for linecard 12
%EX2YD:12 %DIFFSERV-2-
DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for class 5 (Gi 12/22)
entries on portpipe 1 for linecard 12
```

If you exceed the QoS CAM space:

Step	Task
1	Verify that you have configured a CAM profile that allocates 24K entries to the IPv4 system flow region. See View CAM Profiles on page 228.
2	Allocate more entries in the IPv4Flow region to QoS. .

FTOS version 7.4.1 introduced the ability to view the actual CAM usage before applying a service-policy. The command **test cam-usage service-policy** provides this test framework, see Pre-calculating Available QoS CAM Space on page 582.

Note: For troubleshooting other CAM issues see the *E-Series Network Operations Guide*.

11

Dynamic Host Configuration Protocol (DHCP)

Dynamic Host Configuration Protocol (DHCP) is available on platforms: E C S S55 S60 S4810
□

This chapter contains the following sections:

- Protocol Overview
- Implementation Information
- Configuration Tasks
- Configure the System to be a DHCP Server
- Configure the System to be a Relay Agent
- Configure the System for User Port Stacking
- Configure Secure DHCP

Protocol Overview

Dynamic Host Configuration Protocol (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.
- 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**—a network device offering configuration parameters to the client
- **DHCP Client**—a network device requesting configuration parameters from the server

- **Relay agent**—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

DHCP Packet Format and Options

DHCP uses 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; some common options are given in Table 11-1.

Figure 11-1. DHCP Packet Format

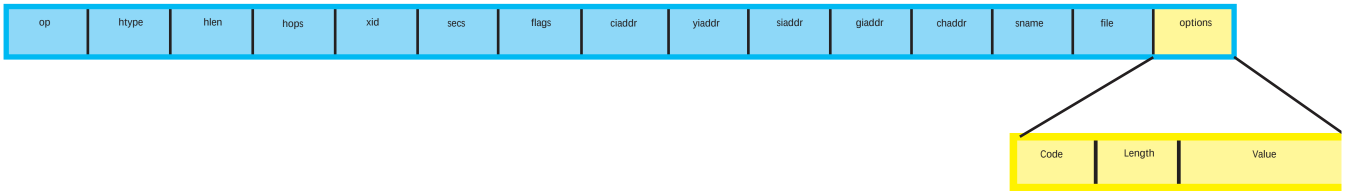


Table 11-1. Common DHCP Options

Option	Code	Description
Subnet Mask	1	Specifies the client’s subnet mask.
Router	3	Specifies the router IP addresses that may serve as the client’s default gateway.
Domain Name Server	6	Specifies the DNS servers that are available to the client.
Domain Name	15	Specifies the domain name that clients should use when resolving hostnames via DNS.
IP Address Lease Time	51	Specifies the amount of time that the client is allowed to use an assigned IP address.
DHCP Message Type	53	1: DHCPDISCOVER 2: DHCPOFFER 3: DHCPREQUEST 4: DHCPDECLINE 5: DHCPACK 6: DHCPNACK 7: DHCPRELEASE 8: DHCPINFORM
Parameter Request List	55	Clients use this option to tell the server which parameters it requires. It is a series of octets where each octet is DHCP option code.
Renewal Time	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	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	255	Signals the last option in the DHCP packet.

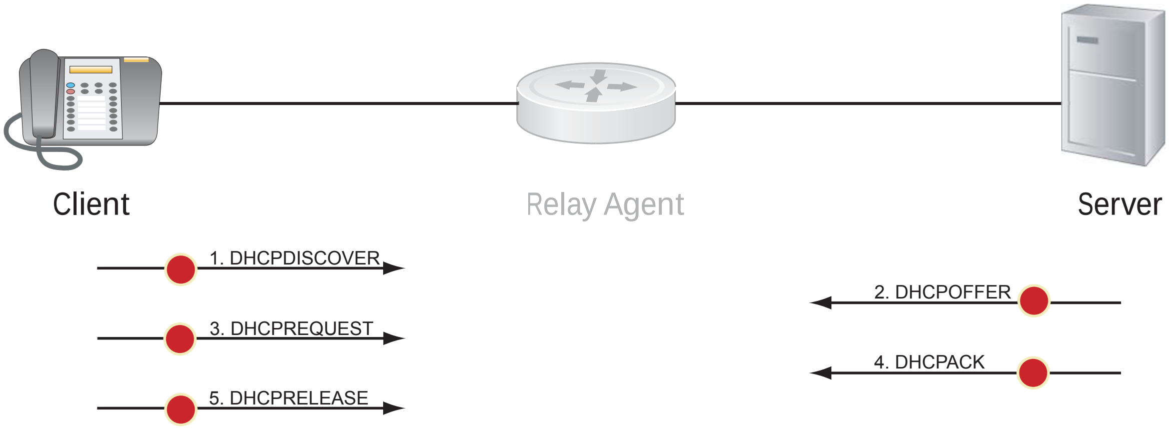
Assigning an IP Address using DHCP

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. Upon receiving a DHCPREQUEST, the server binds the clients' unique identifier (the hardware address plus IP address) to the accepted configuration parameters and stores the data in a database called a *binding table*. The server then broadcasts a **DHCPACK** message, which signals to the client that it may begin using the assigned parameters.
5. When the client leaves the network, or the lease time expires, returns its IP address to the server in a **DHCPRELEASE** message.

There are additional messages that are used in case the DHCP negotiation deviates from the process previously described and shown in the illustration below.

- **DHCPDECLINE—**A client sends this message to the server in response to a DHCPACK if the configuration parameters are unacceptable, for example, if the offered address is already in use. In this case, the client starts the configuration process over by sending a DHCPDISCOVER.
- **DHCPINFORM—**A client uses this message to request configuration parameters when it assigned an IP address manually rather than with DHCP. The server responds by unicast.
- **DHCPNAK—**A server sends this message to the client if it is not able to fulfill a DHCPREQUEST, for example if the requested address is already in use. In this case, the client starts the configuration process over by sending a DHCPDISCOVER.



Implementation Information

- The Dell Force10 implementation of DHCP is based on RFC 2131 and RFC 3046.
- IP Source Address Validation is a sub-feature of DHCP Snooping; FTOS uses ACLs internally to implement this feature and as such, you cannot apply ACLs to an interface which has IP Source Address Validation. If you configure IP Source Address Validation on a member port of a VLAN and then attempt to apply a access list to the VLAN, FTOS displays the first line in Message 1. If you first apply an ACL to a VLAN and then attempt enable IP Source Address Validation on one of its member ports, FTOS displays the second line in Message 1.

Message 1 DHCP Snooping with VLAN ACL Compatibility Error

```
% Error: Vlan member has access-list configured.
% Error: Vlan has an access-list configured.
```

Note: If DHCP snooping is enabled globally and any L2 port is configured, any IP ACL,MAC ACL, or DHCP Source-Address validation ACL won't block DHCP packets .

- FTOS provides 40K entries that can be divided between leased addresses and excluded addresses. By extension, the maximum number of pools you can configure depends on the on the subnet mask that you give to each pool. For example, if all pools were configured for a /24 mask, the total would be 40000/253 (approximately 158). If the subnet is increased, more pools can be configured. The maximum subnet that can be configured for a single pool is /17. FTOS displays an error message for configurations that exceed the allocated memory.
- E-Series supports 16K DHCP Snooping entries across 500 VLANs.
- C-Series, S-Series (S25/S50), S55, S60 and S4810 support 4K DHCP Snooping entries.
- All platforms support Dynamic ARP Inspection on 16 VLANs per system. Refer to Dynamic ARP Inspection.

Note: 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 JumpStart mode, it will not be able to reach the DHCP server, resulting in BMP failure.

Configuration Tasks

- Configure the System to be a DHCP Server
- Configure the System to be a Relay Agent
- Configure Secure DHCP

Configure the System to be a DHCP Server

Configure the System to be a DHCP Server is supported only on platforms: Z and □(S25/S50), S55 , S60 , and S4810

A DHCP server is a network device that has been programmed to provide network configuration parameters to clients upon request. Servers typically serve many clients, making host management much more organized and efficient.

The key responsibilities of DHCP servers are:

1. **Address Storage and Management**: DHCP servers are the owners of the addresses used by DHCP clients.The server stores the addresses and manages their use, keeping track of which addresses have been allocated and which are still available.
2. **Configuration Parameter Storage and Management**: DHCP servers also store and maintain other parameters that are sent to clients when requested. These parameters specify in detail how a client is to operate.
3. **Lease Management**: DHCP servers use leases to allocate addresses to clients for a limited time. The DHCP server maintains information about each of the leases, including lease length.
4. **Responding To Client Requests**: DHCP servers respond to different types of requests from clients, primarily, granting, renewing, and terminating leases.
5. **Providing Administration Services**: The DHCP server includes functionality that allows an administrator to implement policies that govern how DHCP performs its other tasks.

Configuration Tasks

To configure DHCP, an administrator must first set up a DHCP server and provide it with configuration parameters and policy information including IP address ranges, lease length specifications, and configuration data that DHCP hosts need.

Configuring the Dell Force10 system to be a DHCP server is a 3-step process:

1. Configure the Server for Automatic Address Allocation
2. Specify a Default Gateway
3. Enable DHCP Server

Related Configuration Tasks

- Configure a Method of Hostname Resolution
- Create Manual Binding Entries
- Debug DHCP server
- DHCP Clear Commands

Configure the Server for Automatic Address Allocation

This feature is available on C and S (S25/S50), S55, S60, and S4810 platforms only.

Automatic Address Allocation is an address assignment method by which the DHCP server leases an IP address to a client from a pool of available addresses.

Create an IP Address Pool

An address pool is a range of IP addresses that may be assigned by the DHCP server. Address pools are indexed by subnet number. To create an address pool:

Step	Task	Command Syntax	Command Mode
1	Access the DHCP server CLI context.	**ip dhcp server**	CONFIGURATION
2	Create an address pool and give it a name.	**pool** *name*	DHCP
3	Specify the range of IP addresses from which the DHCP server may assign addresses. • *network* is the subnet address. • *prefix-length* specifies the number of bits used for the network portion of the address you specify.	**network** *network lprefix-length* Prefix-length Range: 17 to 31	DHCP <POOL>
4	Display the current pool configuration.	**show config**	DHCP <POOL>

Once an IP address is leased to a client, only that client may release the address. FTOS performs a IP + MAC source address validation to ensure that no client can release another clients address. This is a default behavior and is separate from IP+MAC Source Address Validation.

Exclude Addresses from the Address Pool

The DHCP server assumes that all IP addresses in a DHCP address pool are available for assigning to DHCP clients. You must specify the IP address that the DHCP server should not assign to clients.

Task	Command Syntax	Command Mode
Exclude an address range from DHCP assignment. The exclusion applies to all configured pools.	**excluded-address**	DHCP

Specify an Address Lease Time

Task	Command Syntax	Command Mode
Specify an address lease time for the addresses in a pool.	**lease** {**days** [**hours**] [**minutes**] | **infinite**} Default: 24 hours	DHCP <POOL>

Specify a Default Gateway

The IP address of the default router should be on the same subnet as the client.

Task	Command Syntax	Command Mode
Specify default gateway(s) for the clients on the subnet, in order of preference.	**default-router** *address*	DHCP <POOL>

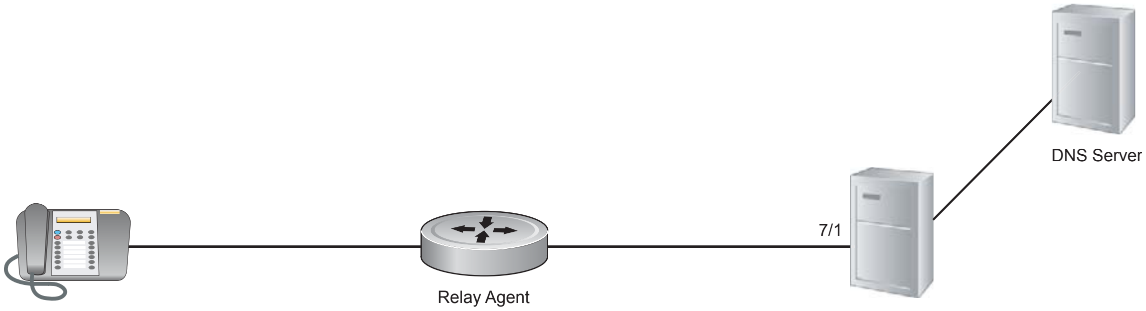
Enable DHCP Server

This feature is available on C and S (S25/S50), S55 , S60 , and S4810 platforms only.

The DHCP server is disabled by default.

Step	Task	Command Syntax	Command Mode
1	Enter the DHCP command-line context.	**ip dhcp server**	CONFIGURATION
2	Enable DHCP server.	**no disable** Default: Disabled	DHCP
3	Display the current DHCP configuration.	**show config**	DHCP

In the illustration below, an IP phone is powered by PoE and has acquired an IP address from the Dell Force10 system, which is advertising LLDP-MED. The leased IP address is displayed using **show ip dhcp binding**, and confirmed with **show lldp neighbors**.



Configure a Method of Hostname Resolution

Dell Force10 systems are capable of providing DHCP clients with parameters for two methods of hostname resolution.

Address Resolution using DNS

A domain is a group of networks. DHCP clients query DNS IP servers when they need to correlate host names to IP addresses.

Step	Task	Command Syntax	Command Mode
1	Create a domain.	**domain-name** *name*	DHCP <POOL>
2	Specify in order of preference the DNS servers that are available to a DHCP client.	**dns-server** *address*	DHCP <POOL>

Address Resolution using NetBIOS WINS

Windows Internet Naming Service (WINS) is a name resolution service that Microsoft DHCP clients use to correlate host names to IP addresses within a group of networks. Microsoft DHCP clients can be one of four types of NetBIOS nodes: broadcast, peer-to-peer, mixed, or hybrid.

Step	Task	Command Syntax	Command Mode
1	Specify the NetBIOS Windows Internet Naming Service (WINS) name servers, in order of preference, that are available to Microsoft Dynamic Host Configuration Protocol (DHCP) clients.	**netbios-name-server** *address*	DHCP <POOL>
2	Specify the NetBIOS node type for a Microsoft DHCP client. Dell Force10 recommends specifying clients as hybrid.	**netbios-node-type** *type*	DHCP <POOL>

Create Manual Binding Entries

An address binding is a mapping between the IP address and Media Access Control (MAC) address of a client. The DHCP server assigns the client an available IP address automatically, and then creates a entry in the binding table. However, the administrator can manually create an entry for a client; manual bindings are useful when you want to guarantee that a particular network device receives a particular IP address. Manual bindings can be considered single-host address pools. There is no limit on the number of manual bindings, but you can only configure one manual binding per host.

Note: FTOS does not prevent you from using a network IP as a host IP; be sure to not use a network IP as a host IP.

To create a manual binding:

Step	Task	Command Syntax	Command Mode
1	Create an address pool	**pool** *name*	DHCP
2	Specify the client IP address.	**host** *address*	DHCP <POOL>
3	Specify the client hardware address. • *hardware-address* is the client MAC address. *type* is the protocol of the hardware platform. The default protocol is Ethernet.	**hardware-address** *hardware-address type*	DHCP <POOL>

Debug DHCP server

Task	Command Syntax	Command Mode
Display debug information for DHCP server.	**debug ip dhcp server [events | packets]**	EXEC Privilege

DHCP Clear Commands

Task	Command Syntax	Command Mode
Clear DHCP binding entries for the entire binding table.	**clear ip dhcp binding**	EXEC Privilege
Clear a DHCP binding entry for an individual IP address.	**clear ip dhcp binding** *ip address*	EXEC Privilege
Clear a DHCP address conflict.	**clear ip dhcp conflict**	EXEC Privilege
Clear DHCP server counters.	**clear ip dhcp server statistics**	EXEC Privilege
Clear the DHCP binding table.	**clear ip dhcp snooping**	EXEC Privilege

Configure the System to be a Relay Agent

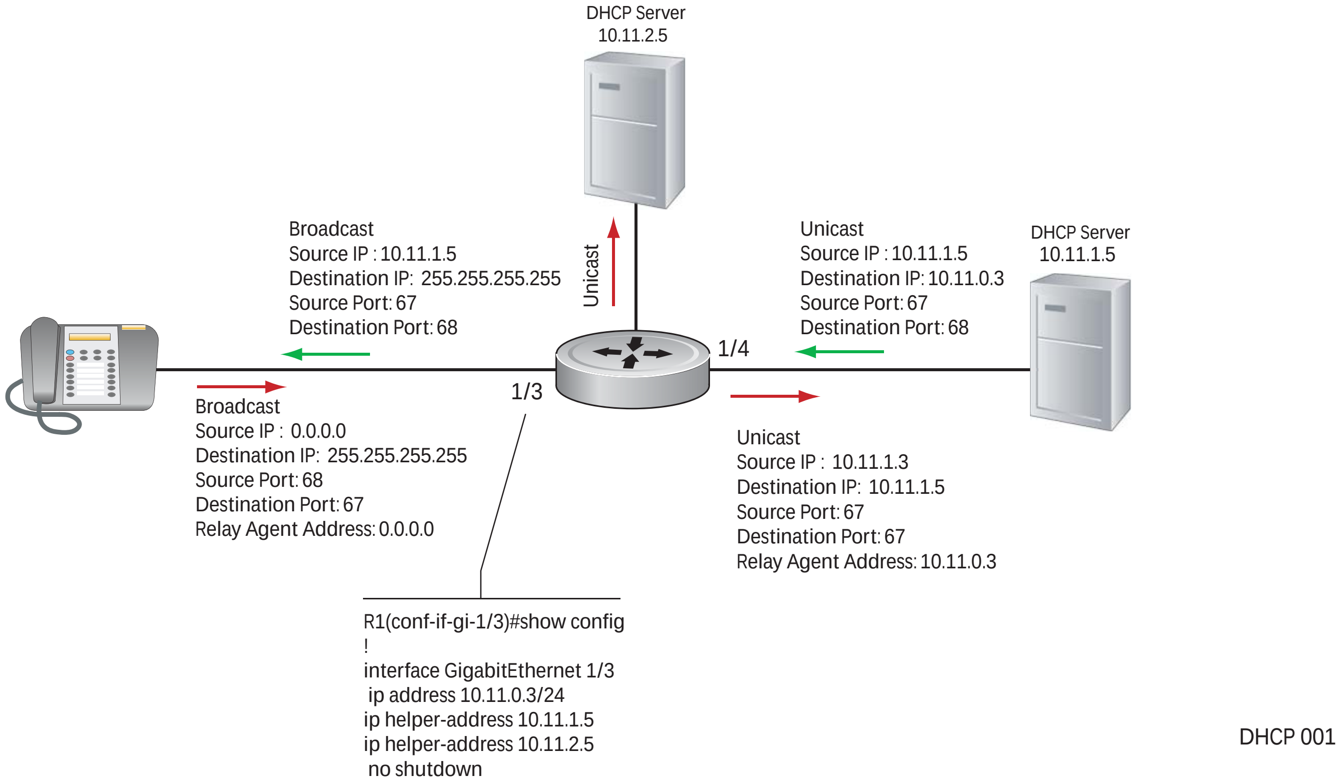
The following feature is available on platforms: ZE S4810

DHCP clients and servers request and offer configuration information via broadcast DHCP messages. Routers do not forward broadcasts, so if there are no DHCP servers on the subnet, the client does not receive a response to its request and therefore cannot access the network.

You can configure an interface on the Dell Force10 system to relay the DHCP messages to a specific DHCP server using the command **ip helper-address** *dhcp-address* from INTERFACE mode, as shown in the illustration below. Specify multiple DHCP servers by entering the **ip helper-address** *dhcp-address* command multiple times.

When **ip helper-address** is configured, the system listens for DHCP broadcast messages on port 67. The system rewrites packets received from the client and forwards it via unicast; the system rewrites the destination IP address and writes its own address as the relay device. Responses from the server are unicast back to the relay agent on port 68 and the relay agent rewrites the destination address and forwards the packet to the client subnet via broadcast.

Note: DHCP Relay is not available on Layer 2 interfaces and VLANs.



```
R1(conf-if-gi-1/3)#show config
!
interface GigabitEthernet 1/3
 ip address 10.11.0.3/24
ip helper-address 10.11.1.5
ip helper-address 10.11.2.5
 no shutdown
```

To view the **ip helper-address** configuration for an interface, use the command **show ip interface** from EXEC privilege mode, as shown in the following example.

```
R1_E600#show ip int gig 1/3
GigabitEthernet 1/3 is up, line protocol is down
Internet address is 10.11.0.1/24
Broadcast address is 10.11.0.255
Address determined by user input
IP MTU is 1500 bytes
Helper address is 192.168.0.1
                  192.168.0.2
Directed broadcast forwarding is disabled
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent
```

Configure the System for User Port Stacking

When you set the DHCP offer on the DHCP server, you can set the stacking-option variable to provide the stack-port detail so a stack can be formed when the units are connected.

Configure Secure DHCP

The following feature is available on platforms: Z E S4810 and Z except where noted.

DHCP as defined by RFC 2131 provides no authentication or security mechanisms. Secure DHCP is a suite of features that protects networks that use dynamic address allocation from spoofing and attacks.

- Option 82
- DHCP Snooping
- Dynamic ARP Inspection
- Source Address Validation

Option 82

RFC 3046 (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** is the interface on which the client-originated message is received.
- **Remote ID** identifies the host from which the message is received. The value of this sub-option is the MAC address of the switch. On the S4810, S55, and S60, the Remote ID can also be the hostname of the switch or an arbitrary string.

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.

Task	Command Syntax	Command Mode
All platforms: Enable Option 82. Remote ID is the MAC of the switch.	**ip dhcp relay information-option**	CONFIGURATION
S4810, S60, S55: Enables Option 82. Remote ID is the hostname of the switch.	**ip dhcp relay information-option remote-id** *hostname*	CONFIGURATION
S4810, S60, S55: Enables Option 82. Remote ID is *remote-id*	**ip dhcp relay information-option remote-id** *remote-id*	CONFIGURATION

DHCP Snooping

DHCP Snooping protects networks from spoofing. In the context of DHCP Snooping, all ports are either trusted or untrusted. By default, all ports are untrusted. Trusted ports are ports through which attackers cannot connect. Manually configure ports connected to legitimate servers and relay agents as trusted.

When DHCP Snooping is enabled, the relay agent builds a binding table—using DHCPACK messages—containing the client MAC address, IP addresses, IP address lease time, port, VLAN ID, and binding type. Every time the relay agent receives a DHCPACK on an trusted port, it adds an entry to the table.

The relay agent then checks all subsequent DHCP client-originated IP traffic (DHCPRELEASE, DHCPNACK, and DHCPDECLINE) against the binding table to ensure that the MAC-IP address pair is legitimate, and that the packet arrived on the correct port; packets that do not pass this check are forwarded to the server for validation. This check-point prevents an attacker from spoofing a client and declining or releasing the real client’s address. Server-originated packets (DHCPOFFER, DHCPACK, DHCPNACK) that arrive on an untrusted port are also dropped. This check-point prevents an attacker from impostering as a DHCP server to facilitate a man-in-the-middle attack.

Binding table entries are deleted when a lease expires, or the relay agent encounters a DHCPRELEASE, DHCPNACK, DHCPDECLINE.

FTOS Behavior: Introduced in FTOS version 7.8.1.0, DHCP Snooping was available for Layer 3 only and dependent on DHCP Relay Agent (**ip helper-address**). FTOS version 8.2.1.0 extends DHCP Snooping to Layer 2, and you do not have to enable relay agent to snoop on Layer 2 interfaces.

FTOS Behavior: Binding table entries are deleted when a lease expires or when the relay agent encounters a DHCPRELEASE. Starting with FTOS Release 8.2.1.2, line cards maintain a list of snooped VLANs. When the binding table is exhausted, DHCP packets are dropped on snooped VLANs, while these packets are forwarded across non-snooped VLANs. Since DHCP packets are dropped, no new IP address assignments are made. However, DHCPRELEASE and DHCPDECLINE packets are allowed so that the DHCP snooping table can decrease in size. Once the table usage falls below the maximum limit of 4000 entries, new IP address assignments are allowed.

Note: DHCP server packets will be dropped on all untrusted interfaces of a system configured for DHCP snooping. To prevent these packets from being dropped, configure **ip dhcp snooping trust** on the server-connected port.

Enable DCHP snooping

Step	Task	Command Syntax	Command Mode
1	Enable DHCP Snooping globally.	**ip dhcp snooping**	CONFIGURATION
2	Specify ports connected to DHCP servers as trusted.	**ip dhcp snooping trust**	INTERFACE
3	Enable DHCP Snooping on a VLAN.	**ip dhcp snooping vlan**	CONFIGURATION

Add a static entry in the binding table

Task	Command Syntax	Command Mode
Add a static entry in the binding table.	**ip dhcp snooping binding mac**	EXEC Privilege

Clear the binding table

Task	Command Syntax	Command Mode
Delete all of the entries in the binding table	**clear ip dhcp snooping binding**	EXEC Privilege

Display the contents of the binding table

Task	Command Syntax	Command Mode
Display the contents of the binding table.	**show ip dhcp snooping**	EXEC Privilege

View the DHCP Snooping statistics with the **show ip dhcp snooping** command.

```
FTOS#show ip dhcp snooping

IP DHCP Snooping                           : Enabled.
IP DHCP Snooping Mac Verification          : Disabled.
IP DHCP Relay Information-option           : Disabled.
IP DHCP Relay Trust Downstream             : Disabled.

Database write-delay (In minutes)          : 0


DHCP packets information
Relay Information-option packets           : 0
Relay Trust downstream packets             : 0
Snooping packets                           : 0

Packets received on snooping disabled L3 Ports   : 0
Snooping packets processed on L2 vlans     : 142

DHCP Binding File Details
Invalid File                               : 0
Invalid Binding Entry                      : 0
Binding Entry lease expired                : 0
List of Trust Ports                        :Te 0/49
List of DHCP Snooping Enabled Vlans        :Vl 10
List of DAI Trust ports                    :Te 0/49
```

Drop DHCP packets on snooped VLANs only

Binding table entries are deleted when a lease expires or the relay agent encounters a DHCPRELEASE.

Starting with FTOS Release 8.2.1.1, line cards maintain a list of snooped VLANs. When the binding table fills, DHCP packets are dropped only on snooped VLANs, while such packets will be forwarded across non-snooped VLANs. Since DHCP packets are dropped, no new IP address assignments are made. However, DHCP Release and Decline packets are allowed so that the DHCP snooping table can decrease in size. Once the table usage falls below the max limit of 4000 entries, new IP address assignments are allowed.

View the number of entries in the table with the **show ip dhcp snooping binding** command. This output displays the snooping binding table created using the ACK packets from the trusted port.

```
FTOS#show ip dhcp snooping binding

Codes :  S - Static D - Dynamic

IP Address       MAC Address        Expires(Sec)  Type  VLAN   Interface
========================================================================
10.1.1.251       00:00:4d:57:f2:50    172800        D    Vl 10    Gi 0/2
10.1.1.252       00:00:4d:57:e6:f6    172800        D    Vl 10    Gi 0/1
10.1.1.253       00:00:4d:57:f8:e8    172740        D    Vl 10    Gi 0/3
10.1.1.254       00:00:4d:69:e8:f2    172740        D    Vl 10    Te 0/50

Total number of Entries in the table : 4
```

Dynamic ARP Inspection

Dynamic ARP inspection prevents ARP spoofing by forwarding only ARP frames that have been validated against the DHCP binding table.

ARP is a stateless protocol that provides no authentication mechanism. Network devices accept ARP requests and replies from any device, and ARP replies are accepted even when no request was sent. If a client receives an ARP message for which a relevant entry already exists in its ARP cache, it overwrites the existing entry with the new information.

The lack of authentication in ARP makes it vulnerable to spoofing. ARP spoofing is a technique attackers use to inject false IP-to-MAC mappings into the ARP cache of a network device. It is used to launch man-in-the-middle (MITM), and denial-of-service (DoS) attacks, among others.

A spoofed ARP message is one in which the MAC address in the sender hardware address field and the IP address in the sender protocol field are strategically chosen by the attacker. For example, in an MITM attack, the attacker sends a client an ARP message containing the attacker’s MAC address and the gateway’s IP address. The client then thinks that the attacker is the gateway, and sends all internet-bound packets to it. Likewise, the attacker sends the gateway an ARP message containing the attacker’s MAC address and the client’s IP address. The gateway then thinks that the attacker is the client and forwards all packets addressed to the client to it. As a result, the attacker is able to sniff all packets to and from the client.

Other attacks using ARP spoofing include:

- broadcast—an attacker can broadcast an ARP reply that specifies FF:FF:FF:FF:FF:FF as the gateway’s MAC address, resulting in all clients broadcasting all internet-bound packets.
- MAC flooding—an attacker can send fraudulent ARP messages to the gateway until the ARP cache is exhausted, after which, traffic from the gateway is broadcast.

- denial of service—an attacker can send a fraudulent ARP messages to a client to associate a false MAC address with the gateway address, which would blackhole all internet-bound packets from the client.

Note: DAI uses entries in the L2SysFlow CAM region, a sub-region of SystemFlow. One CAM entry is required for every DAI-enabled VLAN. You can enable DAI on up to 16 VLANs on a system. However, the ExaScale default CAM profile allocates only 9 entries to the L2SysFlow region for DAI. You can configure 10 to 16 DAI-enabled VLANs by allocating more CAM space to the L2SysFlow region before enabling DAI.

SystemFlow has 102 entries by default. This region is comprised of two sub-regions: L2Protocol and L2SystemFlow. L2Protocol has 87 entries, and L2SystemFlow has 15 entries. Six L2SystemFlow entries are used by Layer 2 protocols, leaving 9 for DAI. L2Protocol can have a maximum of 100 entries, and this region must be expanded to capacity before you can increase the size of L2SystemFlow. This is relevant when you are enabling DAI on VLANs. If, for example, you want to enable DAI on 16 VLANs, you need 7 more entries; in this case, reconfigure the SystemFlow region for 122 entries:

layer-2 eg-acl *value* **fib** *value* **frrp** *value* **ing-acl** *value* **learn** *value* **l2pt** *value* **qos** *value* **system-flow 122**

The logic is as follows:

L2Protocol has 87 entries by default and must be expanded to its maximum capacity, 100 entries, before L2SystemFlow can be increased; therefore 13 more L2Protocol entries are required. L2SystemFlow has 15 entries by default, but only 9 are for DAI; to enable DAI on 16 VLANs, 7 more entries are required. 87 L2Protocol + 13 additional L2Protocol + 15 L2SystemFlow + 7 additional L2SystemFlow equals 122.

Step	Task	Command Syntax	Command Mode
1	Enable DHCP Snooping.		
2	Validate ARP frames against the DHCP Snooping binding table.	**arp inspection**	INTERFACE VLAN

View the number of entries in the ARP database with the **show arp inspection database** command.

```
FTOS#show arp inspection database

Protocol    Address         Age(min)  Hardware Address    Interface   VLAN  CPU
-----------------------------------------------------------------------------
Internet    10.1.1.251          -   00:00:4d:57:f2:50    Gi 0/2     Vl 10  CP
Internet    10.1.1.252          -   00:00:4d:57:e6:f6    Gi 0/1     Vl 10  CP
Internet    10.1.1.253          -   00:00:4d:57:f8:e8    Gi 0/3     Vl 10  CP
Internet    10.1.1.254          -   00:00:4d:69:e8:f2    Te 0/50     Vl 10  CP
FTOS#
```

Use **show arp inspection statistics** command to see how many valid and invalid ARP packets have been processed.

```
FTOS#show arp inspection statistics

Dynamic ARP Inspection (DAI) Statistics
---------------------------------------
Valid ARP Requests                      : 0
Valid ARP Replies                       : 1000
Invalid ARP Requests                    : 1000
Invalid ARP Replies                     : 0
FTOS#
```

Bypass the ARP Inspection

You can configure a port to skip ARP inspection by defining the interface as trusted, which is useful in multi-switch environments. ARPs received on trusted ports bypass validation against the binding table. All ports are untrusted by default.

Task	Command Syntax	Command Mode
Specify an interface as trusted so that ARPs are not validated against the binding table.	**arp inspection-trust**	INTERFACE

FTOS Behavior: Introduced in FTOS version 8.2.1.0, Dynamic ARP Inspection (DAI) was available for Layer 3 only. FTOS version 8.2.1.1 extends DAI to Layer 2.

Source Address Validation

Using the DHCP binding table, FTOS can perform three types of source address validation (SAV):

- IP Source Address Validation prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table.
- DHCP MAC Source Address Validation verifies a DHCP packet's source hardware address matches the client hardware address field (CHADDR) in the payload.
- IP+MAC Source Address Validation verifies that the IP source address and MAC source address are a legitimate pair.

IP Source Address Validation

IP Source Address Validation (SAV) prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table. A spoofed IP packet is one in which the IP source address is strategically chosen to disguise the attacker. For example, using ARP spoofing an attacker can assume a legitimate client's identity and receive traffic addressed to it. Then the attacker can spoof the client's IP address to interact with other clients.

The DHCP binding table associates addresses assigned by the DHCP servers, with the port on which the requesting client is attached. When IP Source Address Validation is enabled on a port, the system verifies that the source IP address is one that is associated with the incoming port. If an attacker is impostering as a legitimate client the source address appears on the wrong ingress port, and the system drops the packet. Likewise, if the IP address is fake, the address will not be on the list of permissible addresses for the port, and the packet is dropped.

Task	Command Syntax	Command Mode
Enable IP Source Address Validation	**ip dhcp source-address-validation**	INTERFACE

Note: If IP Source Guard is enabled using the **ip dhcp source-address-validation** command and there are 187 entries or more in the current DHCP snooping binding table, Source Address Validation (SAV) may not be applied to all entries.
To ensure that SAV is applied correctly to all entries, enable the **ip dhcp source-address-validation** command before adding entries to the binding table.

DHCP MAC Source Address Validation

DHCP MAC Source Address Validation (SAV) validates a DHCP packet's source hardware address against the client hardware address field (CHADDR) in the payload.

FTOS Release 8.2.1.1 ensures that the packet's source MAC address is checked against the CHADDR field in the DHCP header only for packets from snooped VLANs.

Task	Command Syntax	Command Mode
Enable DHCP MAC Source Address Validation.	**ip dhcp snooping verify mac-address**	CONFIGURATION

IP+MAC Source Address Validation

The following feature is available on platforms: C, S and S4810.

IP Source Address Validation validates the IP source address of an incoming packet against the DHCP Snooping binding table. IP+MAC Source Address Validation ensures that the IP source address and MAC source address are a legitimate pair, rather validating each attribute individually. IP+MAC Source Address Validation cannot be configured with IP Source Address Validation.

Step	Task	Command Syntax	Command Mode
1	Allocate at least one FP block to the ipmacacl CAM region.	**cam-acl l2acl**	CONFIGURATION
2	Save the running-config to the startup-config.	**copy running-config startup-config**	EXEC Privilege
3	Reload the system.	**reload**	EXEC Privilege

Step	Task	Command Syntax	Command Mode
4	Enable IP+MAC Source Address Validation.	**ip dhcp source-address-validation ipmac**	INTERFACE

FTOS creates an ACL entry for each IP+MAC address pair in the binding table and applies it to the interface.

Task	Command Syntax	Command Mode
Display the IP+MAC ACL for an interface for the entire system.	**show ip dhcp snooping source-address-validation** [**interface**]	EXEC Privilege

12

Dell Force10 Resilient Ring Protocol

Dell Force10 Resilient Ring Protocol is supported on platforms C E S

Dell Force10 Resilient Ring Protocol (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. The Dell Force10 Resilient Ring Protocol (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.

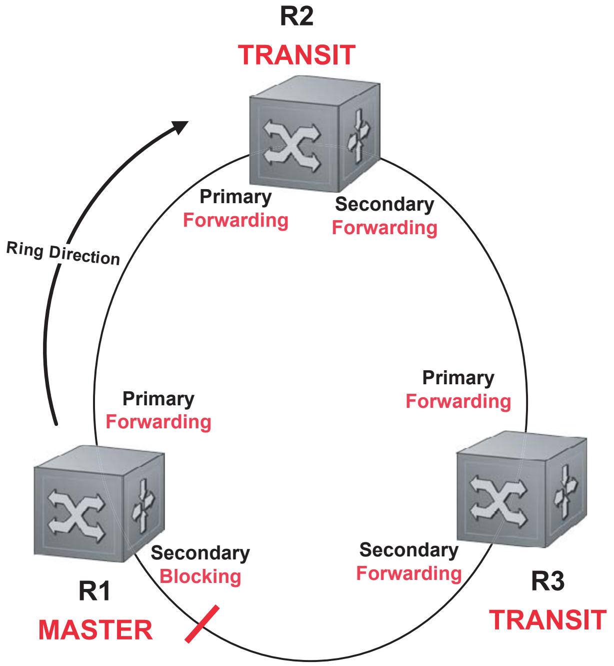
Protocol Overview

FRRP is built on a ring topology. Up to 255 rings can be configured 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, it 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. See Figure 12-1 for a simple example of this FRRP topology. Note that ring direction is determined by the Master node’s Primary and Secondary ports.

Figure 12-1. 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 Ring Health Frame (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. Once the ring is restored, the next status check frame is received on the Master node's Secondary port. This will cause 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

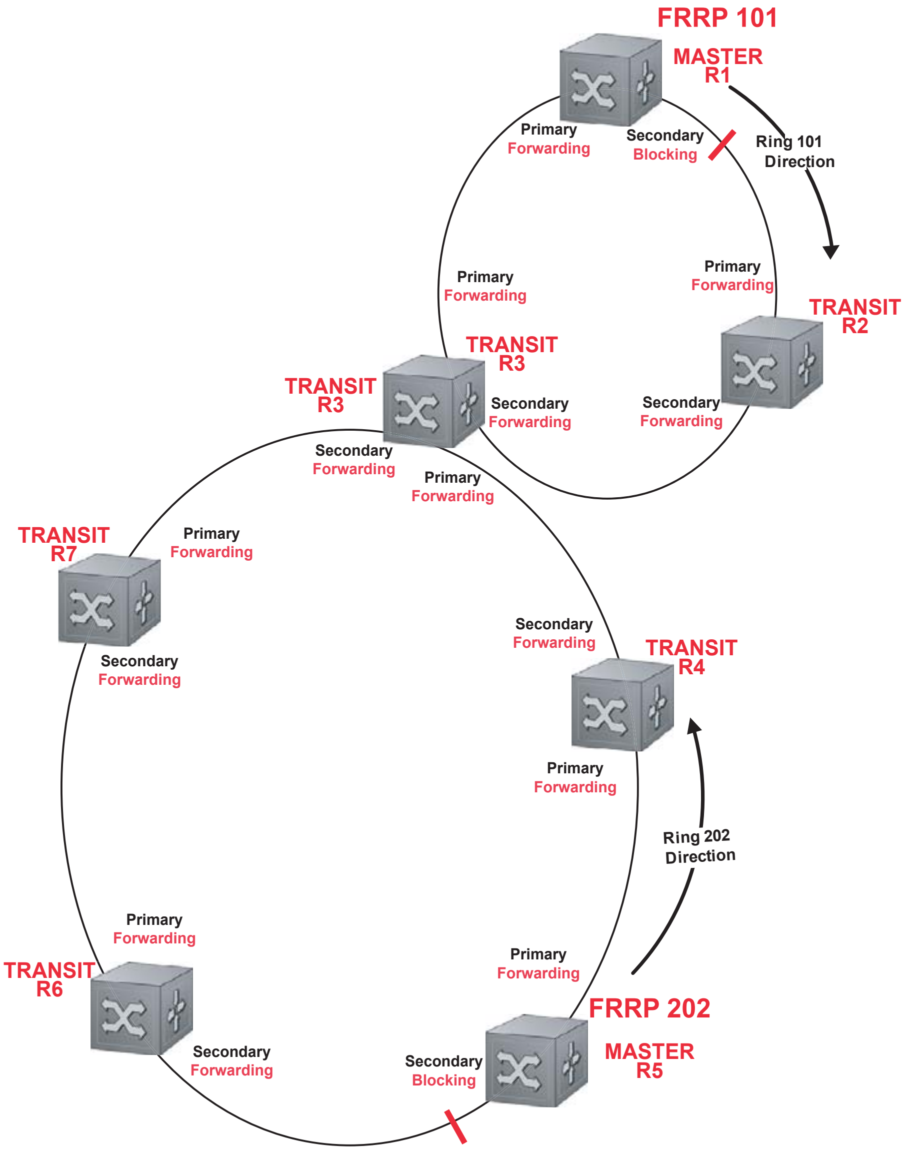
Up to 255 rings allowed per system and multiple rings can be run on one system. However, it is not recommended on the S-Series to have more than 34 rings on the same interface (either a physical interface or a portchannel). More than the recommended number of rings may cause interface instability. Multiple rings can be configured 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 example shown in Figure 12-2, 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 12-2. Example of Multiple Rings Connected by Single Switch



Important FRRP Points

FRRP provides a convergence time that can generally range between 150ms and 1500ms for Layer 2 networks. The master node originates a high-speed frame that circulates around the ring. This frame, appropriately, sets up or breaks down the ring.

- Ring Status Check Frames are transmitted by the Master Node at specified intervals
- Multiple physical rings can be run on the same switch
- One Master node per ring—all other nodes are Transit
- Each node has 2 member interfaces—Primary, Secondary
- No limit to the number of nodes on a ring
- Master node ring port states—blocking, pre-forwarding, forwarding, disabled

- Transit node ring port states—blocking, pre-forwarding, forwarding, 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)
        - — Transmitted and processed by Master node only
    - Topology Change RHF
        - — Triggered updates
        - — Processed at all nodes

Important FRRP Concepts

Table 12-1 lists some important FRRP concepts.

Table 12-1. FRRP Components

Concept	Explanation
Ring ID	Each *ring* has a unique 8-bit ring ID through which the ring is identified (e.g. FRRP 101 and FRRP 202 as shown in Figure 12-2.
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 Ring Health Frames, 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 Ring Health Frames (RHF). The Master node Secondary port receives the RHF frames. 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. • **Forwarding State**—Both ring control and data traffic is passed. When the ring is in Normal operation, the Primary port on the Master node and both Primary and Secondary ports on the Transit nodes are in forwarding state. When the ring is broken, all ring ports are in this state. • **Disabled State**—When the port is disabled or down, or is not on the VLAN.

Table 12-1. FRRP Components

Concept	Explanation
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 3 times the Hello interval rate. The dead interval is configurable in 50 ms increments from 50 ms to 6000 ms.
Ring Status	The state of the FRRP ring. During initialization/configuration, the default ring status is Ring-down (disabled). The Primary and Secondary interfaces, Control VLAN, and Master and Transit node information must be configured for the ring to be up. • **Ring-Up**: Ring is up and operational • **Ring-Down**: Ring is broken or not set up
Ring Health-check Frame (RHF)	Two types of RHFs are generated by the Master node. 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 the 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. It 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 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 Force10 proprietary protocol that does not interoperate with any other vendor.
- Spanning Tree must be disabled on both Primary and Secondary interfaces before FRRP is enabled.
- All ring ports must be Layer 2 ports. This is required for both Master and Transit nodes.
- A VLAN configured as control VLAN for a ring cannot be configured as a control or member VLAN for any other ring.
- The Control VLAN is 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.

- Create the FRRP group
- Configure the Control VLAN
  - Configure Primary and Secondary ports
- Configure and add the Member VLANs
  - Configure Primary and Secondary ports
- Configure the Master node
- Configure a Transit node
- Set FRRP Timers (optional)
- Enable FRRP

Other FRRP related commands are:

- Clear FRRP counters

Create the FRRP group

The FRRP group must be created on each switch in the ring.

Use the commands in the following sequence to create the FRRP group.

Command Syntax	Command Mode	Purpose
protocol frrp *ring-id*	CONFIGURATION	Create the FRRP group with this Ring ID Ring ID: 1-255

Configure 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 complete information about configuring VLANS in Layer 2 mode, see Chapter 20, Layer 2.

Be sure to follow these guidelines:

- All VLANS must be in Layer 2 mode.
- Only ring nodes can be added to the VLAN.
- A Control VLAN can belong to one FRRP group only.
- Control VLAN ports must be tagged.
- All ports on the ring must use the same VLAN ID for the Control VLAN.
- A VLAN cannot be configured 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

Use the commands in the following sequence, on the switch that will act as the Master node, to create the Control VLAN for this FRRP group.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Create a VLAN with this ID number VLAN ID: 1-4094
2	**tagged** *interface slot/ port {range}*	CONFIG-INT-VLAN	Tag the specified interface or range of interfaces to this VLAN. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword GigabitEthernet followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/ port information • For a 10 Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information. Slot/Port, Range: Slot and Port ID for the interface. Range is entered Slot/Port-Port.
3	**interface primary** *int slot/port* **secondary** *int slot/port* **control-vlan** *vlan id*	CONFIG-FRRP	Assign the Primary and Secondary ports, and the Control VLAN for the ports on the ring. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword GigabitEthernet followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/ port information • For a SONET interface, enter the keyword sonet followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information. Slot/Port: Slot and Port ID for the interface. VLAN ID: The VLAN identification of the Control VLAN.
4	**mode** *master*	CONFIG-FRRP	Configure the Master node
5	**member-vlan** *vlan-id {range}*	CONFIG-FRRP	Identify the Member VLANs for this FRRP group VLAN-ID, Range: VLAN IDs for the ring’s Member VLANS.
6	**no disable**	CONFIG-FRRP	Enable FRRP

Configure and add 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 complete information about configuring VLANS in Layer 2 mode, see Chapter 20, Layer 2.

Be sure to follow these guidelines:

- All VLANS must be in Layer 2 mode.
- Control VLAN ports must be tagged. 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.

Use the commands in the following sequence, on all of the Transit switches in the ring, to create the Members VLANs for this FRRP group.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Create a VLAN with this ID number VLAN ID: 1-4094
2	**tagged** *interface slot/ port {range}*	CONFIG-INT-VLAN	Tag the specified interface or range of interfaces to this VLAN. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword GigabitEthernet followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/ port information • For a SONET interface, enter the keyword sonet followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information. Slot/Port, Range: Slot and Port ID for the interface. Range is entered Slot/Port-Port.
3	**interface primary** *int slot/port* **secondary** *int slot/port* **control-vlan** *vlan id*	CONFIG-FRRP	Assign the Primary and Secondary ports, and the Control VLAN for the ports on the ring. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword GigabitEthernet followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/ port information • For a SONET interface, enter the keyword sonet followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information. Slot/Port: Slot and Port ID for the interface. VLAN ID: Identification number of the Control VLAN
4	**mode** *transit*	CONFIG-FRRP	Configure a Transit node
5	**member-vlan** *vlan-id {range}*	CONFIG-FRRP	Identify the Member VLANs for this FRRP group VLAN-ID, Range: VLAN IDs for the ring’s Member VLANs.

Step	Command Syntax	Command Mode	Purpose
6	**no disable**	CONFIG-FRRP	Enable this FRRP group on this switch.

Set FRRP Timers

Step	Command Syntax	Command Mode	Purpose
1	**timer** *{hello-interval|dead-interval} milliseconds*	CONFIG-FRRP	Enter the desired intervals for Hello-Interval or Dead-Interval times. Hello-Interval: 50-2000, in increments of 50 (default is 500) Dead-Interval: 50-6000, in increments of 50 (default is 1500)
		The Dead-Interval time should be set at 3x the Hello-Interval.	

Clear FRRP counters

Use one of the following commands to clear the FRRP counters.

Command Syntax	Command Mode	Purpose
clear frrp *ring-id*	EXEC PRIVELEGED	Clear the counters associated with this Ring ID Ring ID: 1-255
clear frrp	EXEC PRIVELEGED	Clear the counters associated with all FRRP groups

Show FRRP configuration

Use the following command to view the configuration for the FRRP group.

Command Syntax	Command Mode	Purpose
show configuration	CONFIG-FRRP	Show the configuration for this FRRP group

Show FRRP information

Use one of the following commands show general FRRP information.

Command Syntax	Command Mode	Purpose
show frrp *ring-id*	EXEC *or* EXEC PRIVELEGED	Show the information for the identified FRRP group. Ring ID: 1-255

Command Syntax	Command Mode	Purpose
show frrp summary	EXEC *or* EXEC PRIVELEGED	Show the state of all FRRP groups. Ring ID: 1-255

Troubleshooting FRRP

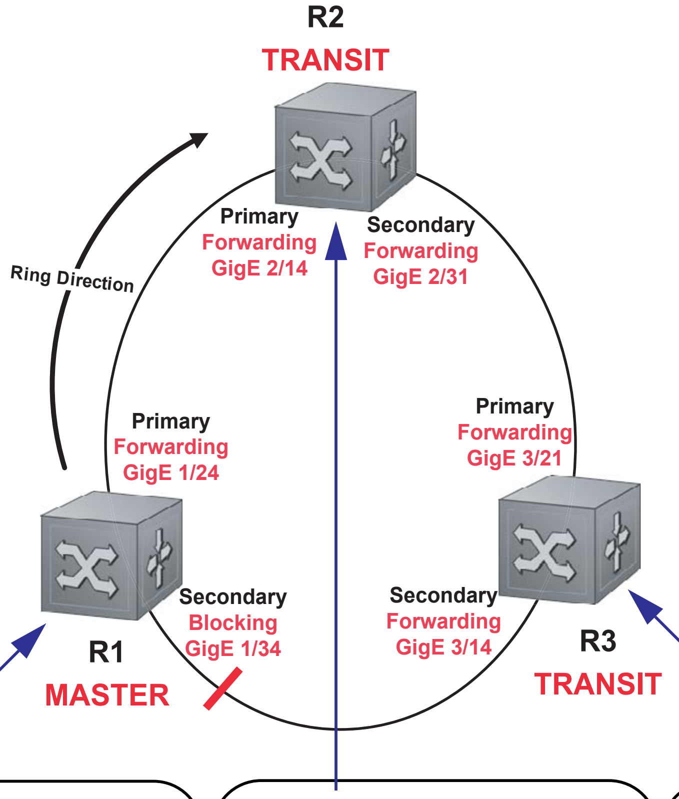
Configuration Checks

- 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.
- FRRP can be configured on Layer 2 interfaces only
- Spanning Tree (if enabled globally) must be disabled on both Primary and Secondary interfaces when FRRP is enabled.
  - When the interface ceases to be a part of any FRRP process, if Spanning Tree is enabled globally, it must be enabled explicitly for the interface.
- The maximum number of rings allowed on a chassis is 255.

Sample Configuration and Topology

Figure 12-3 is an example of a basic FRRP topology. Below the figure are the associated CLI commands.

Figure 12-3. Basic Topology and CLI commands



```
R1 MASTER
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
```

```
R2 TRANSIT
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
```

```
R3 TRANSIT
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
```

13

GARP VLAN Registration Protocol

GARP VLAN Registration Protocol is supported on platform C E S

Protocol Overview

Typical VLAN implementation involves manually configuring each Layer 2 switch that participates in a given VLAN. GARP VLAN Registration Protocol (GVRP), defined by the IEEE 802.1q specification, is a Layer 2 network protocol that provides for automatic VLAN configuration of switches. GVRP-compliant switches use GARP to register and de-register attribute values, such as VLAN IDs, with each other.

GVRP exchanges network VLAN information to allow switches to dynamically forward frames for one or more VLANs. Consequently, GVRP spreads this information and configures the needed VLAN(s) 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; you must 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. Use the **show gvrp statistics** {**interface** *interface* | **summary**} command to display status.
- On the E-Series, C-Series, and non-S60/S55 S-Series, Per-VLAN Spanning Tree (PVST+) or MSTP and GVRP cannot be enabled at the same time, as shown in Figure 13-1. If Spanning Tree and GVRP are both required, implement RSTP. The S60 and the S55 systems do support enabling GVRP and MSTP at the same time.

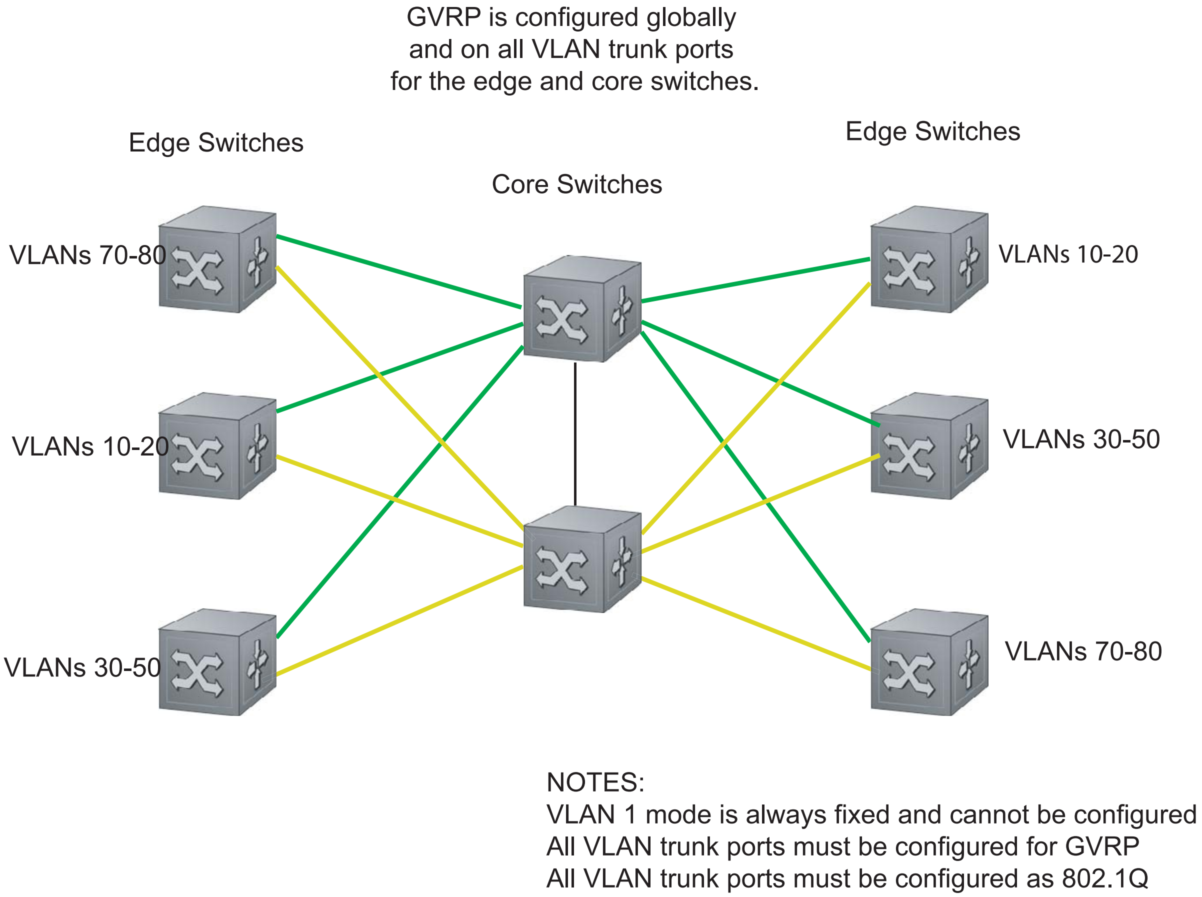
Figure 13-1. GVRP Compatibility Error Message

```
Force10(conf)#protocol spanning-tree pvst
Force10(conf-pvst)#no disable
% Error: GVRP running. Cannot enable PVST.

.........
Force10(conf)#protocol spanning-tree mstp
Force10(conf-mstp)#no disable
% Error: GVRP running. Cannot enable MSTP.

.........

Force10(conf)#protocol gvrp
Force10(conf-gvrp)#no disable
% Error: PVST running. Cannot enable GVRP.
% Error: MSTP running. Cannot enable GVRP.
```

Configuring GVRP

Globally, enable GVRP on each switch to facilitate GVRP communications. 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 Figure 13-2, that kind of port is referred to as a VLAN trunk port, but it is not necessary to specifically identify to FTOS that the port is a trunk port.

Figure 13-2. GVRP Configuration Overview



GVRP is configured globally and on all VLAN trunk ports for the edge and core switches.
Edge Switches
Core Switches
Edge Switches
VLANs 70-80
VLANs 10-20
VLANs 30-50
VLANs 10-20
VLANs 30-50
VLANs 70-80
NOTES:
VLAN 1 mode is always fixed and cannot be configured
All VLAN trunk ports must be configured for GVRP
All VLAN trunk ports must be configured as 802.1Q

Basic GVRP configuration is a 2-step process:

1. Enable GVRP globally. See page 268.
2. Enable GVRP on an interface. See page 268.

Related Configuration Tasks

- Configuring GVRP Registration on page 268
- Configuring a GARP Timer on page 269

Enabling GVRP Globally

Enable GVRP for the entire switch using the command **gvrp enable** in CONFIGURATION mode, as shown in Figure 13-3. Use the **show gvrp brief** command to inspect the global configuration.

Figure 13-3. Enabling GVRP Globally

```
Force10(conf)#protocol gvrp
Force10(config-gvrp)#no disable
Force10(config-gvrp)#show config
!
protocol gvrp
 no disable
Force10(config-gvrp)#
```

Enabling GVRP on a Layer 2 Interface

Enable GVRP on a Layer 2 interface using the command **gvrp enable** in INTERFACE mode, as shown in Figure 13-4. Use **show config** from the INTERFACE mode to inspect the interface configuration, as shown in Figure 13-4, or use the **show gvrp** *interface* command in EXEC or EXEC Privilege mode.

Figure 13-4. Enabling GVRP on a Layer 2 Interface

```
Force10(conf-if-gi-1/21)#switchport
Force10(conf-if-gi-1/21)#gvrp enable
Force10(conf-if-gi-1/21)#no shutdown
Force10(conf-if-gi-1/21)#show config
!
interface GigabitEthernet 1/21
 no ip address
 switchport
 gvrp enable
 no shutdown
```

Configuring GVRP Registration

- **Fixed Registration Mode**: Configuring a port in fixed registration mode allows for manual creation and registration of VLANs, prevents VLAN de-registration, 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 un-configured when it receives a Leave PDU. So, 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. So, set the interface to the registration mode of FORBIDDEN if you do not want the interface to advertise or learn about particular VLANS.

Based on the configuration in the example shown in Figure 13-5, the interface 1/21 will not be removed from VLAN 34 or VLAN 35 despite receiving a GVRP Leave message. Additionally, the interface will not be dynamically added to VLAN 45 or VLAN 46, even if a GVRP Join message is received.

Figure 13-5. Configuring GVRP Registration

```
Force10(conf-if-gi-1/21)#gvrp registration fixed 34,35
Force10(conf-if-gi-1/21)#gvrp registration forbidden  45,46
Force10(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
Force10(conf-if-gi-1/21)#
```

Configuring a GARP Timer

GARP timers must be set to the same values on all devices that are exchanging information using GVRP:

- **Join**: A GARP device reliably transmits Join messages to other devices by sending each Join message two times. Use this parameter to define the interval between the two sending operations of each Join message. The FTOS default is 200ms.
- **Leave**: When a GARP device expects to de-register a piece of attribute information, it will send out a Leave message and start 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 FTOS default is 600ms.
- **LeaveAll**: Upon startup, a GARP device globally starts a LeaveAll timer. Upon expiration of this interval, it will send 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 FTOS default is 10000ms.

Figure 13-6. Configuring GVRP Registration

```
Force10(conf)#garp timer leav 1000
Force10(conf)#garp timers leave-all 5000
Force10(conf)#garp timer join 300

Verification:

Force10(conf)#do show garp timer
GARP Timers     Value (milliseconds)
----------------------------------------
Join Timer        300
Leave Timer       1000
LeaveAll Timer    5000
Force10(conf)#
```

FTOS displays Message 1 if an attempt is made to configure an invalid GARP timer.

Message 1 GARP Timer Error

```
Force10(conf)#garp timers join 300
% Error: Leave timer should be >= 3*Join timer.
```

14

Internet Group Management Protocol

Table 14-1. FTOS Support for IGMP and IGMP Snooping

Feature	Platform
IGMP version 1, 2, and 3	C E S
IGMP Snooping version 2	C E S
IGMP Snooping version 3	C E S

Multicast is premised on identifying many hosts by a single destination IP address; hosts represented by the same IP address are a *multicast group*. 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 PIM) use the information in IGMP messages to discover which groups are active and to populate the multicast routing table.

IGMP Implementation Information

- FTOS supports IGMP versions 1, 2, and 3 based on RFCs 1112, 2236, and 3376, respectively.
- FTOS does not support IGMP version 3 and versions 1 or 2 on the same subnet.
- IGMP on FTOS supports up to 512 interfaces on E-Series, 31 interfaces on C-Series and S-Series, and an unlimited number of groups on all platforms.

> **Note:** The S55 supports up to 95 interfaces.

- Dell Force10 systems cannot serve as an IGMP host or an IGMP version 1 IGMP Querier.
- FTOS automatically enables IGMP on interfaces on which you enable a multicast routing protocol.

IGMP Protocol Overview

IGMP has three versions. Version 3 obsoletes and is backwards-compatible with version 2; version 2 obsoletes version 1.

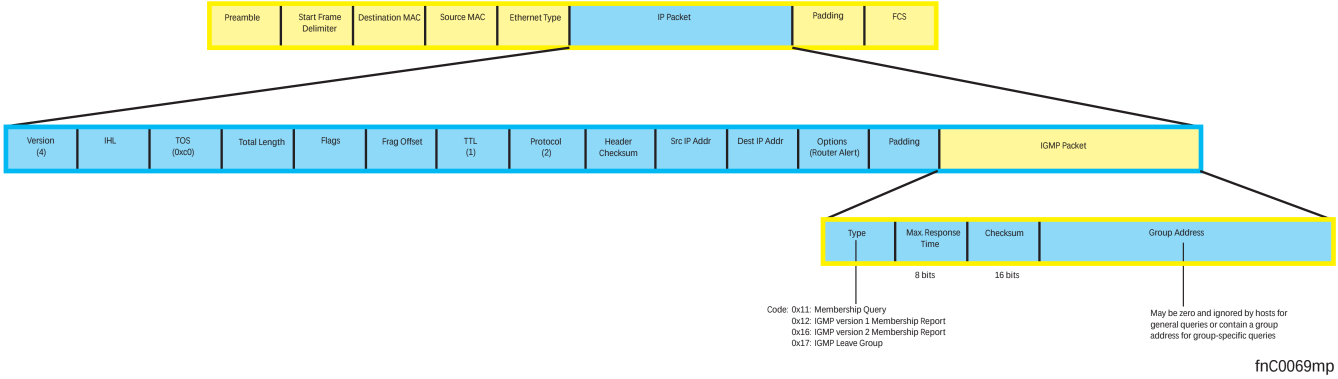
IGMP version 2

IGMP version 2 improves 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, as shown in Figure 14-1.

Figure 14-1. 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

1. One router on a subnet is elected as the querier. The querier periodically multicasts (to all-multicast-systems address 224.0.0.1) a general query to all hosts on the subnet.
2. A host that wants to join a multicast group responds with an IGMP Membership Report that contains the multicast address of the group it wants to join (the packet is addressed to the same group). If multiple hosts want to join the same multicast group, only the report from the first host to respond reaches the querier, and the remaining hosts suppress their responses (see Adjusting Query and Response Timers on page 278 for how the delay timer mechanism works).
3. The querier receives the report for a group and adds the group to the list of multicast groups associated with its outgoing port to the subnet. Multicast traffic for the group is then forwarded to that subnet.

Sending an Unsolicited IGMP Report

A host does not have to wait for a general query to join a group. It may send an unsolicited IGMP Membership Report, also called an IGMP Join message, to the querier.

Leaving a Multicast Group

1. A host sends a membership report of type 0x17 (IGMP Leave message) to the all routers multicast address 224.0.0.2 when it no longer cares about multicast traffic for a particular group.
2. The querier sends a Group-Specific Query to determine whether there are any remaining hosts in the group. There must be at least one receiver in a group on a subnet for a router to forward multicast traffic for that group to the subnet.
3. Any remaining hosts respond to the query according to the delay timer mechanism (see Adjusting Query and Response Timers on page 278). 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. There are differences:

- Version 3 adds the ability to filter by multicast source, which helps multicast routing protocols avoid forwarding traffic to subnets where there are no interested receivers.
- To enable filtering, routers must keep track of more state information, that is, the list of sources that must be filtered. An additional query type, the Group-and-Source-Specific Query, keeps track of state changes, while the Group-Specific and General queries still refresh existing state.
- Reporting is more efficient and robust: hosts do not suppress query responses (non-suppression helps track state and enables the immediate-leave and IGMP Snooping features), state-change reports are retransmitted to insure delivery, and a single membership report bundles multiple statements from a single host, rather than sending an individual packet for each statement.

The version 3 packet structure is different from version 2 to accommodate these protocol enhancements. Queries (Figure 14-2) are still sent to the all-systems address 224.0.0.1, but reports (Figure 14-3) are sent to the all IGMP version 3-capable multicast routers address 244.0.0.22.

Figure 14-2. IGMP version 3 Membership Query Packet Format

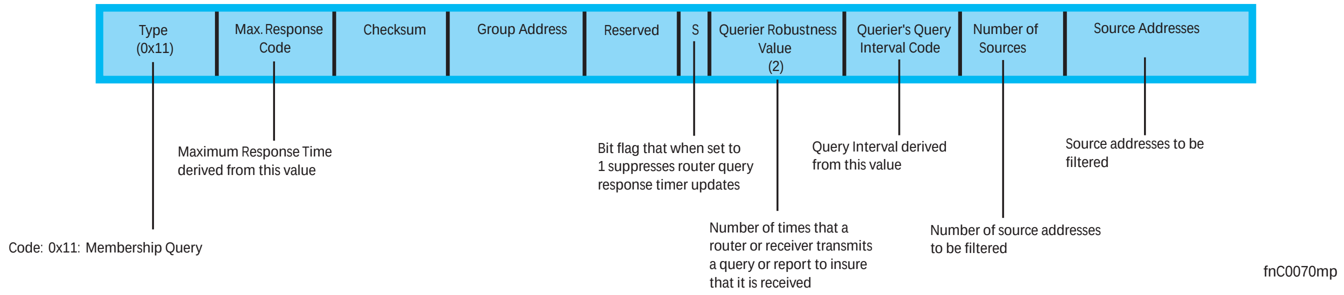
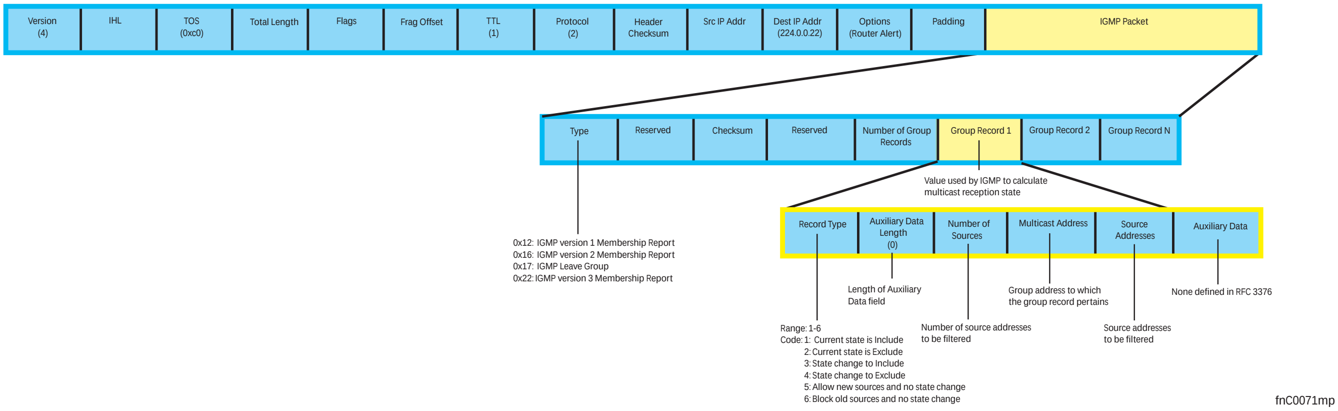

Type (0x11)
Max. Response Code
Checksum
Group Address
Reserved
S
Querier Robustness Value (2)
Querier's Query Interval Code
Number of Sources
Source Addresses
Maximum Response Time derived from this value
Bit flag that when set to 1 suppresses router query response timer updates
Query Interval derived from this value
Source addresses to be filtered
Code: 0x11: Membership Query
Number of times that a router or receiver transmits a query or report to insure that it is received
Number of source addresses to be filtered
fnC0070mp

Figure 14-3. IGMP version 3 Membership Report Packet Format

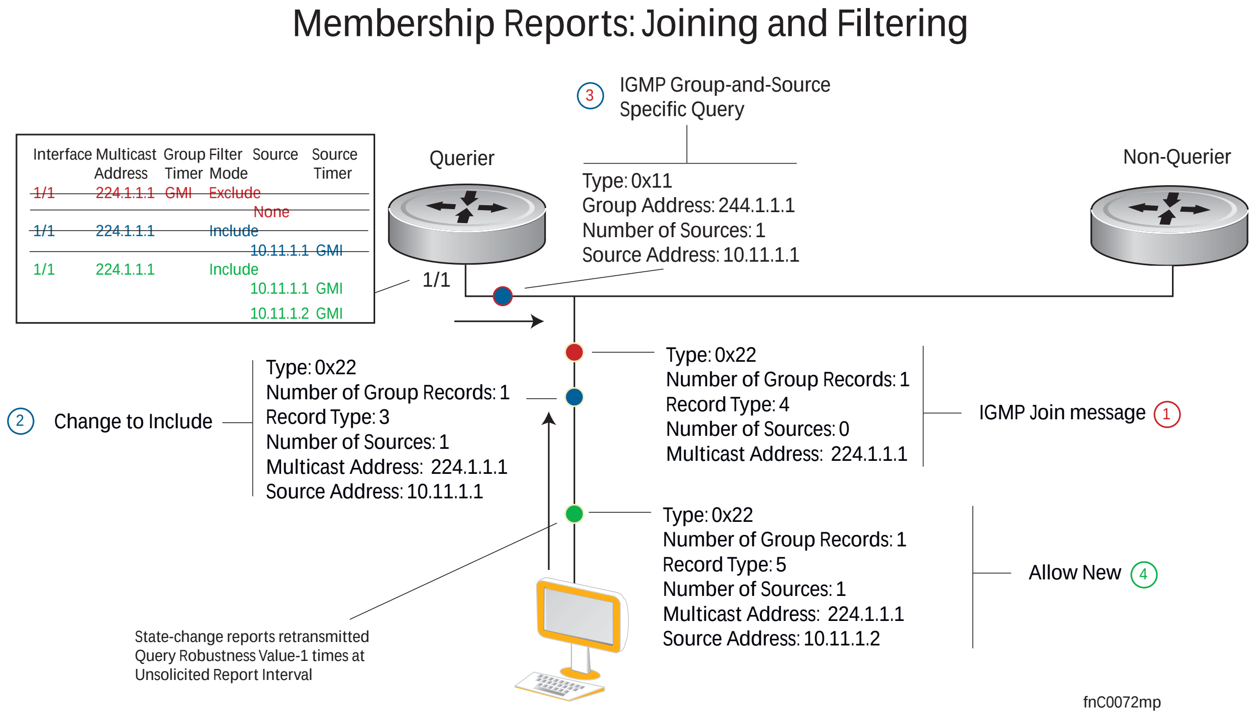

Version (4)
IHL
TOS (0xc0)
Total Length
Flags
Frag Offset
TTL (1)
Protocol (2)
Header Checksum
Src IP Addr
Dest IP Addr (224.0.0.22)
Options (Router Alert)
Padding
IGMP Packet
Type
Reserved
Checksum
Reserved
Number of Group Records
Group Record 1
Group Record 2
Group Record N
Value used by IGMP to calculate multicast reception state
0x12: IGMP version 1 Membership Report
0x16: IGMP version 2 Membership Report
0x17: IGMP Leave Group
0x22: IGMP version 3 Membership Report
Record Type
Auxiliary Data Length (0)
Number of Sources
Multicast Address
Source Addresses
Auxiliary Data
Length of Auxiliary Data field
Group address to which the group record pertains
None defined in RFC 3376
Range: 1-6
Code: 1: Current state is Include
2: Current state is Exclude
3: State change to Include
4: State change to Exclude
5: Allow new sources and no state change
6: Block old sources and no state change
Number of source addresses to be filtered
Source addresses to be filtered
fnC0071mp

Joining and Filtering Groups and Sources

Figure 14-4 shows how multicast routers maintain the group and source information from unsolicited reports.

1. The first unsolicited report from the host indicates that it wants to receive traffic for group 224.1.1.1.
2. The host’s second report indicates that it is only interested in traffic from group 224.1.1.1, source 10.11.1.1. Include messages 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, then the router cannot record the include request. There are no other interested hosts, so the request is recorded. At this point, the multicast routing protocol prunes the tree to all but the specified sources.
3. The host’s third message indicates that it is only interested in traffic from sources 10.11.1.1 and 10.11.1.2. Since 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 14-4. IGMP Membership Reports: Joining and Filtering


Membership Reports: Joining and Filtering
3
IGMP Group-and-Source Specific Query
Querier
Non-Querier
1/1
2
Change to Include
IGMP Join message
1
Allow New
4
fnC0072mp

Interface	Multicast Address	Group Timer	Filter Mode	Source	Source Timer
~~1/1~~	~~224.1.1.1~~	~~GMI~~	~~Exclude~~		
				~~None~~	
~~1/1~~	~~224.1.1.1~~		~~Include~~		
				~~10.11.1.1~~	~~GMI~~
1/1	224.1.1.1		Include		
				10.11.1.1	GMI
				10.11.1.2	GMI

Type: 0x11
Group Address: 244.1.1.1
Number of Sources: 1
Source Address: 10.11.1.1

Type: 0x22
Number of Group Records: 1
Record Type: 4
Number of Sources: 0
Multicast Address: 224.1.1.1

Type: 0x22
Number of Group Records: 1
Record Type: 3
Number of Sources: 1
Multicast Address: 224.1.1.1
Source Address: 10.11.1.1

Type: 0x22
Number of Group Records: 1
Record Type: 5
Number of Sources: 1
Multicast Address: 224.1.1.1
Source Address: 10.11.1.2

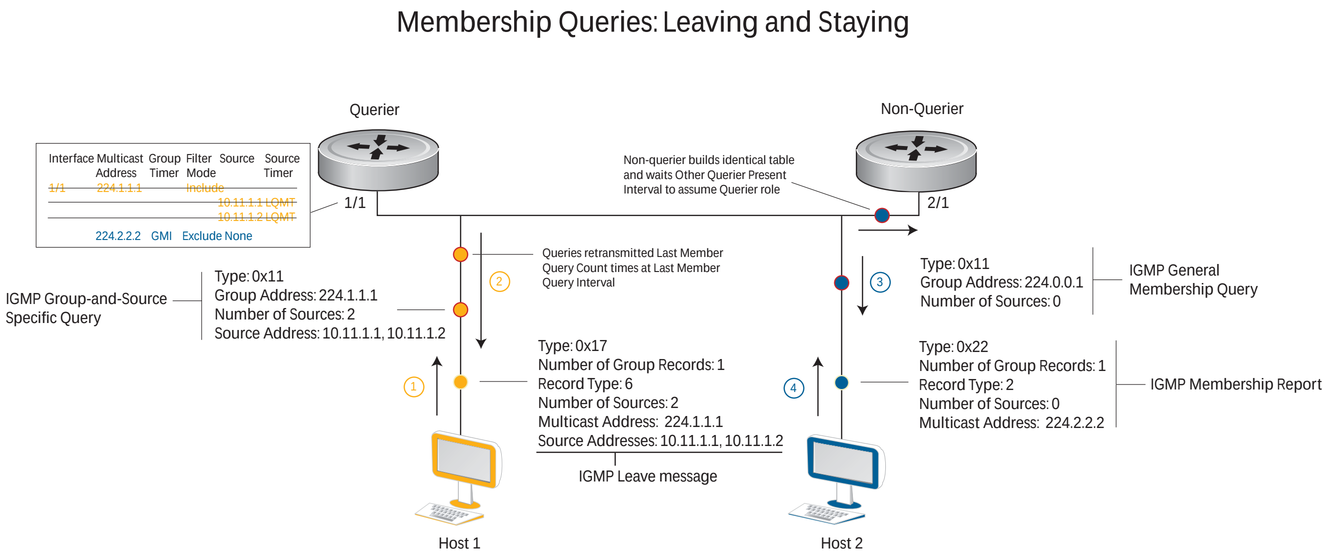
State-change reports retransmitted Query Robustness Value-1 times at Unsolicited Report Interval

Leaving and Staying in Groups

Figure 14-5 shows how multicast routers track and refresh state changes in response to group-and-specific and general queries.

1. Host 1 sends a message indicating it is leaving group 224.1.1.1 and that the include filter for 10.11.1.1 and 10.11.1.2 are no longer necessary.
2. The querier, before making any state changes, sends a group-and-source query to see if any other host is interested in these two sources; queries for state-changes are retransmitted multiple times. If any are, they respond with their current state information and the querier refreshes the relevant state information.
3. Separately in Figure 14-5, the querier sends a general query to 224.0.0.1.
4. Host 2 responds to the periodic general query so the querier refreshes the state information for that group.

Figure 14-5. IGMP Membership Queries: Leaving and Staying in Groups


Membership Queries: Leaving and Staying
Querier
Non-Querier
Interface Multicast Address Group Timer Filter Mode Source Source Timer
~~1/1 224.1.1.1 Include~~
~~10.11.1.1 LQMT~~
~~10.11.1.2 LQMT~~
224.2.2.2 GMI Exclude None
1/1
2/1
Non-querier builds identical table and waits Other Querier Present Interval to assume Querier role
Queries retransmitted Last Member Query Count times at Last Member Query Interval
IGMP Group-and-Source Specific Query
Type: 0x11
Group Address: 224.1.1.1
Number of Sources: 2
Source Address: 10.11.1.1, 10.11.1.2
1
2
3
4
Type: 0x17
Number of Group Records: 1
Record Type: 6
Number of Sources: 2
Multicast Address: 224.1.1.1
Source Addresses: 10.11.1.1, 10.11.1.2
IGMP Leave message
Type: 0x11
Group Address: 224.0.0.1
Number of Sources: 0
IGMP General Membership Query
Type: 0x22
Number of Group Records: 1
Record Type: 2
Number of Sources: 0
Multicast Address: 224.2.2.2
IGMP Membership Report
Host 1
Host 2

Configuring IGMP

Configuring IGMP is a two-step process:

1. Enable multicast routing using the command **ip multicast-routing**.
2. Enable a multicast routing protocol.

Related Configuration Tasks

- Viewing IGMP Enabled Interfaces on page 276
- Selecting an IGMP Version on page 277
- Viewing IGMP Groups on page 277
- Adjusting Timers on page 278
- Configuring a Static IGMP Group on page 279
- Prevent a Host from Joining a Group on page 447
- Enabling IGMP Immediate-leave on page 279
- IGMP Snooping on page 280
- Fast Convergence after MSTP Topology Changes on page 282
- Designating a Multicast Router Interface on page 282

Viewing IGMP Enabled Interfaces

Interfaces that are enabled with PIM-SM are automatically enabled with IGMP. View IGMP-enabled interfaces using the command **show ip igmp interface** command in the EXEC Privilege mode.

Figure 14-6. Viewing IGMP-enabled Interfaces

```
Force10#show ip igmp interface gig 7/16
GigabitEthernet 7/16 is up, line protocol is up
  Internet address is 10.87.3.2/24
  IGMP is enabled on interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 300 seconds
  IGMP max query response time is 10 seconds
  Last member query response interval is 199 ms
  IGMP activity: 0 joins, 0 leaves
  IGMP querying router is 10.87.3.2 (this system)
  IGMP version is 2
Force10#
```

Selecting an IGMP Version

FTOS enables IGMP version 2 by default, which supports version 1 and 2 hosts, but is not compatible with version 3 on the same subnet. If hosts require IGMP version 3, you can switch to IGMP version 3 using the command **ip igmp version** from INTERFACE mode, as shown in Figure 14-7.

Figure 14-7. Selecting an IGMP Version

```
Force10(conf-if-gi-1/13)#ip igmp version 3
Force10(conf-if-gi-1/13)#do show ip igmp interface
GigabitEthernet 1/13 is up, line protocol is down
  Inbound IGMP access group is not set
  Interface IGMP group join rate limit is not set
  Internet address is 1.1.1.1/24
  IGMP is enabled on interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 125 seconds
  IGMP max query response time is 10 seconds
  IGMP last member query response interval is 1000 ms
  IGMP immediate-leave is disabled
  IGMP activity: 0 joins, 0 leaves, 0 channel joins, 0 channel leaves
  IGMP querying router is 1.1.1.1 (this system)
  IGMP version is 3
Force10(conf-if-gi-1/13)#
```

Viewing IGMP Groups

View both learned and statically configured IGMP groups using the command **show ip igmp groups** from EXEC Privilege mode.

Figure 14-8. Viewing Static and Learned IGMP Groups

```
Force10(conf-if-gi-1/0)#do sho ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address    Interface                Uptime    Expires   Last Reporter
224.1.1.1        GigabitEthernet 1/0      00:00:03  Never     CLI
224.1.2.1        GigabitEthernet 1/0      00:56:55  00:01:22  1.1.1.2
```

Adjusting Timers

View the current value of all IGMP timers using the command **show ip igmp interface** from EXEC Privilege mode, as shown in Figure 14-6.

Adjusting Query and Response Timers

The querier periodically sends a general query to discover which multicast groups are active. A group must have at least one host to be active. When a host receives a query, it does not respond immediately, but rather starts a delay timer. The delay time is set to a random value between 0 and the Maximum Response Time. The host sends a response when the timer expires; in version 2, if another host responds before the timer expires, the timer is nullified, and no response is sent.

The Maximum Response Time is the amount of time that the querier waits for a response to a query before taking further action. The querier advertises this value in the query (see Figure 14-1). Lowering this value decreases leave latency but increases response burstiness since all host membership reports must be sent before the Maximum Response Time expires. Inversely, increasing this value decreases burstiness at the expense of leave latency.

- Adjust the period between queries using the command **ip igmp query-interval** from INTERFACE mode.
- Adjust the Maximum Response Time using the command **ip igmp query-max-resp-time** from INTERFACE mode.

When the querier receives a leave message from a host, it sends a group-specific query to the subnet. If no response is received, it sends another. The amount of time that the querier waits to receive a response to the initial query before sending a second one is the Last Member Query Interval (LMQI). The switch waits one LMQI after the second query before removing the group from the state table.

- Adjust the Last Member Query Interval using the command **ip igmp last-member-query-interval** from INTERFACE mode.

Adjusting the IGMP Querier Timeout Value

If there is more than one multicast router on a subnet, only one is elected to be the querier, which is the router that sends queries to the subnet.

1. Routers send queries to the all multicast systems address, 224.0.0.1. Initially, all routers send queries.

2. When a router receives a query it compares the IP address of the interface on which it was received with the source IP address given in the query. If the receiving router IP address is greater than the source address given in the query, the router stops sending queries. By this method, the router with the lowest IP address on the subnet is elected querier and continues to send queries.
3. If a specified amount of time elapses during which other routers on the subnet do not receive a query, those routers assume that the querier is down, and a new querier is elected.

The amount of time that elapses before routers on a subnet assume that the querier is down is the Other Querier Present Interval. Adjust this value using the command **ip igmp querier-timeout** from INTERFACE mode.

Configuring a Static IGMP Group

Configure a static IGMP group using the command **ip igmp static-group**. Multicast traffic for static groups is always forwarded to the subnet even if there are no members in the group.

View the static groups using the command **show ip igmp groups** from EXEC Privilege mode. Static groups have an expiration value of *Never* and a Last Reporter value of *CLI*, as shown in Figure 14-8.

Enabling IGMP Immediate-leave

If the querier does not receive a response to a group-specific or group-and-source query, it sends another (Querier Robustness Value). Then, after no response, it removes the group from the outgoing interface for the subnet.

IGMP Immediate Leave reduces leave latency by enabling a router to immediately delete the group membership on an interface upon receiving a Leave message (it does not send any group-specific or group-and-source queries before deleting the entry). Configure the system for IGMP Immediate Leave using the command **ip igmp immediate-leave**.

View the enable status of this feature using the command **show ip igmp interface** from EXEC Privilege mode, as shown in Figure 14-7.

IGMP Snooping

Multicast packets are addressed with multicast MAC addresses, which represent a group of devices, rather than one unique device. Switches forward multicast frames out of all ports in a VLAN by default, even though there may be only some interested hosts, which is a waste of bandwidth. IGMP Snooping enables switches to use information in IGMP packets to generate a forwarding table that associates ports with multicast groups so that when they receive multicast frames, they can forward them only to interested receivers.

IGMP Snooping Implementation Information

- IGMP Snooping on FTOS uses IP multicast addresses not MAC addresses.
- IGMP Snooping is not supported on stacked VLANs.
- IGMP Snooping is supported on all S-Series stack members,.
- IGMP Snooping reacts to STP and MSTP topology changes by sending a general query on the interface that transitions to the forwarding state.

Configuring IGMP Snooping

Configuring IGMP Snooping is a one-step process. That is, enable it on a switch using the command **ip igmp snooping enable** from CONFIGURATION mode. View the configuration using the command **show running-config** from CONFIGURATION mode, as shown in Figure 14-9. You can disable snooping on for a VLAN using the command **no ip igmp snooping** from INTERFACE VLAN mode.

Figure 14-9. Enabling IGMP Snooping

```
Force10(conf)#ip igmp snooping enable
Force10(conf)#do show running-config igmp
ip igmp snooping enable
Force10(conf)#
```

Related Configuration Tasks

- Enabling IGMP Immediate-leave on page 280
- Disabling Multicast Flooding on page 281
- Specifying a Port as Connected to a Multicast Router on page 281
- Configuring the Switch as Querier on page 281

Enabling IGMP Immediate-leave

Configure the switch to remove a group-port association upon receiving an IGMP Leave message using the command **ip igmp fast-leave** from INTERFACE VLAN mode. View the configuration using the command **show config** from INTERFACE VLAN mode, as shown in Figure 14-10.

Figure 14-10. Enabling IGMP Snooping

```
Force10(conf-if-vl-100)#show config
!
interface Vlan 100
 no ip address
 ip igmp snooping fast-leave
 shutdown
Force10(conf-if-vl-100)#
```

Disabling Multicast Flooding

If the switch receives a multicast packet that has an IP address of a group it has not learned (unregistered frame), the switch floods that packet out of all ports on the VLAN.

On the E-Series, you can configure the switch to only forward unregistered packets to ports on a VLAN that are connected to multicast routers (mrouter ports) using the command **no ip igmp snooping flood** from CONFIGURATION mode. When flooding is disabled, if there are no such ports in the VLAN connected to a multicast router, the switch drops the packets.

On the C-Series and S-Series, when you configure **no ip igmp snooping flood**, the system drops the packets immediately. The system does not forward the frames on mrouter ports, even if they are present. On the C-Series and S-Series, Layer 3 multicast must be disabled (**no ip multicast-routing**) in order to disable multicast flooding.

Specifying a Port as Connected to a Multicast Router

You can statically specify a port in a VLAN as connected to a multicast router using the command **ip igmp snooping mrouter** from INTERFACE VLAN mode.

View the ports that are connected to multicast routers using the command **show ip igmp snooping mrouter** from EXEC Privilege mode.

Configuring the Switch as Querier

Hosts that do not support unsolicited reporting wait for a general query before sending a membership report. When the multicast source and receivers are in the same VLAN, multicast traffic is not routed, and so there is no querier. You must configure the switch to be the querier for a VLAN so that hosts send membership reports, and the switch can generate a forwarding table by snooping.

Configure the switch to be the querier for a VLAN by first assigning an IP address to the VLAN interface, and then using the command **ip igmp snooping querier** from INTERFACE VLAN mode.

- IGMP snooping Querier does not start if there is a statically configured multicast router interface in the VLAN.
- The switch may lose the querier election if it does not have the lowest IP address of all potential queriers on the subnet.

- When enabled, IGMP snooping Querier starts after one query interval in case no IGMP general query (with IP SA lower than its VLAN IP address) is received on any of its VLAN members.

Adjusting the Last Member Query Interval

When the querier receives a leave message from a receiver, it sends a group-specific query out of the ports specified in the forwarding table. If no response is received, it sends another. The amount of time that the querier waits to receive a response to the initial query before sending a second one is the Last Member Query Interval (LMQI). The switch waits one LMQI after the second query before removing the group-port entry from the forwarding table.

Adjust the Last Member Query Interval using the command **ip igmp snooping last-member-query-interval** from INTERFACE VLAN mode.

Fast Convergence after MSTP Topology Changes

When a port transitions to the Forwarding state as a result of an STP or MSTP topology change, FTOS sends a general query out of all ports except the multicast router ports. The host sends a response to the general query and the forwarding database is updated without having to wait for the query interval to expire.

When an IGMP snooping switch is not acting as a Querier it sends out the general query, in response to the MSTP triggered link-layer topology change, with the source IP address of 0.0.0.0 to avoid triggering Querier election.

Designating a Multicast Router Interface

You can designate an interface as a multicast router interface with the command **ip igmp snooping mrouter interface**. FTOS also has the capability of listening in on the incoming IGMP General Queries and designate those interfaces as the multicast router interface when the frames have a non-zero IP source address. All IGMP control packets and IP multicast data traffic originating from receivers is forwarded to multicast router interfaces.

15

Interfaces

This chapter describes interface types, both physical and logical, and how to configure them with FTOS.

10/100/1000 Mbps Ethernet, Gigabit Ethernet, and 10 Gigabit Ethernet interfaces are supported on platforms C E S

SONET interfaces are only supported on platform E and are covered in the SONET/SDH chapter of the E-Series *FTOS Configuration Guide*.

Basic Interface Configuration:

- Interface Types
- View Basic Interface Information
- Enable a Physical Interface
- Physical Interfaces
- Management Interfaces
- VLAN Interfaces
- Loopback Interfaces
- Null Interfaces on page 294
- Port Channel Interfaces

Advanced Interface Configuration:

- Bulk Configuration
- Interface Range Macros on page 309
- Monitor and Maintain Interfaces
- Link Debounce Timer
- Link Dampening
- Ethernet Pause Frames
- Configure MTU Size on an Interface
- Port-pipes on page 320
- Auto-Negotiation on Ethernet Interfaces
- View Advanced Interface Information

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

The user has several options for viewing interface status and configuration parameters. The **show interfaces** command in EXEC mode will list all configurable interfaces on the chassis and 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 a port channel interface is configured, the **show interfaces** command can list 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 FTOS will return to the command prompt.

Figure 15-1 displays the configuration and status information for one interface.

Figure 15-1. show interfaces Command Example

```
Force10#show interfaces tengigabitethernet 1/0
TenGigabitEthernet 1/0 is up, line protocol is up
Hardware is Force10Eth, address is 00:01:e8:05:f3:6a
    Current address is 00:01:e8:05:f3:6a
Pluggable media present, XFP type is 10GBASE-LR.
    Medium is MultiRate, Wavelength is 1310nm
    XFP receive power reading is -3.7685
Interface index is 67436603
Internet address is 65.113.24.238/28
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit, Mode full duplex, Master
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:09:54
Queueing strategy: fifo
Input Statistics:
     0 packets, 0 bytes
     0 Vlans
     0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     0 Multicasts, 0 Broadcasts
```

Use the **show ip interfaces brief** command in the EXEC Privilege mode to view which interfaces are enabled for Layer 3 data transmission. In Figure 15-2, GigabitEthernet interface 1/5 is in Layer 3 mode since an IP address has been assigned to it and the interface's status is operationally `up`.

Figure 15-2. show ip interfaces brief Command Example (Partial)

```
Force10#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
```

Use the **show interfaces configured** command in the EXEC Privilege mode to view only configured interfaces. In Figure 15-2, GigabitEthernet interface 1/5 is in Layer 3 mode since 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. (Figure 158).

Figure 15-3. Interfaces listed in the show running-config Command (Partial)

```
Force10#show running
Current Configuration ...
!
interface GigabitEthernet 7/6
 no ip address
 shutdown
!
interface GigabitEthernet 7/7
 no ip address
 shutdown
!
interface GigabitEthernet 7/8
 no ip address
 shutdown
!
interface GigabitEthernet 7/9
 no ip address
 shutdown
```

Enable a Physical Interface

After determining the type of physical interfaces available, the user may enter the INTERFACE mode by entering the command **interface** *interface slot/port* to enable and configure the interfaces.

To enter the INTERFACE mode, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface** *interface*	CONFIGURATION	Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/port information. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.
2	**no shutdown**	INTERFACE	Enter the **no shutdown** command to enable the interface. If the interface is a SONET interface, enter the **encap ppp** command to enable PPP encapsulation.

To confirm that the interface is enabled, use the **show config** command in the INTERFACE mode.

To leave the INTERFACE mode, use the **exit** command or **end** command.

The user can not delete a physical interface.

Physical Interfaces

The *Management Ethernet interface*, is a single RJ-45 Fast Ethernet port on the Route Processor Module (RPM) of the C-Series and E-Series and on each unit of the S55; it provides dedicated management access to the system. The other S-Series (non-S55) systems supported by FTOS do not have this dedicated management interface, but you can use any Ethernet port configured with an IP address and route.

Line card interfaces support Layer 2 and Layer 3 traffic over the 10/100/1000, Gigabit, and 10-Gigabit Ethernet interfaces. SONET interfaces with PPP encapsulation support Layer 3 traffic. These interfaces (except SONET interfaces with PPP encapsulation) can also become part of virtual interfaces such as VLANs or port channels.

Link detection on ExaScale line cards is interrupt-based rather than poll-based, which enables ExaScale cards to bring up and take down links faster.

For more information on VLANs, see Bulk Configuration on page 306 and for more information on port channels, see Port Channel Interfaces on page 294.

FTOS Behavior: S-Series systems use a single MAC address for all physical interfaces while E-Series and C-Series use a unique MAC address for each physical interface, though this results in no functional difference between these platforms.

Configuration Task List for Physical Interfaces

By default, all interfaces are operationally disabled and traffic will not pass through them.

The following section includes information about optional configurations for physical interfaces:

- Overview of Layer Modes on page 288
- Configure Layer 2 (Data Link) Mode on page 288
- Management Interfaces on page 290
- Auto-Negotiation on Ethernet Interfaces on page 321
- Adjust the keepalive timer on page 323
- Clear interface counters on page 327

Overview of Layer Modes

On all systems running FTOS, 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 15-1. Interfaces Types

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)
SONET (PPP encapsulation)	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)

Configure Layer 2 (Data Link) Mode

Use the **switchport** command in INTERFACE mode to enable Layer 2 data transmissions through an individual interface. The user can not configure switching or Layer 2 protocols such as spanning tree protocol on an interface unless the interface has been set to Layer 2 mode.

Figure 15-4 displays the basic configuration found in a Layer 2 interface.

Figure 15-4. show config Command Example of a Layer 2 Interface

```
Force10(conf-if-po-1)#show config
!
interface Port-channel 1
 no ip address
 switchport
 no shutdown
Force10(conf-if)#
```

To configure an interface in Layer 2 mode, use these commands in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
no shutdown	INTERFACE	Enable the interface.
switchport	INTERFACE	Place the interface in Layer 2 (switching) mode.

For information on enabling and configuring Spanning Tree Protocol, see Chapter 10, Layer 2, on page 47. To view the interfaces in Layer 2 mode, use the command **show interfaces switchport** in the EXEC mode.

Configure Layer 3 (Network) Mode

When you assign an IP address to a physical interface, you place it in Layer 3 mode. Use the **ip address** command and **no shutdown** command in INTERFACE mode to enable Layer 3 mode on an individual interface. 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).

Figure 15-5 shows how the **show config** command displays an example of a Layer 3 interface.

Figure 15-5. show config Command Example of a Layer 3 Interface

```
Force10(conf-if-gi-1/5)#show config
!
interface GigabitEthernet 1/5
 ip address 10.10.10.1 /24
 no shutdown
Force10(conf-if)#
```

If an interface is in the incorrect layer mode for a given command, an error message is displayed to the user. For example, in Figure 15-6, the command **ip address** triggered an error message because the interface is in Layer 2 mode and the **ip address** command is a Layer 3 command only.

Figure 15-6. Error Message When Trying to Add an IP Address to Layer 2 Interface

```
Force10(conf-if)#show config
!
interface GigabitEthernet 1/2
 no ip address
 switchport
 no shutdown
Force10(conf-if)#ip address 10.10.1.1 /24
% Error: Port is in Layer 2 mode Gi 1/2.
Force10(conf-if)#
```

Error message

To determine the configuration of an interface, you can use the **show config** command in INTERFACE mode or the various **show interface** commands in EXEC mode.

To assign an IP address, use both of the following commands in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
no shutdown	INTERFACE	Enable the interface.

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure a primary IP address and mask on the interface. The *ip-address* must be in dotted-decimal format (A.B.C.D) and the *mask* must be in slash format (/xx). Add the keyword **secondary** if the IP address is the interface's backup IP address.

You can only configure one (1) 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 the EXEC mode (Figure 176).

To view IP information on an interface in Layer 3 mode, use the **show ip interface** command in the EXEC Privilege mode (Figure 159).

Figure 15-7. Command Example: show ip interface

```
Force10>show ip int vlan 58
Vlan 58 is up, line protocol is up
Internet address is 1.1.49.1/24
Broadcast address is 1.1.49.255
Address determined by config file
MTU is 1554 bytes
Inbound  access list is not set
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent
```

Management Interfaces

The S55 system supports the Management Ethernet interface as well as the standard S-Series interface on any port. Either method can be used to connect to the system.

Configure Management Interfaces on the E-Series and C-Series and on the S55

On the E-Series, C-Series, and S55 the dedicated Management interface provides management access to the system. You can configure this interface with FTOS, but the configuration options on this interface are limited. Gateway addresses and IP addresses cannot be configured if it appears in the main routing table of FTOS. In addition, Proxy ARP is not supported on this interface.

Note: On the S55, a default IP address is assigned to the Management port. Use this IP address to set your laptop Ethernet port to the same network for test purposes.

To configure a Management interface, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface Managementethernet *interface*	CONFIGURATION	Enter the slot and the port (0). ON the E-Series and C-Series, dual RPMs can be in use. Slot range: C-Series, E-Series: 0-1 S55: 0

To view the Primary RPM Management port, use the **show interface Managementethernet** command in the EXEC Privilege mode. If there are 2 RPMs, the you cannot view information on that interface.

To configure IP addresses on a Management interface, use the following command in the MANAGEMENT INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ip-address mask*	INTERFACE	Configure an IP address and mask on the interface. • *ip-address mask:* enter an address in dotted-decimal format (A.B.C.D), the mask must be in /prefix format (/x)

If there are 2 RPMs on the system, each Management interface must be configured with a different IP address. Unless the **management route** command is configured, you can only access the Management interface from the local LAN. To access the Management interface from another LAN, the **management route** command must be configured to point to the Management interface.

Alternatively, you can use **virtual-ip** to manage a system with one or two RPMs. A virtual IP is an IP address assigned to the system (not to any management interfaces) and is a CONFIGURATION mode command. When a virtual IP address is assigned to the system, the active management interface of the RPM is recognized by the virtual IP address—not by the actual interface IP address assigned to it. During an RPM failover, you do not have to remember the IP address of the new RPM's management interface—the system will still recognizes the virtual-IP address.

Configure Management Interfaces on the S-Series

The user can manage the S-Series from any port. Configure an IP address for the port using the **ip address** command, and enable it using the command **no shutdown**. The user may use the command **description** from INTERFACE mode to note that the interface is the management interface. There is no separate management routing table, so the user must configure all routes in the IP routing table (the **ip route** command).

As shown in Figure 15-8, from EXEC Privilege mode, display the configuration for a given port by entering the command **show interface**, and the routing table with the **show ip route** command.

Figure 15-8. Viewing Management Routes on the S-Series

```
Force10#show int gig 0/48
GigabitEthernet 0/48 is up, line protocol is up
Description: This is the Managment Interface
Hardware is Force10Eth, address is 00:01:e8:cc:cc:ce
    Current address is 00:01:e8:cc:cc:ce
Pluggable media not present
Interface index is 46449666
Internet address is 10.11.131.240/23
[output omitted]
Force10#show ip route

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is 10.11.131.254 to network 0.0.0.0

       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
 *S    0.0.0.0/0          via 10.11.131.254, Gi 0/48           1/0        1d2h
  C    10.11.130.0/23     Direct, Gi 0/48                      0/0        1d2h
Force10#
```

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 on VLANs and Layer 2, refer to Chapter 10, Layer 2, on page 47. See also Chapter 18, VLAN Stacking, on page 367.

Note: To monitor VLAN interfaces, use the Management Information Base for Network Management of TCP/IP-based internets: MIB-II (RFC 1213). Monitoring VLAN interfaces via SNMP is supported only on E-Series.

FTOS 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 on configuring different routing protocols, refer to the chapters on the specific protocol.

A consideration for including VLANs in routing protocols is that the **no shutdown** command must be configured. (For routing traffic to flow, the VLAN must be enabled.)

Note: An IP address cannot be assigned to the Default VLAN, which, by default, is VLAN 1. To assign another VLAN ID to the Default VLAN, use the **default vlan-id** *vlan-id* command.

Assign an IP address to an interface with the following command the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure an IP address and mask on the interface. • *ip-address mask:* enter an address in dotted-decimal format (A.B.C.D) and 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.

Figure 15-9 shows a sample configuration of a VLAN participating in an OSPF process.

Figure 15-9. Sample Layer 3 Configuration of a VLAN

```
interface Vlan 10
 ip address 1.1.1.2/24
 tagged GigabitEthernet 2/2-13
 tagged TenGigabitEthernet 5/0
 ip ospf authentication-key force10
 ip ospf cost 1
 ip ospf dead-interval 60
 ip ospf hello-interval 15
 no shutdown
!
```

Loopback Interfaces

A Loopback interface is a virtual interface in which the software emulates an interface. Packets routed to it are processed locally. Since 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 a Loopback interface, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface loopback *number*	CONFIGURATION	Enter a number as the loopback interface. Range: 0 to 16383.

To view Loopback interface configurations, use the **show interface loopback** *number* command in the EXEC mode.

To delete a Loopback interface, use the **no interface loopback** *number* command syntax in the CONFIGURATION mode.

Many of the same commands found in the physical interface are found in Loopback interfaces.

Null Interfaces

The Null interface is another virtual interface created by the E-Series software. There is only one Null interface. It is always up, but no traffic is transmitted through this interface.

To enter the INTERFACE mode of the Null interface, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface null 0	CONFIGURATION	Enter the INTERFACE mode of the Null interface.

The only configurable command in the INTERFACE mode of the Null interface is the **ip unreachable** command.

Port Channel Interfaces

Port channel interfaces support link aggregation, as described in IEEE Standard 802.3ad.

This section covers the following topics:

- Port channel definition and standards on page 294
- Port channel benefits on page 294
- Port channel implementation on page 295
- Configuration task list for port channel interfaces on page 296

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 FTOS, 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

For the E-Series, 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, the user can create larger-capacity interfaces by utilizing a group of lower-speed links. For example, the user can build a 5-Gigabit interface by aggregating five 1-Gigabit Ethernet interfaces together. If one of the five interfaces fails, traffic is redistributed across the four remaining interfaces.

Port channel implementation

FTOS supports two types of port channels:

- **Static**—Port channels that are statically configured
- **Dynamic**—Port channels that are dynamically configured using Link Aggregation Control Protocol (LACP). For details, see Chapter 19, Link Aggregation Control Protoco.

Table 15-2. Number of Port-channels per Platform

Platform	Port-channels	Members/Channel
E-Series	255	16
C-Series	128	8
S-Series	128	8

Table 15-3. Maximum number of configurable Port-channels

Platform	Port-channels	Members/Channel
E-Series ExaScale	255	64

As soon as a port channel is configured, FTOS 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 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.

10/100/1000 Mbps interfaces in port channels

When both 10/100/1000 interfaces and GigE 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 that interface is enabled, 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, FTOS disables them.

For example, if four interfaces (Gi 0/0, 0/1, 0/2, 0/3) in which Gi 0/0 and Gi 0/3 are set to speed 100 Mb/s and the others are set to 1000 Mb/s, with all interfaces enabled, and you add them to a port channel by entering **channel-member gigabitethernet 0/0-3** while in the port channel interface mode, and FTOS determines if the first interface specified (Gi 0/0) is up. Once it is up, the common speed of the port channel is 100 Mb/s. FTOS disables those interfaces configured with speed 1000 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 Gi 0/0 interface to 1000 Mb/s.

Configuration task list for port channel interfaces

To configure a port channel (LAG), you use the commands similar to those found in physical interfaces. By default, no port channels are configured in the startup configuration.

- Create a port channel (mandatory)
- Add a physical interface to a port channel on page 297 (mandatory)
- Reassign an interface to a new port channel on page 299 (optional)
- Configure the minimum oper up links in a port channel (LAG) on page 300 (optional)
- Add or remove a port channel from a VLAN on page 300 (optional)
- Assign an IP address to a port channel on page 301 (optional)
- Delete or disable a port channel on page 301 (optional)
- Load balancing through port channels on page 302 (optional)

Create a port channel

You can create up to 255 port channels on an E-Series. You can create up to 128 port channels on C-Series and S-Series.

To configure a port channel, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface port-channel** *id-number*	CONFIGURATION	Create a port channel.
2	**no shutdown**	INTERFACE PORT-CHANNEL	Ensure that the port channel is active.

The port channel is now enabled and you can place the port channel in Layer 2 or Layer 3 mode. Use the **switchport** command to place the port channel in Layer 2 mode or configure an IP address to place the port channel in Layer 3 mode.

You can configure a port channel as you would a physical interface by enabling or configuring protocols or assigning access control lists.

Add 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 FTOS disables the interfaces that are not the same speed of the first channel member in the port channel (see 10/100/1000 Mbps interfaces in port channels).

You can add any physical interface to a port channel if the interface configuration is minimal. Only the following commands can be configured 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 S-Series 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 the INTERFACE mode for that interface and enter the **show config** command or from the EXEC Privilege mode, enter the **show running-config interface** *interface* command.

When an interface is added to a port channel, FTOS recalculates the hash algorithm.

To add a physical interface to a port channel, use these commands in the following sequence in the INTERFACE mode of a port channel:

Step	Command Syntax	Command Mode	Purpose
1	**channel-member** *interface*	INTERFACE PORT-CHANNEL	Add the interface to a port channel. The *interface* variable is the physical interface type and slot/port information.
2	**show config**	INTERFACE PORT-CHANNEL	Double check that the interface was added to the port channel.

To view the port channel's status and channel members in a tabular format, use the **show interfaces port-channel brief** (Figure 177) command in the EXEC Privilege mode.

Figure 15-10. show interfaces port-channel brief Command Example

```
Force10#show int port brief

LAG Mode  Status       Uptime    Ports
1   L2L3  up           00:06:03  Gi 13/6    (Up) *
                                 Gi 13/12   (Up)
2   L2L3  up           00:06:03  Gi 13/7    (Up) *
                                 Gi 13/8    (Up)
                                 Gi 13/13   (Up)
                                 Gi 13/14   (Up)
Force10#
```

Figure 15-11 displays 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.

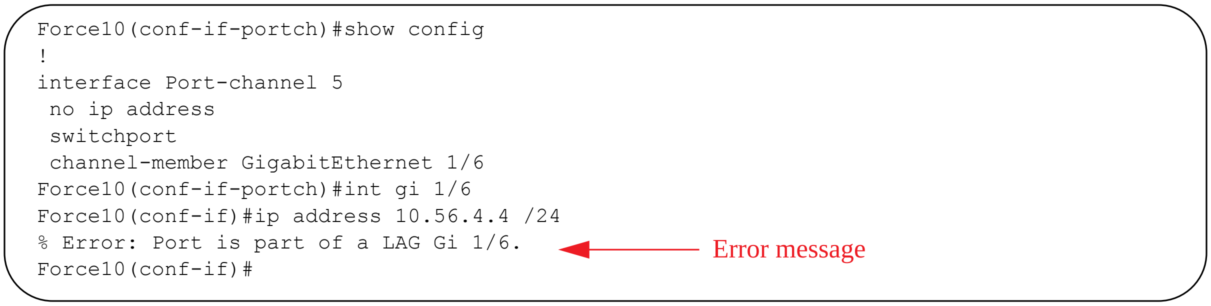
Figure 15-11. show interface port-channel Command Example

```
Force10>show interface port-channel 20
Port-channel 20 is up, line protocol is up
Hardware address is 00:01:e8:01:46:fa
Internet address is 1.1.120.1/24
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 2000 Mbit
Members in this channel:  Gi 9/10 Gi 9/17
ARP type: ARPA, ARP timeout 04:00:00
Last clearing of "show interface" counters 00:00:00
Queueing strategy: fifo
     1212627 packets input, 1539872850 bytes
     Input 1212448 IP Packets, 0 Vlans 0 MPLS
     4857 64-byte pkts, 17570 over 64-byte pkts, 35209 over 127-byte pkts
     69164 over 255-byte pkts, 143346 over 511-byte pkts, 942523 over 1023-byte pkts
     Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
     42 CRC, 0 IP Checksum, 0 overrun, 0 discarded
     2456590833 packets output, 203958235255 bytes, 0 underruns
     Output 1640 Multicasts, 56612 Broadcasts, 2456532581 Unicasts
     2456590654 IP Packets, 0 Vlans, 0 MPLS
     0 throttles, 0 discarded
Rate info (interval 5 minutes):
     Input 00.01Mbits/sec,          2 packets/sec
     Output 81.60Mbits/sec,     133658 packets/sec
Time since last interface status change: 04:31:57

Force10>
```

When more than one interface is added to a Layer 2 port channel, FTOS selects one of the active interfaces in the port channel to be the Primary Port. The primary port replies to flooding and sends protocol 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. As Figure 15-12 illustrates, interface GigabitEthernet 1/6 is part of port channel 5, which is in Layer 2 mode, and an error message appeared when an IP address was configured.

Figure 15-12. Error Message



```
Force10(conf-if-portch)#show config
!
interface Port-channel 5
 no ip address
 switchport
 channel-member GigabitEthernet 1/6
Force10(conf-if-portch)#int gi 1/6
Force10(conf-if)#ip address 10.56.4.4 /24
% Error: Port is part of a LAG Gi 1/6.
Force10(conf-if)#
```

Error message

Reassign 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, you must 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, FTOS recalculates the hash algorithm for the port channel.

To reassign an interface to a new port channel, use these commands in the following sequence in the INTERFACE mode of a port channel:

Step	Command Syntax	Command Mode	Purpose
1	**no channel-member** *interface*	INTERFACE PORT-CHANNEL	Remove the interface from the first port channel.
2	**interface port-channel** id *number*	INTERFACE PORT-CHANNEL	Change to the second port channel INTERFACE mode.
3	**channel-member** *interface*	INTERFACE PORT-CHANNEL	Add the interface to the second port channel.

Figure 15-13 displays an example of moving the GigabitEthernet 1/8 interface from port channel 4 to port channel 3.

Figure 15-13. Command Example from Reassigning an Interface to a Different Port Channel

```
Force10(conf-if-portch)#show config
!
interface Port-channel 4
 no ip address
 channel-member GigabitEthernet 1/8
 no shutdown
Force10(conf-if-portch)#no chann gi 1/8
Force10(conf-if-portch)#int port 5
Force10(conf-if-portch)#channel gi 1/8
Force10(conf-if-portch)#sho conf
!
interface Port-channel 5
 no ip address
 channel-member GigabitEthernet 1/8
 shutdown
Force10(conf-if-portch)#
```

Configure the minimum oper up links in a port channel (LAG)

You can configure the minimum links in a port channel (LAG) that must be in “oper up” status for the port channel to be considered to be in “oper up” status. Use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
minimum-links *number*	INTERFACE	Enter the number of links in a LAG that must be in “oper up” status. Default: 1

Figure 15-14 displays an example of configuring five minimum “oper up” links in a port channel.

Figure 15-14. Example of using the minimum-links Command

```
Force10#config t
Force10(conf)#int po 1
Force10(conf-if-po-1)#minimum-links 5
Force10(conf-if-po-1)#
```

Add or remove 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, you must place the port channel in Layer 2 mode (by using the **switchport** command).

To add a port channel to a VLAN, use either of the following commands:

Command Syntax	Command Mode	Purpose
tagged port-channel *id number*	INTERFACE VLAN	Add the port channel to the VLAN as a tagged interface. An interface with tagging enabled can belong to multiple VLANs.
untagged port-channel *id number*	INTERFACE VLAN	Add the port channel to the VLAN as an untagged interface. An interface without tagging enabled can belong to only one VLAN.

To remove a port channel from a VLAN, use either of the following commands:

Command Syntax	Command Mode	Purpose
no tagged port-channel *id number*	INTERFACE VLAN	Remove the port channel with tagging enabled from the VLAN.
no untagged port-channel *id number*	INTERFACE VLAN	Remove the port channel without tagging enabled from the VLAN.

To see which port channels are members of VLANs, enter the **show vlan** command in the EXEC Privilege mode.

Assign 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 in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure an IP address and mask on the interface. • *ip-address mask:* enter an address in dotted-decimal format (A.B.C.D) and 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.

Delete or disable a port channel

To delete a port channel, you must be in the CONFIGURATION mode and use the **no interface portchannel** *channel-number* command.

When you disable a port channel (using the **shutdown** command) all interfaces within the port channel are operationally down also.

Load balancing through port channels

FTOS 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.

FTOS 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.

E-Series load-balancing

On the E-Series, the default **load-balance** criteria are a 5-tuple, as follows:

- IP source address
- IP destination address
- Protocol type
- TCP/UDP source port
- TCP/UDP destination port

Balancing may be applied to IPv4, switched IPv6, and non-IP traffic. For these traffic types, the IP-header-based hash and MAC-based hash may be applied to packets by using the following methods.

Table 15-4. Hash Methods as Applied to Port Channel Types

Hash (Header Based)	Layer 2 Port Channel	Layer 3 Port Channel
5-tuple	X	X
3-tuple	X	X
Packet-based	X	X
MAC source address (SA) and destination address (DA)	X	

On the E-Series, to change the 5-tuple default to 3-tuple, MAC, or packet-based, use the following command in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
[**no**] **load-balance** [**ip-selection** {**3-tuple | packet-based**}] [**mac**]	CONFIGURATION	To designate a method to balance traffic over a port channel. By default, IP 5-tuple is used to distribute traffic over members port channel. **ip-selection 3-tuple**—Distribute IP traffic based on IP source address, IP destination address, and IP protocol type. **ip-selection packet-based**—Distribute IPV4 traffic based on the IP Identification field in the IPV4 header. **mac**—Distribute traffic based on the MAC source address, and the MAC destination address. See Table 15-6 for more information.

For details on the **load-balance** command, see the IP Routing chapter of the *FTOS Command Reference*.

To distribute IP traffic over an E-Series port channel member, FTOS uses the 5-tuple IP default. The 5-tuple and the 3-tuple hash use the following keys:

Table 15-5. 5-tuple and 3-tuple Keys

Keys	5-tuple	3-tuple
IP source address (lower 32 bits)	X	X
IP destination address (lower 32 bits)	X	X
Protocol type	X	X
TCP/UDP source port	X	
TCP/UDP destination port	X	

Note: For IPV6, only the first 32 bits (LSB) of IP Source Address and IP Destination Address are used for hash generation.

Figure 15-15 shows the configuration and show command for packet-based hashing on the E-Series.

Figure 15-15. Command example: load-balance ip-selection packet-based

```
Force10(conf)#load-balance ip-selection packet-based


Force10#show running-config | grep load
load-balance ip-selection packet-based
Force10#
```

The **load-balance packet based** command can co-exist with **load balance mac** command to achieve the functionality shown in Table 15-6.

IPv4, IPv6, and non-IP traffic handling on the E-Series

The table below presents the combinations of the **load-balance** command and their effect on traffic types.

Table 15-6. The load-balance Commands and Port Channel Types

Configuration Commands	Switched IP Traffic	Routed IP Traffic (IPv4 only)	Switched Non-IP Traffic
Default (IP 5-tuple)	IP 5-tuple (lower 32 bits)	IP 5-tuple	MAC-based
load-balance ip-selection 3-tuple	IP 3-tuple (lower 32 bits)	IP 3-tuple	MAC-based
load-balance ip-selection mac	MAC-based	IP 5-tuple	MAC-based
load-balance ip-selection 3-tuple **load-balance ip-selection mac**	MAC-based	IP 3-tuple	MAC-based
load-balance ip-selection packet-based	Packet based: IPV4 No distribution: IPV6	Packet-based	MAC-based
load-balance ip-selection packet-based **load-balance ip-selection mac**	MAC-based	Packet-based	MAC-based

C-Series and S-Series load-balancing

For LAG hashing on C-Series and S-Series, the source IP, destination IP, source TCP/UDP port, and destination TCP/UDP port are used for hash computation by default. For packets without a Layer 3 header, FTOS automatically uses **load-balance mac source-dest-mac**.

IP hashing or MAC hashing should not be configured at the same time. If you configure an IP and MAC hashing scheme at the same time, the MAC hashing scheme takes precedence over the IP hashing scheme.

To change the IP traffic load balancing default on the C-Series and S-Series, use the following command:

Command Syntax	Command Mode	Purpose
[**no**] **load-balance** {**ip-selection** [**dest-ip** | **source-ip**]} | {**mac** [**dest-mac** | **source-dest-mac** | **source-mac**]} | {**tcp-udp enable**}	CONFIGURATION	Replace the default IP 4-tuple method of balancing traffic over a port channel. You can select one, two, or all three of the following basic hash methods **ip-selection** [**dest-ip | source-ip**]—Distribute IP traffic based on IP destination or source address. **mac** [**dest-mac | source-dest-mac | source-mac**]—Distribute IPV4 traffic based on the destination or source MAC address, or both, along with the VLAN, Ethertype, source module ID and source port ID. **tcp-udp enable**—Distribute traffic based on TCP/UDP source and destination ports.

Hash algorithm

The **load-balance** command discussed above selects the hash criteria applied to port channels.

If even distribution is not obtained with the load-balance command, the **hash-algorithm** command can be used to select the hash scheme for LAG, ECMP and NH-ECMP. The 12 bit Lag Hash can be rotated or shifted till the desired hash is achieved.

The **nh-ecmp** option allows you to change the hash value for recursive ECMP routes independently of non-recursive ECMP routes. This option provides for better traffic distribution over available equal cost links that involve a recursive next hop lookup.

For the E-Series TeraScale and ExaScale, you can select one of 47 possible hash algorithms.

Command Syntax	Command Mode	Purpose
hash-algorithm {*algorithm-number*}| {**ecmp** {**checksum|crc|xor**} [*number*]} **lag** {*checksum*|*crc*|*xor*][*number*]}**nh-ecmp** {[*checksum*|*crc*|*xor]* [*number*]}}| {**linecard** *number* **ip-sa-mask** *value* **ip-da-mask** *value*}	CONFIGURATION	Change the default (0) to another algorithm and apply it to ECMP, LAG hashing, or a particular line card. **Note:** To achieve the functionality of hash-align on the ExaScale platform, do not use CRC as an hash-algorithm method. For ExaScale systems, set the default hash-algorithm method to ensure CRC is not used for LAG. For example, **hash-algorithm ecmp xor lag checksum nh-ecmp checksum** For details on the algorithm choices, see the command details in the IP Routing chapter of the *FTOS Command Reference*.

Note: E-Series systems require the **lag-hash-align** microcode be configured in the in the CAM profile. E-Series TeraScale E T includes this microcode as an option with the Default cam profile. E-Series ExaScale E X systems require that a CAM profile be created and specifically include **lag-hash-align** microcode.

Figure 15-16 shows a sample configuration for the **hash-algorithm** command.

Figure 15-16. Command example: hash-algorithm

```
Force10(conf)#Force10(conf)#hash-algorithm ecmp xor 26 lag crc 26 nh-ecmp checksum 26
Force10(conf)#
```

On C-Series and S-Series, the **hash-algorithm** command is specific to ECMP groups and has different defaults from the E-Series. The default ECMP hash configuration is **crc-lower**. This takes the lower 32 bits of the hash key to compute the egress port. Other options for ECMP hash-algorithms are:

- **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
- **lsb** — always uses the least significant bit of the hash key to compute the egress port

To change to another method, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
hash-algorithm ecmp {**crc-upper**} | {**dest-ip**} | {**lsb**}	CONFIGURATION	Change to another algorithm.

For more on load-balancing, see “Equal Cost Multipath and Link Aggregation Frequently Asked Questions” in the E-Series FAQ section (login required) of iSupport:

https://www.force10networks.com/CSPortal20/KnowledgeBase/ToolTips.aspx

Bulk Configuration

Bulk configuration enables 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 interface range prompt. In the following example, Tengigabit 3/0 and VLAN 1000 do not exist.

Note: When creating an interface range, interfaces appear in the order they were entered and are not sorted.

The **show range** command is available under interface range mode. This command allows you to display all interfaces that have been validated under the interface range context.

The **show configuration** command is also available under the interface range mode. This command allows you to display the running configuration only for interfaces that are part of interface range.

Bulk Configuration Examples

The following are examples of using 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

Figure 15-17. Creating a Single-Range Bulk Configuration

```
Force10(config)# interface range gigabitethernet 5/1 - 23
Force10(config-if-range-gi-5/1-23)# no shutdown
```

Create a multiple-range

Figure 15-18. Creating a Multiple-Range Prompt

```
Force10(conf)#interface range tengigabitethernet 3/0 , gigabitethernet 2/1 - 47 , vlan 1000
Force10(conf-if-range-gi-2/1-47,so-5/0)#
```

Exclude duplicate entries

Duplicate single interfaces and port ranges are excluded from the resulting interface range prompt:

Figure 15-19. Interface Range Prompt Excluding Duplicate Entries

```
Force10(conf)#interface range vlan 1 , vlan 1 , vlan 3 , vlan 3
Force10(conf-if-range-vl-1,vl-3)#
Force10(conf)#interface range gigabitethernet 2/0 - 23 , gigabitethernet 2/0 - 23 , gigab 2/0 - 23
Force10(conf-if-range-gi-2/0-23)#
```

Exclude a smaller port range

If interface range has multiple port ranges, the smaller port range is excluded from prompt:

Figure 15-20. Interface Range Prompt Excluding a Smaller Port Range

```
Force10(conf)#interface range gigabitethernet 2/0 - 23 , gigab 2/1 - 10
Force10(conf-if-range-gi-2/0-23)#
```

Overlap port ranges

If overlapping port ranges are specified, the port range is extended to the smallest start port number and largest end port number:

Figure 15-21. Interface Range Prompt Including Overlapping Port Ranges

```
Force10(conf)#inte ra gi 2/1 - 11 , gi 2/1 - 23
Force10(conf-if-range-gi-2/1-23)#
```

Commas

The example below shows 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.

```
Force10(config-if)# interface range gigabitethernet 5/1 - 23, tengigabitethernet 1/1 - 2
Force10(config-if-range-gi-5/1-23)# no shutdown
Force10(config-if-range-gi-5/1-23)#
```

Figure 15-22. Multiple-Range Bulk Configuration Gigabit Ethernet and Ten-Gigabit Ethernet

Add ranges

The example below shows how to use commas to add VLAN and port-channel interfaces to the range.

Figure 15-23. Multiple-Range Bulk Configuration with VLAN, and Port-channel

```
Force10(config-ifrange-gi-5/1-23-te-1/1-2)# interface range Vlan 2 – 100 , Port 1 – 25
Force10(config-if-range-gi-5/1-23-te-1/1-2-so-5/1-vl-2-100-po-1-25)# no shutdown
Force10(config-if-range)#
```

Interface Range Macros

The user 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, you must define the macro.

To define an interface-range macro, enter this command:

Command Syntax	Command Mode	Purpose
Force10 (config)# **define** *interface-range macro_name* {**vlan** *vlan_ID* - *vlan_ID*} | {{**gigabitethernet** | **tengigabitethernet**} *slot/interface* - *interface*} [**,** {**vlan** *vlan_ID* - *vlan_ID*} {{**gigabitethernet** | **tengigabitethernet**} *slot/interface* - *interface*}]	CONFIGURATION	Defines the interface-range macro and saves it in the running configuration file.

Define the Interface Range

This example shows how to define an interface-range macro named "test" to select Fast Ethernet interfaces 5/1 through 5/4:

```
Force10(config)# define interface-range test gigabitethernet 5/1 - 4
```

Choose an Interface-range Macro

To use an interface-range macro in the **interface range** command, enter this command:

Command Syntax	Command Mode	Purpose
interface range macro *name*	CONFIGURATION	Selects the interfaces range to be configured using the values saved in a named interface-range macro.

The example below shows how to change to the interface-range configuration mode using the interface-range macro named "test".

```
Force10(config)# interface range macro test
Force10(config-if)#
```

Monitor and Maintain 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, etc.

Command Syntax	Command Mode	Purpose
monitor interface *interface*	EXEC Privilege	View the interface’s statistics. Enter the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/ port information. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.

The information (Figure 15-24) displays in a continuous run, refreshing every 2 seconds by default. Use the following keys to manage the output.

m - Change mode	c - Clear screen
l - Page up	a - Page down
T - Increase refresh interval (by 1 second)	t - Decrease refresh interval (by 1 second)
q - Quit	

Figure 15-24. Command Example: monitor interface

```
Force10#monitor interface gi 3/1


 Force10 uptime is 1 day(s), 4 hour(s), 31 minute(s)
   Monitor time: 00:00:00   Refresh Intvl.: 2s

 Interface: Gi 3/1, Disabled, Link is Down, Linespeed is 1000 Mbit

  Traffic statistics:              Current           Rate                Delta
         Input bytes:                    0          0 Bps                    0
        Output bytes:                    0          0 Bps                    0
       Input packets:                    0          0 pps                    0
      Output packets:                    0          0 pps                    0
         64B packets:                    0          0 pps                    0
    Over 64B packets:                    0          0 pps                    0
   Over 127B packets:                    0          0 pps                    0
   Over 255B packets:                    0          0 pps                    0
   Over 511B packets:                    0          0 pps                    0
  Over 1023B packets:                    0          0 pps                    0
 Error statistics:
     Input underruns:                    0          0 pps                    0
        Input giants:                    0          0 pps                    0
     Input throttles:                    0          0 pps                    0
           Input CRC:                    0          0 pps                    0
   Input IP checksum:                    0          0 pps                    0
       Input overrun:                    0          0 pps                    0
    Output underruns:                    0          0 pps                    0
    Output throttles:                    0          0 pps                    0

       m - Change mode                            c - Clear screen
       l - Page up                                a - Page down
       T - Increase refresh interval              t - Decrease refresh interval
       q - Quit

q
Force10#
```

Maintenance using TDR

The Time Domain Reflectometer (TDR) is supported on all Dell Force10 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 10/100/1000 BASE-T modules, use the **tdr-cable-test** command:

Step	Command Syntax	Command Mode	Usage
1	**tdr-cable-test gigabitethernet** *<slot>/<port>*	EXEC Privilege	To test for cable faults on the GigabitEthernet cable. • Between two ports, the user must not start the test on both ends of the cable. • The user must enable the interface before starting the test. • The port should be enabled to run the test or the test prints an error message.
2	**show tdr gigabitethernet** *<slot>/<port>*	EXEC Privilege	Displays TDR test results.

Link Debounce Timer

Link Debounce Timer is supported on platform E

The Link Debounce Timer feature isolates upper layer protocols on Ethernet switches and routers from very short-term, possibly repetitive interface flaps often caused by network jitter on the DWDM equipment connecting the switch and other devices on a SONET ring. The Link Debounce Timer delays link change notifications, thus decreasing traffic loss due to network configuration. All interfaces have a built-in timer to manage traffic. This feature extends the time allowed by the upper layers.

The SONET ring has its own restore time whenever there is a failure. During this time, however, the Ethernet interface connected to the switch will flap. Link Debounce Timer instructs the Ethernet switch to delay the notification of the link change to the upper layers. If the link state changes again within this period, no notification goes to the upper layers, so that the switch remains unaware of the change.

Note: Enabling the link debounce timer causes link up and link down detections to be delayed, resulting in traffic being blackholed during the debouncing period. This situation might affect the convergence and reconvergence of some Layer 2 and Layer 3 protocols.

Important Points to Remember about Link Debounce Timer

- Link Debounce Timer is configurable on physical ports only.
- Only 1G fiber, 10/100/1000 copper, 10G fiber, 10G copper are supported.
- This feature is not supported on management interfaces or SONET interfaces.
- Link Debounce takes effect only when the operational state of the port is up.
- Link Debounce is supported on interfaces that also have link dampening configured.
- Unlike link dampening, link debounce timer does not notify other protocols.

- Changes made do not affect any ongoing debounces. The timer changes take affect from the next debounce onward.

Assign a debounce time to an interface

Command Syntax	Command Mode	Purpose
link debounce time [*milliseconds*]	INTERFACE	Enter the time to delay link status change notification on this interface. Range: 100-5000 ms • Default for Copper is 3100 ms • Default for Fiber is 100 ms

Figure 15-25. Setting Debounce Time

```
Force10(conf)#int gi 3/1
Force10(conf-if-gi-3/1)#link debounce time 150
Force10(conf-if-gi-3/1)#=
```

Show debounce times in an interface

show interface debounce [type**] [**slot/ port**]**	EXEC Privilege	Show the debounce time for the specified interface. Enter the interface type keyword followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/ port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.

Figure 15-26. Showing .Debounce Time

```
Force10#
Force10#show interfaces debounce gigabitethernet 3/1
 Interface                Time(ms)
 GigabitEthernet 3/1      200
Force10#
```

Note: FTOS rounds the entered debounce time up to the nearest hundredth.
Note in Figure 15-25 that the timer was set at 150 ms, but appears as 200 in Figure 15-26.

Disable ports when one only SFM is available (E300 only)

Selected ports can be shut down when a single SFM is available on the E300 system. Each port to be shut down must be configured individually.

When an E300 system boots up and a single SFM is active this configuration, any ports configured with this feature will be shut down. All other ports are booted up.

Similarly, if an SFM fails (or is removed) in an E300 system with two SFM, ports configured with this feature will be shut down. All other ports are treated normally.

When a second SFM is installed or replaced, all ports are booted up and treated as normally. This feature does not take affect until a single SFM is active in the E300 system.

Disable port on one SFM

This feature must be configured for each interface to shut down in the event that an SFM is disabled. Enter the command **disable-on-sfm-failure** from INTERFACE mode to disable the port when only a single SFM is available.

Link Dampening

Interface state changes occur when interfaces are administratively brought up or down or if an interface state changes. Every time an interface changes state or flaps, routing protocols are notified of the status of the routes that are affected by the change in state, and these protocols go through momentous task of re-converging. Flapping therefore puts the status of entire network at risk of transient loops and black holes.

Link dampening minimizes the risk created by flapping by imposing a penalty for each interface flap and decaying the penalty exponentially. Once the penalty exceeds certain threshold, the interface is put in an "error-disabled" state, and for all practical purposes of routing, the interface is deemed to be "down." Once the interface becomes stable and the penalty decays below a certain threshold, the interface comes up again and the routing protocols re-converge.

Link dampening:

- reduces processing on the CPUs by reducing excessive interface flapping.
- improves network stability by penalizing misbehaving interfaces and redirecting traffic
- improves convergence times and stability throughout the network by isolating failures so that disturbances are not propagated.

Important Points to Remember

- Link dampening is not supported on VLAN interfaces
- Link dampening is disabled when the interface is configured for port monitoring
- Link dampening can be applied to Layer 2 and Layer 3 interfaces.
- Link dampening can be configured on individual interfaces in a LAG.

Enable Link Dampening

Enable link dampening using the command **dampening** from INTERFACE mode, as shown in Figure 15-27.

Figure 15-27. Configuring Link Dampening

```
R1(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 ip address 10.10.19.1/24
 dampening 1 2 3 4
 no shutdown
R1(conf-if-gi-1/1)#exit
```

View the link dampening configuration on an interface using the command **show config**, or view dampening information on all or specific dampened interfaces using the command **show interfaces dampening** from EXEC Privilege mode, as shown in Figure 15-28.

Figure 15-28. Viewing all Dampened Interfaces

```
Force10# show interfaces dampening
InterfaceState Flaps  Penalty Half-LifeReuse SuppressMax-Sup
Gi 0/0         Up     0       0       5              750     2500           20
Gi 0/1         Up     2       1200    20             500     1500           300
Gi 0/2         Down   4       850     30             600     2000           120
```

View a dampening summary for the entire system using the command show interfaces dampening summary from EXEC Privilege mode, as shown in Figure 15-29.

Figure 15-29. Viewing a System-wide Dampening Summary

```
Force10# show interfaces dampening summary
20 interfaces are configured with dampening. 3 interfaces are currently suppressed.
Following interfaces are currently suppressed:
Gi 0/2
Gi 3/1
Gi 4/2
Force10#
```

Clear Dampening Counters

Clear dampening counters and accumulated penalties using the command **clear dampening**, as shown in Figure 15-30.

Figure 15-30. Clearing Dampening Counters

```
Force10# clear dampening interface Gi 0/1

Force10# show interfaces dampening GigabitEthernet0/0
InterfaceState Flaps   Penalty Half-LifeReuse SuppressMax-Sup
Gi 0/1  Up      0       0       20              500     1500          300
```

Link Dampening Support for XML

View the output of the following show commands in XML by adding **| display xml** to the end of the command:

- show interfaces dampening
- show interfaces dampening summary
- show interfaces interface x/y

Configure MTU size on an Interface

The E-Series supports a link Maximum Transmission Unit (MTU) of 9252 bytes and maximum IP MTU of 9234 bytes. The link MTU is the frame size of a packet, and 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, FTOS divides the packet into fragments no bigger than the size set in the **ip mtu** command.

In FTOS, MTU is defined as the entire Ethernet packet (Ethernet header + FCS + payload)

Since different networking vendors define MTU differently, check their documentation when planing MTU sizes across a network.

Table 15-7 lists the range for each transmission media.

Table 15-7. MTU Range

Transmission Media	MTU Range (in bytes)
Ethernet	594-9252 = link MTU 576-9234 = IP MTU

Ethernet Pause Frames

Ethernet Pause Frames is supported on platforms C E S

Threshold Settings are supported only on platforms: C S

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 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 that 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 half duplex is already configured: 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.

Threshold Settings

Threshold Settings are supported only on platforms: C S

When the transmission pause is set (**tx on**), 3 thresholds can be set to define the controls more closely. Ethernet Pause Frames flow control can be triggered when either the flow control buffer threshold or flow control packet pointer threshold is reached. The thresholds are:

- Number of flow-control packet pointers: 1-2047 (default = 75)
- Flow-control buffer threshold in KB: 1-2013 (default = 49KB)
- Flow-control discard threshold in KB: 1-2013 (default= 75KB)

The pause is started when *either* the packet pointer or the buffer threshold is met (whichever is met first). When the discard threshold is met, packets are dropped.

The pause ends when *both* the packet pointer and the buffer threshold fall below 50% of the threshold settings.

The discard threshold defines when the interface starts dropping the packet on the interface. This may be necessary when a connected device doesn’t honor the flow control frame sent by S-Series.

The discard threshold should be larger than the buffer threshold so that the buffer holds at least hold at least 3 packets.

Enable Pause Frames

Note: On the C-Series and S-Series (non-S55) platforms, Ethernet Pause Frames TX should be enabled *only after* consulting with the Dell Force10 Technical Assistance Center.

Note: The S55 supports only the **rx** control option. The S55 does not transmit pause frames.

Ethernet Pause Frames flow control must be enabled on all ports on a chassis or a line card. If not, the system may exhibit unpredictable behavior.

On the C-Series and S-Series systems, the flow-control sender and receiver must be on the same port-pipe. Flow control is not supported across different port-pipes on the C-Series or S-Series system.

Command Syntax	Command Mode	Purpose
flowcontrol rx [*off* | *on*] **tx** [*off* | *on*] [*threshold* {<1-2047> <1-2013> <1-2013>}]	INTERFACE	Control how the system responds to and generates 802.3x pause frames on 1 and 10Gig line cards. Defaults: C-Series: **rx off tx off** E-Series: **rx on tx on** S-Series: **rx off tx off** S55: **rx o**ff
	Parameters: **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. **threshold** (C-Series and S-Series only)**:** When **tx on** is configured, the user can set the threshold values for: Number of flow-control packet pointers: 1-2047 (default = 75) Flow-control buffer threshold in KB: 1-2013 (default = 49KB) Flow-control discard threshold in KB: 1-2013 (default= 75KB) Pause control is triggered when either the flow control buffer threshold or flow control packet pointer threshold is reached.	

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 592-9252, with a default of 1500. On the E-Series, the user must enter the **ip mtu** command to manually configure the IP MTU to compensate for the Layer 2 header. The C-Series and S-Series automatically configure the IP MTU.

Table 15-8 lists the various Layer 2 overheads found in FTOS and the number of bytes.

Table 15-8. 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.

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.

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 port pipe is a Dell Force10 specific term for the hardware path that packets follow through a system. Port pipes travel through a collection of circuits (ASICs) built into line cards and RPMs on which various processing events for the packets occur. One or two port pipes process traffic for a given set of physical interfaces or a port-set. The E300 only supports one port pipe per slot. On the E1200 and E600 each slot has two port pipes with following specifications:

- 48 port line rate cards have two port pipes on the line card
- 48 port high density cards have only one port pipe on the line card

Note: All references to the E1200 in this section include the E1200i-AC and E1200i-DC. References to E600 include the E600i.

For the purposes of diagnostics, the major difference between the E-Series platforms is the number of port pipes per slot.

- E1200 and E600—Each slot has two port-pipes. Each portpipe has nine 3.125Gbps channels to the backplane, one to each SFM.
- E300—Each slot has one portpipe. Each port-pipe has eight 3.125Gbps channels to the backplane, with four channels to each SFM.

Table 15-9 presents these platform differences again.

Table 15-9. Platform Differences Concerning Port-pipes

Chassis Type	Port-pipes / Slot	Channels / Port-pipe	Capacity of Each Channel (Gbps)	Raw Slot Capacity (Gbps)
E1200/E1200i-AC/DC	2	9	3.125	56.25
E600/E600i	2	9	3.125	56.25
E300	1	8	3.125	25

Auto-Negotiation on Ethernet Interfaces

Setting speed and duplex mode of Ethernet Interfaces

By default, auto-negotiation of speed and duplex mode is enabled on 10/100/1000 Base-T Ethernet interfaces. Only 10GE interfaces do not support auto-negotiation. When using 10GE interfaces, verify that the settings on the connecting devices are set to no auto-negotiation.

Note: Starting with FTOS 7.8.1.0, when a copper SFP2 module with catalog number GP-SFP2-1T is used in the S25P model of the S-Series, its speed can be manually set with the **speed** command. When the speed is set to 10 or 100 Mbps, the **duplex** command can also be executed.

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 Force10 recommends keeping auto-negotiation enabled. Auto-negotiation should only be disabled on switch ports that attach to devices not capable of supporting negotiation or where connectivity issues arise from interoperability issues.

For 10/100/1000 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** or **auto**.

To discover whether the remote and local interface require manual speed synchronization, and to manually synchronize them if necessary, use the following command sequence (see Figure 15-32 on page 322):

Step	Task	Command Syntax	Command Mode
1	Determine the local interface status. See Figure 15-31.	**show interfaces** [*interface* | **linecard** *slot-number*] **status**	EXEC Privilege
2	Determine the remote interface status.	[Use the command on the remote system that is equivalent to the above command.]	EXEC EXEC Privilege
3	Access CONFIGURATION mode.	**config**	EXEC Privilege
4	Access the port.	**interface** *interface slot*/*port*	CONFIGURATION
5	Set the local port speed.	**speed** {**10** | **100** | **1000** | **auto**}	INTERFACE
6	Optionally, set full- or half-duplex.	**duplex** {**half** | **full**}	INTERFACE
7	Disable auto-negotiation on the port. If the speed was set to 1000, auto-negotiation does not need to be disabled.	**no negotiation auto**	INTERFACE
8	Verify configuration changes.	show config	INTERFACE

Note: The **show interfaces status** command displays link status, but not administrative status. For link and administrative status, use **show ip interface** [**interface | brief | linecard slot-number**] [**configuration**].

Figure 15-31. show interfaces status Command Example

```
Force10#show interfaces status
Port    Description Status Speed      Duplex Vlan
Gi 0/0              Up     1000 Mbit  Auto    --
Gi 0/1              Down   Auto       Auto    1
Gi 0/2              Down   Auto       Auto    1
Gi 0/3              Down   Auto       Auto    --
Gi 0/4 Force10Port  Up     1000 Mbit  Auto   30-130
Gi 0/5              Down   Auto       Auto    --
Gi 0/6              Down   Auto       Auto    --
Gi 0/7              Up     1000 Mbit  Auto   1502,1504,1506-1508,1602
Gi 0/8              Down   Auto       Auto    --
Gi 0/9              Down   Auto       Auto    --
Gi 0/10             Down   Auto       Auto    --
Gi 0/11             Down   Auto       Auto    --
Gi 0/12             Down   Auto       Auto    --
[output omitted]
```

In the example, above, several ports display “Auto” in the Speed field, including port 0/1. In Figure 15-32, the speed of port 0/1 is set to 100Mb and then its auto-negotiation is disabled.

Figure 15-32. Setting Port Speed Example

```
Force10#configure
Force10(config)#interface gig 0/1
Force10(Interface 0/1)#speed 100
Force10(Interface 0/1)#duplex full
Force10(Interface 0/1)#no negotiation auto
Force10(Interface 0/1)#show config
!
interface GigabitEthernet 0/1
no ip address
speed 100
duplex full
no shutdown
```

Setting Auto-Negotiation Options

The **negotiation auto** command provides a **mode** option for configuring an individual port to forced master/forced slave once auto-negotiation is enabled.

Caution: Ensure that only one end of the node is configured as forced-master and the other is configured as forced-slave. If both are configured the same (that is both as forced-master or both as forced-slave), the **show interface** command will flap between an auto-neg-error and forced-master/slave states.

Figure 15-33. Setting Auto-Negotiation Options

```
Force10(conf)# int gi 0/0
Force10(conf-if)#neg auto
Force10(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
```

For details on the **speed**, **duplex**, and **negotiation auto** commands, see the Interfaces chapter of the *FTOS Command Reference*.

Adjust the keepalive timer

Use the **keepalive** command to change the time interval between keepalive messages on the interfaces. 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:

Command Syntax	Command Mode	Purpose
keepalive [*seconds*]	INTERFACE	Change the default interval between keepalive messages.

To view the new setting, use the **show config** command in the INTERFACE mode.

View Advanced Interface Information

Display Only Configured Interfaces

The following options have been implemented for **show [ip | running-config] interfaces** commands for (only) linecard interfaces. When the **configured** keyword is used, only interfaces that have non-default configurations are displayed. Dummy linecard interfaces (created with the **linecard** command) are treated like any other physical interface.

Figure 15-34 lists the possible show commands that have the configured keyword available:

Figure 15-34. show Commands with configured Keyword Examples

```
Force10#show interfaces configured
Force10#show interfaces linecard 0 configured
Force10#show interfaces gigabitEthernet 0 configured
Force10#show ip interface configured
Force10#show ip interface linecard 1 configured
Force10#show ip interface gigabitEthernet 1 configured
Force10#show ip interface br configured
Force10#show ip interface br linecard 1 configured
Force10#show ip interface br gigabitEthernet 1 configured
Force10#show running-config interfaces configured
Force10#show running-config interface gigabitEthernet 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 (Figure 15-35) displays the interface, whether the interface supports IEEE 802.1Q tagging or not, and the VLANs to which the interface belongs.

Figure 15-35. show interfaces switchport Command Example

```
Force10#show interfaces switchport
Name: GigabitEthernet 13/0
802.1QTagged: True
Vlan membership:
Vlan    2


Name: GigabitEthernet 13/1
802.1QTagged: True
Vlan membership:
Vlan    2


Name: GigabitEthernet 13/2
802.1QTagged: True
Vlan membership:
```

Configure Interface Sampling Size

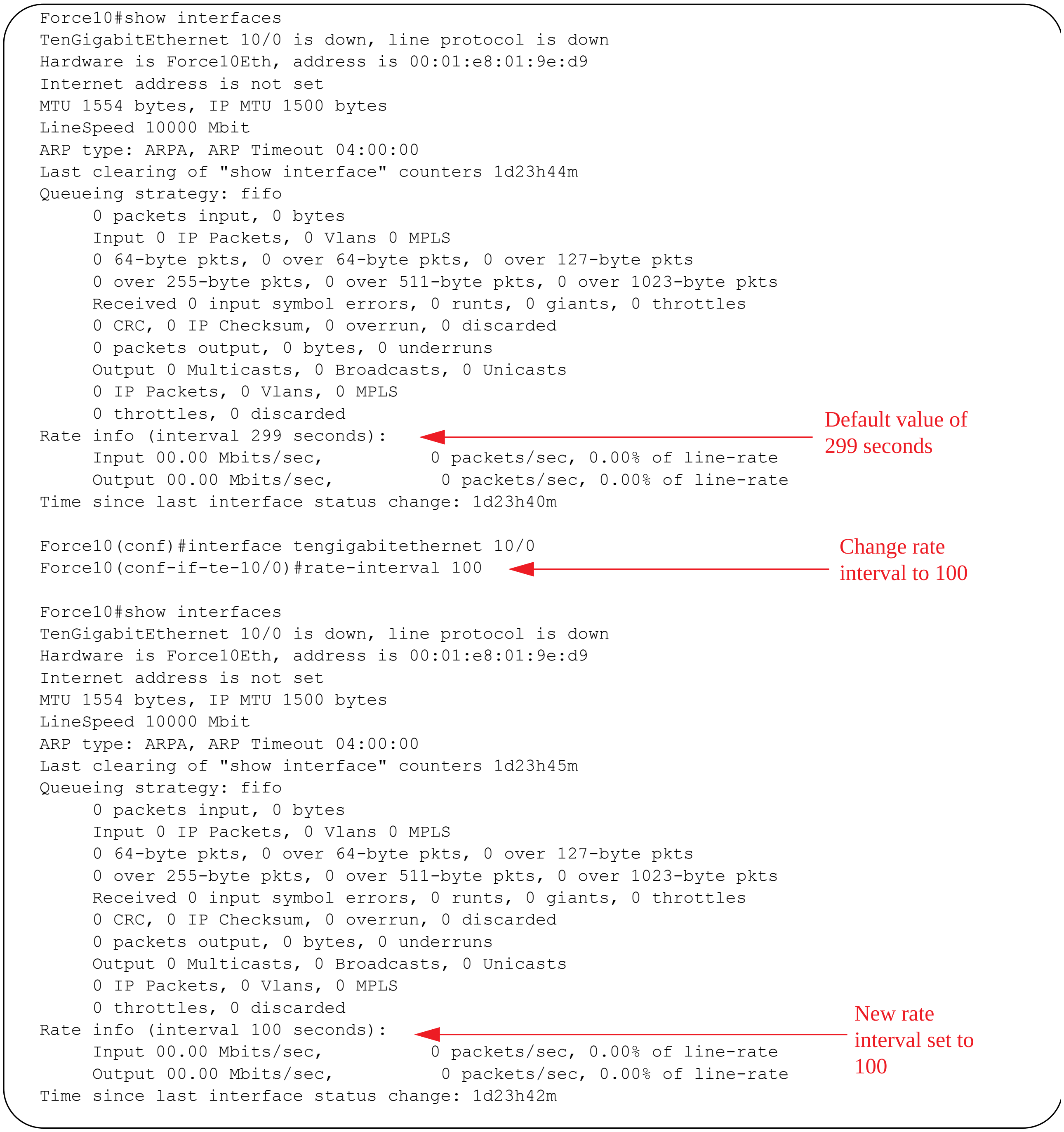
Use the **rate-interval** command, in INTERFACE mode, to configure the number of seconds of traffic statistics to display in the **show interfaces** output.

Although any value between 30 and 299 seconds (the default) can be entered, software polling is done once every 15 seconds. So, for example, if you enter “19”, you will actually get a sample of the past 15 seconds.

All LAG members inherit the rate interval configuration from the LAG.

Figure 15-36 shows how to configure rate interval when changing the default value:

Figure 15-36. Configuring Rate Interval Example



```
Force10#show interfaces
TenGigabitEthernet 10/0 is down, line protocol is down
Hardware is 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

Force10(conf)#interface tengigabitethernet 10/0
Force10(conf-if-te-10/0)#rate-interval 100

Force10#show interfaces
TenGigabitEthernet 10/0 is down, line protocol is down
Hardware is 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 IP Packets, 0 Vlans, 0 MPLS
     0 throttles, 0 discarded
Rate info (interval 100 seconds):
     Input 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
     Output 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
Time since last interface status change: 1d23h42m
```

Default value of 299 seconds
Change rate interval to 100
New rate interval set to 100

Dynamic Counters

By default, counting for the following four applications is enabled:

- IPFLOW
- IPACL
- L2ACL
- L2FIB

For remaining applications, FTOS automatically turns on counting when the application is enabled, and is turned off when the application is disabled. Please note that if more than four counter-dependent applications are enabled on a port pipe, there is an impact on line rate performance.

The following counter-dependent applications are supported by FTOS:

- Egress VLAN
- Ingress VLAN
- Next Hop 2
- Next Hop 1
- Egress ACLs
- ILM
- IP FLOW
- IP ACL
- IP FIB
- L2 ACL
- L2 FIB

Clear interface counters

The counters in the **show interfaces** command are reset by the **clear counters** command. This command does not clear the counters captured by any SNMP program.

To clear the counters, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
clear counters [*interface*] [**vrrp** [*vrid*] | **learning-limit**]	EXEC Privilege	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. (OPTIONAL) Enter the following interface keywords and slot/port or number information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Loopback interface, enter the keyword **loopback** followed by a number from 0 to 16383. • For a Port Channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale. • For the management interface on the RPM, enter the keyword **ManagementEthernet** followed by slot/port information. The slot range is 0-1, and the port range is 0. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS. (OPTIONAL) Enter the keyword **vrrp** to clear statistics for all VRRP groups configured. Enter a number from 1 to 255 as the *vrid*. (OPTIONAL) Enter the keyword **learning-limit** to clear unknown source address (SA) drop counters when MAC learning limit is configured on the interface.

When you enter this command, you must confirm that you want FTOS to clear the interface counters for that interface (Figure 15-37).

Figure 15-37. Clearing an Interface

```
Force10#clear counters gi 0/0
Clear counters on GigabitEthernet 0/0 [confirm]
Force10#
```

16

IPv4 Addressing

IPv4 Addressing is supported on platforms C E S

FTOS supports various IP addressing features. This chapter explains the basics of Domain Name Service (DNS), Address Resolution Protocol (ARP), and routing principles and their implementation in FTOS.

- IP Addresses on page 329
- Directed Broadcast on page 334
- Resolution of Host Names on page 334
- ARP on page 337
- ICMP on page 341
- on page 342

Table 16-1 lists the defaults for the IP addressing features described in this chapter.

Table 16-1. IP Defaults

IP Feature	Default
DNS	Disabled
Directed Broadcast	Disabled
Proxy ARP	Enabled
ICMP Unreachable	Disabled
ICMP Redirect	Disabled

IP Addresses

FTOS supports IP version 4, as described in RFC 791. It also supports classful routing and Variable Length Subnet Masks (VLSM). With VLSM one network can be can configured with different masks. Supernetting, which increases the number of subnets, is also supported. Subnetting is when a mask is added 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 on IP addressing, refer to RFC 791, *Internet Protoco*l.

Implementation Information

In FTOS, you can configure any IP address as a static route except IP addresses already assigned to interfaces.

Note: FTOS 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. FTOS supports RFC 3021 with ARP.

Configuration Task List for IP Addresses

The following list includes the configuration tasks for IP addresses:

- Assign IP addresses to an interface on page 330 (mandatory)
- Configure static routes on page 332 (optional)
- Configure static routes for the management interface on page 333 (optional)

For a complete listing of all commands related to IP addressing, refer to *FTOS Command Line Interface Reference*.

Assign IP addresses to an interface

Assign primary and secondary IP addresses to physical or logical (for example, VLAN or port channel) interfaces to enable IP communication between the E-Series and hosts connected to that interface. In FTOS, you can assign one primary address and up to 255 secondary IP addresses to each interface.

To assign an IP address to an interface, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface** *interface*	CONFIGURATION	Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Loopback interface, enter the keyword **loopback** followed by a number from 0 to 16383. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/port information. The slot range is 0-1 and the port range is 0. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.
2	**no shutdown**	INTERFACE	Enable the interface.
3	**ip address** *ip-address mask* [**secondary**]	INTERFACE	Configure a primary IP address and mask on the interface. • *ip-address mask:* IP address must be in dotted decimal format (A.B.C.D) and the mask must be in slash prefix-length format (/24). Add the keyword **secondary** if the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

To view the configuration, use the **show config** command (Figure 246) in the INTERFACE mode or **show ip interface** in the EXEC privilege mode (Figure 247).

Figure 16-1. show config Command Example in the INTERFACE Mode

```
Force10(conf-if)#show conf
!
interface GigabitEthernet 0/0
 ip address 10.11.1.1/24
 no shutdown
!
Force10(conf-if)#
```

Figure 16-2. show ip interface Command Example

```
Force10#show ip int gi 0/8
GigabitEthernet 0/8 is up, line protocol is up
Internet address is 10.69.8.1/24
Broadcast address is 10.69.8.255
Address determined by config file
MTU is 1554 bytes
Inbound  access list is not set
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent


Force10#
```

Configure static routes

A static route is an IP address that is manually configured and not learned by a routing protocol, such as OSPF. Often static routes are used as backup routes in case other dynamically learned routes are unreachable.

To configure a static route, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip route *ip-address mask* {*ip-address* | *interface* [*ip-address*]} [*distance*] [**permanent**] [**tag** *tag-value*]	CONFIGURATION	Configure a static IP address. 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 followed by slot/port information. • *distance* range: 1 to 255 (optional). • **permanent:** Keep the static route in the routing table (if *interface* option is used) even if the interface with the route is disabled. (optional) • **tag** *tag-value* range: 1 to 4294967295. (optional)

You can enter as many static IP addresses as necessary.

To view the configured routes, use the **show ip route static** command.

Figure 16-3. show ip route static Command Example (partial)

```
Force10#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--
```

FTOS installs a next hop that is on the directly connected subnet of current IP address on the interface (for example, if interface gig 0/0 is on 172.31.5.0 subnet, FTOS installs the static route).

FTOS 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 gig 0/0 has ip address on subnet 2.2.2.0 and if 172.31.5.43 recursively resolves to 2.2.2.0, FTOS installs the static route.

- When interface goes down, FTOS withdraws the route.
- When interface comes up, FTOS re-installs the route.
- When recursive resolution is “broken,” FTOS withdraws the route.
- When recursive resolution is satisfied, FTOS re-installs the route.

Configure static routes for the management interface

When an IP address used by a protocol 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 in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
management route *ip-address mask* {*forwarding-router-address* | **ManagementEthernet** *slot/port*}	CONFIGURATION	Assign a static route to point to the management interface or forwarding router.

To view the configured static routes for the management port, use the **show ip management-route** command in the EXEC privilege mode.

Figure 16-4. show ip management-route Command Example

```
Force10>show ip management-route

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

Force10>
```

Directed Broadcast

By default, FTOS drops directed broadcast packets destined for an interface. This default setting provides some protection against Denial of Service (DOS) attacks.

To enable FTOS to receive directed broadcasts, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip directed-broadcast	INTERFACE	Enable directed broadcast.

To view the configuration, use the **show config** command in the 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 the feature is enabled, the system resolves only host names entered into the host table with the **ip host** command.

- Enable dynamic resolution of host names on page 334
- Specify local system domain and a list of domains on page 335
- DNS with traceroute on page 336

Enable dynamic resolution of host names

By default, dynamic resolution of host names (DNS) is disabled.

To enable DNS, use the following commands in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip domain-lookup	CONFIGURATION	Enable dynamic resolution of host names.
ip name-server *ip-address* [*ip-address2 ... ip-address6*]	CONFIGURATION	Specify up to 6 name servers. The order you entered the servers determines the order of their use.

To view current bindings, use the **show hosts** command.

Figure 16-5. show hosts Command Example

```
Force10>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
Force10>
```

To view the current configuration, use the **show running-config resolve** command.

Specify local system domain and a list of domains

If you enter a partial domain, FTOS 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. FTOS 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 FTOS cannot resolve the domain, it tries the domain name assigned to the local system. If that does not resolve the partial domain, FTOS searches the list of domains configured

To configure a domain name, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip domain-name *name*	CONFIGURATION	Enter up to 63 characters to configure one domain name for the E-Series.

To configure a list of domain names, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip domain-list *name*	CONFIGURATION	Enter up to 63 characters to configure names to complete unqualified host names. Configure this command up to 6 times to specify a list of possible domain names. FTOS searches the domain names in the order they were configured until a match is found or the list is exhausted.

DNS with traceroute

To configure your switch to perform DNS with traceroute, follow the steps below in the CONFIGURATION mode.

Command Syntax	Command Mode	Purpose
ip domain-lookup	CONFIGURATION	Enable dynamic resolution of host names.
ip name-server *ip-address* [*ip-address2 ... ip-address6*]	CONFIGURATION	Specify up to 6 name servers. The order you entered the servers determines the order of their use.
traceroute [*host* | *ip-address*]	CONFIGURATION	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). To keep the default setting for those parameters, press the ENTER key.

Figure 16-6 is an example output of DNS using the traceroute command.

Figure 16-6. Traceroute command example

```
Force10#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
------------------------------------------------------------------------------------------
```

ARP

FTOS uses two forms of address resolution: ARP and Proxy ARP.

Address Resolution Protocol (ARP) runs over Ethernet and enables endstations to learn the MAC addresses of neighbors on an IP network. Over time, FTOS 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 on ARP, see RFC 826, *An Ethernet Address Resolution Protocol*.

In FTOS, 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 on Proxy ARP, refer to RFC 925, *Multi-LAN Address Resolution,* and RFC 1027, *Using ARP to Implement Transparent Subnet Gateways.*

Configuration Task List for ARP

The following list includes configuration tasks for ARP:

- Configure static ARP entries on page 337 (optional)
- Enable Proxy ARP on page 338 (optional)
- Clear ARP cache on page 338 (optional)
- ARP Learning via Gratuitous ARP on page 339
- ARP Learning via ARP Request on page 340
- Configurable ARP Retries on page 341

For a complete listing of all ARP-related commands, refer to .

Configure 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 in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
arp *ip-address mac-address interface*	CONFIGURATION	Configure an IP address and MAC address mapping for an 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.

These entries do not age and can only be removed manually. To remove a static ARP entry, use the **no arp** *ip-address* command syntax.

To view the static entries in the ARP cache, use the **show arp static** command (Figure 253) in the EXEC privilege mode.

Figure 16-7. show arp static Command Example

```
Force10#show arp

Protocol    Address          Age(min)  Hardware Address    Interface  VLAN    CPU
--------------------------------------------------------------------------------
Internet    10.1.2.4               17  08:00:20:b7:bd:32   Ma 1/0      -      CP
Force10#
```

Enable Proxy ARP

By default, Proxy ARP is enabled. To disable Proxy ARP, use **no proxy-arp** command in the interface mode.

To re-enable Proxy ARP, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip proxy-arp	INTERFACE	Re-enable Proxy ARP.

To view if Proxy ARP is enabled on the interface, use the **show config** command in the 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.

Clear ARP cache

To clear the ARP cache of dynamically learnt ARP information, use the following command in the EXEC privilege mode:

Command Syntax	Command Mode	Purpose
clear arp-cache [*interface* | *ip ip-address*] [**no-refresh**]	EXEC privilege	Clear the ARP caches for all interfaces or for a specific interface by entering the following information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number between 1 and 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS. **ip** *ip-address* (OPTIONAL) Enter the keyword ip followed by the IP address of the ARP entry you wish to clear. **no-refresh** (OPTIONAL) Enter the keyword **no-refresh** to delete the ARP entry from CAM. Or use this option with *interface* or **ip** *ip-address* to specify which dynamic ARP entries you want to delete. **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 FTOS, the gratuitous ARP is a request. A Gratuitous ARP Request is an ARP request that is not needed according to the ARP specification, but one that hosts may send to:

- detect IP address conflicts
- inform switches of their presence on a port so that packets can be forwarded
- update the ARP table of other nodes on the network in case of an address change

In the request, the host uses its own IP address in the Sender Protocol Address and Target Protocol Address fields.

In FTOS versions prior to 8.3.1.0, if a gratuitous ARP is received some time after an ARP request is sent, only RP2 installs the ARP information. For example:

1. At time t=0 FTOS sends an ARP request for IP *A.B.C.D*
2. At time t=1 FTOS receives an ARP request for IP *A.B.C.D*
3. At time t=2 FTOS installs an ARP entry for *A.B.C.D* only on RP2.

Beginning with version 8.3.1.0, when a Gratuitous ARP is received, FTOS installs an ARP entry on all 3 CPUs.

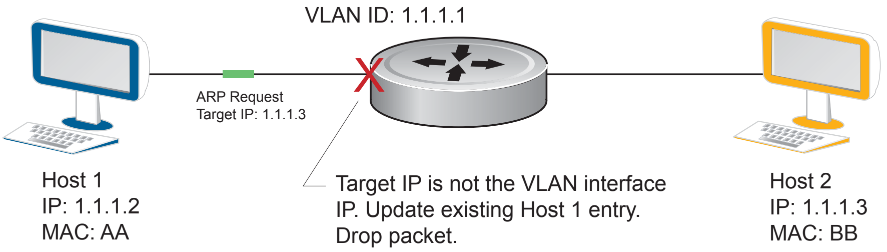
Task	Command Syntax	Command Mode
Enable ARP learning via gratuitous ARP.	**arp learn-enable**	CONFIGURATION

ARP Learning via ARP Request

In FTOS versions prior to 8.3.1.0, FTOS 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.

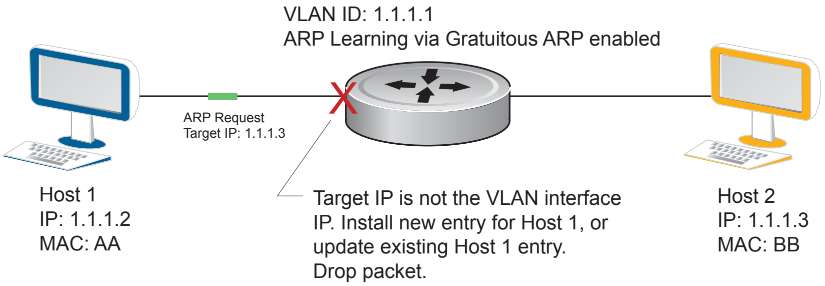
If the Target IP does not match the incoming interface, then the packet is dropped. If there is an existing entry for the requesting host, it is updated.

Figure 16-8. Learning via Gratuitous ARP



Beginning with FTOS version 8.3.1.0, when ARP Learning via Gratuitous ARP is enabled, the system installs a new ARP entry, or updates an existing entry for all received ARP requests.

Figure 16-9. Learning via Gratuitous ARP


VLAN ID: 1.1.1.1
ARP Learning via Gratuitous ARP enabled
ARP Request
Target IP: 1.1.1.3
Host 1
IP: 1.1.1.2
MAC: AA
Target IP is not the VLAN interface
IP. Install new entry for Host 1, or
update existing Host 1 entry.
Drop packet.
Host 2
IP: 1.1.1.3
MAC: BB

Whether ARP Learning via Gratuitous ARP is is enabled or disabled, 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.

Configurable ARP Retries

In FTOS versions prior to 8.3.1.0 the number of ARP retries is set to 5 and is not configurable. After 5 retries, FTOS backs off for 20 seconds before it sends a new request. Beginning with FTOS version 8.3.1.0, the number of ARP retries is configurable. The backoff interval remains at 20 seconds.

Task	Command Syntax	Command Mode
Set the number of ARP retries.	**arp retries** *number* Default: 5 Range: 5-20	CONFIGURATION
Display all ARP entries learned via gratuitous ARP.	**show arp retries**	EXEC Privilege

ICMP

For diagnostics, Internet Control Message Protocol (ICMP) provide 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 Task List for ICMP

Use the following steps to configure ICMP:

- Enable ICMP unreachable messages on page 341
- Enable ICMP redirects on page 342

See the for a complete listing of all commands related to ICMP.

Enable ICMP unreachable messages

By default, ICMP unreachable messages are disabled. When enabled ICMP unreachable messages are created and sent out all interfaces. To disable ICMP unreachable messages, use the **no ip unreachable** command.

To reenable the creation of ICMP unreachable messages on the interface, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip unreachable	INTERFACE	Set FTOS to create and send ICMP unreachable messages on the interface.

To view if ICMP unreachable messages are sent on the interface, use the **show config** command in the 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.

Enable ICMP redirects

Enable ICMP redirects is supported on E platform

By default, ICMP redirect messages are disabled. When enabled, ICMP redirect messages are created and sent out all interfaces. To disable ICMP redirect messages, use the **no ip redirect** command.

To reenable the creation of ICMP redirect messages on the interface, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip redirect	INTERFACE	Set FTOS to create and send ICMP redirect messages on the interface.

To view if ICMP redirect messages are sent on the interface, use the **show config** command in the 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.

Figure 16-10.

17

IPv6 Addressing

IPv6 Addressing, applies to platforms C E S

Note: The IPv6 basic commands are supported on all platforms. However, not all features are supported on all platforms, nor for all releases. See Table 17-2to determine the FTOS version supporting which features and platforms.

IPv6 (Internet Protocol Version 6) is the successor to IPv4. Due to the extremely 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 discussion of the differences between IPv4 and IPv6, and Dell Force10' support of IPv6. This chapter discusses the following, but is not intended to be a comprehensive discussion of IPv6.

- Protocol Overview on page 343
    - Extended Address Space
    - Stateless Autoconfiguration
    - IPv6 Headers
- Implementing IPv6 with FTOS on page 350
    - Table 17-2 FTOS and IPv6 Feature Support
    - ICMPv6
    - Path MTU Discovery
    - IPv6 Neighbor Discovery
    - QoS for IPv6
    - IPv6 Multicast
    - SSH over an IPv6 Transport
- Configuration Task List for IPv6 on page 356

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 IP addresses by using either the MAC address or a private random number to build its unique IP address.

Stateless auto-configuration 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 auto-configuration, network hosts can obtain their IPv6 addresses using Dynamic Host Control Protocol (DHCP) servers via stateful auto-configuration.

Note: FTOS provides the flexibility to add prefixes to advertise responses to RS messages. By default the RA response messages are not sent when an RS message is received. Enable the RA response messages with the **ipv6 nd prefix default** command in INTERFACE mode.

FTOS manipulation of IPv6 stateless auto-configuration 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 Neighbor Discovery (ND) messages are not used to create an IPv6 address.

The router redistribution functionality in Neighbor Discovery Protocol (NDP) is similar to IPv4 router redirect messages. Neighbor Discovery Protocol (NDP) uses ICMPv6 redirect messages (Type 137) to inform nodes that a better router exists on the link.

IPv6 Headers

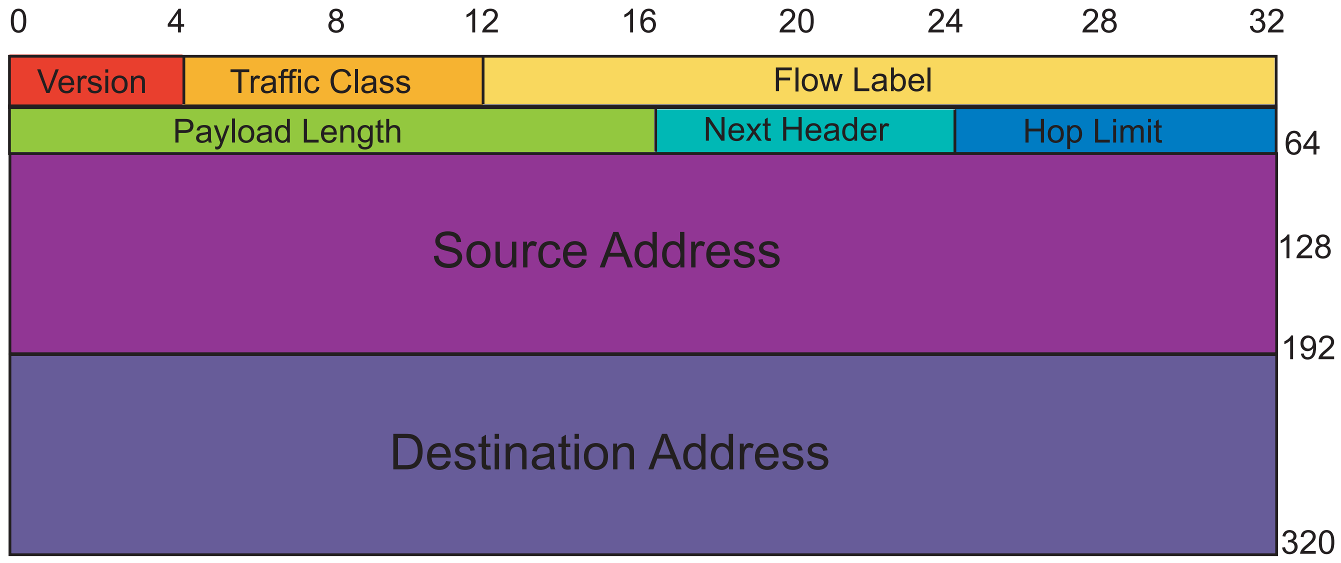
The IPv6 header has a fixed length of 40 bytes. This provides 16 bytes each for Source and Destination information, and 8 bytes for general header information. The IPv6 header includes the following fields:

- 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 show in Figure 17-1.

Figure 17-1. 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 head 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 (Table 17-1). If the next header is a TCP or 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.

Table 17-1. Next Header field values

Value	Description
0	Hop-by Hop option header following
4	IPv4
6	TCP
8	Exterior Gateway Protocol (EGP)
41	IPv6
43	Routing header
44	Fragmentation header
50	Encrypted Security
51	Authentication header

Table 17-1. Next Header field values

Value	Description
59	No Next Header
60	Destinations option header

Note: This is not a comprehensive table of Next Header field values. Refer to the Internet Assigned Numbers Authority (IANA) web page http://www.iana.org/assignments/protocol-numbers for a complete and current listing.

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 IP address for the packet originator.

Destination Address (128 bits)

The Destination Address field contains the intended recipient’s IP 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 bye 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 (Table 17-1).

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) (Table 17-1).

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 shown in Table 17-1.
- 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 1 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 zeros in a group can also be omitted (as in ::1 for localhost).

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. Since 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 IP addresses are manually assigned to a computer by an administrator. Dynamic IP addresses are assigned either randomly or by a server using Dynamic Host Configuration Protocol (DHCP). Even though IP 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 IP addresses. In this case, a DHCP server is used, but it is specifically configured to always assign the same IP address to a particular computer, and never to assign that IP address to another computer. This allows static IP 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 FTOS

FTOS supports both IPv4 and IPv6, and both may be used simultaneously in your system.

Note: Dell Force10 recommends that you use FTOS version 7.6.1.0 or later when implementing IPv6 functionality on an E-Series system.

Table 17-2 lists the FTOS Version in which an IPv6 feature became available for each platform. The sections following the table give some greater detail about the feature. Specific platform support for each feature or functionality is designated by the C E S symbols.

Table 17-2. FTOS and IPv6 Feature Support

Feature and/or Functionality	**FTOS Release Introduction**				**Documentation and Chapter Location**
	E-Series TeraScale	**E-Series ExaScale**	**C-Series**	**S-Series**	
Basic IPv6 Commands	7.4.1	8.2.1	7.8.1	7.8.1	IPv6 Basic Commands in the *FTOS Command Line Interface Reference Guide*
IPv6 Basic Addressing					
IPv6 address types: Unicast	7.4.1	8.2.1	7.8.1	7.8.1	Extended Address Space in this chapter
IPv6 neighbor discovery	7.4.1	8.2.1	7.8.1	7.8.1	IPv6 Neighbor Discovery in this chapter
IPv6 stateless autoconfiguration	7.4.1	8.2.1	7.8.1	7.8.1	Stateless Autoconfiguration in this chapter
IPv6 MTU path discovery	7.4.1	8.2.1	7.8.1	7.8.1	Path MTU Discovery in this chapter
IPv6 ICMPv6	7.4.1	8.2.1	7.8.1	7.8.1	ICMPv6 in this chapter
IPv6 ping	7.4.1	8.2.1	7.8.1	7.8.1	ICMPv6 in this chapter
IPv6 traceroute	7.4.1	8.2.1	7.8.1	7.8.1	ICMPv6 in this chapter
IPv6 Routing					
Static routing	7.4.1	8.2.1	7.8.1	7.8.1	Assign a Static IPv6 Route in this chapter
Route redistribution	7.4.1	8.2.1	7.8.1		OSPF, IS-IS, and IPv6 BGP chapters in the *FTOS Command Line Reference Guide*
Multiprotocol BGP extensions for IPv6	7.4.1	8.2.1	7.8.1		IPv6 BGP in the *FTOS Command Line Reference Guide*
IPv6 BGP MD5 Authentication	8.2.1.0	8.2.1.0	8.2.1.0		IPv6 BGP in the *FTOS Command Line Reference Guide*
OSPF for IPv6 (OSPFv3)	7.4.1	8.2.1	7.8.1		OSPFv3 in the *FTOS Command Line Reference Guide*
Equal Cost Multipath for IPv6	7.4.1	8.2.1	7.8.1	7.8.1	
IPv6 Services and Management					
Telnet client over IPv6 (outbound Telnet)	7.5.1	8.2.1	7.8.1	7.8.1	Telnet with IPv6 in this chapter Control and Monitoring in the *FTOS Command Line Reference Guide*
Telnet server over IPv6 (inbound Telnet)	7.4.1	8.2.1	7.8.1	7.8.1	Telnet with IPv6 in this chapter Control and Monitoring in the *FTOS Command Line Reference Guide*

Table 17-2. FTOS and IPv6 Feature Support

Secure Shell (SSH) client support over IPv6 (outbound SSH) Layer 3 only	7.5.1	8.2.1	7.8.1	7.8.1	SSH over an IPv6 Transport in this chapter
Secure Shell (SSH) server support over IPv6 (inbound SSH) Layer 3 only	7.4.1	8.2.1	7.8.1	7.8.1	SSH over an IPv6 Transport in this chapter
IPv6 Access Control Lists	7.4.1	8.2.1	7.8.1	8.2.1.0	IPv6 Access Control Lists in the *FTOS Command Line Reference Guide*
IPv6 Multicast					
PIM-SM for IPv6	7.4.1	8.2.1			IPv6 Multicast in this chapter; IPv6 PIM in the *FTOS Command Line Reference Guide*
PIM-SSM for IPv6	7.5.1	8.2.1			IPv6 Multicast in this chapter IPv6 PIM in the *FTOS Command Line Reference Guide*
MLDv1/v2	7.4.1	8.2.1			IPv6 Multicast in this chapter Multicast IPv6 in the *FTOS Command Line Reference Guide*
MLDv1 Snooping	7.4.1	8.2.1			IPv6 Multicast in this chapter Multicast IPv6 in the *FTOS Command Line Reference Guide*
MLDv2 Snooping	8.3.1.0	8.3.1.0			IPv6 Multicast in this chapter Multicast IPv6 in the *FTOS Command Line Reference Guide*
IPv6 QoS					
trust DSCP values	7.4.1	8.2.1			QoS for IPv6 in this chapter

ICMPv6

ICMPv6 is supported on platforms C E S

ICMP for IPv6 combines the roles of ICMP, IGMP and ARP in IPv4. Like IPv4, it provides functions for reporting delivery and forwarding errors, and provides a simple echo service for troubleshooting. The FTOS implementation of ICMPv6 is based on RFC 2463.

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 FTOS `ping` and `traceroute` commands extend to support IPv6 addresses. These commands use ICMPv6 Type-2 messages.

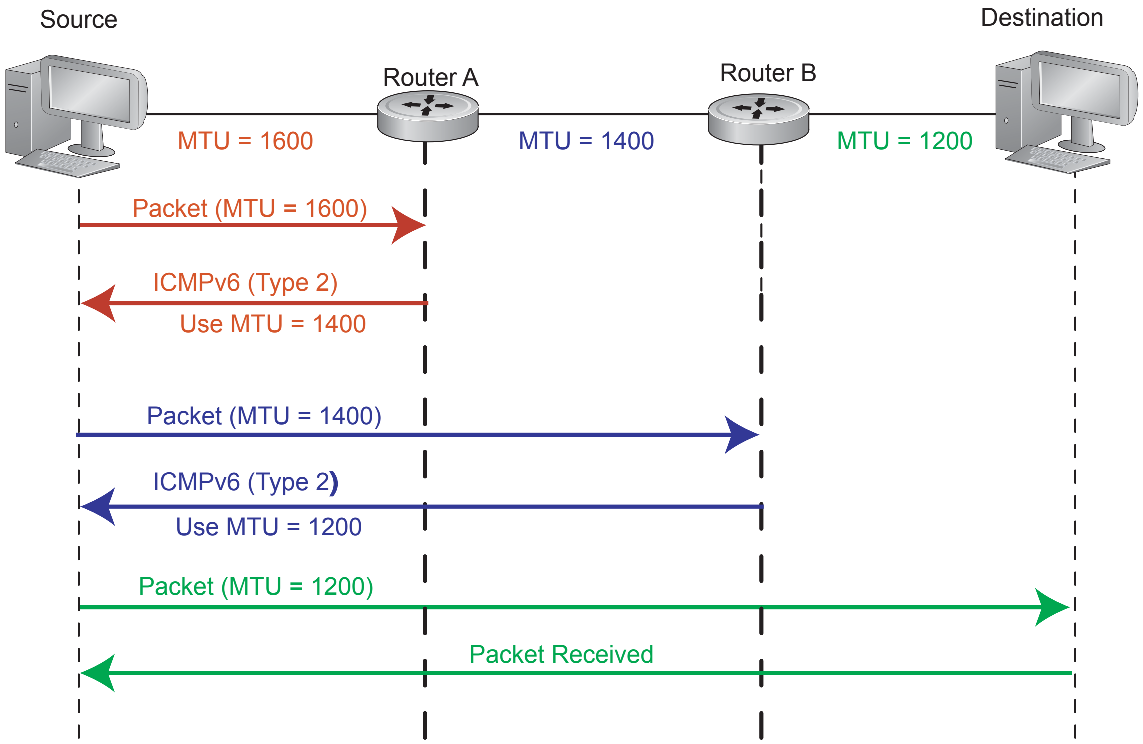
Path MTU Discovery

IPv6 MTU Discovery is supported on platforms C E S

Path MTU (Maximum Transmission Unit) 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 (headers, for example) or underlying per-packet delays remain fixed.

Figure 17-2. MTU Discovery Path



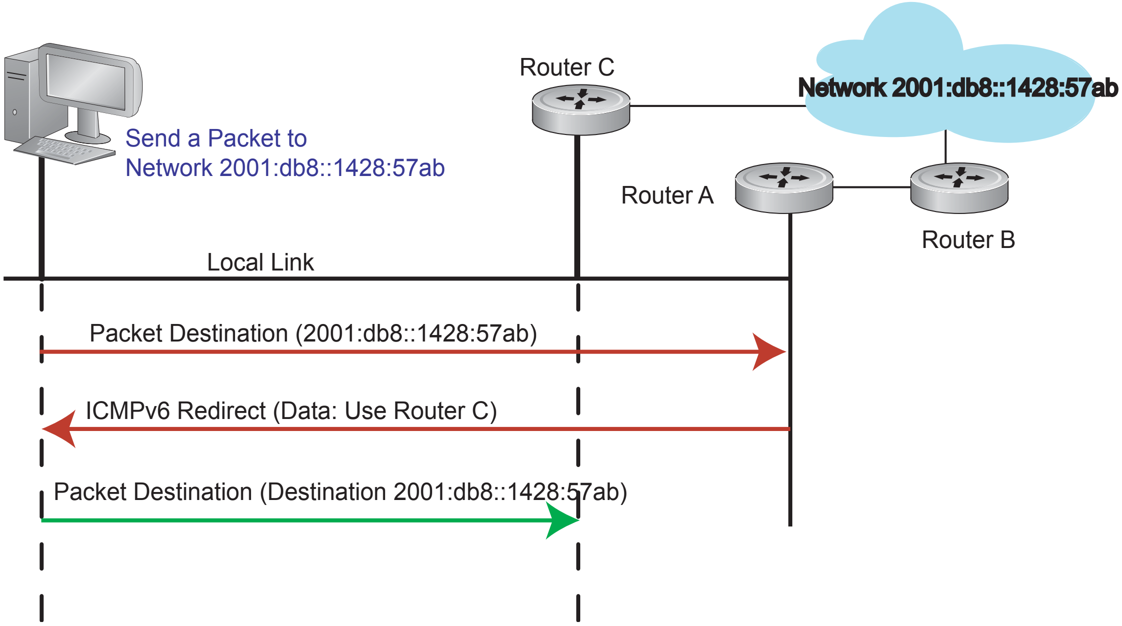
IPv6 Neighbor Discovery

IPv6 NDP is supported on platforms C E S

Neighbor Discovery Protocol (NDP) is a top-level protocol for neighbor discovery on an IPv6 network. In lieu of 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.

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 17-3. NDP Router Redistribution


Router C
Network 2001:db8::1428:57ab
Send a Packet to
Network 2001:db8::1428:57ab
Router A
Router B
Local Link
Packet Destination (2001:db8::1428:57ab)
ICMPv6 Redirect (Data: Use Router C)
Packet Destination (Destination 2001:db8::1428:57ab)

IPv6 Neighbor Discovery of MTU packets

With FTOS 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 **ip nd mtu** command sets the value advertised to routers. It does not set the actual MTU rate. For example, if **ip nd mtu** is set to 1280, the interface will still pass 1500-byte packets, if that is what is set with the **mtu** command.

QoS for IPv6

IPv6 QoS is supported on platforms E

FTOS IPv6 supports quality of service based on DSCP field. You can configure FTOS to honor the DSCP value on incoming routed traffic and forward the packets with the same value.

IPv6 Multicast

IPv6 Multicast is supported only on platform E

FTOS 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, and MLD version 2 is based on version 3 of the IGMP for IPv4. IPv6 multicast for FTOS supports versions 1 and 2
- PIM-SM. Protocol-Independent Multicast-Sparse Mode (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.
- PIM in Source Specific Multicast (PIM-SSM). PIM-SSM protocol is based on the source specific model for forwarding Multicast traffic across multiple domains on the Internet. It is restricted to shortest path trees (SPTs) to specific sources described by hosts using MLD. PIM-SSM is essentially a subset of PIM-SM protocol, which has the capability to join SPTs. The only difference being register states and shared tree states for Multicast groups in SSM range are not maintained. End-hosts use MLD to register their interest in a particular source-group (S,G) pair. PIM-SSM protocol interacts with MLD to construct the multicast forwarding tree rooted at the source S.

Refer to *FTOS Command Line Interface Reference* document chapters Multicast IPv6, and Protocol Independent Multicast (IPv6) for configuration details.

SSH over an IPv6 Transport

IPv6 SSH is supported on platforms C E S

FTOS 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.

Refer to the Security Commands chapter in the *FTOS Command Line Interface Reference* document for SSH configuration details.

Configuration Task List for IPv6

This section contains information regarding the following:

- Change your CAM-Profile on an E-Series system (mandatory)
- Adjust your CAM-Profile on an C-Series or S-Series
- Assign an IPv6 Address to an Interface
- Assign a Static IPv6 Route
- Telnet with IPv6
- SNMP over IPv6
- Show IPv6 Information
- Clear IPv6 Routes

Change your CAM-Profile on an E-Series system

The **cam-profile** command is supported only on platform E

Change your CAM profile to the CAM ipv6-extacl before doing any further IPv6 configuration. Once the CAM profile is changed, save the configuration and reboot your router.

Command Syntax	Command Mode	Purpose
cam-profile ipv6-extacl microcode ipv6-extacl *chassis* | *linecard slot*	EXEC Privileged	Enable the CAM profile with IPv6 extended ACLs on the entire chassis or on a specific linecard *chassis* changes the CAM profile for all linecards in the chassis *linecard slot/port* changes the CAM profile only for the specified slot

Figure 17-4 displays the IPv6 CAM profile summary for a chassis that already has IPv6 CAM profile configured. Figure 17-5 shows the full IPv6 CAM profiles. Refer to Chapter 10, Content Addressable Memory, on page 219 for complete information regarding CAM configuration.

Figure 17-4. Command Example: show cam-profile summary **(E-Series)**

```
Force10#show cam-profile summary

-- Chassis CAM Profile --
                  : Current Settings : Next Boot
Profile Name      : IPV6-ExtACL      : IPV6-ExtACL
MicroCode Name    : IPv6-ExtACL      : IPv6-ExtACL

-- Line card 1 --
                  : Current Settings : Next Boot
Profile Name      : IPV6-ExtACL      : IPV6-ExtACL
MicroCode Name    : IPv6-ExtACL      : IPv6-ExtACL

Force10#
```

Figure 17-5. Command Example: show cam profile **(E-Series**)

```
Force10#show cam-profile

-- Chassis CAM Profile --

CamSize          : 18-Meg
                 : Current Settings : Next Boot
Profile Name     : IPV6-ExtACL      : IPV6-ExtACL
L2FIB            : 32K entries      : 32K entries
L2ACL            : 1K entries       : 1K entries
IPv4FIB          : 192K entries     : 192K entries
IPv4ACL          : 12K entries      : 12K entries
IPv4Flow         : 8K entries       : 8K entries
EgL2ACL          : 1K entries       : 1K entries
EgIPv4ACL        : 1K entries       : 1K entries
Reserved         : 2K entries       : 2K entries
IPv6FIB          : 6K entries       : 6K entries
IPv6ACL          : 3K entries       : 3K entries
IPv6Flow         : 4K entries       : 4K entries
EgIPv6ACL        : 1K entries       : 1K entries
MicroCode Name   : IPv6-ExtACL      : IPv6-ExtACL

-- Line card 1 --
CamSize          : 18-Meg
                 : Current Settings : Next Boot
--More--
```

Adjust your CAM-Profile on an C-Series or S-Series

The **cam-acl** command is supported on platforms C S

Although this is not a mandatory step, if you plan to implement IPv6 ACLs, you must adjust your CAM settings.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. Note that there are 16 FP blocks, but the System Flow requires 3 blocks that cannot be reallocated.

The **ipv6acl** allocation 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.

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

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.

Command Syntax	Command Mode	Purpose
cam-acl { **ipv6acl** }	CONFIGURATION	Allocate space for IPV6 ACLs. Enter the CAM profile name followed by the amount to be allotted. When not selecting the default option, you must 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 cam-acl	EXEC EXEC Privilege	Show the current CAM settings.

Assign an IPv6 Address to an Interface

IPv6 Addresses are supported on platforms C E S

Essentially IPv6 is enabled in FTOS simply by assigning IPv6 addresses to individual router interfaces. IPv6 and IPv4 can be used together on a system, but be sure to differentiate that usage carefully. Use the `ipv6 address` command to assign an IPv6 address to an interface.

Command Syntax	Command Mode	Purpose
`ipv6 address` *ipv6 address/ mask*	CONFIG-INTERFACE	Enter the IPv6 Address for the device. *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128
	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 earlier in this chapter.	

Assign a Static IPv6 Route

IPv6 Static Routes are supported on platforms C E S

Use the `ipv6 route` command to configure IPv6 static routes.

Command Syntax	Command Mode	Purpose
`ipv6 route` *prefix type {slot/ port} forwarding router tag*	CONFIGURATION	Set up IPv6 static routes *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` followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword `GigabitEthernet` followed by the slot/ port information. • For a Gigabit Ethernet interface, enter the keyword `GigabitEthernet` followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword `TenGigabitEthernet` followed by the slot/port information. • For a loopback interface, enter the keyword `loopback` followed by the loopback number • For a linecard interface, enter the keyword `linecard` followed by the slot number • For a port-channel interface, enter the keyword `port-channel` followed by the port-channel number • For a VLAN interface, enter the keyword `vlan` followed by the VLAN ID • For a Null interface, enter the keyword `null` followed by the Null interface number

Telnet with IPv6

IPv6 Telnet is supported on platforms C E S

The Telnet client and server in FTOS support IPv6 connections. You can establish a Telnet session directly to the router using an IPv6 Telnet client, or an IPv6 Telnet connection can be initiated from the router.

Note: Telnet to link local addresses is not supported.

Command Syntax	Command Mode	Purpose
telnet *ipv6 address*	EXEC *or* EXEC Privileged	Enter the IPv6 Address for the device. *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128
	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 earlier in this chapter.	

SNMP over IPv6

SNMP is supported on platforms C E S

Simple Network Management Protocol (SNMP) can be configured over IPv6 transport so that an IPv6 host can perform SNMP queries and receive SNMP notifications from a device running FTOS IPv6. The FTOS SNMP-server commands for IPv6 have been extended to support IPv6. Refer to the *SNMP and SYSLOG* chapter in the *FTOS Command Line Interface Reference* for more information regarding SNMP commands.

- 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

Show IPv6 Information

All of the following show commands are supported on platforms C E S

View specific IPv6 configuration with the following commands.

Command Syntax	Command Mode	Purpose
show ipv6 ?	EXEC *or* EXEC Privileged	List the IPv6 show options

Command Syntax	Command Mode	Purpose

```
Force10#show ipv6 ?
accounting      IPv6 accounting information
cam linecard    IPv6 CAM Entries for Line Card
fib linecard    IPv6 FIB Entries for Line Card
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
Force10#
```

Show an IPv6 Interface

View the IPv6 configuration for a specific interface with the following command.

Command Syntax	Command Mode	Purpose
show ipv6 interface *type {slot/ port}*	EXEC	Show the currently running configuration for the specified interface Enter the keyword interface followed by 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 followed by the slot/ port information. • For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/ port information. • For a 10 Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information. • For a loopback interface, enter the keyword loopback followed by the loopback number • For a linecard interface, enter the keyword linecard followed by the slot number • For a port-channel interface, enter the keyword port-channel followed by the port-channel number • For a VLAN interface, enter the keyword vlan followed by the VLAN ID

Figure 17-6 illustrates the show ipv6 interface command output.

Figure 17-6. Command Example: show ipv6 interface

```
Force10#show ipv6 interface gi 2/2
GigabitEthernet 2/2 is down, line protocol is down
  IPV6 is enabled
  Link Local address: fe80::201:e8ff:fe06:95a3
  Global Unicast address(es):
    3:4:5:6::8, subnet is 3::/24
  Global Anycast address(es):
  Joined Group address(es):
    ff02::1
    ff02::2
    ff02::1:ff00:8
    ff02::1:ff06:95a3
  MTU is 1500
  ICMP redirects are not sent
  DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30 seconds
  ND advertised reachable time is 30 seconds
  ND advertised retransmit interval is 30 seconds
  ND router advertisements are sent every 200 seconds
  ND router advertisements live for 1800 seconds
```

Show IPv6 Routes

View the global IPv6 routing information with the following command.

Command Syntax	Command Mode	Purpose
show ipv6 route *type*	EXEC	Show IPv6 routing information for the specified route type. Enter the keyword: • To display information about a network, enter the ipv6 address (X:X:X:X::X). • To display information about a host, enter the 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.

Figure 17-7 illustrates the show ipv6 route command output.

Figure 17-7. Command Example: show ipv6 route

```
Force10#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
```

Figure 17-8 illustrates the show ipv6 route summary command output.

Figure 17-8. Command Example: show ipv6 route summary

```
Force10#show ipv6 route summary

Route Source              Active Routes   Non-active Routes
connected                 5               0
static                    0               0
```

Figure 17-9 illustrates the show ipv6 route static command output.

Figure 17-9. Command Example: show ipv6 route static

```
Force10#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
```

Show the Running-Configuration for an Interface

View the configuration for any interface with the following command.

Command Syntax	Command Mode	Purpose
show running-config interface *type {slot/port}*	EXEC	Show the currently running configuration for the specified interface Enter the keyword interface followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information. • For the Management interface on the RPM, enter the keyword ManagementEthernet followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information.

Figure 17-10 illustrates the show running-config command output. Note the IPv6 address listed.

Figure 17-10. Command Example: show running-config interface

```
Force10#show run int gi 2/2
!
interface GigabitEthernet 2/2
 no ip address
 ipv6 address 3:4:5:6::8/24
 shutdown
Force10#
```

Clear IPv6 Routes

Use the clear IPv6 route command to clear routes from the IPv6 routing table.

Command Syntax	Command Mode	Purpose
clear ipv6 route {* | *ipv6 address prefix-length*}	EXEC	Clear (refresh) all or a specific routes from the IPv6 routing table. * : all routes *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128
	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 earlier in this chapter.	

18

iSCSI Optimization

This chapter describes how to detect and configure switchports for Dell Compellent arrays. The topics covered in this chapter include:

- iSCSI Optimization Overview
- Detection and Port Configuration for Dell Compellent Arrays

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).

The following iSCSI feature is available on platforms S55 S60 S4810

- Manual configuration to detect Compellent storage arrays where auto-detection is not supported.

Detection and Port Configuration for Dell Compellent Arrays

This feature is available on platforms S55 S60 S4810

Switches support the iscsi profile-compellent command to configure a port connected to a Dell Compellent storage array. The command configures a port for the best iSCSI traffic conditions and must be entered in INTERFACE Configuration mode.

The following message is displayed the first time you use the **iscsi profile-compellent** command to configure a port connected to a Dell Compellent storage array and describes the configuration changes that are automatically performed:

```
%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_AUTO_CONFIG: This switch is being configured for optimal
conditions to support iSCSI traffic which will cause some automatic configuration to occur
including jumbo frames and flow-control on all ports; no storm control and spanning-tree port
fast to be enabled on the port of detection.
```

After you execute the iscsi profile-compellent command, the following actions occur:

- Jumbo frame size is set to 12000 for the S4810 and 9252 for S55 and S60 on all ports and port-channels, if it is not already enabled.
- Spanning-tree portfast is enabled on the interface.
- Unicast storm control is disabled on the interface.

You must enter the iscsi profile-compellent command in INTERFACE configuration mode. For example:

```
FTOS(conf-if-te-o/50# iscsi profile-compellent)
```

Auto-detection of Dell Compellent

To auto-detect iSCSI optimization on a switch connected to a Dell Compellent array, follow these steps:

Step	Task	Command	Command Mode
1	Configure the auto-detection of Dell Compellent arrays on a port. Default: Dell Compellent disk arrays are not detected.	[no] iscsi profile-compellent	INTERFACE

19

Link Aggregation Control Protoco

Link Aggregation Control Protoco is supported on platforms C E S

The major sections in the chapter are:

Introduction to Dynamic LAGs and LACP

A *Link Aggregation Group* (*LAG*), referred to as a *port channel* by FTOS, can provide both load-sharing and port redundancy across line cards. LAGs can be enabled as static or dynamic. The benefits and constraints are basically the same, as described in Port Channel Interfaces on page 294 in Chapter 15, Interfaces.

The unique benefit of a dynamic LAG is that its ports can toggle between participating in the LAG or acting as dedicated ports, whereas ports in a static LAG must be specifically removed from the LAG in order to act alone.

FTOS uses LACP to create dynamic LAGs. LACP provides a standardized means of exchanging information between two systems (also called Partner Systems) and automatically establishes the LAG between the systems. LACP permits the exchange of messages on a link to allow their LACP instances to:

- Reach agreement on the identity of the LAG to which the link belongs.
- Move the link to that LAG.
- Enable the transmission and reception functions in an orderly manner.

The FTOS 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 PDUs across LAN Ethernet links. The protocol packets are only exchanged between ports that are configured as LACP capable.

Important Points to Remember

- LACP enables you to add members to a port channel (LAG) as long as it has no static members. Conversely, if the LAG already contains a statically defined member (**channel-member** command), the **port-channel mode** command is not permitted.
- A static LAG cannot be created if a dynamic LAG using the selected number already exists.
- **No dual membership in static and dynamic LAGs**:
    - If a physical interface is a part of a static LAG, then the command **port-channel-protocol lacp** will be rejected on that interface.
    - If a physical interface is a part of a dynamic LAG, it cannot be added as a member of a static LAG. The command **channel-member gigabitethernet** *x/y* will be rejected in the static LAG interface for that physical interface.
- A dynamic LAG can be created with any type of configuration.
- There is a difference between the **shutdown** and **no interface port-channel**:
    - — The **shutdown** command on LAG “xyz” disables the LAG and retains the user commands. However, the system does not allow the channel number “xyz” to be statically created.
    - — The command **no interface port-channel** *channel-number* deletes the specified LAG, including a dynamically created LAG. This command causes all LACP-specific commands on the member interfaces to be removed. The interfaces are restored to a state that is ready to be configured.

  Note: There will be no configuration on the interface since that condition is required for an interface to be part of a LAG.
- Link dampening can be configured on individual members of a LAG. See Link Debounce Timer on page 312 for more information.

LACP modes

FTOS provides the following three modes for configuration of LACP:

- **Off**—In this state, an interface is not capable of being part of a dynamic LAG. LACP does not run on any port that is configured to be in this state.
- **Active**—In this state, the interface is said to be in the “active negotiating state.” LACP runs on any link that is configured to be in this state. A port in Active state also automatically initiates negotiations with other ports by initiating LACP packets.
- **Passive**—In this state, the interface is not in an active negotiating state, but LACP will run on the link. A port in Passive state also responds to negotiation requests (from ports in Active state). Ports in Passive state respond to LACP packets.

FTOS supports LAGs in the following cases:

- A port in Active state can set up a port channel (LAG) with another port in Active state.
- A port in Active state can set up a LAG with another port in Passive state.

A port in Passive state cannot set up a LAG with another port in Passive state.

LACP Configuration Commands

If aggregated ports are configured with compatible LACP modes (Off, Active, Passive), LACP can automatically link them, as defined in IEEE 802.3, Section 43. The following commands configure LACP:

Command Syntax	Command Mode	Purpose
[**no**] **lacp system-priority** *priority-value*	CONFIGURATION	Configure the system priority. Range: 1– 65535 (the higher the number, the lower the priority) Default: 32768
[**no**] **port-channel-protocol lacp**	INTERFACE	Enable or disable LACP on any LAN port: • Default is “LACP disabled” • This command creates a new context.
[**no**] **port-channel** *number* **mode** [**active** | **passive** | **off**]	LACP	Configure LACP mode. • Default is “LACP active” • **number** cannot statically contain any links
[**no**] **lacp port-priority** *priority-value*	LACP	Configure port priority. • Ranges: 1 – 65535 (the higher the number, the lower the priority) • Default: 32768

LACP Configuration Tasks

The tasks covered in this section are:

- Create a LAG
- Configure the LAG interfaces as dynamic on page 372
- Set the LACP long timeout on page 372
- Monitor and Debugging LACP on page 373
- Configure Shared LAG State Tracking on page 374

Create a LAG

To create a dynamic port channel (LAG), define the LAG and then the LAG interfaces. Use the **interface port-channel** and **switchport** commands, as shown in Figure 19-1, which uses the example of LAG 32:

Figure 19-1. Placing a LAG into the Default VLAN

```
Force10(conf)#interface port-channel 32
Force10(conf-if-po-32)#no shutdown
Force10(conf-if-po-32)#switchport
```

The LAG is in the default VLAN. To place the LAG into a non-default VLAN, use the **tagged** command on the LAG (Figure 19-2):

Figure 19-2. Placing a LAG into a Non-default VLAN

```
Force10(conf)#interface vlan 10
Force10(conf-if-vl-10)#tagged port-channel 32
```

Configure the LAG interfaces as dynamic

After creating a LAG, configure the dynamic LAG interfaces. Figure 19-3 shows ports 3/15, 3/16, 4/15, and 4/16 added to LAG 32 in LACP mode with the command **port-channel-protocol lacp**.

Figure 19-3. Creating a Dynamic LAG Example

```
Force10(conf)#interface Gigabitethernet 3/15
Force10(conf-if-gi-3/15)#no shutdown
Force10(conf-if-gi-3/15)#port-channel-protocol lacp
Force10(conf-if-gi-3/15-lacp)#port-channel 32 mode active
...
Force10(conf)#interface Gigabitethernet 3/16
Force10(conf-if-gi-3/16)#no shutdown
Force10(conf-if-gi-3/16)#port-channel-protocol lacp
Force10(conf-if-gi-3/16-lacp)#port-channel 32 mode active
...
Force10(conf)#interface Gigabitethernet 4/15
Force10(conf-if-gi-4/15)#no shutdown
Force10(conf-if-gi-4/15)#port-channel-protocol lacp
Force10(conf-if-gi-4/15-lacp)#port-channel 32 mode active
...
Force10(conf)#interface Gigabitethernet 4/16
Force10(conf-if-gi-4/16)#no shutdown
Force10(conf-if-gi-4/16)#port-channel-protocol lacp
Force10(conf-if-gi-4/16-lacp)#port-channel 32 mode active
```

The **port-channel 32 mode active** command shown above may be successfully issued as long as there is no existing static channel-member configuration in LAG 32.

Set the LACP long timeout

PDUs are exchanged between port channel (LAG) interfaces to maintain LACP sessions. PDUs are transmitted at either a slow or fast transmission rate, depending upon the LACP timeout value. The timeout value is the amount of time that a LAG interface waits for a PDU from the remote system before bringing the LACP session down. The default timeout value is 1 second; it can be configured to be 30 seconds. Invoking the longer timeout might prevent the LAG from flapping if the remote system is up but temporarily unable to transmit PDUs due to a system interruption.

Note: The 30-second timeout is available for dynamic LAG interfaces only. The **lacp long-timeout** command can be entered for static LAGs, but it has no effect.

To configure the LACP long timeout (Figure 196):

Step	Task	Command Syntax	Command Mode
1	Set the LACP timeout value to 30 seconds.	**lacp long-timeout**	CONFIG-INT-PO

Figure 19-4. Invoking the LACP Long Timeout

```
Force10(conf)# interface port-channel 32
Force10(conf-if-po-32)#no shutdown
Force10(conf-if-po-32)#switchport
Force10(conf-if-po-32)#lacp long-timeout
Force10(conf-if-po-32)#end
Force10# show lacp 32
Port-channel 32 admin up, oper up, mode lacp
Actor System ID: Priority 32768, Address 0001.e800.a12b
Partner System ID: Priority 32768, Address 0001.e801.45a5
Actor Admin Key 1, Oper Key 1, Partner Oper Key 1
LACP LAG 1 is an aggregatable link
A - Active LACP, B - Passive LACP, C - Short Timeout, D - Long Timeout
E - Aggregatable Link, F - Individual Link, G - IN_SYNC, H - OUT_OF_SYNC
I - Collection enabled, J - Collection disabled, K - Distribution enabled L - Distribution disabled,
M - Partner Defaulted, N - Partner Non-defaulted, O - Receiver is in expired state,
P - Receiver is not in expired state
Port Gi 10/6 is enabled, LACP is enabled and mode is lacp
Actor Admin: State ADEHJLMP Key 1 Priority 128
```

Note: View PDU exchanges and the timeout value using the command **debug lacp**. See Monitor and Debugging LACP on page 373.

Monitor and Debugging LACP

The system log (syslog) records faulty LACP actions.

To debug LACP, use the following command:

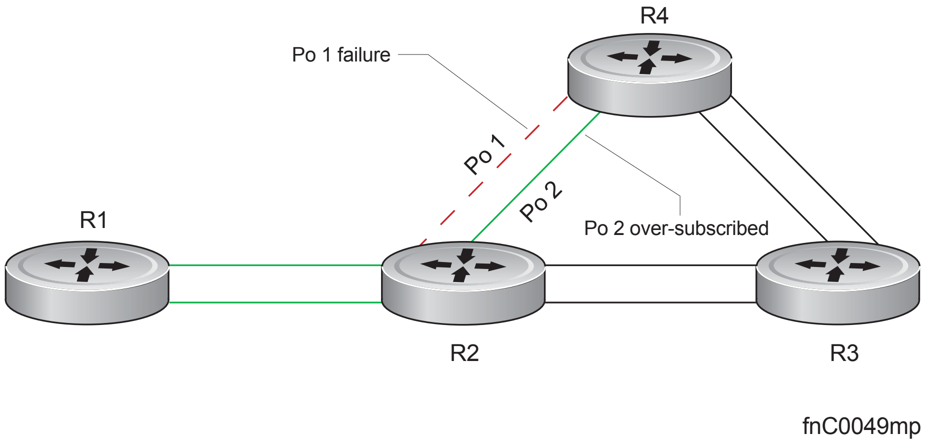
Command Syntax	Command Mode	Purpose
[**no**] **debug lacp** [**config** | **events** | **pdu** [**in** | **out** | [*interface* [**in** | **out**]]]]	EXEC	Debug LACP, including configuration and events.

Shared LAG State Tracking

Shared LAG State Tracking provides the flexibility to bring down a port channel (LAG) based on the operational state of another LAG. At any time, only two LAGs can be a part of a group such that the fate (status) of one LAG depends on the other LAG.

In Figure 19-5, line-rate traffic from R1 destined for R4 follows the lowest-cost route via R2, as shown. Traffic is equally distributed between LAGs 1 and 2. If LAG 1 fails, all traffic from R1 to R4 flows across LAG 2 only. This condition over-subscribes the link, and packets are dropped.

Figure 19-5. LAGs using ECMP without Shared LAG State Tracking



To avoid packet loss, traffic must be re-directed through the next lowest-cost link (R3 to R4). FTOS has the ability to bring LAG 2 down in the event that LAG 1 fails, so that traffic can be re-directed, as described. This is what is meant by Shared LAG State Tracking. To achieve this functionality, you must group LAG 1 and LAG 2 into a single entity, called a *failover group*.

Configure Shared LAG State Tracking

To configure Shared LAG State Tracking, you configure a failover group:

Step	Task	Command	Command Mode
1	Enter port-channel failover group mode.	**port-channel failover-group**	CONFIGURATION
2	Create a failover group and specify the two port-channels that will be members of the group.	**group** *number* **port-channel** *number* **port-channel** *number*	CONFIG-PO-FAILOVER-GRP

In Figure 19-6, LAGs 1 and 2 have been placed into to the same failover group.

Figure 19-6. Configuring Shared LAG State Tracking

```
R2#config
R2(conf)#port-channel failover-group
R2(conf-po-failover-grp)#group 1 port-channel 1 port-channel 2
```

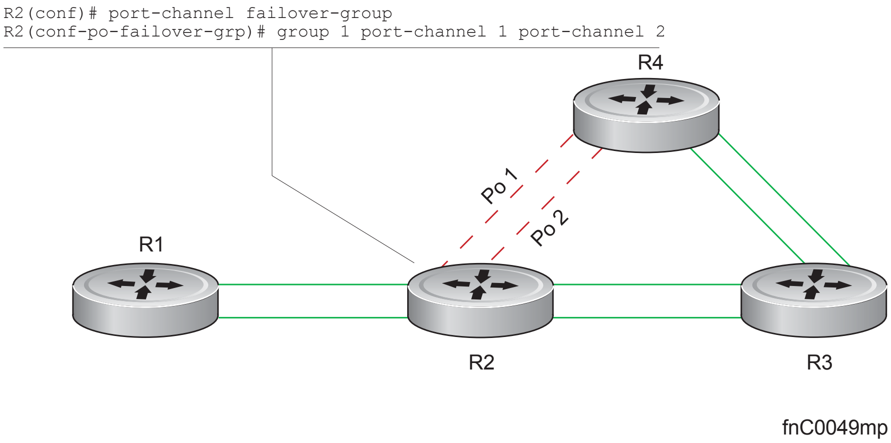
View the failover group configuration using the **show running-configuration po-failover-group** command, as shown in Figure 19-7.

Figure 19-7. Viewing Shared LAG State Tracking in the Running-configuration

```
R2#show running-config po-failover-group
!
port-channel failover-group
 group 1 port-channel 1 port-channel 2
```

In Figure 19-8, LAGs 1 and 2 are members of a failover group. LAG 1 fails and LAG 2 is brought down upon the failure. This effect is logged by Message 1, in which a console message declares both LAGs down at the same time.

Figure 19-8. Shared LAG State Tracking



```
R2(conf)# port-channel failover-group
R2(conf-po-failover-grp)# group 1 port-channel 1 port-channel 2
```

Message 1 Shared LAG State Tracking Console Message

```
2d1h45m: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 1
2d1h45m: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 2
```

View the status of a failover group member using the command **show interface port-channel,** as shown in Figure 19-9.

Figure 19-9. Viewing Status of a Failover Group Member

```
R2#show interface Port-channel 2
Port-channel 2 is up, line protocol is down (Failover-group 1 is down)
Hardware address is 00:01:e8:05:e8:4c, Current address is 00:01:e8:05:e8:4c
Interface index is 1107755010
Minimum number of links to bring Port-channel up is 1
Port-channel is part of failover-group 1
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit
Members in this channel:  Gi 1/17(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:01:28
Queueing strategy: fifo
```

Note: The set of console messages shown in Message 1 appear only if Shared LAG State Tracking is configured on that router (the feature can be configured on one or both sides of a link). For example, in Figure 19-8, if Shared LAG State Tracking is configured on R2 only, then no messages appear on R4 regarding the state of LAGs in a failover group.

Important Points about Shared LAG State Tracking

- This feature is available for static and dynamic LAGs.
- Only a LAG can be a member of a failover group.
- Shared LAG State Tracking can be configured on one side of a link or on both sides.
- If a LAG that is part of a failover group is deleted, the failover group is deleted.
- If a LAG moves to the down state due to this feature, its members may still be in the up state.

Configure LACP as Hitless

Configure LACP as Hitless is supported only on platforms: C E

LACP on Dell Force10 systems can be configured to be hitless. When configured as hitless, there is no noticeable impact on dynamic LAG state upon an RPM failover. Critical LACP state information is synchronized between the two RPMs.

Configure LACP to be hitless using the command **redundancy protocol lacp** from CONFIGURATION mode, as shown in Figure 19-10.

Figure 19-10. Enabling Hitless LACP

```
Force10(conf)#redundancy protocol lacp


Force10#show running-config redundancy
!
redundancy protocol lacp
Force10#
Force10#show running-config interface gigabitethernet 0/12
!
interface GigabitEthernet 0/12
 no ip address
!
 port-channel-protocol LACP
  port-channel 200 mode active
 no shutdown
```

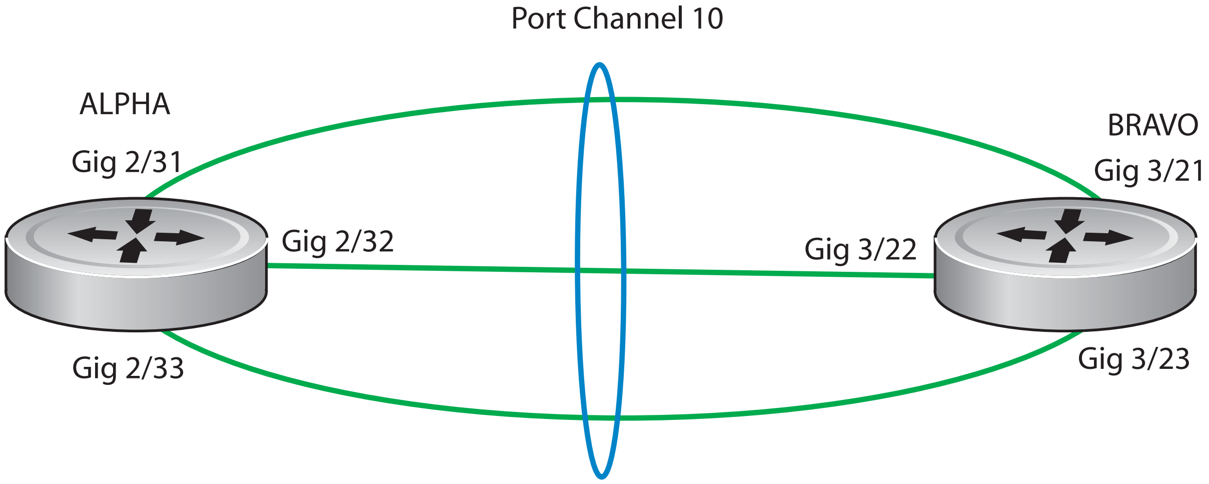
LACP Basic Configuration Example

The screenshots in this section are based on the example topology shown in Figure 19-11. Two routers are named ALPHA and BRAVO, and their hostname prompts reflect those names.

The sections are:

- Configuring a LAG on ALPHA on page 378
- Summary of the configuration on ALPHA on page 381
- Summary of the configuration on BRAVO on page 382

Figure 19-11. LACP Sample Topology

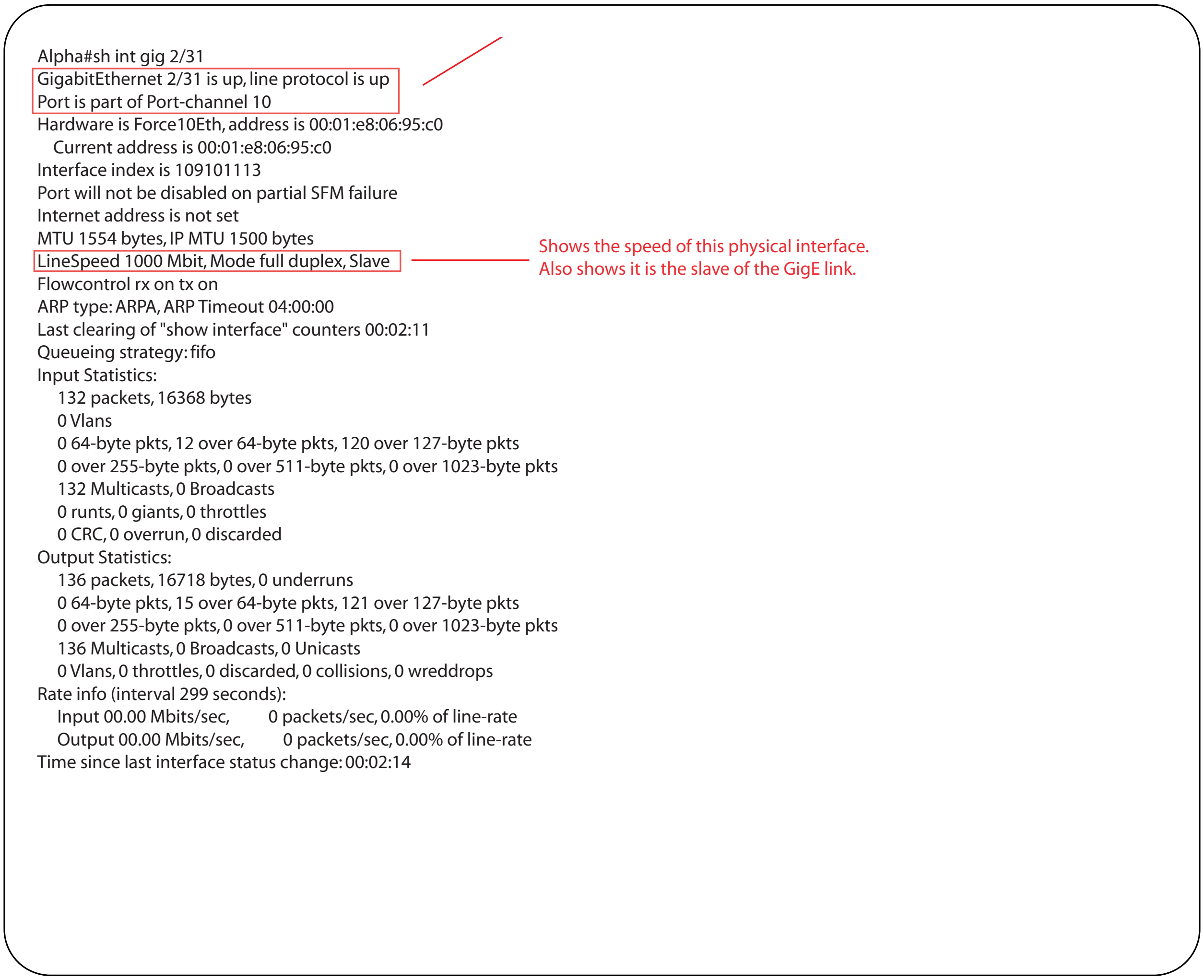


Configuring a LAG on ALPHA

Figure 19-12. Creating a LAG on ALPHA

```
Alpha(conf)#interface port-channel 10
Alpha(conf-if-po-10)#no ip address
Alpha(conf-if-po-10)#switchport
Alpha(conf-if-po-10)#no shutdown
Alpha(conf-if-po-10)#show config
!
interface Port-channel 10
 no ip address
 switchport
 no shutdown
!
Alpha(conf-if-po-10)#
```

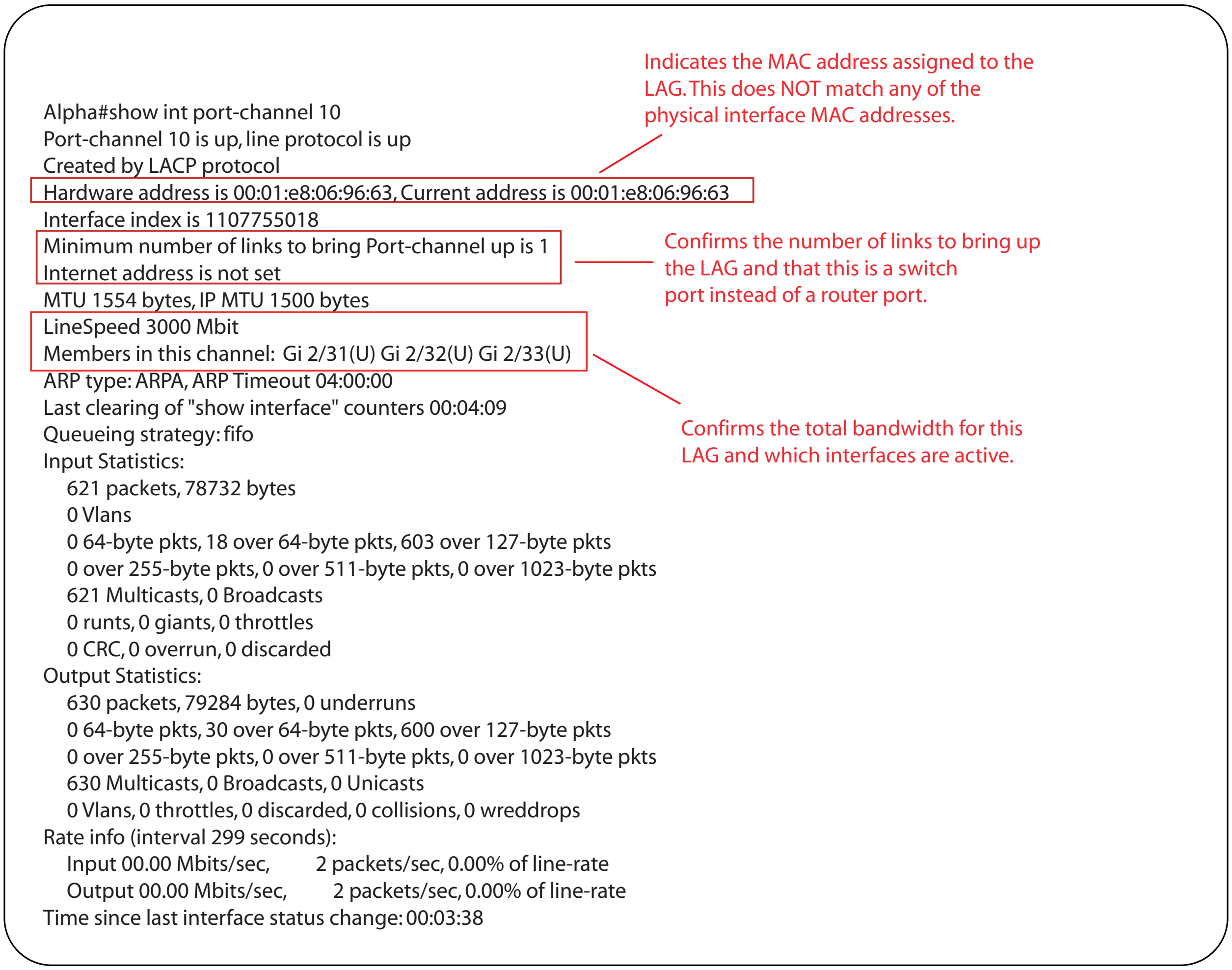
Figure 19-13. Inspecting a LAG Port Configuration on ALPHA



```
Alpha#sh int gig 2/31
GigabitEthernet 2/31 is up, line protocol is up
Port is part of Port-channel 10
Hardware is Force10Eth, address is 00:01:e8:06:95:c0
   Current address is 00:01:e8:06:95:c0
Interface index is 109101113
Port will not be disabled on partial SFM failure
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode full duplex, Slave
Flowcontrol rx on tx on
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:02:11
Queueing strategy: fifo
Input Statistics:
    132 packets, 16368 bytes
    0 Vlans
    0 64-byte pkts, 12 over 64-byte pkts, 120 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    132 Multicasts, 0 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics:
    136 packets, 16718 bytes, 0 underruns
    0 64-byte pkts, 15 over 64-byte pkts, 121 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    136 Multicasts, 0 Broadcasts, 0 Unicasts
    0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
    Input 00.00 Mbits/sec,      0 packets/sec, 0.00% of line-rate
    Output 00.00 Mbits/sec,      0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:02:14
```

Shows the speed of this physical interface.
Also shows it is the slave of the GigE link.

Figure 19-14. Inspecting Configuration of LAG 10 on ALPHA



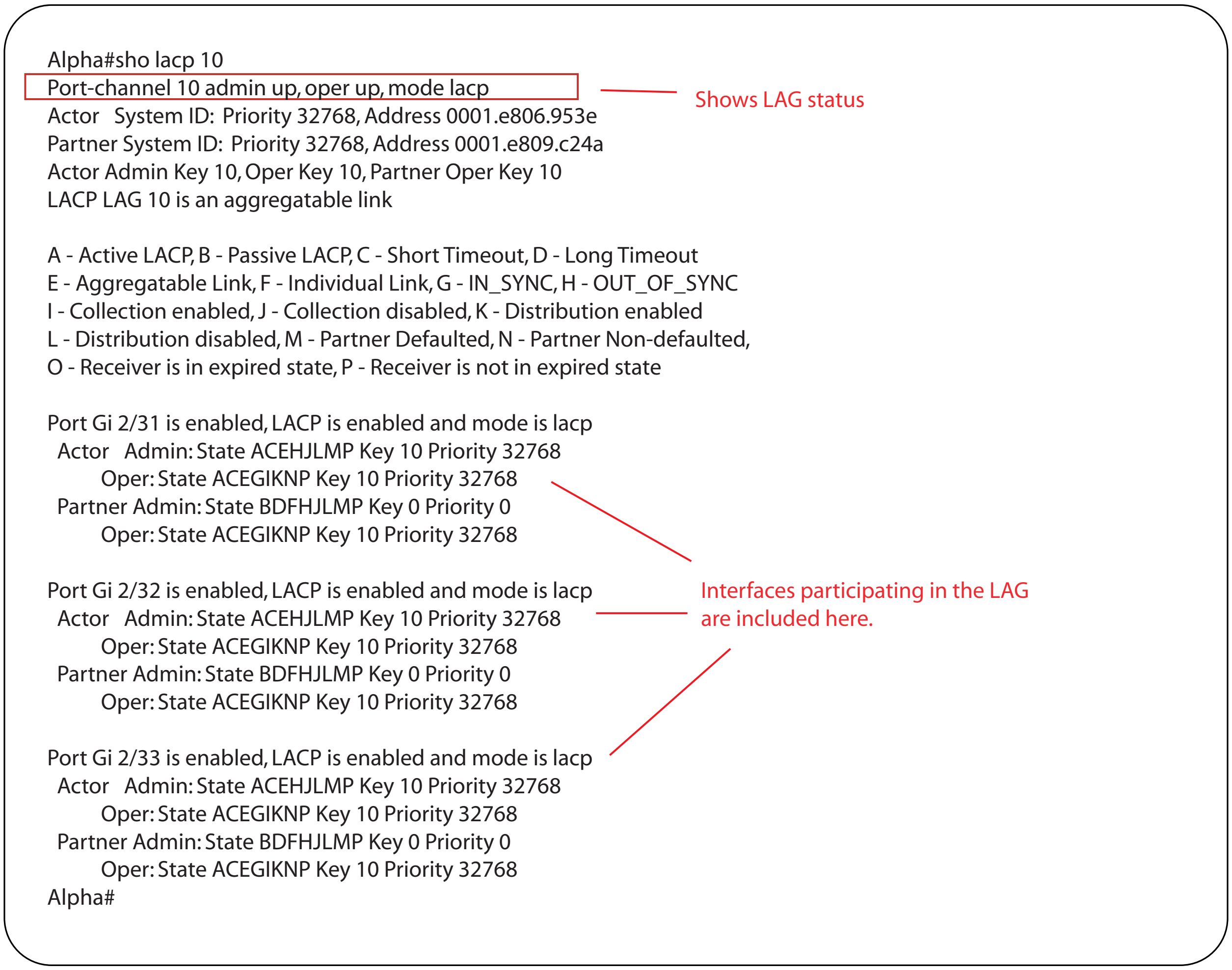
```
Alpha#show int port-channel 10
Port-channel 10 is up, line protocol is up
Created by LACP protocol
Hardware address is 00:01:e8:06:96:63, Current address is 00:01:e8:06:96:63
Interface index is 1107755018
Minimum number of links to bring Port-channel up is 1
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 3000 Mbit
Members in this channel:  Gi 2/31(U) Gi 2/32(U) Gi 2/33(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:04:09
Queueing strategy: fifo
Input Statistics:
    621 packets, 78732 bytes
    0 Vlans
    0 64-byte pkts, 18 over 64-byte pkts, 603 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    621 Multicasts, 0 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics:
    630 packets, 79284 bytes, 0 underruns
    0 64-byte pkts, 30 over 64-byte pkts, 600 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    630 Multicasts, 0 Broadcasts, 0 Unicasts
    0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
    Input 00.00 Mbits/sec,        2 packets/sec, 0.00% of line-rate
    Output 00.00 Mbits/sec,        2 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:03:38
```

Indicates the MAC address assigned to the LAG. This does NOT match any of the physical interface MAC addresses.

Confirms the number of links to bring up the LAG and that this is a switch port instead of a router port.

Confirms the total bandwidth for this LAG and which interfaces are active.

Figure 19-15. Using the show lacp Command to Verify LAG 10 Status on ALPHA



```
Alpha#sho lacp 10
Port-channel 10 admin up, oper up, mode lacp
Actor   System ID:  Priority 32768, Address 0001.e806.953e
Partner System ID:  Priority 32768, Address 0001.e809.c24a
Actor Admin Key 10, Oper Key 10, Partner Oper Key 10
LACP LAG 10 is an aggregatable link

A - Active LACP, B - Passive LACP, C - Short Timeout, D - Long Timeout
E - Aggregatable Link, F - Individual Link, G - IN_SYNC, H - OUT_OF_SYNC
I - Collection enabled, J - Collection disabled, K - Distribution enabled
L - Distribution disabled, M - Partner Defaulted, N - Partner Non-defaulted,
O - Receiver is in expired state, P - Receiver is not in expired state

Port Gi 2/31 is enabled, LACP is enabled and mode is lacp
  Actor   Admin: State ACEHJLMP Key 10 Priority 32768
          Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
          Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 2/32 is enabled, LACP is enabled and mode is lacp
  Actor   Admin: State ACEHJLMP Key 10 Priority 32768
          Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
          Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 2/33 is enabled, LACP is enabled and mode is lacp
  Actor   Admin: State ACEHJLMP Key 10 Priority 32768
          Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
          Oper: State ACEGIKNP Key 10 Priority 32768
Alpha#
```

Shows LAG status

Interfaces participating in the LAG are included here.

Summary of the configuration on ALPHA

Figure 19-16. Summary of the configuration on ALPHA

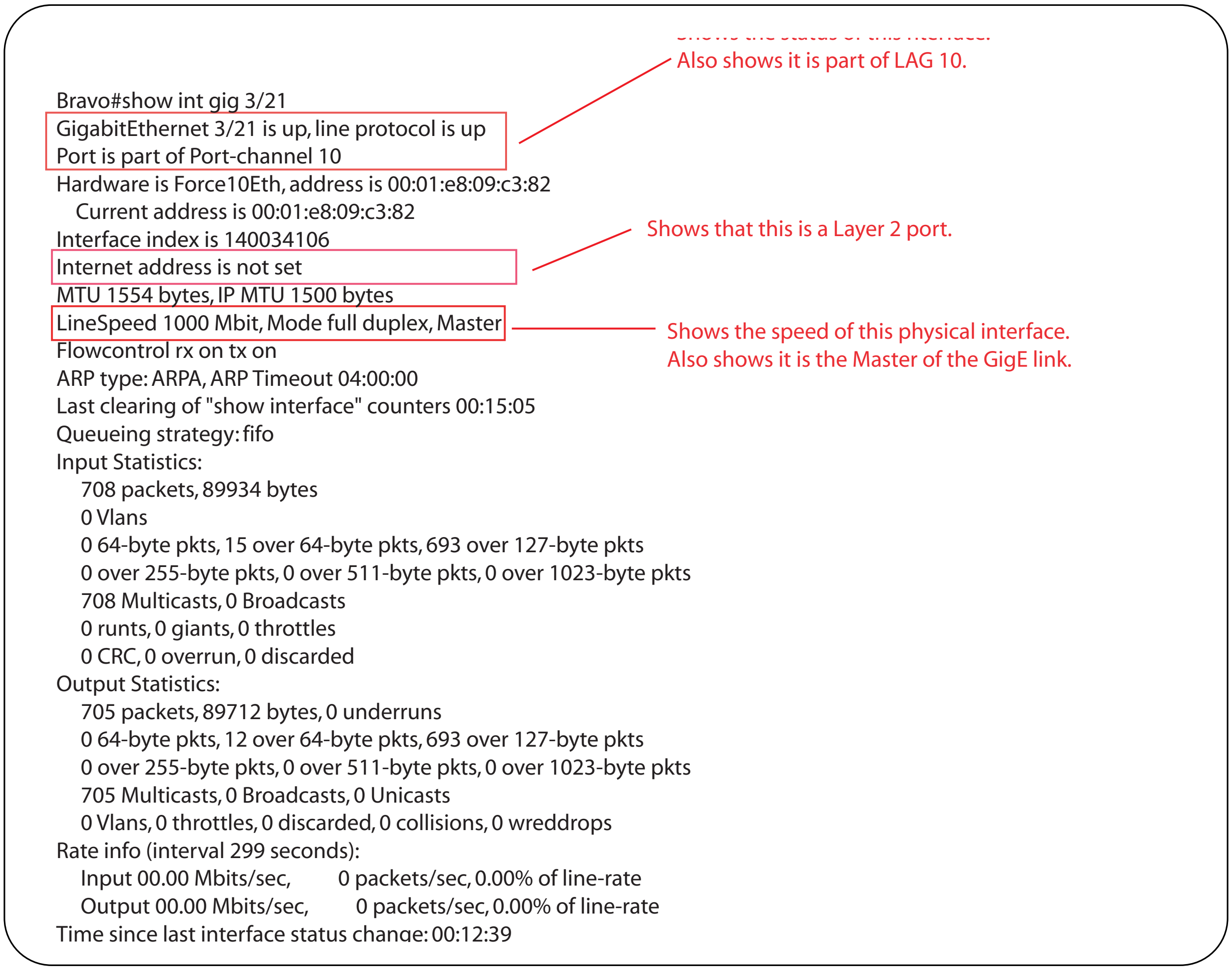
```
Alpha(conf-if-po-10)#int gig 2/31
Alpha(conf-if-gi-2/31)#no ip address
Alpha(conf-if-gi-2/31)#no switchport
Alpha(conf-if-gi-2/31)#shutdown
Alpha(conf-if-gi-2/31)#port-channel-protocol lacp
Alpha(conf-if-gi-2/31-lacp)#port-channel 10 mode active
Alpha(conf-if-gi-2/31-lacp)#no shut
Alpha(conf-if-gi-2/31)#show config

!
interface GigabitEthernet 2/31
 no ip address
!
 port-channel-protocol LACP
  port-channel 10 mode active
 no shutdown
!
Alpha(conf-if-gi-2/31)#

interface Port-channel 10
no ip address
switchport
no shutdown

interface GigabitEthernet 2/31
no ip address
no switchport
switchport
port-channel-protocol LACP
port-channel 10 mode active
no shutdown
```

Summary of the configuration on BRAVO

Figure 19-17. Summary of the configuration on BRAVO

```
Bravo(conf-if-gi-3/21)#int port-channel 10
Bravo(conf-if-po-10)#no ip add
Bravo(conf-if-po-10)#switch
Bravo(conf-if-po-10)#no shut
Bravo(conf-if-po-10)#show config
!
interface Port-channel 10
 no ip address
 switchport
 no shutdown
!
Bravo(conf-if-po-10)#exit

Bravo(conf)#int gig 3/21
Bravo(conf)#no ip address
Bravo(conf)#no switchport
Bravo(conf)#shutdown
Bravo(conf-if-gi-3/21)#port-channel-protocol lacp
Bravo(conf-if-gi-3/21-lacp)#port-channel 10 mode active
Bravo(conf-if-gi-3/21-lacp)#no shut
Bravo(conf-if-gi-3/21)#end

!
interface GigabitEthernet 3/21
 no ip address
!
 port-channel-protocol LACP
  port-channel 10 mode active
 no shutdown
Bravo(conf-if-gi-3/21)#end

int port-channel 10
no ip address
switchport
no shutdown
show config

int gig 3/21
no ip address
no switchport
shutdown
port-channel-protocol lacp
port-channel 10 mode active
no shut
show config
end
```

Figure 19-18. Using the show interface Command to Inspect a LAG Port on BRAVO



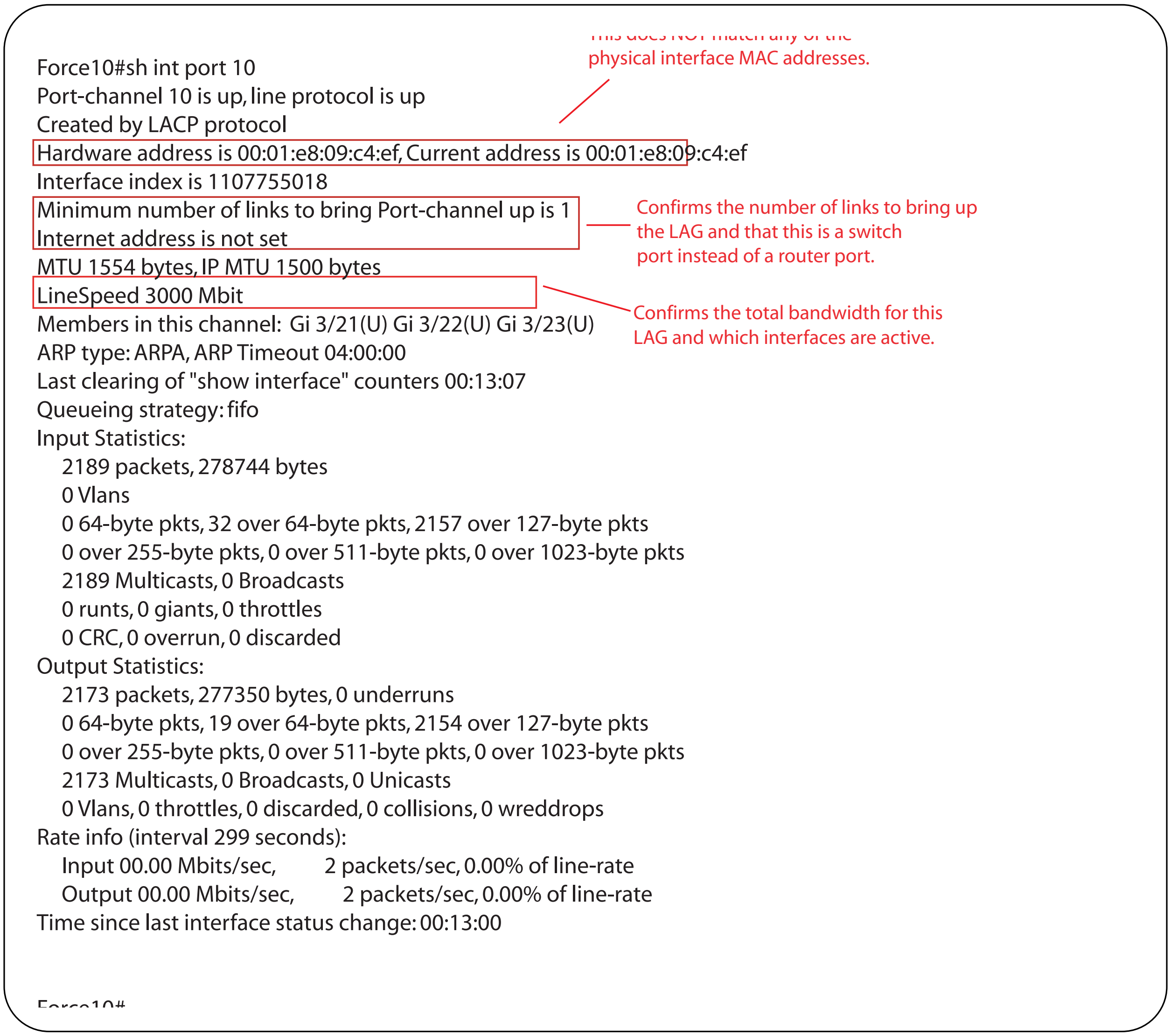
```
Bravo#show int gig 3/21
GigabitEthernet 3/21 is up, line protocol is up
Port is part of Port-channel 10
Hardware is Force10Eth, address is 00:01:e8:09:c3:82
   Current address is 00:01:e8:09:c3:82
Interface index is 140034106
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode full duplex, Master
Flowcontrol rx on tx on
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:15:05
Queueing strategy: fifo
Input Statistics:
    708 packets, 89934 bytes
    0 Vlans
    0 64-byte pkts, 15 over 64-byte pkts, 693 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    708 Multicasts, 0 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics:
    705 packets, 89712 bytes, 0 underruns
    0 64-byte pkts, 12 over 64-byte pkts, 693 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    705 Multicasts, 0 Broadcasts, 0 Unicasts
    0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
    Input 00.00 Mbits/sec,       0 packets/sec, 0.00% of line-rate
    Output 00.00 Mbits/sec,       0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:12:39
```

Shows the status of this interface.
Also shows it is part of LAG 10.

Shows that this is a Layer 2 port.

Shows the speed of this physical interface.
Also shows it is the Master of the GigE link.

Figure 19-19. Using the show interfaces port-channel Command to Inspect LAG 10



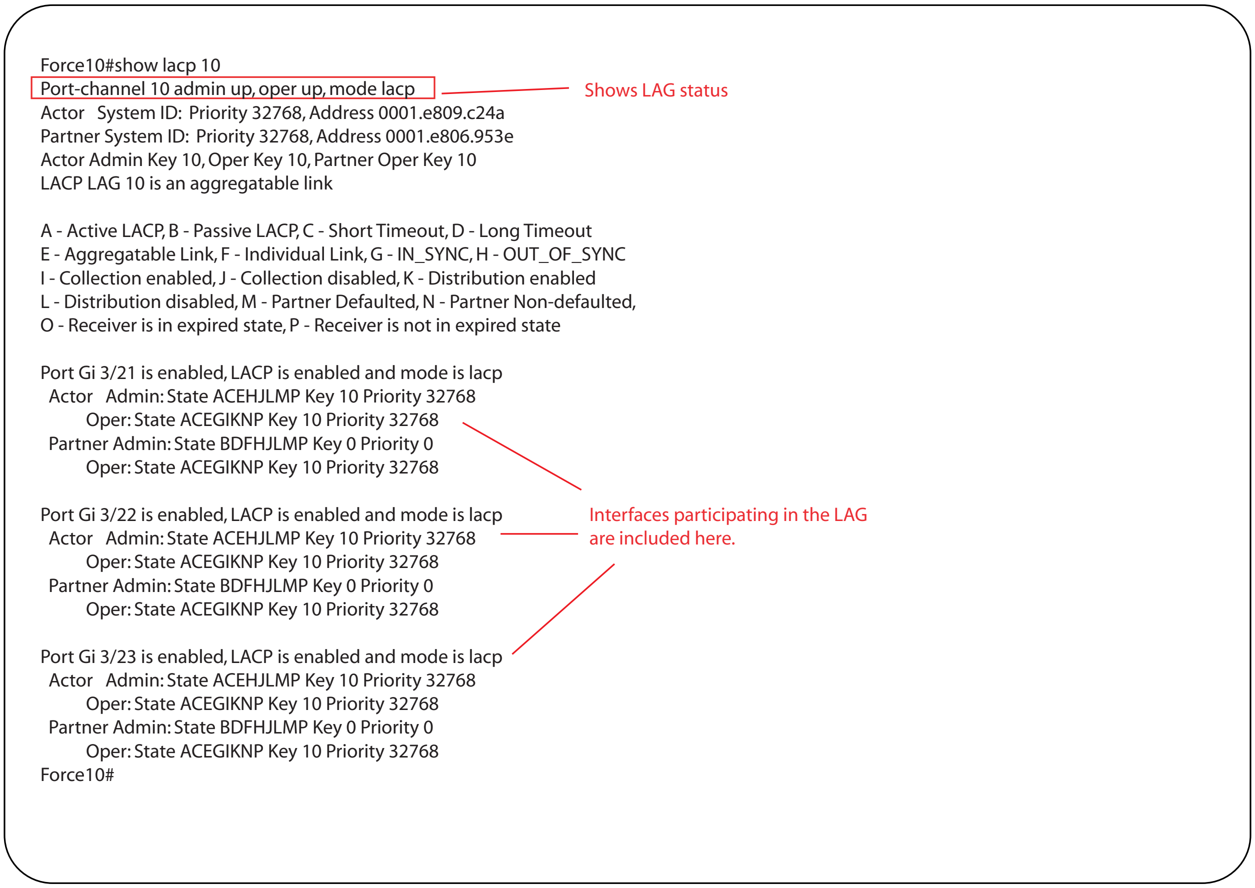
```
Force10#sh int port 10
Port-channel 10 is up, line protocol is up
Created by LACP protocol
Hardware address is 00:01:e8:09:c4:ef, Current address is 00:01:e8:09:c4:ef
Interface index is 1107755018
Minimum number of links to bring Port-channel up is 1
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 3000 Mbit
Members in this channel:  Gi 3/21(U) Gi 3/22(U) Gi 3/23(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:13:07
Queueing strategy: fifo
Input Statistics:
   2189 packets, 278744 bytes
   0 Vlans
   0 64-byte pkts, 32 over 64-byte pkts, 2157 over 127-byte pkts
   0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
   2189 Multicasts, 0 Broadcasts
   0 runts, 0 giants, 0 throttles
   0 CRC, 0 overrun, 0 discarded
Output Statistics:
   2173 packets, 277350 bytes, 0 underruns
   0 64-byte pkts, 19 over 64-byte pkts, 2154 over 127-byte pkts
   0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
   2173 Multicasts, 0 Broadcasts, 0 Unicasts
   0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
   Input 00.00 Mbits/sec,       2 packets/sec, 0.00% of line-rate
   Output 00.00 Mbits/sec,       2 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:13:00


Force10#
```

This does NOT match any of the physical interface MAC addresses.

Confirms the number of links to bring up the LAG and that this is a switch port instead of a router port.

Confirms the total bandwidth for this LAG and which interfaces are active.

Figure 19-20. Using the show lacp Command to Inspect LAG Status



```
Force10#show lacp 10
Port-channel 10 admin up, oper up, mode lacp
Actor   System ID:  Priority 32768, Address 0001.e809.c24a
Partner System ID:  Priority 32768, Address 0001.e806.953e
Actor Admin Key 10, Oper Key 10, Partner Oper Key 10
LACP LAG 10 is an aggregatable link

A - Active LACP, B - Passive LACP, C - Short Timeout, D - Long Timeout
E - Aggregatable Link, F - Individual Link, G - IN_SYNC, H - OUT_OF_SYNC
I - Collection enabled, J - Collection disabled, K - Distribution enabled
L - Distribution disabled, M - Partner Defaulted, N - Partner Non-defaulted,
O - Receiver is in expired state, P - Receiver is not in expired state

Port Gi 3/21 is enabled, LACP is enabled and mode is lacp
  Actor   Admin: State ACEHJLMP Key 10 Priority 32768
           Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 3/22 is enabled, LACP is enabled and mode is lacp
  Actor   Admin: State ACEHJLMP Key 10 Priority 32768
           Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 3/23 is enabled, LACP is enabled and mode is lacp
  Actor   Admin: State ACEHJLMP Key 10 Priority 32768
           Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
           Oper: State ACEGIKNP Key 10 Priority 32768
Force10#
```

Shows LAG status

Interfaces participating in the LAG are included here.

PPP is a connection-oriented protocol that enables layer two links over a variety of different physical layer connections. It is supported on both synchronous and asynchronous lines, and can operate in half-duplex or full-duplex mode. It was designed to carry IP traffic but is general enough to allow any type of network layer datagram to be sent over a PPP connection. As its name implies, it is for point-to-point connections between exactly two devices, and assumes that frames are sent and received in the same order.

20

Layer 2

Layer 2 features are supported on platforms: E □ S4810

This chapter describes the following Layer 2 features:

- Managing the MAC Address Table
- MAC Learning Limit
- NIC Teaming
- Microsoft Clustering
- Configuring Redundant Pairs
- Restricting Layer 2 Flooding
- Restricting Layer 2 Multicast Flooding over Low Speed Ports
- Far-end Failure Detection

Managing the MAC Address Table

FTOS provides the following management activities for the MAC address table:

- Clear the MAC Address Table
- Set the Aging Time for Dynamic Entries
- Configure a Static MAC Address
- Display the MAC Address Table

Clear the MAC Address Table

You may clear the MAC address table of dynamic entries:

Task	Command Syntax	Command Mode
Clear a MAC address table of dynamic entries. • **address** deletes the specified entry • **all** deletes all dynamic entries • **interface** deletes all entries for the specified interface • **vlan** deletes all entries for the specified VLAN	**clear mac-address-table {dynamic | sticky}** {*address* | **all | interface | vlan**}	EXEC Privilege

Set the Aging Time for Dynamic Entries

Learned MAC addresses are entered in the table as dynamic entries, which means that they are subject to aging. For any dynamic entry, if no packet arrives on the switch with the MAC address as the source or destination address within the timer period, the address is removed from the table. The default aging time is 1800 seconds.

Task	Command Syntax	Command Mode
Disable MAC address aging for all dynamic entries.	**mac-address-table aging-time 0**	CONFIGURATION
Specify an aging time.	**mac-address-table aging-time** *seconds* Range: 10-1000000	CONFIGURATION

Set the Aging Time for Dynamic Entries on a VLAN

Set the Aging Time for Dynamic Entries on a VLAN is available only on platform: E

Task	Command Syntax	Command Mode
Specify an aging time.	**mac-address-table aging-time** *seconds* Range: 1-1000000	INTERFACE VLAN

FTOS Behavior: The time elapsed before the configured MAC aging time expires is not precisely as configured. For example, the VLAN configuration **mac-address-table aging-time 1** does not remove dynamic entries from the CAM after precisely 1 second. The actual minimum aging time for entries is approximately 5 seconds because this is the default MAC address table scanning interval. Therefore, MAC aging configurations of less than 5 seconds, as in this example, might be ineffective. Configuring **mac-address-table station-move time-interval 500** solves this limitation. Reducing the scanning interval to the minimum, 500 milliseconds, increases the detection speed, which results in FTOS clearing entries closer to the actual desired aging time.

Configure a Static MAC Address

A static entry is one that is not subject to aging. Static entries must be entered manually:

Task	Command Syntax	Command Mode
Create a static MAC address entry in the MAC address table.	**mac-address-table static**	CONFIGURATION

Display the MAC Address Table

To display the contents of the MAC address table:

Task	Command Syntax	CommandMode
Display the contents of the MAC address table. • **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.	**show mac-address-table** [**address** | **aging-time** [**vlan** *vlan-id*]| **count** | **dynamic** | **interface** | **static | vlan**]	EXEC Privilege

MAC Learning Limit

This section has the following sub-sections:

- mac learning-limit dynamic
- mac learning-limit mac-address-sticky
- mac learning-limit station-move
- Learning Limit Violation Actions
- Station Move Violation Actions
- Recovering from Learning Limit and Station Move Violations
- Per-VLAN MAC Learning Limit

MAC Address Learning Limit is a method of port security on Layer 2 port-channel and physical interfaces, and VLANs. It enables you to set an upper limit on the number of MAC addresses that learned on an interface/VLAN. After the limit is reached, the system drops all traffic from a device with an unlearned MAC address.

FTOS Behavior: When configuring MAC Learning Limit on a port or VLAN the configuration is accepted (becomes part of **running-config** and **show mac learning-limit interface**) before the system verifies that sufficient CAM space exists. If the CAM check fails, a message is displayed:

```
%E90MH:5 %ACL_AGENT-2-ACL_AGENT_LIST_ERROR: Unable to apply  access-list Mac-Limit  on 
GigabitEthernet 5/84
```

In this case, the configuration is still present in the running-config and **show** output. Remove the configuration before re-applying a MAC learning limit with lower value. Also, ensure that Syslog messages can be viewed on your session.

Note: The CAM-check failure message beginning in FTOS version 8.3.1.0 is different from versions 8.2.1.1 and earlier, which read:

```
% Error: ACL returned error

% Error: Remove existing limit configuration if it was configured before
```

To set a MAC learning limit on an interface:

Task	Command Syntax	Command Mode
Specify the number of MAC addresses that the system can learn off a Layer 2 interface.	**mac learning-limit** *address_limit*	INTERFACE

Three options are available with the **mac learning-limit** command: **dynamic**, **no-station-move**, and **station-move**.

Note: An SNMP trap is available for **mac learning-limit station-move**. No other SNMP traps are available for MAC Learning Limit, including limit violations.

mac learning-limit dynamic

The MAC address table is stored on the Layer 2 FIB region of the CAM (and the Layer 2 ACL region on the E-Series). On the C-Series and S-Series the Layer 2 FIB region allocates space for static MAC address entries and dynamic MAC address entries (all MAC address entries on the E-Series are dynamic). When MAC Learning Limit is enabled, entries created on this port are static by default. When you configure the **dynamic** option, learned MAC addresses are stored in the dynamic region and are subject to aging. Entries created before this option is set are not affected.

FTOS Behavior: If you do not configure the **dynamic** option, the C-Series and S-Series do not detect station moves in which a MAC address learned off of a MAC-limited port is learned on another port on same line card. Therefore, FTOS does not take any configured station-move violation action. When a MAC address is relearned on any other linecard (any line card except the one to which the original MAC-limited port belongs), the station-move is detected, and the system takes the configured the violation action.

mac learning-limit mac-address-sticky

Using sticky MAC addresses allows you to associate a specific port with MAC addresses from trusted devices. If sticky MAC is enabled, the specified port will retain any dynamically-learned addresses and prevent them from being transferred or learned on other ports.

If **mac-learning-limit** is configured and sticky MAC is enabled, all dynamically-learned addresses are converted to sticky MAC addresses for the selected port. Any new MAC addresses learned on this port will be converted to sticky MAC addresses.

To save all sticky MAC addresses into a configuration file that can be used as a startup configuration file, use the **write config** command. If the number of existing MAC addresses is fewer than the configured mac learn limit, any additional MAC addresses will be converted to sticky MACs on that interface. To remove all sticky MAC addresses from the running config file, disable sticky MAC and use the **write config** command.

When sticky mac is enabled on an interface, dynamically-learned MAC addresses will not age, even if **mac-learning-limit dynamic** is enabled. If **mac-learning-limit** and **mac-learning-limit dynamic** are configured and sticky MAC is disabled, any dynamically-learned MAC addresses will age.

mac learning-limit station-move

mac learning-limit station-move is available only on platforms: C S Z

The **station-move** option, allows a MAC address already in the table to be learned off of another interface. For example, if you disconnect a network device from one interface and reconnect it to another interface, the MAC address is learned on the new interface. When the system detects this “station move,” the system clears the entry learned on the original interface and installs a new entry on the new interface.

Learning Limit Violation Actions

Learning Limit Violation Actions are supported only on platforms: E S60 S4810.

You can configure the system to take an action when the MAC learning limit is reached on an interface and a new address is received using one of the following options with the **mac learning-limit** command:

Task	Command Syntax	Command Mode
Generate a system log message when the MAC learning limit is exceeded.	**learn-limit-violation log**	INTERFACE
Shut down the interface and generate a system log message when the MAC learning limit is exceeded.	**learn-limit-violation shutdown**	INTERFACE

Station Move Violation Actions

Station Move Violation Actions are supported only on platforms: E S60 , and S4810.

no-station-move is the default behavior. You can configure the system to take an action if a station move occurs using one the following options with the **mac learning-limit** command:.

Task	Command Syntax	Command Mode
Generate a system log message indicating a station move.	**station-move-violation log**	INTERFACE
Shut down the first port to learn the MAC address.	**station-move-violation shutdown-original**	INTERFACE
Shut down the second port to learn the MAC address.	**station-move-violation shutdown-offending**	INTERFACE
Shut down both the first and second port to learn the MAC address.	**station-move-violation shutdown-both**	INTERFACE

To display a list of interfaces configured with MAC learning limit or station move violation actions:

Task	Command Syntax	Command Mode
Display a list of all of the interfaces configured with MAC learning limit or station move violation.	**show mac learning-limit violate-action**	CONFIGURATION

Recovering from Learning Limit and Station Move Violations

After a learning-limit or station-move violation shuts down an interface, you must manually reset it:

Task	Command Syntax	Command Mode
Reset interfaces in ERR_Disabled state caused by a learning limit violation or station move violation.	**mac learning-limit reset**	EXEC Privilege
Reset interfaces in ERR_Disabled state caused by a learning limit violation.	**mac learning-limit reset learn-limit-violation [interface | all]**	EXEC Privilege
Reset interfaces in ERR_Disabled state caused by a station move violation.	**mac learning-limit reset station-move-violation [interface | all]**	EXEC Privilege

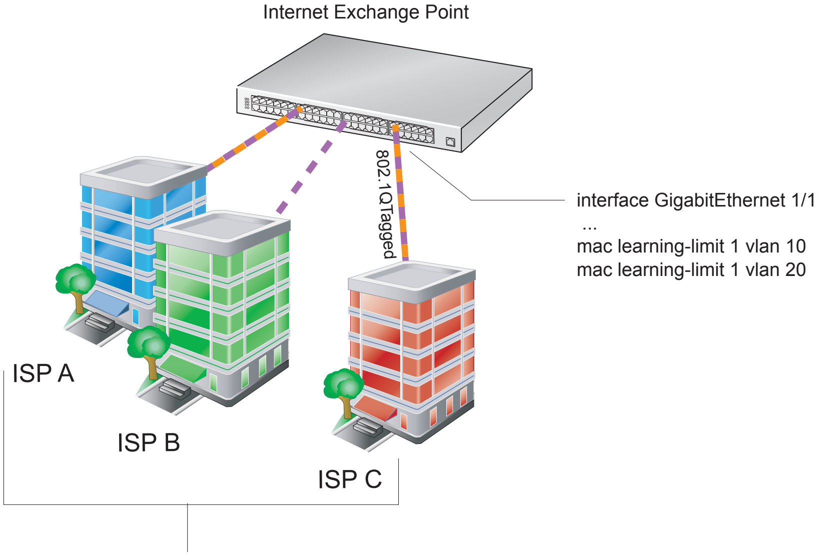
Note: Alternatively, you can reset the interface by shutting it down using the **shutdown** command and then reenabling it using the command **no shutdown**.

Per-VLAN MAC Learning Limit

Per-VLAN MAC Learning Limit is available only on platform: E

An individual MAC learning limit can be configured for each VLAN using Per-VLAN MAC Learning Limit.

One application of Per-VLAN MAC Learning Limit is on access ports. In the following illustration, an Internet Exchange Point (IXP) connects multiple Internet Service Provider (ISP). An IXP can provide several types of services to its customers including public and private peering. Public peering means that all customers are connected to one VLAN. If one ISP wants to peer with another ISP, it establishes a BGP peering session over this VLAN. Private Peering means that the IXP sets up a separate VLAN between two customers that want to peer privately; only the ports of these two ISPs would belong to this VLAN and they would peer via BGP. In the following illustration, Per-VLAN MAC Learning Limit is used on the access ports for the ISPs that have subscribed to private and public peering since these access ports are members of multiple VLANs.



```
interface GigabitEthernet 1/1
 ...
mac learning-limit 1 vlan 10
mac learning-limit 1 vlan 20
```

ISP A, B, and C are all public peers through VLAN 10. In addition, ISP A and C are private peers on a separate VLAN, VLAN 20. Since the access ports for ISP A and C are members of multiple VLANs, Per-VLAN MAC Learning Limit can be applied to those ports.

Task	Command Syntax	Command Mode
Configure a MAC learning limit on a VLAN.	**mac learning-limit** *limit* **vlan** *vlan-id*	INTERFACE
Display the MAC learning limit counters for a VLAN.	**show mac learning-limit** [**interface** *slot/port* [**vlan** *vlan-id*]]	EXEC Privilege

Task	Command Syntax	Command Mode

```
FTOS#show mac learning-limit
Interface      Vlan      Learning       Dynamic        Static         Unknown SA
Slot/port      Id        Limit          MAC count      MAC count      Drops
Gi 5/84        2         2                    0                   0                  0
Gi 5/84        *         5                    0                   0                  0
Gi 5/85        3         3                    0                   0                  0
Gi 5/85        *         10                   0                   0                  0
FTOS#show mac learning-limit interface gig 5/84
Interface      Vlan      Learning       Dynamic        Static         Unknown SA
Slot/port      Id        Limit          MAC count      MAC count      Drops
Gi 5/84        2         2                    0                    0                 0
Gi 5/84        *         5                    0                    0                 0
FTOS#show mac learning-limit interface gig 5/84 vlan 2
Interface      Vlan      Learning       Dynamic        Static         Unknown SA
Slot/port      Id        Limit          MAC count      MAC count      Drops
Gi 5/84        2         2                      0                 0                  0
```

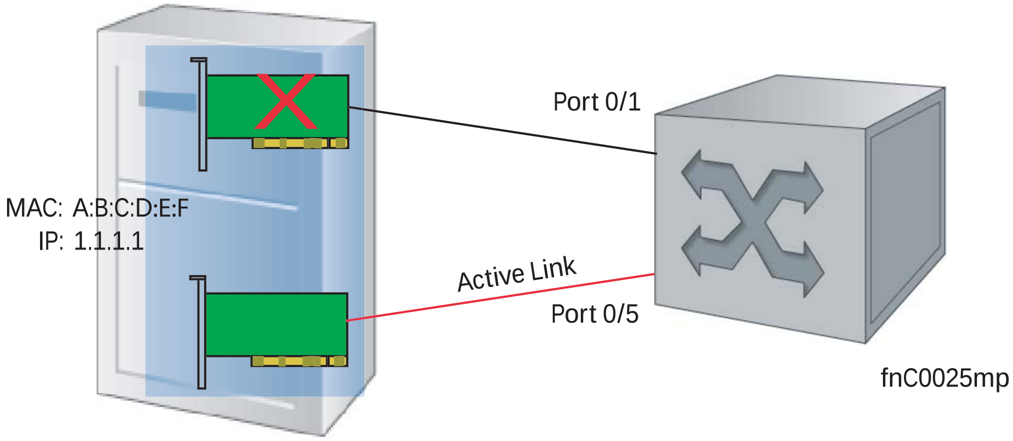
NIC Teaming

NIC Teaming is available on the following platforms: Z □ E □ □

NIC teaming is a feature that allows multiple network interface cards in a server to be represented by one MAC address and one IP address in order to provide transparent redundancy, balancing, and to fully utilize network adapter resources.

The following illustration shows a topology where two NICs have been teamed together. In this case, if the primary NIC fails, traffic switches to the secondary NIC since they are represented by the same set of addresses.

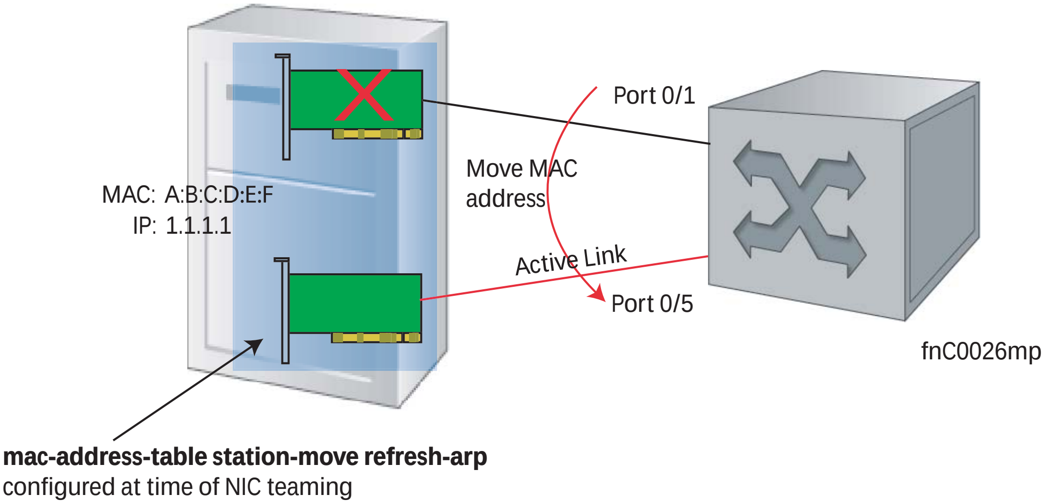
Figure 20-1. Redundant NICs with NIC Teaming



When NIC teaming is employed, consider that the server MAC address is originally learned on Port 0/1 of the switch (Figure 20-2) and Port 0/5 is the failover port. When the NIC fails, the system automatically sends an ARP request for the gateway or host NIC to resolve the ARP and refresh the egress interface. When the ARP is resolved, the same MAC address is learned on the same port where the ARP is resolved (in the above example, this is Port 0/5 of the switch). To ensure the MAC address is disassociated with one port and re-associated with another port in the ARP table, you must configure the command **mac-address-table station-move refresh-arp** on the Dell Force10 switch at the time that NIC teaming is being configured on the server.

Note: If this command is not configured, traffic continues to be forwarded to the failed NIC until the ARP entry on the switch times out.

Figure 20-2. Configuring mac-address-table station-move refresh-arp Command


Port 0/1
Move MAC address
MAC: A:B:C:D:E:F
IP: 1.1.1.1
Active Link
Port 0/5
fnC0026mp
mac-address-table station-move refresh-arp
configured at time of NIC teaming

MAC Move Optimization

MAC Move Optimization is supported only on platform: E

Station-move detection takes 5000ms because this is the interval at which the detection algorithm runs. On the E-Series, you can reduce detection time to as little as 500ms using the command **mac-address-table station-move threshold time-interval** (though at the expense of CPU resources).

threshold is the number of times a station move must be detected in a single interval in order to trigger a system log message. For example, if you configure **mac-address-table station-move threshold 2 time-interval 5000**, and 4 station moves occur in 5000ms, then two log messages are generated.

Microsoft Clustering

Microsoft Clustering is supported only on platform: E

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. FTOS does not recognize server clusters by default; it must be configured to do so.

Default Behavior

When an ARP request is sent to a server cluster, either the active server or all of the servers send a reply, depending on the cluster configuration. If the active server sends a reply, the Dell Force10 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 (Figure 20-3); the virtual MAC address is never learned.

Since 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 (Figure 20-4).

Figure 20-3. Server Clustering: Multiple ARP Replies

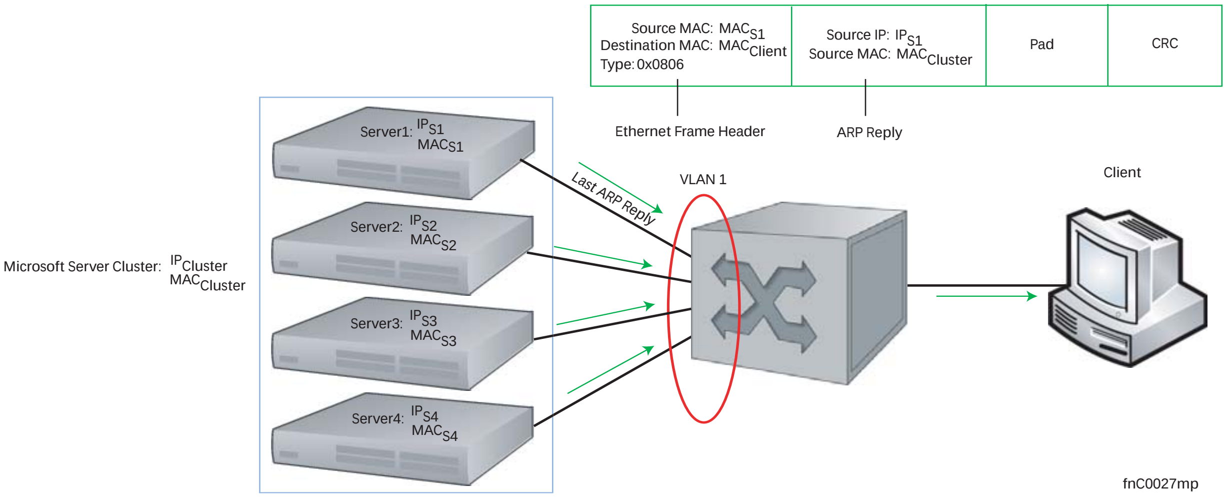
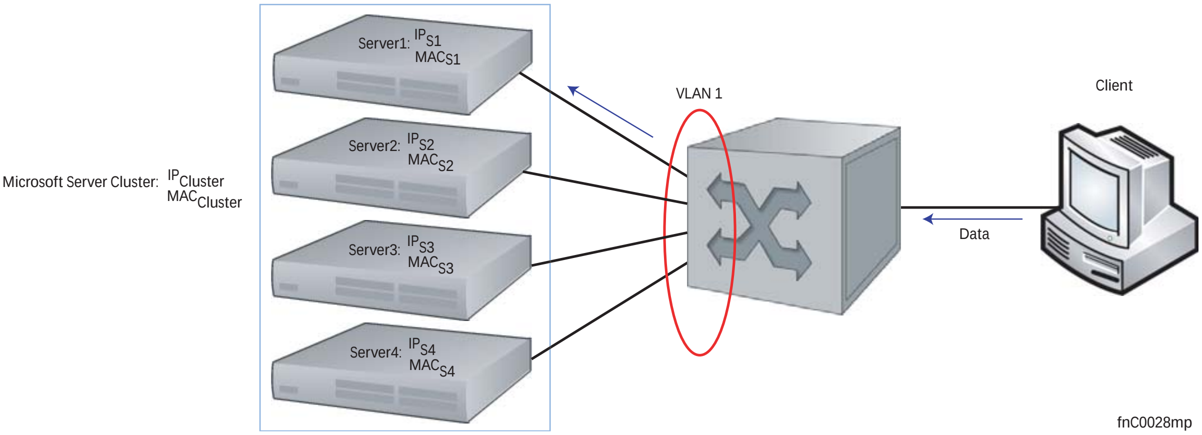


Figure 20-4. Server Clustering: Failover and Balancing Not Preserved

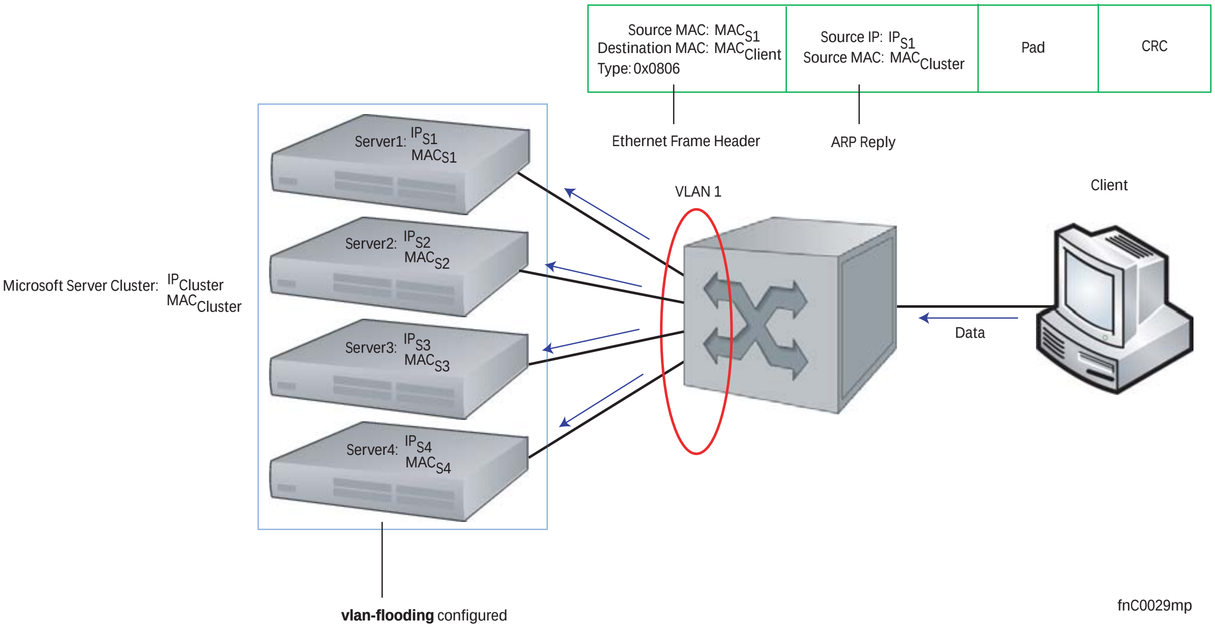


Configuring the Switch for Microsoft Server Clustering

To preserve failover and balancing, the Dell Force10 switch must learn the cluster's virtual MAC address, and it must forward traffic destined for the server cluster out all member ports in the VLAN connected to the cluster. To ensure that this happens, you must configure the command **vlan-flooding** on the Dell Force10 switch at the time that the Microsoft cluster is configured (Figure 20-5).

As shown in Figure 20-5, 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. The **vlan-flooding** command directs the system to discover that there are different MAC addresses in an ARP reply and associate the virtual MAC address with the VLAN connected to the cluster. Then, all traffic destined for the cluster is flooded out of all member ports. Since all of the servers in the cluster receive traffic, failover and balancing are preserved.

Figure 20-5. Server Cluster: Failover and Balancing Preserved with the vlan-flooding Command



Source MAC: MAC$_{S1}$
Destination MAC: MAC$_{Client}$
Type: 0x0806
Source IP: IP$_{S1}$
Source MAC: MAC$_{Cluster}$
Pad
CRC
Ethernet Frame Header
ARP Reply
Server1: IP$_{S1}$ MAC$_{S1}$
Server2: IP$_{S2}$ MAC$_{S2}$
Server3: IP$_{S3}$ MAC$_{S3}$
Server4: IP$_{S4}$ MAC$_{S4}$
Microsoft Server Cluster: IP$_{Cluster}$ MAC$_{Cluster}$
VLAN 1
Client
Data
vlan-flooding configured
fnC0029mp

Enable and Disable VLAN Flooding

- ARP entries already resolved through the VLAN are deleted when the feature is enabled. This ensures that ARP entries across the VLAN are consistent.
- All ARP entries learned after the feature is enabled are deleted when the feature is disabled, and RP2 triggers ARP resolution. The feature is disabled with the command **no vlan-flooding**.
- 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 is not reflected in the output of the show arp command but is reflected in the output of the command **show ipf fib**.

The ARP entries exist in the secondary RPM CAM, so failover has no effect on the feature.

Configuring Redundant Pairs

Configuring Redundant Pairs is supported on platforms: E C S S4810 Z

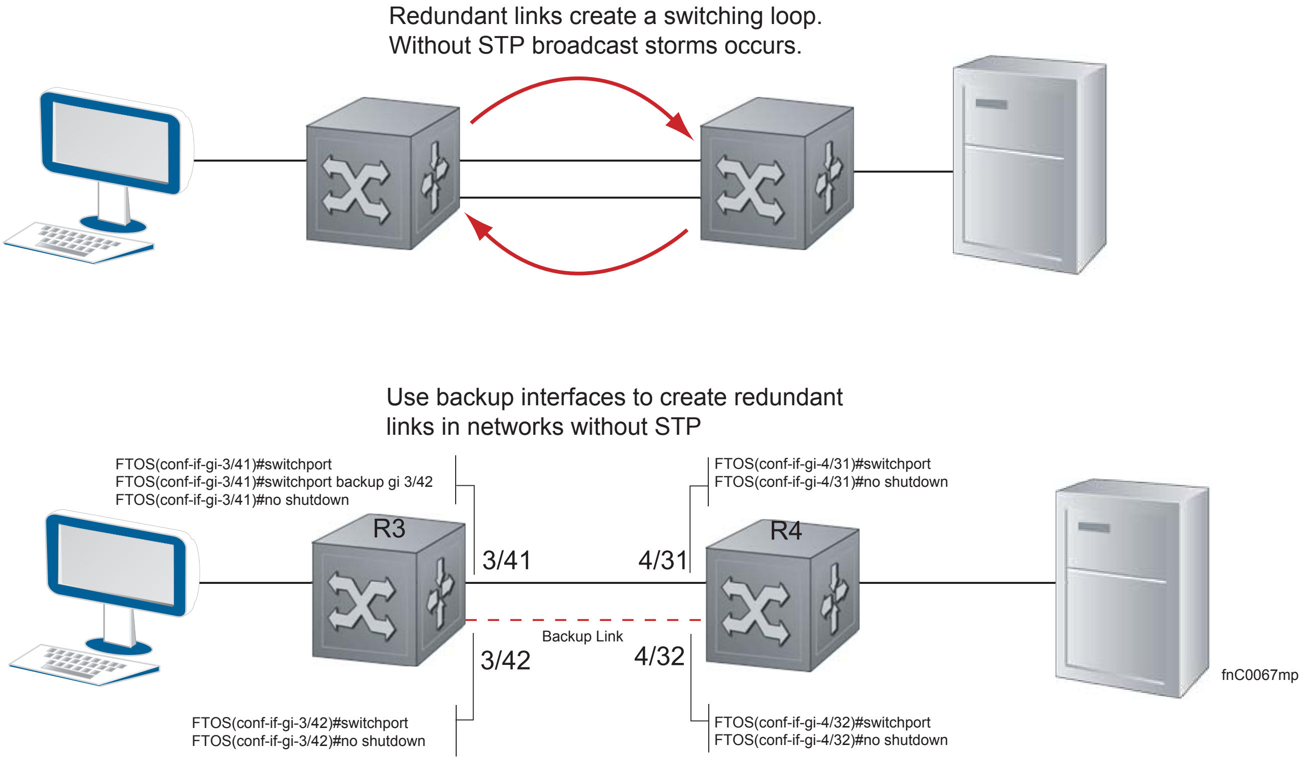
Networks that employ switches that do not support Spanning Tree (STP) — for example, networks with Digital Subscriber Line Access Mutiplexers (DSLAM) — cannot have redundant links between switches because they create switching loops (Figure 20-6). The Redundant Pairs feature enables you to create redundant links in networks that do not use STP by configuring backup interfaces for the interfaces on either side of the primary link.

Note: For details on STP, see Chapter 39, “Spanning Tree Protocol,” on page 701.

Assign a backup interface to an interface using the command **switchport backup**. The backup interface remains in down state until the primary fails, at which point it transitions to up state. If the primary interface fails, and later comes up, it becomes the backup interface for the redundant pair. FTOS supports Gigabit, 10-Gigabit, and 40-Gigabit interfaces as backup interfaces.

You must apply all other configurations to each interface in the redundant pair such that their configurations are *identical*, so that transition to the backup interface in the event of a failure is transparent to rest of the network.

Figure 20-6. Configuring Redundant Layer 2 Pairs without Spanning Tree


Redundant links create a switching loop.
Without STP broadcast storms occurs.
Use backup interfaces to create redundant
links in networks without STP
R3
R4
3/41
4/31
Backup Link
3/42
4/32
fnC0067mp

```
FTOS(conf-if-gi-3/41)#switchport
FTOS(conf-if-gi-3/41)#switchport backup gi 3/42
FTOS(conf-if-gi-3/41)#no shutdown
```

```
FTOS(conf-if-gi-4/31)#switchport
FTOS(conf-if-gi-4/31)#no shutdown
```

```
FTOS(conf-if-gi-3/42)#switchport
FTOS(conf-if-gi-3/42)#no shutdown
```

```
FTOS(conf-if-gi-4/32)#switchport
FTOS(conf-if-gi-4/32)#no shutdown
```

You configure a redundant pair by assigning a backup interface to a primary interface with the **switchport backup interface** command. Initially, the primary interface is active and transmits traffic and the backup interface remains down. If the primary fails for any reason, the backup transitions to an active UP state. If the primary interface fails and later comes back up, it remains as the backup interface for the redundant pair.

FTOS supports only Gigabit, 10-Gigabit, and 40-Gigabit ports and port channels as primary/backup interfaces in redundant pairs. (A port channel is also referred to as a Link Aggregation Group (LAG). See Chapter 15, Interfaces, Port Channel Interfaces on page 294 for more information.) If the interface is a member link of a LAG, the following primary/backup interfaces are also supported:

- primary interface is a physical interface, the backup interface can be a physical interface
- primary interface is a physical interface, the backup interface can be a static or dynamic LAG
- primary interface is a static or dynamic LAG, the backup interface can be a physical interface
- primary interface is a static or dynamic LAG, the backup interface can be a static or dynamic LAG

In a redundant pair, any combination of physical and port-channel interfaces is supported as the two interfaces in a redundant pair. For example, you can configure a static (without LACP) or dynamic (with LACP) port-channel interface as either the primary or backup link in a redundant pair with a physical interface.

To ensure that existing network applications see no difference when a primary interface in a redundant pair transitions to the backup interface, be sure to apply *identical* configurations of other traffic parameters to each interface.

If you remove an interface in a redundant link (remove the line card of a physical interface or delete a port channel with the **no interface port-channel** command), the redundant pair configuration is also removed.

Important Points about Configuring Redundant Pairs

- You may not configure any interface to be a backup for more than one interface, no interface can have more than one backup, and a backup interface may not have a backup interface.
- Neither the active nor the backup interface may be a member of a LAG.
- The active and standby do *not* have to be of the same type (1G, 10G, etc).
- You may not enable any Layer 2 protocol on any interface of a redundant pair or to ports connected to them.

In Figure 20-7, interface 3/41 is a backup interface for 3/42, and 3/42 is in the down state, as shown in message Message 1. If 3/41 fails, 3/42 transitions to the up state, which makes the backup link active. A message similar to Message 1 appears whenever you configure a backup port.

Message 1 Configuring a Backup Layer 2 Port

```
      02:28:04: %RPM0-P:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on Gi 3/41
and Gi 3/42
      02:28:04: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 3/42
      02:28:04: %RPM0-P:CP %IFMGR-5-STATE_ACT_STBY: Changed interface state to standby: Gi
3/42
```

Figure 20-7. CLI for Configuring Redundant Layer 2 Pairs without Spanning Tree

```
FTOS(conf-if-range-gi-3/41-42)#switchport backup interface GigabitEthernet 3/42
FTOS(conf-if-range-gi-3/41-42)#show config
!
interface GigabitEthernet 3/41
 no ip address
 switchport
 switchport backup interface GigabitEthernet 3/42
 no shutdown
!
interface GigabitEthernet 3/42
 no ip address
 switchport
 no shutdown
FTOS(conf-if-range-gi-3/41-42)#
FTOS(conf-if-range-gi-3/41-42)#do show ip int brief | find 3/41
GigabitEthernet 3/41     unassigned      YES Manual up                    up
GigabitEthernet 3/42     unassigned      NO  Manual up                    down
[output omitted]
FTOS(conf-if-range-gi-3/41-42)#interface gig 3/41
FTOS(conf-if-gi-3/41)#shutdown
00:24:53: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 3/41
FTOS(conf-if-gi-3/41)#00:24:55: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to
down: Gi 3/41
00:24:55: %RPM0-P:CP %IFMGR-5-INACTIVE: Changed Vlan interface state to inactive: Vl 1
00:24:55: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Gi 3/42
00:24:55: %RPM0-P:CP %IFMGR-5-ACTIVE: Changed Vlan interface state to active: Vl 1
00:24:55: %RPM0-P:CP %IFMGR-5-STATE_STBY_ACT: Changed interface state from standby to active:
Gi 3/42

FTOS(conf-if-gi-3/41)#do show ip int brief | find 3/41
GigabitEthernet 3/41     unassigned      NO  Manual administratively down down
GigabitEthernet 3/42     unassigned      YES Manual up                    up
[output omitted]
```

Figure 20-8. CLI for Redundant Pair in Port-channel on S4810

```
FTOS#show interfaces port-channel brief
Codes: L - LACP Port-channel


    LAG  Mode  Status       Uptime      Ports
    1    L2    up           00:08:33    Te 0/0     (Up)
    2    L2    up           00:00:02    Te 0/1     (Up)
FTOS#configure
FTOS(conf)#interface port-channel 1
FTOS(conf-if-po-1)#switchport backup interface port-channel 2
Apr 9 00:15:13: %STKUNIT0-M:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on Po 1 and Po 2
Apr 9 00:15:13: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 2
Apr 9 00:15:13: %STKUNIT0-M:CP %IFMGR-5-STATE_ACT_STBY: Changed interface state to standby: Po 2
FTOS(conf-if-po-1)#
FTOS#
FTOS#show interfaces switchport backup
Interface                 Status     Paired Interface          Status
Port-channel 1            Active     Port-chato mannel 2            Standby
Port-channel 2            Standby    Port-channel 1            Active
FTOS#

FTOS(conf-if-po-1)#switchport backup interface tengigabitethernet 0/2
Apr 9 00:16:29: %STKUNIT0-M:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on Po 1 and Te 0/2
FTOS(conf-if-po-1)#
```

Restricting Layer 2 Flooding

Restricting Layer 2 Flooding is supported only on platform: E

When Layer 2 multicast traffic must be forwarded on a VLAN that has multiple ports with different speeds on the same port-pipe, forwarding is limited to the speed of the slowest port. Restricted Layer 2 Flooding prevents slower ports from lowering the throughput of multicast traffic on faster ports by restricting flooding to ports with a speed equal to or above a link speed you specify.

For example, if a VLAN that has an (auto-negotiated) 100M port and a 1G port on the same port-pipe, and you enable Restricted Layer 2 Flooding with a minimum speed of 1G, multicast traffic is only flooded on the 1G port.

Enable Restricted Layer 2 Flooding using the command **restrict-flooding** from INTERFACE VLAN mode.

In combination with **restrict-flooding**, you can use the command **mac-flood-list** from CONFIGURATION mode, without the **min-speed** option, to allow some specific multicast traffic (identified using a MAC address range you specify) to be flooded on all ports regardless of the **restrict-flooding** configuration.

Conversely, if you want all multicast traffic to be flooded on all ports, but some specific traffic to be restricted, use **mac-flood-list** with the **min-speed** option, but without **restrict-flooding** configured. This configuration restricts flooding only for traffic with destination multicast MAC addresses within the multicast MAC address range you specify.

In the following example, flooding of unknown multicast traffic is restricted to 1G ports on VLAN100 using the command **restrict-flooding**. However, the command **mac-flood-list** allows traffic with MAC addresses 01:01:e8:00:00:00 to 01:01:e8:ff:ff:ff to be flooded on all ports regardless of link speed.

Figure 20-9. Restricting Layer 2 Multicast Flooding over Low Speed Ports

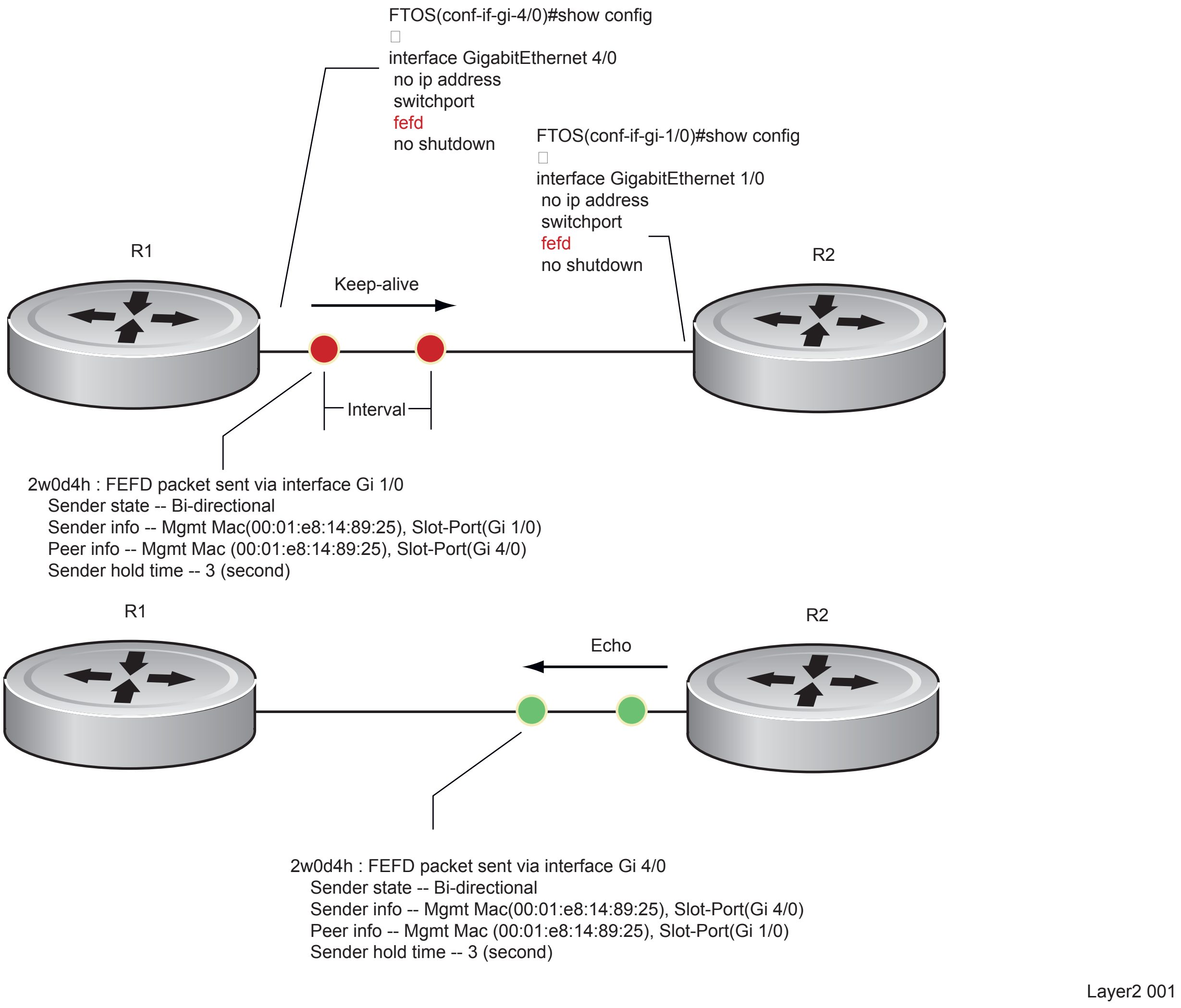
```
FTOS(conf)#$1:01:e8:00:00:00 ff:ff:ff:00:00:00 vlan 100-200,300
FTOS#show run | find mac-flood-list
mac-flood-list 01:01:e8:00:00:00 ff:ff:ff:00:00:00 vlan 100-200,300
[output omitted]
FTOS(conf)#interface vlan 100
FTOS(conf-if-vl-100)#restrict-flooding multicast min-speed 1000
FTOS(conf-if-vl-100)#show config
!
interface Vlan 100
restrict-flooding multicast min-speed 1000
 no shutdown
FTOS(conf-if-vl-100)#
```

Far-end Failure Detection

Far-end Failure Detection is supported on platforms E S4810 Z

Far-end Failure Detection (FEFD) is a protocol that senses remote data link errors in a network. It responds by sending a unidirectional report that triggers an echoed response after a specified time interval. FEFD can be enabled globally or locally on an interface basis. Disabling the global FEFD configuration does not disable the interface configuration.

Figure 20-10. Configuring Far-end Failure Detection



```
FTOS(conf-if-gi-4/0)#show config
□
interface GigabitEthernet 4/0
 no ip address
 switchport
 fefd
 no shutdown
```

```
FTOS(conf-if-gi-1/0)#show config
□
interface GigabitEthernet 1/0
 no ip address
 switchport
 fefd
 no shutdown
```

```
2w0d4h : FEFD packet sent via interface Gi 1/0
   Sender state -- Bi-directional
   Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Gi 1/0)
   Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Gi 4/0)
   Sender hold time -- 3 (second)
```

```
2w0d4h : FEFD packet sent via interface Gi 4/0
   Sender state -- Bi-directional
   Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Gi 4/0)
   Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Gi 1/0)
   Sender hold time -- 3 (second)
```

The report consists of several packets in SNAP format that are sent to the nearest known MAC address.

In the event of a far-end failure, the device stops receiving frames, and after the specified time interval, assumes that the far-end is not available. The connecting line protocol is brought down so that upper layer protocols can detect the neighbor unavailability faster.

FEFD state changes

FEFD has two operational modes, Normal and Aggressive. When Normal mode is enabled on an interface an a far-end failure is detected, no intervention is required to reset the interface to bring it back to an FEFD operational state.When Aggressive mode is enabled on an interface in the same state, manual intervention is required to reset the interface.

FEFD enabled systems (comprised of one or more interfaces) will automatically switch between four different states: Idle, Unknown, Bi-directional, and Err-disabled.

1. An interface on which FEFD is not configured is in Normal mode by default.
2. Once FEFD is enabled on an interface, it transitions to the Unknown state and sends an FEFD packet to the remote end of the link.
3. When the local interface receives the echoed packet from the remote end, the local interface transitions to the Bi-directional state.
4. If the FEFD enabled system is configured to use FEFD in Normal mode and neighboring echoes are not received after three intervals, (each interval can be set between 3 and 300 seconds by the user) the state changes to unknown.
5. If the FEFD system has been set to Aggressive mode and neighboring echoes are not received after three intervals, the state changes to Err-disabled. All interfaces in the Err-disabled state must be manually reset using the **fefd reset** [*interface*] command in EXEC privilege mode (it can be done globally or one interface at a time) before the FEFD enabled system can become operational again.

Table 20-1. State Changes When Configuring FEFD

Local Event	Mode	Local State	Remote State	Local Admin Status	Local Protocol Status	Remote Admin Status	Remote Protocol Status
Shutdown	Normal	Admin Shutdown	Unknown	Down	Down	Up	Down
Shutdown	Aggressive	Admin Shutdown	Err-disabled	Up	Down	Up	Down
FEFD enable	Normal	Bi-directional	Bi-directional	Up	Up	Up	Up
FEFD enable	Aggressive	Bi-directional	Bi-directional	Up	Up	Up	Up
FEFD + FEFD disable	Normal	Locally disabled	Unknown	Up	Down	Up	Down
FEFD + FEFD disable	Aggressive	Locally disabled	Err-disabled	Up	Down	Up	Down
Link Failure	Normal	Unknown	Unknown	Up	Down	Up	Down
Link Failure	Aggressive	Err-disabled	Err-disabled	Up	Down	Up	Down

Important Points to Remember

- FEFD enabled ports are subject to an 8 to 10 second delay during an RPM failover before becoming operational.
- FEFD can be enabled globally or on a per interface basis. Interface FEFD configurations override global FEFD configurations.
- FTOS supports FEFD on physical Ethernet interfaces only, excluding the management interface.

Configuring FEFD

You can configure FEFD for all interfaces from CONFIGURATION mode, or on individual interfaces from INTERFACE mode.

Enable FEFD Globally

To enable FEFD globally on all interfaces enter the command **fefd-global** in CONFIGURATION mode.

Report interval frequency and mode adjustments can be made by supplementing this command as well.

Step	Task	Command Syntax	Command Mode
1	Setup two or more connected interfaces for Layer 2 or Layer 3 use	**ip address** *ip address*, **switchport**	INTERFACE
2	Activate the necessary ports administratively	**no shutdown**	INTERFACE
3	Enable fefd globally	**fefd {interval | mode}**	CONFIGURATION

Entering the **show fefd** command in EXEC privilege mode displays information about the state of each interface.

Figure 20-11. Show FEFD global outputs

```
FTOS#show fefd
FEFD is globally 'ON', interval is 3 seconds, mode is 'Normal'.

INTERFACE       MODE            INTERVAL        STATE
                                (second)
Gi 1/0          Normal          3               Bi-directional
Gi 1/1          Normal          3               Admin Shutdown
Gi 1/2          Normal          3               Admin Shutdown
Gi 1/3          Normal          3               Admin Shutdown

FTOS#show run fefd
!
fefd-global mode normal
fefd-global interval 3
```

Enable FEFD on an Interface

Entering the command **fefd** in INTERFACE mode enables FEFD on a per interface basis. To change the FEFD mode, supplement the **fefd** command in INTERFACE mode by entering the command **fefd** [**mode** {**aggressive** | **normal**}].

To disable FEFD protocol on one interface, enter the command **fefd disable** in INTERFACE mode. Disabling an interface will shut down all protocols working on that interface's connected line, and will not delete your previous FEFD configuration which can be enabled again at any time.

Step	Task	Command Syntax	Command Mode
1	Setup two or more connected interfaces for Layer 2 or Layer 3 use	**ip address** *ip address*, **switchport**	INTERFACE
2	Activate the necessary ports administratively	**no shutdown**	INTERFACE
3	Enable FEFD on each interface	**fefd** {**disable |** **interval** | **mode}**	INTERFACE

Figure 20-12. FEFD enabled interface configuration

```
FTOS(conf-if-gi-1/0)#show config
!
interface GigabitEthernet 1/0
 no ip address
 switchport
 fefd mode normal
 no shutdown

FTOS(conf-if-gi-1/0)#do show fefd | grep 1/0
Gi 1/0         Normal         3             Unknown
```

Debugging FEFD

By entering the command **debug fefd events** in EXEC privilege mode, output is displayed whenever events occur that initiate or disrupt an FEFD enabled connection.

Figure 20-13. Debug FEFD events display

```
FTOS#debug fefd events
FTOS#config
FTOS(conf)#int gi 1/0
FTOS(conf-if-gi-1/0)#shutdown
2w1d22h: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 1/0
FTOS(conf-if-gi-1/0)#2w1d22h : FEFD state on Gi 1/0 changed from ANY to Unknown
2w1d22h: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 1/0
2w1d22h: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 4/0
2w1d22h: %RPM0-P:CP %IFMGR-5-INACTIVE: Changed Vlan interface state to inactive: Vl 1
2w1d22h : FEFD state on Gi 4/0 changed from Bi-directional to Unknown
```

Entering the command **debug fefd packets** in EXEC privilege mode will provide output for each packet transmission over the FEFD enabled connection.

Figure 20-14. Debug FEFD packets display

```
FTOS#debug fefd packets
FTOS#2w1d22h : FEFD packet sent via interface Gi 1/0
    Sender state -- Bi-directional
    Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Gi 1/0)
    Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Gi 4/0)
    Sender hold time -- 3 (second)

2w1d22h : FEFD packet received on interface Gi 4/0
    Sender state -- Bi-directional
    Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Gi 1/0)
    Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Gi 4/0)
    Sender hold time -- 3 (second)
```

During an RPM Failover

In the event that an RPM failover occurs, FEFD will become operationally down on all enabled ports for approximately 8-10 seconds before automatically becoming operational again.

Figure 20-15. FEFD state change during an RPM failover

```
02-05-2009          12:40:38              Local7.Debug     10.16.151.12       Feb 5 07:06:09:
%RPM1-S:CP %RAM-6-FAILOVER_REQ: RPM failover request from active peer: User request.
02-05-2009          12:40:38              Local7.Debug     10.16.151.12       Feb 5 07:06:19:
%RPM1-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Gi 0/45
02-05-2009          12:40:38              Local7.Debug     10.16.151.12       Feb 5 07:06:19:
%RPM1-P:CP %FEFD-5-FEFD-BIDIRECTION-LINK-DETECTED: Interface Gi 0/45 has bidirectional link with its
peer
```

21

Link Layer Discovery Protocol

Link Layer Discovery Protocol is supported only on platforms: C E S

This chapter contains the following sections:

- 802.1AB (LLDP) Overview on page 407
- TIA-1057 (LLDP-MED) Overview on page 410
- Configuring LLDP on page 414

802.1AB (LLDP) Overview

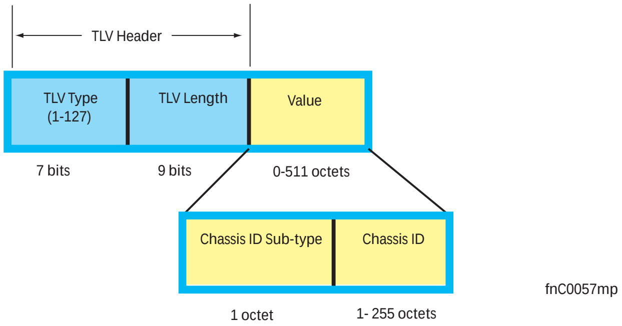
Link Layer Discovery Protocol (LLDP)—defined by IEEE 802.1AB—is a protocol that enables a LAN device to advertise its configuration and receive configuration information from adjacent LLDP-enabled LAN infrastructure devices. The collected information is stored in a management information base (MIB) on each device, and is accessible via SNMP.

Protocol Data Units

Configuration information is exchanged in the form of Type, Length, Value (TLV) segments. Figure 21-1 shows the Chassis ID TLV.

- **Type—**The kind of information included in the TLV
- **Length—**The value, in octets, of the TLV after the Length field
- **Value—**The configuration information that the agent is advertising

Figure 21-1. Type, Length, Value (TLV) Segment



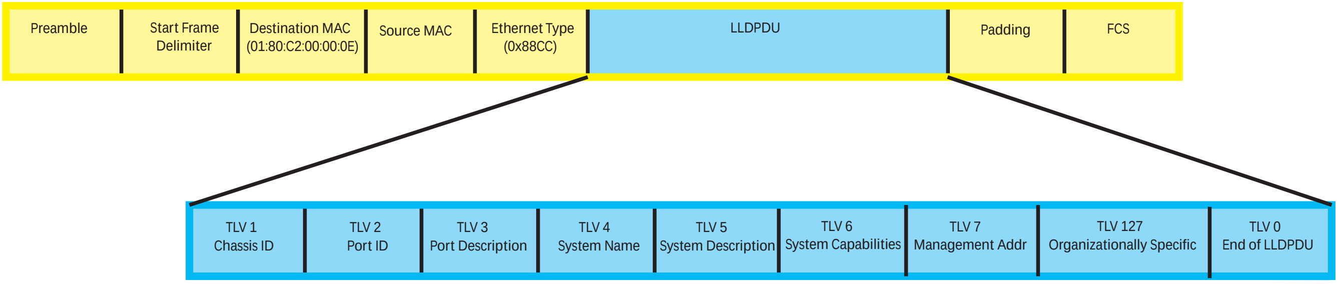
TLVs are encapsulated in a frame called an LLDP Data Unit (LLDPDU) (Figure 21-2), which is transmitted from one LLDP-enabled device to its LLDP-enabled neighbors. LLDP is a one-way protocol. LLDP-enabled devices (LLDP agents) can transmit and/or receive advertisements, but they cannot solicit and do not respond to advertisements.

There are five types of TLVs. All types are mandatory in the construction of an LLDPDU except Optional TLVs. The inclusion of individual Optional TLVs is user configurable.

Table 21-1. Type, Length, Value (TLV) Types

Type	TLV	Description
0	End of LLDPDU	Marks the end of an LLDPDU
1	Chassis ID	An administratively assigned name that identifies the LLDP agent
2	Port ID	An administratively assigned name that identifies a port through which TLVs are sent and received
3	Time to Live	A value that tells the receiving agent how long the information contained in the TLV Value field is 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 21-2. LLDPDU Frame



fnC0047mp

Optional TLVs

FTOS supports the following optional TLVs:

- Management TLVs
- IEEE 802.1 and 802.3 Organizationally Specific TLVs
- TIA-1057 Organizationally Specific TLVs

Management TLVs

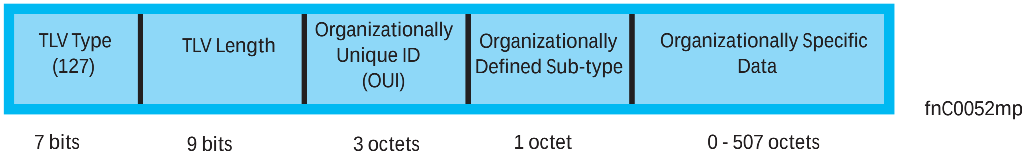
A Management TLV is an Optional TLVs sub-type. This kind of TLV contains essential management information about the sender. The five types are described in Table 21-2.

Organizationally Specific TLVs

Organizationally specific TLVs can be defined by a professional organization or a vendor. They have two mandatory fields (Figure 21-3) in addition to the basic TLV fields (Figure 21-1):

- 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 21-3. Organizationally Specific TLV


TLV Type (127)
TLV Length
Organizationally Unique ID (OUI)
Organizationally Defined Sub-type
Organizationally Specific Data
fnC0052mp
7 bits
9 bits
3 octets
1 octet
0 - 507 octets

IEEE Organizationally Specific TLVs

Eight TLV types have been defined by the IEEE 802.1 and 802.3 working groups (Table 21-2) as a basic part of LLDP; the IEEE OUI is 00-80-C2. You can configure the Dell Force10 system to advertise any or all of these TLVs.

Table 21-2. Optional TLV Types

Type	TLV	Description
Optional TLVs		
4	Port description	A user-defined alphanumeric string that describes the port. FTOS 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 describes the system
7	System capabilities	Identifies the chassis as one or more of the following: repeater, bridge, WLAN Access Point, Router, Telephone, DOCSIS cable device, end station only, or other
8	Management address	Indicates the network address of the management interface. FTOS does not currently support this TLV.
IEEE 802.1 Organizationally Specific TLVs		
127	Port-VLAN ID	On Dell Force10 systems, indicates the untagged VLAN to which a port belongs

Table 21-2. Optional TLV Types

Type	TLV	Description
127	Port and Protocol VLAN ID	On Dell Force10 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. This TLV is supported on C-Series only.
127	Protocol Identity	Indicates the protocols that the port can process. FTOS 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 FTOS implementation of LLDP, but is available and mandatory (non-configurable) in the LLDP-MED implementation.
127	Power via MDI	Dell Force10 supports LLDP-MED protocol, which recommends that Power via MDI TLV be not implemented, and therefore Dell Force10 implements Extended Power via MDI TLV only.
127	Link Aggregation	Indicates whether the link is capable of being aggregated, whether it is currently in a LAG, and the port identification of the LAG. FTOS does not currently support this TLV.
127	Maximum Frame Size	Indicates the maximum frame size capability of the MAC and PHY

TIA-1057 (LLDP-MED) Overview

Link Layer Discovery Protocol—Media Endpoint Discovery (LLDP-MED)—as defined by ANSI/ TIA-1057— provides additional organizationally specific TLVs so that endpoint devices and network connectivity devices can advertise their characteristics and configuration information; the OUI for the Telecommunications Industry Association (TIA) is 00-12-BB.

- **LLDP-MED Endpoint Device**—any device that is on an IEEE 802 LAN network edge can communicate using IP and uses the LLDP-MED framework.
- **LLDP-MED Network Connectivity Device**—any device that provides access to an IEEE 802 LAN to an LLDP-MED endpoint device and supports IEEE 802.1AB (LLDP) and TIA-1057 (LLDP-MED). The Dell Force10 system is an LLDP-MED network connectivity device.

With regard to connected endpoint devices, LLDP-MED provides network connectivity devices with the ability to:

- manage inventory
- manage Power over Ethernet (PoE)
- identify physical location
- identify network policy

LLDP-MED is designed for, but not limited to, VoIP endpoints.

TIA Organizationally Specific TLVs

The Dell Force10 system is an LLDP-MED Network Connectivity Device (Device Type 4). Network connectivity devices are responsible for:

- transmitting an LLDP-MED capabilities TLV to endpoint devices
- storing the information that endpoint devices advertise

Table 21-3 describes the five types of TIA-1057 Organizationally Specific TLVs.

Table 21-3. TIA-1057 (LLDP-MED) Organizationally Specific TLVs

Type	Sub-type	TLV	Description
127	1	LLDP-MED Capabilities	Indicates: • whether the transmitting device supports LLDP-MED • what LLDP-MED TLVs it supports • LLDP device class
127	2	Network Policy	Indicates the application type, VLAN ID, Layer 2 Priority, and DSCP value
127	3	Location Identification	Indicates the physical location of the device expressed in one of three possible formats: • Coordinate Based LCI • Civic Address LCI • Emergency Call Services ELIN
127	4	Extended Power via MDI	Indicates power requirements, priority, and power status
Inventory Management TLVs			Implementation of this set of TLVs is optional in LLDP-MED devices. None or all TLVs must be supported. FTOS does not currently support these TLVs.
127	5	Inventory - Hardware Revision	Indicates the hardware revision of the LLDP-MED device
127	6	Inventory - Firmware Revision	Indicates the firmware revision of the LLDP-MED device
127	7	Inventory - Software Revision	Indicates the software revision of the LLDP-MED device
127	8	Inventory - Serial Number	Indicates the device serial number of the LLDP-MED device
127	9	Inventory - Manufacturer Name	Indicates the manufacturer of the LLDP-MED device
127	10	Inventory - Model Name	Indicates the model of the LLDP-MED device
127	11	Inventory - Asset ID	Indicates a user specified device number to manage inventory
127	12-255	Reserved	—

LLDP-MED Capabilities TLV

The LLDP-MED Capabilities TLV communicates the types of TLVs that the endpoint device and the network connectivity device support. LLDP-MED network connectivity devices must transmit the Network Policies TLV.

- The value of the LLDP-MED Capabilities field in the TLV is a 2 octet bitmap (Figure 21-4), each bit represents an LLDP-MED capability (Table 21-4).
- The possible values of the LLDP-MED Device Type is listed in Table 21-5. The Dell Force10 system is a Network Connectivity device, which is Type 4.

When you enable LLDP-MED in FTOS (using the command **advertise med**) the system begins transmitting this TLV.

Figure 21-4. LLDP-MED Capabilities TLV

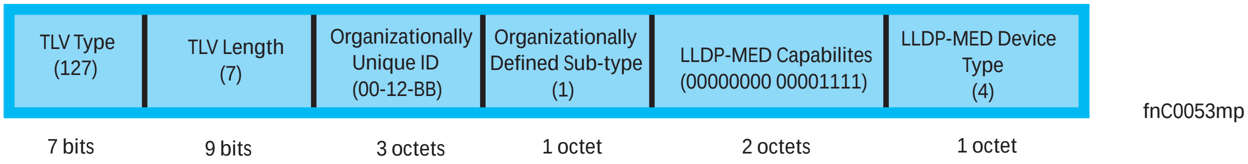


Table 21-4. FTOS LLDP-MED Capabilities

Bit Position	TLV	FTOS 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 21-5. 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, specifically:

- VLAN ID
- VLAN tagged or untagged status
- Layer 2 priority
- DSCP value

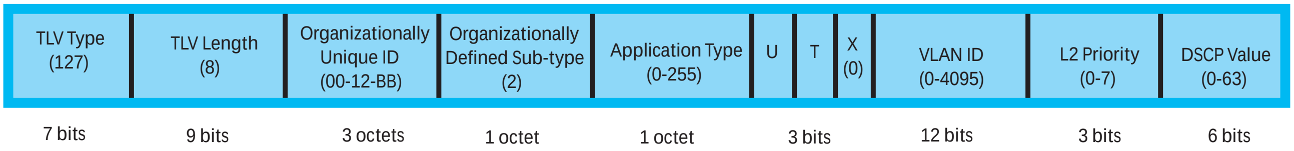
The application type is a represented by an integer (the Type integer in Table 21-6), which indicates a device function for which a unique network policy is defined. An individual LLDP-MED Network Policy TLV is generated for each application type that you specify with the FTOS CLI (Advertising TLVs on page 416).

Note: With regard to Table 21-6, 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 21-6. 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.
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	Softphone is a computer program that enables IP telephony on a computer, rather than using a phone. Specify this application type for this type of endpoint device.
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 broadcast or multicast based video content distribution and other similar applications supporting streaming video services. This does not include video applications relying on TCP with buffering.
8	Video Signaling	Specify this application type only if video control packets use a separate network policy than video data.
9-255	Reserved	—

Figure 21-5. LLDP-MED Policies TLV



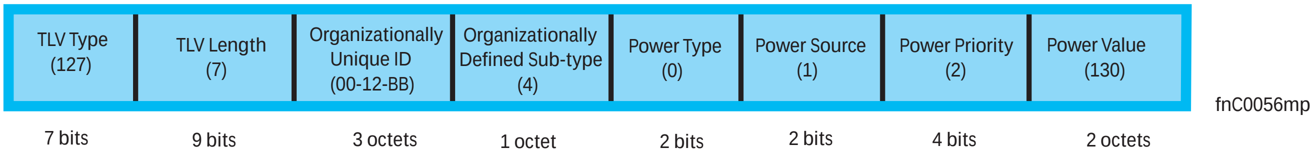
TLV Type (127)	TLV Length (8)	Organizationally Unique ID (00-12-BB)	Organizationally Defined Sub-type (2)	Application Type (0-255)	U	T	X (0)	VLAN ID (0-4095)	L2 Priority (0-7)	DSCP Value (0-63)
7 bits	9 bits	3 octets	1 octet	1 octet	3 bits			12 bits	3 bits	6 bits

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 Sourcing Entity (PSE) or Power Device (PD). The Dell Force10 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 Force10 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 Force10 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 Force10 also honors the power priority value sent by the powered device. However, the CLI configuration takes precedence.
- **Power Value**—Dell Force10 advertises the maximum amount of power that can be supplied on the port. By default it 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 Force10 also honors the power value (power requirement) sent by the powered device when the port is configured for **power inline auto**.

Figure 21-6. Extended Power via MDI TLV



TLV Type (127)	TLV Length (7)	Organizationally Unique ID (00-12-BB)	Organizationally Defined Sub-type (4)	Power Type (0)	Power Source (1)	Power Priority (2)	Power Value (130)
7 bits	9 bits	3 octets	1 octet	2 bits	2 bits	4 bits	2 octets

Configuring LLDP

Configuring LLDP is a two-step process:

1. Enable LLDP globally. See page 416.
2. Advertise TLVs out of an interface. See page 416.

Related Configuration Tasks

- Viewing the LLDP Configuration on page 418

Important Points to Remember

- LLDP is disabled by default.
- Dell Force10 systems support up to 8 neighbors per interface.
- Dell Force10 systems support a maximum of 8000 total neighbors per system. If the number of interfaces multiplied by 8 exceeds the maximum, the system will not configure more than 8000.
- INTERFACE level configurations override all CONFIGURATION level configurations.
- LLDP is not hitless.

LLDP Compatibility

- Spanning Tree and Dell Force10 Ring Protocol “blocked” ports allow LLDPDUs.
- 802.1X controlled ports do not allow LLDPDUs until the connected device is authenticated.

CONFIGURATION versus INTERFACE Configurations

All LLDP configuration commands are available in PROTOCOL LLDP mode, which is a sub-mode of CONFIGURATION mode and INTERFACE mode.

- Configurations made at CONFIGURATION level are global, that is, they affect all interfaces on the system.
- Configurations made at INTERFACE level affect only the specific interface, and they override CONFIGURATION level configurations.

Figure 21-7. Configuration and Interface mode LLDP Commands

```
R1(conf)#protocol lldp
R1(conf-lldp)#?
advertise               Advertise TLVs
disable                 Disable LLDP protocol globally
end                     Exit from configuration mode
exit                    Exit from LLDP configuration mode
hello                   LLDP hello configuration
mode                    LLDP mode configuration (default = rx and tx)
multiplier              LLDP multiplier configuration
no                      Negate a command or set its defaults
show                    Show LLDP configuration
R1(conf-lldp)#exit
R1(conf)#interface gigabitethernet 1/31
R1(conf-if-gi-1/31)#protocol lldp
R1(conf-if-gi-1/31-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
R1(conf-if-gi-1/31-lldp)#
```

Enabling LLDP

LLDP is disabled by default. LLDP can be enabled and disabled globally or per interface. If LLDP is enabled globally, all up interfaces send periodic LLDPDUs. To enable LLDP:

Step	Task	Command	Command Mode
1	Enter Protocol LLDP mode.	**protocol lldp**	CONFIGURATION or INTERFACE
2	Enable LLDP.	**no disable**	PROTOCOL LLDP

Disabling and Undoing LLDP

- Disable LLDP globally or for an interface using the command **disable**.
- Undo an LLDP configuration by preceding 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 will send LLDPDUs with the specified TLVs.

- If you configure an interface, only the interface will send LLDPDUs with the specified TLVs.

If LLDP is configured both globally and at interface level, the interface level configuration overrides the global configuration. To advertise TLVs:

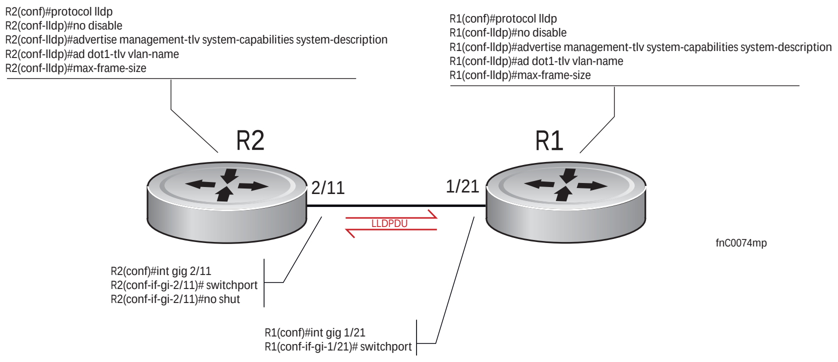
Step	Task	Command	Command Mode
1	Enter LLDP mode.	**protocol lldp**	CONFIGURATION or INTERFACE
2	Advertise one or more TLVs. Include the keyword for each TLV you want to advertise. • For management TLVs: **system-capabilities**, **system-description** • For 802.1 TLVs: **port-protocol-vlan-id**, **port-vlan-id, vlan-name** • For 802.3 TLVs: **max-frame-size** • For TIA-1057 TLVs: **•guest-voice** **•guest-voice-signaling** **•location-identification** **•power-via-mdi** **•softphone-voice** **•streaming-video** **•video-conferencing** **•video-signaling** **•voice** **•voice-signaling**	**advertise {management-tlv | dot1-tlv | dot3-tlv | med}**	PROTOCOL LLDP

*

Note: vlan-name is supported on C-Series and S-Series only.

In Figure 21-8, LLDP is enabled globally. R1 and R2 are transmitting periodic LLDPDUs that contain management, 802.1, and 802.3 TLVs.

Figure 21-8. Configuring LLDP



```
R2(conf)#protocol lldp
R2(conf-lldp)#no disable
R2(conf-lldp)#advertise management-tlv system-capabilities system-description
R2(conf-lldp)#ad dot1-tlv vlan-name
R2(conf-lldp)#max-frame-size
```

```
R1(conf)#protocol lldp
R1(conf-lldp)#no disable
R1(conf-lldp)#advertise management-tlv system-capabilities system-description
R1(conf-lldp)#ad dot1-tlv vlan-name
R1(conf-lldp)#max-frame-size
```

```
R2(conf)#int gig 2/11
R2(conf-if-gi-2/11)# switchport
R2(conf-if-gi-2/11)#no shut
```

```
R1(conf)#int gig 1/21
R1(conf-if-gi-1/21)# switchport
```

Viewing the LLDP Configuration

Display the LLDP configuration using the command **show config** in either CONFIGURATION or INTERFACE mode, as shown in Figure 21-9 and Figure 21-10, respectively

Figure 21-9. Viewing LLDP Global Configurations

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 hello 10
 no disable
R1(conf-lldp)#
```

Figure 21-10. Viewing LLDP Interface Configurations

```
R1(conf-lldp)#exit
R1(conf)#interface gigabitethernet 1/31
R1(conf-if-gi-1/31)#show config
!
interface GigabitEthernet 1/31
 no ip address
 switchport
 no shutdown
R1(conf-if-gi-1/31)#protocol lldp
R1(conf-if-gi-1/31-lldp)#show config
!
 protocol lldp
R1(conf-if-gi-1/31-lldp)#
```

Viewing Information Advertised by Adjacent LLDP Agents

Display brief information about adjacent devices using the command **show lldp neighbors**, as shown in Figure 21-11. Display all of the information that neighbors are advertising using the command **show lldp neighbors detail**, as shown in Figure 21-12.

Figure 21-11. Viewing Brief Information Advertised by Adjacent LLDP Agents

```
R1(conf-if-gi-1/31-lldp)#end
R1(conf-if-gi-1/31)#do show lldp neighbors
 Loc PortID   Rem Host Name      Rem Port Id           Rem Chassis Id
 -------------------------------------------------------------------------

 Gi 1/21      -                  GigabitEthernet 2/11  00:01:e8:06:95:3e
 Gi 1/31      -                  GigabitEthernet 3/11  00:01:e8:09:c2:4a
```

Figure 21-12. Viewing All Information Advertised by Adjacent LLDP Agent

```
R1#show lldp neighbors detail
========================================================================
 Local Interface Gi 1/21 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:  GigabitEthernet 2/11
    Local Port ID: GigabitEthernet 1/21
    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:  Force10 Networks Real Time Operating System Software
     . Force10 Operating System Version: 1.0. Force10 App
     lication Software Version: 7.5.1.0. Copyright (c) 19
     99-Build Time: Thu Aug 9 01:05:51 PDT 2007
    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. You can configure a non-default transmit interval—at CONFIGURATION level or INTERFACE level—using the command **hello** (Figure 21-13).

Figure 21-13. Configuring LLDPDU Transmit and Receive Mode

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#mode ?
rx                      Rx only
tx                      Tx only
R1(conf-lldp)#mode tx
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 mode tx
 no disable
R1(conf-lldp)#no mode
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#
```

Configuring Transmit and Receive Mode

Once LLDP is enabled, Dell Force10 systems transmit *and* receive LLDPDUs by default. You can configure the system—at CONFIGURATION level or INTERFACE level—to transmit only by executing the command **mode tx**, or receive only by executing the command **mode rx**. Return to the default with the **no mode** command (Figure 21-14).

Figure 21-14. Configuring LLDPDU Transmit and Receive Mode

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#mode ?
rx                      Rx only
tx                      Tx only
R1(conf-lldp)#mode tx
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 mode tx
 no disable
R1(conf-lldp)#no mode
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#
```

Configuring 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—at CONFIGURATION level or INTERFACE level—using the command **multiplier**. Return to the default multiplier value using the command **no multiplier** (Figure 21-15).

Figure 21-15. Configuring LLDPDU Time to Live

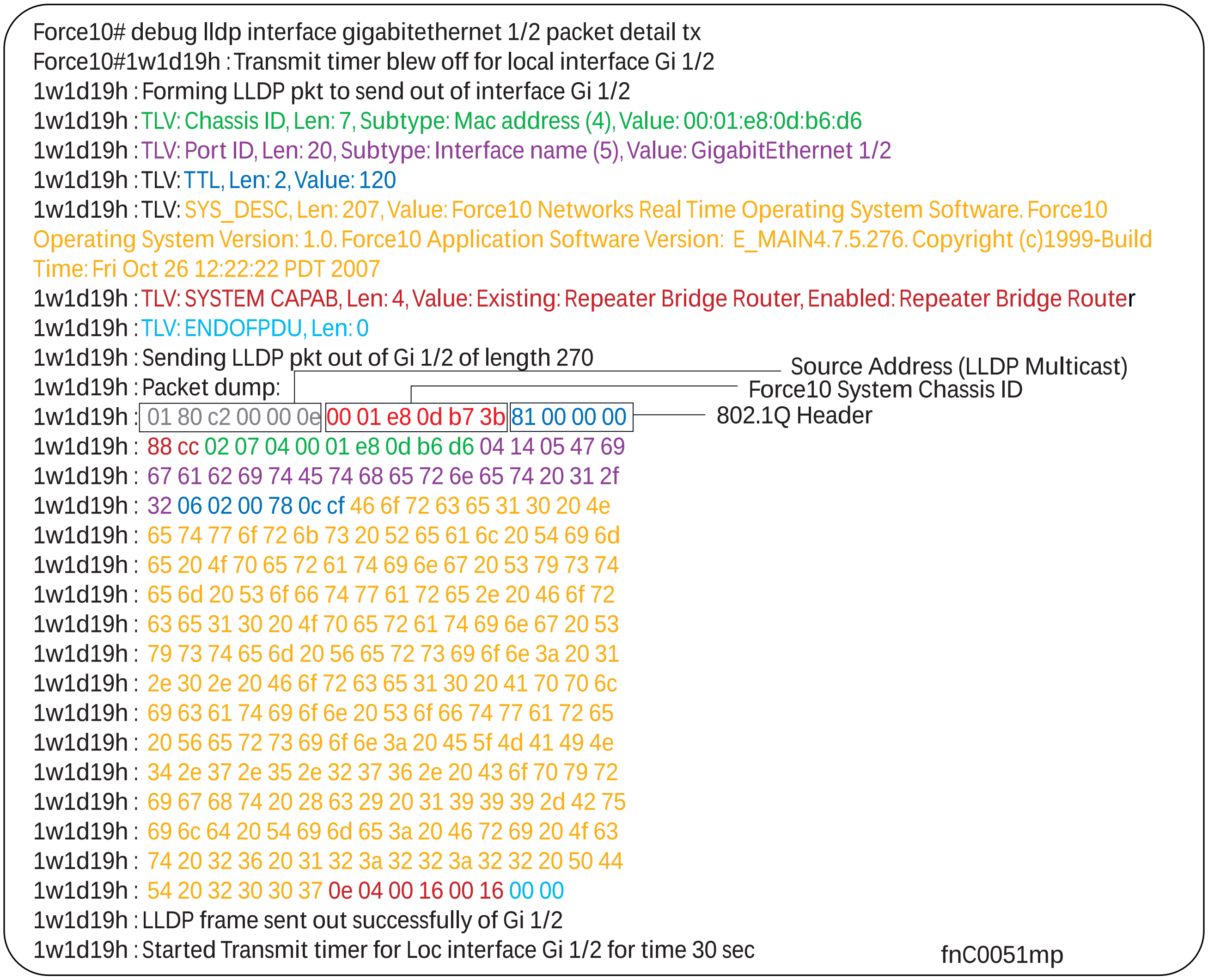
```
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#multiplier ?
<2-10>                  Multiplier (default=4)
R1(conf-lldp)#multiplier 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)#
```

Debugging LLDP

The command **debug lldp** enables you to view the TLVs that your system is sending and receiving.

- Use the **debug lldp brief** command to view a readable version of the TLVs.
- Use the **debug lldp detail** command to view a readable version of the TLVs plus a hexadecimal version of the entire LLDPDU.

Figure 21-16. debug lldp detail—LLDPDU Packet Dissection



```
Force10# debug lldp interface gigabitethernet 1/2 packet detail tx
Force10#1w1d19h : Transmit timer blew off for local interface Gi 1/2
1w1d19h : Forming LLDP pkt to send out of interface Gi 1/2
1w1d19h : TLV: Chassis ID, Len: 7, Subtype: Mac address (4), Value: 00:01:e8:0d:b6:d6
1w1d19h : TLV: Port ID, Len: 20, Subtype: Interface name (5), Value: GigabitEthernet 1/2
1w1d19h : TLV: TTL, Len: 2, Value: 120
1w1d19h : TLV: SYS_DESC, Len: 207, Value: Force10 Networks Real Time Operating System Software. Force10
Operating System Version: 1.0. Force10 Application Software Version:  E_MAIN4.7.5.276. Copyright (c)1999-Build
Time: Fri Oct 26 12:22:22 PDT 2007
1w1d19h : TLV: SYSTEM CAPAB, Len: 4, Value: Existing: Repeater Bridge Router, Enabled: Repeater Bridge Router
1w1d19h : TLV: ENDOFPDU, Len: 0
1w1d19h : Sending LLDP pkt out of Gi 1/2 of length 270
1w1d19h : Packet dump:
1w1d19h : 01 80 c2 00 00 0e 00 01 e8 0d b7 3b 81 00 00 00
1w1d19h : 88 cc 02 07 04 00 01 e8 0d b6 d6 04 14 05 47 69
1w1d19h : 67 61 62 69 74 45 74 68 65 72 6e 65 74 20 31 2f
1w1d19h : 32 06 02 00 78 0c cf 46 6f 72 63 65 31 30 20 4e
1w1d19h : 65 74 77 6f 72 6b 73 20 52 65 61 6c 20 54 69 6d
1w1d19h : 65 20 4f 70 65 72 61 74 69 6e 67 20 53 79 73 74
1w1d19h : 65 6d 20 53 6f 66 74 77 61 72 65 2e 20 46 6f 72
1w1d19h : 63 65 31 30 20 4f 70 65 72 61 74 69 6e 67 20 53
1w1d19h : 79 73 74 65 6d 20 56 65 72 73 69 6f 6e 3a 20 31
1w1d19h : 2e 30 2e 20 46 6f 72 63 65 31 30 20 41 70 70 6c
1w1d19h : 69 63 61 74 69 6f 6e 20 53 6f 66 74 77 61 72 65
1w1d19h : 20 56 65 72 73 69 6f 6e 3a 20 45 5f 4d 41 49 4e
1w1d19h : 34 2e 37 2e 35 2e 32 37 36 2e 20 43 6f 70 79 72
1w1d19h : 69 67 68 74 20 28 63 29 20 31 39 39 39 2d 42 75
1w1d19h : 69 6c 64 20 54 69 6d 65 3a 20 46 72 69 20 4f 63
1w1d19h : 74 20 32 36 20 31 32 3a 32 32 3a 32 32 20 50 44
1w1d19h : 54 20 32 30 30 37 0e 04 00 16 00 16 00 00
1w1d19h : LLDP frame sent out successfully of Gi 1/2
1w1d19h : Started Transmit timer for Loc interface Gi 1/2 for time 30 sec
```

Relevant Management Objects

FTOS supports all IEEE 802.1AB MIB objects.

- Table 21-7 lists the objects associated with received and transmitted TLVs.
- Table 21-8 lists the objects associated with the LLDP configuration on the local agent.
- Table 21-9 lists the objects associated with IEEE 802.1AB Organizationally Specific TLVs.
- Table 21-10 lists the objects associated with received and transmitted LLDP-MED TLVs.

Table 21-7. LLDP Configuration MIB Objects

MIB Object Category	LLDP Variable	LLDP MIB Object	Description
LLDP Configuration	adminStatus	lldpPortConfigAdminStatus	Whether the local LLDP agent is enabled for transmit, receive, or both
	msgTxHold	lldpMessageTxHoldMultiplier	Multiplier value
	msgTxInterval	lldpMessageTxInterval	Transmit Interval value
	rxInfoTTL	lldpRxInfoTTL	Time to Live for received TLVs
	txInfoTTL	lldpTxInfoTTL	Time to Live for transmitted TLVs
Basic TLV Selection	mibBasicTLVsTxEnable	lldpPortConfigTLVsTxEnable	Indicates which management TLVs are enabled for system ports
	mibMgmtAddrInstanceTxEnable	lldpManAddrPortsTxEnable	The management addresses defined for the system and and the ports through which they are enabled for transmission
LLDP Statistics	statsAgeoutsTotal	lldpStatsRxPortAgeoutsTotal	Total number of times that a neighbors 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 21-8. 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

Table 21-8. LLDP System MIB Objects

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
8	Management Address	enabled capabilities	Local	lldpLocSysCapEnabled
			Remote	lldpRemSysCapEnabled
		management address length	Local	lldpLocManAddrLen
			Remote	lldpRemManAddrLen
		management address subtype	Local	lldpLocManAddrSubtype
			Remote	lldpRemManAddrSubtype
		management address	Local	lldpLocManAddr
			Remote	lldpRemManAddr
		interface numbering subtype	Local	lldpLocManAddrIfSubtype
			Remote	lldpRemManAddrIfSubtype
		interface number	Local	lldpLocManAddrIfId
			Remote	lldpRemManAddrIfId
		OID	Local	lldpLocManAddrOID
			Remote	lldpRemManAddrOID

Table 21-9. 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 21-10. 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 lldpXMedPortConfigTLVsTxEnable
			Remote	lldpXMedRemCapSupported, lldpXMedRemConfigTLVsTxEnable
		LLDP-MED Class Type	Local	lldpXMedLocDeviceClass
			Remote	lldpXMedRemDeviceClass

Table 21-10. LLDP-MED System MIB Objects

TLV Sub-Type	TLV Name	TLV Variable	System	LLDP-MED MIB Object
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
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

22

Multiple Spanning Tree Protocol

Multiple Spanning Tree Protocol is supported on platforms: C E S

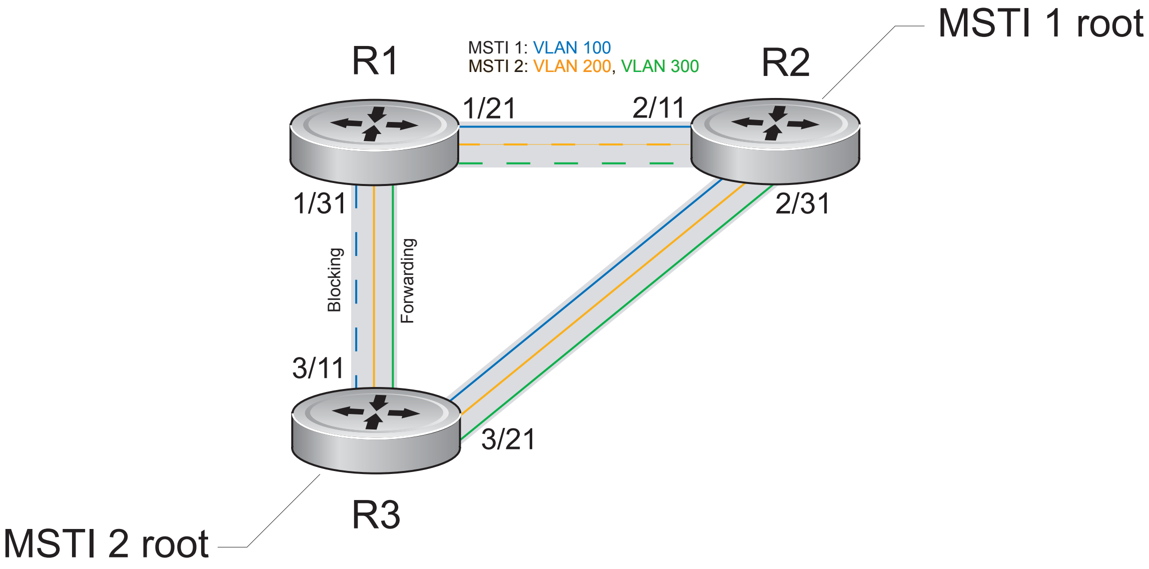
Protocol Overview

Multiple Spanning Tree Protocol (MSTP)—specified in IEEE 802.1Q-2003—is an RSTP-based spanning tree variation that improves on PVST+. MSTP allows multiple spanning tree instances and allows you to map many VLANs to one spanning tree instance to reduce the total number of required instances.

In contrast, PVST+ allows a spanning tree instance for each VLAN. This 1:1 approach is not suitable if you have many VLANs, because each spanning tree instance costs bandwidth and processing resources.

In Figure 282, 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 in Figure 282 demonstrates how you can use MSTP to achieve load balancing.

Figure 22-1. MSTP with Three VLANs Mapped to Two Spanning Tree Instances



FTOS supports three other variations of Spanning Tree, as shown in Table 44.

Table 22-1. FTOS Supported Spanning Tree Protocols

Dell Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Implementation Information

- The FTOS 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.
- FTOS supports only one MSTP region.
- When you enable MSTP, all ports in Layer 2 mode participate in MSTP.
- On the C-Series and S-Series, 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. See page 703.
2. Place the interfaces in VLANs.
3. Enable Multiple Spanning Tree Protocol. See page 429.
4. Create Multiple Spanning Tree Instances, and map VLANs to them. See page 429.

Related Configuration Tasks

- Create Multiple Spanning Tree Instances on page 429
- Add and Remove Interfaces on page 429
- Influence MSTP Root Selection on page 431
- Interoperate with Non-FTOS Bridges on page 431
- Modify Global Parameters on page 432
- Modify Interface Parameters on page 433
- Configure an EdgePort on page 434
- Flush MAC Addresses after a Topology Change on page 435
- Debugging and Verifying MSTP Configuration on page 440
- Preventing Network Disruptions with BPDU Guard on page 711

- SNMP Traps for Root Elections and Topology Changes on page 616
- Configuring Spanning Trees as Hitless on page 713

Enable Multiple Spanning Tree Globally

MSTP is not enabled by default. To enable MSTP:

Step	Task	Command Syntax	Command Mode
1	Enter PROTOCOL MSTP mode.	**protocol spanning-tree mstp**	CONFIGURATION
2	Enable MSTP.	**no disable**	PROTOCOL MSTP

Verify that MSTP is enabled using the **show config** command from PROTOCOL MSTP mode, as shown in Figure 22-2.

Figure 22-2. Verifying MSTP is Enabled

```
Force10(conf)#protocol spanning-tree mstp
Force10(config-mstp)#show config
!
protocol spanning-tree mstp
 no disable
Force10#
```

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.

Add and Remove Interfaces

- To add an interface to the MSTP topology, configure it for Layer 2 and add it to a VLAN. If you previously disabled MSTP on the interface using the command **no spanning-tree 0**, re-enable it using the command **spanning-tree 0**.
- Remove an interface from the MSTP topology using the command **no spanning-tree 0** command. See also Removing an Interface from the Spanning Tree Group on page 707 for BPDU Filtering behavior.

Create Multiple Spanning Tree Instances

A single MSTI provides no more benefit than RSTP. To take full advantage of MSTP you must create multiple MSTIs and map VLANs to them.

Create an MSTI using the command **msti** from PROTOCOL MSTP mode. Specify the keyword **vlan** followed by the VLANs that you want to participate in the MSTI, as shown in Figure 22-3.

Figure 22-3. Mapping VLANs to MSTI Instances

```
Force10(conf)#protocol spanning-tree mstp
Force10(conf-mstp)#msti 1 vlan 100
Force10(conf-mstp)#msti 2 vlan 200-300
Force10(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. View to which instance a VLAN is mapped using the command **show spanning-tree mst vlan** from EXEC Privilege mode, as shown in Figure 22-6.

View the forwarding/discarding state of the ports participating in an MSTI using the command **show spanning-tree msti** from EXEC Privilege mode, as shown in Figure 22-4.

Figure 22-4. Viewing MSTP Port States

```
Force10#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 Gi 1/21

Port 374 (GigabitEthernet 1/21) is root Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.374
Designated root has priority 32768, address 0001.e806.953e
Designated bridge has priority 32768, address 0001.e806.953e
Designated port id is 128.374, designated path cost 20000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 93671, received 46843
The port is not in the Edge port mode

Port 384 (GigabitEthernet 1/31) is alternate Discarding
Port path cost 20000, Port priority 128, Port Identifier 128.384
Designated root has priority 32768, address 0001.e806.953e
Designated bridge has priority 32768, address 0001.e809.c24a
Designated port id is 128.384, designated path cost 20000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 39291, received 7547
The port is not in the Edge port mode
```

Influence MSTP Root Selection

MSTP determines the root bridge, but you can assign one bridge a lower priority to increase the probability that it will become the root bridge.

To change the bridge priority:

Task	Command Syntax	Command Mode
Assign a number as the bridge priority. A lower number increases the probability that the bridge becomes the root bridge. **Range**: 0-61440, in increments of 4096 **Default**: 32768	**msti** *instance* **bridge-priority** *priority*	PROTOCOL MSTP

The simple configuration Figure 22-1 by default yields the same forwarding path for both MSTIs. Figure 22-5, shows how R3 is assigned bridge priority 0 for MSTI 2, which elects a different root bridge than MSTI 2. View the bridge priority using the command **show config** from PROTOCOL MSTP mode, also shown in Figure 22-5.

Figure 22-5. Changing the Bridge Priority

```
R3(conf-mstp)#msti 2 bridge-priority 0
1d2h51m: %RPM0-P:RP2 %SPANMGR-5-STP_ROOT_CHANGE: MSTP root changed for instance 2. My
Bridge ID: 0:0001.e809.c24a Old Root: 32768:0001.e806.953e New Root: 0:0001.e809.c24a

R3(conf-mstp)#show config
!
protocol spanning-tree mstp
 no disable
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
 MSTI 2 bridge-priority 0
```

Interoperate with Non-FTOS Bridges

FTOS 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 on FTOS is null.
- **Revision** is a two-byte number. The default revision number on FTOS 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 will match on all Dell Force10 FTOS equipment. If you have non-FTOS equipment that will participate in MSTP, ensure these values to match on all the equipment.

Note: Some non-FTOS equipment may implement a non-null default region name. SFTOS, for example, uses the Bridge ID, while others may use a MAC address.

To change the region name or revision:

Task	Command Syntax	Command Mode
Change the region name.	**name** *name*	PROTOCOL MSTP
Change the region revision number. • Range: 0 to 65535 • Default: 0	**revision** *number*	PROTOCOL MSTP

View the current region name and revision using the command **show spanning-tree mst configuration** from EXEC Privilege mode, as shown in Figure 22-6.

Figure 22-6. Viewing the MSTP Region Name and Revision

```
Force10(conf-mstp)#name my-mstp-region
Force10(conf-mstp)#exit
Force10(conf)#do show spanning-tree mst config
MST region name: my-mstp-region
Revision: 0
MSTI     VID
   1     100
   2     200-300
```

Modify 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** is the amount of time an interface waits in the Listening State and the Learning State before it transitions to the Forwarding State.
- **Hello-time** is the time interval in which the bridge sends MSTP Bridge Protocol Data Units (BPDUs).
- **Max-age** is the length of time the bridge maintains configuration information before it refreshes that information by recomputing the MST topology.
- **Max-hops** is the maximum number of hops a BPDU can travel before a receiving switch discards it.

Note: Dell Dell Force10 recommends that only experienced network administrators change MSTP parameters. Poorly planned modification of MSTP parameters can negatively impact network performance.

To change MSTP parameters, use the following commands on the root bridge:

Task	Command Syntax	Command Mode
Change the forward-delay parameter. • Range: 4 to 30 • Default: 15 seconds	**forward-delay** *seconds*	PROTOCOL MSTP

Task	Command Syntax	Command Mode
Change the hello-time parameter. **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**hello-time** *seconds*	PROTOCOL MSTP
Change the max-age parameter. Range: 6 to 40 Default: 20 seconds	**max-age** *seconds*	PROTOCOL MSTP
Change the max-hops parameter. Range: 1 to 40 Default: 20	**max-hops** *number*	PROTOCOL MSTP

View the current values for MSTP parameters using the **show running-config spanning-tree mstp** command from EXEC privilege mode.

Figure 22-7. Viewing the Current Values for MSTP Parameters

```
Force10(conf-mstp)#forward-delay 16
Force10(conf-mstp)#exit
Force10(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
Force10(conf)#
```

Modify 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 will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

Table 22-2 lists the default values for port cost by interface.

Table 22-2. MSTP Default Port Cost Values

Port Cost	Default Value
100-Mb/s Ethernet interfaces	200000
1-Gigabit Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
Port Channel with 100 Mb/s Ethernet interfaces	180000
Port Channel with 1-Gigabit Ethernet interfaces	18000
Port Channel with 10-Gigabit Ethernet interfaces	1800

To change the port cost or priority of an interface:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 200000 Default: see Table 22-2.	**spanning-tree msti** *number* **cost** *cost*	INTERFACE
Change the port priority of an interface. Range: 0 to 240, in increments of 16 Default: 128	**spanning-tree msti** *number* **priority** *priority*	INTERFACE

View the current values for these interface parameters using the command **show config** from INTERFACE mode. See Figure 22-8.

Configure 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 shutdown 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 will drop 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 it is enabled on an interface connected to a network.

To enable EdgePort on an interface, use the following command:

Task	Command Syntax	Command Mode
Enable EdgePort on an interface.	**spanning-tree mstp edge-port** [**bpduguard |** **shutdown-on-violation**]	INTERFACE

Verify that EdgePort is enabled on a port using the command **show config** from the INTERFACE mode, as shown in Figure 22-8.

FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

The error disabled state can be cleared with any of the following methods:

- Perform an **shutdown** command on the interface.
- Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).
- Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
- Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Figure 22-8. Configuring EdgePort

```
Force10(conf-if-gi-3/41)#spanning-tree mstp edge-port
Force10(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
Force10(conf-if-gi-3/41)#
```

Flush MAC Addresses after a Topology Change

FTOS 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 command **tc-flush-standard**, which flushes MAC addresses upon every topology change notification. View the enable status of this feature using the command **show running-config spanning-tree mstp** from EXEC Privilege mode.

MSTP Sample Configurations

The running-configurations in Figure 22-10, Figure 22-11, and Figure 22-11 support the topology shown in Figure 22-9. The configurations are from FTOS systems. An S50 system using SFTOS, configured as shown Figure 22-13, could be substituted for an FTOS router in this sample following topology and MSTP would function as designed.

Figure 22-9. MSTP with Three VLANs Mapped to Two Spanning Tree Instances

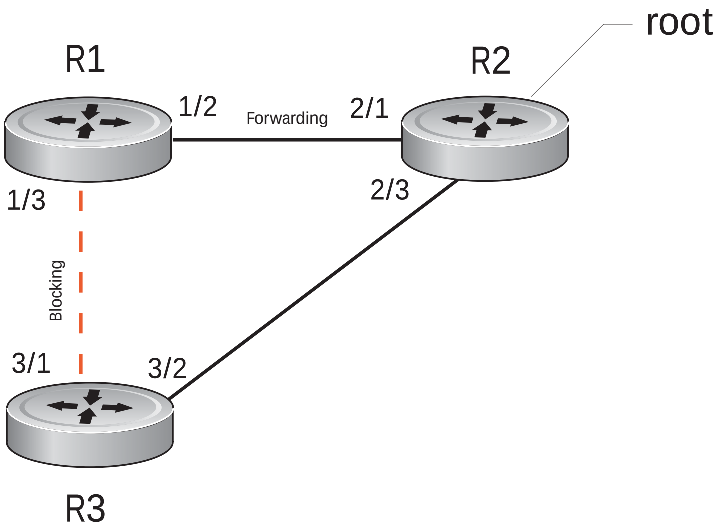
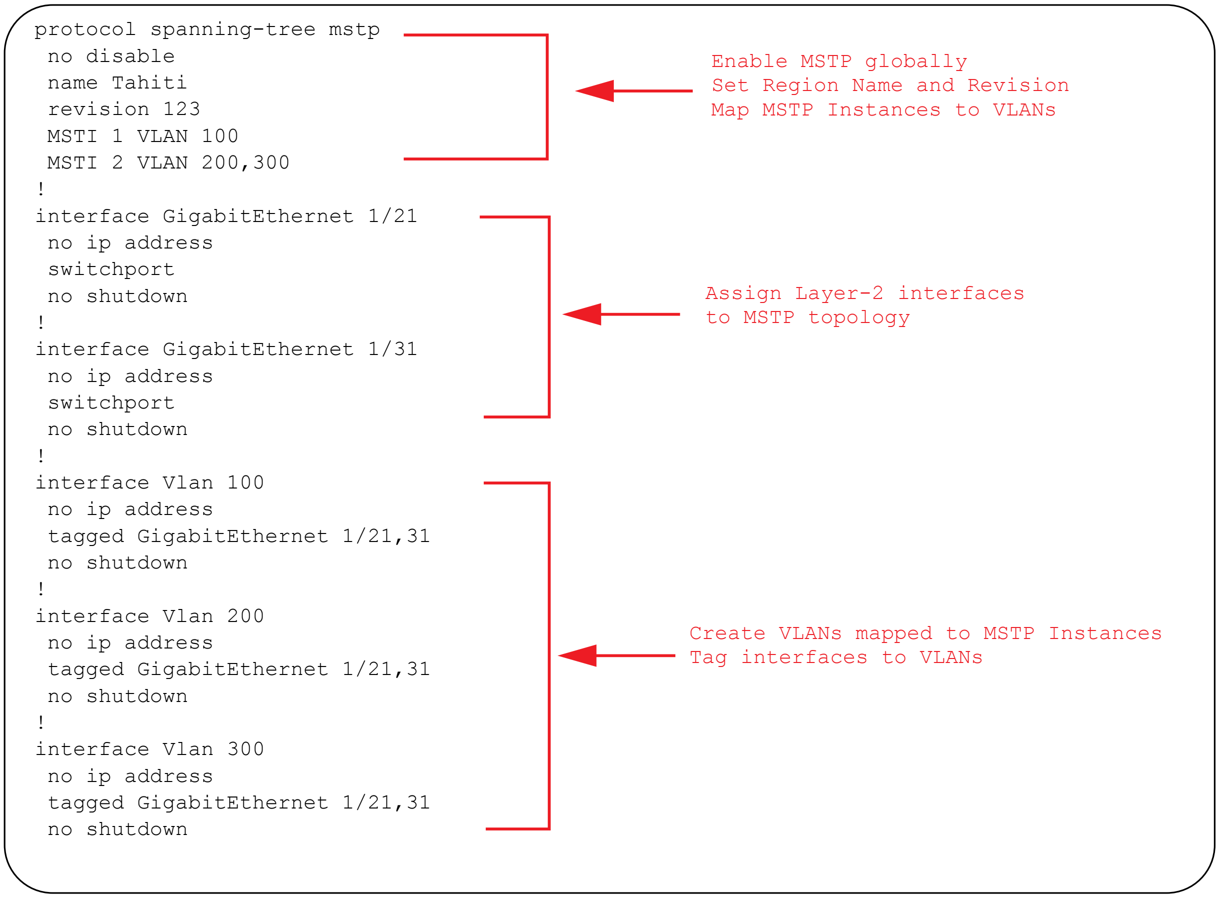


Figure 22-10. Router 1 Running-configuration



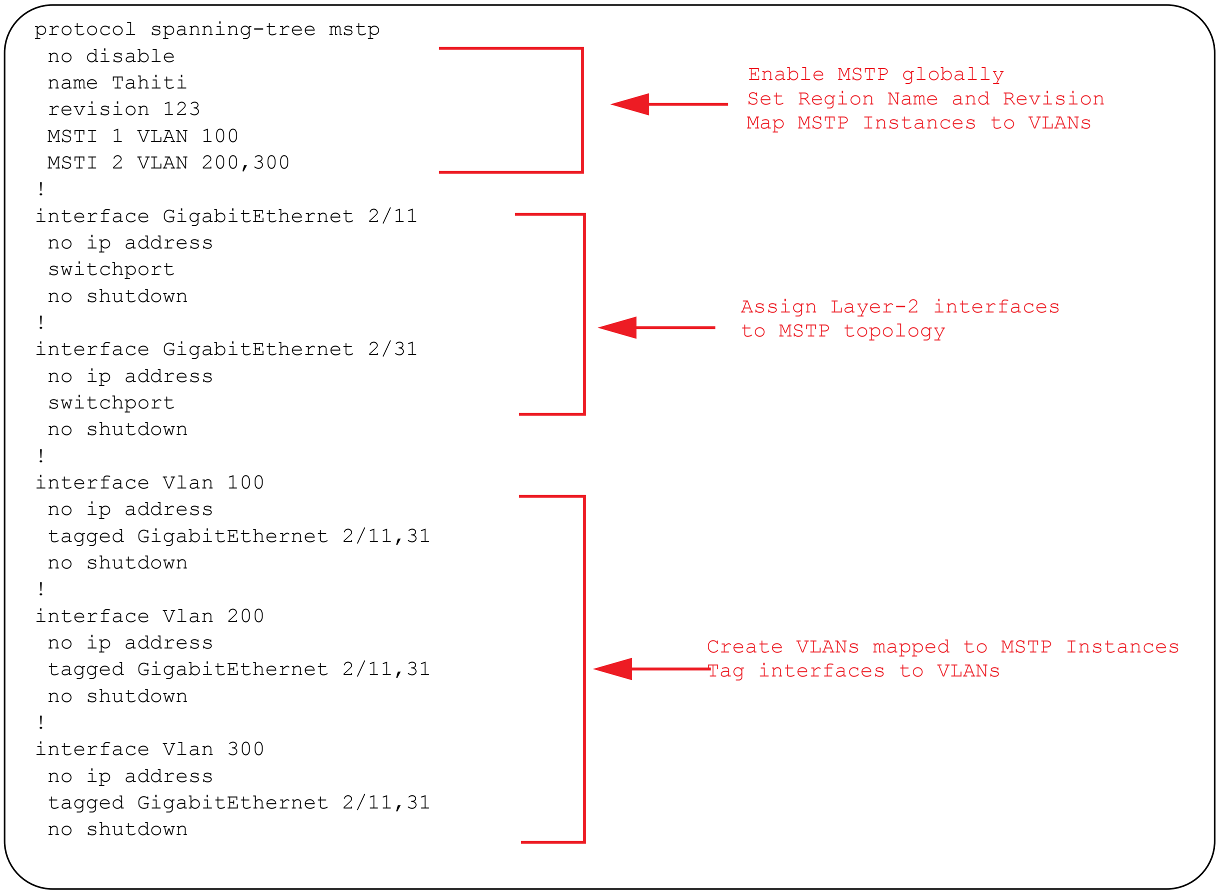
```
protocol spanning-tree mstp
 no disable
 name Tahiti
 revision 123
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
!
interface GigabitEthernet 1/21
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/31
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 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
```

Enable MSTP globally
Set Region Name and Revision
Map MSTP Instances to VLANs

Assign Layer-2 interfaces
to MSTP topology

Create VLANs mapped to MSTP Instances
Tag interfaces to VLANs

Figure 22-11. Router 2 Running-configuration



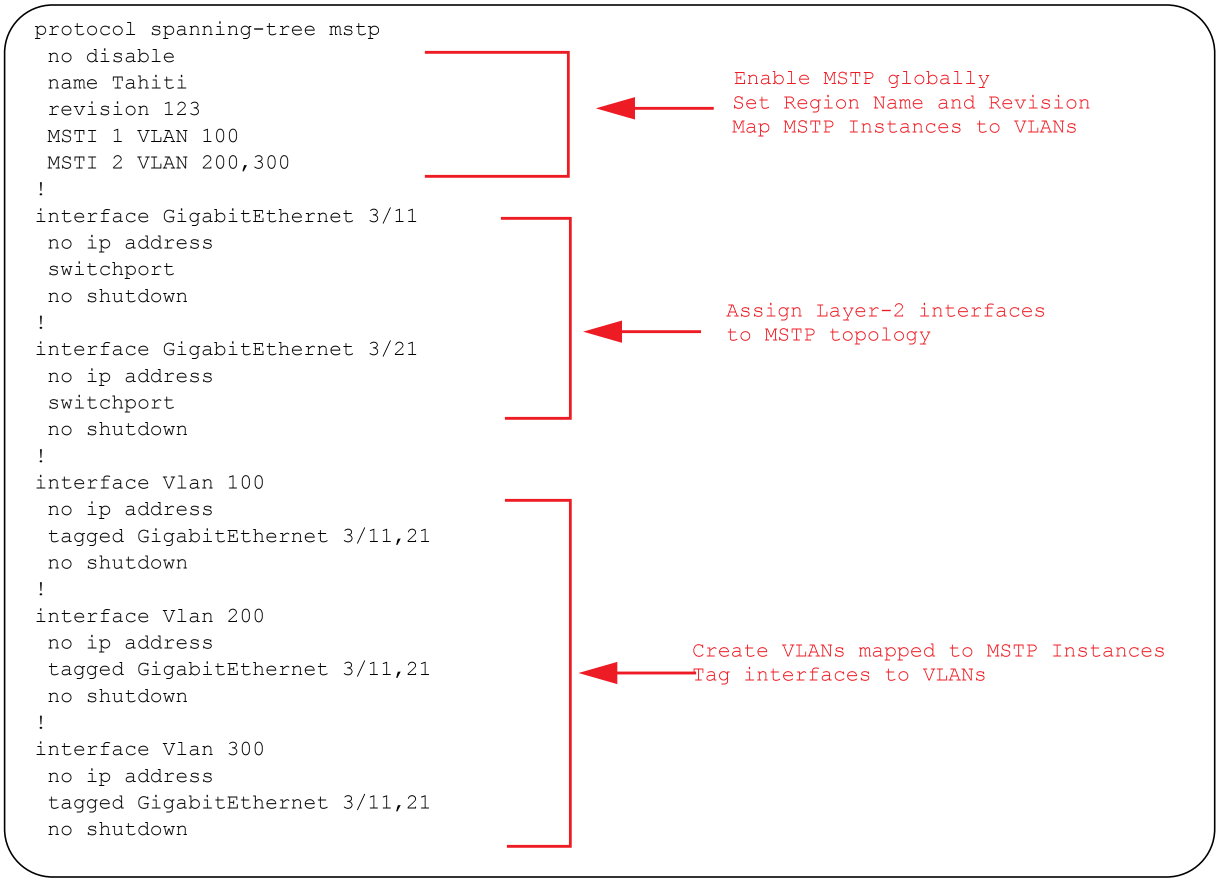
```
protocol spanning-tree mstp
 no disable
 name Tahiti
 revision 123
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
!
interface GigabitEthernet 2/11
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 2/31
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 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
```

Enable MSTP globally
Set Region Name and Revision
Map MSTP Instances to VLANs

Assign Layer-2 interfaces
to MSTP topology

Create VLANs mapped to MSTP Instances
Tag interfaces to VLANs

Figure 22-12. Router 3 Running-configuration



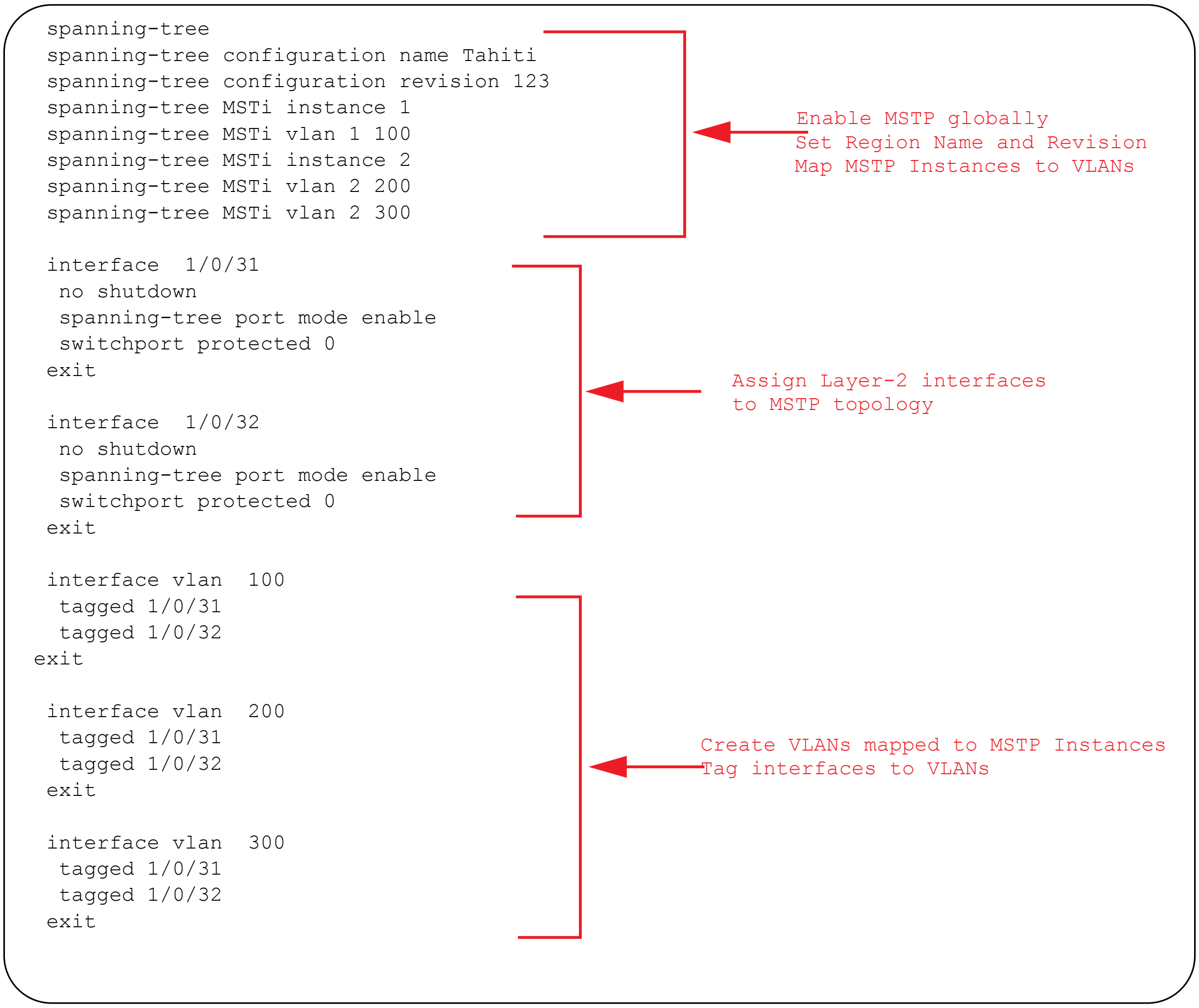
```
protocol spanning-tree mstp
 no disable
 name Tahiti
 revision 123
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
!
interface GigabitEthernet 3/11
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 3/21
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 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
```

Enable MSTP globally
Set Region Name and Revision
Map MSTP Instances to VLANs

Assign Layer-2 interfaces
to MSTP topology

Create VLANs mapped to MSTP Instances
Tag interfaces to VLANs

Figure 22-13. SFTOS Example Running-Configuration



```
 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

 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

 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
```

Enable MSTP globally
Set Region Name and Revision
Map MSTP Instances to VLANs

Assign Layer-2 interfaces
to MSTP topology

Create VLANs mapped to MSTP Instances
Tag interfaces to VLANs

Debugging and Verifying MSTP Configuration

Display BPDUs using the command **debug spanning-tree mstp bpdu** from EXEC Privilege mode. Display MSTP-triggered topology change messages **debug spanning-tree mstp events**.

Figure 22-14. Displaying BPDUs and Events

```
Force10#debug spanning-tree mstp bpdu
1w1d17h : MSTP: Sending BPDU on Gi 1/31 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x68
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 20000
Regional Bridge Id: 32768:0001.e809.c24a, CIST Port Id: 128:384
Msg Age: 2, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver3 Len: 96
Name: my-mstp-region, Rev: 0, Int Root Path Cost: 20000
Rem Hops: 19, Bridge Id: 32768:0001.e80d.b6d6
E1200#1w1d17h : INST 1: Flags: 0x28, Reg Root: 32768:0001.e809.c24a, Int Root Co
         Brg/Port Prio: 32768/128, Rem Hops: 19
INST 2: Flags: 0x68, Reg Root: 4096:0001.e809.c24a, Int Root Cost: 20000
         Brg/Port Prio: 32768/128, Rem Hops: 19
[output omitted]
Force10#debug spanning-tree mstp events
1w1d17h : MSTP: TC flag set in the incoming BPDU on port Gi 1/31 for instance 0

1w1d17h : MSTP: TC flag set in the incoming BPDU on port Gi 1/31 for instance 0

1w1d17h : MSTP: TC flag set in the incoming BPDU on port Gi 1/31 for instance 0
```

Examine your individual routers to ensure all the necessary parameters match.

1. Region Name
2. Region Version
3. VLAN to Instance mapping

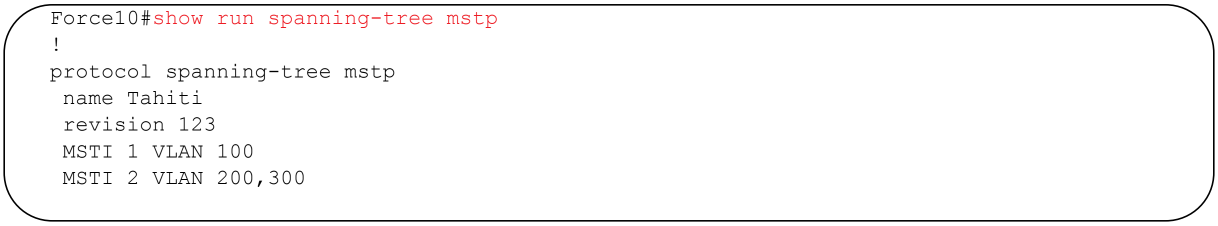
The **show spanning-tree mst** commands will show various portions of the MSTP configuration. To view the overall MSTP configuration on the router, use the **show running-configuration spanning-tree mstp** in the EXEC Privilege mode (output sample shown in Figure 22-15).

Use the **debug spanning-tree mstp bpdu** command to monitor and verify that the MSTP configuration is connected and communicating as desired (output sample shown in Figure 22-16).

Key items to look for in the debug report:

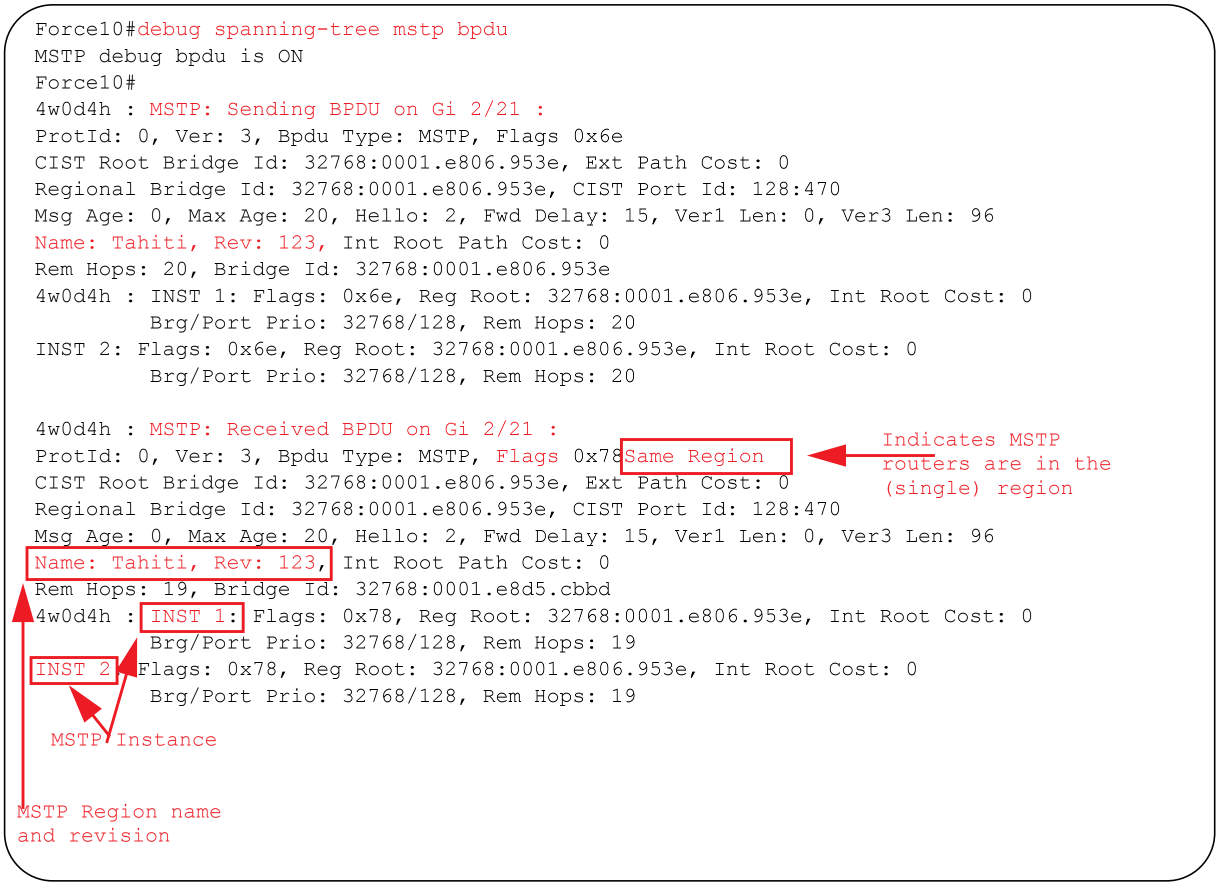
- MSTP flags indicate communication received from the same region.
  - In Figure 22-16, the output shows that the MSTP routers are located in the same region.
  - Does the debug log indicate that packets are coming from a “Different Region” (Figure 22-17)? If so, one of the key parameters is not matching.
- 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.
  - Use the show commands to verify the VLAN to MSTP Instance mapping.
  - Are there “extra” MSTP Instances in the Sending or Received logs? That may mean that an additional MSTP Instance was configured on one router but not the others.

Figure 22-15. Sample Output for show running-configuration spanning-tree mstp command



```
Force10#show run spanning-tree mstp
!
protocol spanning-tree mstp
 name Tahiti
 revision 123
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
```

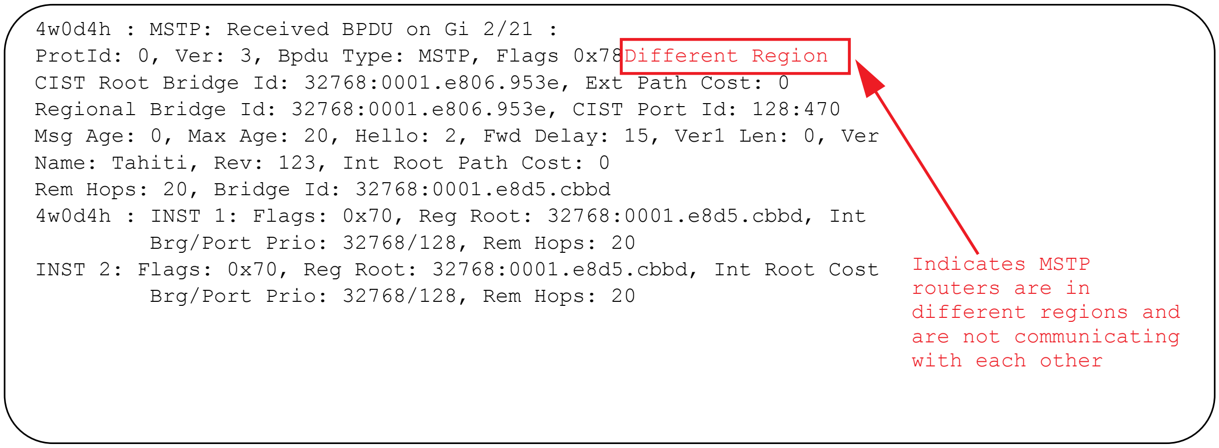
Figure 22-16. Displaying BPDUs and Events - Debug Log of Successful MSTP Configuration



```
Force10#debug spanning-tree mstp bpdu
MSTP debug bpdu is ON
Force10#
4w0d4h : MSTP: Sending BPDU on Gi 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 Gi 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
```

Indicates MSTP routers are in the (single) region
MSTP Instance
MSTP Region name and revision

Figure 22-17. Displaying BPDUs and Events - Debug Log of Unsuccessful MSTP Configuration



```
4w0d4h : MSTP: Received BPDU on Gi 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
```

Indicates MSTP routers are in different regions and are not communicating with each other

23

Multicast Features

Multicast Features are supported on platforms: C E S

This chapter contains the following sections:

- Enable IP Multicast on page 443
- Multicast with ECMP on page 444
- First Packet Forwarding for Lossless Multicast on page 445
- Multicast Policies on page 446
- Multicast Traceroute on page 453
- Multicast Quality of Service on page 453
- Optimize the E-Series for Multicast Traffic on page 454
- Tune the Central Scheduler for Multicast on page 454

FTOS supports the following multicast protocols:

- PIM Sparse-Mode on page 497
- PIM Source-Specific Mode on page 505
- Internet Group Management Protocol on page 271

Implementation Information

- Multicast is not supported on secondary IP addresses.

Enable IP Multicast

Enable IP Multicast is supported on platforms C E S

Prior to enabling any multicast protocols, you must enable multicast routing.

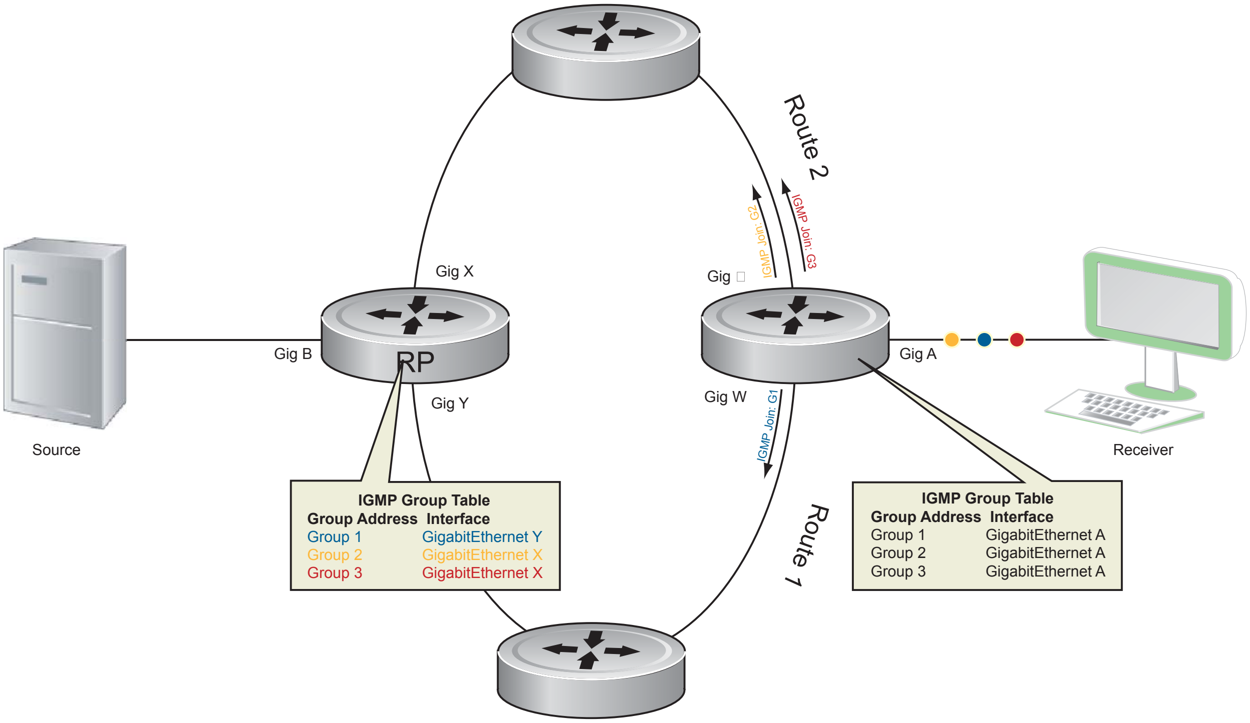
Task	Command Syntax	Command Mode
Enable multicast routing.	**ip multicast-routing**	CONFIGURATION

Multicast with ECMP

Dell Force10 multicast uses Equal-cost Multi-path (ECMP) routing to load-balance multiple streams across equal cost links. When creating the shared-tree Protocol Independent Multicast (PIM) uses routes from all configured routing protocols to select the best route to the rendezvous point (RP). If there are multiple, equal-cost paths, the PIM selects the route with the least number of currently running multicast streams. If multiple routes have the same number of streams, PIM selects the first equal-cost route returned by the Route Table Manager (RTM).

In Figure 23-1, the receiver joins three groups. The last-hop DR initially has two equal-cost routes to the RP with no streams, so it non-deterministically selects Route 1 for the Group 1 IGMP Join message. Route 1 then has one stream associated with it, so the last-hop DR sends the Group 2 Join by Route 2. It then non-deterministically selects Route 2 for the Group 3 Join since both routes already have one multicast stream.

Figure 23-1. Multicast with ECMP



IGMP Group Table

Group Address	Interface
Group 1	GigabitEthernet Y
Group 2	GigabitEthernet X
Group 3	GigabitEthernet X

IGMP Group Table

Group Address	Interface
Group 1	GigabitEthernet A
Group 2	GigabitEthernet A
Group 3	GigabitEthernet A

Implementation Information

- Because protocol control traffic in FTOS is redirected using the MAC address, and multicast control traffic and multicast data traffic might map to the same MAC address, FTOS might forward data traffic with certain MAC addresses to the CPU in addition to control traffic.

As the upper five bits of an IP Multicast address are dropped in the translation, 32 different multicast group IDs all map to the same Ethernet address. For example, 224.0.0.5 is a well known IP address for OSPF that maps to the multicast MAC address 01:00:5e:00:00:05. However, 225.0.0.5, 226.0.0.5, etc., map to the same multicast MAC address. The Layer 2 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 ones below.

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
VRRP	01:00:5e:00:00:12
PIM-SM	01:00:5e:00:00:0d

- The FTOS 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 multicast routing is enabled.

First Packet Forwarding for Lossless Multicast

Beginning with FTOS version 7.8.1.0 for the E-Series TeraScale, version 8.2.1.0 for E-Series ExaScale, and version 8.3.1.0 on all other FTOS platforms, all initial multicast packets are forwarded to receivers to achieve lossless multicast.

In previous versions, when the Dell Force10 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 FTOS versions above, when the Dell Force10 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 Force10 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

FTOS offers parallel Multicast features for IPv4 and IPv6.

- IPv4 Multicast Policies on page 446
- IPv6 Multicast Policies on page 451

IPv4 Multicast Policies

- Limit the Number of Multicast Routes on page 446
- Prevent a Host from Joining a Group on page 447
- Rate Limit IGMP Join Requests on page 449
- Prevent a PIM Router from Forming an Adjacency on page 449
- Prevent a Source from Registering with the RP on page 449
- Prevent a PIM Router from Processing a Join on page 451

Limit the Number of Multicast Routes

Task	Command Syntax	Command Mode
Limit the total number of multicast routes on the system.	**ip multicast-limit** Range: 1-50000 Default: 15000	CONFIGURATION

When the limit is reached, FTOS does not process any IGMP or 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, you must increase the limit by at least 10% for IGMP and MLD to resume.
- If the limit is decreased after it is reached, FTOS does not clear the existing sessions. Entries are cleared upon a timeout (you may also clear entries using **clear ip mroute**).

Note: FTOS 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, FTOS displays Message 1.

Message 1 Multicast Route Limit Error

```
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.
```

Note: The IN-L3-McastFib CAM partition is used to store multicast routes and is a separate hardware limit that is exists per port-pipe. Any software-configured limit might be superseded 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 set by the **ip multicast-limit** is reached.

Prevent a Host from Joining a Group

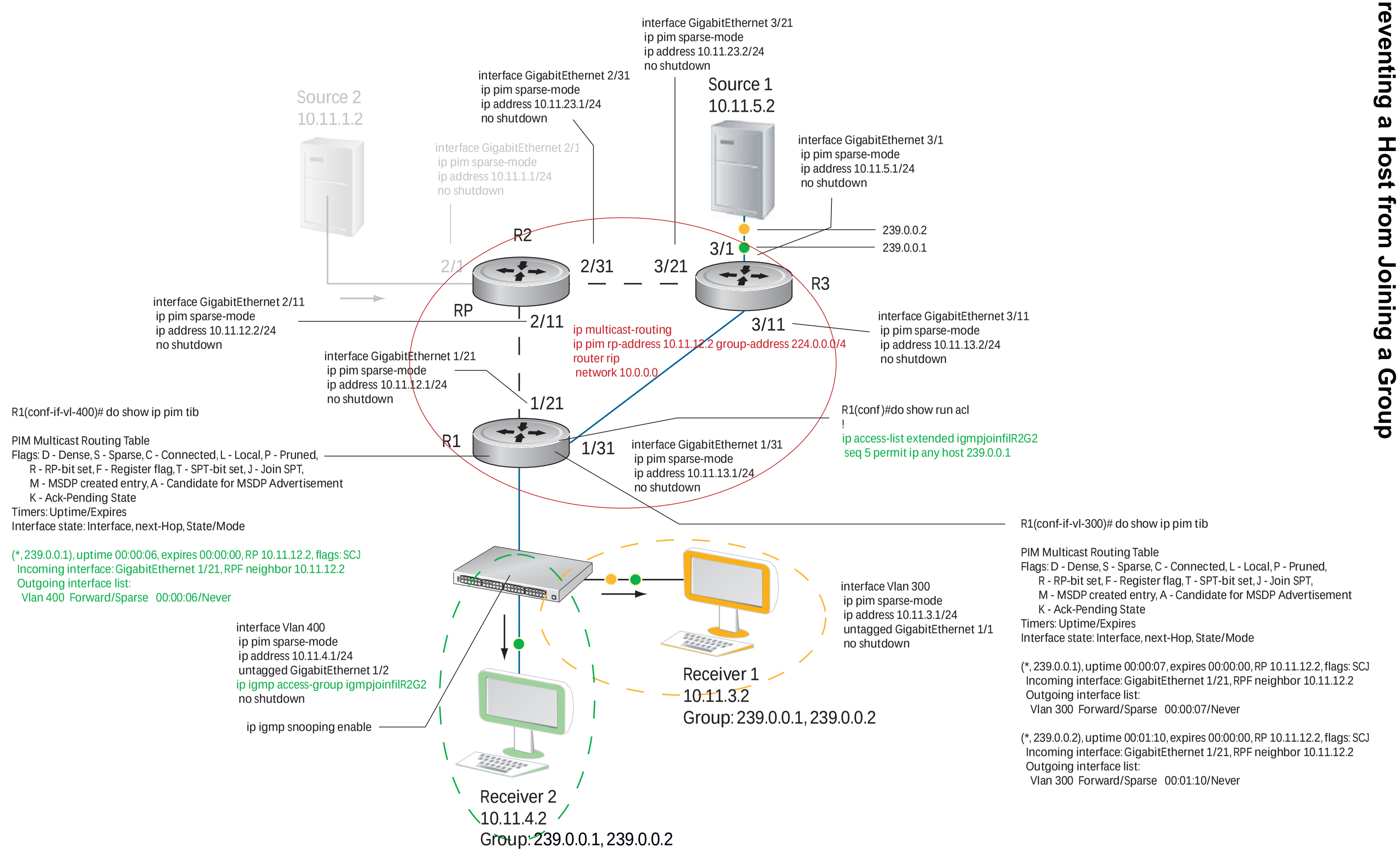
You can prevent a host from joining a particular group by blocking specific IGMP reports. Create an extended access list containing the permissible source-group pairs. Use the command **ip igmp access-group** *access-list-name* from INTERFACE mode to apply the access list.

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 Figure 23-2), since IGMPv2 hosts do not know in advance who the source is for the group in which they are interested.

FTOS Behavior: Do not enter the command **ip igmp access-group** before creating the access-list. If you do, upon entering your first *deny* rule, FTOS 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 command **ip igmp access-group** before creating the access-list, prevent FTOS from clearing the routing table by entering a *permit any* rule with high sequence number before you enter any other rules.

In Figure 23-2, 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 23-2. Preventing a Host from Joining a Group



```
interface GigabitEthernet 3/21
 ip pim sparse-mode
 ip address 10.11.23.2/24
 no shutdown
```

```
interface GigabitEthernet 2/31
 ip pim sparse-mode
 ip address 10.11.23.1/24
 no shutdown
```

```
interface GigabitEthernet 2/1
 ip pim sparse-mode
 ip address 10.11.1.1/24
 no shutdown
```

```
interface GigabitEthernet 3/1
 ip pim sparse-mode
 ip address 10.11.5.1/24
 no shutdown
```

```
interface GigabitEthernet 2/11
 ip pim sparse-mode
 ip address 10.11.12.2/24
 no shutdown
```

```
interface GigabitEthernet 3/11
 ip pim sparse-mode
 ip address 10.11.13.2/24
 no shutdown
```

```
ip multicast-routing
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
router rip
 network 10.0.0.0
```

```
interface GigabitEthernet 1/21
 ip pim sparse-mode
 ip address 10.11.12.1/24
 no shutdown
```

```
interface GigabitEthernet 1/31
 ip pim sparse-mode
 ip address 10.11.13.1/24
 no shutdown
```

```
R1(conf)#do show run acl
!
ip access-list extended igmpjoinfilR2G2
 seq 5 permit ip any host 239.0.0.1
```

```
R1(conf-if-vl-400)# do 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,
    M - MSDP created entry, A - Candidate for MSDP Advertisement
    K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.0.0.1), uptime 00:00:06, expires 00:00:00, RP 10.11.12.2, flags: SCJ
 Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
 Outgoing interface list:
  Vlan 400  Forward/Sparse   00:00:06/Never
```

```
interface Vlan 400
 ip pim sparse-mode
 ip address 10.11.4.1/24
 untagged GigabitEthernet 1/2
ip igmp access-group igmpjoinfilR2G2
 no shutdown
```

```
ip igmp snooping enable
```

```
interface Vlan 300
 ip pim sparse-mode
 ip address 10.11.3.1/24
 untagged GigabitEthernet 1/1
 no shutdown
```

```
R1(conf-if-vl-300)# do 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,
    M - MSDP created entry, A - Candidate for MSDP Advertisement
    K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.0.0.1), uptime 00:00:07, expires 00:00:00, RP 10.11.12.2, flags: SCJ
 Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
 Outgoing interface list:
  Vlan 300  Forward/Sparse   00:00:07/Never

(*, 239.0.0.2), uptime 00:01:10, expires 00:00:00, RP 10.11.12.2, flags: SCJ
 Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
 Outgoing interface list:
  Vlan 300  Forward/Sparse   00:01:10/Never
```

Rate Limit 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 using the command **ip igmp group-join-limit** from INTERFACE mode. 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.

View the enable status of this feature using the command **show ip igmp interface** from EXEC Privilege mode.

Prevent a PIM Router from Forming an Adjacency

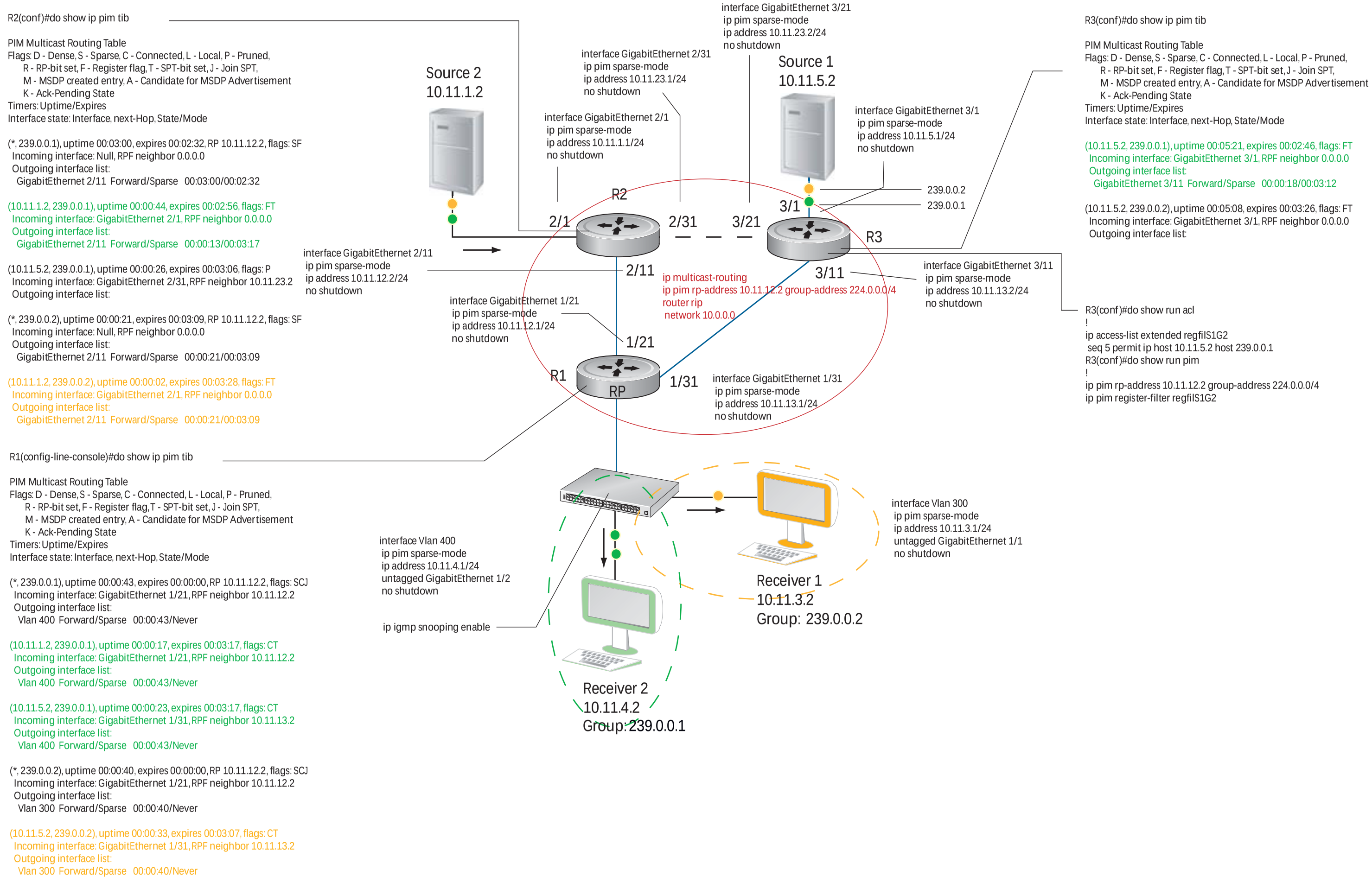
To prevent a router from participating in Protocol Independent Multicast (PIM) (for example, to configure stub multicast routing), use the **ip pim neighbor-filter** command from INTERFACE mode.

Prevent a Source from Registering with the RP

Use the command **ip pim register-filter** from CONFIGURATION mode to prevent a source from transmitting to a particular group. This command prevents the PIM source DR from sending register packets to RP for the specified multicast source and group; if the source DR never sends register packets to the RP, no hosts can ever discover the source and create an SPT to it.

In Figure 23-3, 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 23-3. Preventing a Source from Transmitting to a Group



```
R2(conf)#do 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,
      M - MSDP created entry, A - Candidate for MSDP Advertisement
      K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.0.0.1), uptime 00:03:00, expires 00:02:32, RP 10.11.12.2, flags: SF
 Incoming interface: Null, RPF neighbor 0.0.0.0
 Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse   00:03:00/00:02:32

(10.11.1.2, 239.0.0.1), uptime 00:00:44, expires 00:02:56, flags: FT
 Incoming interface: GigabitEthernet 2/1, RPF neighbor 0.0.0.0
 Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse   00:00:13/00:03:17

(10.11.5.2, 239.0.0.1), uptime 00:00:26, expires 00:03:06, flags: P
 Incoming interface: GigabitEthernet 2/31, RPF neighbor 10.11.23.2
 Outgoing interface list:

(*, 239.0.0.2), uptime 00:00:21, expires 00:03:09, RP 10.11.12.2, flags: SF
 Incoming interface: Null, RPF neighbor 0.0.0.0
 Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse   00:00:21/00:03:09

(10.11.1.2, 239.0.0.2), uptime 00:00:02, expires 00:03:28, flags: FT
 Incoming interface: GigabitEthernet 2/1, RPF neighbor 0.0.0.0
 Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse   00:00:21/00:03:09
```

```
R1(config-line-console)#do 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,
      M - MSDP created entry, A - Candidate for MSDP Advertisement
      K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.0.0.1), uptime 00:00:43, expires 00:00:00, RP 10.11.12.2, flags: SCJ
 Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
 Outgoing interface list:
  Vlan 400  Forward/Sparse   00:00:43/Never

(10.11.1.2, 239.0.0.1), uptime 00:00:17, expires 00:03:17, flags: CT
 Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
 Outgoing interface list:
  Vlan 400  Forward/Sparse   00:00:43/Never

(10.11.5.2, 239.0.0.1), uptime 00:00:23, expires 00:03:17, flags: CT
 Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
 Outgoing interface list:
  Vlan 400  Forward/Sparse   00:00:43/Never

(*, 239.0.0.2), uptime 00:00:40, expires 00:00:00, RP 10.11.12.2, flags: SCJ
 Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
 Outgoing interface list:
  Vlan 300  Forward/Sparse   00:00:40/Never

(10.11.5.2, 239.0.0.2), uptime 00:00:33, expires 00:03:07, flags: CT
 Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
 Outgoing interface list:
  Vlan 300  Forward/Sparse   00:00:40/Never
```

```
interface GigabitEthernet 2/11
 ip pim sparse-mode
 ip address 10.11.12.2/24
 no shutdown
```

```
interface GigabitEthernet 2/1
 ip pim sparse-mode
 ip address 10.11.1.1/24
 no shutdown
```

```
interface GigabitEthernet 2/31
 ip pim sparse-mode
 ip address 10.11.23.1/24
 no shutdown
```

```
interface GigabitEthernet 3/21
 ip pim sparse-mode
 ip address 10.11.23.2/24
 no shutdown
```

```
interface GigabitEthernet 3/1
 ip pim sparse-mode
 ip address 10.11.5.1/24
 no shutdown
```

```
interface GigabitEthernet 1/21
 ip pim sparse-mode
 ip address 10.11.12.1/24
 no shutdown
```

```
ip multicast-routing
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
router rip
 network 10.0.0.0
```

```
interface GigabitEthernet 3/11
 ip pim sparse-mode
 ip address 10.11.13.2/24
 no shutdown
```

```
interface GigabitEthernet 1/31
 ip pim sparse-mode
 ip address 10.11.13.1/24
 no shutdown
```

```
interface Vlan 400
 ip pim sparse-mode
 ip address 10.11.4.1/24
 untagged GigabitEthernet 1/2
 no shutdown
```

```
ip igmp snooping enable
```

```
interface Vlan 300
 ip pim sparse-mode
 ip address 10.11.3.1/24
 untagged GigabitEthernet 1/1
 no shutdown
```

```
R3(conf)#do 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,
      M - MSDP created entry, A - Candidate for MSDP Advertisement
      K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(10.11.5.2, 239.0.0.1), uptime 00:05:21, expires 00:02:46, flags: FT
 Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
 Outgoing interface list:
  GigabitEthernet 3/11  Forward/Sparse   00:00:18/00:03:12

(10.11.5.2, 239.0.0.2), uptime 00:05:08, expires 00:03:26, flags: FT
 Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
 Outgoing interface list:
```

```
R3(conf)#do show run acl
!
ip access-list extended regfilS1G2
 seq 5 permit ip host 10.11.5.2 host 239.0.0.1
R3(conf)#do show run pim
!
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim register-filter regfilS1G2
```

Prevent a PIM Router from Processing a Join

Permit or deny PIM Join/Prune messages on an interface using an extended IP access list. Use the command **ip pim join-filter** to prevent the PIM SM router from creating state based on multicast source and/or group.

IPv6 Multicast Policies

IPv6 Multicast Policies is available only on platform: E

- Limit the Number of IPv6 Multicast Routes on page 451
- Prevent an IPv6 Neighbor from Forming an Adjacency on page 452
- Prevent an IPv6 Source from Registering with the RP on page 452
- Prevent an IPv6 PIM Router from Processing an IPv6 Join on page 452

Limit the Number of IPv6 Multicast Routes

You can limit the total number of IPv6 multicast routes on the system. The maximum number of multicast entries allowed on each line card is determined by the CAM profile. Multicast routes are stored in the IN-V6-McastFib CAM region, which has a fixed number of entries. Any limit configured via the CLI is superseded by this hardware limit. The opposite is also true; the CAM might not be exhausted at the time the CLI-configured route limit is reached.

Task	Command Syntax	Command Mode
Limit the total number of IPv6 multicast routes on the system.	**ipv6 multicast-limit** Range: 1-50000 Default: 15000	CONFIGURATION

Prevent an IPv6 Neighbor from Forming an Adjacency

Task	Command Syntax	Command Mode
Prevent a router from participating in PIM.	**ipv6 pim neighbor-filter** *access-list*	CONFIGURATION

```
Force10(conf)#ipv6 pim neighbor-filter NEIGH_ACL
Force10(conf)#ipv6 access-list NEIGH_ACL
Force10(conf-ipv6-acl)#show config
!
ipv6 access-list NEIGH_ACL
 seq 5 deny ipv6 host fe80::201:e8ff:fe0a:5ad any
 seq 10 permit ipv6 any any
Force10(conf-ipv6-acl)#
```

Prevent an IPv6 Source from Registering with the RP

Task	Command Syntax	Command Mode
Configured on the source DR, prevent the source DR from sending register packets to the RP for specific sources and groups.	**ipv6 pim register-filter** *access-list*	CONFIGURATION

```
Force10(conf)#ipv6 pim register-filter REG-FIL_ACL
Force10(conf)#ipv6 access-list REG-FIL_ACL
Force10(conf-ipv6-acl)#deny ipv6 165:87:34::10/128 ff0e::225:1:2:0/112
Force10(conf-ipv6-acl)#permit ipv6 any any
Force10(conf-ipv6-acl)#exit
```

Prevent an IPv6 PIM Router from Processing an IPv6 Join

Task	Command Syntax	Command Mode
Permit or deny PIM Join/Prune messages on an interface using an access list. This command prevents the PIM-SM router from creating state based on multicast source and/or group.	**ipv6 pim join-filter***access-list* [**in** | **out**]	INTERFACE

```
Force10(conf)#ipv6 access-list JOIN-FIL_ACL
Force10(conf-ipv6-acl)#permit ipv6 165:87:34::0/112 ff0e::225:1:2:0/112
Force10(conf-ipv6-acl)#permit ipv6 any ff0e::230:1:2:0/112
Force10(conf-ipv6-acl)#permit ipv6 165:87:32::0/112 any
Force10(conf-ipv6-acl)#exit
Force10(conf)#interface gigabitethernet 0/84
Force10(conf-if-gi-0/84)#ipv6 pim join-filter JOIN-FIL_ACL in
Force10(conf-if-gi-0/84)#ipv6 pim join-filter JOIN-FIL_ACL out
```

Multicast Traceroute

Multicast Traceroute is supported only on platform: E

MTRACE is an IGMP-based tool that prints that network path that a multicast packet takes from a source to a destination, for a particular group. FTOS has mtrace client and mtrace transmit functionality.

- **MTRACE Client**—an mtrace client transmits mtrace queries and prints out the details received responses.
- **MTRACE Transit**—when a Dell Force10 system is an intermediate router between the source and destination in an MTRACE query, FTOS computes the RPF neighbor for the source, fills in the request, and forwards the request to the RPF neighbor. While computing the RPF neighbor, static mroutes and mBGP routes are preferred over unicast routes. When a Dell Force10 system is the last hop to the destination, FTOS sends a response to the query.

Task	Command Syntax	Command Mode
Print the network path that a multicast packet takes from a multicast source to receiver, for a particular group.	**mtrace** *multicast-source-address multicast-receiver-address multicast-group-address*	EXEC Privilege

Figure 23-4. Tracing a Multicast Route

```
Force10#mtrace 10.11.5.2 10.11.3.2 239.0.0.1
Type Ctrl-C to abort.
Mtrace from 10.11.5.2 to 10.11.3.2 via group 239.0.0.1
From source (?) to destination (?)
Querying full reverse path...
 0  10.11.3.2
-1  10.11.3.1  PIM  Reached RP/Core [default]
-2  10.11.5.2
```

Multicast Quality of Service

Multicast Quality of Service is supported only on platform: E

The Quality of Service (QoS) features available for unicast traffic can be applied to multicast flows. The following QoS features are available:

- Policy-based QoS—Classifying, rate policing, and marking ingress traffic
- WRED
  - See also Allocate More Buffer Memory for Multicast WRED on page 454.

Optimize the E-Series for Multicast Traffic

Optimize the E-Series for Multicast Traffic is supported only on platform: E

The default hardware settings for the E-series are for unicast applications like data centers and ISP networks. This means that the E-Series gives priority to unicast data forwarding rather than multicast data forwarding. For multicast intensive applications like trading, Dell Force10 recommends reconfiguring some default settings.

You may do one or more for the following to optimize the E-Series for your multicast application:

- Tune the Central Scheduler for Multicast on page 454
- Allocate More Buffer Memory for Multicast WRED
- Allocate More Bandwidth to Multicast using Egress WFQ

Allocate More Buffer Memory for Multicast WRED

Allocate more buffer memory to multicast WRED (Weighted Random Early Detection) for bursty multicast traffic that might temporarily become oversubscribed. For example, the example WRED profile in Figure 31-14 on page 580 allocates multicast traffic a minimum of 40 megabytes (out of 80 megabytes) of buffer memory and up to 60 megabytes.

Figure 23-5. Allocating More Bandwidth for Multicast WRED

```
Force10(Conf)#queue egress multicast linecard all wred-profile Egress
Force10(conf)#wred-profile Egress
Force10(conf-wred)# threshold min 40960 max 61440
```

Allocate More Bandwidth to Multicast using Egress WFQ

Egress Weighted Fair Queuing (WFQ) determines per port the ratio of egress bandwidth allocated to multicast, replication, and unicast traffic. By default, FTOS provides 1/64 to multicast, 1/64 to replication, and 62/64 for unicast, which is shared between 8 unicast queues. Allocate more bandwidth for multicast using the command **queue egress multicast linecard** (from CONFIGURATION mode) with the keyword **bandwidth-percent**. For example, allocate 80% of egress bandwidth to multicast on all line cards using the command **queue egress multicast linecard all bandwidth-percent 80.**

Tune the Central Scheduler for Multicast

The Central Scheduler is responsible for scheduling unicast and multicast packets via the Terabit backplane. The default configuration of the Central Scheduler is optimized for network environments that forward primarily unicast traffic—80% of the scheduler weight is for unicast traffic and 20% is for multicast traffic.

FTOS provides the ability to adjust the scheduling weight for multicast traffic. For example, if the majority of your traffic is multicast, the default configuration might yield greater latency. In this case, allocate more backplane bandwidth for multicast using the command **queue multicast bandwidth-percent** from CONFIGURATION mode. View your configuration using the command **show queue backplane multicast bandwidth-percentage**.

Figure 23-6. Tuning the Central Scheduler for Multicast

```
Force10#show queue backplane multicast bandwidth-percent
Configured multicast bandwidth percentage is 80
```

24

Open Shortest Path First (OSPFv2 and OSPFv3)

Open Shortest Path First version 2 (OSPF for IPv4) is supported on platforms C E S

Open Shortest Path First version 3 (OSPF for IPv6) is supported on platforms C E

This chapter is intended to provide a general description of OSPFv2 (OSPF for IPv4) and OSPFv3 (OSPF for IPv6) as supported in the Dell Force10 Operating System (FTOS). It is not intended to provide a complete

Note: The fundamental mechanisms of OSPF (flooding, DR election, area support, SPF calculations, etc.) are the same between OSPFv2 and OSPFv3. Where there are differences between the two versions, they are identified and clarified. Except where identified, the information in this chapter applies to both protocol versions.

This chapter includes the following topics:

- Protocol Overview
- Implementing OSPF with FTOS
  - Fast Convergence (OSPFv2, IPv4 only)
  - Multi-Process OSPF (OSPFv2, IPv4 only)
  - RFC-2328 Compliant OSPF Flooding
  - OSPF ACK Packing
  - OSPF Adjacency with Cisco Routers
- Configuration Information
  - Configuration Task List for OSPFv2 (OSPF for IPv4)
  - Configuration Task List for OSPFv3 (OSPF for IPv6)
- Sample Configurations for OSPFv2

OSPF protocol standards are listed in the Appendix 47, Standards Compliance chapter.

Protocol Overview

Open Shortest Path First (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 SPF algorithm (Shortest Path First 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 Autonomous System 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 Link State Advertisements (LSAs).

OSPFv3 runs on a per-link basis instead of on a per-IP-subnet basis. All neighbors on all link types are identified by Router ID (RID). 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. OSPFv3 removes this inconsistency, and all neighbors on all link types are identified by RID.

Note: OSPFv3 is not backward-compatible with OSPFv2; they can co-exist. To use OSPF with both IPv4 and IPv6, you must run both OSPFv2 and OSPFv3.

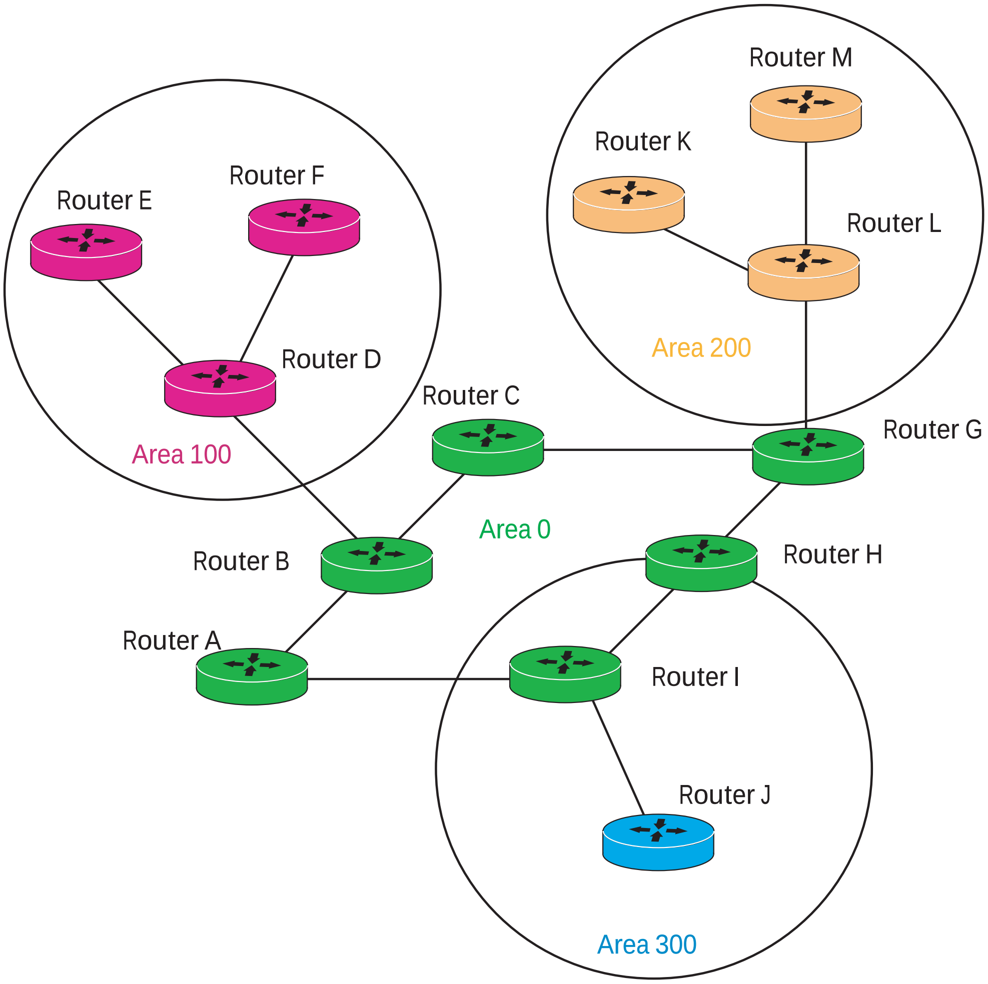
Autonomous System (AS) Areas

OSPF operate 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.

An AS can be divided 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, 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 24-1. 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 Autonomous System (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 an default area number. All other areas can have their Area ID assigned in the configuration.

Figure 24-1 shows 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 that all routers within an assigned Stub area must be configured as stubby, and no 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. Stubby areas cannot be traversed by a virtual link.

A **Not-So-Stubby** Area (NSSA) can import AS external route information and send it to the Backbone. It cannot received external AS information from the Backbone or other areas. It can be traversed by a virtual link.

Totally Stubby Areas are referred to as No Summary areas in FTOS.

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.

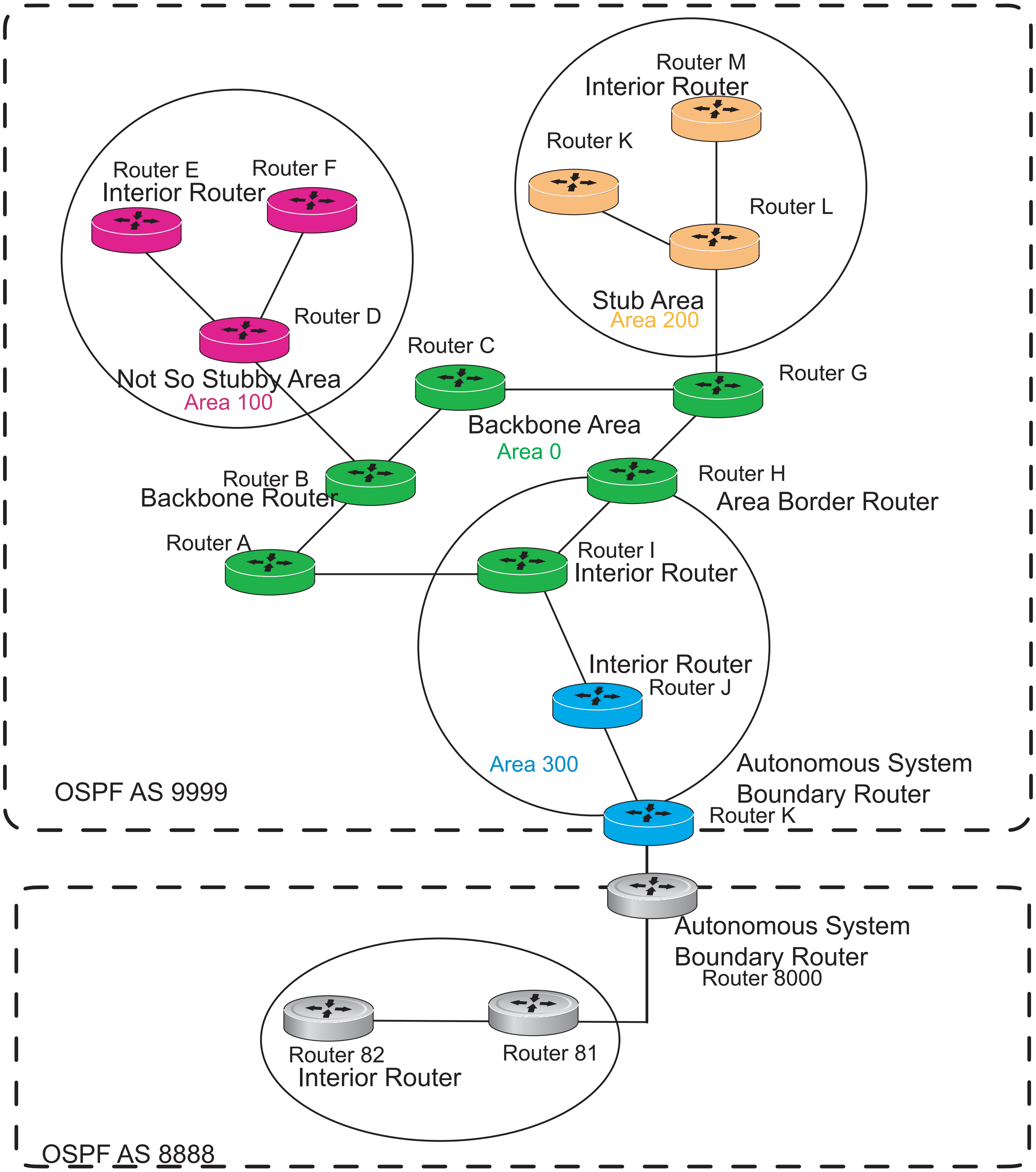
Router Types

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 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). The router ID does not have to be associated with a valid IP address. However, Dell Force10 recommends that the router ID and the router's IP address reflect each other, to make troubleshooting easier.

Figure 24-2gives some examples of the different router designations.

Figure 24-2. OSPF Routing Examples



Backbone Router (BR)

A Backbone Router (BR) is part of the OSPF Backbone, Area 0. This includes all Area Border Routers (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 Figure 24-2 above.

Area Border Router (ABR)

Within an AS, an Area Border (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 Area Border Router (ABR) takes information it has learned on one of its attached areas and can summarize it before sending it out on other areas it is connected to.

An ABR can connect to many areas in an AS, and is considered a member of each area it connects to.

Autonomous System Border Router (ASBR)

The Autonomous System Border Area Router (ASBR) connects to more than one AS, and exchanges information with the routers in other ASs. Generally the ASBR connects to a non-Interior Gate Protocol (IGP) such as BGP or uses static routes.

Internal Router (IR)

The Internal Router (IR) has adjacencies with ONLY routers in the same area, as Router E, M and I are shown in Figure 24-2.

Designated and Backup Designated Routers

OSPF elects a Designated Router and a Backup Designated router. Among other things, the designated router 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 Designated Router (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, it sends it to the Designated Router (DR) and Backup Designated Router (BDR). The DR sends the update out to all other routers in the area.
- The Backup Designated Router (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 discussed earlier. The Designated and Backup Designated Routers are configurable in FTOS. If no DR or BDR is defined in FTOS, 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, then 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 is 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.

- OSPFv3 can treat LSAs as having link-local flooding scope, or store and flood them as if they are understood, while ignoring them in their own SPF algorithms.
- OSPFv2 always discards unknown LSA types.

The LSA types supported by Dell Force10 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 Designated Router (DR) in an area lists which routers are joined together 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 Area Border Router (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 Area Border Router will (ABR) flood the information for the router (i.e. the Autonomous System Border Router (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
    - Routers in a Not-So-Stubby-Area (NSSA) do not receive external LSAs from Area Border Routers (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 Area Border Router 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.

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 Designated Router
- 3: connection to a stub network IP network/subnet number
- 4: virtual link neighboring router ID

Virtual Links

In the case in which an area cannot be directly connected to Area 0, you must configure a virtual link between that area and Area 0. The two endpoints of a virtual link are ABRs, and the virtual link must be configured in both routers. The common non-backbone area to which the two routers belong is called a transit area. A virtual link specifies the transit area and the router ID of the other virtual endpoint (the other ABR).

A Virtual Link cannot be configured through a Stub Area or NSSA.

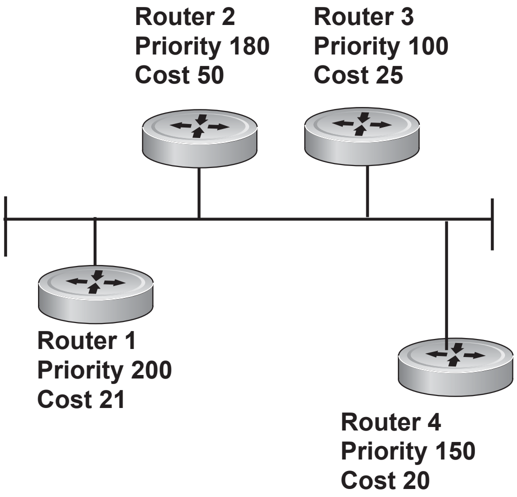
Router Priority and Cost

Router priority and cost is the method the system uses to “rate” the routers. For example, if not assigned, the system will select the router with the highest priority as the DR. The second highest priority is the BDR.

Priority is a numbered rating 0-255. The higher the number, the higher the priority.

Cost is a numbered rating 1-65535. 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 24-3. Priority and Costs Example



Router 1 selected by the system as DR.
Router 2 selected by the system as BDR.

If R1 fails, the system □subtracts□ 21 from R1□s priority number. R1□s new priority is 179.

R2 as both the selected BDR and the now-highest priority, becomes the DR.

If R3 fails, the system □subtracts□ 50 from its priority. R2□s new priority is 130.

R4 is now the highest priority and becomes the DR.

Implementing OSPF with FTOS

FTOS supports up to 10,000 OSPF routes. Within that 10,000 up to 8,000 routes can be designated as external and up to 2,000 designated as inter/intra area routes.

FTOS version 7.8.1.0 and later support multiple OSPF processes (OSPF MP).

Prior to 7.8.1.0, FTOS supports 1 OSPFv2 and 1 OSPFv3 process ID per system. Recall that OSPFv2 and OSPFv3 can coexist but must be configured individually.

FTOS supports Stub areas, Totally Stub (No Summary) and Not So Stubby Areas (NSSAs) and supports the following LSAs, as discussed earlier in this document.

- Router (type 1)
- Network (type 2)
- Network Summary (type 3)
- AS Boundary (type 4)
- AS External (type 5)

- NSSA External (type 7)
- Opaque Link-local (type 9)

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. FTOS enables you to accept and originate LSAa as soon as they are available to speed up route information propagation.

Note that the faster the convergence, the more frequent the route calculations and updates. This will impact CPU utilization and may impact adjacency stability in larger topologies.

Multi-Process OSPF (OSPFv2, IPv4 only)

Multi-Process OSPF is supported on platforms C E and S with FTOS version 7.8.1.0 and later, and is supported on OSPFv2 with IPv4 only.

Multi-Process OSPF allows multiple OSPFv2 processes on a single router. Multiple OSPFv2 processes allow for isolating routing domains, supporting multiple route policies and priorities in different domains, and creating smaller domains for easier management.

- The E-Series supports up to 28 OSPFv2 processes.
- The C-Series supports up to 6 OSPFv2 processes.
- The S50 and S25 support up to 4 OSPFv2 processes.
- The S55, S60, and S4810 support up to 16 OSPFv2 processes.
- The Z9000 supports up to 3 OSPFv2 processes.

Each OSPFv2 process has a unique process ID and must have an associated Router ID. There must be an equal number of interfaces must be in Layer-3 mode for the number of processes created. For example, if 5 OSPFv2 processes are created on a system, there must be at least 5 interfaces assigned in Layer-3 mode.

Each OSPFv2 process is independent. If one process loses adjacency, the other processes continue to function/

Processing SNMP and Sending SNMP Traps

Though there are may be several OSPFv2 processes, only one process can process SNMP requests and send SNMP traps. The **mib-binding** command identifies one of the OSPVFv2 processes as the process responsible for SNMP management. If the **mib-binding** command is not specified, the first OSPFv2 process created manages the SNMP processes and traps.

RFC-2328 Compliant OSPF Flooding

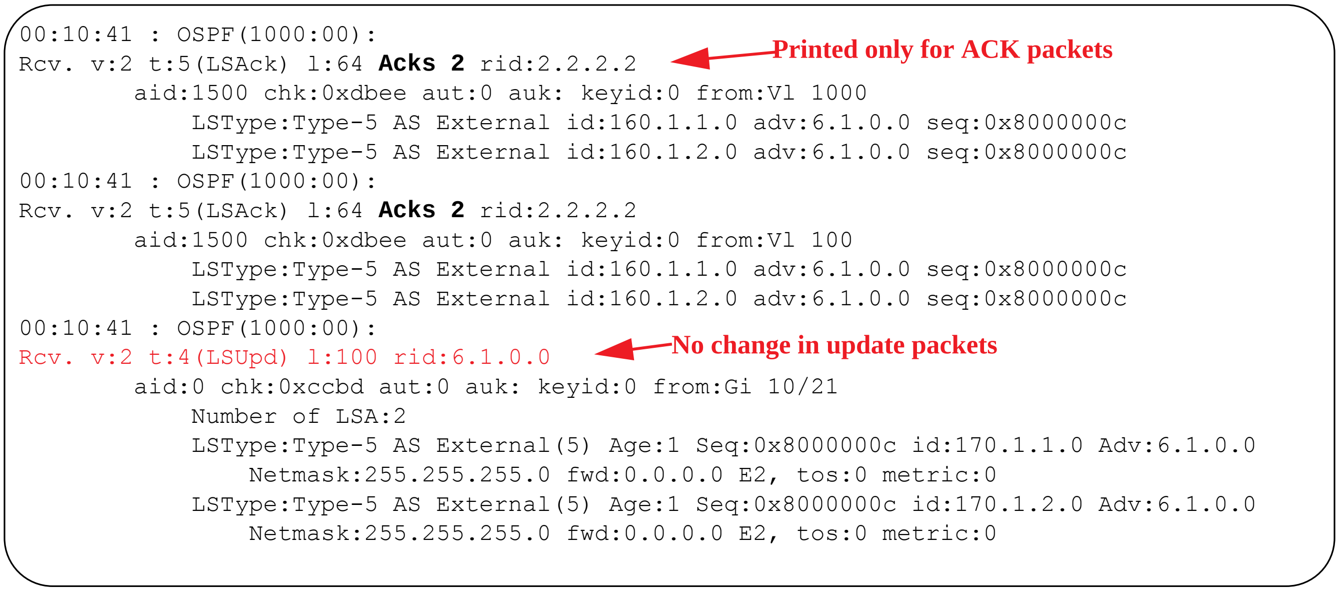
In OSPF, flooding is the most resource-consuming task. The flooding algorithm described in RFC 2328 requires that OSPF flood LSAs on all interfaces, as governed by LSA's flooding scope. (Refer to Section 13 of the RFC.) When multiple direct links connect two routers, the RFC 2328 flooding algorithm generates significant redundant information across all links.

By default, FTOS implements an enhanced flooding procedure which dynamically and intelligently detects when to optimize flooding. Wherever possible, the OSPF task attempts to reduce flooding overhead by selectively flooding on a subset of the interfaces between two routers.

If RFC 2328 flooding behavior is required, enable it by using the command **flood-2328** in ROUTER OSPF mode. When enabled, this command configures FTOS to flood LSAs on all interfaces.

Confirm RFC 2328 flooding behavior by using the command **debug ip ospf packet** and look for output similar to the following:

Figure 24-4. Enabling RFC-2328 Compliant OSPF Flooding



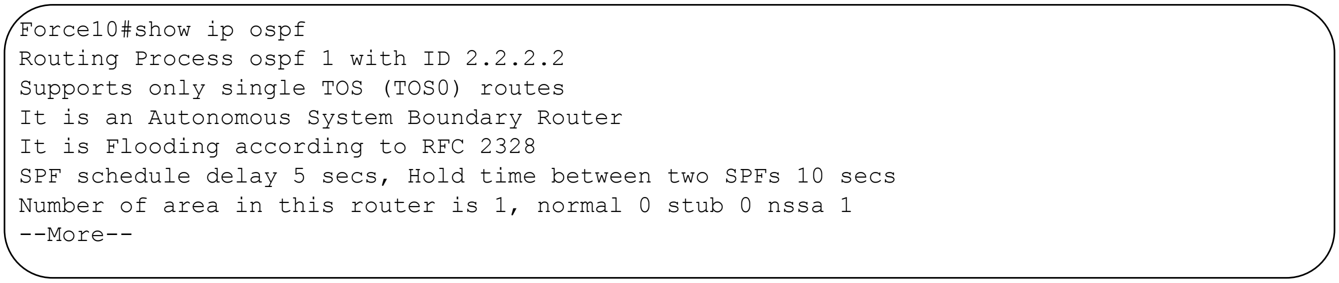
```
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:5(LSAck) l:64 Acks 2 rid:2.2.2.2
        aid:1500 chk:0xdbee aut:0 auk: keyid:0 from:Vl 1000
            LSType:Type-5 AS External id:160.1.1.0 adv:6.1.0.0 seq:0x8000000c
            LSType:Type-5 AS External id:160.1.2.0 adv:6.1.0.0 seq:0x8000000c
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:5(LSAck) l:64 Acks 2 rid:2.2.2.2
        aid:1500 chk:0xdbee aut:0 auk: keyid:0 from:Vl 100
            LSType:Type-5 AS External id:160.1.1.0 adv:6.1.0.0 seq:0x8000000c
            LSType:Type-5 AS External id:160.1.2.0 adv:6.1.0.0 seq:0x8000000c
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:4(LSUpd) l:100 rid:6.1.0.0
        aid:0 chk:0xccbd aut:0 auk: keyid:0 from:Gi 10/21
            Number of LSA:2
            LSType:Type-5 AS External(5) Age:1 Seq:0x8000000c id:170.1.1.0 Adv:6.1.0.0
                Netmask:255.255.255.0 fwd:0.0.0.0 E2, tos:0 metric:0
            LSType:Type-5 AS External(5) Age:1 Seq:0x8000000c id:170.1.2.0 Adv:6.1.0.0
                Netmask:255.255.255.0 fwd:0.0.0.0 E2, tos:0 metric:0
```

Printed only for ACK packets

No change in update packets

In FTOS Version, 7.5.1.0 use **show ip ospf** to confirm that RFC-2328 compliant OSPF flooding is enabled, as shown below.

Figure 24-5. Enabling RFC-2328 Compliant OSPF Flooding



```
Force10#show ip ospf
Routing Process ospf 1 with ID 2.2.2.2
Supports only single TOS (TOS0) routes
It is an Autonomous System Boundary Router
It is Flooding according to RFC 2328
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Number of area in this router is 1, normal 0 stub 0 nssa 1
--More--
```

OSPF ACK Packing

The OSPF ACK Packing feature bundles multiple LS acknowledgements in a single packet, significantly reducing the number of ACK packets transmitted when the number of LSAs increases. This feature also enhances network utilization and reduces the number of small ACK packets sent to a neighboring router. OSPF ACK packing is enabled by default, and non-configurable.

OSPF Adjacency with Cisco Routers

To establish an OSPF adjacency between Dell Force10 and Cisco routers, the hello interval and dead interval must be the same on both routers. In FTOS 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 FTOS. 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 as well.

To ensure equal intervals between the routers, manually set the dead interval of the Dell Force10 router to match the Cisco configuration. Use the command "**ip ospf dead-interval <x>**" in interface mode:

Figure 24-6. Command Example: ip ospf intervals

```
Force10(conf)#int gi 2/2
Force10(conf-if-gi-2/2)#ip ospf hello-interval 20
Force10(conf-if-gi-2/2)#ip ospf dead-interval 80

Force10(conf-if-gi-2/2)#
```

Dead Interval
Set at 4x
Hello Interval

Figure 24-7. OSPF Configuration with intervals set

```
Force10 (conf-if-gi-2/2)#ip ospf dead-interval 20
Force10 (conf-if-gi-2/2)#do show ip os int gi1/3
GigabitEthernet 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)
Force10 (conf-if-gi-2/2)#
```

Dead Interval
Set at 4x
Hello Interval

For more information regarding this functionality or for assistance, go to www.force10networks.com/support.

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.

OSPF must be configured 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)

Open Shortest Path First version 2 (OSPF for IPv4) is supported on platforms C E S

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.

The following configuration steps include two mandatory steps and several optional ones:

- Enable OSPFv2 (mandatory)
- Enable Multi-Process OSPF
- Assign an OSPFv2 area (mandatory)
- Enable OSPFv2 on interfaces
- Configure stub areas
- Enable passive interfaces
- Enable fast-convergence
- Change OSPFv2 parameters on interfaces
- Enable OSPFv2 authentication
- Enable graceful restart
- Configure virtual links
- Redistribute routes
- Troubleshooting OSPFv2

For a complete listing of all commands related to OSPFv2, refer to the OSPF section in the *FTOS Command Line Interface* document.

Enable OSPFv2

Assign an IP address to an interface (physical or Loopback) to enable Layer 3 routing. By default OSPF, like all routing protocols, is disabled.

You *must* configure at least one interface for Layer 3 before enabling OSPFv2 globally.

If implementing, Multi-Process OSPF, you must create an equal number of Layer 3 enabled interfaces and OSPF Process IDs. For example, if you create 4 OSPFv2 process IDs, you must have 4 interfaces with Layer 3 enabled.

Use these commands on one of the interfaces to enable OSPFv2 routing.

Step	Command Syntax	Command Mode	Usage
1	**ip address** *ip-address mask*	CONFIG-INTERFACE	Assign an IP address to an interface. Format: A.B.C.D/M
		If using a Loopback interface, refer to Loopback Interfaces on page 293.	
2	**no shutdown**	CONFIG-INTERFACE	Enable the interface.

Return to CONFIGURATION mode to enable the OSPF process. The OSPF Process ID is the identifying number assigned to the OSPF process, and the Router ID is the IP address associated with the OSPF process. .

Command Syntax	Command Mode	Usage
router ospf *process-id [vrf {vrf name}]*	CONFIGURATION	Enable the OSPFv2 process globally. Range: 0-65535 *vrf name*: Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.

Once the OSPF process and the VRF are tied together, the OSPF Process ID cannot be used again in the system.

If you try to enter an OSPF Process ID, or if you try to enable more OSPF processes than available Layer 3 interfaces, prior to assigning an IP address to an interface and setting the no shutdown command, you will see the following message.

Message 1

```
C300(conf)#router ospf 1
% Error: No router ID available.
```

In CONFIGURATION ROUTER OSPF mode, assign the Router ID. The Router ID is not required to be the router's IP address. Dell Force10 recommends using the IP address as the Router ID for easier management and troubleshooting.

Command Syntax	Command Mode	Usage
router-id *ip address*	CONFIG-ROUTER-OSPF-id	Assign the Router ID for the OSPFv2 process. IP Address: A.B.C.D

Use the **no router ospf** *process-id* command syntax in the CONFIGURATION mode to disable OSPF.

Use the **clear ip ospf** *process-id* command syntax in EXEC Privilege mode to reset the OSPFv2 process.

Use the **show ip ospf** *process-id* command in EXEC mode (Figure 408) to view the current OSPFv2 status.

Figure 24-8. Command Example: show ip ospf *process-id*

```
Force10#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
Force10#
```

Enable Multi-Process OSPF

Multi-Process OSPF allows multiple OSPFv2 processes on a single router. The following list shows the number of processes supported on each platform type.

- The E-Series supports up to 30 OSPFv2 processes.
- The C-Series supports up to 6 OSPFv2 processes.
- The S50 and S25 support up to 4 OSPFv2 processes.
- The S55, S60, and S4810 support up to 16 OSPFv2 processes.
- The Z9000 supports up to 3 OSPFv2 processes.

Follow the same steps as above, when configuring a single OSPF process. Repeat them as often as necessary for the desired number of processes. Once the process is created, all other configurations apply as usual,

Step	Command Syntax	Command Mode	Usage
1	**ip address** *ip-address mask*	CONFIG-INTERFACE	Assign an IP address to an interface. Format: A.B.C.D/M
		If using a Loopback interface, refer to Loopback Interfaces on page 293.	
2	**no shutdown**	CONFIG-INTERFACE	Enable the interface.

Return to CONFIGURATION mode to enable the OSPF process. The OSPF Process ID is the identifying number assigned to the OSPF process, and the Router ID is the IP address associated with the OSPF process. .

Command Syntax	Command Mode	Usage
router ospf *process-id* [**vrf** *{vrf name}]*	CONFIGURATION	Enable the OSPFv2 process globally. Range: 0-65535 *vrf name*: Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.

Once the OSPF process and the VRF are tied together, the OSPF Process ID cannot be used again in the system.

If you try to enable more OSPF processes than available Layer 3 interfaces you will see the following message.

Message 2

```
C300(conf)#router ospf 1
% Error: No router ID available.
```

In CONFIGURATION ROUTER OSPF mode, assign the Router ID. The Router ID is not required to be the router's IP address. Dell Force10 recommends using the IP address as the Router ID for easier management and troubleshooting.

Command Syntax	Command Mode	Usage
router-id *ip address*	CONFIG-ROUTER-OSPF-id	Assign the Router ID for the OSPFv2 process. IP Address: A.B.C.D

Use the **no router ospf** *process-id* command syntax in the CONFIGURATION mode to disable OSPF.

Use the **clear ip ospf** *process-id* command syntax in EXEC Privilege mode to reset the OSPFv2 process.

Assign an OSPFv2 area

After OSPFv2 is enabled, 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, you must also 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 since it is already included in the first network address.

When configuring the **network** command, you must configure a network address and mask that is a superset of the IP subnet configured on the Layer-3 interface to be used for OSPFv2.

Use this command in CONFIGURATION ROUTER OSPF mode to set up each neighbor and OSPF area. The Area can be assigned by a number or with an IP interface address.

Command Syntax	Command Mode	Usage
network *ip-address mask* **area** *area-id*	CONFIG-ROUTER-OSPF-id	Enable OSPFv2 on an interface and assign an network address range to a specific OSPF area. IP Address Format: A.B.C.D/M Area ID Range: 0-65535 or A.B.C.D/M

Enable OSPFv2 on interfaces

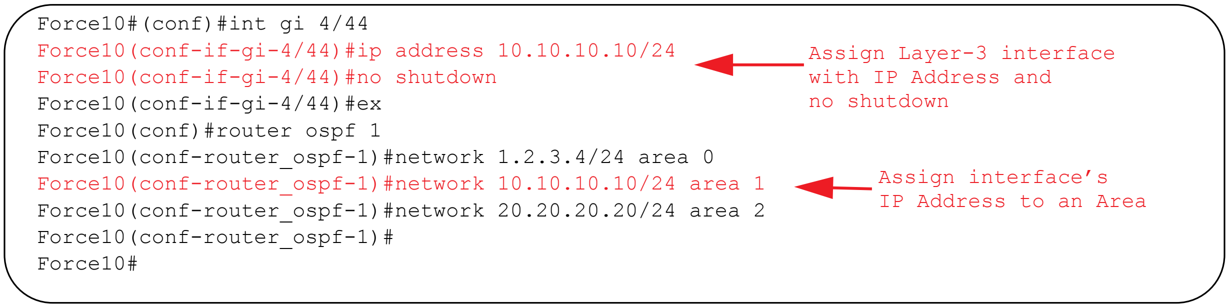
Each interface must have OSPFv2 enabled on it. It must be configured for Layer 3 protocol, and not be shutdown. OSPFv2 can also be assigned to a loopback interface as a virtual interface.

OSPF functions and features, such as MD5 Authentication, Grace Period, Authentication Wait Time, etc, are assigned on a per interface basis.

Note: If using features like MD5 Authentication, ensure all the neighboring routers are also configured for MD5.

Figure 24-9 presents an example of assigning an IP address to an interface and then assigning an OSPFv2 area that includes that Layer-3 interface's IP address.

Figure 24-9. Configuring an OSPF Area Example



```
Force10#(conf)#int gi 4/44
Force10(conf-if-gi-4/44)#ip address 10.10.10.10/24
Force10(conf-if-gi-4/44)#no shutdown
Force10(conf-if-gi-4/44)#ex
Force10(conf)#router ospf 1
Force10(conf-router_ospf-1)#network 1.2.3.4/24 area 0
Force10(conf-router_ospf-1)#network 10.10.10.10/24 area 1
Force10(conf-router_ospf-1)#network 20.20.20.20/24 area 2
Force10(conf-router_ospf-1)#
Force10#
```

Assign Layer-3 interface with IP Address and no shutdown

Assign interface's IP Address to an Area

Dell Force10 recommends that the OSPFv2 Router ID be the interface IP addresses for easier management and troubleshooting.

Use the **show config** command in CONFIGURATION ROUTER OSPF mode to view the configuration.

OSPF, by default, sends hello packets out to all physical interfaces assigned an IP address that are a subset of a network on which OSPF is enabled. Use the **show ip ospf interface** command (Figure 410) to view the interfaces currently active and the areas assigned to the interfaces.

Figure 24-10. Command Example: show ip ospf *process-id* interface

```
Force10>show ip ospf 1 interface

GigabitEthernet 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

GigabitEthernet 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)
Force10>
```

Loopback interfaces also assist in the OSPF process. OSPF will pick the highest interface address as the router-id and a loopback interface address has a higher precedence than other interface addresses.

Figure 24-11 gives an example of the **show ip ospf** *process-id interface* command with a Loopback interface.

Figure 24-11. Command Example: show ip ospf *process-id* interface

```
Force10#show ip ospf 1 int

GigabitEthernet 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.
Force10#
```

Configure 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 Area Border Router (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.

Use these commands in the following sequence, starting in EXEC Privilege mode to configure a stub area.

Step	Command Syntax	Command Mode	Usage
1	**show ip ospf** *process-id [***vrf** *vrf name]* **database database-summary**	EXEC Privilege	Review all areas after they were configured to determine which areas are NOT receiving type 5 LSAs. *vrf name*: Show only the OSPF information tied to the VRF process.
2	**configure**	EXEC Privilege	Enter the CONFIGURATION mode.
3	**router ospf** *process-id* **[vrf** *{vrf name}]*	CONFIGURATION	Enter the ROUTER OSPF mode. Process ID is the ID assigned when configuring OSPFv2 globally (page 58). *vrf name*: Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.
4	**area** *area-id* **stub** [**no-summary**]	CONFIG-ROUTER-OSPF-id	Configure the area as a stub area. Use the **no-summary** keywords to prevent transmission in to the area of summary ASBR LSAs. Area ID is the number or IP address assigned when creating the Area (page 60).

Use the **show ip ospf database** *process-id* **database-summary** command syntax (Figure 413) in the EXEC Privilege mode To view which LSAs are transmitted.

Figure 24-12. Command Example: show ip ospf *process-id* database database-summary

```
Force10#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
Force10#
```

To view information on areas, use the **show ip ospf** *process-id* command in the EXEC Privilege mode.

Enable 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 will neither send nor receive routing updates, the network on that interface will still be included in OSPF updates sent via other interfaces.

Use the following command in the ROUTER OSPF mode to suppress the interface's participation on an OSPF interface. This command stops the router from sending updates on that interface.

Command Syntax	Command Mode	Usage
passive-interface {default | interface}	CONFIG-ROUTER-OSPF-id	Specify whether all or some of the interfaces will be passive. **Default** enabled passive interfaces on ALL interfaces in the OSPF process. Entering the physical interface type, slot, and number enable passive interface on only the identified interface. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information (**e.g. passive-interface gi 2/1**). • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale (**e.g. passive-interface po 100**) • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information (**e.g. passive-interface ten 2/3**). • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 (**e.g. passive-interface vlan 2222)**. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.
	The default keyword sets all interfaces on this OSPF process as passive. The passive interface can be removed from select interfaces using the `no passive-interface interface` command while **passive interface default** is configured.	

To enable both receiving and sending routing updates, enter the `no passive-interface interface` command.

When you configure a passive interface, the **show ip ospf** *process-id* **interface** command (Figure 413) adds the words "`passive interface`" to indicate that hello packets are not transmitted on that interface.

Figure 24-13. Command Example: show ip ospf *process-id* interface

```
Force10#show ip ospf 34 int

GigabitEthernet 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

GigabitEthernet 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
Loopback interface is treated as a stub Host.
Force10#
```

Interface is not running the OSPF protocol.

Enable fast-convergence

The fast-convergence CLI sets the minimum origination and arrival LSA parameters to zero (0), allowing rapid route calculation. When fast-convergence is disabled, origination and arrival LSA parameters are set to 5 seconds and 1 second, respectively.

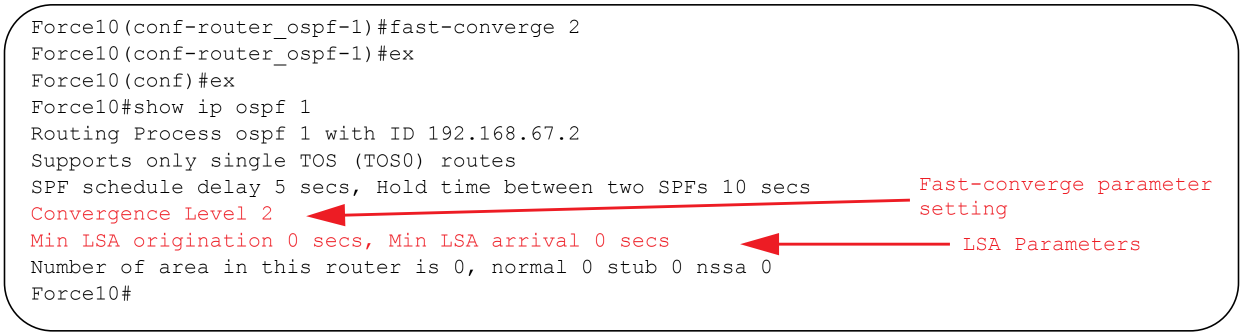
Setting the convergence parameter (1-4) indicates the actual convergence level. Each convergence setting adjusts the LSA parameters to zero, but the fast-convergence parameter setting allows for even finer tuning of the convergence speed. The higher the number, the faster the convergence. Use the following command in the ROUTER OSPF mode to enable or disable fast-convergence.

Command Syntax	Command Mode	Usage
fast-convergence {*number*}	CONFIG-ROUTER-OSPF-id	Enable OSPF fast-convergence and specify the convergence level. **Parameter: 1-4** The higher the number, the faster the convergence. When disabled, the parameter is set at 0 (Figure 24-15).

Note: A higher convergence level can result in occasional loss of OSPF adjacency. Generally, convergence level 1 meets most convergence requirements. Higher convergence levels should only be selected following consultation with Dell Force10 technical support.

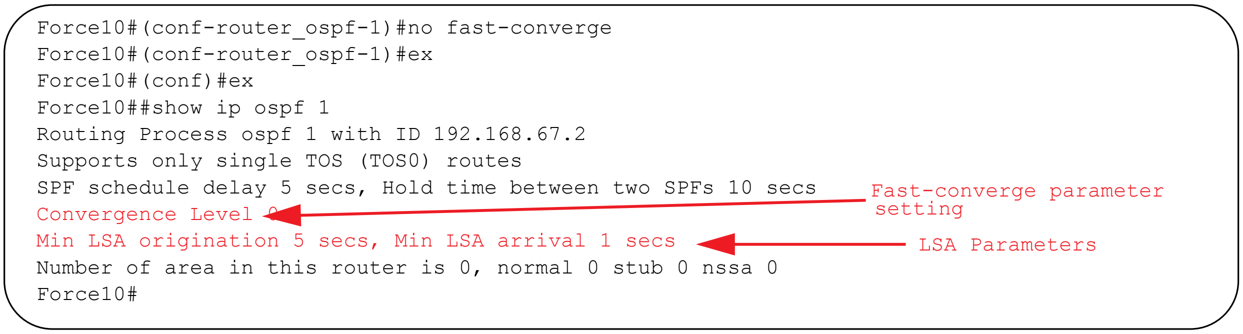
Figure 24-14 shows the convergence settings when fast-convergence is enabled and Figure 24-15 shows settings when fast-convergence is disabled. These displays appear with the **show ip ospf** command.

Figure 24-14. Command Example: show ip ospf *process-id* (fast-convergence enabled)



```
Force10(conf-router_ospf-1)#fast-converge 2
Force10(conf-router_ospf-1)#ex
Force10(conf)#ex
Force10#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
Force10#
```

Figure 24-15. Command example: show ip ospf *process-id* (fast-convergence disabled)



```
Force10#(conf-router_ospf-1)#no fast-converge
Force10#(conf-router_ospf-1)#ex
Force10#(conf)#ex
Force10##show ip ospf 1
Routing Process ospf 1 with ID 192.168.67.2
Supports only single TOS (TOS0) routes
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Convergence Level 0
Min LSA origination 5 secs, Min LSA arrival 1 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
Force10#
```

Change OSPFv2 parameters on interfaces

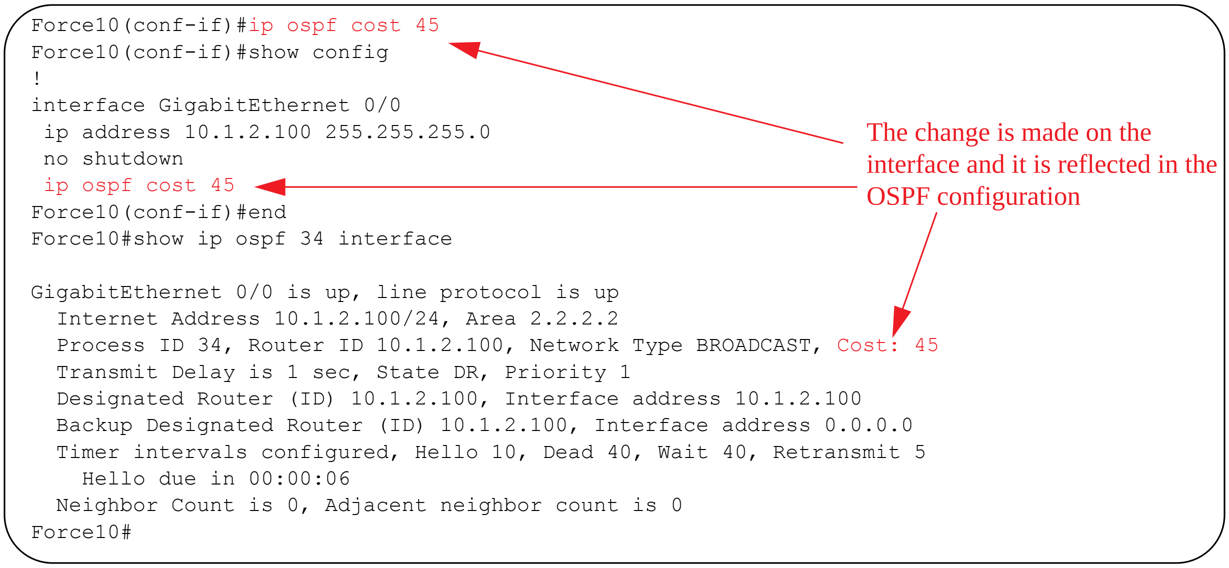
In FTOS, you can modify the OSPF settings on the interfaces. Some interface parameter values must be consistent across all interfaces to avoid routing errors. For example, you must set the same time interval for the hello packets on all routers in the OSPF network to prevent misconfiguration of OSPF neighbors.

Use any or all of the following commands in CONFIGURATION INTERFACE mode to change OSPFv2 parameters on the interfaces:

Command Syntax	Command Mode	Usage
ip ospf *cost*	CONFIG-INTERFACE	Change the cost associated with OSPF traffic on the interface. Cost: 1 to 65535 (default depends on the interface speed).
ip ospf dead-interval *seconds*	CONFIG-INTERFACE	Change the time interval the router waits before declaring a neighbor dead. Configure Seconds range: 1 to 65535 (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.
ip ospf hello-interval *seconds*	CONFIG-INTERFACE	Change the time interval between hello-packet transmission. Seconds range: from 1 to 65535 (default is 10 seconds). The hello interval must be the same on all routers in the OSPF network.
ip ospf message-digest-key *keyid* **md5** *key*	CONFIG-INTERFACE	Use the MD5 algorithm to produce a message digest or key, which is sent instead of the key. Keyid range: 1 to 255 Key: a character string Be sure to write down or otherwise record the Key. You cannot learn the key once it is configured. You must be careful when changing this key.
ip ospf priority *number*	CONFIG-INTERFACE	Change the priority of the interface, which is used to determine the Designated Router for the OSPF broadcast network. Number range: 0 to 255 (the default is 1).
ip ospf retransmit-interval *seconds*	CONFIG-INTERFACE	Change the retransmission interval between LSAs. Seconds range: from 1 to 65535 (default is 5 seconds). The retransmit interval must be the same on all routers in the OSPF network.
ip ospf transmit-delay *seconds*	CONFIG-INTERFACE	Change the wait period between link state update packets sent out the interface. Seconds range: from 1 to 65535 (default is 1 second). The transmit delay must be the same on all routers in the OSPF network.

Use the **show config** command in CONFIGURATION INTERFACE mode (Figure 24-16) to view interface configurations. Use the **show ip ospf interface** command in EXEC mode to view interface status in the OSPF process.

Figure 24-16. Changing the OSPF Cost Value on an Interface



```
Force10(conf-if)#ip ospf cost 45
Force10(conf-if)#show config
!
interface GigabitEthernet 0/0
 ip address 10.1.2.100 255.255.255.0
 no shutdown
 ip ospf cost 45
Force10(conf-if)#end
Force10#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
Force10#
```

The change is made on the interface and it is reflected in the OSPF configuration

Enable OSPFv2 authentication

Use the following commands in CONFIGURATION INTERFACE mode to enable or change various OSPF authentication parameters:

Command Syntax	Command Mode	Usage
ip ospf authentication-key *key*	CONFIG-INTERFACE	Set clear text authentication scheme on the interface. Configure a *key* that is a text string no longer than eight characters. All neighboring routers must share the same password to exchange OSPF information.
ip ospf auth-change-wait-time *seconds*	CONFIG-INTERFACE	Set the authentication change wait time in *seconds* between 0 and 300 for the interface. This 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.

Enable graceful restart

Graceful Restart is enabled for the global OSPF process. Use these commands to configure OSPF graceful restart.

The Dell Force10 implementation of OSPF 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, OSPF graceful restart is disabled.

You enable OSPF graceful restart in CONFIGURATION ROUTER OSPF mode.

Command Syntax	Command Mode	Usage
graceful-restart grace-period *seconds*	CONFIG-ROUTER-OSPF-id	Enable OSPF graceful-restart globally and set the grace period. Seconds range: between 40 and 3000 This is the period of time that an OSPF router's neighbors will advertise it as fully adjacent, regardless of the synchronization state, during a graceful restart. OSPF terminates this process when the grace period ends.
graceful-restart helper-reject *router-id*	CONFIG-ROUTER-OSPF-id	Enter the Router ID of the OSPF helper router from which the router does not accept graceful restart assistance. This applies to the specified router only. IP Address: A.B.C.D
graceful-restart mode [planned-only | unplanned-only]	CONFIG-ROUTER-OSPF-id	Specify the operating mode in which graceful-restart functions. FTOS supports the following options: • Planned-only. The OSPF router supports graceful-restart for planned restarts only. A planned restart is when the user manually enters 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 OSPF router supports graceful-restart for only unplanned restarts. During an unplanned restart, OSPF sends out a Grace LSA once the secondary RPM comes online. By default, OSPF supports both planned and unplanned restarts. Selecting one or the other mode restricts OSPF to the single selected mode.
graceful-restart role [helper-only | restart-only]	CONFIG-ROUTER-OSPF-id	Configure the graceful restart role or roles that this OSPF router performs. FTOS supports the following options: • Helper-only. The OSPF router supports graceful-restart only as a helper router. • Restart-only. The OSPF router supports graceful-restart only during unplanned restarts. By default, OSPF supports both restarting and helper roles. Selecting one or the other role restricts OSPF to the single selected role.

When you configure a graceful restart, the **show run ospf** command (Figure 24-17) displays information similar to the following.

Figure 24-17. Command Example: show run ospf (partial)

```
Force10#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
Force10#
```

Use the following command to disable OSPF graceful-restart after you have enabled it.

Command Syntax	Command Mode	Usage
no graceful-restart grace-period	CONFIG-ROUTER-OSPF-id	Disable OSPF graceful-restart. Returns OSPF graceful-restart to its default state.

For more information on OSPF graceful restart, refer to the *FTOS Command Line Interface Reference Guide*.

Configure virtual links

Areas within OSPF must be connected to the backbone area (Area ID 0.0.0.0). If an OSPF area does not have a direct connection to the backbone, at least one virtual link is required. Virtual links must be configured on an ABR connected to the backbone.

- hello-interval: help packet
- retransmit-interval: LSA retransmit interval
- transmit-delay: LSA transmission delay
- dead-interval: dead router detection time
- authentication-key: authentication key
- message-digest-key: MD5 authentication key

Use the following command in CONFIGURATION ROUTER OSPF mode to configure virtual links.

Command Syntax	Command Mode	Usage
area *area-id* **virtual-link** *router-id* [hello-interval *seconds* | retransmit-interval *seconds* | transmit-delay *seconds* | dead-interval *seconds* | authentication-key *key* | message-digest-key *keyid* md5 *key*]	CONFIG-ROUTER-OSPF-id	Configure the optional parameters of a virtual link: • Area ID: assigned earlier (0-65535 or A.B.C.D) • Router ID: IP address associated with the virtual link neighbor • Hello Interval Seconds: 1-8192 (default 10) • Retransmit Interval Seconds: 1-3600 (default 5) • Transmit Delay Seconds: 1-3600 (default 1) • Dead Interval Seconds: 1-8192 (default 40) • Authentication Key: 8 characters • Message Digest Key: 1-255 • MD5 Key: 16 characters Only the Area ID and Router ID require configuration to create a virtual link. If no other parameter is entered, the defaults are used. Use EITHER the Authentication Key or the Message Digest (MD5) key.

Use the **show ip ospf** *process-id* **virtual-links** command (Figure 24-18) in the EXEC mode to view the virtual link.

Figure 24-18. Command Example: show ip ospf *process-id* virtual-links

```
Force10#show ip ospf 1 virtual-links

Virtual Link to router 192.168.253.5 is up
    Run as demand circuit
    Transit area 0.0.0.1, via interface GigabitEthernet 13/16, Cost of using 2
    Transmit Delay is 1 sec, State POINT_TO_POINT,
        Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
        Hello due in 00:00:02
Force10#
```

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, and 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.

Command Syntax	Command Mode	Usage
ip prefix-list *prefix-name*	CONFIGURATION	Create a prefix list and assign it a unique name. You are in PREFIX LIST mode.

Command Syntax	Command Mode	Usage
seq *sequence-number* {deny |permit} *ip-prefix* [ge min-prefix-length] [le max-prefix-length]	CONFIG- PREFIX LIST	Create a prefix list with a sequence. number and a deny or permit action. The optional parameters are: **ge** min-prefix-length: is the minimum prefix length to be matched (0 to 32). le *max-prefix-length:* is the maximum prefix length to be matched (0 to 32).

For configuration information on prefix lists, refer to *IP Access Control Lists, Prefix Lists, and Route-maps* chapter in the *FTOS Configuration Guide*.

Use the following commands in CONFIGURATION-ROUTER OSPF mode to apply prefix lists to incoming or outgoing OSPF routes

Command Syntax	Command Mode	Usage
distribute-list *prefix-list-name* **in** [*interface*]	CONFIG-ROUTER-OSPF-id	Apply a configured prefix list to incoming OSPF routes.
distribute-list *prefix-list-name* **out** [**connected** | **isis** | **rip** | **static**]	CONFIG-ROUTER-OSPF-id	Assign a configured prefix list to outgoing OSPF routes.

Redistribute routes

You can add routes from other routing instances or protocols to the OSPF process. With the **redistribute** command syntax, 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.

Use the following command in CONFIGURATION- ROUTER-OSPF mode to redistribute routes:

Command Syntax	Command Mode	Usage
redistribute {**bgp** | **connected** | **isis** | **rip** | **static**} [**metric** *metric-value* | **metric-type** *type-value*] [**route-map** *map-name*] [**tag** *tag-value*]	CONFIG-ROUTER-OSPF-id	Specify which routes will be redistributed into OSPF process. Configure the following required and optional parameters: • **bgp, connected**, **isis**, **rip**, or **static**: enter one of the keyword to redistribute those routes. **rip** is supported only on E-Series. • **metric** *metric-value* range: 0 to 4294967295. • **metric-type** *metric-type*: 1 for OSPF external route type 1 or 2 for OSPF external route type 2. • **route-map** *map-name*: enter a name of a configured route map. • **tag** *tag-value* range: 0 to 4294967295.

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

Figure 24-19. Command Example: show config

```
Force10(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
Force10(conf-router_ospf)#
```

Troubleshooting OSPFv2

FTOS has several tools to make troubleshooting easier. Be sure to check the following, as these are typical issues that interrupt an OSPFv2 process. Note that this is not a comprehensive list, just some examples of typical troubleshooting checks.

- Has OSPF been enabled 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 virtual links
- show routes

Use the **show running-config ospf** command to see the state of all the enabled OSPFv2 processes.

Command Syntax	Command Mode	Usage
show running-config ospf	EXEC Privilege	View the summary of all OSPF process IDs enables on the router.

Figure 24-20. Command Example: show running-config ospf

```
Force10#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
Force10#
```

Use the following commands in EXEC Privilege mode to get general route and links status information.

Command Syntax	Command Mode	Usage
show ip route summary	EXEC Privilege	View the summary information of the IP routes
show ip ospf database	EXEC Privilege	View the summary information for the OSPF database

Use the following command in EXEC Privilege mode to view the OSPFv2 configuration for a neighboring router:

Command Syntax	Command Mode	Usage
show ip ospf neighbor	EXEC Privilege	View the configuration of OSPF neighbors connected to the local router.

Use the following command in EXEC Privilege mode to configure the debugging options of an OSPFv2 process:

Command Syntax	Command Mode	Usage
debug ip ospf *process-id* [**event** | **packet** | **spf**]	EXEC Privilege	View debug messages. To view debug messages for a specific OSPF process ID, enter **debug ip ospf** *process-id.* 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 shortest path first (spf) information.

Configuration Task List for OSPFv3 (OSPF for IPv6)

Open Shortest Path First version 3 (OSPF for IPv6) is supported on platforms C E

The configuration options of OSPFv3 are the same as those for OSPFv2, but may be configured with differently labeled commands. Process IDs and areas need to be specified. Interfaces and addresses need to be included in the process. Areas can be defined 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, they must be assigned to OSPF areas.

TheOSPFv3 **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 that the OSPFv2 **network area** command can enable OSPF on multiple interfaces with the single command, while the OSPFv3 **ipv6 ospf area** command must be configured on each interface that will be running OSPFv3.

All IPv6 addresses on an interface are included in the OSPFv3 process that is created on the interface.

OSPFv3 for IPv6 is enabled by specifying an OSPF Process ID and an Area in the INTERFACE mode. If an OSPFv3 process has not yet been created, it is created automatically. All IPv6 addresses configured on the interface are included in the specified OSPF process.

Note: IPv6 and OSPFv3 do *not* support Multi-Process OSPF. Only a single OSPFv3 process is can be enabled.

- Enable IPv6 Unicast Routing
- Assign IPv6 addresses on an interface
- Assign Area ID on interface
- Assign OSPFv3 Process ID and Router ID Globally
- Configure stub areas
- Configure Passive-Interface
- Redistribute routes
- Configure a default route

Enable IPv6 Unicast Routing

Command Syntax	Command Mode	Usage
ipv6 unicast routing	CONFIGURATION	Enables IPv6 unicast routing globally.

Assign IPv6 addresses on an interface

Command Syntax	Command Mode	Usage
ipv6 address *ipv6 address*	CONF-INT-type slot/port	Assign IPv6 address to the interface. IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). FORMAT: A:B:C::F/128
no shutdown	CONF-INT-type slot/port	Bring the interface up.

Assign Area ID on interface

Command Syntax	Command Mode	Usage
ipv6 ospf *process-id* **area** *area-id*	CONF-INT-type slot/port	Assign the OSPFv3 process and an OSPFv3 area to this interface. *process-id*: The Process ID number assigned above. *area-id*: the area ID for this interface.

The **ipv6 ospf area** command enables OSPFv3 on an interface and places the interface in the specified area. Additionally, it creates the OSPFv3 process with ID on the router. OSPFv2 required 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 OSPFv2 on an interface. Note that the OSPFv2 **network area** command can enable OSPFv2 on multiple interfaces with the single command, whereas the OSPFv3 **ipv6 ospf area** command must be configured on each interface that will be running OSPFv3.

Assign OSPFv3 Process ID and Router ID Globally

Command Syntax	Command Mode	Usage
ipv6 router ospf {*process ID}*	CONFIGURATION	Enable the OSPFv3 process globally and enter OSPFv3 mode. Range: 0-65535
router-id *{number}*	CONF-IPV6-ROUTER-OSPF	Assign the Router ID for this OSPFv3 process *number*: IPv4 address Format: A.B.C.D
	Note: The router-id for an OSPFv3 router is entered as an IPv4 IP address.	

Configure stub areas

Command Syntax	Command Mode	Usage
area *area-id* **stub** [**no-summary**]	CONF-IPV6-ROUTER-OSPF	Configure the area as a stub area. Use the **no-summary** keywords to prevent transmission in to the area of summary ASBR LSAs. *Area ID* is a number or IP address assigned when creating the Area. The Area ID can be represented as a number between 0 – 65536 if a dotted decimal format is assigned, rather than an IP address.

Configure Passive-Interface

Use the following command to suppress the interface's participation on an OSPFv3 interface. This command stops the router from sending updates on that interface.

Command Syntax	Command Mode	Usage
passive-interface {type slot/port}	CONF-IPV6-ROUTER-OSPF	Specify whether some or all some of the interfaces will be passive. **Interface** identifies the specific interface that will be passive. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information (**e.g. passive-interface gi 2/1**). • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale (**e.g. passive-interface po 100**) • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information (**e.g. passive-interface ten 2/3**). • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 (**e.g. passive-interface vlan 2222)**. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

To enable both receiving and sending routing updates, enter the `no passive-interface interface` command.

When you configure a passive interface, the **show ipv6 ospf interface** command adds the words "`passive interface`" to indicate that hello packets are not transmitted on that interface.

Redistribute routes

You can add routes from other routing instances or protocols to the OSPFv3 process. With the **redistribute** command syntax, you can include RIP, static, or directly connected routes in the OSPF process.

Command Syntax	Command Mode	Usage
redistribute {**bgp** | **connected** | **static**} [**metric** *metric-value* | **metric-type** *type-value*] [**route-map** *map-name*] [**tag** *tag-value*]	CONF-IPV6-ROUTER-OSPF	Specify which routes will be redistributed into OSPF process. Configure the following required and optional parameters: • **bgp, connected**, or **static**: enter one of the keyword to redistribute those routes. • **metric** *metric-value* range: 0 to 4294967295. • **metric-type** *metric-type*: 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* range: 0 to 4294967295.

Configure a default route

Configure FTOS to generate a default external route into the OSPFv3 routing domain.

Command Syntax	Command Mode	Usage
default-information originate [**always** [**metric** *metric-value*] [**metric-type** *type-value*]] [**route-map** *map-name*]	CONF-IPV6-ROUTER-OSPF	Specify the information for the default route. Configure the following required and optional parameters: • **always**: indicate that default route information must always be advertised • **metric** *metric-value* range: 0 to 4294967295. • **metric-type** *metric-type*: 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.

Troubleshooting OSPFv3

FTOS has several tools to make troubleshooting easier. Be sure to check the following, as these are typical issues that interrupt the OSPFv3 process. Note that this is not a comprehensive list, just some examples of typical troubleshooting checks.

- Has OSPF been enabled 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 ipv6 interfaces
- show ipv6 protocols
- debug IPv6 OSPF events and/or packets
- show ipv6 neighbors
- show virtual links
- show ipv6 routes

Use the following commands in EXEC Privilege mode to get general route and links status information.

Command Syntax	Command Mode	Usage
show ipv6 route summary	EXEC Privilege	View the summary information of the IPv6 routes
show ipv6 ospf database	EXEC Privilege	View the summary information for the OSPFv3 database

Use the following command in EXEC Privilege mode to view the OSPF configuration for a neighboring router:

Command Syntax	Command Mode	Usage
show ipv6 ospf neighbor	EXEC Privilege	View the configuration of OSPFv3 neighbors.

Use the following command in EXEC Privilege mode to configure the debugging options of an OSPFv3 process:

Command Syntax	Command Mode	Usage
debug ipv6 ospf packet {type slot/port}	EXEC Privilege	View debug messages for all OSPFv3 interfaces. • **packet**: view OSPF packet information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information (**e.g. passive-interface gi 2/1**). • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale (**e.g. passive-interface po 100**) • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information (**e.g. passive-interface ten 2/3**). • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 (**e.g. passive-interface vlan 2222)**. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Sample Configurations for OSPFv2

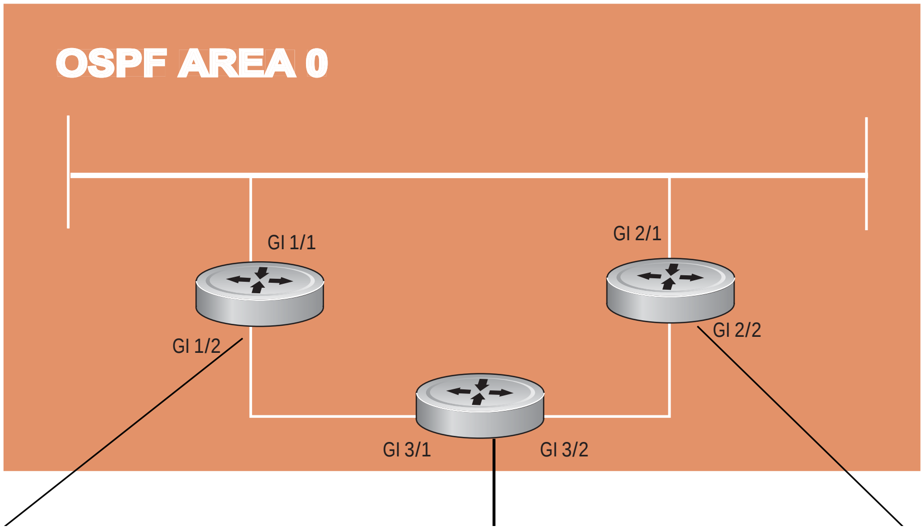
The following configurations are examples for enabling OSPFv2. These are not comprehensive directions. They are intended to give you a some guidance with typical configurations.

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.

Basic OSPFv2 Router Topology

The following illustration is a sample basic OSPFv2 topology.

Figure 24-21. Basic topology and CLI commands for OSPFv2



```
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
```

```
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
```

```
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
```

25

PIM Sparse-Mode

PIM Sparse-Mode is supported on platforms: C E S

PIM-Sparse Mode (PIM-SM) is a multicast protocol that forwards multicast traffic to a subnet only upon 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

- The Dell Force10 implementation of PIM-SM is based on the IETF *Internet Draft draft-ietf-pim-sm-v2-new-05*.
- C-Series supports a maximum of 31 PIM interfaces and 4K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors C-Series can have.
- S-Series 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 S-Series 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 upon receiving the first multicast packet.
- FTOS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.
- FTOS supports PIM-SM on physical, VLAN, and port-channel interfaces.
- FTOS supports 2000 IPv6 multicast forwarding entries, with up to 128 PIM-SSM neighbors/ interfaces.
- PIM-SM on VLAN interfaces is supported on the E-Series on TeraScale platforms only.
- IPv6 Multicast is not supported on 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. Once a receiver receives traffic from the RP, PM-SM switches to shortest path trees (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 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. Upon 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.

Refusing 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. Upon 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.

Sending 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. Once 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. Upon 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 will receive 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.

FTOS 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. FTOS 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 Points to Remember

- If a loopback interface with a /32 mask is used as the RP, you must enable PIM Sparse-mode on the interface.

Configure PIM-SM

Configuring PIM-SM is a two-step process:

1. Enable multicast routing using the command **ip multicast-routing** from CONFIGURATION mode.
2. Select a Rendezvous Point.
3. Enable PIM-SM on an interface. See page 500.

Related Configuration Tasks

- Configurable S,G Expiry Timers on page 501
- Configure a Static Rendezvous Point on page 502
- Configure a Designated Router on page 503
- Create Multicast Boundaries and Domains on page 504

Enable PIM-SM

You must enable PIM-SM on each participating interface:

Step	Task	Command	Command Mode
1	Enable multicast routing on the system.	**ip multicast-routing**	CONFIGURATION
2	Enable PIM-Sparse Mode	**ip pim sparse-mode**	INTERFACE

Display which interfaces are enabled with PIM-SM using the command **show ip pim interface** from EXEC Privilege mode, as shown in Figure 25-1.

Figure 25-1. Viewing PIM-SM Enabled Interfaces

```
Force10#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
Force10#
```

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.

Display PIM neighbors for each interface using the command **show ip pim neighbor** from EXEC Privilege mode, as shown in Figure 25-2.

Figure 25-2. Viewing PIM Neighbors Command Example

```
Force10#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
Force10#
```

Display the PIM routing table using the command **show ip pim tib** from EXEC privilege mode, as shown in Figure 25-3.

Figure 25-3. Viewing the PIM Multicast Routing Table

```
Force10#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--
```

Configurable 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 a expiry time for a particular entry. If both are configured, the ACL supercedes the global configuration for the specified entries.

When an expiry time created, deleted, or updated, the changes are applied when the keep alive timer refreshes.

To configure a global expiry time:

Task	Command	Command Mode
Enable global expiry timer for S, G entries Range 211-86400 seconds Default: 210	**ip pim sparse-mode sg-expiry-timer** *seconds*	CONFIGURATION

Configure the expiry time for a particular (S,G) entry:

Step	Task	Command Syntax	Command Mode
1	Create an Extended ACL	**ip access-list extended** *access-list-name*	CONFIGURATION
2	Specify the source and group to which the timer will be applied using extended ACLs with permit rules only.	[**seq** *sequence-number*] **permit ip** *source-address/mask* **| any** | **host** *source-address*} {*destination-address/mask* | **any** | **host** *destination-address*}	CONFIG-EXT-NACL

Step	Task	Command Syntax	Command Mode
3	Set the expiry time for a specific (S,G) entry (Figure 25-4). Range 211-86400 seconds Default: 210	**ip pim sparse-mode sg-expiry-timer** *seconds* **sg-list** *access-list-name*	CONFIGURATION

Note: The expiry time configuration is nullified, and the default global expiry time is used if:

- an ACL is specified for an 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.

Figure 25-4. Configuring an (S,G) Expiry Time

```
Force10(conf)#ip access-list extended SGtimer
Force10(config-ext-nacl)#permit ip 10.1.2.3/24 225.1.1.0/24
Force10(config-ext-nacl)#permit ip any 232.1.1.0/24
Force10(config-ext-nacl)#permit ip 100.1.1.0/16 any
Force10(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
Force10(config-ext-nacl)#exit

Force10(conf)#ip pim sparse-mode sg-expiry-timer 1800 sg-list SGtimer
```

Display the expiry time configuration using the **show running-configuration** [**acl | pim**] command from EXEC Privilege mode.

Configure a Static Rendezvous Point

The rendezvous point 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 using the command **ip pim rp-address**, as shown in Figure 25-5.

Figure 25-5. Electing a Rendezvous Point

```
Force10#sh run int loop0
!
interface Loopback 0
 ip address 1.1.1.1/32
 ip pim sparse-mode
 no shutdown
Force10#sh run pim
!
ip pim rp-address 1.1.1.1 group-address 224.0.0.0/4
```

Override Bootstrap Router Updates

PIM-SM routers need to 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.

If you have configured a static RP for a group, use the option **override** with the command **ip pim rp-address** to override bootstrap router updates with your static RP configuration. If you do not use this option, the RPs advertised in the BSR updates take precedence over any statically configured RPs.

Display the assigned RP for a group using the command **show ip pim rp** from EXEC privilege mode, as shown in Figure 25-6.

Figure 25-6. Displaying the Rendezvous Point for a Multicast Group

```
Force10#show ip pim rp
Group             RP
225.0.1.40        165.87.50.5
226.1.1.1         165.87.50.5
```

Display the assigned RP for a group range (group-to-RP mapping) using the command **show ip pim rp mapping** command in EXEC privilege mode

Figure 25-7. Display the Rendezvous Point for a Multicast Group Range

```
Force10#show ip pim rp mapping
PIM Group-to-RP Mappings
Group(s): 224.0.0.0/4, Static
  RP: 165.87.50.5, v2
```

Configure a Designated Router

Multiple PIM-SM routers might be connected to a single 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 using the command **ip pim dr-priority priority-value** from INTERFACE mode.
- Change the interval at which a router sends hello messages using the command **ip pim query-interval seconds** from INTERFACE mode.
- Display the current value of these parameter using the command **show ip pim interface** EXEC Privilege mode.

Create 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, use the **ip pim bsr-border** command. This command is applied to the subsequent inbound and outbound updates. Already existing BSR advertisements are removed by timeout.

Remove candidate RP advertisements using the **clear ip pim rp-mapping** command.

- .

26

PIM Source-Specific Mode

PIM Source-Specific Mode is supported on platforms: C E S

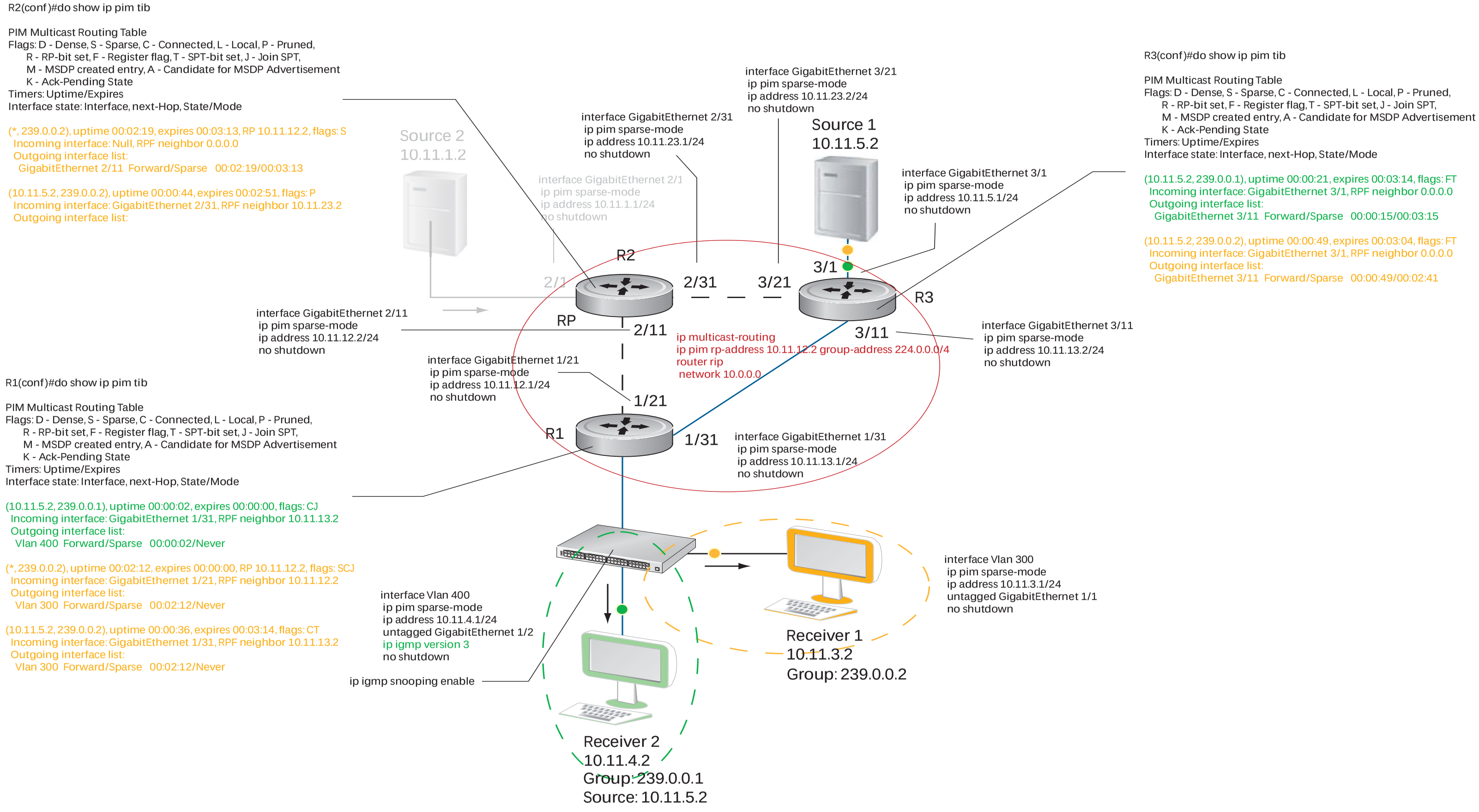
PIM-Source-Specific Mode (PIM-SSM) is a multicast protocol that forwards multicast traffic from a single source to a subnet. In the other versions of Protocol Independent Multicast (PIM), a receiver subscribes to a group only. The receiver receives traffic not just from the source in which it is interested but from all sources sending to that group. PIM-SSM requires that receivers specify the sources in which they are interested using IGMPv3 include messages to avoid receiving unwanted traffic.

PIM-SSM is more efficient than PIM-SM because it immediately creates shortest path trees (SPT) to the source rather than first using shared trees. PIM-SM requires a shared tree rooted at the RP because IGMPv2 receivers do not know about the source sending multicast data. Multicast traffic passes from the source to the receiver through the RP, until the receiver learns the source address, at which point it switches to the SPT. PIM-SSM uses IGMPv3. Since receivers subscribe to a source and group, the RP and shared tree is unnecessary, so only SPTs are used. On Dell Force10 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 should 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.

In Figure 26-1, Receiver 1 is an IGMPv2 host. The packets for group 239.0.0.2 travel to it first via the RP, then by the SPT. Receiver 2 is an IGMPv3 host. The packets for group 239.0.0.1 travel only via the STP.

Figure 26-1. PIM-SM with IGMPv2 versus PIM-SM with IGMPv3



```
R2(conf)#do 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,
     M - MSDP created entry, A - Candidate for MSDP Advertisement
     K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.0.0.2), uptime 00:02:19, expires 00:03:13, RP 10.11.12.2, flags: S
  Incoming interface: Null, RPF neighbor 0.0.0.0
  Outgoing interface list:
   GigabitEthernet 2/11  Forward/Sparse   00:02:19/00:03:13

(10.11.5.2, 239.0.0.2), uptime 00:00:44, expires 00:02:51, flags: P
  Incoming interface: GigabitEthernet 2/31, RPF neighbor 10.11.23.2
  Outgoing interface list:
```

```
interface GigabitEthernet 2/1
 ip pim sparse-mode
 ip address 10.11.1.1/24
 no shutdown
```

```
interface GigabitEthernet 2/31
 ip pim sparse-mode
 ip address 10.11.23.1/24
 no shutdown
```

```
interface GigabitEthernet 3/21
 ip pim sparse-mode
 ip address 10.11.23.2/24
 no shutdown
```

```
interface GigabitEthernet 3/1
 ip pim sparse-mode
 ip address 10.11.5.1/24
 no shutdown
```

```
R3(conf)#do 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,
     M - MSDP created entry, A - Candidate for MSDP Advertisement
     K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(10.11.5.2, 239.0.0.1), uptime 00:00:21, expires 00:03:14, flags: FT
  Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
  Outgoing interface list:
   GigabitEthernet 3/11  Forward/Sparse   00:00:15/00:03:15

(10.11.5.2, 239.0.0.2), uptime 00:00:49, expires 00:03:04, flags: FT
  Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
  Outgoing interface list:
   GigabitEthernet 3/11  Forward/Sparse   00:00:49/00:02:41
```

```
interface GigabitEthernet 2/11
 ip pim sparse-mode
 ip address 10.11.12.2/24
 no shutdown
```

```
ip multicast-routing
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
router rip
 network 10.0.0.0
```

```
interface GigabitEthernet 3/11
 ip pim sparse-mode
 ip address 10.11.13.2/24
 no shutdown
```

```
interface GigabitEthernet 1/21
 ip pim sparse-mode
 ip address 10.11.12.1/24
 no shutdown
```

```
R1(conf)#do 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,
     M - MSDP created entry, A - Candidate for MSDP Advertisement
     K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(10.11.5.2, 239.0.0.1), uptime 00:00:02, expires 00:00:00, flags: CJ
  Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
  Outgoing interface list:
   Vlan 400  Forward/Sparse   00:00:02/Never

(*, 239.0.0.2), uptime 00:02:12, expires 00:00:00, RP 10.11.12.2, flags: SCJ
  Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
  Outgoing interface list:
   Vlan 300  Forward/Sparse   00:02:12/Never

(10.11.5.2, 239.0.0.2), uptime 00:00:36, expires 00:03:14, flags: CT
  Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
  Outgoing interface list:
   Vlan 300  Forward/Sparse   00:02:12/Never
```

```
interface GigabitEthernet 1/31
 ip pim sparse-mode
 ip address 10.11.13.1/24
 no shutdown
```

```
interface Vlan 400
 ip pim sparse-mode
 ip address 10.11.4.1/24
 untagged GigabitEthernet 1/2
 ip igmp version 3
 no shutdown

ip igmp snooping enable
```

```
interface Vlan 300
 ip pim sparse-mode
 ip address 10.11.3.1/24
 untagged GigabitEthernet 1/1
 no shutdown
```

Implementation Information

- The Dell Force10 implementation of PIM-SSM is based on RFC 3569.
- C-Series supports a maximum of 31 PIM interfaces and 4K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors C-Series can have.
- S-Series 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 S-Series can have.
- FTOS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.

Important Points to Remember

- The default SSM range is 232/8 always. Applying an SSM range does not overwrite the default range. Both the default range and SSM range are effective even when the default range is not added to the SSM ACL.
- Extended ACLs cannot be used for configuring SSM range. Be sure to create the ACL first and then apply it to the SSM range.
- The default range is always supported, so range can never be smaller than the default.

Configure PIM-SM

Configuring PIM-SSM is a one-step process:

1. Configure PIM-SM. See page 497.
2. Enable PIM-SSM for a range of addresses. See page 507.

Related Configuration Tasks

- Use PIM-SSM with IGMP version 2 Hosts on page 508

Enable PIM-SSM

To enable PIM-SSM:

Step	Task	Command Syntax	Command Mode
1	Create an ACL that uses permit rules to specify what range of addresses should use SSM. You must at least include one rule, **permit 232.0.0.0/8**, which is the default range for PIM-SSM.	**ip access-list standard** *name*	CONFIGURATION

Step	Task	Command Syntax	Command Mode
2	Enter the command **ip pim ssm-range** and specify the ACL you created.	**ip pim ssm-range** *acl-name*	CONFIGURATION

Display address ranges in the PIM-SSM range using the command **show ip pim ssm-range** from EXEC Privilege mode.

Figure 26-2. Enabling PIM-SSM

```
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.

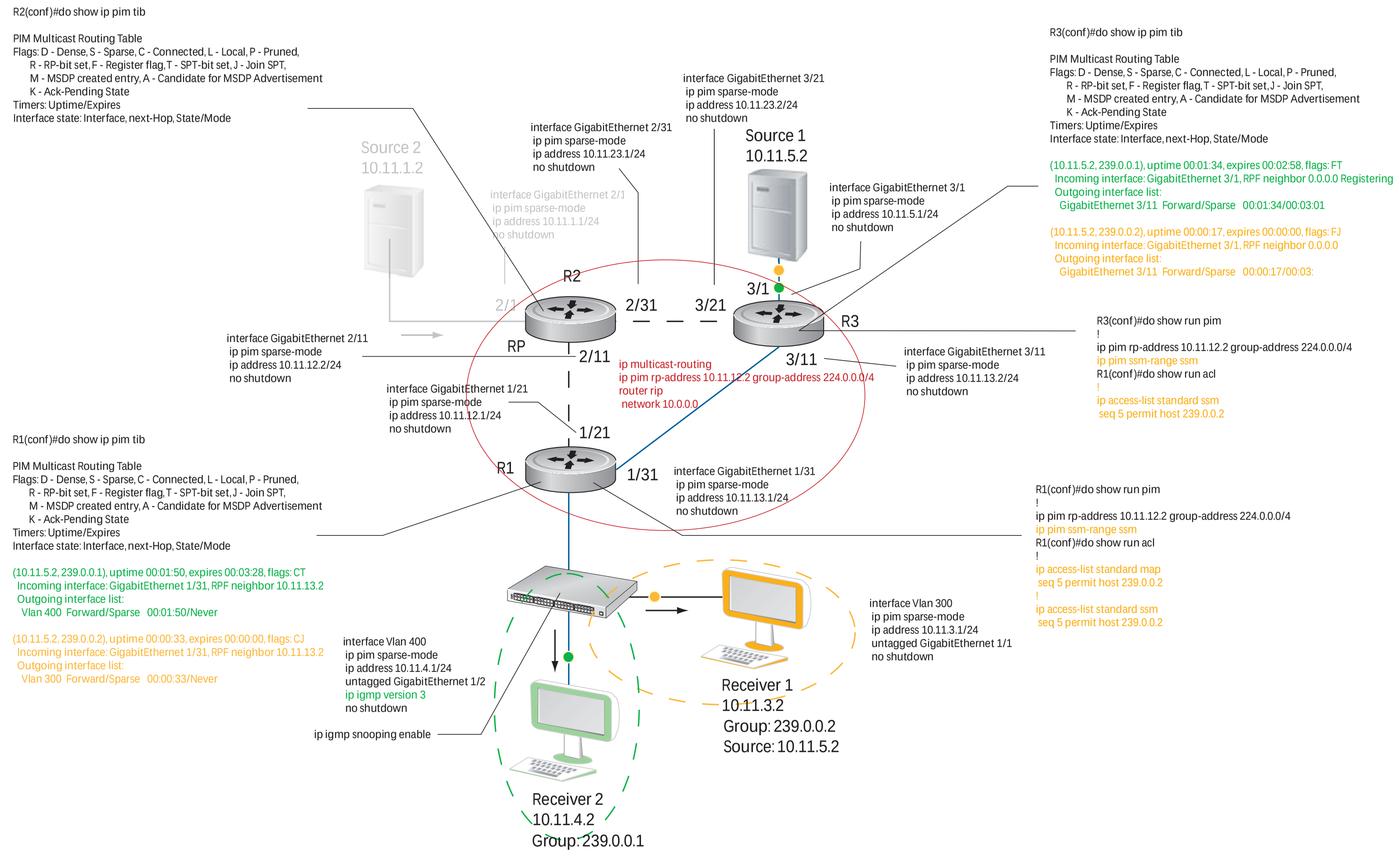
Translate (*,G) entries to (S,G) entries using the command **ip igmp ssm-map** *acl source* from CONFIGURATION mode. In a standard access list, specify the groups or the group ranges that you want to map to a source. Then, specify the multicast source.

- When a SSM map is in place and FTOS 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, FTOS removes the corresponding (S,G) states that it created and reestablishes 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, FTOS displays an error message. If you apply an extended ACL before you create it, FTOS accepts the configuration, but when the ACL is later defined, FTOS ignores the ACL and the stated mapping has no effect.

Display the source to which a group is mapped using the command **show ip igmp ssm-map** [*group*], as shown in Figure 26-4 on page 510. If 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, then the display is a list of groups currently in the IGMP group table that have a group-to-source mapping.

Display the list of sources mapped to a group currently in the IGMP group table using the command **show ip igmp groups** *group* **detail**, as shown in Figure 26-4 on page 510.

Figure 26-3. Using PIM-SM with IGMPv2 versus PIM-SSM with IGMPv2



```
R2(conf)#do 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,
    M - MSDP created entry, A - Candidate for MSDP Advertisement
    K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode
```

```
interface GigabitEthernet 2/31
 ip pim sparse-mode
 ip address 10.11.23.1/24
 no shutdown
```

```
interface GigabitEthernet 3/21
 ip pim sparse-mode
 ip address 10.11.23.2/24
 no shutdown
```

```
interface GigabitEthernet 2/1
 ip pim sparse-mode
 ip address 10.11.1.1/24
 no shutdown
```

```
interface GigabitEthernet 3/1
 ip pim sparse-mode
 ip address 10.11.5.1/24
 no shutdown
```

```
R3(conf)#do 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,
    M - MSDP created entry, A - Candidate for MSDP Advertisement
    K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(10.11.5.2, 239.0.0.1), uptime 00:01:34, expires 00:02:58, flags: FT
 Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0 Registering
 Outgoing interface list:
  GigabitEthernet 3/11  Forward/Sparse   00:01:34/00:03:01

(10.11.5.2, 239.0.0.2), uptime 00:00:17, expires 00:00:00, flags: FJ
 Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
 Outgoing interface list:
  GigabitEthernet 3/11  Forward/Sparse   00:00:17/00:03:
```

```
interface GigabitEthernet 2/11
 ip pim sparse-mode
 ip address 10.11.12.2/24
 no shutdown
```

```
ip multicast-routing
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
router rip
 network 10.0.0.0
```

```
interface GigabitEthernet 3/11
 ip pim sparse-mode
 ip address 10.11.13.2/24
 no shutdown
```

```
R3(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
```

```
interface GigabitEthernet 1/21
 ip pim sparse-mode
 ip address 10.11.12.1/24
 no shutdown
```

```
R1(conf)#do 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,
    M - MSDP created entry, A - Candidate for MSDP Advertisement
    K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(10.11.5.2, 239.0.0.1), uptime 00:01:50, expires 00:03:28, flags: CT
 Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
 Outgoing interface list:
  Vlan 400  Forward/Sparse   00:01:50/Never

(10.11.5.2, 239.0.0.2), uptime 00:00:33, expires 00:00:00, flags: CJ
 Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
 Outgoing interface list:
  Vlan 300  Forward/Sparse   00:00:33/Never
```

```
interface GigabitEthernet 1/31
 ip pim sparse-mode
 ip address 10.11.13.1/24
 no shutdown
```

```
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
```

```
interface Vlan 400
 ip pim sparse-mode
 ip address 10.11.4.1/24
 untagged GigabitEthernet 1/2
 ip igmp version 3
 no shutdown
```

```
interface Vlan 300
 ip pim sparse-mode
 ip address 10.11.3.1/24
 untagged GigabitEthernet 1/1
 no shutdown
```

```
ip igmp snooping enable
```

Figure 26-4. 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
```

27

Power over Ethernet

Power over Ethernet (PoE) is supported only on platforms: C S

This chapter contains the following major sections:

- Configuring Power over Ethernet on page 512
- Deploying VOIP on page 521
- Deploying VOIP on page 521

FTOS supports Power over Ethernet (PoE), as described by IEEE 802.3af . IEEE 802.3af specifies that a maximum of 15.4 Watts can be transmitted to Ethernet devices over the signal pairs of an unshielded twisted pair (UTP) cable. PoE is useful in networks with IP phones and wireless access points because separate power supplies for *powered devices* (PD) are not needed.

Table 27-2 describes the classes of powered devices defined by IEEE 802.3af:

Table 27-1. PoE Classes of Powered Devices

Class	Power Range (Watts)	Classification Current (mA)
0	0.44 to 12.95	< 5.0
1	0.44 to 3.84	10.5
2	3.84 to 6.49	18.5
3	6.49 to 12.95	28
4	Reserved	40

Note: FTOS treats Class 0, Class 3, and Class 4 powered devices the same. Class 4 is meant for IEEE802.3at compliant devices which require >12.95 Watts. Currently FTOS treats Class 4 devices as Class 3.

FTOS supports PoE on all copper ports on the C-Series and on the S25V and S50V models of the S-Series. The C-Series and S-Series transmit power to connected IEEE 802.3af-compliant powered devices through ports that have been configured to supply PoE. Those platforms also support the protocols LLDP and LLDP-MED, which help optimize power distribution to PoE devices. See Chapter 46, Link Layer Discovery Protocol, on page 861.

For the C-Series, FTOS requires that a minimum number of AC power supplies (PSU) be installed before PoE can be enabled, and some PSUs are reserved for PoE redundancy, as described in Table 27-2.

Note: The C-Series can provide PoE only through its AC power supplies.

Table 27-2. PoE Ports per Power Supply Unit in the C-Series*

Number of Power Supply Units	Max PoE Ports on C300	Max PoE Ports on C150
1	—	—
2	—	System Redundancy
3	System Redundancy	96
4	96	192
5	192	PoE Redundancy
6	288	PoE Redundancy
7	384	N/A
8	PoE redundancy	N/A

FTOS Behavior: Table 27-2 provides the maximum number of PoE ports per PSU, based on the assumption that each port deliver 15.4W. In many cases, the PD requires <15.4W. Typical IP Phones require only 3-10 Watts. So, if the ports are configured optimally, more PDs can be powered with fewer PSUs.

On the C-Series, though each PSU used for PoE (units 4-7 on the C300, and 3-4 on the C150) provides 1200 Watts of power, each actually makes available 1478.40 Watts for PoE. This is possible because each unit, once installed, borrows 278.40 Watts from the system redundancy power supply. If a power supply used for PoE is removed, PoE ports are shut down so that the system redundancy PSU retains is capability.

Note: The S25V and S50V models contain AC power supplies in order to support PoE. You can also add the external Dell Force10 470W Redundant Power Supply to power more PoE devices. For details, see Deploying VOIP on page 521 and see the **power budget** command in the Power Over Ethernet (PoE) chapter of the *FTOS Command Reference for the S-Series*.

Configuring Power over Ethernet

Configuring PoE is a two-step process:

1. Connect the IEEE 802.3af-compliant powered device directly to a port.
2. Enable PoE on the port, as described next.

Related Configuration Tasks

- Manage Ports using Power Priority and the Power Budget on page 515
- Monitor the Power Budget on page 518
- Manage Power Priorities on page 519
- Recover from a Failed Power Supply on page 520
- Deploying VOIP on page 521

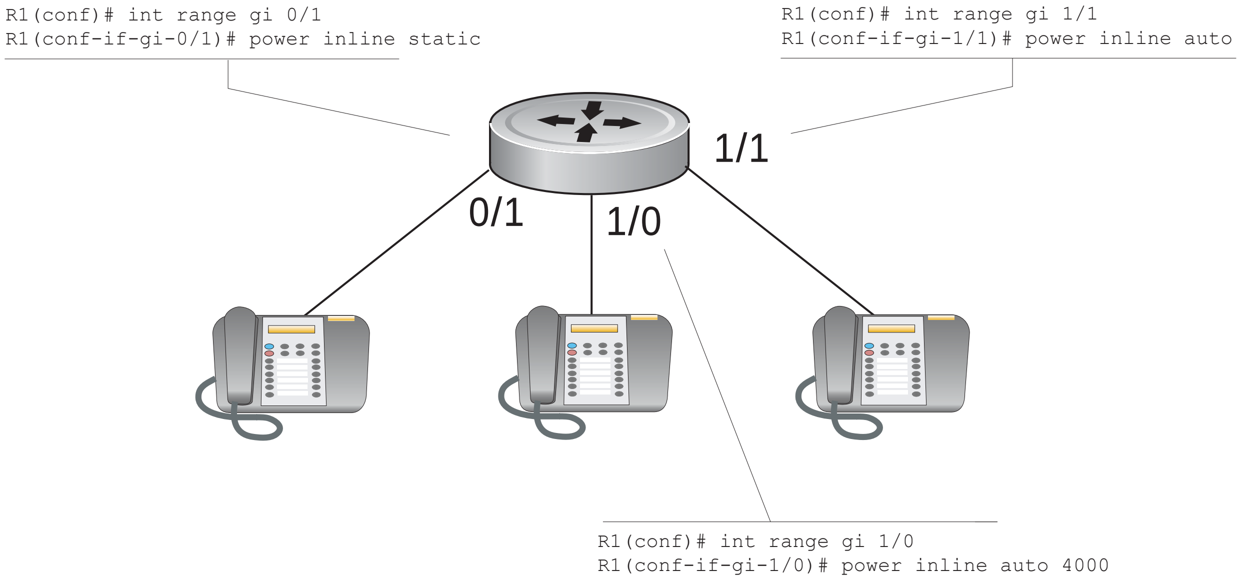
Enabling PoE on a Port

PoE is disabled by default. Enable PoE on a port from INTERFACE mode using the command **power inline** {**auto** [*max_milliwatts*] **|** **static** [*max_milliwatts*]}.

- The **power inline auto** command allows the port to determine the amount of power that a connected Class 1–4 powered device requires, and supply it. See Table 27-1 on page 511.
- The **power inline static** command without the qualifier guarantees 15.4W to the powered device.
- You can limit the maximum amount of power (in milliwatts) available to a powered device with the command **power inline auto** *max_milliwatts* or with **power inline static** *max_milliwatts*
- Disable PoE on a port using the **no power inline** command.

Ports configured with **power inline auto** have a lower priority for access to power than those configured with **power inline static**. As a second layer of priority setting, use the [**no**] **power inline priority** command. Use the **power inline static** *max_milliwatts* command to avoid allocating more power than necessary to a port because allocated power is made unavailable to other ports regardless of whether it is consumed. Typical IP phones use 3-10 Watts.

Figure 27-1. Enabling PoE



```
R1(conf)# int range gi 0/1
R1(conf-if-gi-0/1)# power inline static
```

```
R1(conf)# int range gi 1/1
R1(conf-if-gi-1/1)# power inline auto
```

```
R1(conf)# int range gi 1/0
R1(conf-if-gi-1/0)# power inline auto 4000
```

View the amount of power that a port is consuming using the **show power inline** command from EXEC privilege mode.

Figure 27-2. PoE Allocation Displayed with show power inline Command (example from C-Series)

```
Force10#show power inline

Interface   Admin   Oper   Inline Power   Inline Power    Class        User
                           Allocated      Consumed                    Priority
                           (Watts)         (Watts)
--------- -----     ----   ------------   ------------    -----      ----------
Gi 0/40     auto      on        7.00         3.20             2            Low
Gi 0/41     auto      on        0.00         0.00        NO_DEVICE         High
Gi 0/40     auto      on        7.00         3.20             2            Low
```

Table 27-3 describes the fields that the **show power inline** command displays:

Table 27-3. show power inline Field Description

Field	Port Number
Interface	Displays all PoE-enabled ports.
Admin	Displays the administrative mode of the interface: • *auto* indicates that power is supplied according to the requirements of the powered device. • *static* indicates that the maximum configured amount of power is supplied to the powered device.
Oper	Displays the system PoE operational status for the port. This signifies the readiness of a port to supply power. It is “on” when the system is able to supply sufficient power to port. It will be “off” when there is no power available for port. This column does not reflect the PDs operational status.
Inline Power Allocated	Displays the amount of power that can be allocated to a port if sufficient power is available. When sufficient power is not available for particular port based on the priority logic, then FTOS sets the PoE Oper status for the port is to “off” and inline power is not supplied to that port. If you insert an additional power supply, or when the priority of the port is increased, then the system supplies the allocated power to the port.
Inline Power Consumed	Displays the amount of power that a powered device is consuming.
Class	Displays the type of powered device: Class 0, Class 1, Class 2, Class 3, or Class 4. Displays NO_DEVICE if no PD is connected.

View the total power consumption of the chassis using the **show power detail** command from EXEC privilege mode.

Figure 27-3. PoE Consumed, Allocated, and Available with show power detail Command

```
R1#show power detail
Catalog            slot    Logic Power      Inline Power      Inline Power
Name               Id      Consumed         Allocated         Consumed
                             (Watts)          (Watts)           (Watts)
----------------------------------------------------------------------------
EX4PB              0         200              0.00              0.00
RPM                0         200              0.00              0.00
E48VB              7         150              35.8              7.14
CC-C300-FAN        -         100              0.00              0.00

Total Inline Power Available: 1478.40 W
Total Inline Power Used      :   35.8 W
Total Inline Power Remaining: 1442.6 W
```

Table 27-4 describes the fields that the **show power detail** command displays.

Table 27-4. show power detail Field Description

Field	Port Number
Catalog Name	Displays the Dell Force10 catalog number of the line card, RPM, and fan tray.
Slot ID	Displays the slot number in which the component in installed.
Logic Power Consumed	Displays the total amount of power that the chassis component is consuming for basic functionality.
Inline Power Allocated	Displays the sum of inline power allocated for ports in a line card.
Inline Power Consumed	Displays the sum of inline power consumed by all ports in a PoE line card.
Total Inline Power Available	Displays the total inline power that is available in the system for supply to PoE ports.
Total Inline Power Used	Displays the sum of inline power allocated across all PoE line cards.
Total Inline Power Remaining	Displays the difference (“Total Inline Power Available” minus “Total Inline Power Used”).

Manage Ports using Power Priority and the Power Budget

The allocation and return of power on ports depends on the total inline power available in the system and the power priority calculation.

Determine the Power Priority for a Port

FTOS uses a sophisticated port prioritization algorithm for determining which ports receive PoE so that PoE ports are powered up/down deterministically.

FTOS uses the following four parameters, in order, for defining the power priority for a port:

1. the **power-inline** mode: **static** or **auto**,
2. the **power-inline priority** configuration,
3. the LLDP-MED priority sent by the PD in the Extended Power-via-MDI TLV,
4. and the port's slot and port number.

FTOS maintains a sorted list of PoE ports based these four parameters. Static ports have a higher weight than auto mode ports, so all static ports always stay on top of all auto ports regardless of the other 3 parameters. Within the set of static ports, FTOS attempts to order them based on the second parameter **power-inline priority**, the default of which is "Low". If FTOS finds multiple ports with the same **power-inline priority**, it breaks the tie using the third parameter, the LLDP-MED Priority advertised by the PD, which like **power-inline priority** could be "Critical," "High," or "Low". After this, if FTOS still finds a tie, priority is based on the fourth parameter which is the ports position in the chassis; there cannot be a tie based on this parameter.

This sorted list is dynamically updated by FTOS when:

- a user changes the **power-inline** mode or priority
- the PD advertises a different LLDP-MED priority
- the PD is connected or disconnected

FTOS always uses this sorted list of ports for allocation. When an additional PSU is added, additional ports are powered based on this list, and PSU is removed, this same list is used to remove power from the lowest priority ports.

power-inline mode

FTOS allows ports to be configured in one of two modes: **auto** and **static**.

auto: Ports configured for auto mode manage the power allocation by themselves. There is no prior reservation of power made on these ports. When no PD is connected on this port, the power allocated is zero. Once a PD is connected, FTOS detects its PoE class dynamically and the maximum power for its class is allocated to the port. The PD then boots using this allocated power. After bootup, if the PD is LLDP-MED capable, it might send in Extended Power via MDI TLV to the system. In this case, the Dell Force10 switch revises the power allocation to the value that the PD requests via LLDP-MED. The advertised Power Requirement from the PD could be less than or greater than the currently allocated value.

Ports configured for auto mode with the *max_milliwatts* option allocate power the same way, but the allocation never exceeds the specified maximum. If *max_milliwatts* is greater than the PoE class maximum the system allocates only the class maximum. Note that if a PD has class maximum that is greater than *max_milliwatts*, the system allocates no power, and the PD does not power up.

static: Ports configured in static mode reserve a fixed power allocation whether a device is connected or not. By default 15.4W is allocated, but this is user-configurable with the *max_milliwatts* option. No dynamic PoE class detection is performed on static ports, and Extended Power via MDI TLVs have no effect.

Extended Power-via-MDI TLV

The PD sends three pieces of information in the LLDP-MED Extended Power-via-MDI TLV:

1. Power Requirement: FTOS honors this and uses it for power allocation.
2. Power Priority—Critical, High, or Low: FTOS honors this information and uses it for power priority calculation.
3. External Power Source: FTOS does not use this information.

Determine the Affect of a Port on the Power Budget

The PoE power budget is affected differently depending on how PoE is enabled and whether a device is connected:

1. When you configure a port with **power inline auto** *without* the *max_milliwatts* power limit option, power is only allocated after you connect a device to the port.
    - When you connect a device, the maximum power for the device class is allocated if there is sufficient power in the budget. See Table 27-1 on page 511.
    - If there is not enough power in the budget, the configuration is maintained and the port waits for power to become available.
    - If the device advertises its power requirement through LLDP-MED, then FTOS allocates the required amount and returns the remaining amount to the budget.

Note: LLDP-MED TLVs are only honored if the port is configured with **power inline auto** (with or without the *max_milliwatts* option).

2. When you configure a port with **power inline auto** *with* the power limit option *max_milliwatts*, power is only allocated after you connect a device to the port.
    - If the maximum power for the device class is *less* than the power limit you specified, FTOS allocates the required amount and returns the remaining amount to the budget.
    - If there is not enough power in the budget, the configuration is maintained and the port waits for power to become available.
    - If the maximum power for the device class is *more than* than the power limit you specified, FTOS does not allocate any power.

Note: When a port is configured with **power inline auto** (with or without the *max_milliwatts* option) and the PoE device is disconnected, the allocated power is returned to the power budget.

3. When you configure a port with **power inline static** *without* the power limit option (*max_milliwatts*), FTOS allocates 15.4W (subject to availability and priority) to the port whether or not a device is connected.
4. When you configure a port with **power inline static** *with* the power limit option (*max_milliwatts*), FTOS allocates the specified number of Watts.

- If there is not enough power in the budget, the configuration is maintained and port waits for power to become available.
- If the maximum power for the device class is *more than* than the power limit you specified, FTOS does not allocate any power.

Monitor the Power Budget

The power budget is the amount of power available from the installed PSUs minus the power required to operate the chassis. Use the **show power inline** (Figure 27-2 on page 514) and **show power detail** (Figure 27-3 on page 515) commands to help you determine if power is available for additional PoE ports (1478.40 Watts are supplied per C-Series PSU; max of 790W on S-Series with load-sharing external DC PSU).

Enabling PoE on more ports than is supported by the power budget produces one of these results:

- If the newly PoE-enabled port has a lower priority, then the command is accepted, but power is not allocated to the port. In this case, the following message is displayed.

Message 1 Insufficient Power to Enable PoE

```
%Warning: Insufficient power to enable. POE oper-status set to OFF for port <linecard/
portnumber>
```

- If the newly PoE-enabled port has a higher priority, then the CLI is accepted, and power is terminated on the lowest priority port in the chassis. If another power supply is added to the system at a later time, both ports receive power.
- If all of the lower priority ports combined cannot meet the power requirements of the newly enabled port, then the CLI is accepted, but power on the lower priority ports is not terminated, and no power is supplied to the port.

The second result in this scenario is true even if a powered device is not connected to the port. Power can be allocated to a port, thus subtracting it from the power budget and making it unavailable to other ports, but that power does not have to be consumed.

Manage Power Priorities

PoE-enabled ports have power access priorities based first on their configuration and then by line card and port number. The default prioritization is presented in Table 27-5.

Note: For S-Series, where Table 27-5 refers to "line cards with the lowest slot number", substitute "S-Series stack members with the lowest unit ID".)

Table 27-5. PoE Ports Priorities

Configuration	Port Number	Priority
Ports configured with **power inline static**	Ports with the lowest port numbers in line cards with the lowest slot number	1
	Ports with the lowest port numbers	2
Ports configured with **power inline auto**	Ports with the lowest port numbers in line cards with the lowest slot number	3
	Ports with the lowest port numbers	4

You can augment the default prioritization using the command **[no] power inline priority** {**critical** | **high** | **low**}, where **critical** is the highest priority, and **low** is the lowest. FTOS ignores any LLDP-MED priority on this port if you configure a priority with this command. If you do not configure a port priority with this command, FTOS honors any LLDP-MED priority.

In general, priority is assigned in this order:

1. **power inline** [**static | auto**] setting: **power inline static** ports have a higher priority than **power inline auto** ports
2. **power inline priority** {**critical | high | low**} setting or LLDP-MED TLV, if **power inline priority** is not configured
3. slot ID
4. port ID

Recover from a Failed Power Supply

If ports are PoE-enabled, and a PSU fails, power might be terminated on some ports to compensate for the power loss. This does not affect PoE individual port configurations.

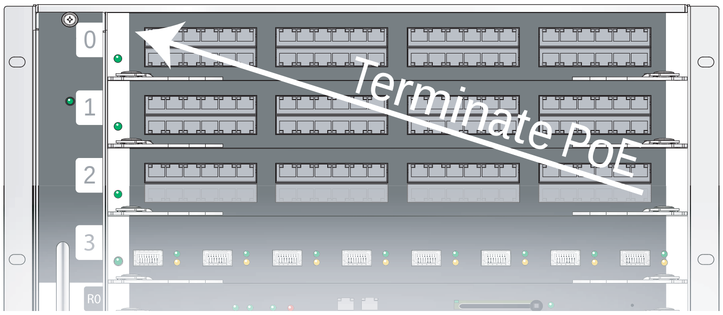
For C-Series, use the **show power supply** command to display PSU status (Figure 27-4).
For S-Series, see the Power over Ethernet (PoE) chapter in the *FTOS Command Reference for the S-Series* for an example of the output of the **show power inline** output and its field descriptions.

Figure 27-4. show power supply Command Example

```
R1#show power supply

Power         Model
Supply        Number            Type      Status
------------------------------------------------
PS0           --                --         Absent
PS1           CC-C300-PWR-AC    AC          Active
PS2           CC-C300-PWR-AC    AC         Fail
PS3           CC-C300-PWR-AC    AC         Remote Off
PS4           --                --         Absent
PS5           --                --         Absent
PS6           --                --         Absent
PS7           --                --         Absent
```

If power must be terminated for some ports, the order in which ports are affected is based on priority. Ports with the lowest priority are terminated first (see Manage Power Priorities on page 519).

Figure 27-5. Order of PoE Termination

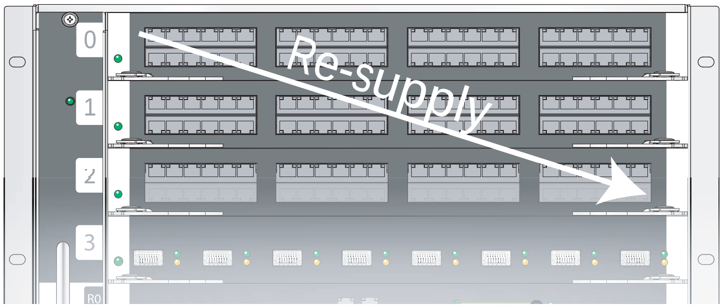


For the configuration in Figure 27-2:

- Power for ports 7/1 and 7/2 is terminated first because it is configured with **inline power auto**.
- Power for port 7/2 is terminated before PoE for port 7/1 because port 7/1 has a lower port number.
- Power for port 7/0 is terminated last because it is configured with **inline power static**.

When a failed PSU is replaced and there is sufficient power for PoE, power is automatically re-supplied for previously configured PoE ports, and power is supplied first to ports with the highest priority.

Figure 27-6. Order of PoE Re-Supply



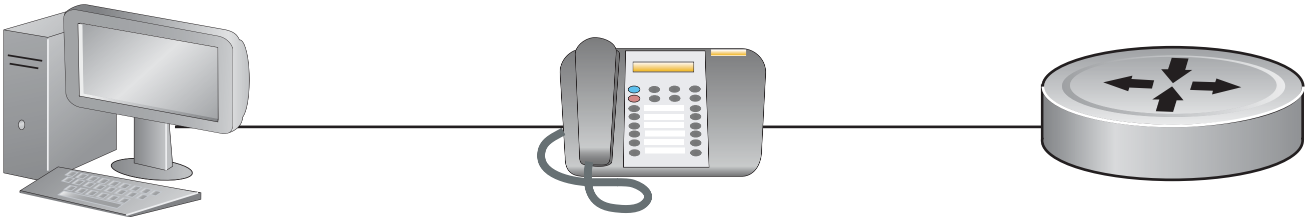
Deploying VOIP

VoIP phones on the market today follow the same basic boot and operations process:

1. Wait for an LLDP from the Ethernet switch.
2. Obtain an IP address from a DHCP server.
3. Send an LLDP-MED frame to the switch.
4. Wait for an LLDP-MED frame from the switch and read the Network Policy TLV to get the VLAN ID, Layer 2 Priority, and DSCP value.
5. Download applications and software from the call manager.
6. After configuration, send voice packets as tagged frames and data packets as untagged frames.

Figure 27-7 shows a basic configuration for a deployment in which the end workstation plugs into an IP phone for its Ethernet connection.

Figure 27-7. Office VOIP Deployment



Create VLANs for an Office VOIP Deployment

The phone requires one tagged VLAN for VOIP service and one untagged VLAN for PC data, as shown in Figure 27-7. You may configure voice signaling on the voice VLAN, but some implementations might need an additional tagged VLAN for this traffic; Figure 27-8 adds an additional tagged VLAN for voice signaling. The example is from a C-Series, but an S-Series would be configured in the same way.

Figure 27-8. Creating VLANs for an Office VOIP Deployment

```
Force10#show running-config interface configured
!
interface GigabitEthernet 6/0
 no ip address
 no shutdown
!
interface GigabitEthernet 6/10
 no ip address
 portmode hybrid
 switchport!
 power inline auto
 no shutdown
!
interface Vlan 100
 description "Data VLAN"
 no ip address
 untagged GigabitEthernet 6/10-11,22-23,46-47
 shutdown
!
interface Vlan 200
 description "Voice VLAN"
 no ip address
 tagged GigabitEthernet 6/10-11,22-23,46-47
 shutdown
!
interface Vlan 300
 description "Voice Signaling VLAN"
 no ip address
 tagged GigabitEthernet 6/10-11,22-23,46-47
 shutdown
```

Configure LLDP-MED for an Office VOIP Deployment

VOIP deployments may optionally use LLDP-MED. LLDP-MED advertises VLAN, dot1P, and DSCP configurations on the switch so that you do not need to manually configure every phone with this information. See Chapter 21, Link Layer Discovery Protocol. Based on the configuration in Figure 27-9, the phone will initiate a DHCP request on the advertised voice VLAN, VLAN 200.

Figure 27-9. LLDP Configuration for Office VOIP Deployment

```
Force10#show running-config lldp
protocol lldp
 advertise med
 advertise med voice 200 6 46
 advertise med voice-signaling 300 5 28
 no disable
Force10#show lldp neighbors
 Loc PortID        Rem Chassis Id              Rem Port Id
 -------------------------------------------------------------------------

 Gi 6/10           0.0.0.0                     001B0CDBA109:P1
 Gi 6/11           0.0.0.0                     001AA2197992:P1
 Gi 6/22           0.0.0.0                     08:00:0f:22:7f:83
 Gi 6/23           0.0.0.0                     08:00:0f:23:de:a9
```

Configure Quality of Service for an Office VOIP Deployment

There are multiple ways you can use QoS to map ingress phone and PC traffic so that you can give them each a different quality of service. See Chapter 31, Quality of Service.

Honor the incoming DSCP value

On both the C-Series or S-Series, if you know traffic originating from the phone is tagged with the DSCP value of 46 (EF), you might make the associated queue a strict priority queue, as shown in Figure 27-10; on the C-Series and S-Series, FTOS maps DSCP 46 to queue 2 (see Table 31-5 on page 573 in the QoS chapter.)

Figure 27-10. Honoring the DSCP Value on Incoming Voice Data

```
Force10#sh run policy-map-input
!
policy-map-input HonorDSCP
 trust diffserv
Force10#sh run int gigabitethernet 6/11
!
interface GigabitEthernet 6/11
 description "IP Phone X"
 no ip address
 portmode hybrid
 switchport
 service-policy input HonorDSCP
 power inline auto
 no shutdown
Force10#sh run | grep strict-priority
strict-priority unicast 2
```

Honor the incoming dot1p value

On the C-Series, if you know traffic originating from the phone is tagged with a dot1p value of 5, you might make the associated queue a strict priority queue, as shown in Figure 27-11; on the C-Series, FTOS maps dot1p priority 5 to queue 2.

Figure 27-11. Honoring the Dot1P Value on Incoming Voice Traffic

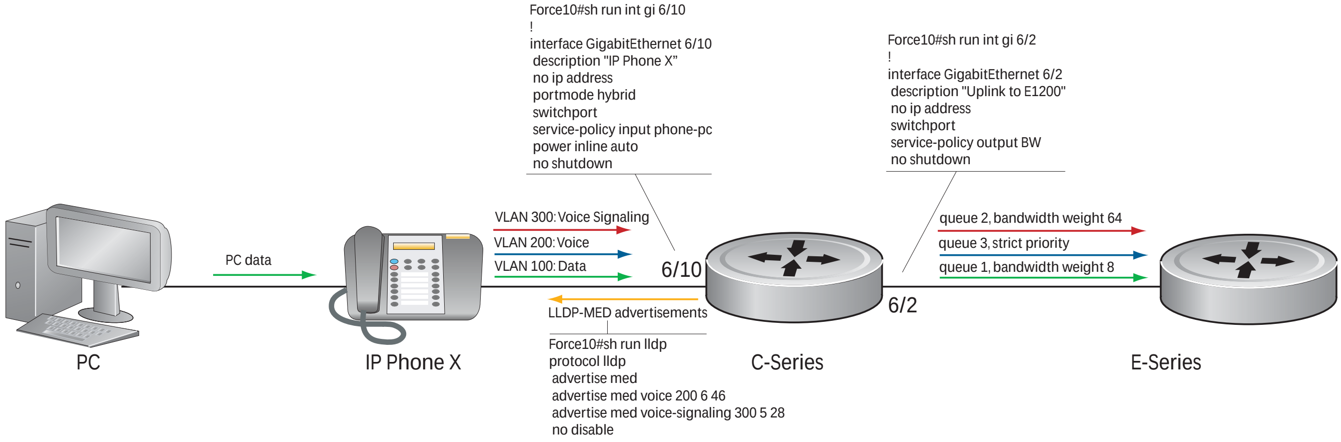
```
Force10#sh run int gi 6/10
!
interface GigabitEthernet 6/10
 description "IP Phone X"
 no ip address
 portmode hybrid
 switchport
 service-class dynamic dot1p
 power inline auto
 no shutdown
Force10#sh run | grep strict-priority
strict-priority unicast 2
```

Classifying VOIP traffic and applying QoS policies

Avoid congestion and give precedence to voice and signaling traffic by classifying traffic based on subnet and using strict priority and bandwidth weights on egress, as outlined in the steps below.

Figure 27-12 depicts the topology and shows the configuration for a C-Series. The steps are the same on an S-Series. Figure 27-13 on page 525 is a screenshot showing some of the steps and the resulting running-config.

Figure 27-12. Classifying VOIP Traffic and Applying QoS Policies for an Office VOIP Deployment



```
Force10#sh run int gi 6/10
!
interface GigabitEthernet 6/10
 description "IP Phone X"
 no ip address
 portmode hybrid
 switchport
 service-policy input phone-pc
 power inline auto
 no shutdown
```

```
Force10#sh run int gi 6/2
!
interface GigabitEthernet 6/2
 description "Uplink to E1200"
 no ip address
 switchport
 service-policy output BW
 no shutdown
```

```
Force10#sh run lldp
protocol lldp
 advertise med
 advertise med voice 200 6 46
 advertise med voice-signaling 300 5 28
 no disable
```

Step	Task	Command	Command Mode
1	Create three standard or extended access-lists, one each for voice, voice signaling, and PC data, and place each in its own match-any class-map.	**ip access-list** **class-map match-any**	CONFIGURATION CLASS-MAP
2	Create an input policy-map containing all three class-maps, and assign each class-map a different service queue.	**policy-map-input** **service-queue**	CONFIGURATION POLICY-MAP-IN
3	Create two input QoS policies, one each for PC data and voice signaling. Assign a different bandwidth weight to each policy.	**qos-policy-out** bandwidth-weight	CONFIGURATION QOS-POLICY-IN
4	Create an output policy map containing both QoS policies, and assign them to different service queues.	**policy-map-out** **service-queue**	CONFIGURATION POLICY-MAP-OUT
5	Assign a strict priority to unicast traffic in queue 3.	strict-priority	CONFIGURATION
6	Apply the input policy map you created in Step 2 to the interface connected to the phone, and apply the output policy map you created in Step 4 to the interface connected your desired next-hop router.	service-policy	INTERFACE

Figure 27-13 on page 525 is a screenshot showing some of the steps, above, and the resulting running-config.

Figure 27-13. Classifying VOIP Traffic and Applying QoS Policies for an Office VOIP Deployment

```
Force10#sh run acl
!
ip access-list extended pc-subnet
 seq 5 permit ip 201.1.1.0/24 any
!
ip access-list extended phone-signalling
 seq 5 permit ip 192.1.1.0/24 host 192.1.1.1
!
ip access-list extended phone-subnet
 seq 5 permit ip 192.1.1.0/24 any
Force10#sh run class-map
!
class-map match-any pc-subnet
 match ip access-group pc-subnet
!
class-map match-any phone-signalling
 match ip access-group phone-signalling
!
class-map match-any phone-subnet
 match ip access-group phone-subnet
Force10#sh run policy-map-input
!
policy-map-input phone-pc
 service-queue 1 class-map pc-subnet
 service-queue 2 class-map phone-signalling
 service-queue 3 class-map phone-subnet
Force10#sh run qos-policy-output
!
qos-policy-output data
 bandwidth-weight 8
!
qos-policy-output signalling
 bandwidth-weight 64
Force10#sh run policy-map-output
!
policy-map-output BW
 service-queue 1 qos-policy data
 service-queue 2 qos-policy signalling
Force10#sh run | grep strict-p
strict-priority unicast 3
Force10#sh run int gi 6/10
!
interface GigabitEthernet 6/10
 description "IP Phone X”
 no ip address
 portmode hybrid
 switchport
 service-policy input phone-pc
 power inline auto
 no shutdown
Force10#sh run int gi 6/2
!
interface GigabitEthernet 6/2
 description "Uplink to E1200"
 no ip address
 switchport
 service-policy output BW
 no shutdown
```

28

Port Monitoring

Port Monitoring is supported on platforms: C E S

Port Monitoring is a feature that copies all incoming or outgoing packets on one port and forwards (mirrors) them to to another port. The source port is the monitored port (MD) and the destination port is the monitoring port (MG). Port Monitoring functionality is different between platforms, but the behavior is the same, with highlighted exceptions.

This chapter is divided into the following sections:

Important Points to Remember

- Port Monitoring is supported on physical ports only; VLAN and port-channel interfaces do not support port monitoring.
- The Monitored (source, “MD”) and Monitoring ports (destination, “MG”) must be on the same switch.
- In general, a monitoring port should have **no ip address** and **no shutdown** as the only configuration; FTOS permits a limited set of commands for monitoring ports; display them using the command **?**. A monitoring port also 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). The following error is displayed if you try to assign a monitored port to more than one monitoring port.

```
Force10(conf)#mon ses 1
Force10(conf-mon-sess-1)#$gig 0/0 destination gig 0/60 direction both
Force10(conf-mon-sess-1)#do show mon ses
     SessionID       Source       Destination     Direction      Mode         Type
     ---------       ------       -----------     ---------      ----         ----
             1       Gi 0/0       Gi 0/60         both           interface    Port-based
Force10(conf-mon-sess-1)#mon ses 2
Force10(conf-mon-sess-2)#source gig 0/0 destination gig 0/61 direction both
% Error: MD port is already being monitored.
```

- The C-Series and S-Series may only have four destination ports per port-pipe. There is no limitation on the total number of monitoring sessions.

Table 28-1 lists the maximum number of monitoring sessions per system. For the C-Series and S-Series, the total number of sessions is derived by consuming a unique destination port in each session, in each port-pipe.

Table 28-1. Maximum Number of Monitoring Sessions per System

System	Maximum Sessions	System	Maximum Sessions
C150	∞ (Note)	E1200/E1200i (TeraScale)	28
C300	∞ (Note)	E1200i (ExaScale)	∞
S50V, S50N	∞ (Note)	E600/E600i (TeraScale)	14
S25P	∞ (Note)	E600i (ExaScale)	∞
		E300	6

Note: On the C-Series and S-Series, there is no limit to the number of monitoring sessions per system, provided that there are only 4 destination ports per port-pipe. If each monitoring session has a unique destination port, then the maximum number of session is 4 per port-pipe.

Port Monitoring on E-Series

Both the E-Series TeraScale and E-Series ExaScale support the following.

- FTOS supports one destination (MG) port per monitoring session. The same destination port (MG) can be used in another monitoring session.
- One destination (MG) port can monitor up to 28 source (MD) ports.
- A port cannot be defined as both a source (MD) and a destination (MG) port (Message 1).

Message 1 Cannot define source (MD) and destination (MG) on same port

```
% Error: MD port is already being monitored.
```

E-Series TeraScale

The E-Series TeraScale system supports 1 monitoring session per port-pipe. E-Series TeraScale supports a maximum of 28 port pipes.

On the E-Series TeraScale, FTOS supports a single source-destination statement in a monitor session (Message 2). E-Series TeraScale supports only one source and one destination port per port-pipe (Message 3). Therefore, the E-Series TeraScale supports as many monitoring sessions as there are port-pipes in the system.

Message 2 Multiple Source-Destination Statements Error Message on E-Series TeraScale

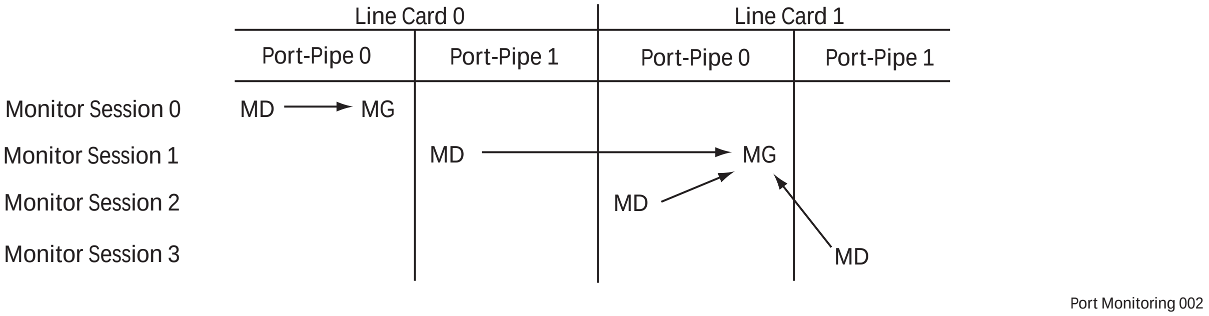
```
% Error: Remove existing monitor configuration.
```

Message 3 One Source/Destination Port per Port-pipe Error Message on E-Series TeraScale

```
% Error: Some port from this port pipe is already configured as MD.
% Error: Some port from this port pipe is already configured as MG.
```

Figure 28-1 illustrates a possible port monitoring configuration on the E-Series.

Figure 28-1. Port Monitoring Configurations on the E-Series



E-Series ExaScale

FTOS on E-Series ExaScale supports a single destination (MG) port monitoring multiple multiple source (MD) ports in one monitor session. One monitor session can have only one destination (MG) port. The same destination (MG) port can be uses with multiple monitoring sessions.

There is no restriction on the number of source (MD) or destination (MG) ports on the chassis because there is no port-pipe restriction on the E-Series ExaScale system.

Port Monitoring on C-Series and S-Series

The C-Series and S-Series support multiple source-destination statements in a monitor session, but there may only be one destination port in a monitoring session (Message 4).

Message 4 One Destination Port in a Monitoring Session Error Message on C-Series and S-Series

```
% Error: Only one MG port is allowed in a session.
```

The number of source ports FTOS allows within a port-pipe is equal to the number of physical ports in the port-pipe (n). However, n number of ports may only have four different destination ports (Message 5).

Figure 28-2. Number of Monitoring Ports on the C-Series and S-Series

```
Force10#show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
Force10(conf)#mon ses 300
Force10(conf-mon-sess-300)#source gig 0/17 destination gig 0/4 direction tx
% Error: Exceeding max MG ports for this MD port pipe.
Force10(conf-mon-sess-300)#
Force10(conf-mon-sess-300)#source gig 0/17 destination gig 0/1 direction tx
Force10(conf-mon-sess-300)#do show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
           300       Gi 0/17      Gi 0/1          tx             interface   Port-based
Force10(conf-mon-sess-300)#
```

In Figure 28-2, ports 0/13, 0/14, 0/15, and 0/16 all belong to the same port-pipe. They are pointing to four different destinations (0/1, 0/2, 0/3, and 0/37). Now it is not possible for another source port from the same port-pipe (for example, 0/17) to point to another new destination (for example, 0/4). If you attempt to configure another destination, Message 5 appears. However, you can configure another monitoring session that uses one of previously used destination ports, as shown in Figure 28-3.

Figure 28-3. Number of Monitoring Ports on the C-Series and S-Series

```
Force10(conf)#mon ses 300
Force10(conf-mon-sess-300)#source gig 0/17 destination gig 0/4 direction tx
% Error: Exceeding max MG ports for this MD port pipe.
Force10(conf-mon-sess-300)#
Force10(conf-mon-sess-300)#source gig 0/17 destination gig 0/1 direction tx
Force10(conf-mon-sess-300)#do show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
           300       Gi 0/17      Gi 0/1          tx             interface   Port-based
```

In Figure 28-4, 0/25 and 0/26 belong to Port-pipe 1. This port-pipe again has the same restriction of only four destination ports, new or used.

Figure 28-4. Number of Monitoring Ports on the C-Series and S-Series

```
Force10(conf-mon-sess-300)#do show mon session
     SessionID        Source        Destination     Direction      Mode        Type
     ---------        ------        -----------     ---------      ----        ----
             0        Gi 0/13       Gi 0/1          rx             interface   Port-based
            10        Gi 0/14       Gi 0/2          rx             interface   Port-based
            20        Gi 0/15       Gi 0/3          rx             interface   Port-based
            30        Gi 0/16       Gi 0/37         rx             interface   Port-based
           100        Gi 0/25       Gi 0/38         tx             interface   Port-based
           110        Gi 0/26       Gi 0/39         tx             interface   Port-based
           300        Gi 0/17       Gi 0/1          tx             interface   Port-based
Force10(conf-mon-sess-300)#
```

A source port may only be monitored by one destination port (Message 6), but a destination port may monitor more than one source port. Given these parameters, Figure 28-1 illustrates conceptually the possible port monitoring configurations on the C-Series and S-Series.

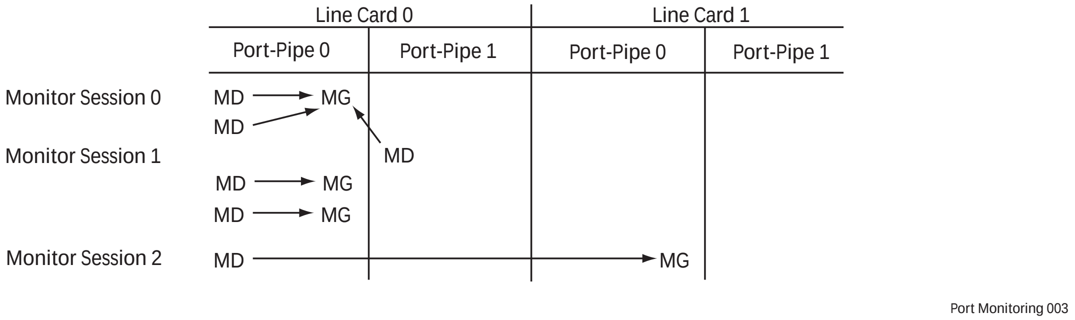
Message 5 One Destination Port in a Monitoring Session Error Message on C-Series and S-Series

```
% Error: Exceeding max MG ports for this MD port pipe.
```

Message 6 One Destination Port per Source Port Error Message

```
% Error: MD port is already being monitored.
```

Figure 28-5. Port Monitoring Configurations on the C-Series and S-Series



FTOS Behavior: On the C-Series and S-Series, 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 gig 6/0 destination gig 6/1 direction tx*, if the MD port gigabitethernet 6/0 is an untagged member of any VLAN, all monitored frames that the MG port gigabitethernet 6/1 receives are tagged with the VLAN ID of the MD port. Similarly, if BPDUs are transmitted, the MG port receives them tagged with the VLAN ID 4095. This behavior might result in a difference between the number of egress packets on the MD port and monitored packets on the MG port.

FTOS Behavior: The C-Series and S-Series continue to mirror outgoing traffic even after an MD participating in Spanning Tree Protocol transitions from the forwarding to blocking.

Configuring Port Monitoring

To configure port monitoring:

Step	Task	Command Syntax	Command Mode
1	Verify that the intended monitoring port has no configuration other than **no shutdown**, as shown in Figure 28-6.	**show interface**	EXEC Privilege
2	Create a monitoring session using the command monitor session from CONFIGURATION mode, as shown in Figure 28-6.	**monitor session**	CONFIGURATION
3	Specify the source and destination port and direction of traffic, as shown in Figure 28-6.	**source**	MONITOR SESSION

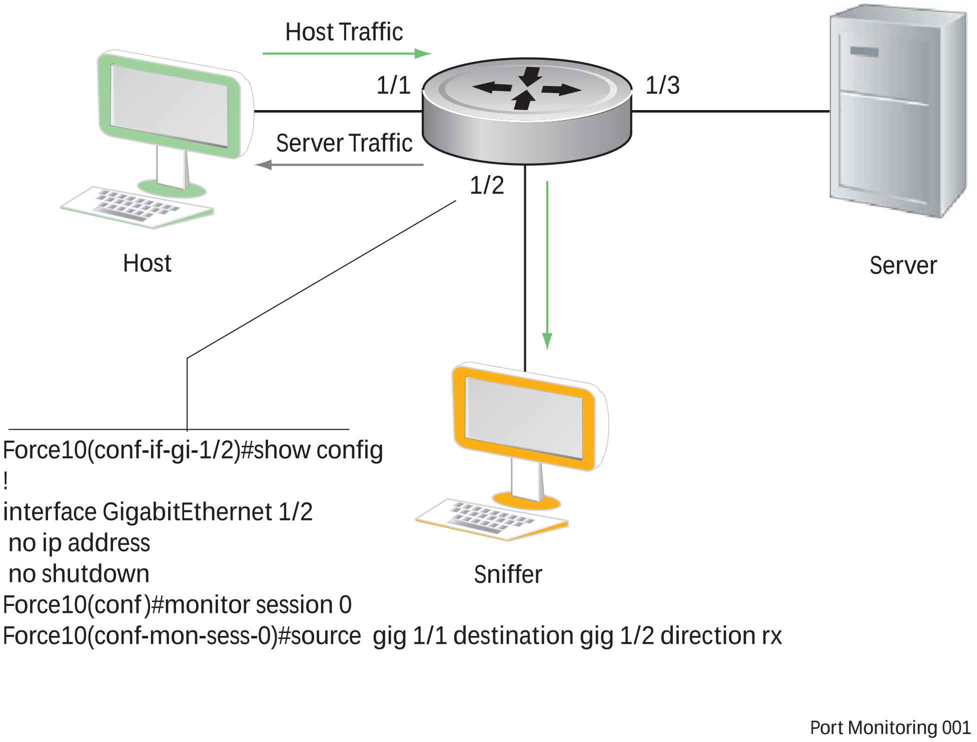
Display monitor sessions using the command **show monitor session** from EXEC Privilege mode, as shown in Figure 28-6.

Figure 28-6. Configuring Port-based Monitoring

```
Force10(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
 no ip address
 no shutdown
Force10(conf-if-gi-1/2)#exit
Force10(conf)#monitor session 0
Force10(conf-mon-sess-0)#source gig 1/1 dest gig 1/2 direction rx
Force10(conf-mon-sess-0)#exit
Force10(conf)#do show monitor session 0
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 1/1       Gi 1/2          rx             interface   Port-based
Force10(conf)#
```

In Figure 28-7, the host and server are exchanging traffic which passes through interface gigabitethernet 1/1. Interface gigabitethernet 1/1 is the monitored port and gigabitethernet 1/2 is the monitoring port, which is configured to only monitor traffic received on gigabitethernet 1/1 (host-originated traffic).

Figure 28-7. Port Monitoring Example



```
Force10(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
 no ip address
 no shutdown
Force10(conf)#monitor session 0
Force10(conf-mon-sess-0)#source  gig 1/1 destination gig 1/2 direction rx
```

Flow-based Monitoring

Flow-based Monitoring is supported only on platform E

Flow-based monitoring conserves bandwidth by monitoring only specified traffic instead 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 may specify traffic using standard or extended access-lists.

To configure flow-based monitoring:

Step	Task	Command Syntax	Command Mode
4	Enable flow-based monitoring for a monitoring session.	**flow-based enable**	MONITOR SESSION
5	Define in an access-list rules that include the keyword **monitor**. FTOS only considers for port monitoring traffic matching rules with the keyword **monitor**. See Chapter 7, Access Control Lists (ACL), Prefix Lists, and Route-maps.	**ip access-list**	CONFIGURATION
6	Apply the ACL to the monitored port. See Chapter 7, Access Control Lists (ACL), Prefix Lists, and Route-maps.	**ip access-group access-list**	INTERFACE

View an access-list that you applied to an interface using the command **show ip accounting access-list** from EXEC Privilege mode, as shown in Figure 28-8.

Figure 28-8. Configuring Flow-based Monitoring

```
Force10(conf)#monitor session 0
Force10(conf-mon-sess-0)#flow-based enable
Force10(conf)#ip access-list ext testflow
Force10(config-ext-nacl)#seq 5 permit icmp any any count bytes monitor
Force10(config-ext-nacl)#seq 10 permit ip 102.1.1.0/24 any count bytes monitor
Force10(config-ext-nacl)#seq 15 deny udp any any count bytes
Force10(config-ext-nacl)#seq 20 deny tcp any any count bytes
Force10(config-ext-nacl)#exit
Force10(conf)#interface gig 1/1
Force10(conf-if-gi-1/1)#ip access-group testflow in
Force10(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 ip address 10.11.1.254/24
 ip access-group testflow in
 shutdown
Force10(conf-if-gi-1/1)#exit
Force10(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)
```

29

Private VLANs

FTOS 7.8.1.0 adds a Private VLAN (PVLAN) feature for the C-Series and S-Series: C S

For syntax details on the commands discussed in this chapter, see the Private VLANs Commands chapter in the *FTOS Command Reference*.

This chapter contains the following major sections:

Private VLANs extend the FTOS security suite by providing Layer 2 isolation between ports within the same VLAN. A private VLAN 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 private VLAN 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, while 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 is has one or more ports that are also isolated from each other.

Private VLAN Concepts

The VLAN types in a private VLAN (PVLAN) include:

Community VLAN — A *community VLAN* is 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 — An *isolated VLAN* is 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—A *primary VLAN* is the base VLAN of a private VLAN:

- 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 *secondary VLAN* is a subdomain of the primary VLAN. There are two types of secondary VLAN — community VLAN and isolated VLAN.

PVLAN port types:

- **Community port:** A *community port* is, by definition, 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**: A *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:** An *isolated port* is, by definition, a port that, in Layer 2, can only communicate with promiscuous ports that are in the same PVLAN.
- **Promiscuous port:** A *promiscuous port* is, by definition, 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**: A *trunk port*, by definition, carries traffic between switches:
  - A trunk port in a PVLAN is always tagged.
  - Primary or secondary VLAN traffic is carried by the trunk port in tagged mode. 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 details on port channels, see Port Channel Interfaces on page 294 in Chapter 15, Interfaces.

For an introduction to VLANs, see Chapter 20, Layer 2.

Private VLAN Commands

The commands dedicated to supporting the Private VLANs feature are:

Table 29-1. Private VLAN Commands

Task	Command Syntax	Command Mode
Enable/disable Layer 3 communication between secondary VLANs.	[**no**] **ip local-proxy-arp** **Note**: Even after **ip-local-proxy-arp** is disabled (**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.	INTERFACE VLAN
Set the mode of the selected VLAN to community, isolated, or primary.	[**no**] **private-vlan mode** {**community** | **isolated | primary**}	INTERFACE VLAN
Map secondary VLANs to the selected primary VLAN.	[**no**] **private-vlan mapping secondary-vlan** *vlan-list*	INTERFACE VLAN
Display type and status of PVLAN interfaces.	**show interfaces private-vlan** [**interface** *interface*]	EXEC EXEC Privilege
Display PVLANs and/or interfaces that are part of a PVLAN.	**show vlan private-vlan** [**community** | *interface* | **isolated** | **primary |** *primary_vlan* | **interface** *interface*]	EXEC EXEC Privilege
Display primary-secondary VLAN mapping.	**show vlan private-vlan mapping**	EXEC EXEC Privilege
Set the PVLAN mode of the selected port.	**switchport mode private-vlan** {**host | promiscuous | trunk**}	INTERFACE

Note: Secondary VLANs are Layer 2 VLANs, so even if they are operationally down while primary VLANs are operationally up, Layer 3 traffic will still be transmitted across secondary VLANs.

The outputs of the following commands are augmented in FTOS 7.8.1.0 to provide PVLAN data:

- **show arp**: See the IP Routing Commands chapter in the *FTOS Command Reference*.
- **show vlan**: See the Layer 2 Commands chapter in the *FTOS Command Reference*.

Private VLAN Configuration Task List

The following sections contain the procedures that configure a private VLAN:

- Creating PVLAN ports
- Creating a Primary VLAN on page 541
- Creating a Community VLAN on page 542
- Creating an Isolated VLAN on page 542

Creating PVLAN ports

Private VLAN ports are those that will be assigned to the private VLAN (PVLAN).

Step	Command Syntax	Command Mode	Purpose
1	**interface** *interface*	CONFIGURATION	Access the INTERFACE mode for the port that you want to assign to a PVLAN.
2	**no shutdown**	INTERFACE	Enable the port.
3	**switchport**	INTERFACE	Set the port in Layer 2 mode.
4	**switchport mode private-vlan** {**host | promiscuous | trunk**}	INTERFACE	Select the PVLAN mode: • **host** (port in isolated or community VLAN) • **promiscuous** (intra-VLAN communication port) • **trunk** (inter-switch PVLAN hub port)

For interface details, see Enable a Physical Interface on page 286 in Chapter 15, Interfaces.

Note: Interfaces that are configured as PVLAN ports cannot be added to regular VLANs. Conversely, “regular” ports (ports not configured as PVLAN ports) cannot be added to PVLANs.

Figure 29-1 shows the use of the **switchport mode private-vlan** command on a port and on a port channel:

Figure 29-1. Examples of switchport mode private-vlan Command

```
Force10#conf
Force10(conf)#interface GigabitEthernet 2/1
Force10(conf-if-gi-2/1)#switchport mode private-vlan promiscuous

Force10(conf)#interface GigabitEthernet 2/2
Force10(conf-if-gi-2/2)#switchport mode private-vlan host

Force10(conf)#interface GigabitEthernet 2/3
Force10(conf-if-gi-2/3)#switchport mode private-vlan trunk

Force10(conf)#interface GigabitEthernet 2/2
Force10(conf-if-gi-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.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Access the INTERFACE VLAN mode for the VLAN to which you want to assign the PVLAN interfaces.
2	**no shutdown**	INTERFACE VLAN	Enable the VLAN.
3	**private-vlan mode primary**	INTERFACE VLAN	Set the PVLAN mode of the selected VLAN to primary.
4	**private-vlan mapping secondary-vlan** *vlan-list*	INTERFACE VLAN	Map secondary VLANs to the selected primary VLAN. 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	**tagged** *interface* or **untagged** *interface*	INTERFACE VLAN	Add promiscuous ports as tagged or untagged interfaces. Add PVLAN trunk ports to the VLAN only as tagged interfaces. Interfaces can be entered singly or in range format, either comma-delimited (*slot/port,port,port*) or hyphenated (*slot/port-port).* Only promiscuous ports or PVLAN trunk ports can be added to the PVLAN (no host or regular ports).
6	**ip address** *ip address*	INTERFACE VLAN	(OPTIONAL) Assign an IP address to the VLAN.
7	**ip local-proxy-arp**	INTERFACE VLAN	(OPTIONAL) Enable/disable Layer 3 communication between secondary VLANs.

Note: If a promiscuous or host port is untagged in a VLAN and it receives a tagged packet in the same VLAN, the packet will NOT be 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.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Access the INTERFACE VLAN mode for the VLAN that you want to make a community VLAN.
2	**no shutdown**	INTERFACE VLAN	Enable the VLAN.
3	**private-vlan mode community**	INTERFACE VLAN	Set the PVLAN mode of the selected VLAN to community.
4	**tagged** *interface* or **untagged** *interface*	INTERFACE VLAN	Add one or more host ports to the VLAN. The interfaces can be entered singly or in range format, either comma-delimited (*slot/port,port,port*) or hyphenated (*slot/ port-port).* Only host (isolated) ports can be added to the VLAN.

Creating an Isolated VLAN

An isolated VLAN is a secondary VLAN of a primary VLAN. Its ports can only talk with the promiscuous ports in that primary VLAN.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Access the INTERFACE VLAN mode for the VLAN that you want to make an isolated VLAN.
2	**no shutdown**	INTERFACE VLAN	Enable the VLAN.
3	**private-vlan mode isolated**	INTERFACE VLAN	Set the PVLAN mode of the selected VLAN to isolated.
4	**tagged** *interface* or **untagged** *interface*	INTERFACE VLAN	Add one or more host ports to the VLAN. The interfaces can be entered singly or in range format, either comma-delimited (*slot/port,port,port*) or hyphenated (*slot/ port-port).* Only ports defined as **host** can be added to the VLAN.

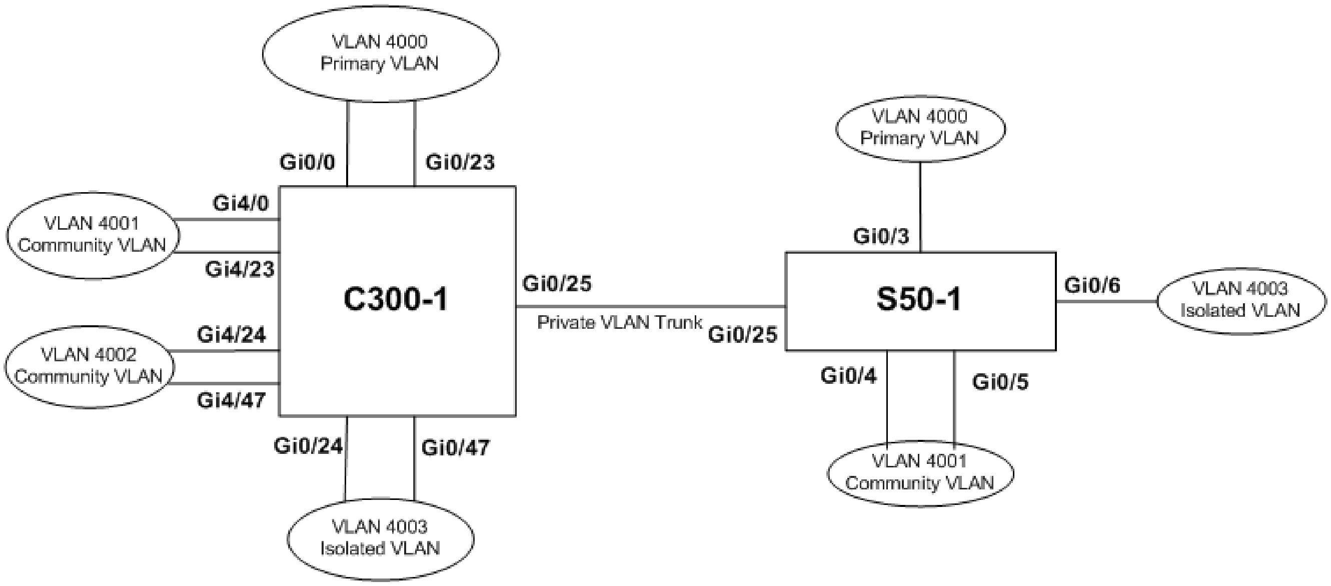
Figure 29-2 shows the use of the PVLAN commands that are used in VLAN INTERFACE mode to configure the PVLAN member VLANs (primary, community, and isolated VLANs):

Figure 29-2. Configuring VLANs for a Private VLAN

```
Force10#conf
Force10(conf)# interface vlan 10
Force10(conf-vlan-10)# private-vlan mode primary
Force10(conf-vlan-10)# private-vlan mapping secondary-vlan 100-101
Force10(conf-vlan-10)# untagged Gi 2/1
Force10(conf-vlan-10)# tagged Gi 2/3

Force10(conf)# interface vlan 101
Force10(conf-vlan-101)# private-vlan mode community
Force10(conf-vlan-101)# untagged Gi 2/10

Force10(conf)# interface vlan 100
```

Private VLAN Configuration Example

Figure 29-3. Sample Private VLAN Topology



The following configuration is based on the example diagram, above:

On C300-1:

- Gi 0/0 and Gi 23 are configured as promiscuous ports, assigned to the primary VLAN, VLAN 4000.
- Gi 0/25 is configured as a PVLAN trunk port, also assigned to the primary VLAN 4000.
- Gi 0/24 and Gi 0/47 are configured as host ports and assigned to the isolated VLAN, VLAN 4003.
- Gi 4/0 and Gi 23 are configured as host ports and assigned to the community VLAN, VLAN 4001.
- Gi 4/24 and Gi 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 command **ip local-proxy-arp** is invoked in the primary VLAN.

Note: Even after **ip-local-proxy-arp** is disabled (**no ip-local-proxy-arp**) in a secondary VLAN, Layer 3 communication may happen between some secondary VLAN hosts, until the ARP timeout happens on those secondary VLAN hosts.

In parallel, on S50-1:

- Gi 0/3 is a promiscuous port and Gi 0/25 is a PVLAN trunk port, assigned to the primary VLAN 4000.
- Gi 0/4-6 are host ports. Gi 0/4 and Gi 0/5 are assigned to the community VLAN 4001, while Gi 0/6 is assigned to the isolated VLAN 4003.

The result is that:

- The S50V ports would have the same intra-switch communication characteristics as described above for the C300.
- For transmission between switches, tagged packets originating from host PVLAN ports in one secondary VLAN and destined for host PVLAN ports in the other switch travel through the promiscuous ports in the local VLAN 4000 and then through the trunk ports (0/25 in each switch).

Inspecting the Private VLAN Configuration

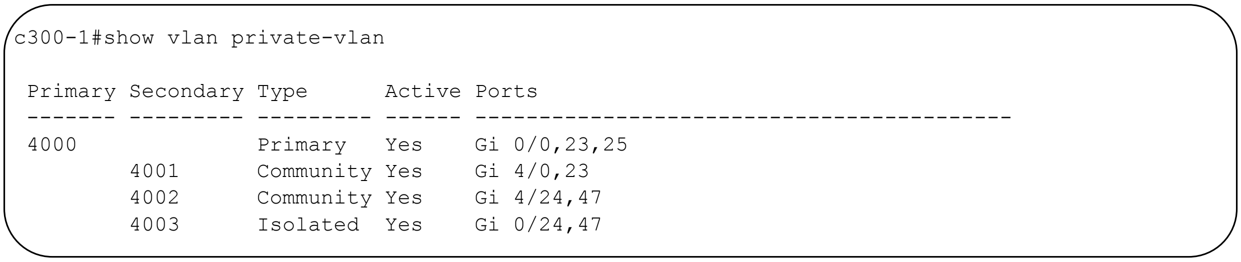
The standard methods of inspecting configurations also apply in PVLANs:

- Within the INTERFACE and INTERFACE VLAN modes, use the **show config** command to display the specific interface configuration.
- Inspect the running-config, and, with the **grep** pipe option (**show running-config** | **grep** *string*), you can display a specific part of the running-config. Figure 29-8 shows the PVLAN parts of the running-config from the S50V switch in the topology diagram shown in Figure 29-3, above.
- You can also use one of three **show** commands that are specific to the Private VLAN feature:
  - **show interfaces private-vlan** [**interface** *interface*]: Display the type and status of the configured PVLAN interfaces. See the example output in the Security chapter of the *FTOS Command Reference*.
  - **show vlan private-vlan** [**community** | *interface* | **isolated** | **primary |** *primary_vlan* | **interface** *interface*]: Display the configured PVLANs or interfaces that are part of a PVLAN. Figure 29-4 shows the results of using the command without command options on the C300 switch in the topology diagram shown in Figure 29-3, above, while Figure 29-5 shows the results on the S50V.

- **show vlan private-vlan mapping**: Display the primary-secondary VLAN mapping. See the example output from the S50V, above, in Figure 29-6.

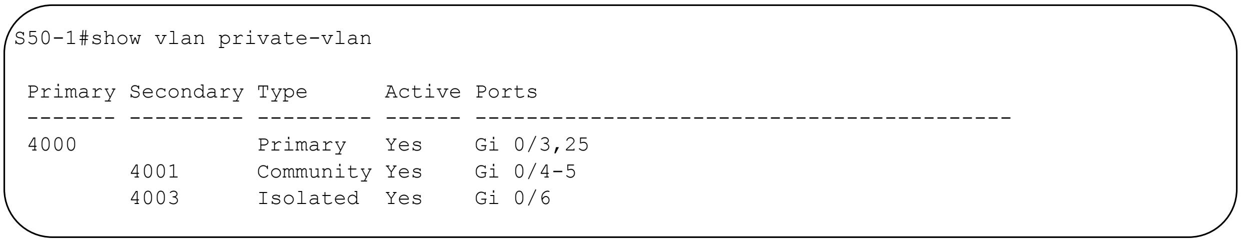
- Two **show** commands revised to display PVLAN data are:
  - **show arp**
  - **show vlan**: See revised output in Figure 29-7.

Figure 29-4. show vlan private-vlan Example Output from C300



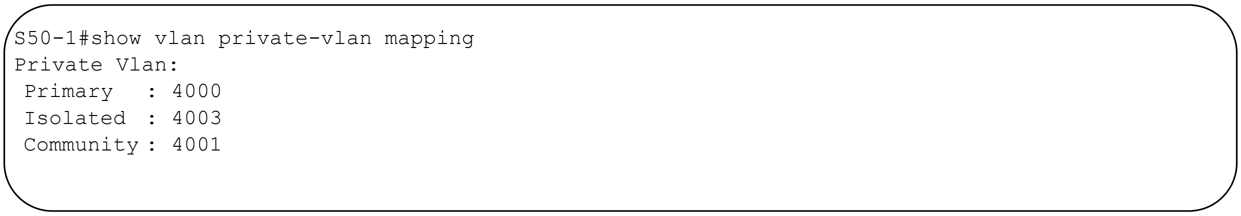
```
c300-1#show vlan private-vlan

 Primary Secondary Type      Active Ports
 ------- --------- --------- ------ ------------------------------------------
 4000              Primary   Yes    Gi 0/0,23,25
         4001      Community Yes    Gi 4/0,23
         4002      Community Yes    Gi 4/24,47
         4003      Isolated  Yes    Gi 0/24,47
```

Figure 29-5. show vlan private-vlan Example Output from S50V



```
S50-1#show vlan private-vlan

 Primary Secondary Type      Active Ports
 ------- --------- --------- ------ ------------------------------------------
 4000              Primary   Yes    Gi 0/3,25
         4001      Community Yes    Gi 0/4-5
         4003      Isolated  Yes    Gi 0/6
```

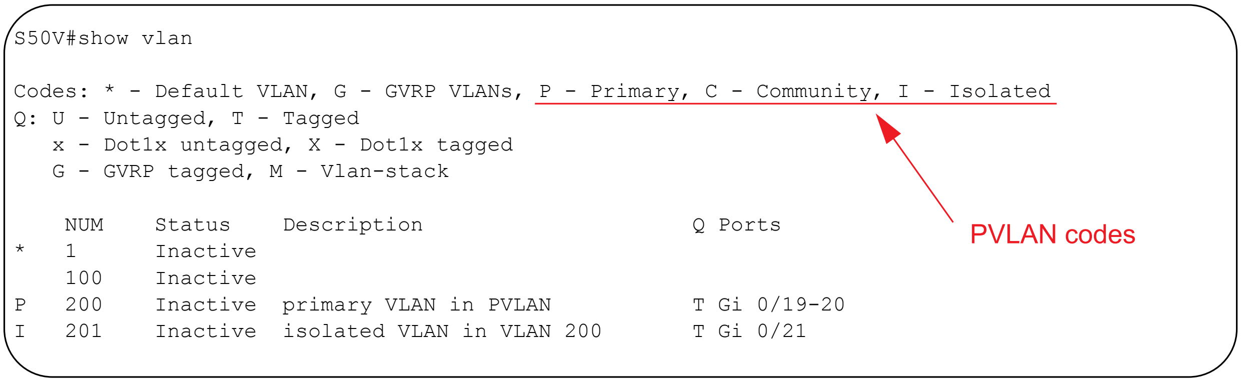
Figure 29-6. show vlan private-vlan mapping Example Output from S50V



```
S50-1#show vlan private-vlan mapping
Private Vlan:
 Primary   : 4000
 Isolated  : 4003
 Community : 4001
```

In the following screenshot, note the addition of the PVLAN codes — P, I, and C — in the left column:

Figure 29-7. show vlan Example Output from S50V



```
S50V#show vlan

Codes: * - Default VLAN, G - GVRP VLANs, P - Primary, C - Community, I - Isolated
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

    NUM    Status    Description                     Q Ports
*   1      Inactive
    100    Inactive
P   200    Inactive  primary VLAN in PVLAN           T Gi 0/19-20
I   201    Inactive  isolated VLAN in VLAN 200       T Gi 0/21
```

PVLAN codes

Figure 29-8. Example running-config Output of PVLAN Configuration from S50V

```
!
interface GigabitEthernet 0/3
 no ip address
 switchport
 switchport mode private-vlan promiscuous
 no shutdown
!
interface GigabitEthernet 0/4
 no ip address
 switchport
 switchport mode private-vlan host
 no shutdown
!
interface GigabitEthernet 0/5
 no ip address
 switchport
 switchport mode private-vlan host
 no shutdown
!
interface GigabitEthernet 0/6
 no ip address
 switchport
 switchport mode private-vlan host
 no shutdown
!
interface GigabitEthernet 0/25
 no ip address
 switchport
 switchport mode private-vlan trunk
 no shutdown
!
interface Vlan 4000
 private-vlan mode primary
 private-vlan mapping secondary-vlan 4001-4003
 no ip address
 tagged GigabitEthernet 0/3,25
 no shutdown
!
interface Vlan 4001
 private-vlan mode community
```

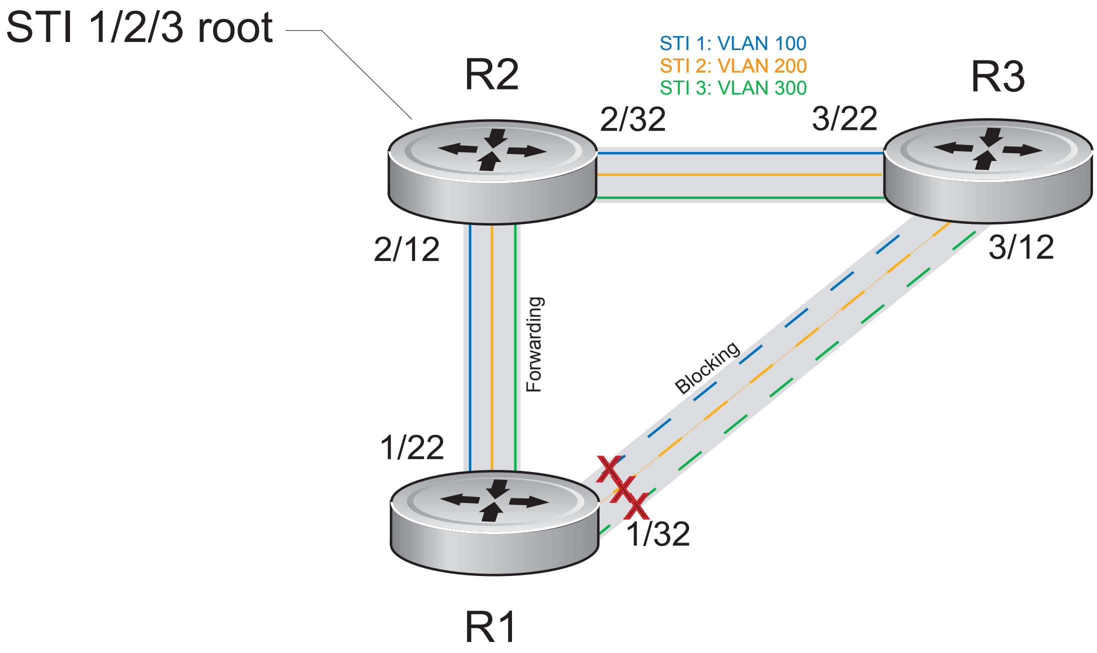
Per-VLAN Spanning Tree Plus

Per-VLAN Spanning Tree Plus is supported platforms: C E S

Protocol Overview

Per-VLAN Spanning Tree Plus (PVST+) is a variation of Spanning Tree—developed by a third party—that allows you to configure a separate Spanning Tree instance for each VLAN. For more information on Spanning Tree, see Chapter 39, Spanning Tree Protocol.

Figure 30-1. Per-VLAN Spanning Tree



FTOS supports three other variations of Spanning Tree, as shown in Table 30-1.

Table 30-1. FTOS Supported Spanning Tree Protocols

Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Implementation Information

- The FTOS implementation of PVST+ is based on IEEE Standard 802.1d.
- The FTOS implementation of PVST+ uses IEEE 802.1s costs as the default costs (Table 30-2). Other implementations use IEEE 802.1d costs as the default costs if you are using Force10 systems in a multi-vendor network, verify that the costs are values you intended.
-
- On the C-Series and S-Series, 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+. See page 549.
4. Optionally, for load balancing, select a non-default bridge-priority for a VLAN. See page 549.

Related Configuration Tasks

- Modify Global PVST+ Parameters on page 551
- Modify Interface PVST+ Parameters on page 552
- Configure an EdgePort on page 553
- Flush MAC Addresses after a Topology Change on page 435
- Preventing Network Disruptions with BPDU Guard on page 711
- SNMP Traps for Root Elections and Topology Changes on page 713
- Configuring Spanning Trees as Hitless on page 713
- PVST+ in Multi-vendor Networks on page 554
- PVST+ Extended System ID on page 554
- PVST+ Sample Configurations on page 555

Enable PVST+

When you enable PVST+, FTOS instantiates STP on each active VLAN. To enable PVST+ globally:

Step	Task	Command Syntax	Command Mode
1	Enter PVST context.	**protocol spanning-tree pvst**	PROTOCOL PVST
2	Enable PVST+.	**no disable**	PROTOCOL PVST

Disable PVST+

Task	Command Syntax	Command Mode
Disable PVST+ globally.	**disable**	PROTOCOL PVST
Disable PVST+ on an interface, or remove a PVST+ parameter configuration.	**no spanning-tree pvst**	INTERFACE

Display your PVST+ configuration by entering the command **show config** from PROTOCOL PVST context, as shown in fig.

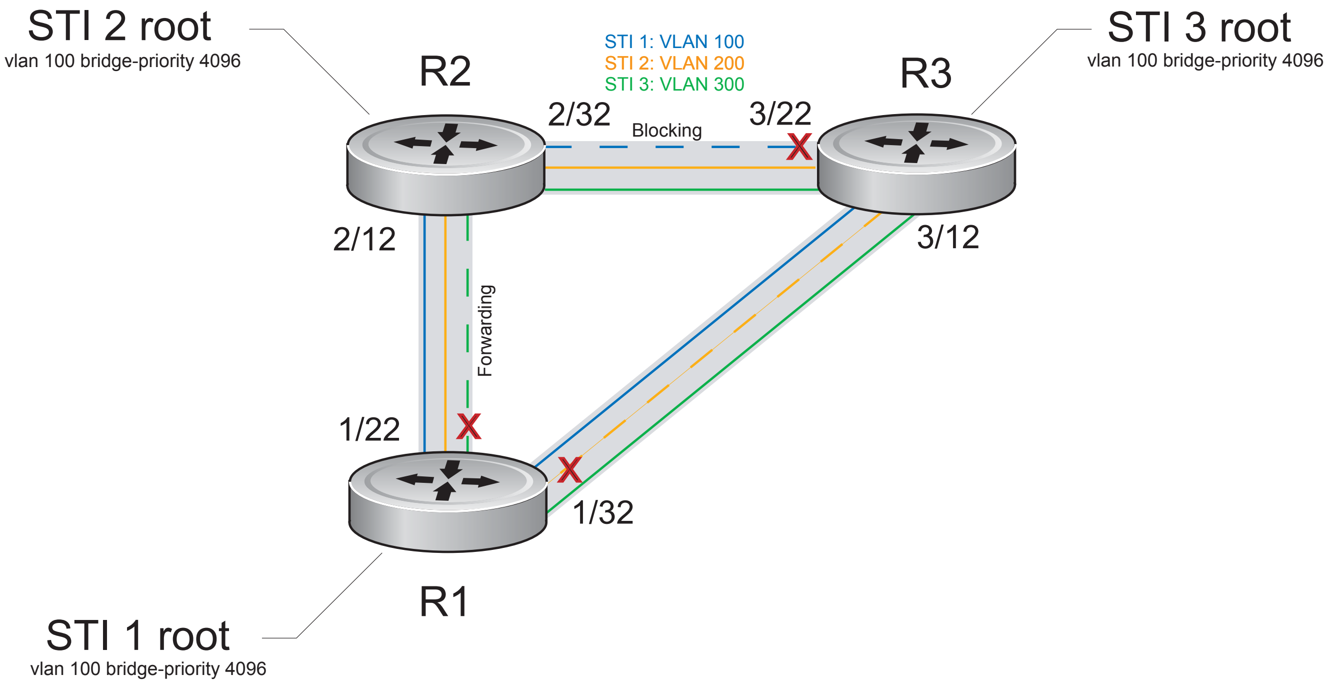
Figure 30-2. Display the PVST+ Configuration

```
Force10_E600(conf-pvst)#show config verbose
!
protocol spanning-tree pvst
 no disable
 vlan 100 bridge-priority 4096
```

Influence PVST+ Root Selection

In Figure 30-1, all VLANs use the same forwarding topology because R2 is elected the root, and all GigabitEthernet ports have the same cost. Figure 30-3 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 30-3. Load Balancing with PVST+



The bridge with the bridge value for bridge priority is elected root. Since all bridges use the default priority (until configured otherwise), lowest MAC address is used as a tie-breaker. Assign bridges a low non-default value for bridge priority to increase the likelihood that it will be selected as the STP root.

Task	Command Syntax	Command Mode
Assign a bridge priority. Range: 0 to 61440 Default: 32768	**vlan bridge-priority**	PROTOCOL PVST

Display the PVST+ forwarding topology by entering the command **show spanning-tree pvst** [**vlan** *vlan-id*] from EXEC Privilege mode, as shown in Figure 30-4.

Figure 30-4. Display the PVST+ Forwarding Topology

```
Force10_E600(conf)#do show spanning-tree pvst vlan 100
VLAN 100
Root Identifier has priority 4096, Address 0001.e80d.b6d6
Root Bridge hello time 2, max age 20, forward delay 15
Bridge Identifier has priority 4096, Address 0001.e80d.b6d6
Configured hello time 2, max age 20, forward delay 15
We are the root of VLAN 100
Current root has priority 4096, Address 0001.e80d.b6d6
Number of topology changes 5, last change occurred 00:34:37 ago on Gi 1/32

Port 375 (GigabitEthernet 1/22) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.375
Designated root has priority 4096, address 0001.e80d.b6:d6
Designated bridge has priority 4096, address 0001.e80d.b6:d6
Designated port id is 128.375 , designated path cost 0
Number of transitions to forwarding state 2
BPDU sent 1159, received 632
The port is not in the Edge port mode

Port 385 (GigabitEthernet 1/32) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.385
Designated root has priority 4096, address 0001.e80d.b6:d6
Designated bridge has priority 4096, address 0001.e80d.b6:d6
Designated port id is 128.385 , designated path cost 0
```

Modify 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** is the amount of time an interface waits in the Listening State and the Learning State before it transitions to the Forwarding State.
- **Hello-time** is the time interval in which the bridge sends Bridge Protocol Data Units (BPDUs).
- **Max-age** is the length of time the bridge maintains configuration information before it refreshes that information by recomputing the PVST+ topology.

To change PVST+ parameters, use the following commands on the root bridge:

Task	Command Syntax	Command Mode
Change the forward-delay parameter. • Range: 4 to 30 • Default: 15 seconds	**vlan forward-delay**	PROTOCOL PVST
Change the hello-time parameter. **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**vlan hello-time**	PROTOCOL PVST

Task	Command Syntax	Command Mode
Change the max-age parameter. Range: 6 to 40 Default: 20 seconds	**vlan max-age**	PROTOCOL PVST

The values for global PVST+ parameters are given in the output of the command **show spanning-tree pvst**, as shown in Figure 30-4.

Modify Interface PVST+ 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 will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

Table 30-2 lists the default values for port cost by interface.

Table 30-2. PVST+ Default Port Cost Values

Port Cost	Default Value
100-Mb/s Ethernet interfaces	200000
1-Gigabit Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
Port Channel with 100 Mb/s Ethernet interfaces	180000
Port Channel with 1-Gigabit Ethernet interfaces	18000
Port Channel with 10-Gigabit Ethernet interfaces	1800

Note: The FTOS implementation of PVST+ uses IEEE 802.1s costs as the default costs. Other implementations use IEEE 802.1d costs as the default costs if you are using Dell Force10 systems in a multi-vendor network, verify that the costs are values you intended.

To change the port cost or priority of an interface:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 200000 Default: see Table 30-2.	**spanning-tree pvst vlan cost**	INTERFACE

Task	Command Syntax	Command Mode
Change the port priority of an interface. Range: 0 to 240, in increments of 16 Default: 128	**spanning-tree pvst vlan priority**	INTERFACE

The values for interface PVST+ parameters are given in the output of the command **show spanning-tree pvst**, as shown in Figure 30-4.

Configure 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 shutdown 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 will drop 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 it is enabled on an interface connected to a network.

To enable EdgePort on an interface, use the following command:

Task	Command Syntax	Command Mode
Enable EdgePort on an interface.	**spanning-tree pvst edge-port [bpduguard | shutdown-on-violation**]	INTERFACE

The EdgePort status of each interface is given in the output of the command **show spanning-tree pvst**, as shown in Figure 30-4.

FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

The error disabled state can be cleared with any of the following methods:

- Perform an **shutdown** command on the interface.
- Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).
- Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
- Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

PVST+ in Multi-vendor Networks

Some non-Dell Force10 systems which have hybrid ports participating in PVST+ transmit two kinds of BPDUs: an 802.1D BPDU and an untagged PVST+ BPDU.

Dell Force10 systems do not expect PVST+ BPDU (tagged or untagged) on an untagged port. If this happens, FTOS places the port in error-disable state. This behavior might result in the network not converging. To prevent FTOS from executing this action, use the command **no spanning-tree pvst err-disable cause invalid-pvst-bpdu**. After you configure this command, if the port receives a PVST+ BPDU, the BPDU is dropped, and the port remains operational.

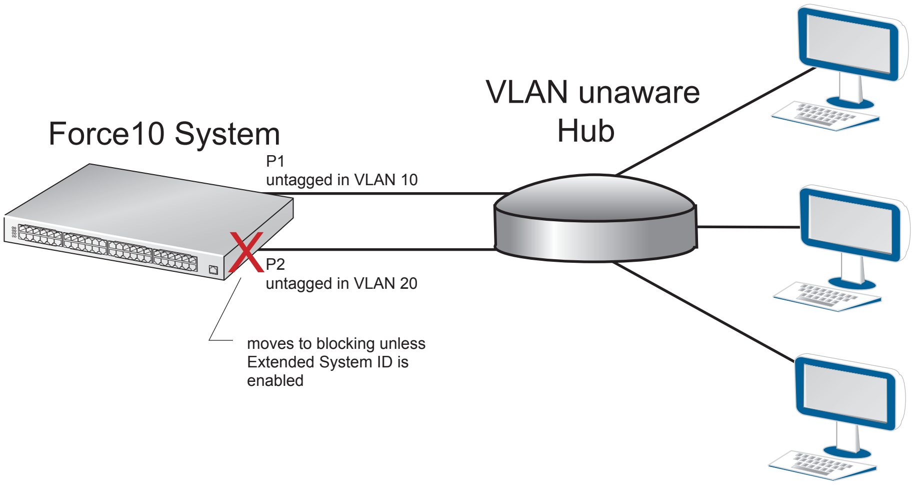
PVST+ Extended System ID

In Figure 30-5, 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 the above scenario, however, PVST+ can be employed to avoid potential misconfigurations.

If PVST+ is enabled on the Dell Force10 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 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 forwarding state.

Figure 30-5. PVST+ with Extend System ID



Task	Command Syntax	Command Mode
Augment the Bridge ID with the VLAN ID.	**extend system-id**	PROTOCOL PVST

```
Force10(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

Figure 30-6, Figure 30-7, and Figure 30-8 provide the running configurations for the the topology shown in Figure 30-3.

Figure 30-6. PVST+ Sample Configuration: R1 Running-configuration

```
interface GigabitEthernet 1/22
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 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 GigabitEthernet 1/22,32
 no shutdown
!
interface Vlan 200
 no ip address
 tagged GigabitEthernet 1/22,32
 no shutdown
!
interface Vlan 300
 no ip address
 tagged GigabitEthernet 1/22,32
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 100 bridge-priority 4096
```

Figure 30-7. PVST+ Sample Configuration: R2 Running-configuration

```
interface GigabitEthernet 2/12
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 2/32
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 2/12,32
 no shutdown
!
interface Vlan 200
 no ip address
 tagged GigabitEthernet 2/12,32
 no shutdown
!
interface Vlan 300
 no ip address
 tagged GigabitEthernet 2/12,32
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 200 bridge-priority 4096
```

Figure 30-8. PVST+ Sample Configuration: R3 Running-configuration

```
interface GigabitEthernet 3/12
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 3/22
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 3/12,22
 no shutdown
!
interface Vlan 200
 no ip address
 tagged GigabitEthernet 3/12,22
 no shutdown
!
interface Vlan 300
 no ip address
 tagged GigabitEthernet 3/12,22
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 300 bridge-priority 4096
```

31

Quality of Service

Quality of Service (QoS) is supported on platforms: C E S

Differentiated service is accomplished by classifying and queuing traffic, and assigning priorities to those queues.

The C-Series traffic has eight queues per port. Four queues are for data traffic and four are for control traffic. All queues are serviced using the Deficit Round Robin scheduling algorithm. You can only manage queuing prioritization on egress.

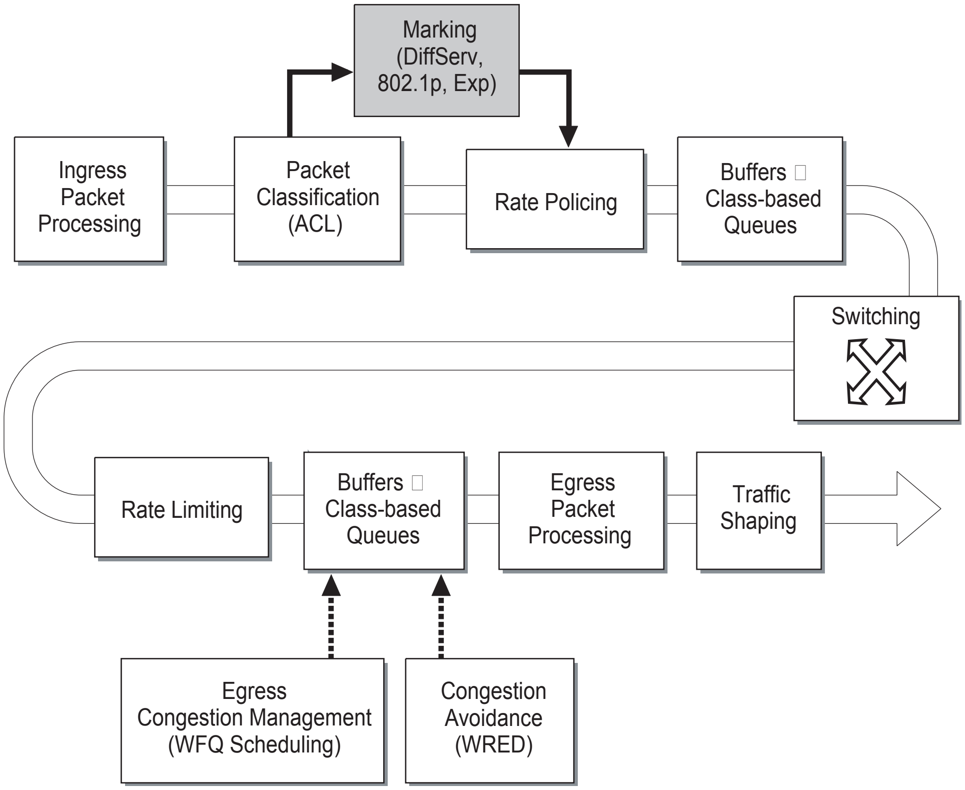
Table 31-1. FTOS Support for Port-based, Policy-based, and Multicast QoS Features

Feature	Platform	Direction
Port-based QoS Configurations	C E S	Ingress + Egress
Set dot1p Priorities for Incoming Traffic	C E S	Ingress
Honor dot1p Priorities on Ingress Traffic	C E S	
Configure Port-based Rate Policing	C E S	
Configure Port-based Rate Limiting	E	Egress
Configure Port-based Rate Shaping	C E S	
Policy-based QoS Configurations	C E S	Ingress + Egress
Classify Traffic	C E S	Ingress
Create a Layer 3 class map	C E S	
Set DSCP values for egress packets based on flow	C E S	
Create a Layer 2 class map	C E S	
Create a QoS Policy	C E S	Ingress + Egress
Create an input QoS policy	C E S	Ingress
Configure policy-based rate policing	C E S	
Set a DSCP value for egress packets	C E S	
Set a dot1p value for egress packets	C E S	

Table 31-1. FTOS Support for Port-based, Policy-based, and Multicast QoS Features

Feature	Platform	Direction
Create an output QoS policy	C E S	Egress
Configure policy-based rate limiting	E	
Configure policy-based rate shaping	C E S	
Allocate bandwidth to queue	C E S	
Specify WRED drop precedence	E	
Create Policy Maps	C E S	Ingress + Egress
Create Input Policy Maps	C E S	Ingress
Honor DSCP values on ingress packets	C E S	
Honoring dot1p values on ingress packets	E C S	
Create Output Policy Maps	C E S	Egress
Specify an aggregate QoS policy	C E S	
QoS Rate Adjustment	C E S	
Strict-priority Queueing	C E S	—
Weighted Random Early Detection	E	Egress
Create WRED Profiles	E	
Configure WRED for Storm Control	E	
Allocating Bandwidth to Multicast Queues	E	Egress
Pre-calculating Available QoS CAM Space	C E S	—
Viewing QoS CAM Entries	E	—

Figure 31-1. Dell Force10 QoS Architecture


Marking (DiffServ, 802.1p, Exp)
Ingress Packet Processing
Packet Classification (ACL)
Rate Policing
Buffers Class-based Queues
Switching
Rate Limiting
Buffers Class-based Queues
Egress Packet Processing
Traffic Shaping
Egress Congestion Management (WFQ Scheduling)
Congestion Avoidance (WRED)

Implementation Information

Dell Force10' 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:

- Set dot1p Priorities for Incoming Traffic on page 562
- Configure Port-based Rate Policing on page 563
- Configure Port-based Rate Limiting on page 564

- Configure Port-based Rate Shaping on page 565
- Storm Control on page 741

Set dot1p Priorities for Incoming Traffic

Change the priority of incoming traffic on the interface using the command **dot1p-priority** from INTERFACE mode, as shown in Figure 31-2. FTOS places traffic marked with a priority in a queue based on Table 31-2. 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 interfaces in a port-channel.

FTOS Behavior: The C-Series and S-Series distribute eight dot1p priorities across four data queues. This is different from the E-Series, which distributes eight dot1p priorities across eight queues (Table 31-2).

Table 31-2. dot1p-priority values and queue numbers

dot1p	E-Series Queue Number	C-Series Queue Number	S-Series Queue Number
0	2	1	1
1	0	0	0
2	1	0	0
3	3	1	1
4	4	2	2
5	5	2	2
6	6	3	3
7	7	3	3

Figure 31-2. Configuring dot1p Priority on an Interface

```
Force10#config
Force10(conf)#interface gigabitethernet 1/0
Force10(conf-if)#switchport
Force10(conf-if)#dot1p-priority 1
Force10(conf-if)#end
Force10#
```

Honor dot1p Priorities on Ingress Traffic

By default FTOS does not honor dot1p priorities on ingress traffic. Use the command **service-class dynamic dot1p** from INTERFACE mode to honor dot1p priorities on ingress traffic, as shown in Figure 31-3. You can configure this feature on physical interfaces and port-channels, but you cannot configure it on individual interfaces in a port channel.

On the C-Series and S-Series 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. See Mapping dot1p values to service queues on page 576.

Note: You cannot configure **service-policy input** and **service-class dynamic dot1p** on the same interface.

Figure 31-3. service-class dynamic dot1p Command Example

```
Force10#config t
Force10(conf)#interface gigabitethernet 1/0
Force10(conf-if)#service-class dynamic dot1p
Force10(conf-if)#end
Force10#
```

Priority-tagged Frames on the Default VLAN

Priority-tagged Frames on the Default VLAN is available only on platforms: E X C S

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 **service-class dynamic dot1p** or **trust dot1p** is configured.

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 dot1p priority if **service-class dynamic dotp** or **trust dot1p** are configured. When priority-tagged frames ingress a tagged port, the frames are dropped because for a tagged port the default VLAN is 0.

FTOS Behavior: Hybrid ports can receive untagged, tagged, and priority tagged frames. The rate metering calculation might be inaccurate for untagged ports, since 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.

Configure Port-based Rate Policing

Rate policing ingress traffic on an interface using the command **rate police** from INTERACE mode, as shown in Figure 31-4. If the interface is a member of a VLAN, you may specify the VLAN for which ingress packets are policed.

Figure 31-4. Rate Policing Ingress Traffic

```
Force10#config t
Force10(conf)#interface gigabitethernet 1/0
Force10(conf-if)#rate police 100 40 peak 150 50
Force10(conf-if)#end
Force10#
```

Figure 31-5. Displaying your Rate Policing Configuration

```
Force10#show interfaces gigabitEthernet 1/2 rate police
  Rate police 300 (50) peak 800 (50)
    Traffic Monitor 0: normal 300 (50) peak 800 (50)
      Out of profile yellow 23386960 red 320605113
    Traffic Monitor 1: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 2: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 3: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 4: normal NA peak NA
      Out of profile yellow 0 red 0
```

Configure Port-based Rate Limiting

Configure Port-based Rate Limiting is supported only on platform E

FTOS Behavior: On the C-Series and S-Series, rate shaping is effectively rate limiting because of its smaller buffer size.

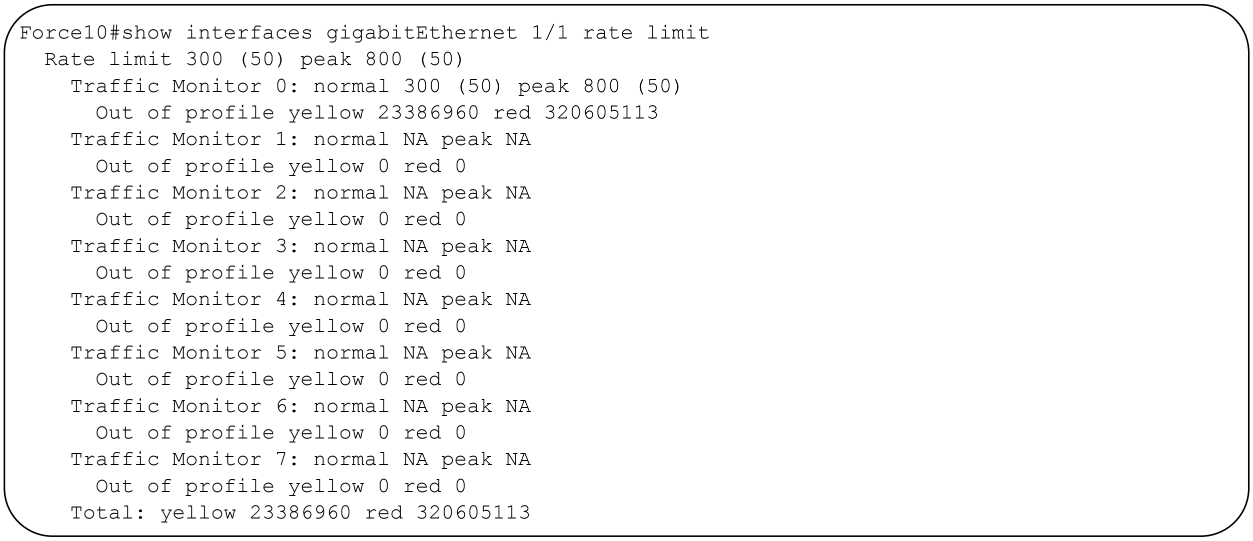
Rate limit egress traffic on an interface using the command rate limit from INTERFACE mode, as shown in Figure 31-6. If the interface is a member of a VLAN, you may specify the VLAN for which egress packets are rate limited.

Figure 31-6. Rate Limiting Egress Traffic

```
Force10#config t
Force10(conf)#interface gigabitethernet 1/0
Force10(conf-if)#rate limit 100 40 peak 150 50
Force10(conf-if)#end
Force10#
```

Display how your rate limiting configuration affects traffic using the keyword **rate limit** with the command **show interfaces**, as shown in Figure 31-7.

Figure 31-7. Displaying How Your Rate Limiting Configuration Affects Traffic



```
Force10#show interfaces gigabitEthernet 1/1 rate limit
  Rate limit 300 (50) peak 800 (50)
    Traffic Monitor 0: normal 300 (50) peak 800 (50)
      Out of profile yellow 23386960 red 320605113
    Traffic Monitor 1: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 2: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 3: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 4: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 5: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 6: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 7: normal NA peak NA
      Out of profile yellow 0 red 0
    Total: yellow 23386960 red 320605113
```

Configure Port-based Rate Shaping

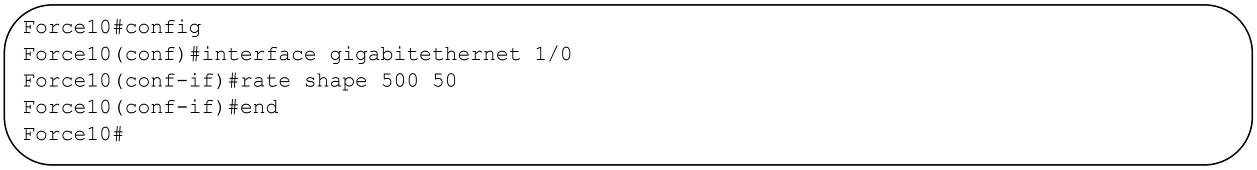
Configure Port-based Rate Limiting is supported only on platform C E S

FTOS Behavior: On the C-Series and S-Series, rate shaping is effectively rate limiting because of its smaller buffer size. On the S55, rate shaping on tagged ports is slightly greater than the configured rate and rate shaping on untagged ports is slightly less than configured rate.

Rate shaping buffers, rather than drops, traffic exceeding the specified rate until the buffer is exhausted. If any stream exceeds the configured bandwidth on a continuous basis, it can consume all of the buffer space that is allocated to the port.

- Apply rate shaping to outgoing traffic on a port using the command **rate shape** from INTERFACE mode, as shown in Figure 31-8.
- Apply rate shaping to a queue using the command **rate-shape** from QoS Policy mode.

Figure 31-8. Applying Rate Shaping to Outgoing Traffic

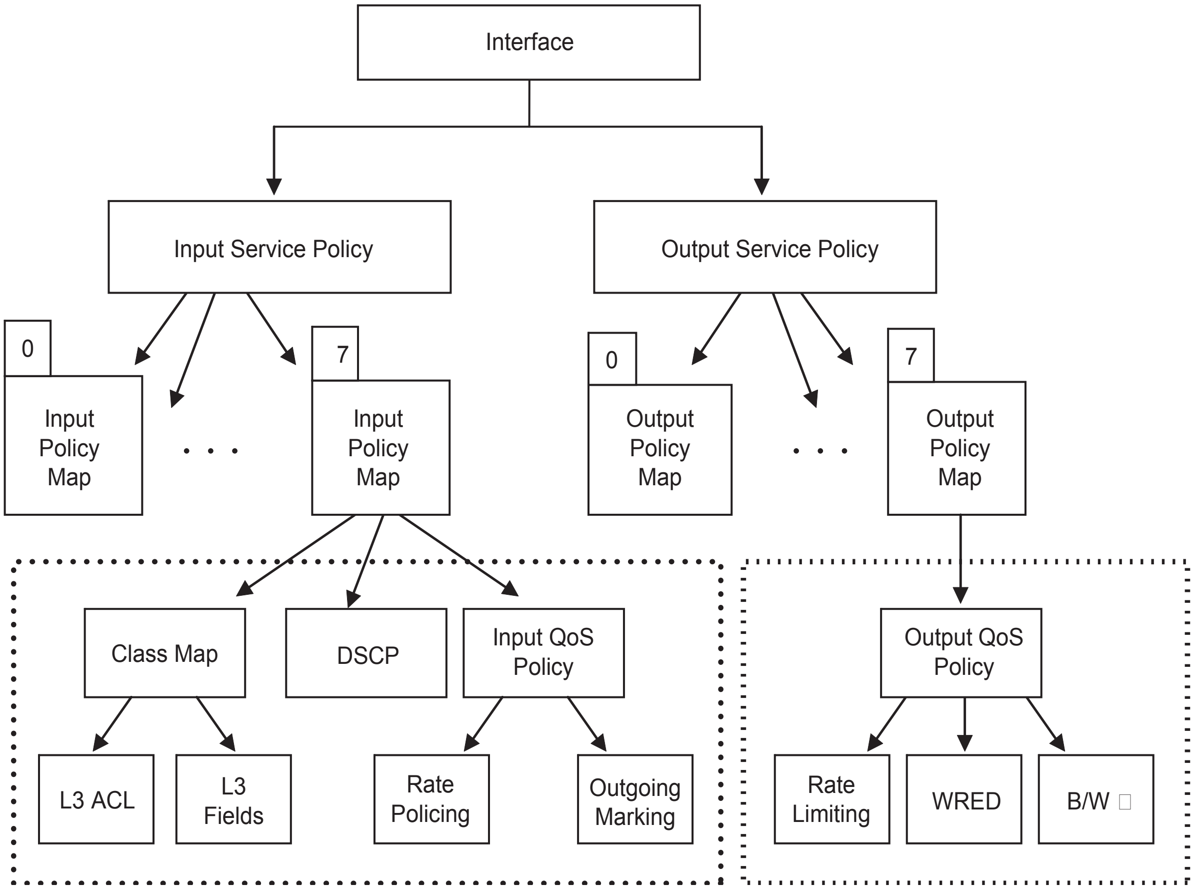


```
Force10#config
Force10(conf)#interface gigabitethernet 1/0
Force10(conf-if)#rate shape 500 50
Force10(conf-if)#end
Force10#
```

Policy-based QoS Configurations

Policy-based QoS configurations consist of the components shown in Figure 31-9.

Figure 31-9. Constructing Policy-based QoS Configurations



Classify Traffic

Class maps differentiate traffic so that you can apply separate quality of service policies to each class. For both class maps, Layer 2 and Layer 3, FTOS matches packets against match criteria in the order that you configure them.

Create a Layer 3 class map

A Layer 3 class map differentiates ingress packets based on DSCP value or IP precedence, and characteristics defined in an IP ACL. 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.

1. Create a match-any class map using the command **class-map match-any** or a match-all class map using the command **class-map match-all** from CONFIGURATION mode, as shown in Figure 31-10.
2. Once you create a class-map, FTOS places you in CLASS MAP mode. From this mode, specify your match criteria using the command **match ip**, as shown in Figure 31-10. Match-any class maps allow up to five ACLs, and match-all class-maps allow only one ACL.
3. After you specify your match criteria, link the class-map to a queue using the command **service-queue** from POLICY MAP mode, as shown in Figure 31-10.

Figure 31-10. Using the Order Keyword in ACLs

```
Force10(conf)#ip access-list standard acl1
Force10(config-std-nacl)#permit 20.0.0.0/8
Force10(config-std-nacl)#exit
Force10(conf)#ip access-list standard acl2
Force10(config-std-nacl)#permit 20.1.1.0/24 order 0
Force10(config-std-nacl)#exit
Force10(conf)#class-map match-all cmap1
Force10(conf-class-map)#match ip access-group acl1
Force10(conf-class-map)#exit
Force10(conf)#class-map match-all cmap2
Force10(conf-class-map)#match ip access-group acl2
Force10(conf-class-map)#exit
Force10(conf)#policy-map-input pmap
Force10(conf-policy-map-in)#service-queue 7 class-map cmap1
Force10(conf-policy-map-in)#service-queue 4 class-map cmap2
Force10(conf-policy-map-in)#exit
Force10(conf)#interface gig 1/0
Force10(conf-if-gi-1/0)#service-policy input pmap
```

Create a Layer 2 class map

All class maps are Layer 3 by default; you can create a Layer 2 class map by specifying the option **layer2** 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.

1. Create a match-any class map using the command **class-map match-any** or a match-all class map using the command **class-map match-all** from CONFIGURATION mode, and enter the keyword **layer2**.
2. Once you create a class-map, FTOS places you in CLASS MAP mode. From this mode, specify your match criteria using the command **match mac**. Match-any class maps allow up to five access-lists, and match-all class-maps allow only one. You can match against only one VLAN ID.
3. After you specify your match criteria, link the class-map to a queue using the command **service-queue** from POLICY MAP mode.

Determine the order in which ACLs are used to classify traffic

When you link class-maps to queues using the command **service-queue**, FTOS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities). For example, in Figure 31-10, class-map *cmap2* is matched against ingress packets before *cmap1*.

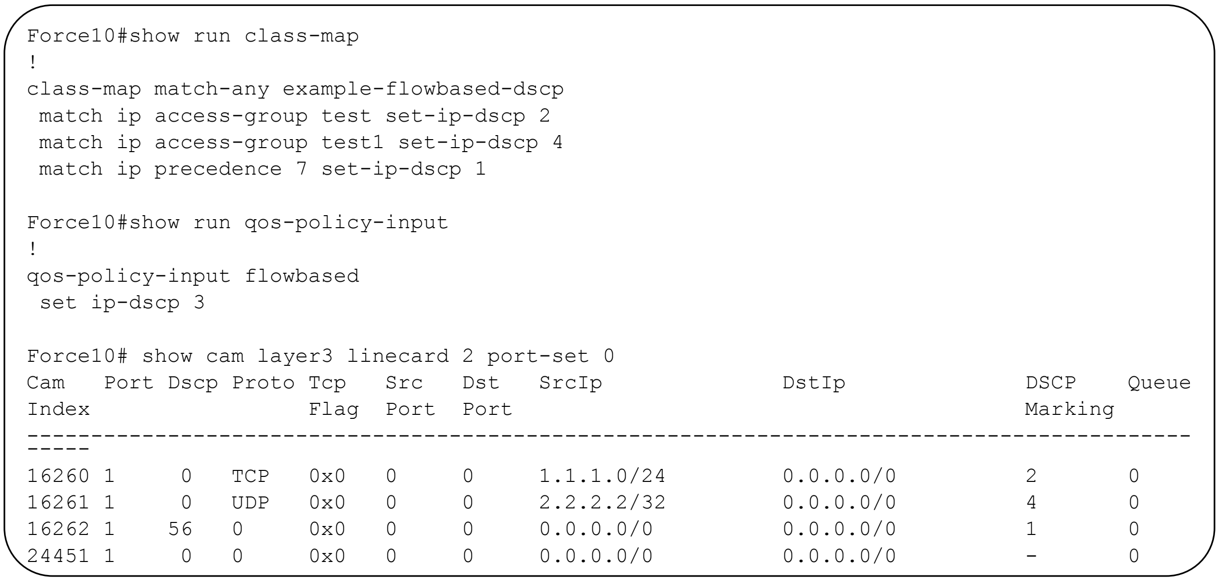
ACLs *acl1* and *acl2* have overlapping rules because the address range 20.1.1.0/24 is within 20.0.0.0/8. Therefore, (without the keyword **order**) packets within the range 20.1.1.0/24 match positive against *cmap1* and are buffered in queue 7, though you intended for these packets to match positive 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, as shown in Figure 31-10. The order can range from 0 to 254. FTOS 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 254.

Set DSCP values for egress packets based on flow

Match-any Layer 3 flows may have several match criteria. All flows that that match at least one of the match criteria are mapped to the same queue since they are in the same class map. Setting a DSCP value from QOS-POLICY-IN mode (see Set a DSCP value for egress packets on page 570) 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 CLASS-MAP mode using the option **set-ip-dscp** with the **match** command so that matching flows within a class map can have *different* DSCP values, as shown in Figure 31-11. The values you set from CLASS-MAP mode override the value you QoS input policy DSCP value, and packets matching the rule are marked with the specified value.

Figure 31-11. Marking Flows in the Same Queue with Different DSCP Values



```
Force10#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

Force10#show run qos-policy-input
!
qos-policy-input flowbased
 set ip-dscp 3

Force10# show cam layer3 linecard 2 port-set 0
Cam   Port Dscp Proto Tcp   Src   Dst   SrcIp              DstIp              DSCP    Queue
Index                 Flag  Port  Port                                        Marking
-------------------------------------------------------------------------------------------
-----
16260 1     0   TCP   0x0   0     0     1.1.1.0/24         0.0.0.0/0          2       0
16261 1     0   UDP   0x0   0     0     2.2.2.2/32         0.0.0.0/0          4       0
16262 1    56   0     0x0   0     0     0.0.0.0/0          0.0.0.0/0          1       0
24451 1     0   0     0x0   0     0     0.0.0.0/0          0.0.0.0/0          -       0
```

Display configured class maps and match criteria

Display all class-maps or a specific class map using the command **show qos class-map** from EXEC Privilege mode.

FTOS 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. Below, traffic is classified in two Queues, 1 and 2. Class-map ClassAF1 is “match any,” and ClassAF2 is “match all”.

```
Force10#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
Force10#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
!
class-map match-all ClassAF2
 match ip access-group AF2
 match ip dscp 18
Force10#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
Force10#show cam layer3-qos interface gigabitethernet 4/49
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
```

Above, 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 the explicit “deny any” rule is removed from all three ACLs, the CAM reflects exactly the desired classification.

```
Force10#show cam layer3-qos interface gigabitethernet 4/49
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.

- Layer 3 QoS input policies allow you to rate police and set a DSCP or dot1p value.
- Layer 2 QoS input policies allow you to rate police and set a dot1p value.

Output QoS policies regulate Layer 3 egress traffic. The regulation mechanisms for output QoS policies are rate limiting, rate shaping, and WRED.

Note: When changing a "service-queue" configuration in a QoS policy map, all QoS rules are deleted and re-added automatically to ensure that the order of the rules is maintained. As a result, the Matched Packets value shown in the "show qos statistics" command is reset.

Create an input QoS policy

To create an input QoS policy:

1. Create a Layer 3 input QoS policy using the command **qos-policy-input** from CONFIGURATION mode. Create a Layer 2 input QoS policy by specifying the keyword **layer2** after the command **qos-policy-input.**
2. Once you create an input QoS policy, do one or more of the following:
   - Configure policy-based rate policing
   - Set a DSCP value for egress packets
   - Set a dot1p value for egress packets

Configure policy-based rate policing

Rate police ingress traffic using the command rate-police from QOS-POLICY-IN mode.

Set a DSCP value for egress packets

Set a DSCP value for egress packets based on ingress QoS classification, as shown in Figure 31-2. The 6 bits that are used for DSCP are also used to identify the queue in which traffic is buffered. When you set a DSCP value, FTOS displays an informational message advising you of the queue to which you should apply the QoS policy (using the command **service-queue** from POLICY-MAP-IN mode). If you apply the QoS policy to a queue *other than* the one specified in the informational message, FTOS replaces the first 3 bits in the DSCP field with the queue ID you specified.

Figure 31-12. Marking DSCP Values for Egress Packets

```
Force10#config
Force10(conf)#qos-policy-input my-input-qos-policy
Force10(conf-qos-policy-in)#set ip-dscp 34
% Info: To set the specified DSCP value 34 (100-010 b) the QoS policy must be mapped to queue
4 (100 b).
Force10(conf-qos-policy-in)#show config
!
qos-policy-input my-input-qos-policy
 set ip-dscp 34
Force10(conf-qos-policy-in)#end

Force10#
```

Set a dot1p value for egress packets

Set a dot1p value for egress packets using the command **set mac-dot1p** from QOS-POLICY-IN mode.

Create an output QoS policy

To create an output QoS policy:

1. Create an output QoS policy using the command **qos-policy-output** from CONFIGURATION mode.
2. Once you configure an output QoS policy, do one or more of the following
   - Configure policy-based rate limiting
   - Configure policy-based rate shaping
   - Allocate bandwidth to queue
   - Specify WRED drop precedence

Configure policy-based rate limiting

Configure policy-based rate limiting is supported only on platform E

Policy-based rate limiting is configured the same way as port-based rate limiting except that the command from QOS-POLICY-OUT mode is **rate-limit** rather than **rate limit** as it is in INTERFACE mode.

Configure policy-based rate shaping

Rate shape egress traffic using the command **rate-shape** from QOS-POLICY-OUT mode.

Allocate bandwidth to queue

The E-Series schedules unicast, multicast, and replication traffic for egress based on the Weighted Fair Queuing algorithm. The C-Series and S-Series schedule packets for egress based on Deficit Round Robin (DRR). These strategies both offer a guaranteed data rate.

To allocate an amount bandwidth to a queue using the command **bandwidth-percentage** on the E-Series.

To allocate bandwidth to queues on the C-Series and S-Series, assign each queue a weight ranging from 1 to 1024, in increments of 2^n , using the command **bandwidth-weight**. Table 31-3 shows the default bandwidth weights for each queue, and their equivalent percentage which is derived by dividing the bandwidth weight by the sum of all queue weights.

Table 31-3. Default Bandwidth Weights for C-Series and S-Series

Queue	Default Weight	Equivalent Percentage
0	1	6.67%
1	2	13.33%
2	4	26.67%
3	8	53.33%

There are two key differences between allocating bandwidth by weight on the C-Series and S-Series and allocating bandwidth by percentage on the E-Series:

1. Assigning a weight to one queue affects the amount of bandwidth that is allocated to other queues. Therefore, whenever you are allocating bandwidth to one queue, Dell Force10 recommends that you evaluate your bandwidth requirements for all other queues as well.
2. Because you are required to choose a bandwidth weight in increments of 2^n you may not be able to achieve exactly a target bandwidth allocation.

Table 31-4 shows an example of choosing bandwidth weights for all four queues to achieve a target bandwidth allocation.

Table 31-4. Assigning Bandwidth Weights for the C-Series and S-Series

Queue	Weight	Equivalent Percentage	Target Allocation
0	1	0.44%	1%
1	64	28.44%	25%
2	128	56.89%	60%
3	32	14.22%	14%

Specify WRED drop precedence

Specify WRED drop precedence is supported only on platform E

Specify a WRED profile to yellow and/or green traffic using the command **wred** from QOS-POLICY-OUT mode. See Apply a WRED profile to traffic on page 579.

Create Policy Maps

There are two types of policy maps: input and output.

Create Input Policy Maps

There are two types of input policy-maps: Layer 3 and Layer 2.

1. Create a Layer 3 input policy map using the command **policy-map-input** from CONFIGURATION mode. Create a Layer 2 input policy map by specifying the keyword **layer2** with the **policy-map-input** command.
2. Once you create an input policy map, do one or more of the following:
   - Apply a class-map or input QoS policy to a queue
   - Apply an input QoS policy to an input policy map
   - Honor DSCP values on ingress packets
   - Honoring dot1p values on ingress packets
3. Apply the input policy map to an interface. See page 60.

FTOS Behavior: On ExaScale, FTOS cannot classify protocol traffic on a Layer 2 interface using Layer 3 policy map. The packets always take the default queue, Queue 0, and cannot be rate-policed.

Apply a class-map or input QoS policy to a queue

Assign an input QoS policy to a queue using the command **service-queue** from POLICY-MAP-IN mode.

Apply an input QoS policy to an input policy map

Apply an input QoS policy to an input policy map using the command **policy-aggregate** from POLICY-MAP-IN mode.

Honor DSCP values on ingress packets

FTOS provides the ability to honor DSCP values on ingress packets using Trust DSCP feature. Enable this feature using the command **trust diffserv** from POLICY-MAP-IN mode. Table 31-5 lists the standard DSCP definitions, and indicates to which queues FTOS maps DSCP values. When Trust DSCP is configured the matched packets and matched bytes counters are not incremented in **show qos statistics**.

Table 31-5. Default DSCP to Queue Mapping

DSCP/CP hex range (XXX)	**DSCP Definition**	**Traditional IP Precedence**	**E-Series Internal Queue ID**	**C-Series Internal Queue ID**	**S-Series Internal Queue ID**	**DSCP/CP** decimal
111XXX		Network Control	7	3	3	48–63
110XXX		Internetwork Control	6	3	3	
101XXX	EF (Expedited Forwarding)	CRITIC/ECP	5	2	2	32–47
100XXX	AF4 (Assured Forwarding)	Flash Override	4	2	2	

DSCP/CP hex range (XXX)	**DSCP Definition**	**Traditional IP Precedence**	**E-Series Internal Queue ID**	**C-Series Internal Queue ID**	**S-Series Internal Queue ID**	**DSCP/CP** decimal
011XXX	AF3	Flash	3	1	1	16–31
010XXX	AF2	Immediate	2	1	1	
001XXX	AF1	Priority	1	0	0	0–15
000XXX	BE (Best Effort)	Best Effort	0	0	0	

Honoring dot1p values on ingress packets

FTOS provides the ability to honor dot1p values on ingress packets with the Trust dot1p feature. Enable Trust dot1p using the command **trust dot1p** from POLICY-MAP-IN mode. Table 31-6 specifies the queue to which the classified traffic is sent based on the dot1p value.

Table 31-6. Default dot1p to Queue Mapping

dot1p	**E-Series Queue ID**	**C-Series Queue ID**	**S-Series Queue ID**
0	2	1	1
1	0	0	0
2	1	0	0
3	3	1	1
4	4	2	2
5	5	2	2
6	6	3	3
7	7	3	3

The dot1p value is also honored for frames on the default VLAN; see Priority-tagged Frames on the Default VLAN.

Fall Back to trust diffserve or dot1p

Fall Back to trust diffserve or dot1p is available only on platforms: E

When using QoS service policies with multiple class maps, you can configure FTOS to use the incoming DSCP or dot1p marking as a secondary option for packet queuing in the event that no match occurs in the class maps.

When class-maps are used, 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 the three class maps:

```
!
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
!
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 above configuration 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.

You can optionally classify packets using their DSCP marking, instead of placing packets in Queue 0, if no match occurs. In the above example, if no match occurs against match-all qos-BE1, the classification logic continues:

4. Queue the packet according to the DSCP marking. The DSCP to Queue mapping will be as per the Table 31-5.

The behavior is similar for **trust dot1p fallback** in a Layer2 input policy map; the dot1p-to-queue mapping is according to Table 31-6.

To enable Fall Back to trust diffserve or dot1p:

Task	Command Syntax	Command Mode
Classify packets according to their DSCP value as a secondary option in case no match occurs against the configured class maps.	**trust** {**diffserve | dot1p**} **fallback**	POLICY-MAP-IN

Mapping dot1p values to service queues

Mapping dot1p values to service queues is available only on platforms: C S

On the C-Series and S-Series all traffic is by default mapped to the same queue, Queue 0. If you honor dot1p on ingress, then you can create service classes based the queueing strategy in Table 31-6 using the command **service-class dynamic dot1p** from INTERFACE mode. You may apply this queuing strategy globally by entering this command from CONFIGURATION mode.

- All dot1p traffic is mapped to Queue 0 unless **service-class dynamic dot1p** is enabled on an interface or globally.
- Layer 2 or Layer 3 service policies supersede dot1p service classes.

Guaranteeing bandwidth to dot1p-based service queues

Guarantee a minimum bandwidth to queues globally from CONFIGURATION mode with the command **service-class bandwidth-weight**. The command is applied in the same way as the bandwidth-weight command in an output QoS policy (see Allocate bandwidth to queue on page 571). The **bandwidth-weight** command in QOS-POLICY-OUT mode supersedes the **service-class bandwidth-weight command.**

Apply an input policy map to an interface

Apply an input policy map to an interface using the command **service-policy input** from INTERFACE mode. Specify the keyword **layer2** if the policy map you are applying a Layer 2 policy map; in this case, the INTERFACE must be in switchport mode. You can apply the same policy map to multiple interfaces, and you can modify a policy map after you apply it.

- You cannot apply a class-map and QoS policies to the same interface.
- You cannot apply an input Layer 2 QoS policy on an interface you also configure with **vlan-stack access.**
- If you apply a service policy that contains an ACL to more than one interface, FTOS uses ACL optimization to conserves CAM space. The ACL Optimization behavior detects when an ACL already exists in the CAM and rather than writing it to the CAM multiple times.

Create Output Policy Maps

Create Output Policy Maps is supported only on platform E

1. Create an output policy map using the command **policy-map-output** from CONFIGURATION mode.
2. Once you create an output policy map, do one or more of the following:
   - Apply an output QoS policy to a queue
   - Specify an aggregate QoS policy
   - Apply an output policy map to an interface
3. Apply the policy map to an interface. See page 61.

Apply an output QoS policy to a queue

Apply an output QoS policy to queues using the command **service-queue** from INTERFACE mode.

Specify an aggregate QoS policy

Specify an aggregate QoS policy using the command **policy-aggregate** from POLICY-MAP-OUT mode.

Apply an output policy map to an interface

Apply an input policy map to an interface using the command **service-policy output** from INTERFACE mode. You can apply the same policy map to multiple interfaces, and you can modify a policy map after you apply it.

QoS Rate Adjustment

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)

By default, while rate limiting, policing, and shaping, FTOS 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. 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.

Task	Command Syntax	Command Mode
Include a specified number of bytes of packet overhead to include in rate limiting, policing, and shaping calculations. For example, to include the Preamble and SFD, enter **qos-rate-adjust** 8. For variable length overhead fields you must know the number of bytes you want to include.	**qos-rate-adjust** *overhead-bytes* Default: Disabled C-Series and S-Series Range: 1-31 E-Series Range: 1-144	CONFIGURATION

Strict-priority Queueing

You can assign strict-priority to one unicast queue, 1-7, using the command **strict-priority** from CONFIGURATION mode. Strict-priority means that FTOS dequeues all packets from the assigned queue before servicing any other queues.

- The **strict-priority** supersedes **bandwidth-percentage** an **bandwidth-weight** percentage configurations.
- A queue with strict-priority can starve other queues in the same port-pipe.
- On the E-Series, this configuration is applied to the queue on both ingress and egress.

Weighted Random Early Detection

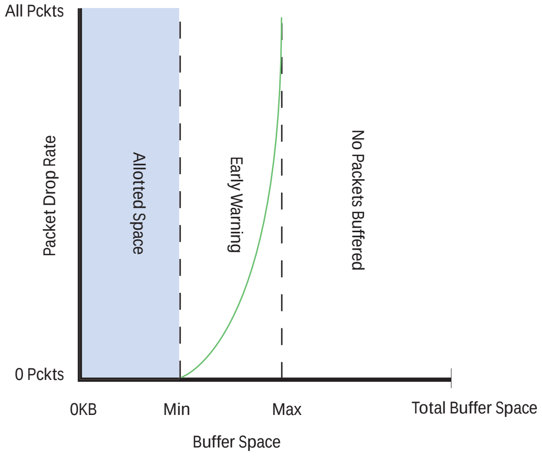
Weighted Random Early Detection is supported only on platform E

Weighted Random Early Detection (WRED) congestion avoidance mechanism that 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 BTM (ingress or egress) can be consumed by only one or a few types of traffic, leaving no space for other types. A WRED profile can be applied 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 will be dropped randomly at an exponential rate until the maximum threshold is reached (Figure 31-13); this is the “early detection” part of WRED. If the maximum threshold—2000KB, for example—is reached, then all incoming packets are dropped until less than 2000KB of buffer space is consumed by the specified traffic.

Figure 31-13. Packet Drop Rate for WREDI



fnC0045mp

You can create a custom WRED profile or use on of the five pre-defined profiles listed in Table 31-7.

Table 31-7. Pre-defined WRED Profiles

Default Profile Name	Minimum Threshold	Maximum Threshold
wred_drop	0	0
wred_ge_y	1024	2048
wred_ge_g	2048	4096
wred_teng_y	4096	8192
wred_teng_g	8192	16384

Create WRED Profiles

To create a WRED profile:

1. Create a WRED profile using the command **wred** from CONFIGURATION mode.
2. The command **wred** places you in WRED mode. From this mode, specify minimum and maximum threshold values using the command **threshold**.

Apply a WRED profile to traffic

Once you create a WRED profile you must specify to which traffic FTOS should apply the profile.

FTOS 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 Force10 uses the first three bits of this field (DP) to determine the drop precedence. DP values of 110 and 100 map to yellow, and all other values map to green. If you do not configure FTOS to honor DSCP values on ingress (Honor DSCP values on ingress packets on page 573) see all traffic defaults to green drop precedence.

Assign a WRED profile to either yellow or green traffic from QOS-POLICY-OUT mode using the command **wred**.

Configure WRED for Storm Control

Configure WRED for Storm Control is supported only on platform E

Storm control limits the percentage of the total bandwidth that broadcast traffic can consume on an interface (if configured locally) or on all interfaces (if configured globally). For **storm-control broadcast 50 out**, the total bandwidth that broadcast traffic can consume on egress on a 1Gbs interface is 512Mbs. The method by which packets are selected to be dropped is the "tail-drop" method, where packets exceeding the specified rate are dropped.

WRED can be used in combination with storm control to regulate broadcast and unknown-unicast traffic. This feature is available through an additional option in command **storm-control** [**broadcast** | **unknown-unicast**] at CONFIGURATION. See the *FTOS Command Line Reference* for information on using this command.

Using the command **storm-control broadcast 50 out wred-profile**, for example, first the total bandwidth that broadcast traffic can consume is reduced to 50% of line rate. Even though broadcast traffic is restricted, the rate of outgoing broadcast traffic might be greater than other traffic, and if so, broadcast packets would consume too much buffer space. So, the **wred-profile** option is added to limit the amount of buffer space that broadcast traffic can consume.

Display Default and Configured WRED Profiles

Display default and configured WRED profiles and their threshold values using the command **show qos wred-profile** from EXEC mode, as shown in Figure 31-14.

Figure 31-14. Displaying WRED Profiles

```
Force10#show qos wred-profile

Wred-profile-name        min-threshold   max-threshold
wred_drop                0               0
wred_ge_y                1000            2000
wred_ge_g                2000            4000
wred_teng_y              4000            8000
wred_teng_g              8000            16000
```

Display WRED Drop Statistics

Display the number of packets FTOS dropped by WRED Profile using the command **show qos statistics** from EXEC Privilege mode, as shown in Figure 31-15.

Figure 31-15. show qos statistics Command Example

```
 Force10#show qos statistics wred-profile
Interface Gi 5/11
Queue#  Drop-statistic  WRED-name        Min      Max      Dropped Pkts

  0     Green           WRED1            10       100      51623
        Yellow          WRED2            20       100      51300
        Out of Profile                                     0
  1     Green           WRED1            10       100      52082
        Yellow          WRED2            20       100      51004
        Out of Profile                                     0
  2     Green           WRED1            10       100      50567
        Yellow          WRED2            20       100      49965
        Out of Profile                                     0
  3     Green           WRED1            10       100      50477
        Yellow          WRED2            20       100      49815
        Out of Profile                                     0
  4     Green           WRED1            10       100      50695
        Yellow          WRED2            20       100      49476
        Out of Profile                                     0
  5     Green           WRED1            10       100      50245
        Yellow          WRED2            20       100      49535
        Out of Profile                                     0
  6     Green           WRED1            10       100      50033
        Yellow          WRED2            20       100      49595
        Out of Profile                                     0
  7     Green           WRED1            10       100      50474
        Yellow          WRED2            20       100      49522
        Out of Profile                                     0
Force10#
```

FTOS Behavior: The C-Series fetches the per-queue packet count via class-maps. The count is the number of packets matching the ACL entries in class-map. Every time the class-map or policy-map is modified, the ACL entries are re-written to the Forwarding Processor, and the queue statistics are cleared. This behavior is different from the E-Series. The E-Series fetches the packet count directly from counters at each queue, which allows queue statistics to persist until explicitly cleared via the CLI.

Allocating Bandwidth to Multicast Queues

Allocating Bandwidth to Multicast Queues is supported on platform: E

The E-Series has 128 multicast queues per port-pipe, which are transparent, and eight unicast queues per port. You can allocate a specific bandwidth percentage per port-pipe to multicast traffic using the command **queue egress multicast bandwidth-percentage** from CONFIGURATION mode.

- If you configure **bandwidth-percentage** for unicast only, 1/8 of the port bandwidth is reserved for multicast, and the remaining bandwidth is distributed based on your configuration.
- If you configure multicast bandwidth, after assigning the specified amount of bandwidth to multicast the remaining bandwidth is distributed according to the WFQ algorithm.
- If you configure **bandwidth-percentage** for both unicast and multicast, then bandwidth is assigned based on your configuration for multicast *then* unicast (based on the remaining available bandwidth), and the remaining bandwidth is distributed among the other queues.

For example, if you configure 70% bandwidth to multicast, 80% bandwidth to one queue in unicast and 0 % to all remaining unicast queues, then first, FTOS assigns 70% bandwidth to multicast, then FTOS derives the 80% bandwidth for unicast from the remaining 30% of total bandwidth.

Pre-calculating Available QoS CAM Space

Pre-calculating Available QoS CAM Space is supported on platforms: C E S

Before version 7.3.1 there was no way to measure the number of CAM entries a policy-map would consume (the number of CAM entries that a rule uses is not predictable; 1 to 16 entries might be used per rule depending upon its complexity). Therefore, it was possible to apply to an interface a policy-map that requires more entries than are available. In this case, the system writes as many entries as possible, and then generates an CAM-full error message (Message 1). The partial policy-map configuration might cause unintentional system behavior.

Message 1 QoS CAM Region Exceeded

```
%EX2YD:12 %DIFFSERV-2-DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for
class 2 (Gi 12/20) entries on portpipe 1 for linecard 12
%EX2YD:12 %DIFFSERV-2-
DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for class 5 (Gi 12/22)
entries on portpipe 1 for linecard 12
```

The command **test cam-usage** enables you to verify that there are enough available CAM entries *before* applying a policy-map to an interface so that you avoid exceeding the QoS CAM space and partial configurations. This command measures the size of the specified policy-map and compares it to the available CAM space in a partition for a specified port-pipe.

Test the policy-map size against the CAM space for a specific port-pipe or all port-pipes using these commands:

- **test cam-usage service-policy input** *policy-map* {**linecard | stack-unit** } *number* **port-set** *number*
- **test cam-usage service-policy input** *policy-map* {**linecard | stack-unit** } *all*

The output of this command, shown in Figure 31-16, displays:

- the estimated number of CAM entries the policy-map will consume
- whether or not the policy-map can be applied
- the number of interfaces in a port-pipe to which the policy-map can be applied

Specifically:

- **Available CAM** is the available number of CAM entries in the specified CAM partition for the specified line card or stack-unit port-pipe.
- **Estimated CAM** is the estimated number of CAM entries that the policy will consume when it is applied to an interface.

- **Status** indicates whether or not the specified policy-map can be completely applied to an interface in the port-pipe.
  - **Allowed** indicates that the policy-map can be applied because the estimated number of CAM entries is less or equal to the available number of CAM entries. The number of interfaces in the port-pipe to which the policy-map can be applied is given in parenthesis.
  - **Exception** indicates that the number of CAM entries required to write the policy-map to the CAM is greater than the number of available CAM entries, and therefore the policy-map cannot be applied to an interface in the specified port-pipe.

Note: The command **show cam-usage** provides much of the same information as **test cam-usage**, but whether or not a policy-map can be successfully applied to an interface cannot be determined without first measuring how many CAM entries the policy-map would consume; the command **test cam-usage** is useful because it provides this measurement.

Figure 31-16. test cam-usage Command Example

```
Force10# test cam-usage service-policy input pmap_l2 linecard 0 port-set 0

Linecard | Port-pipe | CAM Partition | Available CAM | Estimated CAM | Status
===============================================================================
0           0            L2ACL           500            200        Allowed(2)
```

Viewing QoS CAM Entries

Viewing QoS CAM Entries is supported only on platform E

- View Layer 2 QoS CAM entries using the command **show cam layer3-qos** from EXEC Privilege mode.
- View Layer 3 QoS CAM entries using the command **show cam layer2-qos** from EXEC Privilege mode.

32

Routing Information Protocol

Routing Information Protocol is supported only on platforms: C E S

RIP is supported on the S-Series following the release of FTOS version 7.8.1.0, and on the C-Series with FTOS versions 7.6.1.0 and after.

Routing Information Protocol (RIP) is based on a distance-vector algorithm, it tracks distances or hop counts to nearby routers when establishing network connections.

- Protocol Overview on page 585
- Implementation Information on page 586
- Configuration Information on page 586
- RIP Configuration Example on page 594

RIP protocol standards are listed in the Appendix 47, 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 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 VLSM or 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

FTOS supports both versions of RIP and allows you to configure one version globally and the other version or both versions on the interfaces. The C-Series and E-Series both support 1,000 RIP routes.

Table 32-1 displays the defaults for RIP in FTOS.

Table 32-1. RIP Defaults in FTOS

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 FTOS. 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. All devices within the RIP network must be configured to support RIP if they are to participate in the RIP.

Configuration Task List for RIP

- Enable RIP globally on page 587 (mandatory)
- Configure RIP on interfaces on page 588 (optional)
- Control RIP routing updates on page 588 (optional)

- Set send and receive version on page 590 (optional)
- Generate a default route on page 592 (optional)
- Control route metrics on page 593 (optional)
- Summarize routes on page 593 (optional)
- Control route metrics on page 593
- Debug RIP on page 594

For a complete listing of all commands related to RIP, refer to the *FTOS Command Reference.*

Enable RIP globally

By default, RIP is not enabled in FTOS. To enable RIP, use the following commands in sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**router rip**	CONFIGURATION	Enter ROUTER RIP mode and enable the RIP process on FTOS.
2	**network** *ip-address*	ROUTER RIP	Assign an IP network address as a RIP network to exchange routing information. You can use this command multiple times to exchange RIP information with as many RIP networks as you want.

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 FTOS default is to send RIPv1, and to receive RIPv1 and RIPv2. To change the RIP version globally, use the **version** command in the ROUTER RIP mode.

When RIP is enabled, you can view the global RIP configuration by using the **show running-config** command in the EXEC mode or the **show config** command (Figure) in the ROUTER RIP mode.

Figure 32-1. show config Command Example in ROUTER RIP mode

```
Force10(conf-router_rip)#show config
!
router rip
 network 10.0.0.0
Force10(conf-router_rip)#
```

When the RIP process has learned the RIP routes, use the **show ip rip database** command in the EXEC mode to view those routes (Figure 385).

Figure 32-2. show ip rip database Command Example (Partial)

```
Force10#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 the 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 are enabled and configured 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.

Control 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, you must 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, in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
neighbor *ip-address*	ROUTER RIP	Define a specific router to exchange RIP information between it and the Dell Force10 system. You can use this command multiple times to exchange RIP information with as many RIP networks as you want.
passive-interface *interface*	ROUTER RIP	Disable a specific interface from sending or receiving RIP routing information.

Another method of controlling RIP (or any routing protocol) routing information is to filter the information through a prefix list. A prefix lists is applied to incoming or outgoing routes. Those routes must meet the conditions of the prefix list; if not, FTOS drops the route. Prefix lists are globally applied on all interfaces running RIP. Configure the prefix list in the PREFIX LIST mode prior to assigning it to the RIP process.

For configuration information on prefix lists, see Chapter 17, IP Access Control Lists, Prefix Lists, and Route-maps, on page 47.

To apply prefix lists to incoming or outgoing RIP routes, use the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
distribute-list *prefix-list-name* **in**	ROUTER RIP	Assign a configured prefix list to all incoming RIP routes.
distribute-list *prefix-list-name* **out**	ROUTER RIP	Assign a configured prefix list to all outgoing RIP routes.

In addition to filtering routes, you can add routes from other routing instances or protocols to the RIP process. With the **redistribute** command syntax, you can include OSPF, static, or directly connected routes in the RIP process.

To add routes from other routing instances or protocols, use any of the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
redistribute {**connected** | **static**} [**metric** *metric-value*] [**route-map** *map-name*]	ROUTER RIP	Include directly connected or user-configured (static) routes in RIP. • *metric* range: 0 to 16 • *map-name*: name of a configured route map.

Command Syntax	Command Mode	Purpose
redistribute isis [**level-1** | **level-1-2** | **level-2**] [**metric** *metric-value*] [**route-map** *map-name*]	ROUTER RIP	Include IS-IS routes in RIP. • *metric* range: 0 to 16 • *map-name*: name of a configured route map. **Note:** IS-IS is not supported on the S-Series platform.
redistribute ospf *process-id* [**match external** {**1 | 2**} | **match internal**] [**metric** *value*] [**route-map** *map-name*]	ROUTER RIP	Include specific OSPF routes in RIP. Configure the following parameters: • *process-id* range: 1 to 65535 • *metric* range: 0 to 16 • *map-name*: name of a configured route map.

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

Set send and receive version

To specify the RIP version, use the **version** command in the 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 the INTERFACE mode.

To change the RIP version globally in FTOS, use the following command in the ROUTER RIP mode:

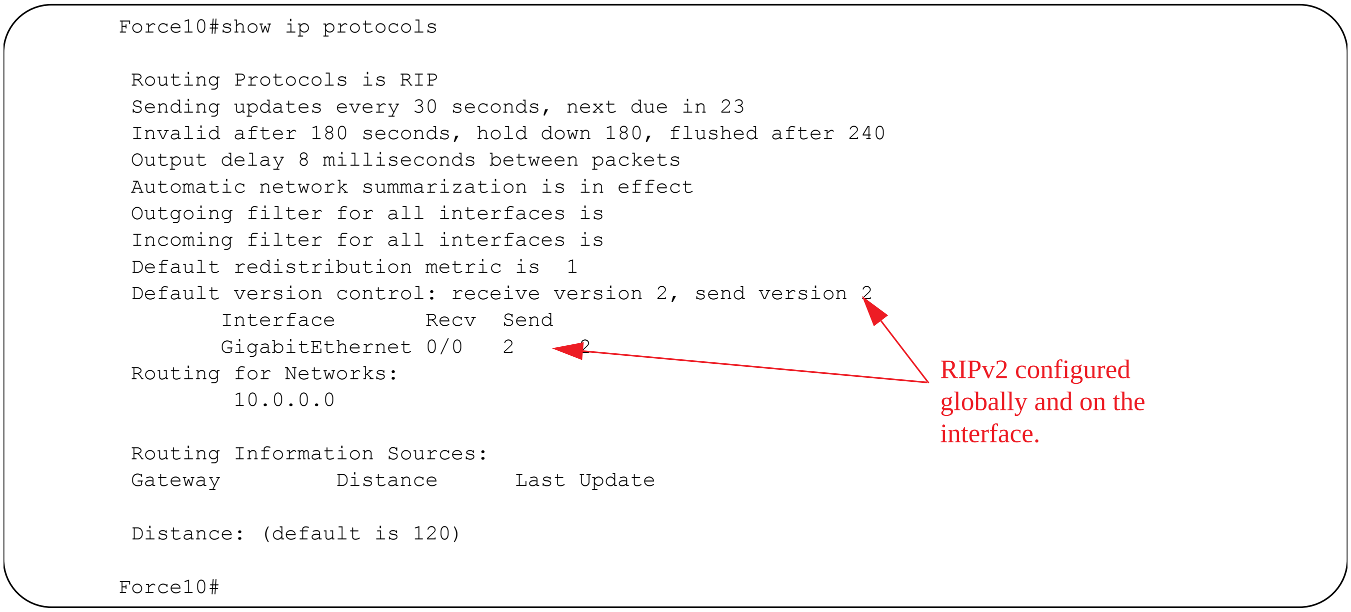
Command Syntax	Command Mode	Purpose
version {**1** | **2**}	ROUTER RIP	Set the RIP version sent and received on the system.

You can set one RIP version globally on the 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.

Use the **show config** command in the ROUTER RIP mode to see whether the **version** command is configured. You can also use the **show ip protocols** command in the EXEC mode to view the routing protocols configuration.

Figure 32-3 shows an example of the RIP configuration after the ROUTER RIP mode **version** command is set to RIPv2. When the ROUTER RIP mode **version** command is set, the interface (GigabitEthernet 0/0) participating in the RIP process is also set to send and receive RIPv2.

Figure 32-3. show ip protocols Command Example



```
Force10#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)

Force10#
```

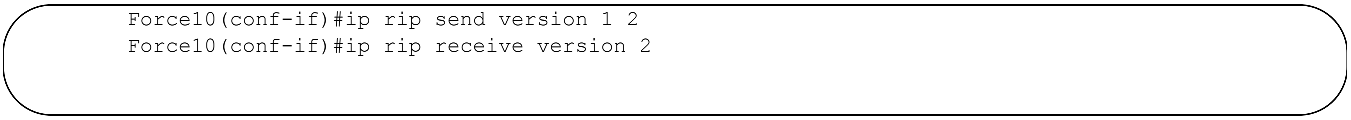
RIPv2 configured globally and on the interface.

To configure the interfaces to send or receive different RIP versions from the RIP version configured globally, use either of the following commands in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip rip receive version [**1**] [**2**]	INTERFACE	Set the RIP version(s) received on that interface.
ip rip send version [**1**] [**2**]	INTERFACE	Set the RIP version(s) sent out on that interface.

To configure an interface to receive or send both versions of RIP, include 1 and 2 in the command syntax. Figure 32-4 displays the command syntax for sending both RIPv1 and RIPv2 and receiving only RIPv2.

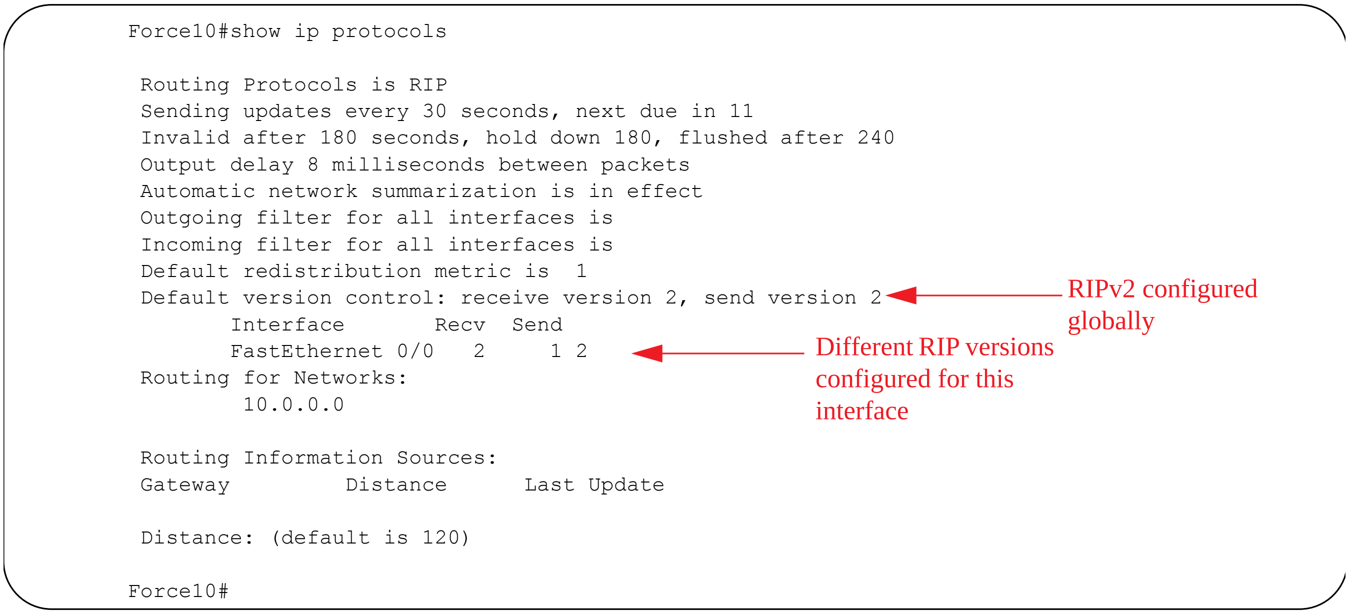
Figure 32-4. Configuring an interface to send both versions of RIP



```
Force10(conf-if)#ip rip send version 1 2
Force10(conf-if)#ip rip receive version 2
```

The **show ip protocols** command example Figure 32-5 confirms that both versions are sent out that interface. This interface no longer sends and receives the same RIP versions as FTOS does globally.

Figure 32-5. show ip protocols Command Example



```
Force10#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)

Force10#
```

RIPv2 configured globally

Different RIP versions configured for this interface

Generate 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 the ROUTER RIP mode to generate a default route into RIP. In FTOS, default routes received in RIP updates from other routes are advertised if the **default-information originate** command is configured.

To configure FTOS to generate a default route, use the following command in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
default-information originate [**always**] [**metric** *value*] [**route-map** *route-map-name*]	ROUTER RIP	Specify the generation of a default route in RIP. Configure the following parameters: • **always**: enter this keyword to always generate a default route. • *value* range: 1 to 16. • *route-map-name*: name of a configured route map.

Use the **show config** command in the ROUTER RIP mode to confirm that the default route configuration is completed.

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 the 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 command **autosummary** requires no other configuration commands. To disable automatic route summarization, in the ROUTER RIP mode, enter **no autosummary**.

Note: If the **ip split-horizon** command is enabled on an interface, then the system does not advertise the summarized address.

Control 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, you must 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. Use the command to assign different weights to routes so that the ones with the lower weight or administrative distance assigned are preferred.

To set route metrics, use either of the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
distance *weight* [*ip-address mask* [*access-list-name*]]	ROUTER RIP	Apply a weight to all routes or a specific route and ACL. Configure the following parameters: • *weight* range: 1 to 255 (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*: name of a configured IP ACL.
offset *access-list-name* {**in** | **out**} *offset* [*interface*]	ROUTER RIP	Apply an additional number to the incoming or outgoing route metrics. Configure the following parameters: • *access-list-name*: the name of a configured IP ACL • *offset* range: 0 to 16. • *interface*: the type, slot, and number of an interface.

Use the **show config** command in the ROUTER RIP mode to view configuration changes.

Debug RIP

The **debug ip rip** command enables RIP debugging. When debugging is enabled, you can view information on RIP protocol changes or RIP routes.

To enable RIP debugging, use the following command in the EXEC privilege mode:

Command Syntax	Command Mode	Purpose
debug ip rip [*interface* | **database** | **events** | **trigger**]	EXEC privilege	Enable debugging of RIP.

Figure 32-6 shows the confirmation when the debug function is enabled.

Figure 32-6. debug ip rip Command Example

```
Force10#debug ip rip
RIP protocol debug is ON
Force10#
```

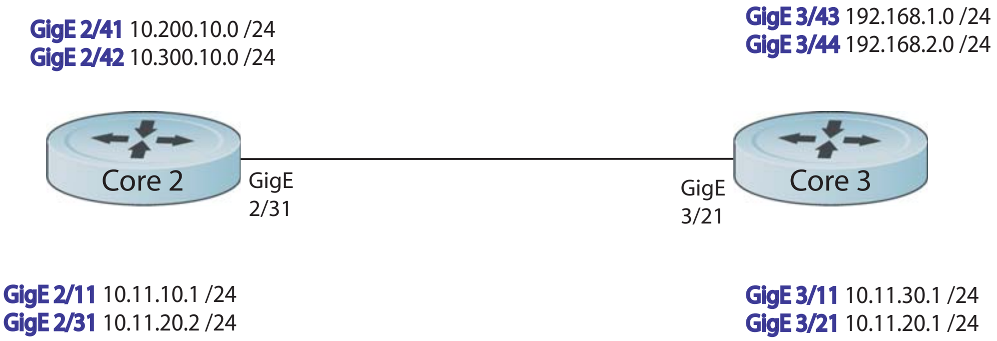
To disable RIP, use the **no debug ip rip** command.

RIP Configuration Example

The example in this section shows the command sequence to configure RIPv2 on the two routers shown in Figure 32-7 — “Core 2” and “Core 3”. The host prompts used in the example screenshots reflect those names. The screenshots are divided into the following groups of command sequences:

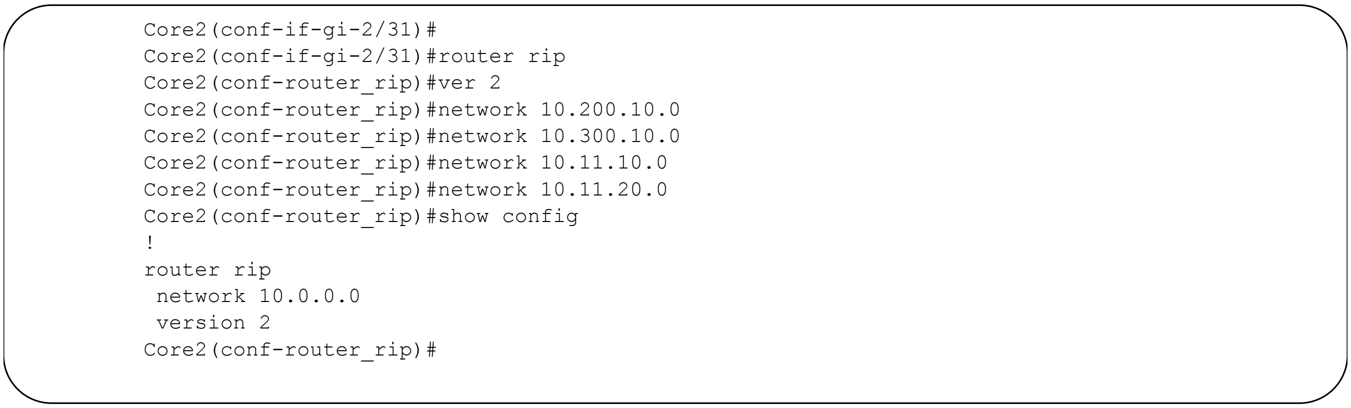
- Configuring RIPv2 on Core 2 on page 595
- Core 2 Output on page 595
- RIP Configuration on Core 3 on page 597
- Core 3 RIP Output on page 597
- RIP Configuration Summary on page 599

Figure 32-7. RIP Topology Example



Configuring RIPv2 on Core 2

Figure 32-8. Configuring RIPv2 on Core 2



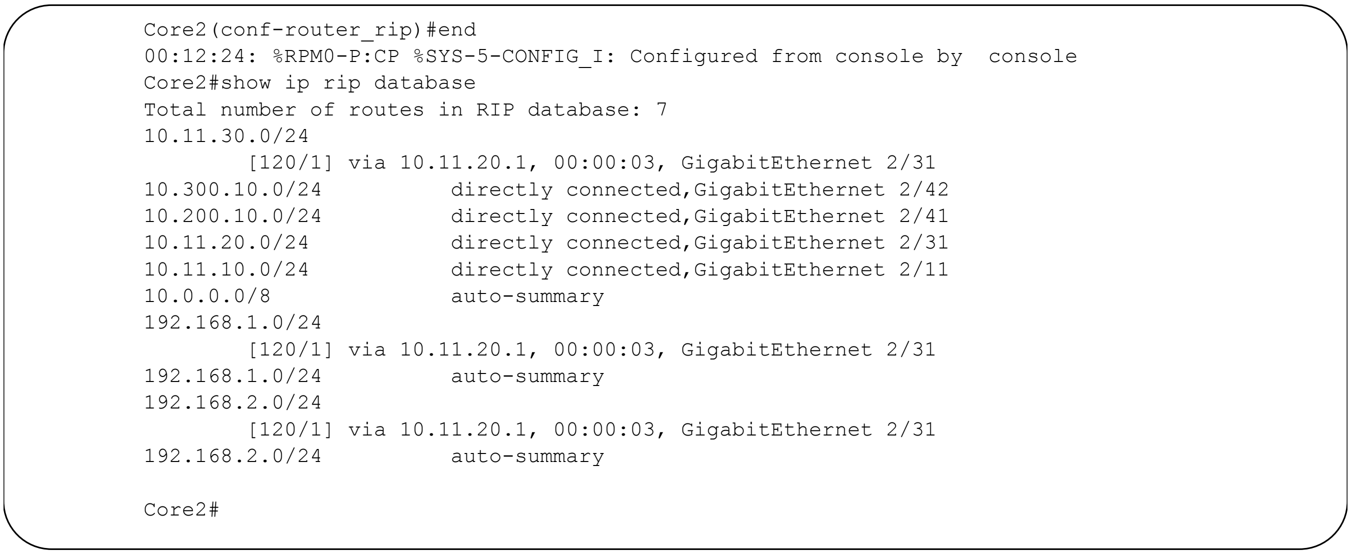
```
Core2(conf-if-gi-2/31)#
Core2(conf-if-gi-2/31)#router rip
Core2(conf-router_rip)#ver 2
Core2(conf-router_rip)#network 10.200.10.0
Core2(conf-router_rip)#network 10.300.10.0
Core2(conf-router_rip)#network 10.11.10.0
Core2(conf-router_rip)#network 10.11.20.0
Core2(conf-router_rip)#show config
!
router rip
 network 10.0.0.0
 version 2
Core2(conf-router_rip)#
```

Core 2 Output

The screenshots in this section are:

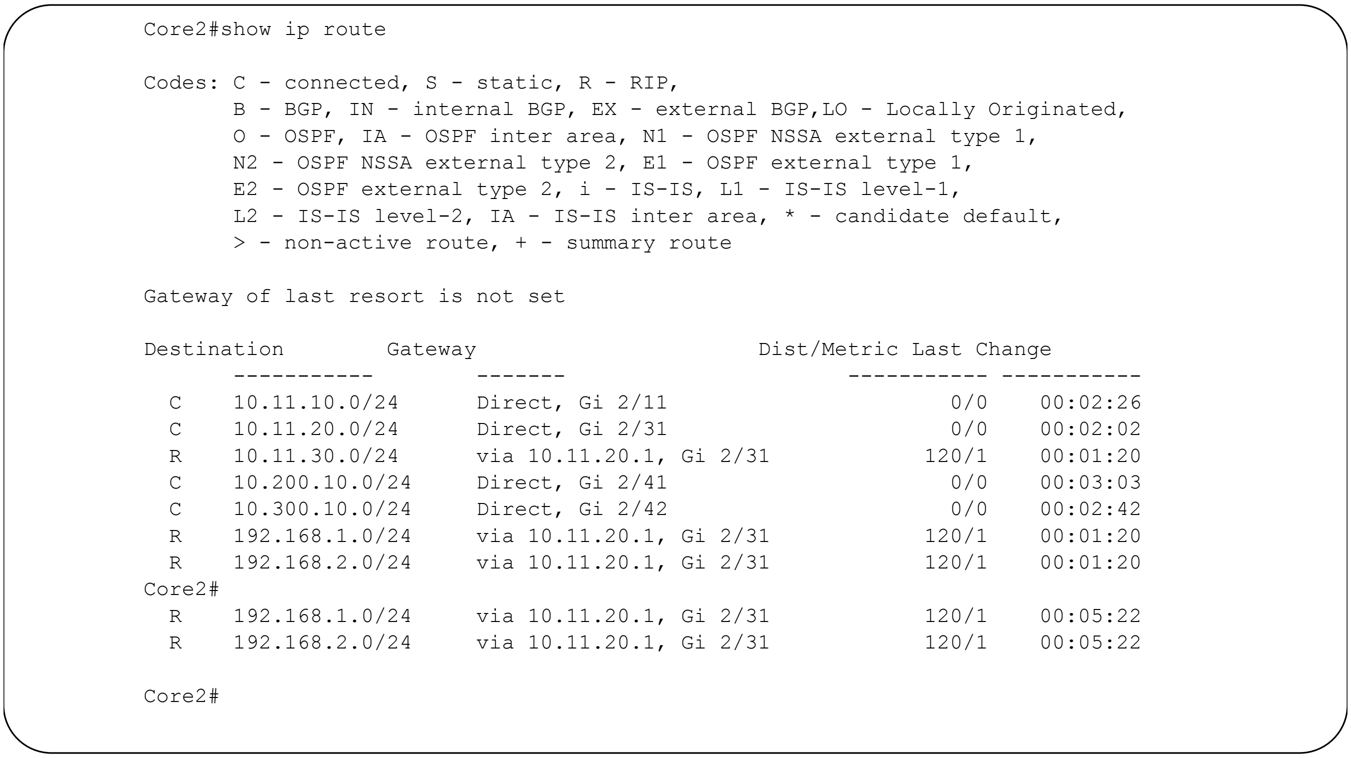
- Figure 32-9: Using **show ip rip database** command to display Core 2 RIP database
- Figure 32-10: Using **show ip route** command to display Core 2 RIP setup
- Figure 32-11: Using **show ip protocols** command to display Core 2 RIP activity

Figure 32-9. Example of RIP Configuration Response from Core 2



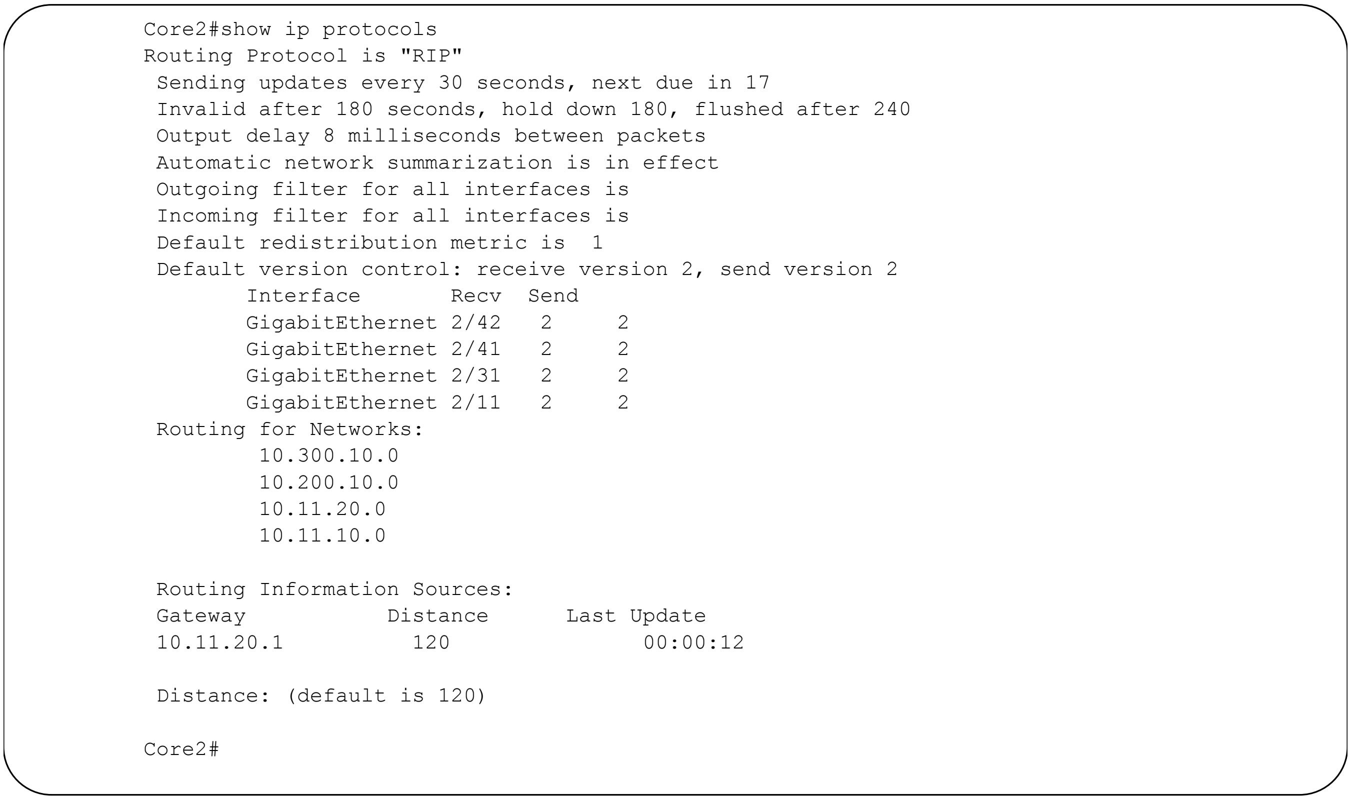
```
Core2(conf-router_rip)#end
00:12:24: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by  console
Core2#show ip rip database
Total number of routes in RIP database: 7
10.11.30.0/24
        [120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
10.300.10.0/24          directly connected,GigabitEthernet 2/42
10.200.10.0/24          directly connected,GigabitEthernet 2/41
10.11.20.0/24           directly connected,GigabitEthernet 2/31
10.11.10.0/24           directly connected,GigabitEthernet 2/11
10.0.0.0/8              auto-summary
192.168.1.0/24
        [120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
192.168.1.0/24          auto-summary
192.168.2.0/24
        [120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
192.168.2.0/24          auto-summary

Core2#
```

Figure 32-10. Using show ip route Command to Show RIP Configuration on Core 2



```
Core2#show ip route

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  C    10.11.10.0/24      Direct, Gi 2/11                      0/0    00:02:26
  C    10.11.20.0/24      Direct, Gi 2/31                      0/0    00:02:02
  R    10.11.30.0/24      via 10.11.20.1, Gi 2/31            120/1    00:01:20
  C    10.200.10.0/24     Direct, Gi 2/41                      0/0    00:03:03
  C    10.300.10.0/24     Direct, Gi 2/42                      0/0    00:02:42
  R    192.168.1.0/24     via 10.11.20.1, Gi 2/31            120/1    00:01:20
  R    192.168.2.0/24     via 10.11.20.1, Gi 2/31            120/1    00:01:20
Core2#
  R    192.168.1.0/24     via 10.11.20.1, Gi 2/31            120/1    00:05:22
  R    192.168.2.0/24     via 10.11.20.1, Gi 2/31            120/1    00:05:22

Core2#
```

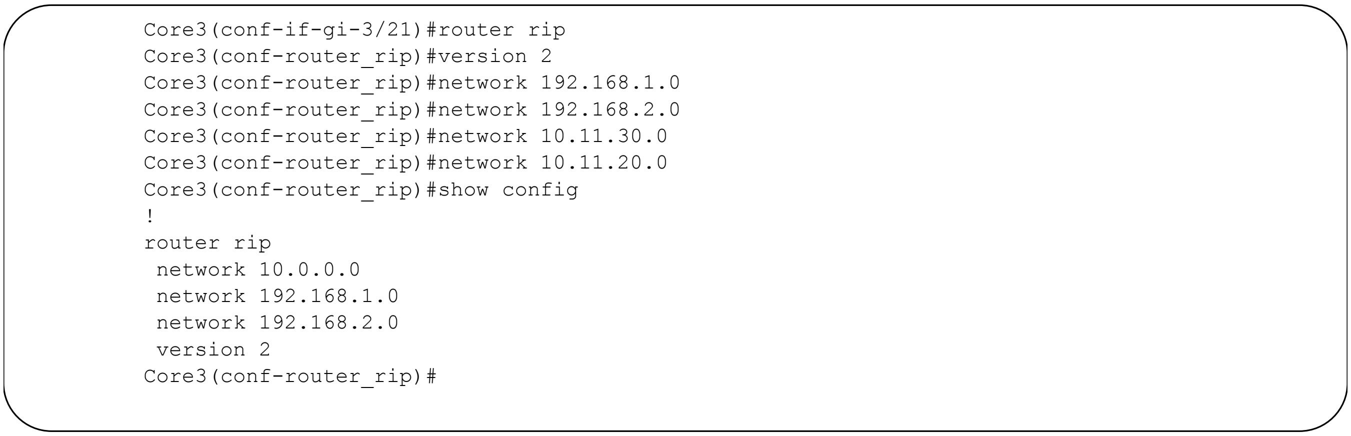
Figure 32-11. Using show ip protocols Command to Show RIP Configuration Activity on Core 2



```
Core2#show ip protocols
Routing Protocol is "RIP"
 Sending updates every 30 seconds, next due in 17
 Invalid after 180 seconds, hold down 180, flushed after 240
 Output delay 8 milliseconds between packets
 Automatic network summarization is in effect
 Outgoing filter for all interfaces is
 Incoming filter for all interfaces is
 Default redistribution metric is  1
 Default version control: receive version 2, send version 2
        Interface       Recv  Send
        GigabitEthernet 2/42   2     2
        GigabitEthernet 2/41   2     2
        GigabitEthernet 2/31   2     2
        GigabitEthernet 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 Core 3

Figure 32-12. RIP Configuration on Core 3



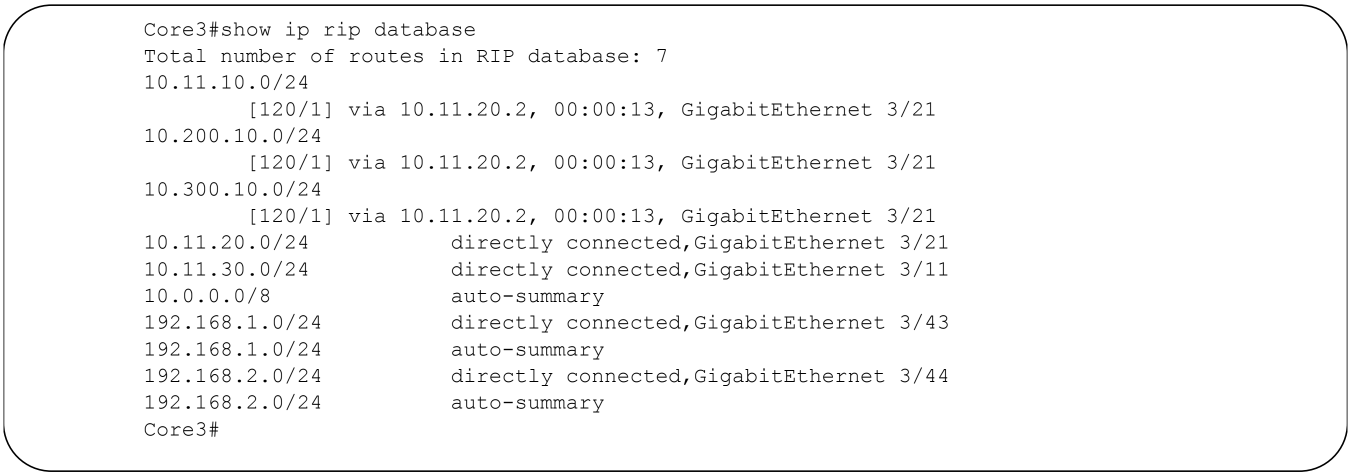
```
Core3(conf-if-gi-3/21)#router rip
Core3(conf-router_rip)#version 2
Core3(conf-router_rip)#network 192.168.1.0
Core3(conf-router_rip)#network 192.168.2.0
Core3(conf-router_rip)#network 10.11.30.0
Core3(conf-router_rip)#network 10.11.20.0
Core3(conf-router_rip)#show config
!
router rip
 network 10.0.0.0
 network 192.168.1.0
 network 192.168.2.0
 version 2
Core3(conf-router_rip)#
```

Core 3 RIP Output

The screenshots in this section are:

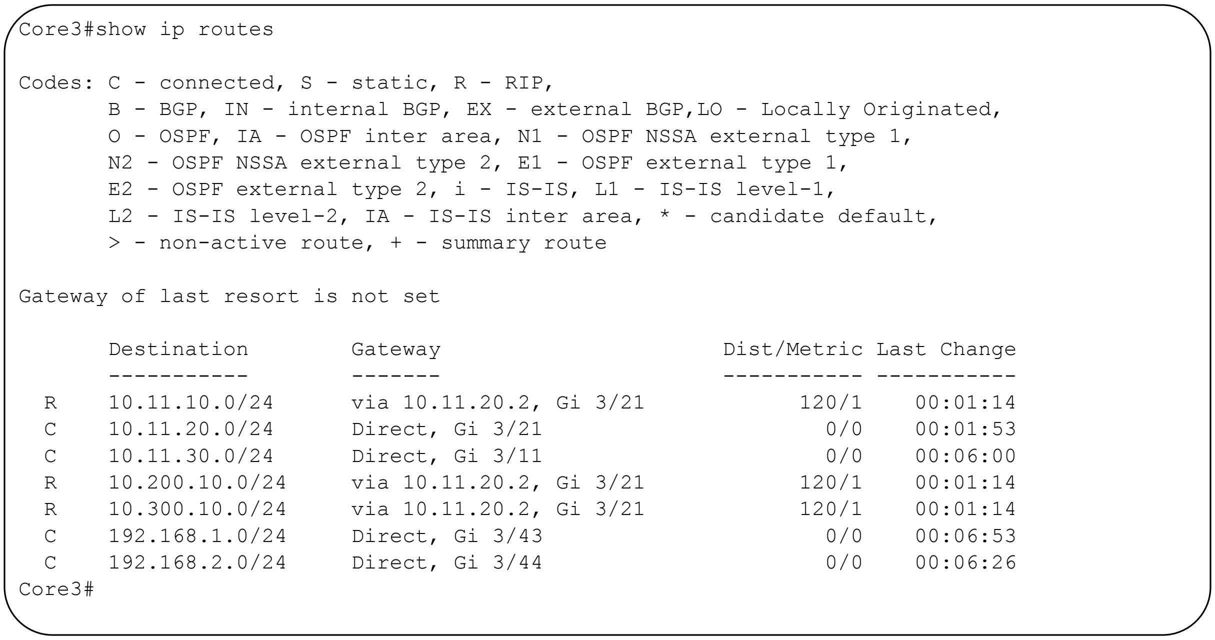
- Figure 32-13: Using **show ip rip database** command to display Core 3 RIP database
- Figure 32-14: Using **show ip route** command to display Core 3 RIP setup
- Figure 32-15: Using **show ip protocols** command to display Core 3 RIP activity

Figure 32-13. Using show ip rip database Command for Core 3 RIP Setup



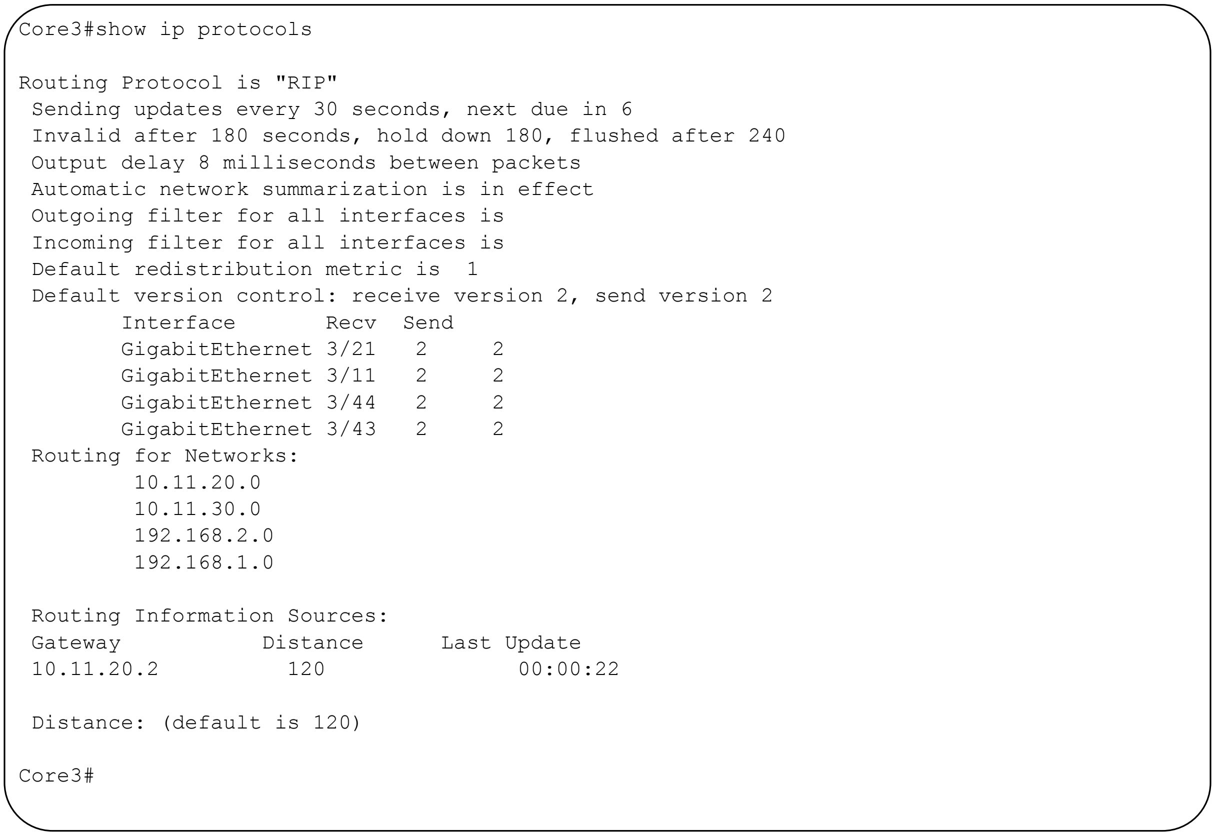
```
Core3#show ip rip database
Total number of routes in RIP database: 7
10.11.10.0/24
        [120/1] via 10.11.20.2, 00:00:13, GigabitEthernet 3/21
10.200.10.0/24
        [120/1] via 10.11.20.2, 00:00:13, GigabitEthernet 3/21
10.300.10.0/24
        [120/1] via 10.11.20.2, 00:00:13, GigabitEthernet 3/21
10.11.20.0/24           directly connected,GigabitEthernet 3/21
10.11.30.0/24           directly connected,GigabitEthernet 3/11
10.0.0.0/8              auto-summary
192.168.1.0/24          directly connected,GigabitEthernet 3/43
192.168.1.0/24          auto-summary
192.168.2.0/24          directly connected,GigabitEthernet 3/44
192.168.2.0/24          auto-summary
Core3#
```

Figure 32-14. Using show ip routes for Core 3 RIP Setup



```
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, Gi 3/21            120/1    00:01:14
  C    10.11.20.0/24     Direct, Gi 3/21                      0/0    00:01:53
  C    10.11.30.0/24     Direct, Gi 3/11                      0/0    00:06:00
  R    10.200.10.0/24    via 10.11.20.2, Gi 3/21            120/1    00:01:14
  R    10.300.10.0/24    via 10.11.20.2, Gi 3/21            120/1    00:01:14
  C    192.168.1.0/24    Direct, Gi 3/43                      0/0    00:06:53
  C    192.168.2.0/24    Direct, Gi 3/44                      0/0    00:06:26
Core3#
```

Figure 32-15. Using show ip protocols Command to Show RIP Configuration Activity on Core 3



```
Core3#show ip protocols

Routing Protocol is "RIP"
 Sending updates every 30 seconds, next due in 6
 Invalid after 180 seconds, hold down 180, flushed after 240
 Output delay 8 milliseconds between packets
 Automatic network summarization is in effect
 Outgoing filter for all interfaces is
 Incoming filter for all interfaces is
 Default redistribution metric is  1
 Default version control: receive version 2, send version 2
        Interface       Recv  Send
        GigabitEthernet 3/21   2     2
        GigabitEthernet 3/11   2     2
        GigabitEthernet 3/44   2     2
        GigabitEthernet 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

Figure 32-16. Summary of Core 2 RIP Configuration Using Output of show run Command

```
!
interface GigabitEthernet 2/11
 ip address 10.11.10.1/24
 no shutdown
!
interface GigabitEthernet 2/31
 ip address 10.11.20.2/24
 no shutdown

!
interface GigabitEthernet 2/41
 ip address 10.200.10.1/24
 no shutdown

!
interface GigabitEthernet 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
```

Figure 32-17. Summary of Core 3 RIP Configuration Using Output of show run Command

```
!
interface GigabitEthernet 3/11
 ip address 10.11.30.1/24
 no shutdown

!
interface GigabitEthernet 3/21
 ip address 10.11.20.1/24
 no shutdown

!
interface GigabitEthernet 3/43
 ip address 192.168.1.1/24
 no shutdown

!
interface GigabitEthernet 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
```

33

Remote Monitoring

Remote Monitoring is supported on platform C E S

This chapter describes the Remote Monitoring (RMON):

- Implementation on page 601
- Fault Recovery on page 602

Remote Monitoring (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 Force10 Ethernet Interfaces.

RMON operates with SNMP and monitors all nodes on a 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 MIBs.

Implementation

You must configure SNMP prior to setting up RMON. For a complete SNMP implementation discussion, refer to Chapter 6, Simple Network Management Protocol (SNMP), on page 47.

Configuring RMON requires using the RMON CLI and includes the following tasks:

- Set rmon alarm
- Configure an RMON event
- Configure RMON collection statistics
- Configure RMON collection history
- Enable an RMON MIB collection history group

RMON implements the following standard RFCs (for details see Appendix 47, Standards Compliance):

- RFC-2819
- RFC-3273
- RFC-3434

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.

Line Card Down—The same as Interface Down (see above).

RPM Down, RPM Failover—Master and standby RPMs run the RMON sampling process in the background. Therefore, when an RPM goes down, the other RPM maintains the sampled data—the new master RPM provides the same sampled data as did the old master—as long as the master RPM had been running long enough to sample all the data.

NMS backs up all the long-term data collection, and displays the failover downtime from the performance graph.

Chassis Down—When a chassis goes down, all sampled data is lost. But the RMON configurations are saved in the configuration file, and the sampling process continues after the chassis returns to operation.

Platform Adaptation—RMON supports all Dell Force10 chassis and all Dell Force10 Ethernet Interfaces.

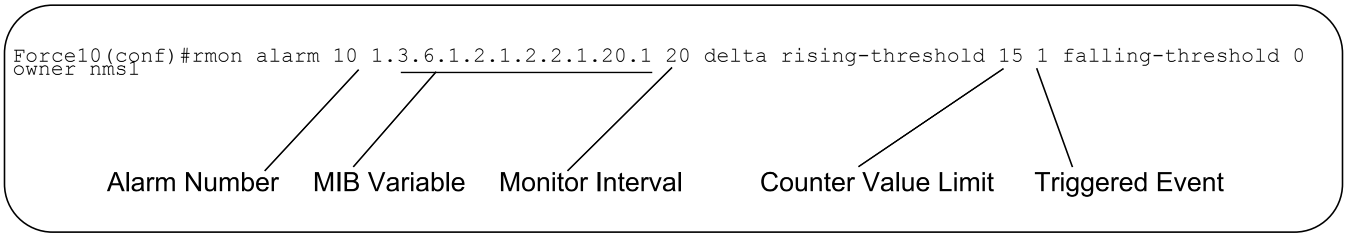
Set rmon alarm

To set an alarm on any MIB object, use the **rmon alarm** or **rmon hc-alarm** command in GLOBAL CONFIGURATION mode. To disable the alarm, use the **no** form of this command:

Command Syntax	Command Mode	Purpose
[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]**	CONFIGURATION	Set an alarm on any MIB object. Use the **no** form of this command to disable the alarm. Configure the alarm using the following optional parameters: • *number*: Alarm number, should be 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 the 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 is the *alarmSampleType* in the RMON Alarm table. • **absolute**: Tests each MIB variable directly, this 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 is a 32-bits value, for **rmon hc-alarm** command this 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 is a 32-bits value, for **rmon hc-alarm** command this is a 64bits 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 is the alarmOwner object in the alarmTable of the RMON MIB. Default is a null-terminated string.

The following example configures an RMON alarm using the **rmon alarm** command.

Figure 33-1. rmon alarm Command Example



```
Force10(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
```

The above 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 a 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.

Configure an RMON event

To add an event in the RMON event table, use the **rmon event** command in GLOBAL CONFIGURATION mode. To disable RMON on the interface, use the **no** form of this command:

Command Syntax	Command Mode	Purpose
[no] rmon event number [log] [trap community] [description string] [owner string]	CONFIGURATION	*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, the value must 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. 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.

The following example shows the **rmon event** command.

Figure 33-2. rmon event Command Example

```
Force10(conf)#rmon event 1 log trap eventtrap description "High ifOutErrors" owner nms1
```

The above configuration example creates RMON event number 1, with the description “High ifOutErrors”, and generates a log entry when the event is triggered by an alarm. 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”.

Configure RMON collection statistics

To enable RMON MIB statistics collection on an interface, use the RMON collection statistics command in interface configuration mode. To remove a specified RMON statistics collection, use the **no** form of this command.

Command Syntax	Command Mode	Purpose
[**no**] **rmon collection statistics** {**controlEntry** *integer*} [**owner** *ownername*]	CONFIGURATION INTERFACE (config-if)	**controlEntry**: Specifies the RMON group of statistics using a value. *integer*: A value from 1 to 65,535 that identifies the RMON Statistics Table. The value must be unique in the RMON Statistic Table. **owner**: (Optional) Specifies the name of the owner of the RMON group of statistics. *ownername*: (Optional) Records the name of the owner of the RMON group of statistics. Default is a null-terminated string

The following command enables the RMON statistics collection on the interface, with an ID value of 20 and an owner of “john.”

Figure 33-3. rmon collection statistics Command Example

```
Force10(conf-if-mgmt)#rmon collection statistics controlEntry 20 owner john
```

Configure 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. To remove a specified RMON history group of statistics collection, use the **no** form of this command.

Command Syntax	Command Mode	Purpose
[no] rmon collection history {controlEntry integer} [owner ownername] [buckets bucket-number] **[interval seconds]**	CONFIGURATION INTERFACE (config-if)	**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.Default is a null-terminated string. *ownername*: (Optional) Records the name of the owner of the RMON group of statistics. **buckets**: (Optional) Specifies the maximum number of buckets desired for the RMON collection history group of statistics. *bucket-number*: (Optional) A value associated with the number of buckets specified for the RMON collection history group of statistics. The value is limited to from 1 to 1000. 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). Default is 1,800 as defined in RFC-2819.

Enable an RMON MIB collection history group

The following 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.

Figure 33-4. rmon collection history Command Example

```
Force10(conf-if-mgmt)#rmon collection history controlEntry 20 owner john
```

34

Rapid Spanning Tree Protocol

Rapid Spanning Tree Protocol is supported on platforms: C E S

Protocol Overview

Rapid Spanning Tree Protocol (RSTP) is a Layer 2 protocol—specified by IEEE 802.1w—that is essentially the same as Spanning-Tree Protocol (STP) but provides faster convergence and interoperability with switches configured with STP and MSTP.

FTOS supports three other variations of Spanning Tree, as shown in Table 34-1.

Table 34-1. FTOS Supported Spanning Tree Protocols

Dell Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Configuring Rapid Spanning Tree

Configuring Rapid Spanning Tree is a two-step process:

1. Configure interfaces for Layer 2. See page 48.
2. Enable Rapid Spanning Tree Protocol. See page 49.

Related Configuration Tasks

- Add and Remove Interfaces on page 612
- Modify Global Parameters on page 612
- Modify Interface Parameters on page 613
- Configure an EdgePort on page 614
- Preventing Network Disruptions with BPDU Guard on page 711
- Influence RSTP Root Selection on page 615

- Configuring Spanning Trees as Hitless on page 713
- SNMP Traps for Root Elections and Topology Changes on page 616
- Fast Hellos for Link State Detection on page 616
- Flush MAC Addresses after a Topology Change on page 435

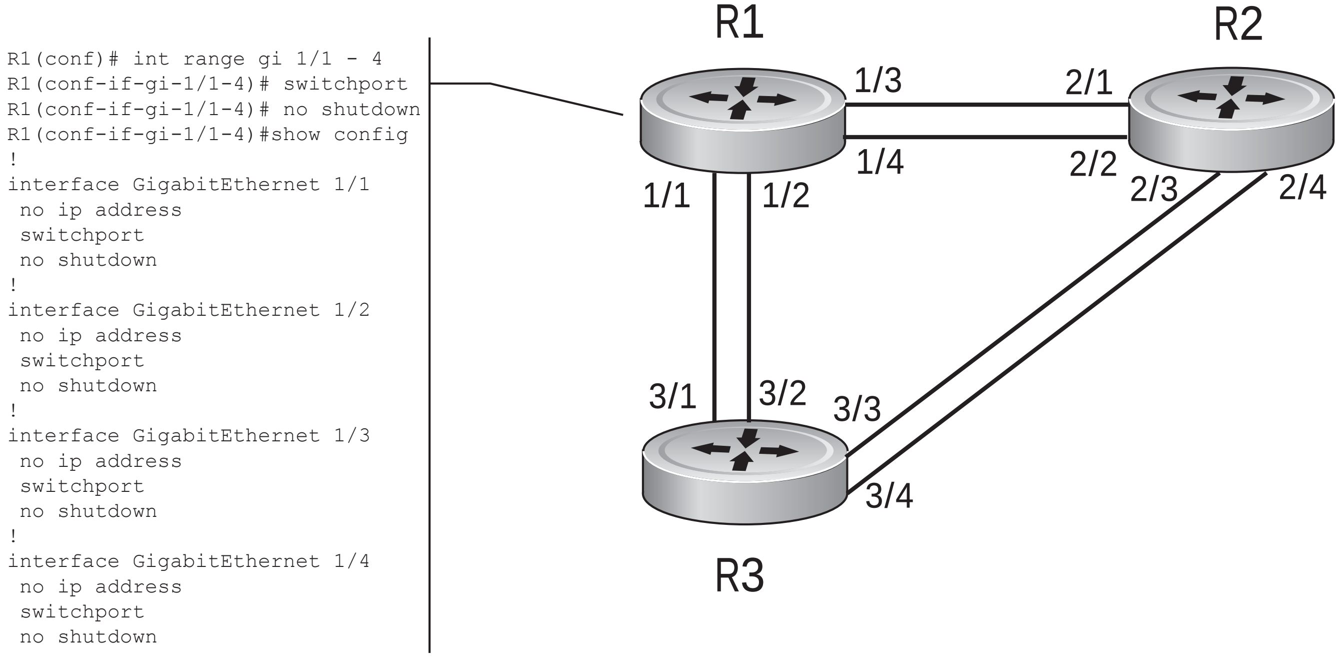
Important Points to Remember

- RSTP is disabled by default.
- FTOS supports only one Rapid Spanning Tree (RST) instance.
- All interfaces in VLANs and all enabled interfaces in Layer 2 mode are automatically added to the RST topology.
- Avoid using the **range** command to add a large group of ports to a large group of VLANs; adding a group of ports to a range of VLANs sends multiple messages to the RSTP task. When using the **range** command, Dell Force10 recommends limiting the range to 5 ports and 40 VLANs.

Configure Interfaces for Layer 2 Mode

All interfaces on all bridges that will participate in Rapid Spanning Tree must be in Layer 2 and enabled.

Figure 34-1. Configuring Interfaces for Layer 2 Mode



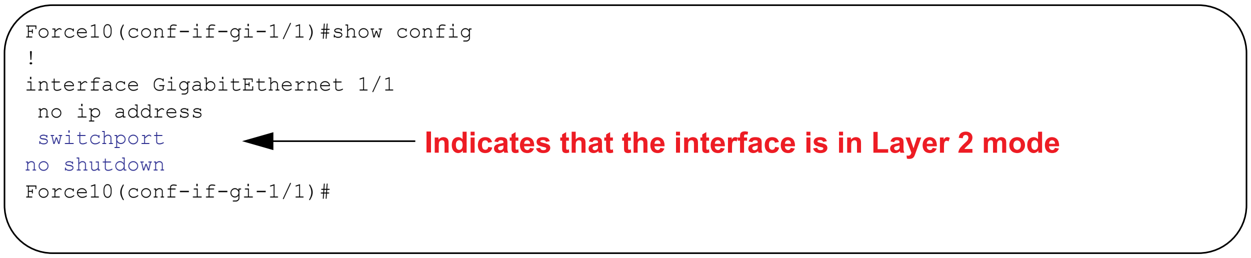
```
R1(conf)# int range gi 1/1 - 4
R1(conf-if-gi-1/1-4)# switchport
R1(conf-if-gi-1/1-4)# no shutdown
R1(conf-if-gi-1/1-4)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/2
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/3
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/4
 no ip address
 switchport
 no shutdown
```

To configure the interfaces for Layer 2 and then enable them:

Step	Task	Command Syntax	Command Mode
1	If the interface has been assigned an IP address, remove it.	**no ip address**	INTERFACE
2	Place the interface in Layer 2 mode.	**switchport**	INTERFACE
3	Enable the interface.	**no shutdown**	INTERFACE

Verify that an interface is in Layer 2 mode and enabled using the **show config** command from INTERFACE mode.

Figure 34-2. Verifying Layer 2 Configuration



```
Force10(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
no shutdown
Force10(conf-if-gi-1/1)#
```

Indicates that the interface is in Layer 2 mode

Enable Rapid Spanning Tree Protocol Globally

Rapid Spanning Tree Protocol must be enabled globally on all participating bridges; it is not enabled by default.

To enable Rapid Spanning Tree globally for all Layer 2 interfaces:

Step	Task	Command Syntax	Command Mode
1	Enter the PROTOCOL SPANNING TREE RSTP mode.	**protocol spanning-tree rstp**	CONFIGURATION
2	Enable Rapid Spanning Tree.	**no disable**	PROTOCOL SPANNING TREE RSTP

Note: To disable RSTP globally for all Layer 2 interfaces, enter the **disable** command from PROTOCOL SPANNING TREE RSTP mode.

Verify that Rapid Spanning Tree is enabled using the **show config** command from PROTOCOL SPANNING TREE RSTP mode.

Figure 34-3. Verifying RSTP is Enabled

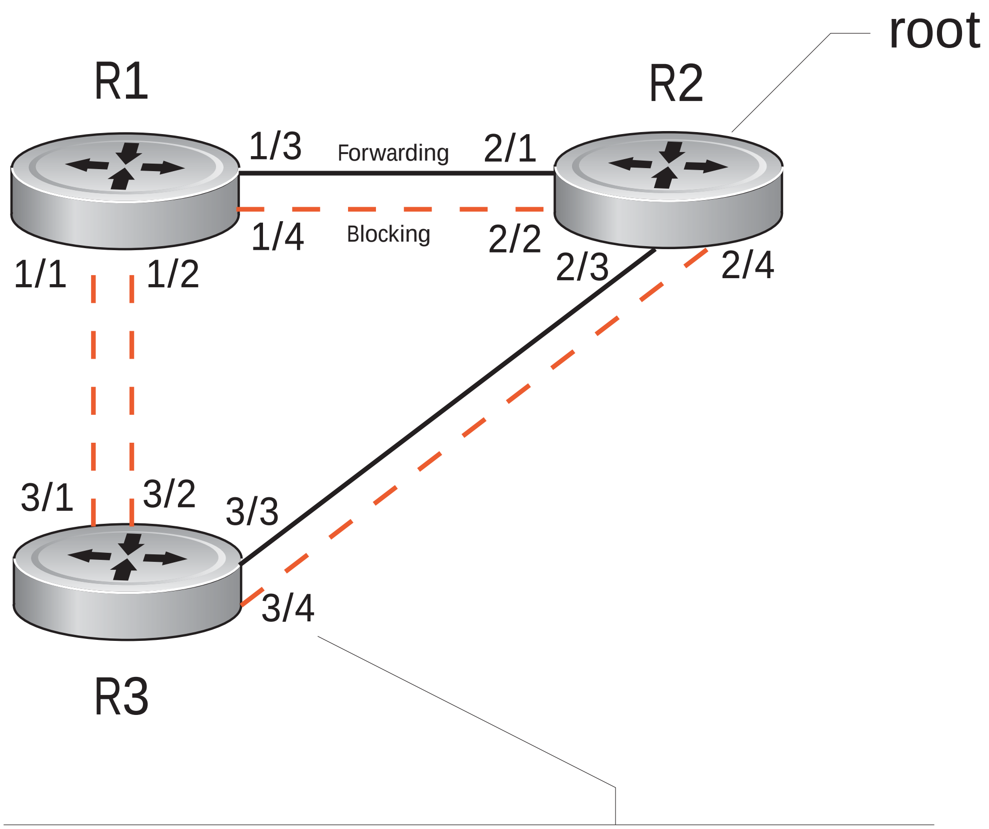
```
Force10(conf-rstp)#show config
!
protocol spanning-tree rstp
 no disable
Force10(conf-rstp)#
```

Indicates that Rapid Spanning Tree is enabled

When you enable Rapid Spanning Tree, 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.

Figure 34-4. Rapid Spanning Tree Enabled Globally



```
Port 684 (GigabitEthernet 4/43) is alternate Discarding
Port path cost 20000, Port priority 128, Port Identifier 128.684
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.684, designated path cost 20000
Number of transitions to forwarding state 0
BPDU : sent 3, received 219
The port is not in the Edge port mode
```

View the interfaces participating in Rapid Spanning Tree using 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.

Figure 34-5. show spanning-tree rstp Command Example

```
Force10#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 Gi 1/26

Port 377 (GigabitEthernet 2/1) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.377
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.377, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 9
The port is not in the Edge port mode

Port 378 (GigabitEthernet 2/2) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.378
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.378, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 2
The port is not in the Edge port mode

Port 379 (GigabitEthernet 2/3) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.379
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.379, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 5
The port is not in the Edge port mode

Port 380 (GigabitEthernet 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
```

Confirm that a port is participating in Rapid Spanning Tree using the **show spanning-tree rstp brief** command from EXEC privilege mode.

Figure 34-6. show spanning-tree rstp brief Command Example

```
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
---------- -------- ---- ------- --- ------- -------------------- --------
Gi 3/1     128.681  128  20000   BLK 20000   32768 0001.e80b.88bd 128.469
Gi 3/2     128.682  128  20000   BLK 20000   32768 0001.e80b.88bd 128.470
Gi 3/3     128.683  128  20000   FWD 20000   32768 0001.e801.cbb4 128.379
Gi 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
---------- ------ -------- ---- ------- --- ------- --------- ----
Gi 3/1     Altr   128.681  128  20000   BLK 20000   P2P       No
Gi 3/2     Altr   128.682  128  20000   BLK 20000   P2P       No
Gi 3/3     Root   128.683  128  20000   FWD 20000   P2P       No
Gi 3/4     Altr   128.684  128  20000   BLK 20000   P2P       No
R3#
```

Add and Remove Interfaces

- 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**, re-enable it using the command **spanning-tree 0**.
- Remove an interface from the Rapid Spanning Tree topology using the command **no spanning-tree 0**. See also Removing an Interface from the Spanning Tree Group on page 707 for BPDU Filtering behavior.

Modify Global Parameters

You can modify Rapid 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 the Rapid Spanning Tree group.

- **Forward-delay** is the amount of time an interface waits in the Listening State and the Learning State before it transitions to the Forwarding State.
- **Hello-time** is the time interval in which the bridge sends RSTP Bridge Protocol Data Units (BPDUs).
- **Max-age** is the length of time the bridge maintains configuration information before it refreshes that information by recomputing the RST topology.

Note: Dell Force10 recommends that only experienced network administrators change the Rapid Spanning Tree group parameters. Poorly planned modification of the RSTG parameters can negatively impact network performance.

Table 34-2 displays the default values for RSTP.

Table 34-2. RSTP Default Values

RSTP Parameter		Default Value
Forward Delay		15 seconds
Hello Time		2 seconds
Max Age		20 seconds
Port Cost	100-Mb/s Ethernet interfaces	200000
	1-Gigabit Ethernet interfaces	20000
	10-Gigabit Ethernet interfaces	2000
	Port Channel with 100 Mb/s Ethernet interfaces	180000
	Port Channel with 1-Gigabit Ethernet interfaces	18000
	Port Channel with 10-Gigabit Ethernet interfaces	1800
Port Priority		128

To change these parameters, use the following commands, on the root bridge:

Task	Command Syntax	Command Mode
Change the forward-delay parameter. • Range: 4 to 30 • Default: 15 seconds	**forward-delay** *seconds*	PROTOCOL SPANNING TREE RSTP
Change the hello-time parameter. **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**hello-time** *seconds*	PROTOCOL SPANNING TREE RSTP
Change the max-age parameter. Range: 6 to 40 Default: 20 seconds	**max-age** *seconds*	PROTOCOL SPANNING TREE RSTP

View the current values for global parameters using the **show spanning-tree rstp** command from EXEC privilege mode. See Figure 34-5.

Modify Interface Parameters

On interfaces in Layer 2 mode, you can set the port cost and port priority values.

- **Port cost** is a value that is based on the interface type. The default values are listed in Table 34-2. The greater the port cost, the less likely the port will be selected to be a forwarding port.

- **Port priority** influences the likelihood that a port will be 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:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 65535 Default: see Table 34-2.	**spanning-tree rstp cost** *cost*	INTERFACE
Change the port priority of an interface. Range: 0 to 15 Default: 128	**spanning-tree rstp priority** *priority-value*	INTERFACE

View the current values for interface parameters using the **show spanning-tree rstp** command from EXEC privilege mode. See Figure 34-5.

Configure 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 shutdown 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 will drop 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 it is enabled on an interface connected to a network.

To enable EdgePort on an interface, use the following command:

Task	Command Syntax	Command Mode
Enable EdgePort on an interface.	**spanning-tree rstp edge-port** [**bpduguard |** **shutdown-on-violation**]	INTERFACE

Verify that EdgePort is enabled on a port using the **show spanning-tree rstp** command from the EXEC privilege mode or the **show config** command from INTERFACE mode; Dell Force10 recommends using the **show config** command, as shown in Figure 34-7.

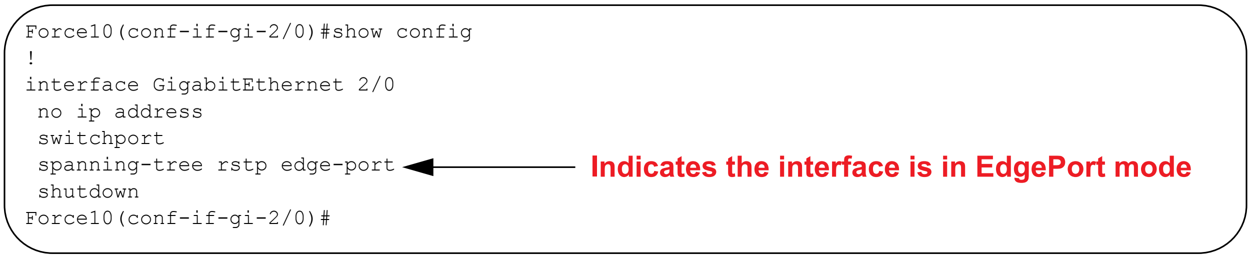
FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

The error disabled state can be cleared with any of the following methods:

- Perform an **shutdown** command on the interface.
- Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).
- Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
- Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Figure 34-7. EdgePort Enabled on Interface



```
Force10(conf-if-gi-2/0)#show config
!
interface GigabitEthernet 2/0
 no ip address
 switchport
 spanning-tree rstp edge-port
 shutdown
Force10(conf-if-gi-2/0)#
```

Influence RSTP Root Selection

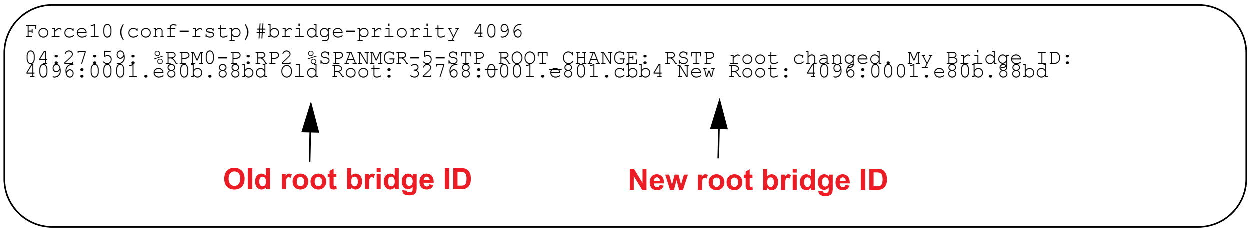
The Rapid Spanning Tree Protocol determines the root bridge, but you can assign one bridge a lower priority to increase the likelihood that it will be selected as the root bridge.

To change the bridge priority, use the following command:

Task	Command Syntax	Command Mode
Assign a number as the bridge priority or designate it as the primary or secondary root. *priority-value* range: 0 to 65535. The lower the number assigned, the more likely this bridge will become the root bridge. The default is 32768. Entries must be multiples of 4096.	**bridge-priority** *priority-value*	PROTOCOL SPANNING TREE RSTP

A console message appears when a new root bridge has been assigned. Figure 34-8 shows the console message after the **bridge-priority** command is used to make R2 the root bridge.

Figure 34-8. bridge-priority Command Example



```
Force10(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

Enable SNMP traps for RSTP, MSTP, and PVST+ collectively using the command **snmp-server enable traps xstp**.

Fast Hellos for Link State Detection

Fast Hellos for Link State Detection is available only on platform: S

Use RSTP Fast Hellos to achieve sub-second link-down detection so that convergence is triggered faster. The standard RSTP link-state detection mechanism does not offer the same low link-state detection speed.

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.

Task	Command Syntax	Command Mode
Configure a hello time on the order of milliseconds.	**hello-time milli-second** *interval* Range: 50 - 950 milliseconds	PROTOCOL RSTP

```
Force10(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.
Note: When millisecond hellos are configured, the default hello interval of 2 seconds is still used for edge ports; the millisecond hello interval is not used.

35

Security

Security features are supported on platforms C E S

This chapter discusses several ways to provide access security to the Dell Force10 system. Platform-specific features are identified by the C, E or S icons (as shown below).

For details on all commands discussed in this chapter, see the Security Commands chapter in the *FTOS Command Reference*.

AAA Accounting

AAA Accounting is part of the AAA security model (Accounting, Authentication, and Authorization), which includes services for authentication, authorization, and accounting. For details on commands related to AAA security, refer to the Security chapter in the *FTOS Command Reference*.

AAA Accounting enables tracking of services that users are accessing and the amount of network resources being consumed by those services. When AAA Accounting is enabled, the network server reports user activity to the security server in the form of accounting records. Each accounting record is comprised of accounting 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 apply that list to various interfaces.

Configuration Task List for AAA Accounting

The following sections present the AAA Accounting configuration tasks:

- Enable AAA Accounting on page 618 (mandatory)
- Suppress AAA Accounting for null username sessions on page 619 (optional)
- Configure Accounting of EXEC and privilege-level command usage on page 619 (optional)
- Configure AAA Accounting for terminal lines on page 619 (optional)
- Monitor AAA Accounting on page 619 (optional)

Enable AAA Accounting

The **aaa accounting** command enables you to create a record for any or all of the accounting functions monitored. To enable AAA accounting, perform the following task in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa accounting {**system** | **exec** | **command** *level*} {*default* | *name*} {**start-stop** | **wait-start** | **stop-only**} {**tacacs+**}	CONFIGURATION	Enable AAA Accounting and create a record for monitoring the accounting function. The variables are: • **system**—sends accounting information of any other AAA configuration • **exec**—sends accounting information when a user has logged in to the EXEC mode • **command** *level*—sends accounting of commands executed at the specified privilege level • *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, FTOS supports only TACACS+

Suppress AAA Accounting for null username sessions

When AAA Accounting is activated, the FTOS software issues accounting records for all users on the system, including users whose username string, because of protocol translation, is NULL. 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, perform the following task in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa accounting suppress null-username	CONFIGURATION	Prevent accounting records from being generated for users whose username string is NULL

Configure Accounting of EXEC and privilege-level command usage

The network access server monitors the accounting functions defined in the TACACS+ attribute/value (AV) pairs.

In the following sample configuration, AAA accounting is set to track all usage of EXEC commands and commands on privilege level 15.

```
Force10(conf)#aaa accounting exec default start-stop tacacs+
Force10(conf)#aaa accounting command 15 default start-stop tacacs+
```

System accounting can use only the default method list:

aaa accounting system default start-stop tacacs+

Configure AAA Accounting for terminal lines

Use the following commands to enable accounting with a named method list for a specific terminal line (where com15 and execAcct are the method list names):

```
Force10(config-line-vty)# accounting commands 15 com15
Force10(config-line-vty)# accounting exec execAcct
```

Monitor AAA Accounting

FTOS 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, perform the following task in Privileged EXEC mode:

Command Syntax	Command Mode	Purpose
show accounting	CONFIGURATION	Step through all active sessions and print all the accounting records for the actively accounted functions.

Figure 35-1. show accounting Command Example for AAA Accounting

```
Force10#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
```

AAA Authentication

FTOS supports a distributed client/server system implemented through Authentication, Authorization, and Accounting (AAA) to help secure networks against unauthorized access. In the Dell Force10 implementation, the Dell Force10 system acts as a RADIUS or TACACS+ client and sends authentication requests to a central RADIUS or TACACS+ server that contains all user authentication and network service access information.

Dell Force10 uses local usernames/passwords (stored on the Dell Force10 system) or AAA for login authentication. With AAA, you can specify the security protocol or mechanism for different login methods and different users. In FTOS, 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.

Configuration Task List for AAA Authentication

The following sections provide the configuration tasks:

- Configure login authentication for terminal lines
- Configure AAA Authentication login methods on page 621
- Enable AAA Authentication on page 622
- AAA Authentication—RADIUS on page 622

For a complete listing of all commands related to login authentication, refer to the Security chapter in the *FTOS Command Reference*.

Configure login authentication for terminal lines

You can assign up to five authentication methods to a method list. FTOS 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, FTOS applies the next method list until the user either passes or fails the authentication. If the user fails a method list, FTOS does not apply the next method list.

Configure AAA Authentication login methods

To configure an authentication method and method list, use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**aaa authentication login** {*method-list-name* | **default**} *method1* [... *method4*]	CONFIGURATION	Define an authentication method-list (*method-list-name*) or specify the **default**. The **default** method-list is applied to all terminal lines. Possible methods are: • **enable**—use the password defined by the **enable secret** or **enable password** command in the CONFIGURATION mode. • **line**—use the password defined by the password command in the LINE mode. • **local**—use the username/password database defined in the local configuration. • **none**—no authentication. • **radius**—use the RADIUS server(s) configured with the radius-server host command. • **tacacs+**—use the TACACS+ server(s) configured with the tacacs-server host command
2	**line** {**aux 0** | **console 0** | **vty** *number* [... *end-number*]}	CONFIGURATION	Enter the LINE mode.
3	**login authentication** {*method-list-name* | **default**}	LINE	Assign a *method-list-name* or the **default** list to the terminal line.

FTOS 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, FTOS 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 in the event that network-wide issue prevents access to these servers.

To view the configuration, use the **show config** command in the LINE mode or the **show running-config** in the EXEC Privilege mode.

Note: Dell Force10 recommends that you use the **none** method only as a backup. This method does not authenticate users. The **none** and **enable** methods do not work with SSH.

You can create multiple method lists and assign them to different terminal lines.

Enable AAA Authentication

To enable AAA authentication, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa authentication enable {*method-list-name* | **default**} *method1* [... *method4*]	CONFIGURATION	• **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 the default list is not set, only the local enable is checked. This has the same effect as issuing: **aaa authentication enable default enable**

AAA Authentication—RADIUS

To enable authentication from the RADIUS server, and use TACACS as a backup, use the following commands:

Step	Command Syntax	Command Mode	Purpose
1	**aaa authentication enable default radius tacacs**	CONFIGURATION	To enable RADIUS and to set up TACACS as backup.
2	**radius-server host x.x.x.x key some-password**	CONFIGURATION	To establish host address and password.
3	**tacacs-server host x.x.x.x key some-password**	CONFIGURATION	To establish host address and password.

To get enable authentication from the RADIUS server, and use TACACS as a backup, issue the following commands:

```
Force10(config)# aaa authentication enable default radius tacacs
Radius and TACACS server has to be properly setup for this.
Force10(config)# radius-server host x.x.x.x key <some-password>
```

To use local authentication for enable secret on console, while using remote authentication on VTY lines, perform the following steps:

```
Force10(config)# aaa authentication enable mymethodlist radius
tacacs
Force10(config)# line vty 0 9
```

Server-side configuration

TACACS+: When using TACACS+, Dell Force10 sends an initial packet with service type SVC_ENABLE, and then, a second packet with just the password. The TACACS server must have an entry for username $enable$.

RADIUS: When using RADIUS authentication, FTOS 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

FTOS 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, FTOS 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 FTOS, you can configure a privilege level for users who need limited access to the system.

Every command in FTOS is assigned a privilege level of 0, 1 or 15. You can configure up to 16 privilege levels in FTOS. FTOS is pre-configured with 3 privilege levels and you can configure 13 more. The three pre-configured levels are:

- **Privilege level 1**—is the default level for the 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 FTOS.

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. Refer to Configure a username and password on page 624 for more information on configuring user names.

By default, commands in FTOS 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, you must log in to the router, enter the **enable** command for privilege level 15 (this is the default level for the command) and then enter the CONFIGURATION mode.

You can configure passwords to control access to the box and assign different privilege levels to users. FTOS 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.

- Configure a username and password on page 624 (mandatory)
- Configure the enable password command on page 625 (mandatory)
- Configure custom privilege levels on page 626 (mandatory)
- Specify LINE mode password and privilege on page 628 (optional)
- Enable and disabling privilege levels on page 628 (optional)

For a complete listing of all commands related to FTOS privilege levels and passwords, refer to the Security chapter in the *FTOS Command Reference*.

Configure a username and password

In FTOS, you can assign a specific username to limit user access to the system.

To configure a username and password, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
username *name* [**access-class** *access-list-name*] [**nopassword | password** [*encryption-type*] *password*] [**privilege** *level*]	CONFIGURATION	Assign a user name and password. Configure the optional and required parameters: • *name:* Enter a text string up to 63 characters long. • **access-class** *access-list-name:* Enter the name of a configured IP ACL. • **nopassword:** Do not require the user to enter a password. • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a string. • **privilege** *level* range: 0 to 15.

To view usernames, use the **show users** command in the EXEC Privilege mode.

Configure the enable password command

To configure FTOS, you must use the **enable** command to enter the EXEC Privilege level 15. After entering the command, FTOS requests that you enter a password. Privilege levels are not assigned to passwords, rather passwords are assigned to a privilege level. A password for any privilege level can always be changed. To change to a different privilege level, enter the **enable** command, followed by 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 in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
enable password [**level** *level*] [*encryption-mode*] *password*	CONFIGURATION	Configure a password for a privilege level. Configure the optional and required parameters: • **level** *level:* Specify a level 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 the EXEC Privilege mode.

In custom-configured privilege levels, the **enable** command is always available. No matter what privilege level you entered FTOS, you can enter the **enable 15** command to access and configure all CLI.

Configure 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 FTOS, commands have certain privilege levels. With the privilege command, the default level can be changed or you can reset their privilege level back to the default.

- Assign the launch keyword (for example, **configure**) for the keyword's command mode.
- If you assign only the first keyword to the privilege level, all commands beginning with that keyword are also assigned to the privilege level. If you enter the entire command, the software assigns the privilege level to that command only.

To assign commands and passwords to a custom privilege level, you must be in privilege level 15 and use these commands in the following sequence in the CONFIGURATION mode:

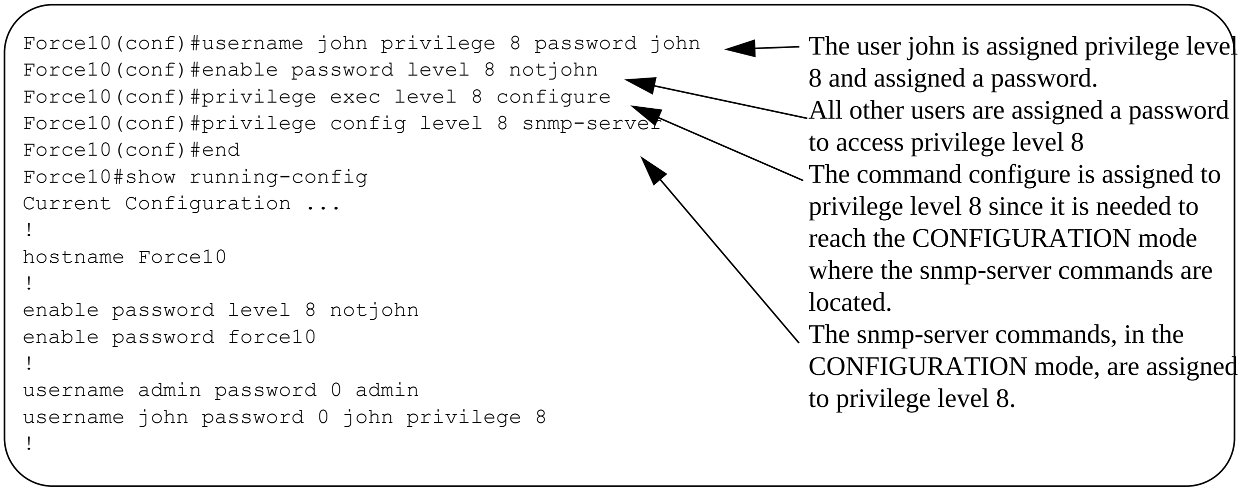
Step	Command Syntax	Command Mode	Purpose
1	**username** *name* [**access-class** *access-list-name*] [**privilege** *level*] [**nopassword | password** [*encryption-type*] *password*]	CONFIGURATION	Assign a user name and password. Configure the optional and required parameters: • *name:* Enter a text string (up to 63 characters). • **access-class** *access-list-name:* Enter the name of a configured IP ACL. • **privilege** *level* range: 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 string.
2	**enable password** [**level** *level*] [*encryption-mode*] *password*	CONFIGURATION	Configure a password for privilege level. Configure the optional and required parameters: • **level** *level:* Specify a level 0 to 15. Level 15 includes all levels. • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a string up to 25 characters long. To change only the password for the enable command, configure only the *password* parameter.

Step	Command Syntax	Command Mode	Purpose
3	**privilege** *mode* {**level** *level command* | **reset** *command*}	CONFIGURATION	Configure level and commands for a mode or reset a command’s level. Configure the following required and optional parameters: • *mode:* Enter a keyword for the modes (exec, configure, interface, line, route-map, router) • **level** *level* range: 0 to 15. Levels 0, 1 and 15 are pre-configured. Levels 2 to 14 are available for custom configuration. • *command:* A FTOS CLI keyword (up to 5 keywords allowed). • **reset:** Return the command to its default privilege mode.

To view the configuration, use the **show running-config** command in the EXEC Privilege mode.

Figure 35-2 is an example of a configuration to allow a user “john” to view only the EXEC mode commands and all **snmp-serve**r commands. Since the **snmp-server** commands are “enable” level commands and, by default, found in the CONFIGURATION mode, you must also assign the launch command for the CONFIGURATION mode, **configure**, to the same privilege level as the **snmp-server** commands.

Figure 35-2. Configuring a Custom Privilege Level



```
Force10(conf)#username john privilege 8 password john
Force10(conf)#enable password level 8 notjohn
Force10(conf)#privilege exec level 8 configure
Force10(conf)#privilege config level 8 snmp-server
Force10(conf)#end
Force10#show running-config
Current Configuration ...
!
hostname Force10
!
enable password level 8 notjohn
enable password force10
!
username admin password 0 admin
username john password 0 john privilege 8
!
```

The user john is assigned privilege level 8 and assigned a password.

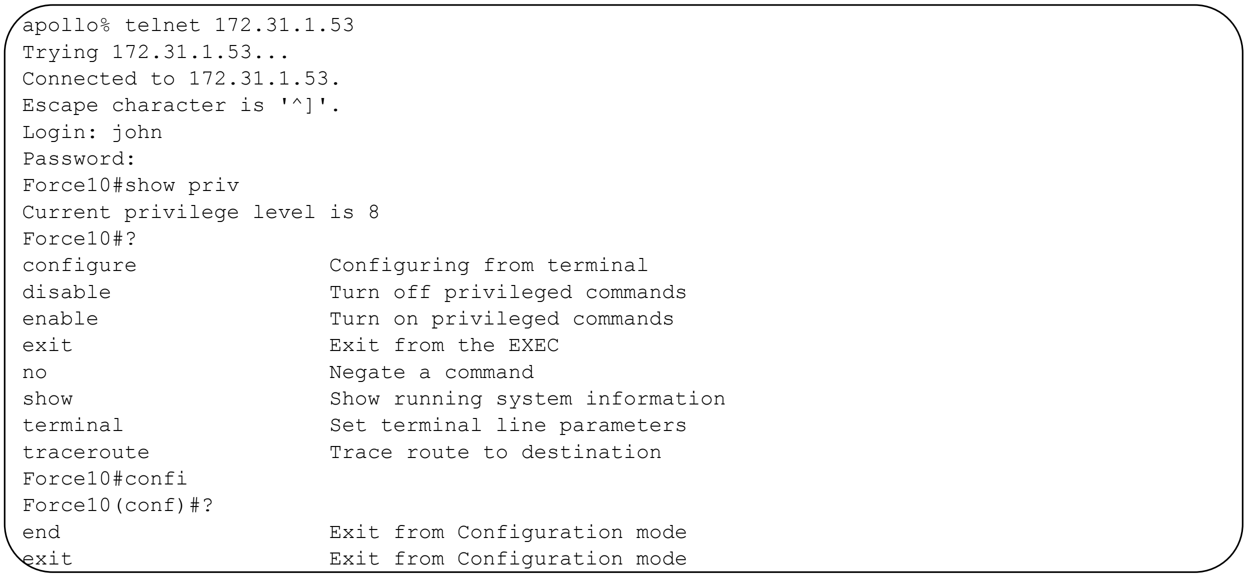
All other users are assigned a password to access privilege level 8

The command configure is assigned to privilege level 8 since it is needed to reach the CONFIGURATION mode where the snmp-server commands are located.

The snmp-server commands, in the CONFIGURATION mode, are assigned to privilege level 8.

Figure 35-3 is a screen shot of the Telnet session for user “john”. The **show privilege** command output confirms that “john” is in privilege level 8. In the EXEC Privilege mode, “john” can access only the commands listed. In CONFIGURATION mode, “john” can access only the **snmp-server** commands.

Figure 35-3. User john's Login and the List of 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:
Force10#show priv
Current privilege level is 8
Force10#?
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
Force10#confi
Force10(conf)#?
end                     Exit from Configuration mode
exit                    Exit from Configuration mode
```

Specify LINE mode password and privilege

You can specify a password authentication of all users on different *terminal* lines. The user's privilege level will be the same as the privilege level assigned to the terminal line, unless a more specific privilege level is is assigned to the user.

To specify a password for the terminal line, use the following commands, in any order, in the LINE mode:

Command Syntax	Command Mode	Purpose
privilege level *level*	LINE	Configure a custom privilege level for the terminal lines. • **level** *level* range: 0 to 15. Levels 0, 1 and 15 are pre-configured. Levels 2 to 14 are available for custom configuration.
password [*encryption-type*] *password*	LINE	Specify either a plain text or encrypted 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 the LINE mode.

Enable and disabling privilege levels

Enter the **enable** or **enable privilege-level** command in the EXEC Privilege mode to set a user's security level. If you do not enter a privilege level, FTOS sets it to 15 by default.

To move to a lower privilege level, enter the command **disable** followed by the **level-number** you wish to set for the user in the EXEC Privilege mode. 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 Force10 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, the error can be viewed 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 on RADIUS, refer to RFC 2865, *Remote Authentication Dial-in User Service*.

RADIUS Authentication and Authorization

FTOS 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 authorization is enabled, 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 authorization is enabled by the RADIUS server, the server returns the following information to the client:

- Idle time
- ACL configuration information
- Auto-command
- Privilege level

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

The RADIUS server can specify an ACL. If an ACL is configured on the RADIUS server, and if that ACL is present, user may be allowed access based on that ACL. If the ACL is absent, authorization fails, and a message is logged indicating the this.

RADIUS can specify an ACL for the user if both of the following are true:

- If an ACL is absent
- 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. To do this, use the command **auto-command**. The auto-command is executed when the user is authenticated and before the prompt appears to the user.

Set access to privilege levels through RADIUS

Through the RADIUS server, you can use the command **privilege level** to configure a privilege level for the user to enter into when they connect to a session.This value is configured on the client system.

Configuration Task List for RADIUS

To authenticate users using RADIUS, at least one RADIUS server must be specified 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.

- Define a aaa method list to be used for RADIUS on page 631 (mandatory)
- Apply the method list to terminal lines on page 631 (mandatory except when using default lists)
- Specify a RADIUS server host on page 632 (mandatory)
- Set global communication parameters for all RADIUS server hosts on page 632 (optional)

- Monitor RADIUS on page 633 (optional)

For a complete listing of all FTOS commands related to RADIUS, refer to the Security chapter in the *FTOS Command Reference*.

Note: RADIUS authentication and authorization are done in a single step. Hence, authorization cannot be used independent of authentication. However, if RADIUS authorization is configured and authentication is not, then 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 the LINE mode or the **show running-config** command in the EXEC Privilege mode.

Define a AAA method list to be used for RADIUS

To configure RADIUS to authenticate or authorize users on the system, you must 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, enter one of the following commands in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa authentication login *method-list-name* **radius**	CONFIGURATION	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.
aaa authorization exec {*method-list-name* | **default**} **radius tacacs+**	CONFIGURATION	Create methodlist with RADIUS and TACACS+ as authorization methods. Typical order of methods: RADIUS, TACACS+, Local, None. If authorization is denied by RADIUS, the session ends (**radius** should not be the last method specified).

Apply the method list to terminal lines

To enable RADIUS AAA login authentication for a method list, you must apply it to a terminal line. To configure a terminal line for RADIUS authentication and authorization, enter the following commands:

Command Syntax	Command Mode	Purpose
line {**aux 0** | **console 0** | **vty** *number* [*end-number*]}	CONFIGURATION	Enter the LINE mode.
login authentication {*method-list-name* | **default**}	LINE	Enable AAA login authentication for the specified RADIUS method list. This procedure is mandatory if you are not using default lists.
authorization exec *methodlist*	CONFIGURATION	To use the methodlist.

Specify 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 in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
radius-server host {*hostname* | *ip-address*} [**auth-port** *port-number*] [**retransmit** *retries*] [**timeout** *seconds*] [**key** [*encryption-type*] *key*]	CONFIGURATION	Enter the host name or IP address of the RADIUS server host. Configure the optional communication parameters for the specific host: • **auth-port** *port-number* range: 0 to 65335. Enter a UDP port number. The default is 1812. • **retransmit** *retries* range: 0 to 100. Default is 3. • **timeout** *seconds* range: 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 these optional parameters are not configured, 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 multiple RADIUS server hosts are configured, FTOS attempts to connect with them in the order in which they were configured. When FTOS 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 Set global communication parameters for all RADIUS server hosts on page 632.

To view the RADIUS configuration, use the **show running-config radius** command in the EXEC Privilege mode.

To delete a RADIUS server host, use the **no radius-server host** {*hostname* | *ip-address*} command.

Set 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 both global and specific host parameters are configured, the specific host parameters override the global parameters for that RADIUS server host.

To set global communication parameters for all RADIUS server hosts, use any or all of the following commands in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
radius-server deadtime *seconds*	CONFIGURATION	Set a time interval after which a RADIUS host server is declared dead. • *seconds* range: 0 to 2147483647. Default: 0 seconds
radius-server key [*encryption-type*] *key*	CONFIGURATION	Configure a key for all RADIUS communications between the system and RADIUS server hosts. • *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.
radius-server retransmit *retries*	CONFIGURATION	Configure the number of times FTOS retransmits RADIUS requests. • *retries* range: 0 to 100. Default is 3 retries.
radius-server timeout *seconds*	CONFIGURATION	Configure the time interval the system waits for a RADIUS server host response. • *seconds* range: 0 to 1000. Default is 5 seconds.

To view the configuration of RADIUS communication parameters, use the **show running-config** command in the EXEC Privilege mode.

Monitor RADIUS

To view information on RADIUS transactions, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
debug radius	EXEC Privilege	View RADIUS transactions to troubleshoot problems.

TACACS+

FTOS 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:

- Choose TACACS+ as the Authentication Method
- Monitor TACACS+
- TACACS+ Remote Authentication and Authorization on page 635
- TACACS+ Remote Authentication and Authorization on page 635
- Specify a TACACS+ server host on page 636
- Choose TACACS+ as the Authentication Method on page 634

For a complete listing of all commands related to TACACS+, refer to the Security chapter in the *FTOS Command Reference*.

Choose 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, you must specify at least one TACACS+ server for the system to communicate with and configure TACACS+ as one of your authentication methods.

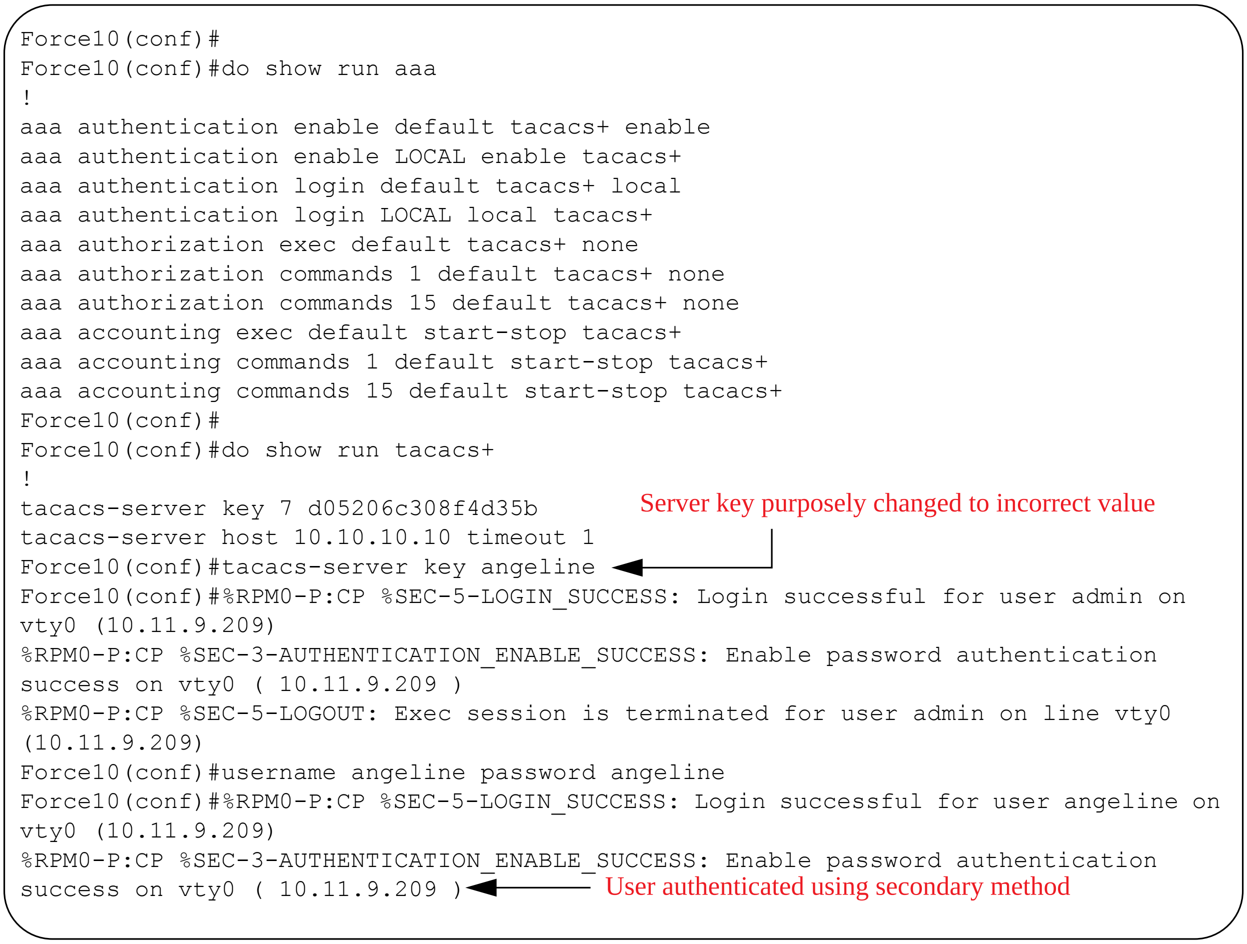
To select TACACS as the login authentication method, use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**tacacs-server host** {*ip-address* | *host*}	CONFIGURATION	Configure a TACACS+ server host. Enter the IP address or host name of the TACACS+ server. Use this command multiple times to configure multiple TACACS+ server hosts.
2	**aaa authentication login** {*method-list-name* | **default**} **tacacs+** [...*method3*]	CONFIGURATION	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 The **tacacs+** method should not be the last method specified.
3	**line** {**aux 0** | **console 0** | **vty** *number* [*end-number*]}	CONFIGURATION	Enter the LINE mode.
4	**login authentication** {*method-list-name* | **default**}	LINE	Assign the *method-list* to the terminal line.

To view the configuration, use the **show config** in the LINE mode or the **show running-config tacacs+** command in the EXEC Privilege mode.

If authentication fails using the primary method, FTOS employs the second method (or third method, if necessary) automatically. For example, if the TACACS+ server is reachable, but the server key is invalid, FTOS proceeds to the next authentication method. In Figure 35-4, the TACACS+ is incorrect, but the user is still authenticated by the secondary method.

Figure 35-4. Failed Authentication



```
Force10(conf)#
Force10(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+
Force10(conf)#
Force10(conf)#do show run tacacs+
!
tacacs-server key 7 d05206c308f4d35b
tacacs-server host 10.10.10.10 timeout 1
Force10(conf)#tacacs-server key angeline
Force10(conf)#%RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user admin on
vty0 (10.11.9.209)
%RPM0-P:CP %SEC-3-AUTHENTICATION_ENABLE_SUCCESS: Enable password authentication
success on vty0 ( 10.11.9.209 )
%RPM0-P:CP %SEC-5-LOGOUT: Exec session is terminated for user admin on line vty0
(10.11.9.209)
Force10(conf)#username angeline password angeline
Force10(conf)#%RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user angeline on
vty0 (10.11.9.209)
%RPM0-P:CP %SEC-3-AUTHENTICATION_ENABLE_SUCCESS: Enable password authentication
success on vty0 ( 10.11.9.209 )
```

Server key purposely changed to incorrect value

User authenticated using secondary method

Monitor TACACS+

To view information on TACACS+ transactions, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
debug tacacs+	EXEC Privilege	View TACACS+ transactions to troubleshoot problems.

TACACS+ Remote Authentication and Authorization

FTOS 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, then FTOS ignores the access class you have configured for the VTY line. FTOS instead gets this access class information from the TACACS+ server. FTOS needs to know the username and password of the incoming user before it can fetch the access class from the server. A user, therefore, will at least see the login prompt. If the access class denies the connection, FTOS closes the Telnet session immediately.

Figure 35-5 demonstrates how to configure the **access-class** from a TACACS+ server. This causes the configured access-class on the VTY line to be ignored. If you have configured a **deny10** ACL on the TACACS+ server, FTOS downloads it and applies it. If the user is found to be coming from the 10.0.0.0 subnet, FTOS also immediately closes the Telnet connection. Note, that no matter where the user is coming from, they see the login prompt.

Figure 35-5. Specify a TACACS+ server host

```
Force10#
Force10(conf)#
Force10(conf)#ip access-list standard deny10
Force10(conf-ext-nacl)#permit 10.0.0.0/8
Force10(conf-ext-nacl)#deny any
Force10(conf)#
Force10(conf)#aaa authentication login tacacsmethod tacacs+
Force10(conf)#aaa authentication exec tacacsauthorization tacacs+
Force10(conf)#tacacs-server host 25.1.1.2 key force10
Force10(conf)#
Force10(conf)#line vty 0 9
Force10(config-line-vty)#login authentication tacacsmethod
Force10(config-line-vty)#authorization exec tacauthor
Force10(config-line-vty)#
Force10(config-line-vty)#access-class deny10
Force10(config-line-vty)#end
```

When configuring a TACACS+ server host, you can set different communication parameters, such as the the key password.

To specify a TACACS+ server host and configure its communication parameters, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
tacacs-server host {*hostname* | *ip-address*} [**port** *port-number*] [**timeout** *seconds*] [**key** *key*]	CONFIGURATION	Enter the host name or IP address of the TACACS+ server host. Configure the optional communication parameters for the specific host: • **port** *port-number* range: 0 to 65335. Enter a TCP port number. The default is 49. • **timeout** *seconds* range: 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 should be the last parameter configured. If these optional parameters are not configured, the default global values are applied.

To specify multiple TACACS+ server hosts, configure the **tacacs-server host** command multiple times. If multiple TACACS+ server hosts are configured, FTOS 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 the 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:
Force10#
Force10#
```

Command Authorization

The AAA command authorization feature configures FTOS to send each configuration command to a TACACS server for authorization before it is added to the running configuration.

By default, the AAA authorization commands configure the system to check both EXEC mode and CONFIGURATION mode commands. Use the command **no aaa authorization config-commands** to enable only EXEC mode command checking.

If rejected by the AAA server, the command is not added to the running config, and messages similar to Message 1 are displayed.

Message 1 Configuration Command Rejection

```
04:07:48: %RPM0-P:CP %SEC-3-SEC_AUTHORIZATION_FAIL: Authorization failure  Command 
authorization failed for user (denyall) on vty0 ( 10.11.9.209 )
```

Protection from TCP Tiny and Overlapping Fragment Attacks

Tiny and overlapping fragment attack is a class of attack where configured ACL entries—denying TCP port-specific traffic—can be bypassed, and traffic can be sent to its destination although denied by the ACL. RFC 1858 and 3128 proposes a countermeasure to the problem. This countermeasure is configured into the line cards and enabled by default.

SCP and SSH

Secure Shell (SSH) is a protocol for secure remote login and other secure network services over an insecure network. FTOS is compatible with SSH versions 1.5 and 2, both the client and server modes. SSH sessions are encrypted and use authentication. For details on command syntax, see the Security chapter in the *FTOS Command Line Interface Reference*.

SCP is a remote file copy program that works with SSH and is supported by FTOS.

Note: The Windows-based WinSCP client software is not supported for secure copying between a PC and an FTOS-based system. Unix-based SCP client software is supported.

To use the SSH client, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
ssh {*hostname*} [**-l** *username* | **-p** *port-number* | **-v** {**1** | **2**}	EXEC Privilege	Open an SSH connection specifying the hostname, username, port number, and version of the SSH client. *hostname* is the IP address or hostname of the remote device. • Enter an IPv4 address in dotted decimal format (A.B.C.D).

To enable the SSH server for version 1 and 2, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip ssh server {**enable | port** *port-number*}	CONFIGURATION	Configure the Dell Force10 system as an SCP/SSH server.

To enable the SSH server for version 1 or 2 only, use the following command:

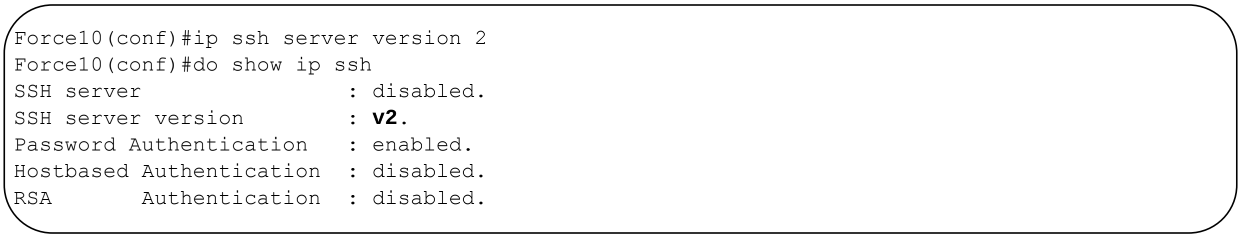
Command Syntax	Command Mode	Purpose
ip ssh server version {**1|2**}	CONFIGURATION	Configure the Dell Force10 system as an SSH server that uses only version 1 or 2.

To view the SSH configuration, use the following command in EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
show ip ssh	EXEC Privilege	Display SSH connection information.

Figure 35-6 on page 639 shows the use of the command **ip ssh server version 2** to enable SSH version 2, and the **show ip ssh** command to confirm the setting.

Figure 35-6. Specifying an SSH version



```
Force10(conf)#ip ssh server version 2
Force10(conf)#do show ip ssh
SSH server                : disabled.
SSH server version        : v2.
Password Authentication   : enabled.
Hostbased Authentication  : disabled.
RSA       Authentication  : disabled.
```

To disable SSH server functions, enter **no ip ssh server enable**.

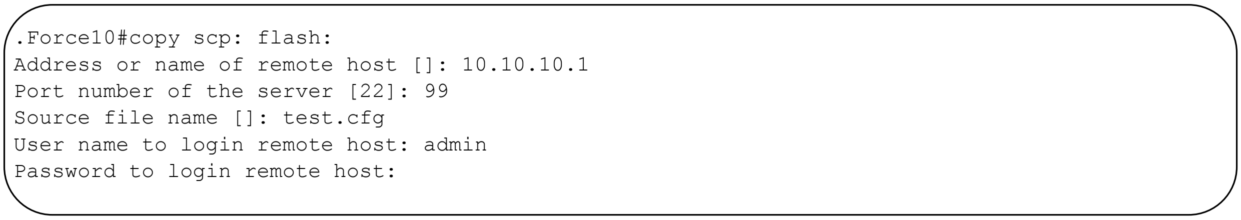
Using SCP with SSH to copy a software image

To use Secure Copy (SCP) to copy a software image through an SSH connection from one switch to another, use the following procedure:

Step	Task	Command Syntax	Command Mode
1	On Chassis One, set the SSH port number (port 22 by default).	**ip ssh server port** *number*	CONFIGURATION
2	On Chassis One, enable SSH.	**ip ssh server enable**	CONFIGURATION
3	On Chassis Two, invoke SCP.	**copy scp: flash:**	CONFIGURATION
4	On Chassis Two, in response to prompts, enter the path to the desired file and enter the port number specified in Step 1.		EXEC Privilege

This 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:

Figure 35-7. Using SCP to copy from an SSH Server on another Switch



```
.Force10#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:
```

Other SSH-related commands 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 hostbased-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 to be used for host-based authentication.
- **ip ssh rhostsfile:** Specify the rhost file to be used for host-based authorization.
- **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.
- **ssh-peer-rpm**: Open an SSH connection to the peer RPM.

Secure Shell Authentication

Secure Shell (SSH) is disabled by default. Enable it using the command **ip ssh server enable**.

SSH supports three methods of authentication:

- SSH Authentication by Password on page 640
- RSA Authentication of SSH on page 641
- Host-based SSH Authentication on page 641

Important Points to Remember for SSH Authentication

- If more than one method is enabled, the order in which the methods are preferred is based on the *ssh_config* file on the Unix machine.
- When all the three authentication methods are enabled, password authentication is the backup method when the RSA method fails.
- The files *known_hosts* and *known_hosts2* are generated when a user tries to SSH using version 1 or version 2, respectively.

SSH Authentication by Password

Authenticate an SSH client by prompting for a password when attempting to connect to the Dell Force10 system. This is the simplest methods of authentication and uses SSH version 1.

Enable SSH password authentication using the command **ip ssh password-authentication enable** from CONFIGURATION mode. View your SSH configuration using the command **show ip ssh** from EXEC Privilege mode.

Figure 35-8. Enabling SSH Password Authentication

```
Force10(conf)#ip ssh server enable
                % Please wait while SSH Daemon initializes ... done.
Force10(conf)#ip ssh password-authentication enable
Force10#sh ip ssh
SSH server                    : enabled.
Password  Authentication      : enabled.
Hostbased Authentication     : disabled.
RSA       Authentication     : disabled.
```

RSA Authentication of SSH

The following procedure authenticates an SSH client based on an RSA key using RSA authentication. This method uses SSH version 2:

Step	Task	Command Syntax	Command Mode
1	On the SSH client (Unix machine), generate an RSA key, as shown in Figure 35-9.		

Figure 35-9. Generating RSA Keys

```
admin@Unix_client#ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/admin/.ssh/id_rsa):
/home/admin/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/admin/.ssh/id_rsa.
Your public key has been saved in /home/admin/.ssh/id_rsa.pub.
```

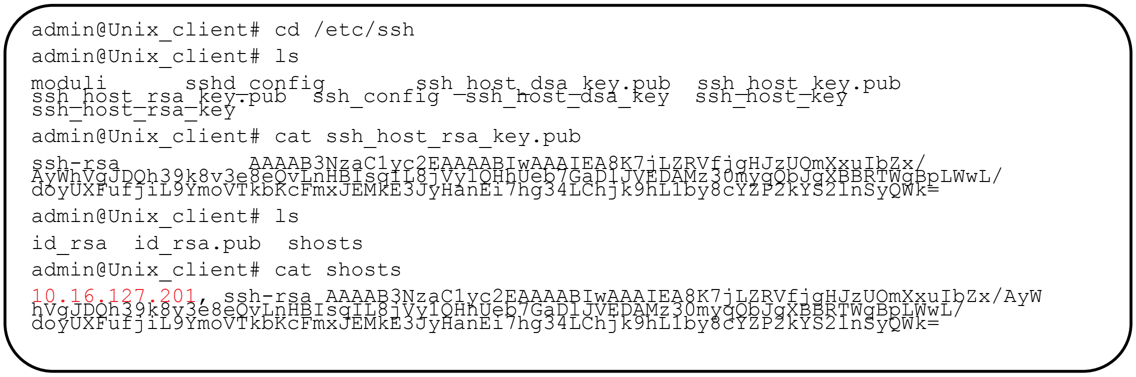
Step	Task	Command Syntax	Command Mode
2	Copy the public key *id_rsa.pub* to the Dell Force10 system.		
3	Disable password authentication if enabled.	**no ip ssh password-authentication enable**	CONFIGURATION
4	Enable RSA authentication.	ip ssh rsa-authentication enable	EXEC Privilege
5	Bind the public keys to RSA authentication.	**ip ssh rsa-authentication my-authorized-keys flash://** *public_key*	EXEC Privilege

Host-based SSH Authentication

Authenticate a particular host. This method uses SSH version 2.

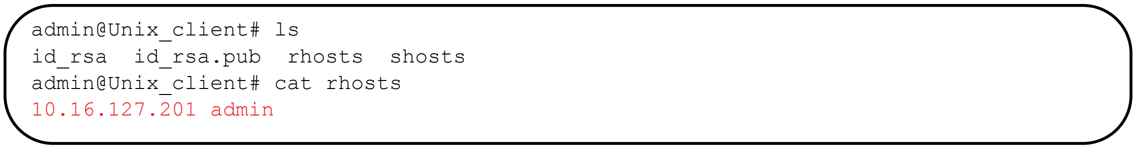
To configure host-based authentication:

<table>
<tr><th>Step</th><th>Task</th><th>Command Syntax</th><th>Command Mode</th></tr>
<tr><td>1</td><td colspan="3">Configure RSA Authentication. See RSA Authentication of SSH, above.</td></tr>
<tr><td>2</td><td>Create shosts by copying the public RSA key to the to the file <i>shosts</i> in the diretory <i>.ssh</i>, and write the IP address of the host to the file.</td><td><b>cp /etc/ssh/ssh_host_rsa_key.pub /.ssh/shosts</b></td><td></td></tr>
<tr><td></td><td colspan="3"><b>Figure 35-10. Creating shosts</b>



<pre>
admin@Unix_client# cd /etc/ssh
admin@Unix_client# ls
moduli   sshd_config   ssh_host_dsa_key.pub ssh_host_key.pub
ssh_host_rsa_key.pub ssh_config ssh_host_dsa_key ssh_host_key
ssh_host_rsa_key
admin@Unix_client# cat ssh_host_rsa_key.pub
ssh-rsa   AAAAB3NzaC1yc2EAAAABIwAAAIEA8K7jLZRVfjgHJzUOmXxuIbZx/
AyWhVgJDQh39k8v3e8eQvLnHBIsqIL8jVy1QHhUeb7GaDlJVEDAMz30myqQbJgXBBRTWgBpLWwL/
doyUXFufjiL9YmoVTkbKcFmxJEMkE3JyHanEi7hg34LChjk9hL1by8cYZP2kYS2lnSyQWk=
admin@Unix_client# ls
id_rsa id_rsa.pub shosts
admin@Unix_client# cat shosts
10.16.127.201, ssh-rsa AAAAB3NzaC1yc2EAAAABIwAAAIEA8K7jLZRVfjgHJzUOmXxuIbZx/AyW
hVgJDQh39k8y3e8eQvLnHBIsqIL8jVy1QHhUeb7GaDlJVEDAMz30myqQbJgXBBRTWgBpLWwL/
doyUXFufjiL9YmoVTkbKcFmxJEMkE3JyHanEi7hg34LChjk9hL1by8cYZP2kYS2lnSyQWk=
</pre></td></tr>
<tr><td>3</td><td colspan="3">Create a list of IP addresses and usernames that are permitted to SSH in a file called <i>rhosts</i>, as shown in Figure 35-11.

<b>Figure 35-11. Creating rhosts</b>



<pre>
admin@Unix_client# ls
id_rsa id_rsa.pub rhosts shosts
admin@Unix_client# cat rhosts
10.16.127.201 admin
</pre></td></tr>
<tr><td>4</td><td colspan="3">Copy the file shosts and rhosts to the Dell Force10 system.</td></tr>
<tr><td>5</td><td>Disable password authentication and RSA authentication, if configured</td><td>• <b>no ip ssh password-authentication</b>
• <b>no ip ssh rsa-authentication</b></td><td>• CONFIGURATION
• EXEC Privilege</td></tr>
<tr><td>6</td><td>Enable host-based authentication.</td><td><b>ip ssh hostbased-authentication enable</b></td><td>CONFIGURATION</td></tr>
<tr><td>7</td><td>Bind shosts and rhosts to host-based authentication.</td><td><b>ip ssh pub-key-file flash://<i>filename</i></b>
<b>ip ssh rhostsfile flash://<i>filename</i></b></td><td>CONFIGURATION</td></tr>
</table>

Client-based SSH Authentication

SSH from the chassis to the SSH client using using the command **ssh** *ip_address*. This method uses SSH version 1 or version 2. If the SSH port is a non-default value, use the command **ip ssh server port** *number*, to change the default port number. You may only change the port number when SSH is disabled. When must then still use the **-p** option with the command **ssh**.

Figure 35-12. Client-based SSH Authentication

```
Force10#ssh 10.16.127.201 ?
-l                         User name option
-p                        SSH server port option (default 22)
-v                        SSH protocol version
```

Troubleshooting SSH

- You may not bind *id_rsa.pub* to RSA authentication while logged in via the console. In this case, Message 2 appears.

Message 2 RSA Authentication Error

```
%Error: No username set for this term.
```

- Host-based authentication must be enabled on the server (Dell Force10 system) and the client (Unix machine). Message 3 appears if you attempt to log in via SSH and host-based is disabled on the client. In this case, verify that host-based authentication is set to “Yes” in the file *ssh_config* (root permission is required to edit this file).

Message 3 Host-based Authentication Error

```
permission denied (host based)
```

- If the IP address in the RSA key does not match the IP address from which you attempt to log in, Message 4 appears. In this case, verify that the name and IP address of the client is contained in the file */etc/hosts*.

Message 4 RSA Authentication Error

```
getname info 8 failed
```

Telnet

To use Telnet with SSH, you must first enable SSH, as described above.

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.

Use the [**no**] **ip telnet server enable** command to enable or disable the Telnet daemon.

```
Force10(conf)#ip telnet server enable
Force10(conf)#no ip telnet server enable
```

VTY Line and Access-Class Configuration

Various methods are available to restrict VTY access in FTOS. These depend on which authentication scheme you use — line, local, or remote:

Table 35-1. 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 FTOS 5.2.1.0 and later)
RADIUS	YES	NO	YES (with FTOS 6.1.1.0 and later)

FTOS provides several ways to configure access classes for VTY lines, including:

- VTY Line Local Authentication and Authorization on page 644
- VTY Line Remote Authentication and Authorization on page 645

VTY Line Local Authentication and Authorization

FTOS 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

Line authentication can be assigned on a per-VTY basis; it is a simple password authentication, using an access-class as authorization.

Local authentication is configured globally. You configure access classes on a per-user basis.

FTOS can assign different access classes to different users by username. Until users attempt to log in, FTOS 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. Once users identify themselves, FTOS retrieves the access class from the local database and applies it. (FTOS also subsequently can close the connection if a user is denied access).

Figure 35-13 shows how to allow or deny a Telnet connection to a user. Users will see a login prompt, even if they cannot login. No access class is configured for the VTY line. It defaults from the local database.

Figure 35-13. Example Access-Class Configuration Using Local Database

```
Force10(conf)#user gooduser password abc privilege 10 access-class permitall
Force10(conf)#user baduser password abc privilege 10 access-class denyall
Force10(conf)#
Force10(conf)#aaa authentication login localmethod local
Force10(conf)#
Force10(conf)#line vty 0 9
Force10(config-line-vty)#login authentication localmethod
Force10(config-line-vty)#end
```

Note: See also the section Chapter 7, Access Control Lists (ACL), Prefix Lists, and Route-maps.

VTY Line Remote Authentication and Authorization

FTOS retrieves the access class from the VTY line.

The Dell Force10 OS takes the access class from the VTY line and applies it to ALL users. FTOS 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, FTOS immediately applies it. If the access-class is **deny all** or **deny for the incoming subnet**, FTOS closes the connection without displaying the login prompt. Figure 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.

Figure 35-14. Example Access Class Configuration Using TACACS+ Without Prompt

```
Force10(conf)#ip access-list standard deny10
Force10(conf-ext-nacl)#permit 10.0.0.0/8
Force10(conf-ext-nacl)#deny any
Force10(conf)#
Force10(conf)#aaa authentication login tacacsmethod tacacs+
Force10(conf)#tacacs-server host 256.1.1.2 key force10
Force10(conf)#
Force10(conf)#line vty 0 9
Force10(config-line-vty)#login authentication tacacsmethod
Force10(config-line-vty)#
Force10(config-line-vty)#access-class deny10
Force10(config-line-vty)#end
(same applies for radius and line authentication)
```

VTY MAC-SA Filter Support

FTOS supports MAC access lists which permit or deny users based on their source MAC address. With this approach, you can implement a security policy based on the source MAC address.

To apply a MAC ACL on a VTY line, use the same **access-class** command as IP ACLs (Figure 35-15). Figure 35-15 shows how to deny incoming connections from subnet 10.0.0.0 without displaying a login prompt.

Figure 35-15. Example Access Class Configuration Using TACACS+ Without Prompt

```
Force10(conf)#mac access-list standard sourcemac
Force10(config-std-mac)#permit 00:00:5e:00:01:01
Force10(config-std-mac)#deny any
Force10(conf)#
Force10(conf)#line vty 0 9
Force10(config-line-vty)#access-class sourcemac
Force10(config-line-vty)#end
```

36

Service Provider Bridging

Service Provider Bridging is supported on platforms: C E S

This chapter contains the following major sections:

- VLAN Stacking on page 647
- VLAN Stacking Packet Drop Precedence on page 658
- Dynamic Mode CoS for VLAN Stacking on page 660
- Layer 2 Protocol Tunneling on page 663
- Provider Backbone Bridging on page 667

VLAN Stacking

VLAN Stacking is supported on platforms: C E S

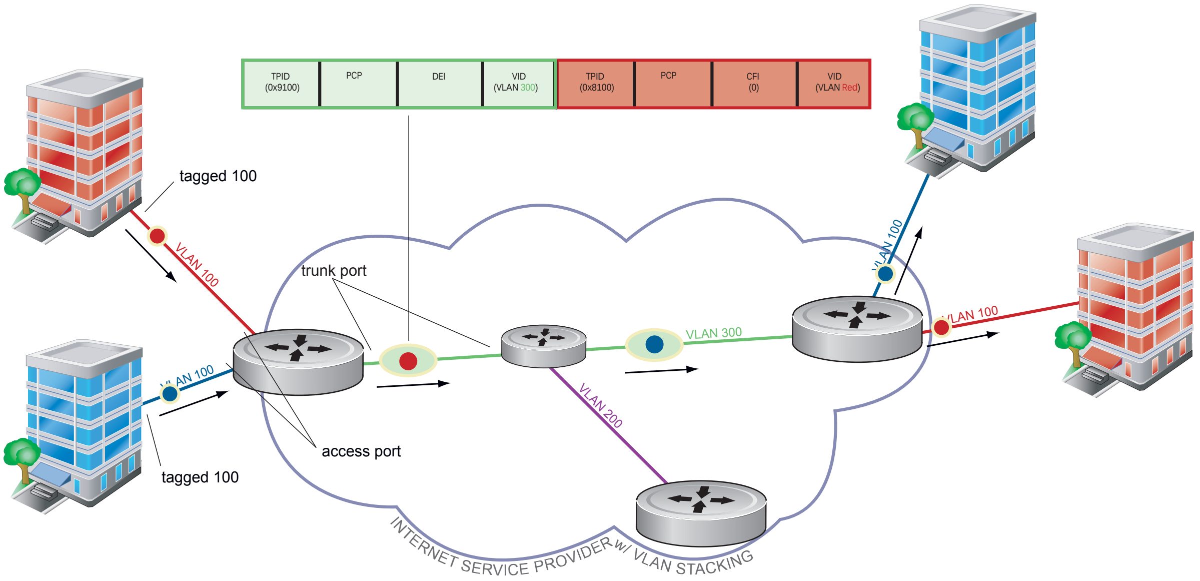
VLAN Stacking, also called Q-in-Q, is defined in IEEE 802.1ad—Provider Bridges, which is an amendment to IEEE 802.1Q—Virtual Bridged Local Area Networks. It enables service providers to use 802.1Q architecture to offer separate VLANs to customers with no coordination between customers, and minimal coordination between customers and the provider.

Using only 802.1Q VLAN tagging all customers would have to use unique VLAN IDs to ensure that traffic is segregated, and customers and the service provider would have to coordinate to ensure that traffic mapped correctly across the provider network. Even under ideal conditions, customers and the provider would still share the 4094 available VLANs.

Instead, 802.1ad allows service providers to add their own VLAN tag to frames traversing the provider network. The provider can then differentiate customers even if they use the same VLAN ID, and providers can map multiple customers to a single VLAN to overcome the 4094 VLAN limitation. Forwarding decisions in the provider network are based on the provider VLAN tag only, so the provider can map traffic through the core independently; the customer and provider need only coordinate at the provider edge.

In 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 (Figure 36-1).

Figure 36-1. 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, these interfaces must be added to a non-default VLAN-Stack-enabled VLAN.
- Dell Force10 cautions against using the same MAC address on different customer VLANs, on the same VLAN-Stack VLAN.
- You can ping across a trunk port only if both systems on the link are an E-Series. You cannot ping across the link if one or both of the systems is a C-Series or S-Series.
- This limitation becomes relevant if you enable the port as a multi-purpose port (carrying single-tagged and double-tagged traffic).

Configure VLAN Stacking

Configuring VLAN-Stacking is a three-step process:

1. Create access and trunk ports. See page 649.
2. Assign access and trunk ports to a VLAN. See page 649.
3. Make the VLAN VLAN-stacking capable.

Related Configuration Tasks

- Configure the Protocol Type Value for the Outer VLAN Tag on page 650
- FTOS Options for Trunk Ports on page 651
- Debug VLAN Stacking on page 652
- VLAN Stacking in Multi-vendor Networks on page 652

Create Access and Trunk Ports

An **access port** is a port on the service provider edge that directly connects to the customer. An access port may belong to only one service provider VLAN.

A **trunk port** is 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.

To create access and trunk ports:

Step	Task	Command Syntax	Command Mode
1	Assign the role of access port to a Layer 2 port on a provider bridge that is connected to a customer.	**vlan-stack access**	INTERFACE
2	Assign the role of trunk port to a Layer 2 port on a provider bridge that is connected to another provider bridge.	vlan-stack trunk	INTERFACE
3	Assign all access ports and trunk ports to service provider VLANs.	**member**	INTERFACE VLAN

Display the VLAN-Stacking configuration for a switchport using the command **show config** from INTERFACE mode, as shown in Figure 36-2.

Figure 36-2. Displaying the VLAN-Stack Configuration on a Layer 2 Port

```
Force10#show run interface gi 7/0
!
interface GigabitEthernet 7/0
 no ip address
 switchport
 vlan-stack access
 no shutdown
Force10#show run interface gi 7/12
!
interface GigabitEthernet 7/12
 no ip address
 switchport
 vlan-stack trunk
 no shutdown
```

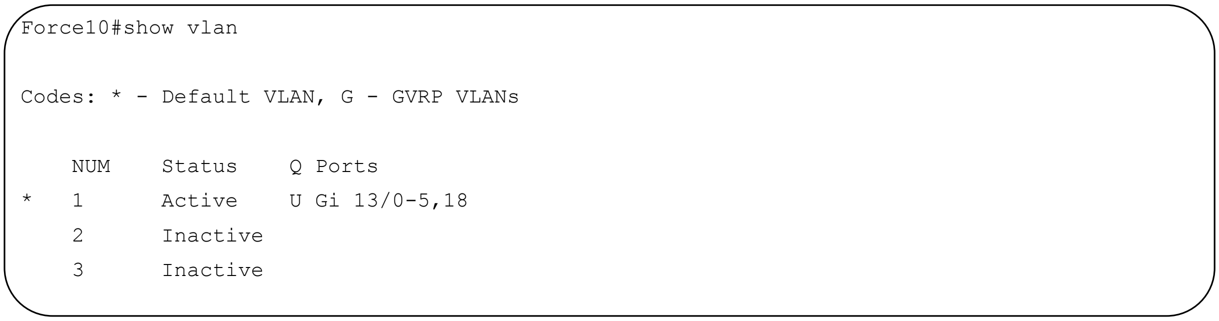
Enable VLAN-Stacking for a VLAN

To enable VLAN-Stacking for a VLAN:

Task	Command Syntax	Command Mode
Enable VLAN-Stacking for the VLAN.	INTERFACE VLAN	**vlan-stack compatible**

Display the status and members of a VLAN using the **show vlan** command from EXEC Privilege mode. Members of a VLAN-Stacking-enabled VLAN are marked with an *M* in column *Q*.

Figure 36-3. Display the Members of a VLAN-Stacking-enabled VLAN



```
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Active    U Gi 13/0-5,18
    2      Inactive
    3      Inactive
```

Configure the Protocol Type Value for the Outer VLAN Tag

The Tag Protocol Identifier (TPID) field of the S-Tag is user-configurable:

Task	Command Syntax	Command Mode
Select a value for the S-Tag TPID. Default: 9100	CONFIGURATION	**vlan-stack protocol-type**

Display the S-Tag TPID for a VLAN using the command **show running-config** from EXEC privilege mode. FTOS displays the S-Tag TPID only if it is a non-default value.

FTOS 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.

Step	Task	Command Syntax	Command Mode
1	Configure a trunk port to carry untagged, single-tagged, and double-tagged traffic by making it a hybrid port. **Note:** Note: On the C-Series and S-Series, a trunk port can be added to an 802.1Q VLAN as well as a Stacking VLAN only when the TPID 0x8100.	**portmode hybrid**	INTERFACE
2	Add the port to a 802.1Q VLAN as tagged or untagged.	[**tagged** | **untagged**]	INTERFACE VLAN

In Figure 36-4 GigabitEthernet 0/1 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.

Figure 36-4. Hybrid Port as VLAN-Stack Trunk Port and as Member of other VLANs

```
Force10(conf)#int gi 0/1
Force10(conf-if-gi-0/1)#portmode hybrid
Force10(conf-if-gi-0/1)#switchport
Force10(conf-if-gi-0/1)#vlan-stack trunk
Force10(conf-if-gi-0/1)#show config
!
interface GigabitEthernet 0/1
 no ip address
 portmode hybrid
 switchport
 vlan-stack trunk
 shutdown
Force10(conf-if-gi-0/1)#interface vlan 100
Force10(conf-if-vl-100)#untagged gigabitethernet 0/1
Force10(conf-if-vl-100)#interface vlan 101
Force10(conf-if-vl-101)#tagged gigabitethernet 0/1
Force10(conf-if-vl-101)#interface vlan 103
Force10(conf-if-vl-103)#vlan-stack compatible
Force10(conf-if-vl-103-stack)#member gigabitethernet 0/1
Force10(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
```

Debug VLAN Stacking

To debug the internal state and membership of a VLAN and its ports, use the **debug member** command, as shown in Figure 36-5. The port notations in Figure 36-5 are as follows:

- **MT** — stacked trunk
- **MU** — stacked access port
- **T**— 802.1Q trunk port
- **U**— 802.1Q access port
- **NU**— Native VLAN (untagged)

Figure 36-5. Example of Output of debug member vlan and debug member port

```
Force10# 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)

Force10#debug member port gigabitethernet 2/47
vlan id   : 603 (MT), 100(T), 101(NU)
Force10#
```

VLAN Stacking in Multi-vendor Networks

The first field in the VLAN tag is the Tag Protocol Identifier (TPID), which is two bytes. In a VLAN-stacking network, once 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 two-byte value; FTOS uses 0x9100 (Figure 36-6) while non-Dell Force10 systems might use a different value.

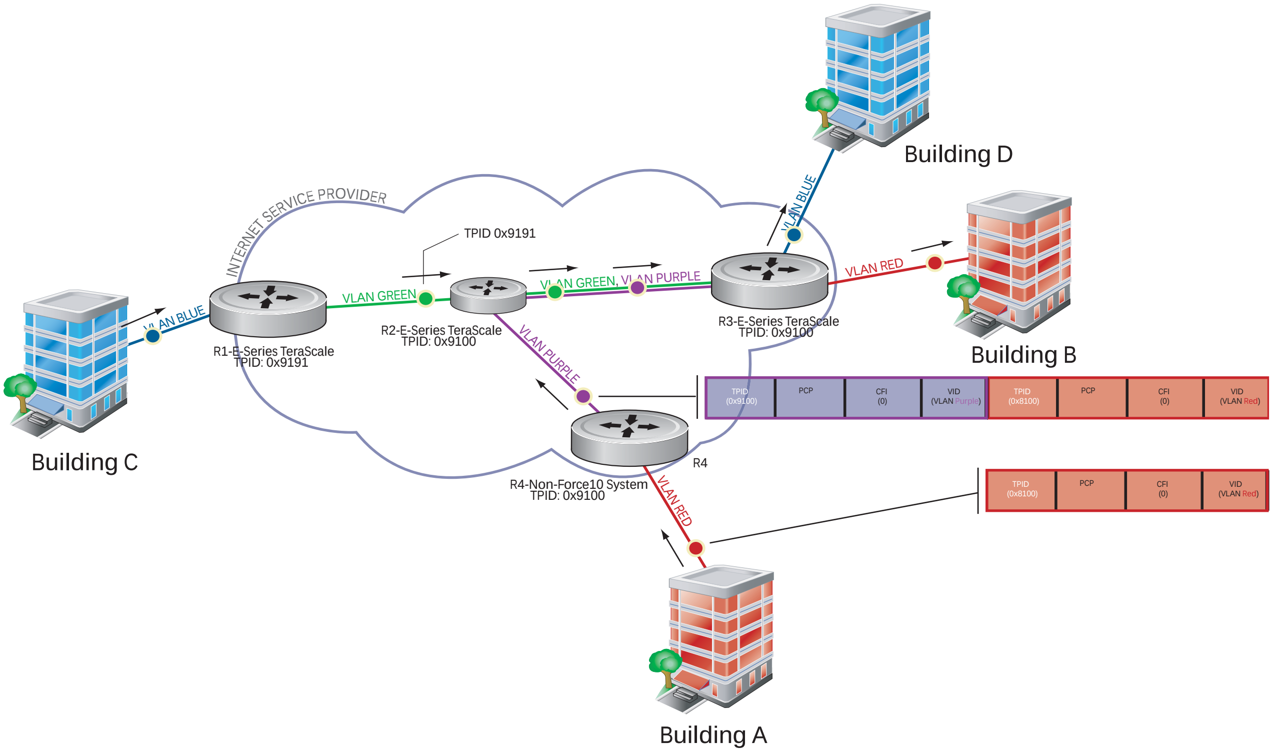
If the next-hop system’s TPID does not match the outer-tag TPID of the incoming frame, the system drops the frame. For example, in Figure 36-6, the frame originating from Building A is tagged VLAN RED, and then double-tagged VLAN PURPLE on egress at R4. The TPID on the outer tag is 0x9100. R2’s TPID must also be 0x9100, and it is, so R2 forwards the frame.

Given the matching-TPID requirement, there are limitations when you employ Dell Force10 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 with E-Series TeraScale Systems

The default TPID for the outer VLAN tag is 0x9100. Although the TPID field is 16 bits, E-Series TeraScale only uses the first eight to make forwarding decisions, and as such makes no distinction between 0x8100 and any other TPID beginning with 0x81, for example, 0x8181. You can configure the first eight bits of the TPID using the command **vlan-stack protocol-type** command. In Figure 36-6, the frame originating from Building C is tagged 0x9191 on egress of R1. R2’s TPID is 0x9100, but it its an E-Series TeraScale system and makes no distinction between 0x9191 and 0x9100, so it forwards the frame.

Figure 36-6. TPID Match and First-byte Match on the E-Series TeraScale

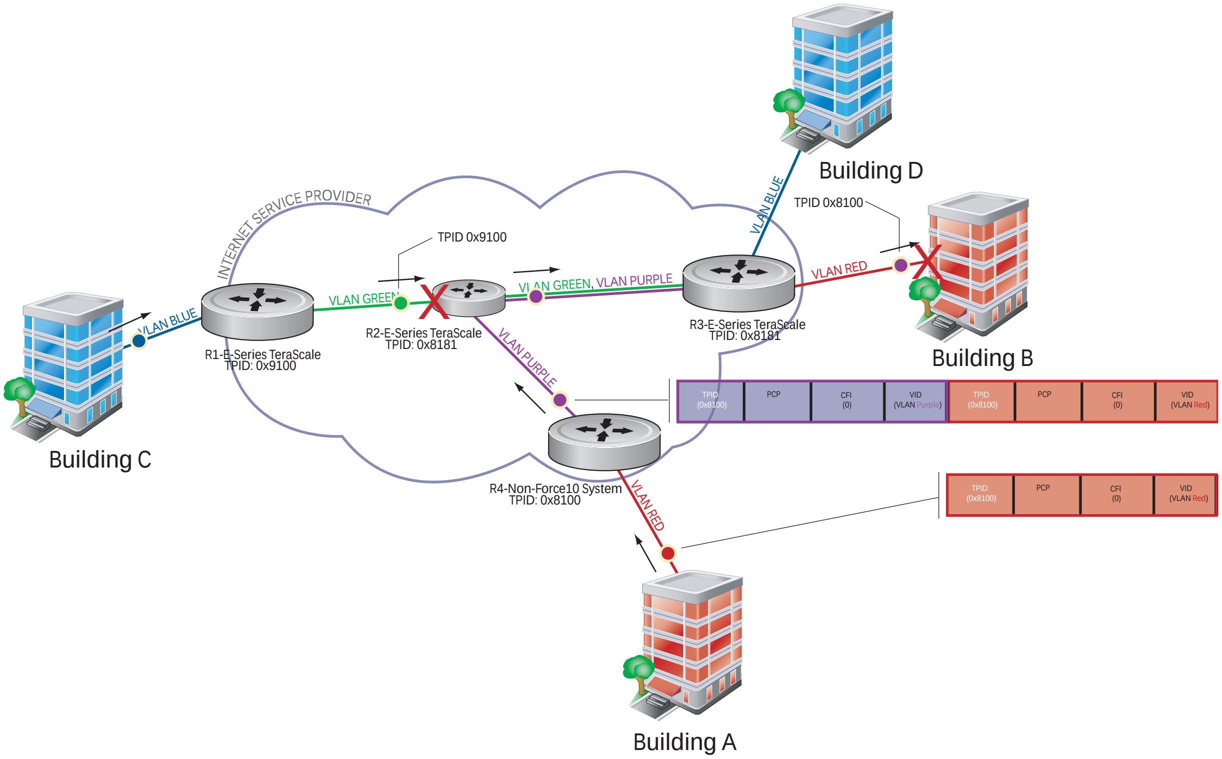


TPID 0x8100 on E-Series TeraScale Systems

E-Series TeraScale treats TPID 0x8100 as a normal VLAN even when on the outer tag. E-Series TeraScale makes forwarding decisions based strictly on the protocol type, without regard for whether the port is an access port. Therefore, when the outer tag has TPID 0x8100, the system does not remove it from frames egressing an access port. Still, although the frames cannot be decapsulated, the system is able to switch them. In Figure 36-7, the frame originating from Building A is double tagged on egress at R4 and is switched towards Building B, but is not decapsulated on egress at R2 because its TPID is 0x8181.

FTOS Behavior: The E-Series ExaScale and TeraScale forward frames with TPID 0x8100 even when its own TPID is not 0x8100. This behavior is required to service ARP and PVST packets, which use TPID 0x8100.

Figure 36-7. TPID Mismatch and 0x8100 Match on the E-Series TeraScale



VLAN Stacking with E-Series ExaScale Systems

E-Series ExaScale, beginning with FTOS version 8.2.1.0, allows you to configure both bytes of the 2-byte TPID. TeraScale systems allow you to configure the first byte only and thus, the system did not differentiate between TPIDs with a common first byte. For example 0x9100 and 0x91A8 were treated as the same TPID. In Figure 36-6, R2 forwards the frame with TPID 0x9191 which originated from Building C. In contrast, R2 drops the frame with TPID 0x9191 originating from Building C in Figure 36-8 because the frames TPID does not match both bytes of its own TPID.

FTOS Behavior: The E-Series ExaScale and TeraScale forwards frames with TPID 0x8100 even when its own TPID is not 0x8100. This behavior is required to service ARP and PVST packets, which use TPID 0x8100.

Figure 36-8. First-byte TPID Match on the E-Series ExaScale

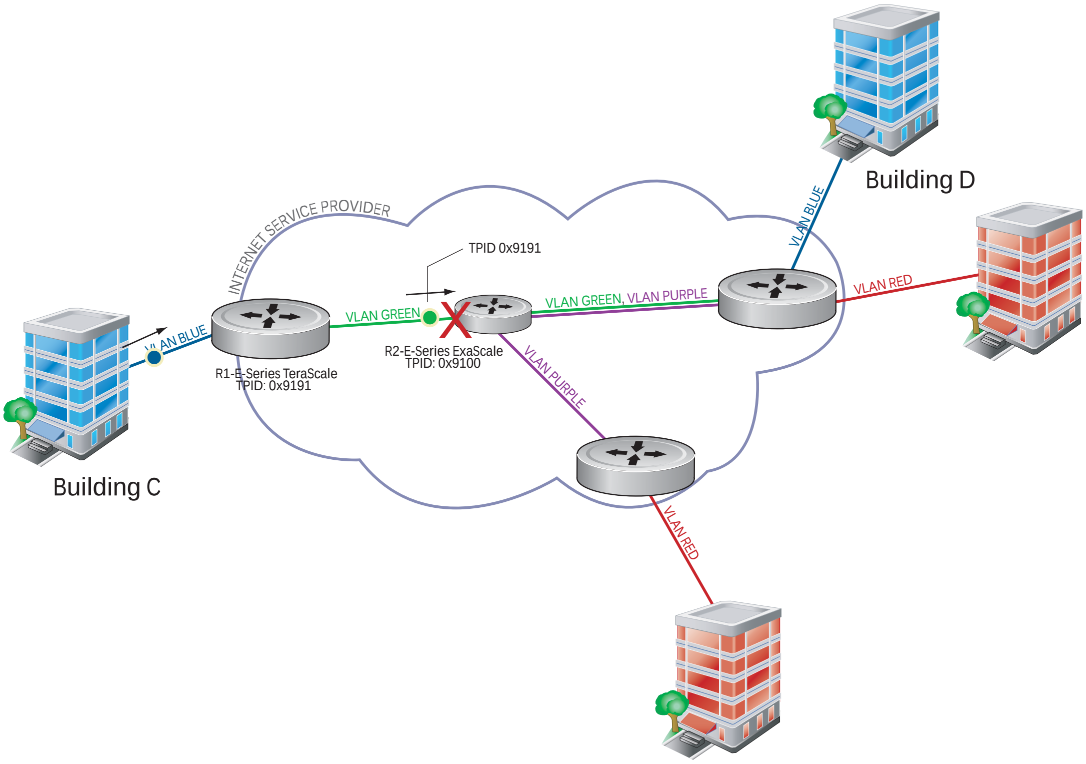


Table 36-1 details the outcome of matched and mis-matched TPIDs in a VLAN-stacking network with the E-Series.

Table 36-1. E-Series Behaviors for Mis-matched TPID

Network Position	Incoming Packet TPID	System TPID	Match Type	TeraScale Behavior	ExaScale Behavior
Core	0xUV**WX**	0xUV**YZ**	1st-byte match	switch as 0xUV**YZ**	drop
	0xUVWZ	0xQRST	mismatch	drop	drop
Egress Access Point	0xUV**WX**	0xUV**YZ**	1st-byte match	switch as 0xUV**YZ**	drop
	0x81**WX**	0x81**YZ**	1st-byte match	switch as is (no decapsulation)	drop
	0xUVWZ	0xQRST	mismatch	drop	drop

VLAN Stacking with C-Series and S-Series

The default TPID for the outer VLAN tag is 0x9100. Beginning with FTOS version 8.2.1.0, both the C-Series and S-Series allow 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 Figure 36-9. Versions 8.2.1.0 and later differentiate between 0x9100 and 0x91XY, as shown in Figure 36-11.

You can configure the first eight bits of the TPID using the command **vlan-stack protocol-type**.

The TPID on the C-Series and S-Series systems 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, then 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 in Figure 36-11. For the same traffic types, if you configure TPID 0x8100, then 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 in Figure 36-11.

Therefore, a mismatched TPID results in the port not differentiating between tagged and untagged traffic.

Figure 36-9. Single and Double-tag TPID Match on the C-Series and S-Series

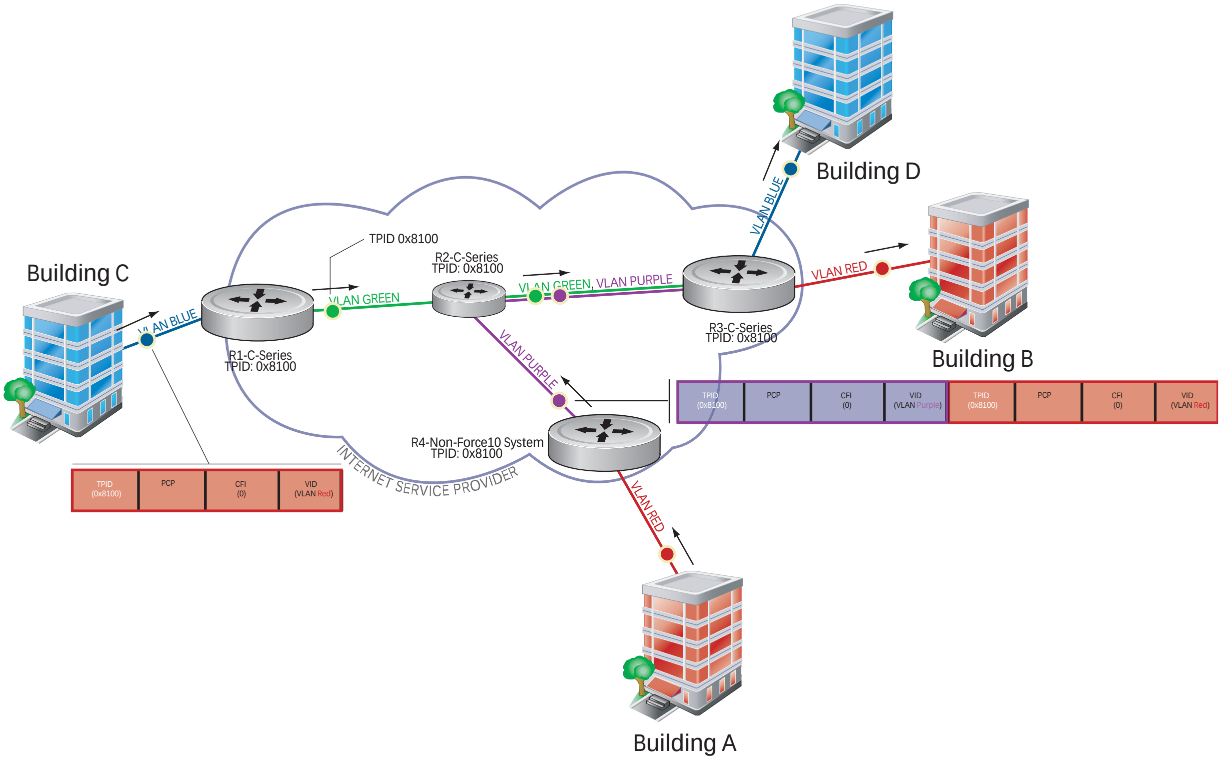


Figure 36-10. Single and Double-tag First-byte TPID Match on C-Series and S-Series

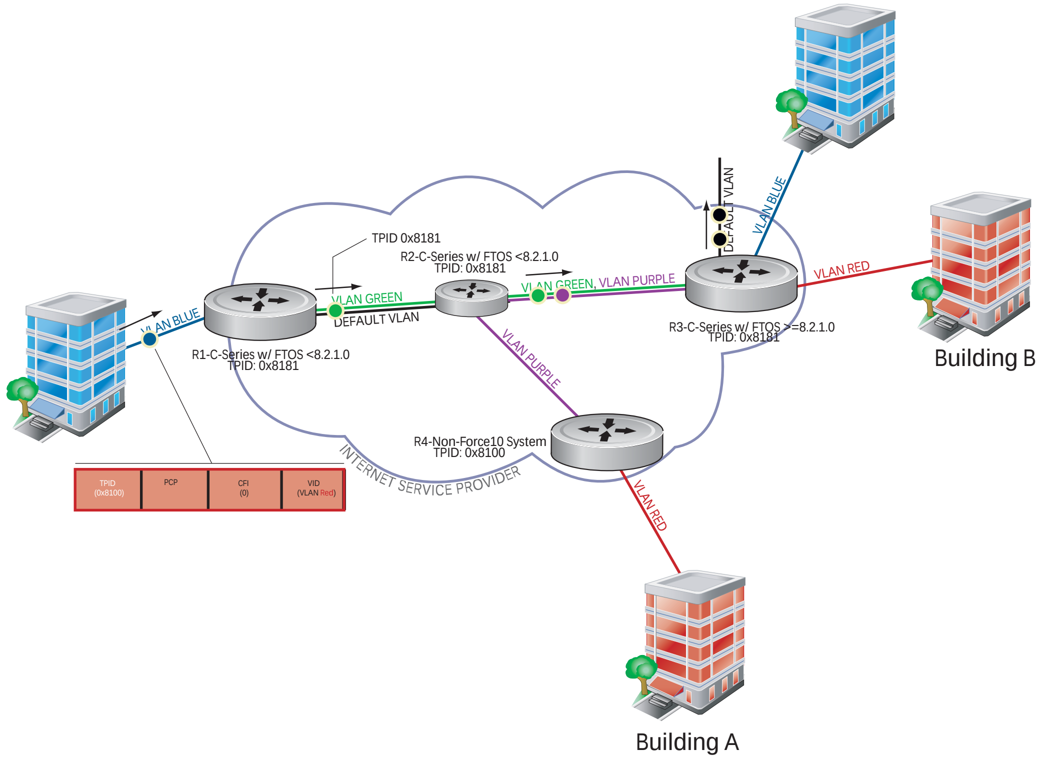


Figure 36-11. Single and Double-tag TPID Mismatch on the C-Series and S-Series

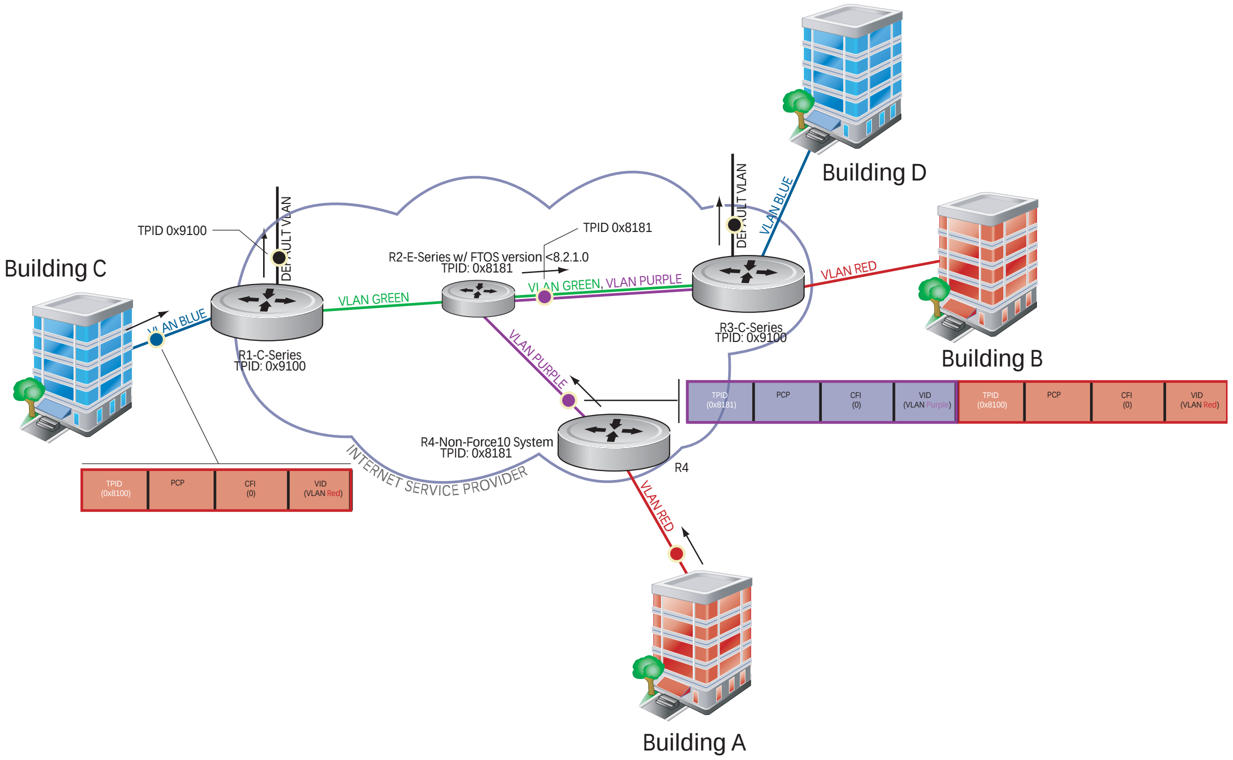


Table 36-2 details the outcome of matched and mismatched TPIDs in a VLAN-stacking network with the C-Series and S-Series.

Table 36-2. C-Series and S-Series Behaviors for Mis-matched TPID

Network Position	Incoming Packet TPID	System TPID	Match Type	Pre-8.2.1.0	8.2.1.0+
Ingress Access Point	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	single-tag (0x8100)	0xUVWX	single-tag mismatch	switch to default VLAN	switch to default VLAN
		0x8100	single-tag match	switch to VLAN	switch to VLAN
		0x81XY	single-tag first-byte match	switch to VLAN	switch to default VLAN
Core	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	double-tag 0xUVWX	0xUVWX	double-tag match	switch to VLAN	switch to VLAN
		0xUVYZ	double-tag first-byte match	switch to VLAN	switch to default VLAN
		0xQRST	double-tag mismatch	switch to default VLAN	switch to default VLAN
Egress Access Point	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	double-tag 0xUVWX	0xUVWX	double-tag match	switch to VLAN	switch to VLAN
		0xUVYZ	double-tag first-byte match	switch to VLAN	switch to default VLAN
		0xQRST	double-tag mismatch	switch to default VLAN	switch to default VLAN

VLAN Stacking Packet Drop Precedence

VLAN Stacking Packet Drop Precedence is available only on platform: C S

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.

Enable Drop Eligibility

You must enable Drop Eligibility globally before you can honor or mark the DEI value.

Task	Command Syntax	Command Mode
Make packets eligible for dropping based on their DEI value. By default, packets are colored green, and DEI is marked 0 on egress.	**dei enable**	CONFIGURATION

When Drop Eligibility is enabled, DEI mapping or marking takes place according to the defaults. In this case, the CFI is affected according to Table 36-3.

Table 36-3. 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

Honor the Incoming DEI Value

To honor the incoming DEI value, you must explicitly map the DEI bit to an FTOS 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).

Task	Command Syntax	Command Mode
Honor the incoming DEI value by mapping it to an FTOS drop precedence. You may enter the command once for 0 and once for 1. Packets with an unmapped DEI value are colored green.	**dei honor** {**0** | **1**} {**green** | **red** | **yellow**}	INTERFACE
Display the DEI-honoring configuration.	**show interface dei-honor** [**interface** *slot/ port* | **linecard** *number* **port-set** *number*]	EXEC Privilege

Task	Command Syntax	Command Mode

```
Force10#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
```

Mark Egress Packets with a DEI Value

On egress, you can set the DEI value according to a different mapping than ingress (see Honor the Incoming DEI Value).

Task	Command Syntax	Command Mode
Set the DEI value on egress according to the color currently assigned to the packet.	**dei mark** {**green** | **yellow**} {**0** | **1**}	INTERFACE
Display the DEI-marking configuration.	**show interface dei-mark** [*interface slot/ port* | **linecard** *number* **port-set** *number*]	EXEC Privilege

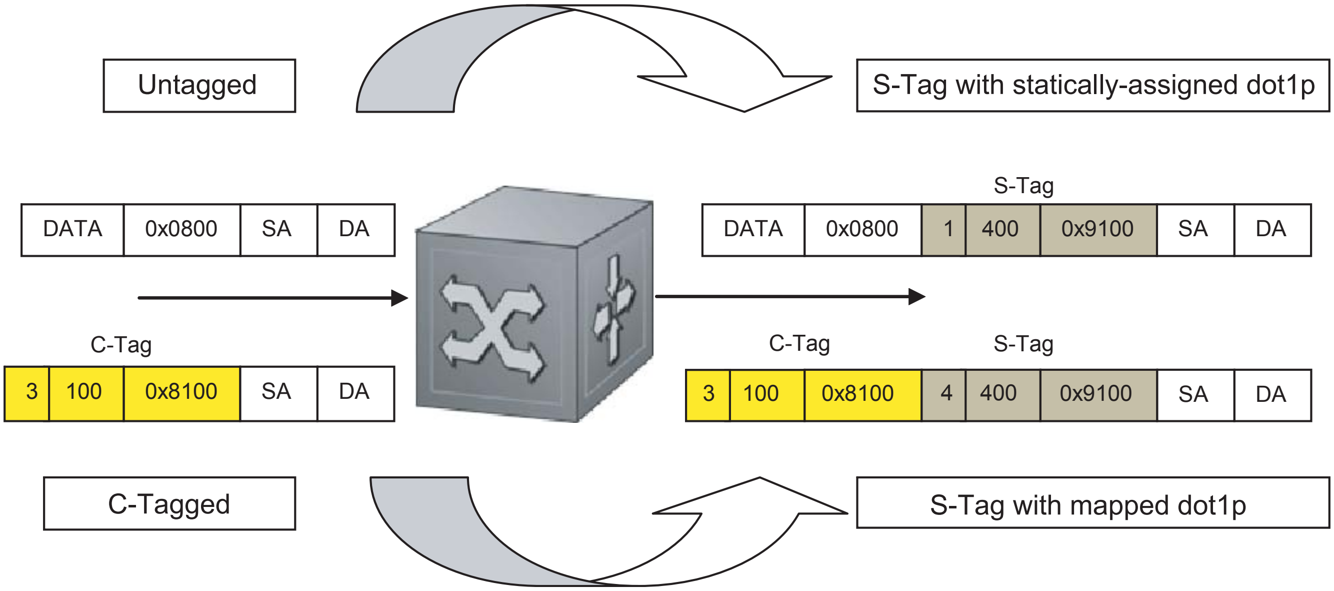
```
Force10#show interface dei-mark

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

Dynamic Mode CoS for VLAN Stacking is available only on platforms: C S

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 36-12. Statically and Dynamically Assigned dot1p for VLAN Stacking



When configuring Dynamic Mode CoS, you have two options:

a 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.

b 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, then 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 up to 8 entries to be installed 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 CAM tables.

FTOS Behavior: For Option A above, 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 will be rate policed at 40Mbps (qos-policy-input for queue 3) since class-map "a" of Queue 3 also matches the traffic. This behavior is expected.

```
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 configuration below, 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**.

A policy map output with rate shape for different queues can also be used.

```
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
```

To map C-Tag dot1p values to S-Tag dot1p values and mark the frames accordingly:

Step	Task	Command Syntax	Command Mode
1	Allocate CAM space to enable queuing frames according to the C-Tag or the S-Tag. **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.	**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* Default: 0 FP blocks for vman-qos and vman-qos-dual-fp	CONFIGURATION
2	The new CAM configuration is stored in NVRAM and takes effect only after a save and reload.	**copy running-config startup-config** **reload**	EXEC Privilege
3	Map C-Tag dot1p values to a S-Tag dot1p value. C-Tag values may be separated by commas, and dashed ranges are permitted. Dynamic Mode CoS overrides any Layer 2 QoS configuration in case of conflicts.	**vlan-stack dot1p-mapping c-tag-dot1p** *values* **sp-tag-dot1p** *value*	INTERFACE

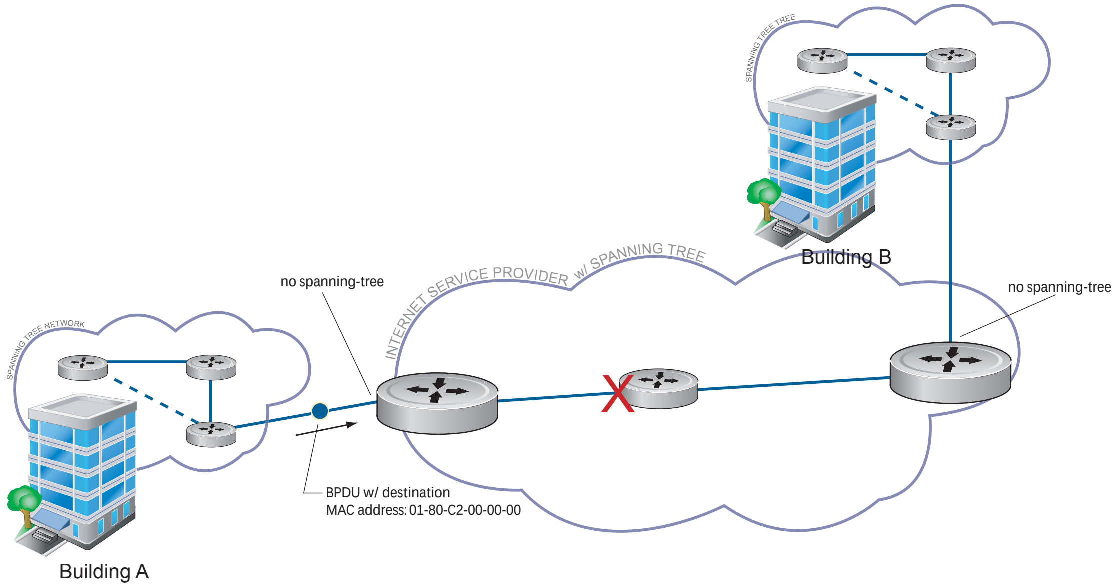
Note: Since **dot1p-mapping** marks *and* queues packets, the only remaining applicable QoS configuration is rate metering. You may use Rate Shaping or Rate Policing.

Layer 2 Protocol Tunneling

Layer 2 Protocol Tunneling (L2PT) is supported on platforms: C E S

Spanning Tree 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 LAN recognize this address and process the BPDU. When VLAN stacking is used to connect physically separate regions of a network, BPDUs attempting to traverse the intermediate network might be consumed and subsequently dropped because the intermediate network itself might be using Spanning Tree (Figure 36-13).

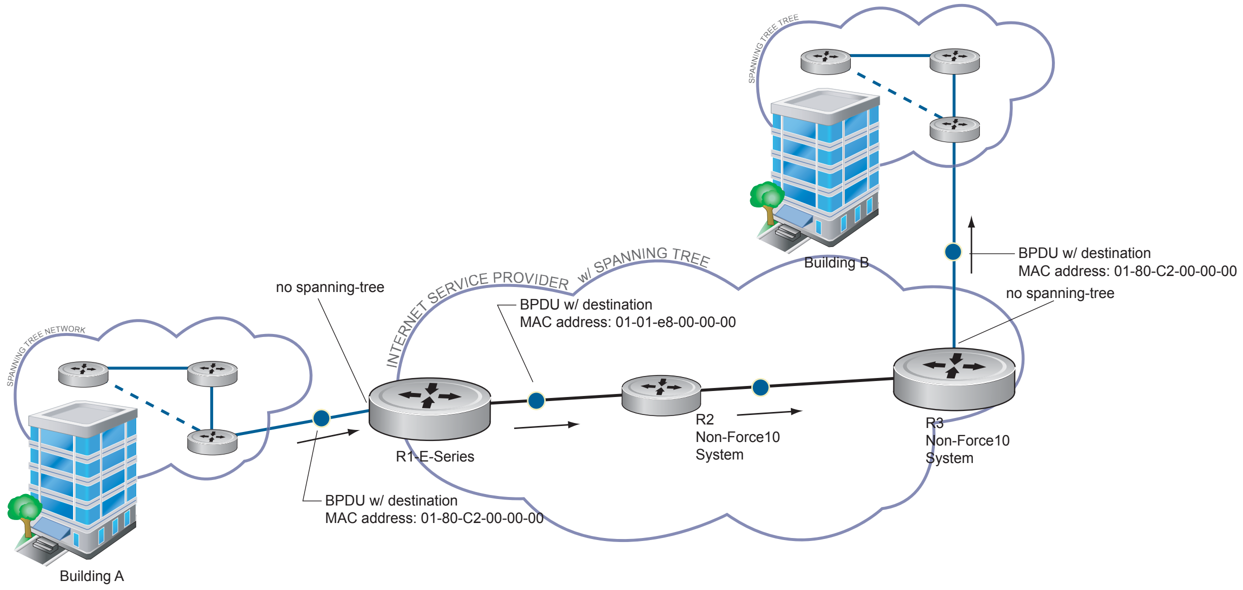
Figure 36-13. VLAN Stacking without L2PT



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. Since 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 (Figure 36-14).

FTOS Behavior: In FTOS versions prior to 8.2.1.0, the MAC address that Dell Force10 systems use to overwrite the Bridge Group Address on ingress was non-configurable. The value of the L2PT MAC address was the Dell Force10-unique MAC address, 01-01-e8-00-00-00. As such, with these FTOS versions, Dell Force10 systems are required at the egress edge of the intermediate network because only FTOS could recognize the significance of the destination MAC address and rewrite it to the original Bridge Group Address. In FTOS version 8.2.1.0 and later, the L2PT MAC address is user-configurable, so you can specify an address that non-Dell Force10 systems can recognize and rewrite the address at egress edge.

Figure 36-14. VLAN Stacking with L2PT



Implementation Information

- L2PT is available for STP, RSTP, MSTP, and PVST+ BPDUs.
- No protocol packets are tunneled when VLAN Stacking is enabled.
- L2PT requires the default CAM profile.

Enable Layer 2 Protocol Tunneling

Step	Task	Command Syntax	Command Mode
1	Verify that the system is running the default CAM profile; you must use this CAM profile for L2PT.	**show cam-profile**	EXEC Privilege
2	Enable protocol tunneling globally on the system.	**protocol-tunnel enable**	CONFIGURATION
3	Tunnel BPDUs the VLAN.	**protocol-tunnel stp**	INTERFACE VLAN

Specify a Destination MAC Address for BPDUs

By default, FTOS uses a Dell Force10-unique MAC address for tunneling BPDUs. You can configure another value.

Task	Command Syntax	Command Mode
Overwrite the BPDU with a user-specified destination MAC address when BPDUs are tunneled across the provider network. Default: 01:01:e8:00:00:00	**protocol-tunnel destination-mac**	CONFIGURATION

Rate-limit BPDUs on the E-Series

In order to rewrite the destination MAC address on BPDUs, they are forwarded to the RPM. You can rate-limit BPDUs to protect the RPM, in which case the system drops BPDUs when the threshold is reached.

Task	Command Syntax	Command Mode
Set a maximum rate at which the RPM will process BPDUs for L2PT. Default: 75 pps E-Series Range: 75 to 3000 pps	**protocol-tunnel rate-limit**	CONFIGURATION

Rate-limit BPDUs on the C-Series and S-Series

CAM space is allocated in sections called Field Processor (FP) blocks.

There are total 13 user-configurable FP blocks on the C-Series and S-Series. The default number of blocks for L2PT is 0; you must allocate at least one to enable BPDU rate-limiting.

Step	Task	Command Syntax	Command Mode
1	Create at least one FP group for L2PT. See CAM Allocation on page 226 for details on this command.	**cam-acl l2acl**	CONFIGURATION
2	Save the running-config to the startup-config.	**copy running-config startup-config**	EXEC Privilege
3	Reload the system.	**reload**	EXEC Privilege
4	Set a maximum rate at which the RPM will process BPDUs for L2PT. Default: no rate limiting C-Series Range: 64 to 640 kbps S-Series Range: 64 to 320 kbps	**protocol-tunnel rate-limit**	VLAN STACKING

Debug Layer 2 Protocol Tunneling

Task	Command Syntax	Command Mode
Display debugging information for L2PT.	**debug protocol-tunnel**	EXEC Privilege

Provider Backbone Bridging

Provider Backbone Bridging is supported only on platforms: C S

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 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.

Task	Command Syntax	Command Mode
Use the Provider Bridge Group address as the destination MAC address in BPDUs. The xstp keyword applies this functionality to STP, RSTP, and MSTP; this functionality is not available for PVST+.	**bpdu-destination-mac-address** [**stp** | **gvrp**] **provider-bridge-group**	CONFIGURATION

37

sFlow

Configuring sFlow is supported on platforms C E S

Overview

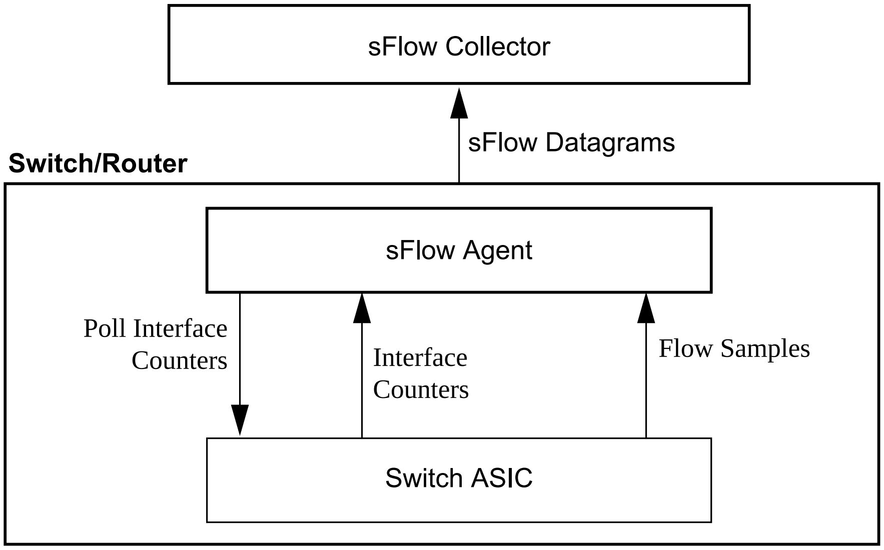
FTOS supports sFlow version 5. sFlow is a standard-based sampling technology embedded within switches and routers which is used to monitor network traffic. It is designed to provide traffic monitoring for high speed networks with many switches and routers. sFlow uses two types of sampling:

- 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 consists of information on, but not limited to, packet header, ingress and egress interfaces, sampling parameters, and interface counters.

Packet sampling is typically done by the ASIC. sFlow Collector analyses the sFlow datagrams received from different devices and produces a network-wide view of traffic flows.

Figure 37-1. sFlow Traffic Monitoring System



Implementation Information

Dell Force10' sFlow is designed so that the hardware sampling rate is per line card port-pipe and is decided based upon all the ports in that port-pipe. If sFlow is not enabled on any port specifically, then the global sampling rate is downloaded to that port and is to calculate the port-pipe's lowest sampling rate. This design supports, then, the possibility that sFlow might be configured on that port in the future. Back-off is triggered based on the port-pipe's hardware sampling rate.

For example, if port 1 in a the port-pipe has sFlow configured with a 16384 sampling rate while port 2 in the port-pipe has sFlow configured but no sampling rate set, FTOS applies a global sampling rate of 512 to port 2. The hardware sampling rate on the port-pipe is ten set at 512 because that is the lowest configured rate on the port-pipe. When a high traffic situation occurs, a back-off is triggered and the hardware sampling rate is backed-off from 512 to 1024. Note that port 1 maintains its sampling rate of 16384; port 1 is unaffected because it maintain its configured sampling rate of 16484.

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 FTOS implementation of the sFlow MIB supports sFlow configuration via **snmpset**.
- Collection through management interface is supported on E-Series only
- Dell Force10 recommends that the sFlow Collector be connected to the Dell Force10 chassis through a line card port rather than the RPM Management Ethernet port.
- E-Series TeraScale sFlow sampling is done on a per-port-pipe basis.
- E-Series ExaScale, C-Series , and S-Series sFlow sampling is done on a per-port basis.

- FTOS exports all sFlow packets to the collector. A small sampling rate can equate to a large number of exported packets. A backoff mechanism will automatically be 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, will always be zero.
- Community list and local preference fields are not filled in extended gateway element in 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 packet being sampled is redirected using PBR (Policy-Based Routing), sFlow datagram may contain incorrect extended gateway/router information.
- Source VLAN field in the extended switch element will not be packed in case of routed packet.
- Destination VLAN field in the extended switch element will not be packed in case of Multicast packet.
- On the S-Series, up to 700 packets can be sampled and processed per second.
- On the C-Series up to 1000 packets can be sampled and processed per second.
- On the E-Series, the maximum number of packets that can be sampled and processed per second is:
    - 7500 packets when no extended information packing is enabled.
    - 1000 packets when only extended-switch information packing is enabled.
    - 1600 packets when extended-router and/or extended-gateway information packing is enabled.

Enable and Disable sFlow

By default, sFlow is *disabled* globally on the system. To enable sFlow globally, use the **sflow enable** command in CONFIGURATION mode. Use the **no** version of this command to disable sFlow globally.

Command Syntax	Command Mode	Usage
[**no**] **sflow enable**	CONFIGURATION	Enable sFlow globally.

Enable and Disable on an Interface

By default, sFlow is *disabled* on all interfaces. To enable sFlow on a specific interface, use the **sflow enable** command in INTERFACE mode. Use the **no** version of this command to disable sFlow on an interface. This CLI is supported on physical ports and LAG ports.

Command Syntax	Command Mode	Usage
[**no**] **sflow enable**	INTERFACE	Enable sFlow on an interface.

sFlow Show Commands

FTOS includes the following sFlow display commands:

- Show sFlow Globally on page 49
- Show sFlow on an Interface on page 50
- Show sFlow on a Line Card on page 50

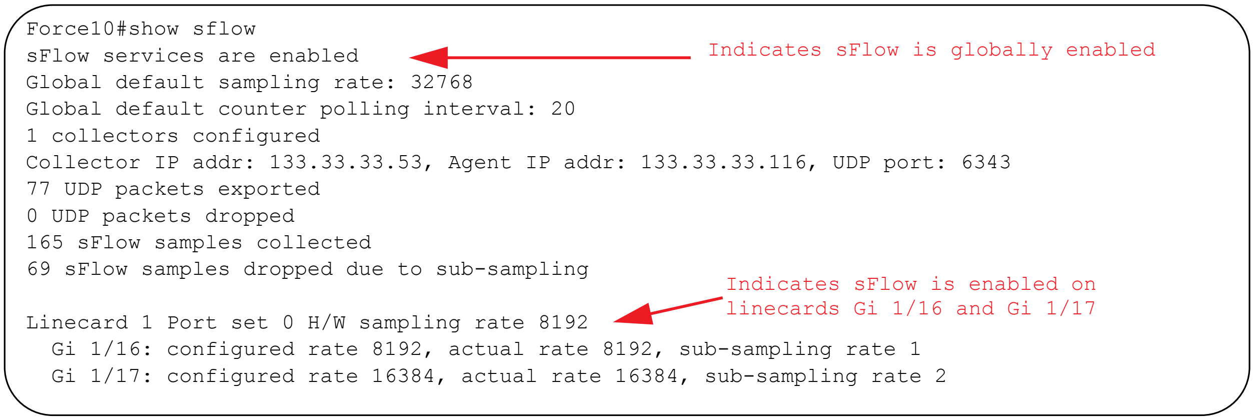
Show sFlow Globally

Use the following command to view sFlow statistics:

Command Syntax	Command Mode	Purpose
show sflow	EXEC	Display sFlow configuration information and statistics.

Figure 37-2 is a sample output from the **show sflow** command:

Figure 37-2. Command Example: show sflow



```
Force10#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

Linecard 1 Port set 0 H/W sampling rate 8192
  Gi 1/16: configured rate 8192, actual rate 8192, sub-sampling rate 1
  Gi 1/17: configured rate 16384, actual rate 16384, sub-sampling rate 2
```

Indicates sFlow is globally enabled

Indicates sFlow is enabled on linecards Gi 1/16 and Gi 1/17

Show sFlow on an Interface

Use the following command to view sFlow information on a specific interface:

Command Syntax	Command Mode	Purpose
show sflow interface *interface-name*	EXEC	Display sFlow configuration information and statistics on a specific interface.

Figure 37-3 is a sample output from the **show sflow interface** command.

Figure 37-3. Command Example: show sflow interface

```
Force10#show sflow interface gigabitethernet 1/16
Gi 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
```

The configuration, shown in Figure 37-2, is also displayed in the running configuration (Figure 37-4):

Figure 37-4. Command Example: show running-config interface

```
Force10#show running-config interface gigabitethernet 1/16
!
interface GigabitEthernet 1/16
 no ip address
 mtu 9252
 ip mtu 9234
 switchport
 sflow enable
 sflow sample-rate 8192
 no shutdown
```

Show sFlow on a Line Card

Use the following command to view sFlow statistitics on a specified line card:

Command Syntax	Command Mode	Purpose
show sflow linecard *slot-number*	EXEC	Display sFlow configuration information and statistics on the specified interface.

Figure 37-5 is a sample output from the **show sflow linecard** command:

Figure 37-5. Command Example: show sflow linecard

```
Force10#show sflow linecard 1
Linecard 1
  Samples rcvd from h/w            :165
  Samples dropped for sub-sampling :69
  Total UDP packets exported       :77
  UDP packets exported via RPM     :77
  UDP packets dropped              :
```

Specify Collectors

The **sflow collector** command allows identification of sFlow Collectors to which sFlow datagrams are forwarded. The user can specify up to two sFlow collectors. If two Collectors are specified, the samples are sent to both.

Collection through Management interface is supported on platform: E.

Command Syntax	Command Mode	Usage
sflow collector *ip-address* **agent-addr** *ip-address* [*number* [**max-datagram-size** *number*]] | [**max-datagram-size** *number*]	CONFIGURATION	Identify sFlow collectors to which sFlow datagrams are forwarded. Default UDP port: 6343 Default max-datagram-size: 1400

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 to be 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.

The polling interval can be configured globally (in CONFIGURATION mode) or by interface (in INTERFACE mode) by executing the interval command:

Command Syntax	Command Mode	Usage
sflow polling-interval *interval value*	CONFIGURATION or INTERFACE	Change the global default counter polling interval. *interval value*—in seconds. Range: 15 to 86400 seconds Default: 20 seconds

Sampling Rate

Sampling Rate is supported on platform E T .

The sFlow sampling rate is the number of packets that are skipped before the next sample is taken. sFlow does not have time-based packet sampling.

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 number and re-enter the command. (For more information on values in power-of-2, see Sub-sampling on page 675.)

The sample rate can be configured globally or by interface using the sample rate command:

Command Syntax	Command Mode	Usage
[**no**] **sflow sample-rate** *sample-rate*	CONFIGURATION or INTERFACE	Change the global or interface sampling rate. Rate must be entered in factors of 2 (eg, 4096, 8192). *sample-rate* range: 256-8388608 for C-Series and S-Series 2-8388608 for E-Series

Sub-sampling

Sub-sampling is available only on platform: E T

The sFlow sample rate is not the frequency of sampling, but the number of packets that are skipped before the next sample is taken. Although a sampling rate can be configured for each port, TeraScale line cards can support only a single sampling rate per port-pipe.

Therefore, 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, sFlow Agent takes the lowest numerical value of the sampling rate of all the ports within the port-pipe, and configures all ports to this value. 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 allows the smallest sampling rate possible to be configured on the hardware, and also allows all other sampling rates to be available through sub-sampling.

For example, if Gig 1/0 and 1/1 are in a port-pipe, and they are configured with a sampling rate of 4096 on interface Gig 1/0, and 8192 on Gig 1/1, 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. Configure interface Gig 1/0 to a sub-sampling rate of 1 to achieve an actual rate of 4096.
3. Configure interface Gig 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 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 CPU condition is cleared. This is as per sFlow version 5 draft. Once the back-off changes the sample-rate, users 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. The actual sampling-rate of the interface and the configured sample-rate can be viewed 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.

Extended sFlow

Extended sFlow is supported fully on platform E

Platforms C and S support **extended-switch** information processing *only*.

Extended sFlow packs additional information in the sFlow datagram depending on the type of sampled packet. The following options can be enabled:

- **extended-switch** — 802.1Q VLAN ID and 802.1p priority information
- **extended-router** — Next-hop and source and destination mask length.
- **extended-gateway** — Source and destination AS number and the BGP next-hop.

Note: The entire AS path is not included. BGP community-list and local preference information are not included. These fields are assigned default values and are not interpreted by the collector.

Use the command **sflow** [**extended-switch**] [**extended-router**] [**extended-gateway**] **enable** command. By default packing of any of the extended information in the datagram is disabled.

Use the command **show sflow** to confirm that extended information packing is enabled, as shown in Figure 37-6.

Figure 37-6. Confirming that Extended sFlow is Enabled

```
Force10#show sflow
sFlow services are enabled
Global default sampling rate: 4096
Global default counter polling interval: 15
Global extended information enabled: gateway, router, 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

Linecard 1 Port set 0 H/W sampling rate 8192
  Gi 1/16: configured rate 8192, actual rate 8192, sub-sampling rate 1
  Gi 1/17: configured rate 16384, actual rate 16384, sub-sampling rate 2

Linecard 3 Port set 1 H/W sampling rate 16384
  Gi 3/40: configured rate 16384, actual rate 16384, sub-sampling rate 1
```

Extended sFlow settings show all 3 types are enabled

If none of the extended information is enabled, the **show** output is as shown in Figure 37-7.

Figure 37-7. Confirming that Extended sFlow is Disabled

```
Force10#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
```

Indicates no Extended sFlow types enabled.

Important Points to Remember

- The IP destination address has to be learned via BGP in order to export extended-gateway data, prior to FTOS version 7.8.1.0.
- If the IP destination address is not learned via BGP the Dell Force10 system does not export extended-gateway data, prior to FTOS version 7.8.1.0.
- FTOS 7.8.1.0 and later enhances the sFlow implementation for real time traffic analysis on the E-Series to provide extended gateway information in cases where the destination IP addresses are learned by different routing protocols, and for cases where the destination is reachable over ECMP.
- If the IP source address is learned via IGP then *srcAS* and *srcPeerAS* are zero.
- The srcAS and srcPeerAS might be zero even though the IP source address is learned via BGP. The Dell Force10 system packs the srcAS and srcPeerAS information only if the route is learned via BGP and it is reachable via the ingress interface of the packet.

The previous points are summarized in following table.

Table 37-1. Extended Gateway Summary

IP SA	IP DA	srcAS and srcPeerAS	dstAS and dstPeerAS	Description
static/connected/IGP	static/connected/IGP	—	—	Extended gateway data is not exported because there is no AS information.
static/connected/IGP	BGP	0	Exported	src_as & src_peer_as are zero because there is no AS information for IGP.
BGP	static/connected/IGP	— Exported	— Exported	Prior to FTOS version 7.8.1.0, extended gateway data is not be exported because IP DA is not learned via BGP. 7.8.1.0 allows extended gateway information in cases where the source and destination IP addresses are learned by different routing protocols, and for cases where is source is reachable over ECMP.
BGP	BGP	Exported	Exported	Extended gateway data is packed.

38

Simple Network Management Protocol

Simple Network Management Protocol is supported on platforms C E S

Protocol Overview

Network management stations use Simple Network Management Protocol (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*.

Implementation Information

- FTOS supports SNMP version 1 as defined by RFC 1155, 1157, and 1212, SNMP version 2c as defined by RFC 1901, and SNMP version 3 as defined by RFC 2571.
- FTOS supports up to 15 trap receivers.
- The FTOS implementation of the sFlow MIB supports sFlow configuration via SNMP sets.
- SNMP traps for STP and MSTP state changes are based on BRIDGE MIB (RFC 1483) for STP and IEEE 802.1 *draft ruzin-mstp-mib-02* for MSTP.

Configure 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 your Dell Force10 system using SNMP. Also, these configurations use SNMP version 2c.

Configuring SNMP requires only a single step:

1. Create a community. See page 680.

Related Configuration Tasks

The following list contains configuration tasks for SNMP:

- Read Managed Object Values on page 681
- Write Managed Object Values on page 682
- Subscribe to Managed Object Value Updates using SNMP on page 683
- Copy Configuration Files on page 113
- Manage VLANs using SNMP on page 691
- Enable and Disable a Port using SNMP on page 696
- Fetch Dynamic MAC Entries using SNMP on page 696
- Deriving Interface Indices on page 697

Important Points to Remember

- Typically, 5-second timeout and 3-second retry values on an SNMP server are sufficient for both LAN and WAN applications. If you experience a timeout with these values, increase the timeout value to greater than 3 seconds, and increase the retry value to greater than 2 on your SNMP server.
- Group ACLs override user ACLs in SNMPv3 configurations when both are configured and the user is part of the group.

Create a Community

The management station generates requests to either retrieve or alter the value of a management object and is called the *SNMP manage*r. 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.

FTOS enables SNMP automatically when you create an SNMP community and displays Message 1. You must specify whether members of the community may only retrieve values (read), or retrieve and alter values (read-write).

To create an SNMP community:

Task	Command	Command Mode
Choose a name for the community.	**snmp-server community** *name* {**ro** | **rw**}	CONFIGURATION

Message 1 SNMP Enabled

```
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
```

View your SNMP configuration, using the command **show running-config snmp** from EXEC Privilege mode, as shown in Figure 38-1.

Figure 38-1. Creating an SNMP Community

```
Force10#snmp-server community my-snmp-community ro
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
Force10#do show running-config snmp
!
snmp-server community mycommunity ro
Force10#
```

Read 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.

There are several Unix SNMP commands that read data:

Task	Command
Read the value of a single managed object, as shown in Figure 38-2.	**snmpget -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}
Read the value of the managed object directly below the specified object, as shown in Figure 38-3.	**snmpgetnext -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}
Read the value of many objects at once, as shown in Figure 38-4.	**snmpwalk -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}

Figure 38-2. 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
```

Figure 38-3. Reading the Value of the Next Managed Object in the MIB

```
> 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: S50V_7.7
```

Task	Command

Figure 38-4. Reading the Value of Many Managed Objects at Once

```
> snmpwalk -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1
SNMPv2-MIB::sysDescr.0 = STRING: Force10 Networks Real Time Operating System Software
Force10 Operating System Version: 1.0
Force10 Application Software Version: E_MAIN4.7.6.350
Copyright (c) 1999-2007 by Force10 Networks, Inc.
Build Time: Mon May 12 14:02:22 PDT 2008
SNMPv2-MIB::sysObjectID.0 = OID: SNMPv2-SMI::enterprises.6027.1.3.1
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32920954) 3 days, 19:26:49.54
SNMPv2-MIB::sysContact.0 = STRING:
SNMPv2-MIB::sysName.0 = STRING: S50V_7.7
SNMPv2-MIB::sysLocation.0 = STRING:
SNMPv2-MIB::sysServices.0 = INTEGER: 4
```

Write Managed Object Values

You may only alter (write) a managed object value if your management station is a member of the same community as the SNMP agent, and the object is writable.

To write or write-over the value of a managed object:

Task	Command
To write or write-over the value of a managed object, as shown in Figure 38-5.	**snmpset -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}

Figure 38-5. Writing over the Current Value of a Managed Object

```
> snmpset -v 2c -c mycommunity 10.11.131.161 sysName.0 s "R5"
SNMPv2-MIB::sysName.0 = STRING: R5
```

Configure Contact and Location Information using SNMP

You may configure system contact and location information from the Dell Force10 system or from the management station using SNMP.

To configure system contact and location information from the Dell Force10 system:

Task	Command	Command Mode
Identify the system manager along with this person's contact information (e.g E-mail address or phone number). You may use up to 55 characters. **Default**: None	**snmp-server contact** *text*	CONFIGURATION

Task	Command	Command Mode
Identify the physical location of the system. For example, San Jose, 350 Holger Way, 1st floor lab, rack A1-1. You may use up to 55 characters. **Default**: None	**snmp-server location** *text*	CONFIGURATION

To configure the system from the manumitting station using SNMP:

Task	Command	Command Mode
Identify the system manager along with this person's contact information (e.g E-mail address or phone number). You may use up to 55 characters. **Default**: None	**snmpset -v** *version* **-c** *community agent-ip sysContact.0* **s "***contact-info***"**	CONFIGURATION
Identify the physical location of the system. For example, San Jose, 350 Holger Way, 1st floor lab, rack A1-1. You may use up to 55 characters. **Default**: None	**snmpset -v** *version* **-c** *community agent-ip sysLocation.0* **s "***location-info***"**	CONFIGURATION

Subscribe to Managed Object Value Updates using SNMP

By default, the Dell Force10 system displays some unsolicited SNMP messages (traps) upon certain events and conditions. You can also configure the system to send the traps to a management station. Traps cannot be saved on the system.

FTOS supports the following three sets of traps:

- **RFC 1157-defined traps**: coldStart, warmStart, linkDown, linkUp, authenticationFailure, egpNeighbborLoss
- **Dell Force10 enterpriseSpecific environment traps**: fan, supply, temperature
- **Dell Force10 enterpriseSpecific protocol traps**: bgp, ecfm, stp, xstp,

To configure the system to send SNMP notifications:

Step	Task	Command	Command Mode
1	Configure the Dell Force10 system send notifications to an SNMP server.	**snmp-server host** *ip-address*	CONFIGURATION

Step	Task	Command	Command Mode
2	Specify which traps the Dell Force10 system sends to the trap receiver. • Enable all Dell Force10 enterpriseSpecific and RFC-defined traps using the command **snmp-server enable traps** from CONFIGURATION mode. • Enable all of the RFC-defined traps using the command **snmp-server enable traps snmp** from CONFIGURATION mode.	**snmp-server enable traps**	CONFIGURATION
3	Specify the interfaces out of which FTOS sends SNMP traps.	**snmp-server trap-source**	CONFIGURATION

Table 38-1 lists the traps the RFC-defined SNMP traps and the command used to enable each. Note that the coldStart and warmStart traps are enabled using a single command.

Table 38-1. RFC 1157 Defined SNMP Traps on FTOS

Command Option	Trap
snmp authentication	`SNMP_AUTH_FAIL:SNMP Authentication failed.Request with invalid community string.`

snmp coldstart	`SNMP_COLD_START: Agent Initialized - SNMP COLD_START.`
	`SNMP_WARM_START: Agent Initialized - SNMP WARM_START.`

snmp linkdown	`PORT_LINKDN:changed interface state to down:%d`
snmp linkup	`PORT_LINKUP:changed interface state to up:%d`

Enable a subset of Dell Force10 enterpriseSpecific SNMP traps using one of the listed command options Table 38-2 with the command **snmp-server enable traps**. Note that the **envmon** option enables all environment traps including those that are enabled with the **envmon supply**, **envmon temperature**, and **envmon fan** options.

Table 38-2. Dell Force10 Enterprise-specific SNMP Traps

Command Option	Trap

envmon	`CARD_SHUTDOWN: %sLine card %d down - %s`
	`CARD_DOWN: %sLine card %d down - %s`

	`LINECARDUP: %sLine card %d is up`
	`CARD_MISMATCH: Mismatch: line card %d is type %s - type %s required.`

	`RPM_STATE: RPM1 is in Active State`
	`RPM_STATE: RPM0 is in Standby State`

	`RPM_DOWN: RPM 0 down - hard reset`
	`RPM_DOWN: RPM 0 down - card removed`

	`HOT_FAILOVER: RPM Failover Completed`
	`SFM_DISCOVERY: Found SFM 1`
	`SFM_REMOVE: Removed SFM 1`
	`MAJOR_SFM: Major alarm: Switch fabric down`
	`MAJOR_SFM_CLR: Major alarm cleared: Switch fabric up`

Table 38-2. Dell Force10 Enterprise-specific SNMP Traps

Command Option	Trap
	MINOR_SFM: MInor alarm: No working standby SFM
	MINOR_SFM_CLR: Minor alarm cleared: Working standby SFM present
	TASK SUSPENDED: SUSPENDED - svce:%d - inst:%d - task:%s
	RPM0-P:CP %CHMGR-2-CARD_PARITY_ERR
	ABNORMAL_TASK_TERMINATION: CRASH - task:%s %s
	CPU_THRESHOLD: Cpu %s usage above threshold. Cpu5SecUsage (%d)
	CPU_THRESHOLD_CLR: Cpu %s usage drops below threshold. Cpu5SecUsage (%d)
	MEM_THRESHOLD: Memory %s usage above threshold. MemUsage (%d)
	MEM_THRESHOLD_CLR: Memory %s usage drops below threshold. MemUsage (%d)
	DETECT_STN_MOVE: Station Move threshold exceeded for Mac %s in vlan %d
envmon supply	PEM_PRBLM: Major alarm: problem with power entry module %s
	PEM_OK: Major alarm cleared: power entry module %s is good
	MAJOR_PS: Major alarm: insufficient power %s
	MAJOR_PS_CLR: major alarm cleared: sufficient power
	MINOR_PS: Minor alarm: power supply non-redundant
	MINOR_PS_CLR: Minor alarm cleared: power supply redundant
envmon temperature	MINOR_TEMP: Minor alarm: chassis temperature
	MINOR_TEMP_CLR: Minor alarm cleared: chassis temperature normal (%s %d temperature is within threshold of %dC)
	MAJOR_TEMP: Major alarm: chassis temperature high (%s temperature reaches or exceeds threshold of %dC)
	MAJOR_TEMP_CLR: Major alarm cleared: chassis temperature lower (%s %d temperature is within threshold of %dC)
envmon fan	FAN_TRAY_BAD: Major alarm: fantray %d is missing or down
	FAN_TRAY_OK: Major alarm cleared: fan tray %d present
	FAN_BAD: Minor alarm: some fans in fan tray %d are down
	FAN_OK: Minor alarm cleared: all fans in fan tray %d are good
xstp	%SPANMGR-5-STP_NEW_ROOT: New Spanning Tree Root, Bridge ID Priority 32768, Address 0001.e801.fc35.
	%SPANMGR-5-STP_TOPOLOGY_CHANGE: Bridge port GigabitEthernet 11/38 transitioned from Forwarding to Blocking state.
	%SPANMGR-5-MSTP_NEW_ROOT_BRIDGE: Elected root bridge for instance 0.
	%SPANMGR-5-MSTP_NEW_ROOT_PORT: MSTP root changed to port Gi 11/38 for instance 0. My Bridge ID: 40960:0001.e801.fc35 Old Root: 40960:0001.e801.fc35 New Root: 32768:00d0.038a.2c01.
	%SPANMGR-5-MSTP_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e801.fc35 Mstp Instance Id 0 port Gi 11/38 transitioned from forwarding to discarding state.
ecfm	%ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000
	%ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000
	%ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain provider at Level 4 VLAN 3000
	%ECFM-5-ECFM_REMOTE_ALARM: Remote CCM Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000
	%ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000

Table 38-2. Dell Force10 Enterprise-specific SNMP Traps

Command Option	Trap
<cr>	SNMP Copy Config Command Completed %RPM0-P:CP %SNMP-4-RMON_RISING_THRESHOLD: RMON rising threshold alarm from SNMP OID <oid> %RPM0-P:CP %SNMP-4-RMON_FALLING_THRESHOLD: RMON falling threshold alarm from SNMP OID <oid> %RPM0-P:CP %SNMP-4-RMON_HC_RISING_THRESHOLD: RMON high-capacity rising threshold alarm from SNMP OID <oid>

Copy Configuration Files Using SNMP

Use SNMP from a remote client to:

- copy the running-config file to the startup-config file, or
- copy configuration files from the Dell Force10 system to a server
- copy configuration files from a server to the Dell Force10 system

The relevant MIBs for these functions are:

Table 38-3. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Object Values	Description
copySrcFileType	.1.3.6.1.4.1.6027.3.5.1.1.1.2	1 = FTOS file 2 = running-config 3 = startup-config	Specifies the type of file to copy from. Valid values are: • If the copySrcFileType is running-config or startup-config, the default copySrcFileLocation is flash. • If the copySrcFileType is a binary file, the copySrcFileLocation and copySrcFileName must also be specified.
copySrcFileLocation	.1.3.6.1.4.1.6027.3.5.1.1.1.3	1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp	Specifies the location of source file. • If the copySrcFileLocation is FTP or SCP, copyServerAddress, copyUserName, and copyUserPassword must be specified.
copySrcFileName	.1.3.6.1.4.1.6027.3.5.1.1.1.4	Path (if file is not in current directory) and filename.	Specifies name of the file. • If copySourceFileType is set to running-config or startup-config, copySrcFileName is not required.
copyDestFileType	.1.3.6.1.4.1.6027.3.5.1.1.1.5	1 = FTOS file 2 = running-config 3 = startup-config	Specifies the type of file to copy to. • If the copySourceFileType is running-config or startup-config, the default copyDestFileLocation is flash. • If the copyDestFileType is a binary the copyDestFileLocation and copyDestFileName must be specified.

Table 38-3. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Object Values	Description
copyDestFileLocation	.1.3.6.1.4.1.6027.3.5.1.1.1.6	1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp	Specifies the location of destination file. • If the copyDestFileLocation is FTP or SCP, copyServerAddress, copyUserName, and copyUserPassword must be specified.
copyDestFileName	.1.3.6.1.4.1.6027.3.5.1.1.1.7	Path (if file is not in default directory) and filename.	Specifies the name of destination file.
copyServerAddress	.1.3.6.1.4.1.6027.3.5.1.1.1.8	IP Address of the server	The IP address of the server. • If the copyServerAddress is specified so must copyUserName, and copyUserPassword.
copyUserName	.1.3.6.1.4.1.6027.3.5.1.1.1.9	Username for the server.	Username for for the FTP, TFTP, or SCP server. • If the copyUserName is specified so must copyUserPassword.
copyUserPassword	.1.3.6.1.4.1.6027.3.5.1.1.1.10	Password for the server.	Password for the FTP, TFTP, or SCP server.

To copy a configuration file:

<table>
<tr><th>Step</th><th>Task</th><th>Command Syntax</th><th>Command Mode</th></tr>
<tr><td>1</td><td>Create an SNMP community string with read/ write privileges.</td><td><b>snmp-server community</b> <i>community-name</i> <b>rw</b></td><td>CONFIGURATION</td></tr>
<tr><td>2</td><td colspan="3">Copy the <i>f10-copy-config.mib</i> MIB from the Dell Force10 iSupport webpage to the server to which you are copying the configuration file.</td></tr>
<tr><td>3</td><td colspan="3">On the server, use the command <b>snmpset</b> as shown:
<b>snmpset -v</b> <i>snmp-version</i> <b>-c</b> <i>community-name</i> <b>-m</b> <i>mib_path</i>/<b>f10-copy-config.mib</b> <i>force10system-ip-address mib-object.index</i> {<b>i</b> | <b>a</b> | <b>s</b>} <i>object-value...</i>
• Every specified object must have an object value, which must be preceded by the keyword <b>i</b>. See Table 6 for valid values.
• <i>index</i> must be unique to all previously executed <b>snmpset</b> commands. If an index value has been used previously, a message like the one in Message 3 appears. In this case, increment the index value and enter the command again.
• Use as many MIB Objects in the command as required by the MIB Object descriptions in Table 6 to complete the command. See Table 7 or examples.
<b>Note:</b> You can use the entire OID rather than the object name. Use the form: <i>OID</i>.<i>index</i> <b>i</b> <i>object-value</i>, as shown in Figure 57.</td></tr>
</table>

Message 2 snmpset Index Value Error

```
Error in packet.
Reason: notWritable (that object does not support modification)
Failed object: FORCE10-COPY-CONFIG-MIB::copySrcFileType.101
```

Table 7 shows examples of using the command **snmpset** to copy a configuration. These examples assume that:

- the server OS is Unix
- you are using SNMP version 2c
- the community name is public, and
- the file *f10-copy-config.mib* is in the current directory or in the snmpset tool path.

Note: In Unix, enter the command **snmpset** for help using this command. Place the file *f10-copy-config.mib* the directory from which you are executing the **snmpset** command or in the snmpset tool path.

Table 38-4. Copying Configuration Files via SNMP

Task
Copy the running-config to the startup-config using the following command from the Unix machine: **snmpset -v 2c -c public -m ./f10-copy-config.mib** *force10system-ip-address* **copySrcFileType.***index* **i 2 copyDestFileType.***index* **i 3**
Figure 56 show the command syntax using MIB object names, and Figure 57 shows the same command using the object OIDs. In both cases, the object is followed by a unique index number.

Figure 38-6. Copying Configuration Files via SNMP using Object-Name Syntax

```
> snmpset -v 2c -r 0 -t 60 -c public -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.101 i
2 copyDestFileType.101 i 3
FORCE10-COPY-CONFIG-MIB::copySrcFileType.101 = INTEGER: runningConfig(2)
FORCE10-COPY-CONFIG-MIB::copyDestFileType.101 = INTEGER: startupConfig(3)
```

Figure 38-7. Copying Configuration Files via SNMP using OID Syntax

```
> snmpset -v 2c -c public -m ./f10-copy-config.mib 10.10.10.10
.1.3.6.1.4.1.6027.3.5.1.1.1.1.2.100 i 2 .1.3.6.1.4.1.6027.3.5.1.1.1.1.5.100 i 3
FORCE10-COPY-CONFIG-MIB::copySrcFileType.100 = INTEGER: runningConfig(2)
FORCE10-COPY-CONFIG-MIB::copyDestFileType.100 = INTEGER: startupConfig(3)
```

Table 38-4. Copying Configuration Files via SNMP

Task

Copy the startup-config to the running-config using the following command from a Unix machine:
snmpset -c private -v 2c *force10system-ip-address* **copySrcFileType.***index* **i 3 copyDestFileType.***index* **i 2**

Figure 38-8. Copying Configuration Files via SNMP using Object-Name Syntax

```
> snmpset -c public -v 2c -m ./f10-copy-config.mib 10.11.131.162 copySrcFileType.7 i 3
copyDestFileType.7 i 2
FORCE10-COPY-CONFIG-MIB::copySrcFileType.7 = INTEGER: runningConfig(3)
FORCE10-COPY-CONFIG-MIB::copyDestFileType.7 = INTEGER: startupConfig(2)
```

Figure 38-9. Copying Configuration Files via SNMP using OID Syntax

```
>snmpset -c public -v 2c 10.11.131.162 .1.3.6.1.4.1.6027.3.5.1.1.1.1.2.8 i 3
.1.3.6.1.4.1.6027.3.5.1.1.1.1.5.8 i 2
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.2.8 = INTEGER: 3
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.5.8 = INTEGER: 2
```

Copy the startup-config to the server via FTP using the following command from the Unix machine:

snmpset -v 2c -c public -m ./f10-copy-config.mib *force10system-ip-address* **copySrcFileType.***index* **i 2 copyDestFileName.***index* **s** *filepathIfilename* **copyDestFileLocation.***index* **i 4 copyServerAddress.***index* **a** *server-ip-address* **copyUserName.***index* **s** *server-login-id* **copyUserPassword.***index* **s** *server-login-password*

- *server-ip-address* must be preceded by the keyword **a**.
- values for copyUsername and copyUserPassword must be preceded by the keyword **s**.

Figure 38-10. Copying Configuration Files via SNMP and FTP to a Remote Server

```
> snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.110 i 2
copyDestFileName.110 s /home/startup-config copyDestFileLocation.110 i 4 copyServerAddress.110
a 11.11.11.11 copyUserName.110 s mylogin copyUserPassword.110 s mypass
FORCE10-COPY-CONFIG-MIB::copySrcFileType.110 = INTEGER: runningConfig(2)
FORCE10-COPY-CONFIG-MIB::copyDestFileName.110 = STRING: /home/startup-config
FORCE10-COPY-CONFIG-MIB::copyDestFileLocation.110 = INTEGER: ftp(4)
FORCE10-COPY-CONFIG-MIB::copyServerAddress.110 = IpAddress: 11.11.11.11
FORCE10-COPY-CONFIG-MIB::copyUserName.110 = STRING: mylogin
FORCE10-COPY-CONFIG-MIB::copyUserPassword.110 = STRING: mypass
```

Copy the startup-config to the server via TFTP using the following command from the Unix machine:

Note: Verify that the file exists and its permissions are set to 777, and specify the relative path to the TFTP root directory.

snmpset -v 2c -c public -m ./f10-copy-config.mib *force10system-ip-address* **copySrcFileType.***index* **i 3 copyDestFileType.***index* **i 1 copyDestFileName.***index* **s** *filepathIfilename* **copyDestFileLocation.***index* **i 3 copyServerAddress.***index* **a** *server-ip-address*

Table 38-4. Copying Configuration Files via SNMP

Task

Figure 38-11. Copying Configuration Files via SNMP and TFTP to a Remote Server

```
.snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10
copySrcFileType.4 i 3
copyDestFileType.4 i 1
copyDestFileLocation.4 i 3
copyDestFileName.4 s /home/myfilename
copyServerAddress.4 a 11.11.11.11
```

Copy a binary file from the server to the startup-configuration on the Dell Force10 system via FTP using the following command from the Unix server:

snmpset -v 2c -c public -m ./f10-copy-config.mib *force10system-ip-address* **copySrcFileType.***index* **i 1 copySrcFileLocation.***index* **i 4 copySrcFileName.***index* **s** *filepath/filename* **copyDestFileType.***index* **i 3 copyServerAddress.***index* **a** *server-ip-address* **copyUserName.***index* **s** *server-login-id* **copyUserPassword.***index* **s** *server-login-password*

Figure 38-12. Copying Configuration Files via SNMP and FTP from a Remote Server

```
> snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.10 i 1
copySrcFileLocation.10 i 4 copyDestFileType.10 i 3 copySrcFileName.10 s /home/myfilename
copyServerAddress.10 a 172.16.1.56 copyUserName.10 s mylogin copyUserPassword.10 s mypass
```

Dell Force10 provides additional MIB Objects to view copy statistics. These are provided in Table 8.

Table 38-5. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Values	Description
copyState	.1.3.6.1.4.1.6027.3.5.1.1.1.11	1= running 2 = successful 3 = failed	Specifies the state of the copy operation.
copyTimeStarted	.1.3.6.1.4.1.6027.3.5.1.1.1.12	Time value	Specifies the point in the up-time clock that the copy operation started.
copyTimeCompleted	.1.3.6.1.4.1.6027.3.5.1.1.1.13	Time value	Specifies the point in the up-time clock that the copy operation completed.
copyFailCause	.1.3.6.1.4.1.6027.3.5.1.1.1.14	1 = bad file name 2 = copy in progress 3 = disk full 4 = file exists 5 = file not found 6 = timeout 7 = unknown	Specifies the reason the copy request failed.
copyEntryRowStatus	.1.3.6.1.4.1.6027.3.5.1.1.1.15	Row status	Specifies the state of the copy operation. Uses CreateAndGo when you are performing the copy. The state is set to active when the copy is completed.

To obtain a value for any of the MIB Objects in Table 8:

Step	Task
1	Get a copy-config MIB object value. **snmpset -v 2c -c public -m ./f10-copy-config.mib** *force10system-ip-address* [*OID***.***index* | *mib-object***.***index* • *index* is the index value used in the **snmpset** command used to complete the copy operation.
	Note: You can use the entire OID rather than the object name. Use the form: *OID***.***index*, as shown in Figure 62.

Figure 61 and Figure 62 are examples of using the command **snmpget** to obtain a MIB object value. These examples assume that:

- the server OS is Unix
- you are using SNMP version 2c
- the community name is public, and
- the file *f10-copy-config.mib* is in the current directory.

Note: In Unix, enter the command **snmpset** for help using this command.

Figure 61 shows the command syntax using MIB object names, and Figure 62 shows the same command using the object OIDs. In both cases, the object is followed by same index number used in the **snmpset** command.

Figure 38-13. Obtaining MIB Object Values for a Copy Operation using Object-name Syntax

```
>snmpget-v2c-cprivate-m./f10-copy-config.mib10.11.131.140 copyTimeCompleted.110
FORCE10-COPY-CONFIG-MIB::copyTimeCompleted.110 = Timeticks: (1179831) 3:16:38.31
```

Figure 38-14. Obtaining MIB Object Values for a Copy Operation using OID Syntax

```
> snmpget -v 2c -c private 10.11.131.140 .1.3.6.1.4.1.6027.3.5.1.1.1.1.13.110
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.13.110 = Timeticks: (1179831) 3:16:38.31
```

Manage VLANs using SNMP

The qBridgeMIB managed objects in the Q-BRIDGE-MIB, defined in RFC 2674, enable you to use SNMP manage VLANs.

Create a VLAN

Use the dot1qVlanStaticRowStatus object to create a VLAN. The snmpset operation in Figure 38-15 creates VLAN 10 by specifying a value of 4 for instance 10 of the dot1qVlanStaticRowStatus object.

Figure 38-15. Creating a VLAN using SNMP

```
> snmpset -v2c -c mycommunity 123.45.6.78 .1.3.6.1.2.1.17.7.1.4.3.1.5.10 i 4
SNMPv2-SMI::mib-2.17.7.1.4.3.1.5.10 = INTEGER: 4
```

Assign a VLAN Alias

Write a character string to the dot1qVlanStaticName object to assign a name to a VLAN, as shown in Figure 38-16.

Figure 38-16. Assign a VLAN Alias using SNMP

```
[Unix system output]

> snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.1.1107787786 s "My 
VLAN"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.1.1107787786 = STRING: "My VLAN"

[Force10 system output]

Force10#show int vlan 10
Vlan 10 is down, line protocol is down
Vlan alias name is: My VLAN
Address is 00:01:e8:cc:cc:ce, Current address is 00:01:e8:cc:cc:ce
Interface index is 1107787786
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 01:01:00
Queueing strategy: fifo
Time since last interface status change: 01:01:00
```

Display the Ports in a VLAN

FTOS identifies VLAN interfaces using an interface index number that is displayed in the output of the command **show interface vlan**, as shown in Figure 38-17.

Figure 38-17. Identifying the VLAN Interface Index Number

```
Force10(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
Last clearing of "show interface" counters 00:12:42
Queueing strategy: fifo
Time since last interface status change: 00:12:42
```

To display the ports in a VLAN, send an **snmpget** request for the object dot1qStaticEgressPorts using the interface index as the instance number, as shown for an S-Series in Figure 38-18.

Figure 38-18. Display 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 00 00 00
00 00 00 00 00 00 00 00 00
```

The table that the Dell Force10 system sends in response to the **snmpget** request is a table that contains hexadecimal (hex) pairs, each pair representing a group of eight ports.

- On the E-Series, 12 hex pairs represents a line card. Twelve pairs accommodates the greatest currently available line card port density, 96 ports.
- On the C-Series, 28 hex pairs represents a line card. Twenty-eight pairs accommodates the greatest currently available line card port density, 28 ports per port-pipe, and any remaining hex pairs are unused.
- On the S-Series, 7 hex pairs represents a stack unit. Seven pairs accommodates the greatest number of ports available on an S-Series, 56 ports. The last stack unit is assigned 8 pairs; the eighth pair is unused.

The first hex pair, 00 in Figure 38-18, represents ports 1-7 in Stack Unit 0. The next pair to the right represents ports 8-15. 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:

- On the E-Series and C-Series each position in the 8-character string is for one port, starting with Port 0 at the left end of the string, and ending with Port 7 at the right end. A 0 indicates that the port is not a member of the VLAN; a 1 indicates VLAN membership.
- On the S-Series, 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.

Figure 38-18 shows the output for an S-Series. All hex pairs are 00, indicating that no ports are assigned to VLAN 10. In Figure 38-19, Port 0/2 is added to VLAN 10 as untagged. And the first hex pair changes from 00 to 04.

Figure 38-19. Displaying Ports in a VLAN using SNMP

```
[Force10 system output]

R5(conf)#do show vlan id 10

Codes: * - Default VLAN, G - GVRP VLANs
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

    NUM    Status    Description                        Q Ports
    10     Inactive                                     U Gi 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
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 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 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 above, the left-most position in the string represents Port 1. The next position from the left represents Port 2 and has a value of 1, indicating that Port 0/2 is in VLAN 10. The remaining positions are 0, so those ports are not in the VLAN.

Note that 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.

Add Tagged and Untagged Ports to a VLAN

The value dot1qVlanStaticEgressPorts object is an array of all VLAN members.

The dot1qVlanStaticUntaggedPorts object is an array of only untagged VLAN members. All VLAN members that are not in dot1qVlanStaticUntaggedPorts are tagged.

- To add a tagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts object, as shown in Figure 38-20.
- To add an untagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts objects, as shown in Figure 38-21.

Note: Whether adding a tagged or untagged port, you must specify values for both dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts.

In Figure 38-20, Port 0/2 is added as an untagged member of VLAN 10.

Figure 38-20. Adding Untagged Ports to a VLAN using SNMP

```
>snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786 x "40 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00"
.1.3.6.1.2.1.17.7.1.4.3.1.4.1107787786 x "40 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
SNMPv2-SMI::mib-2.17.7.1.4.3.1.4.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
```

In Figure 38-21, Port 0/2 is added as a tagged member of VLAN 10.

Figure 38-21. Adding Tagged Ports to a VLAN using SNMP

```
>snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786 x "40 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00"
.1.3.6.1.2.1.17.7.1.4.3.1.4.1107787786 x "00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
SNMPv2-SMI::mib-2.17.7.1.4.3.1.4.1107787786 = Hex-STRING: 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
```

Enable and Disable a Port using SNMP

Step	Task	Command Syntax	Command Mode
1	Create an SNMP community on the Dell Force10 system.	**snmp-server community**	CONFIGURATION
2	From the Dell Force10 system, identify the interface index of the port for which you want to change the admin status. Or, from the management system, use the **snmpwwalk** command to identify the interface index.	**show interface**	EXEC Privilege

3	Enter the command **snmpset** to change the admin status using either the object descriptor or the OID. Choose integer 1 to change the admin status to Up, or 2 to change the admin status to Down.
	snmpset with descriptor: snmpset -v *version* **-c** *community* **agent-ip ifAdminStatus.***ifindex* **i** {**1** | **2**}
	snmpset with OID: snmpset -v *version* **-c** *community* **agent-ip .1.3.6.1.2.1.2.2.1.7.***ifindex* **i** {**1** | **2**}		

Fetch Dynamic MAC Entries using SNMP

Table 38-6. MIB Objects for Fetching Dynamic MAC Entries in the Forwarding Database

MIB Object	OID	Description	MIB
dot1dTpFdbTable	.1.3.6.1.2.1.17.4.3	List the learned unicast MAC addresses on the default VLAN.	Q-BRIDGE MIB
dot1qTpFdbTable	.1.3.6.1.2.1.17.7.1.2.2	List the learned unicast MAC addresses on non-default VLANs.	
dot3aCurAggFdbTable	.1.3.6.1.4.1.6027.3.2.1.1.5	List the learned MAC addresses of aggregated links (LAG).	F10-LINK-AGGREGATION-MIB

In Figure 38-22, R1 has one dynamic MAC address, learned off of port GigabitEthernet 1/21, which a member of the default VLAN, VLAN 1. The SNMP walk returns the values for dot1dTpFdbAddress, dot1dTpFdbPort, and dot1dTpFdbStatus.

Each object is comprised 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 GigabitEthernet 1/21, the manager returns the integer 118. The maximum line card port density on the E-Series is 96 ports, and line card numbering begins with 0; GigabitEthernet 1/21 is the 21st port on Line Card 1, and 96 + 21 yields 118.

Figure 38-22. Fetching Dynamic MAC Addresses on the Default VLAN

```
-----------------------------MAC Addresses on Force10 System-------------------------------
R1_E600#show mac-address-table
VlanId     Mac Address           Type   Interface        State
 1      00:01:e8:06:95:ac       Dynamic Gi 1/21         Active
------------------------------Query from Management Station--------------------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.4.3.1
SNMPv2-SMI::mib-2.17.4.3.1.1.0.1.232.6.149.172 = Hex-STRING: 00 01 E8 06 95 AC
SNMPv2-SMI::mib-2.17.4.3.1.2.0.1.232.6.149.172 = INTEGER: 118
SNMPv2-SMI::mib-2.17.4.3.1.3.0.1.232.6.149.172 = INTEGER: 3
```

In Figure 38-23, GigabitEthernet 1/21 is moved to VLAN 1000, a non-default VLAN. Use the objects dot1qTpFdbTable to fetch the MAC addresses learned on non-default VLANs. The instance number is the VLAN number concatenated with the decimal conversion of the MAC address.

Figure 38-23. Fetching Dynamic MAC Addresses on Non-default VLANs

```
-----------------------------MAC Addresses on Force10 System-------------------------------
R1_E600#show mac-address-table
VlanId     Mac Address           Type   Interface        State
 1000   00:01:e8:06:95:ac       Dynamic Gi 1/21         Active
------------------------------Query from Management Station--------------------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.7.1.2.2.1
SNMPv2-SMI::mib-2.17.7.1.2.2.1.2.1000.0.1.232.6.149.172 = INTEGER: 118
SNMPv2-SMI::mib-2.17.7.1.2.2.1.3.1000.0.1.232.6.149.172 = INTEGER: 3
```

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.

Figure 38-24. Fetching Dynamic MAC Addresses on the Default VLANDell Force10

```
-----------------------------MAC Addresses on Force10 System-------------------------------
R1_E600(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

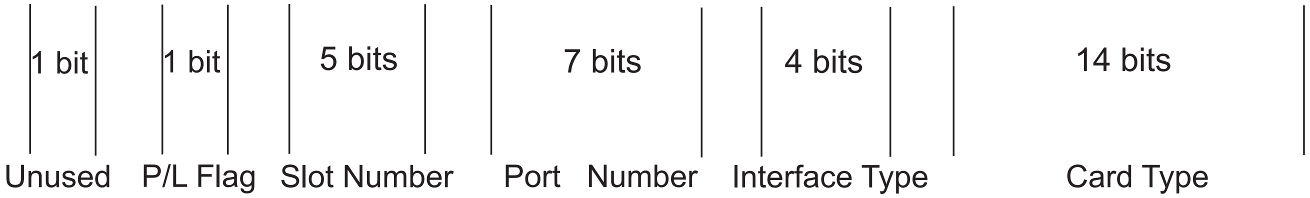
FTOS assigns an interface number to each (configured or unconfigured) physical and logical interface. Display the interface index number using the command **show interface** from EXEC Privilege mode, as shown in Figure 38-25.

Figure 38-25. Display the Interface Index Number

```
Force10#show interface gig 1/21
GigabitEthernet 1/21 is up, line protocol is up
Hardware is Force10Eth, address is 00:01:e8:0d:b7:4e
    Current address is 00:01:e8:0d:b7:4e
Interface index is 72925242
[output omitted]
Force10#show linecard all | grep 1
  1   online         online      E48TF    E48TF    7.7.1.1     48
```

The interface index is a binary number with bits that indicate the slot number, port number, interface type, and card type of the interface. FTOS converts this binary index number to decimal, and displays it in the output of the **show interface** command.

Figure 38-26. Interface Index Binary Calculations

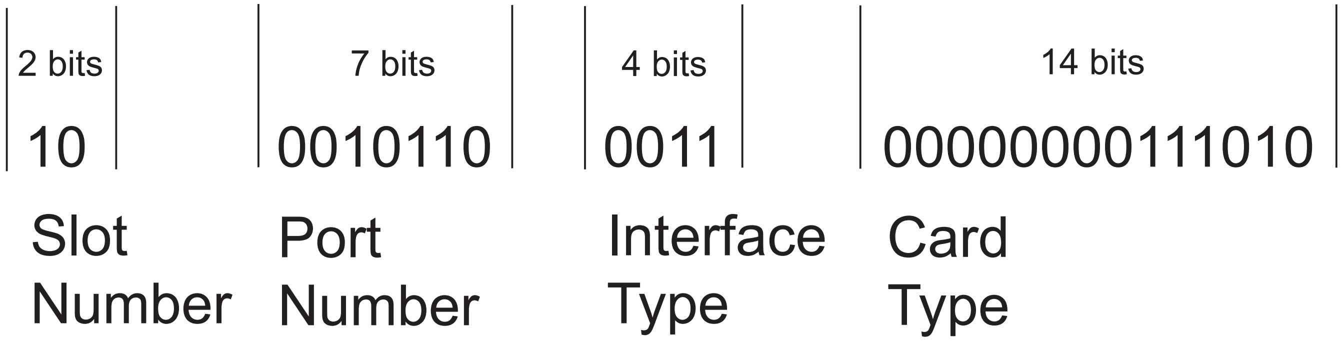


Starting from the least significant bit (LSB):

- 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 72925242 is 100010110001100000000111010 in binary. The binary interface index for GigabitEthernet 1/21 of a 48-port 10/100/1000Base-T line card with RJ-45 interface is shown in Figure 38-27. 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.

Figure 38-27. Binary Representation of Interface Index



For interface indexing, slot and port numbering begins with the binary one. If the Dell Force10 system begins slot and port numbering from 0, then the binary 1 represents slot and port 0. For example, the index number in Figure 38-27 gives the binary 2 for the slot number, though interface GigabitEthernet 1/21 belongs to Slot 1. This is because the port for this example is on an E-Series which begins numbering slots from 0. You must subtract 1 from the slot number 2, which yields 1, the correct slot number for interface 1/21.

Note that the interface index does not change if the interface reloads or fails over. On the S-Series, if the unit is renumbered (for any reason) the interface index will change during a reload.

39

Spanning Tree Protocol

Spanning Tree Protocol is supported on platforms: C E S

Protocol Overview

Spanning Tree Protocol (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 enables you to implement redundant paths, which can be activated upon 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.

FTOS supports three other variations of Spanning Tree, as shown here:

Table 39-1. FTOS Supported Spanning Tree Protocols

Dell Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Configuring Spanning Tree

Configuring Spanning Tree is a two-step process:

1. Configure interfaces for Layer 2. See page 49.
2. Enable Spanning Tree Protocol. See page 704.

Related Configuration Tasks

- Adding an Interface to the Spanning Tree Group on page 707
- Removing an Interface from the Spanning Tree Group on page 707
- Modifying Global Parameters on page 707
- Modifying Interface STP Parameters on page 708

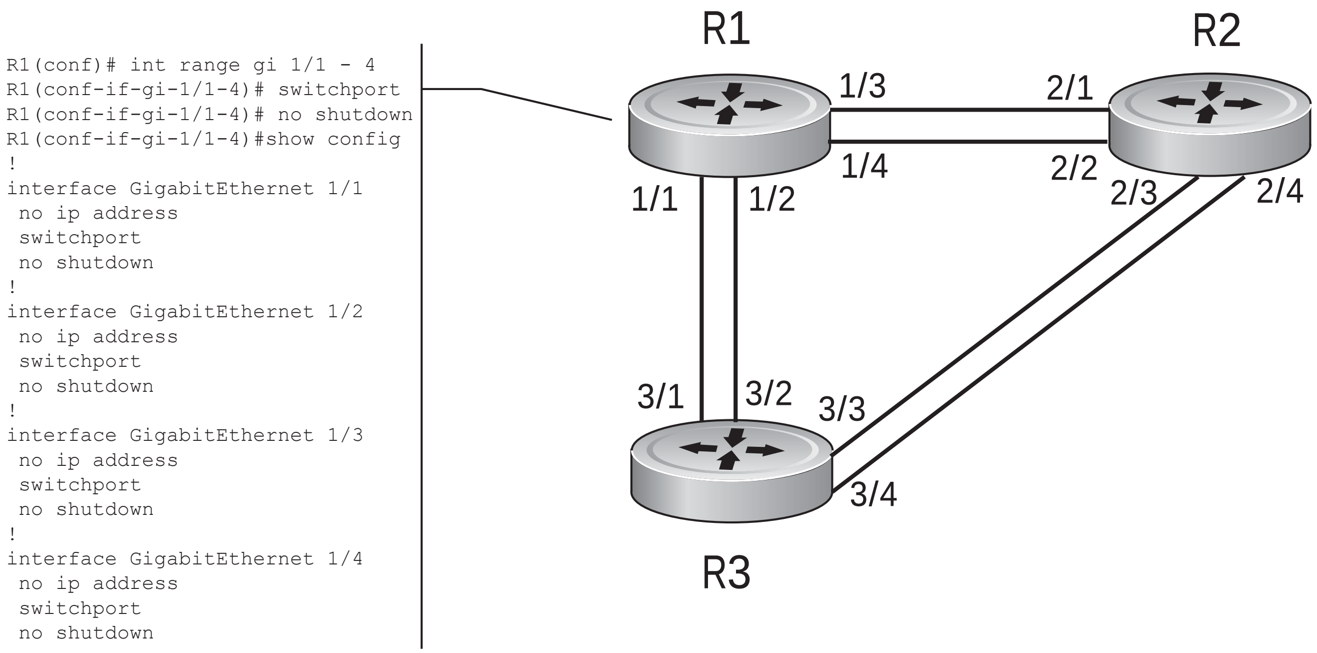
Important Points to Remember

- Spanning Tree Protocol (STP) is disabled by default.
- FTOS supports only one Spanning Tree instance (0). For multiple instances, you must enable MSTP, or PVST+. You may only enable one flavor of Spanning Tree at any one time.
- All ports in 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 STP is enabled, enable the port and configure it for Layer 2 using the command switchport.
- The IEEE Standard 802.1D allows eight bits for port ID and eight bits for priority. However, the eight bits for port ID provide port IDs for only 256 ports and the C-Series can contain 336 ports. To accommodate the increased number of ports, FTOS uses four bits from priority field in the port ID field.This implementation affects the Bridge MIB (RFC 1493), and you must interpret objects such as *dot1dStpPortDesignatedPort* object by using the first four bits as the priority and the last 12 bits as the port ID.

Configuring Interfaces for Layer 2 Mode

All interfaces on all switches that will participate in Spanning Tree must be in Layer 2 mode and enabled.

Figure 39-1. Example of Configuring Interfaces for Layer 2 Mode



```
R1(conf)# int range gi 1/1 - 4
R1(conf-if-gi-1/1-4)# switchport
R1(conf-if-gi-1/1-4)# no shutdown
R1(conf-if-gi-1/1-4)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/2
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/3
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/4
 no ip address
 switchport
 no shutdown
```

To configure the interfaces for Layer 2 and then enable them:

Step	Task	Command Syntax	Command Mode
1	If the interface has been assigned an IP address, remove it.	**no ip address**	INTERFACE
2	Place the interface in Layer 2 mode.	**switchport**	INTERFACE
3	Enable the interface.	**no shutdown**	INTERFACE

Verify that an interface is in Layer 2 mode and enabled using the **show config** command from INTERFACE mode.

Figure 39-2. Verifying Layer 2 Configuration

```
Force10(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
no shutdown
Force10(conf-if-gi-1/1)#
```

Indicates that the interface is in Layer 2 mode

Enabling Spanning Tree Protocol Globally

Spanning Tree Protocol must be enabled globally; it is not enabled by default.

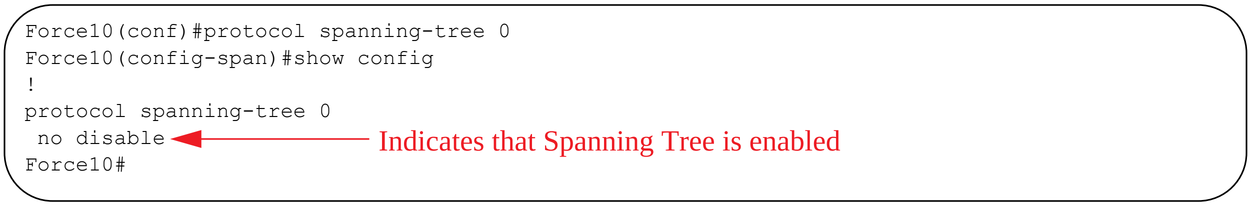
To enable Spanning Tree globally for all Layer 2 interfaces:

Step	Task	Command Syntax	Command Mode
1	Enter the PROTOCOL SPANNING TREE mode.	**protocol spanning-tree 0**	CONFIGURATION
2	Enable Spanning Tree.	**no disable**	PROTOCOL SPANNING TREE

Note: To disable STP globally for all Layer 2 interfaces, enter the **disable** command from PROTOCOL SPANNING TREE mode.

Verify that Spanning Tree is enabled using the **show config** command from PROTOCOL SPANNING TREE mode.

Figure 39-3. Verifying STP is Enabled



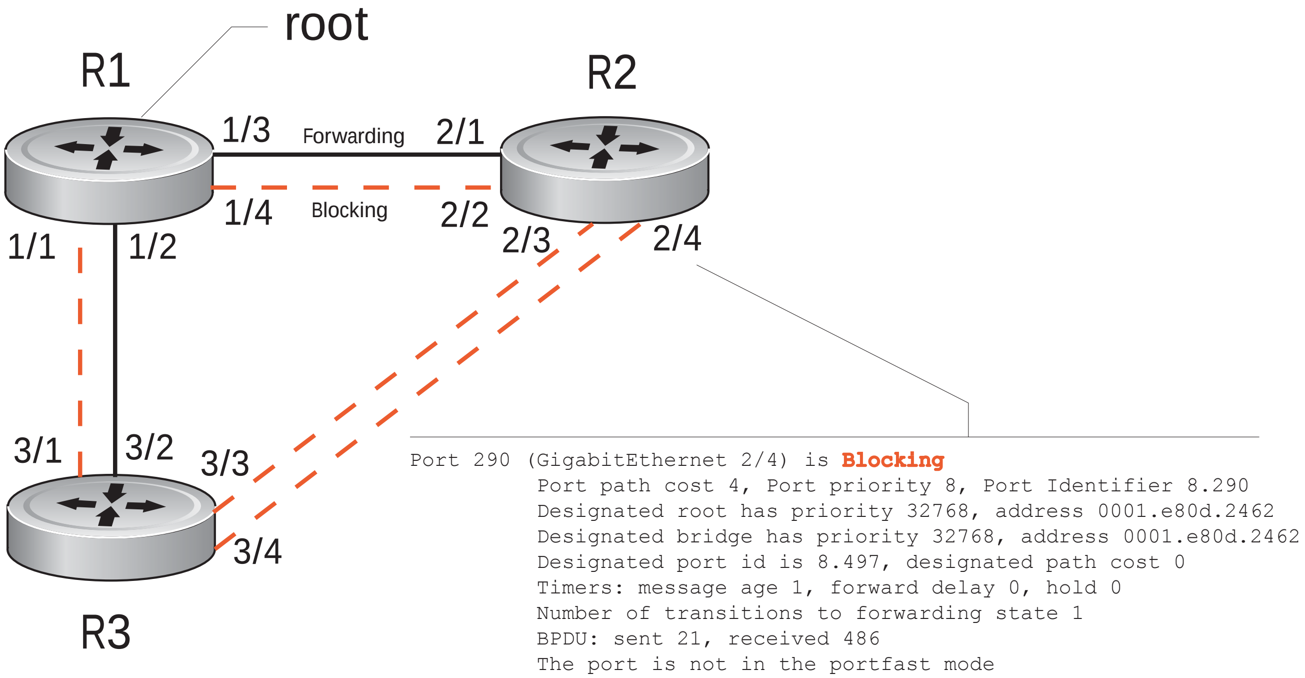
```
Force10(conf)#protocol spanning-tree 0
Force10(config-span)#show config
!
protocol spanning-tree 0
 no disable
Force10#
```

Indicates that Spanning Tree is enabled

When you enable Spanning Tree, 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 39-4. Spanning Tree Enabled Globally



```
Port 290 (GigabitEthernet 2/4) is Blocking
          Port path cost 4, Port priority 8, Port Identifier 8.290
          Designated root has priority 32768, address 0001.e80d.2462
          Designated bridge has priority 32768, address 0001.e80d.2462
          Designated port id is 8.497, 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
```

View the Spanning Tree configuration and the interfaces that are participating in STP using 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.

Figure 39-5. show spanning-tree 0 Command Example

```
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--
```

Confirm that a port is participating in Spanning Tree using the **show spanning-tree 0 brief** command from EXEC privilege mode.

Figure 39-6. show spanning-tree brief Command Example

```
Force10#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
-------------- ------ ---- ---- --- -----    -----------------     ------
Gi 1/1          8.496    8    4 DIS     0    32768 0001.e80d.2462  8.496
Gi 1/2          8.497    8    4 DIS     0    32768 0001.e80d.2462  8.497
Gi 1/3          8.513    8    4 FWD     0    32768 0001.e80d.2462  8.513
Gi 1/4          8.514    8    4 FWD     0    32768 0001.e80d.2462  8.514
Force10#
```

Adding an Interface to the Spanning Tree Group

To add a Layer 2 interface to the Spanning Tree topology:

Task	Command Syntax	Command Mode
Enable Spanning Tree on a Layer 2 interface.	**spanning-tree 0**	INTERFACE

Removing an Interface from the Spanning Tree Group

To remove a Layer 2 interface from the Spanning Tree topology:

Task	Command Syntax	Command Mode
Disable Spanning Tree on a Layer 2 interface.	**no spanning-tree 0**	INTERFACE

In FTOS versions prior to 7.6.1.0, the command **no spanning tree** disables Spanning Tree on the interface, however, BPDUs are still forwarded to the RPM, where they are dropped. Beginning in FTOS version 7.6.1.0, the command **no spanning tree** disables Spanning Tree on the interface, and incoming BPDUs are dropped at the line card instead of at the RPM, which frees processing resources. This behavior is called Layer 2 BPDU filtering and is available for STP, RSTP, PVST+, and MSTP.

Modifying Global Parameters

You can modify 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 Spanning Tree.

Note: Dell Force10 recommends that only experienced network administrators change the Spanning Tree parameters. Poorly planned modification of the Spanning Tree parameters can negatively impact network performance.

Table 39-2 displays the default values for Spanning Tree.

Table 39-2. STP Default Values

STP Parameter	Default Value
Forward Delay	15 seconds
Hello Time	2 seconds
Max Age	20 seconds

Table 39-2. STP Default Values

STP Parameter		Default Value
Port Cost	100-Mb/s Ethernet interfaces	19
	1-Gigabit Ethernet interfaces	4
	10-Gigabit Ethernet interfaces	2
	Port Channel with 100 Mb/s Ethernet interfaces	18
	Port Channel with 1-Gigabit Ethernet interfaces	3
	Port Channel with 10-Gigabit Ethernet interfaces	1
Port Priority		8

To change STP global parameters:

Task	Command Syntax	Command Mode
Change the forward-delay parameter (the wait time before the interface enters the *forwarding* state). • Range: 4 to 30 • Default: 15 seconds	**forward-delay** *seconds*	PROTOCOL SPANNING TREE
Change the hello-time parameter (the BPDU transmission interval). **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**hello-time** *seconds*	PROTOCOL SPANNING TREE
Change the max-age parameter (the refresh interval for configuration information that is generated by recomputing the Spanning Tree topology). Range: 6 to 40 Default: 20 seconds	**max-age** *seconds*	PROTOCOL SPANNING TREE

View the current values for global parameters using the **show spanning-tree 0** command from EXEC privilege mode. See Figure 39-5.

Modifying Interface STP Parameters

You can set the port cost and port priority values of interfaces in Layer 2 mode.

- **Port cost** is a value that is based on the interface type. The greater the port cost, the less likely the port will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

The default values are listed in Table 39-2.

To change the port cost or priority of an interface:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 65535 Default: see Table 39-2.	**spanning-tree 0 cost** *cost*	INTERFACE
Change the port priority of an interface. Range: 0 to 15 Default: 8	**spanning-tree 0 priority** *priority-value*	INTERFACE

View the current values for interface parameters using the **show spanning-tree 0** command from EXEC privilege mode. See Figure 39-5.

Enabling PortFast

The PortFast feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner. Interfaces forward frames by default until they receive a BPDU that indicates that they should behave otherwise; they do not go through the Learning and Listening states. The **bpduguard shutdown-on-violation** option causes the interface hardware to be shutdown 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 will drop packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation.

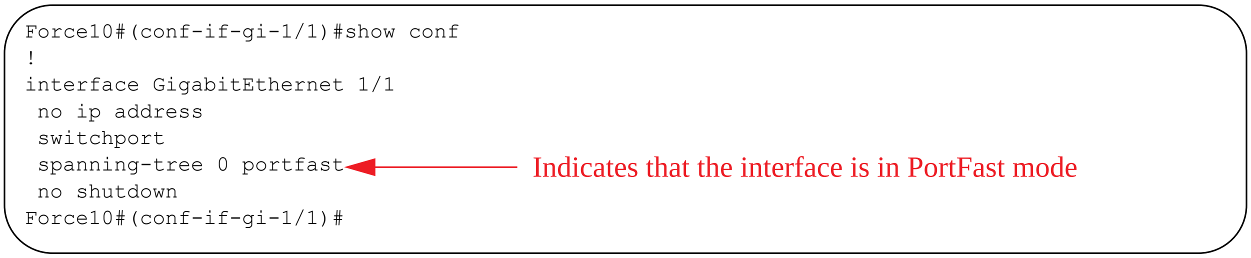
△ **Caution:** Enable PortFast only on links connecting to an end station. PortFast can cause loops if it is enabled on an interface connected to a network.

To enable PortFast on an interface:

Task	Command Syntax	Command Mode
Enable PortFast on an interface.	**spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]	INTERFACE

Verify that PortFast is enabled on a port using the **show spanning-tree** command from the EXEC privilege mode or the **show config** command from INTERFACE mode; Dell Force10 recommends using the **show config** command, as shown in Figure 39-7.

Figure 39-7. PortFast Enabled on Interface



```
Force10#(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 spanning-tree 0 portfast
 no shutdown
Force10#(conf-if-gi-1/1)#
```

Indicates that the interface is in PortFast mode

Preventing Network Disruptions with BPDU Guard

The Portfast (and Edgeport, in the case of RSTP, PVST+, and MSTP) feature should be configured 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 effect the STP topology. The BPDU Guard feature blocks an edgeport upon receiving a BPDU to prevent network disruptions, and FTOS displays Message 1. Enable BPDU Guard using the option **bpduguard** when enabling PortFast or EdgePort. The **bpduguard shutdown-on-violation** option causes the interface hardware to be shutdown 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.

Figure 39-8 shows a scenario in which an edgeport might unintentionally receive a BPDU. The port on the Dell Force10 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 BPDU Guard is enabled, when the edge port receives the BPDU, the BPDU will be dropped, the port will be blocked, and a console message will be generated.

Message 1 BPDU Guard Error

```
3w3d0h: %RPM0-P:RP2 %SPANMGR-5-BPDU_GUARD_RX_ERROR: Received Spanning Tree BPDU on BPDU guard
port. Disable GigabitEthernet 3/41.
```

Note: Note that *unless* the **shutdown-on-violation** option is enabled, spanning-tree only *drops packets* after a BPDU violation; the physical interface remains up, as shown below.

```
Force10(conf-if-gi-0/7)#do show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 32768, Address 0001.e805.fb07
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 32768, Address 0001.e85d.0e90
Configured hello time 2, max age 20, forward delay 15

Interface                                        Designated
 Name      PortID   Prio Cost    Sts Cost        Bridge ID        PortID
---------- -------- ---- ------- --- ------- -------------------- --------
Gi 0/6     128.263  128  20000   FWD 20000   32768 0001.e805.fb07 128.653
Gi 0/7     128.264  128  20000   EDS 20000   32768 0001.e85d.0e90 128.264

Interface
 Name      Role   PortID   Prio Cost    Sts Cost    Link-type Edge
---------- ------ -------- ---- ------- --- ------- --------- ----
Gi 0/6     Root   128.263  128  20000   FWD 20000   P2P       No
Gi 0/7     ErrDis 128.264  128  20000   EDS 20000   P2P       No
Force10(conf-if-gi-0/7)#do show ip int br gi 0/7
Interface                IP-Address      OK  Method Status                Protocol
GigabitEthernet 0/7      unassigned      YES Manual up                    up
```

FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

The error disabled state can be cleared with any of the following methods:

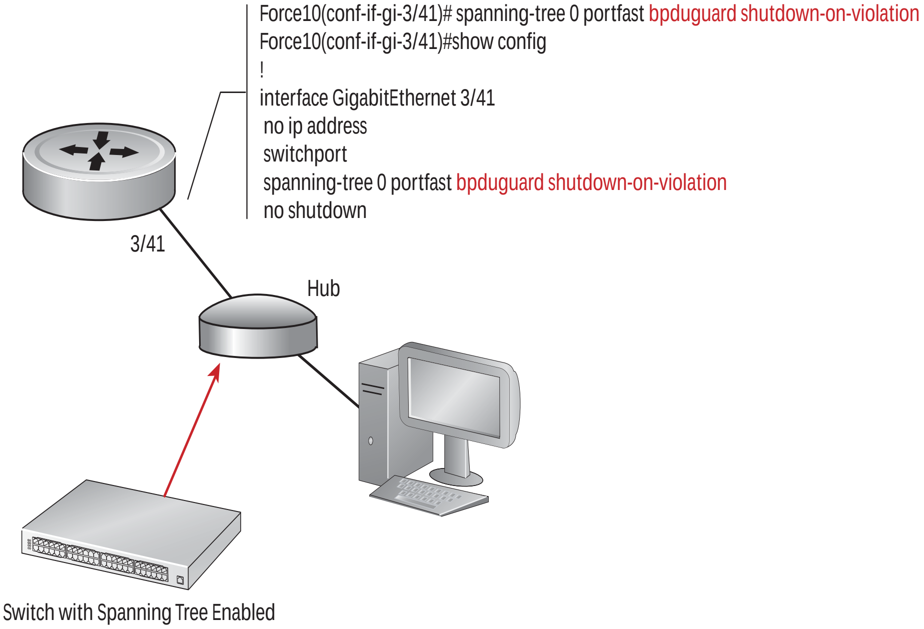
•Perform an **shutdown** command on the interface.

•Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).

•Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).

•Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Figure 39-8. Enabling BPDU Guard



```
Force10(conf-if-gi-3/41)#spanning-tree 0 portfast bpduguard shutdown-on-violation
Force10(conf-if-gi-3/41)#show config
!
interface GigabitEthernet 3/41
 no ip address
 switchport
 spanning-tree 0 portfast bpduguard shutdown-on-violation
 no shutdown
```

FTOS Behavior: BPDU Guard and BPDU filtering (see Removing an Interface from the Spanning Tree Group on page 707) both block BPDUs, but are two separate features.

BPDU Guard:

- is used on edgeports and blocks all traffic on edgeport if it receives a BPDU
- drops the BPDU after it reaches the RPM and generates a console message

BPDU Filtering:

- disables Spanning Tree on an interface
- drops all BPDUs at the line card without generating a console message

STP Root Selection

The Spanning Tree Protocol determines the root bridge, but you can assign one bridge a lower priority to increase the likelihood that it will be selected as 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:

Task	Command Syntax	Command Mode
Assign a number as the bridge priority or designate it as the root or secondary root. *priority-value* range: 0 to 65535. The lower the number assigned, the more likely this bridge will become the root bridge. The default is 32768. • The primary option specifies a bridge priority of 8192. • The secondary option specifies a bridge priority of 16384.	**bridge-priority** {*priority-value* | **primary** | **secondary**}	PROTOCOL SPANNING TREE

View only the root information using the **show spanning-tree root** command (see Figure 39-9) from EXEC privilege mode.

Figure 39-9. show spanning-tree root Command Example

```
Force10#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
Force10#
```

SNMP Traps for Root Elections and Topology Changes

- Enable SNMP traps for Spanning Tree state changes using the command **snmp-server enable traps stp**.
- Enable SNMP traps for RSTP, MSTP, and PVST+ collectively using the command **snmp-server enable traps xstp**.

Configuring Spanning Trees as Hitless

Configuring Spanning Trees as Hitless is supported only on platforms: C E

You can configure Spanning Tree (STP), Rapid Spanning Tree (RSTP), Multiple Spanning Tree (MSTP), and Per-Vlan Spanning Tree (PVST+) to be hitless (all or none must be configured as hitless). When configured as hitless, critical protocol state information is synchronized between RPMs so that RPM failover is seamless, and no topology change is triggered.

Configure all Spanning Tree types to be hitless using the command **redundancy protocol xstp** from CONFIGURATION mode, as shown in Figure 39-10.

Figure 39-10. Configuring all Spanning Tree Types to be Hitless

```
Force10(conf)#redundancy protocol xstp
Force10#show running-config redundancy
!
redundancy protocol xstp
Force10#
```

40

Stacking S-Series Switches

Stacking S-Series Switches is supported on platform S

Using the FTOS stacking feature, multiple S-Series switch units can be interconnected with stacking interfaces. The stack becomes manageable as a single switch through the stack management unit.

This chapter contains the following sections:

- S-Series Stacking Overview
- Important Points to Remember
- S-Series Stacking Configuration Tasks

S-Series Stacking Overview

The S55 supports up to 12 stack members with FTOS version 8.3.5.1.

A stack is analogous to an E-Series or C-Series system with redundant RPMs and multiple line cards. FTOS elects a primary and secondary management unit, and all other units are member units. The forwarding database resides on the primary, and all other units maintain a sychnronized local copy. Each unit in the stack makes forwarding decisions based on their local copy.

FTOS presents all of the units like line cards; for example, to access GigabitEthernet Port 1 on Stack Unit 0, enter **interface gigabitethernet 0/1** from CONFIGURATION mode.

High Availability on S-Series Stacks

S-Series stacks have primary and secondary management units analogous to Dell Force10 Route Processor Modules (Message 40-1). The management units synchronize the running configuration and protocol states so that the system fails over in the event of a hardware or software fault on the primary. In such an event, or when the primary is removed, the secondary unit becomes the stack manager and FTOS elects a new secondary. FTOS resets the failed management unit, and once online, it becomes a member unit; the remaining members remain online.

Figure 40-1. S55 Stack Manager Redundancy

```
Force10#show redundancy

--------- Stack-unit Status  ---------
 Mgmt ID:                          1
 Stack-unit ID:                    0
 Stack-unit Redundancy Role:       Primary
 Stack-unit State:                 Active
 Stack-unit SW Version:            SD8.3.5.1
 Link to Peer:                     Up

--------- PEER Stack-unit Status  --------------------
 Stack-unit State:                 Standby
 Peer stack-unit ID:               1
 Stack-unit SW Version:            SD8.3.5.1

--------- Stack-unit Redundancy Configuration  --------------
 Primary Stack-unit:               mgmt-id    1
 Auto Data Sync:                   Full
 Failover Type:                    Hot Failover
 Auto reboot Stack-unit:           Disabled
 Auto failover limit:              3 times in 60 minutes

--------- Stack-unit Failover Record  ---------------------
 Failover Count:                   3
 Last failover timestamp:          Oct 30 2010 04:15:36
 Last failover Reason:             User request
 Last failover type:               Hot Failover

--------- Last Data Block Sync Record:  ----------------------
 Stack Unit Config:         succeeded  Oct 30 2010 04:16:13
   Start-up Config:         succeeded  Oct 30 2010 04:16:13
 Runtime Event Log:         succeeded  Oct 30 2010 04:16:13
    Running Config:         succeeded  Oct 30 2010 04:16:13
           ACL Mgr:         succeeded  Oct 30 2010 04:16:13
              LACP:         no block sync done
               STP:         no block sync done
              SPAN:         no block sync done
Force10#
```

Management Unit Selection on S-Series Stacks

FTOS has a selection algorithm to decide which stack units will be the primary and secondary management units. During the bootup of a single unit or the stack, FTOS compares the priority values of all of the units in the stack and elects the unit with the numerically highest priority the primary, and the next highest priority the secondary.

For example, if you add a powered standalone unit to a stack, either the standalone unit or the stack reloads (excluding the new unit), depending on which has the higher priority, the new unit or the existing stack manager. If the new unit has the higher priority, it becomes the new stack manager after the stack reloads.

All switches have a default priority of 1; if a priority tie occurs, the system with the highest MAC address supersedes, as shown in Figure 40-2.

Figure 40-2. Electing the S55 Stack Manager

```
Force10# #show system  brief

Stack MAC : 00:01:e8:55:00:85

Reload Type : normal-reload

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
 0    Member       online         S55         S55         8.3.5.1      52
 1    Management   online         S55         S55         8.3.5.1      52
 2    Standby      online         S55         S55         8.3.5.1      52
 3    Member       not present
 4    Member       not present
 5    Member       not present
 6    Member       not present
 7    Member       not present
Force10#show system stack-unit 0 | grep priority
Master priority : 0
Force10#show system stack-unit 1 | grep priority
Master priority : 0
Force10#show system stack-unit 2 | grep priority
Master priority : 0
Force10#show system stack-unit 0| grep "Burned In MAC"
Burned In MAC   : 00:01:e8:d5:ef:81
Force10#show system stack-unit 0 | grep "Burned In MAC"
Burned In MAC   : 00:01:e8:d5:ef:81
Force10#show system stack-unit 1 | grep "Burned In MAC"
Burned In MAC   : 00:01:e8:d5:f9:6f
Force10#show system stack-unit 2 | grep "Burned In MAC"
```

MAC Addressing on S-Series Stacks

The S-Series has three MAC addressees: the chassis MAC, interface MAC, and null interface MAC. All interfaces in the stack use the interface MAC address of the management unit (stack manager), and the chassis MAC for the stack is the master’s chassis MAC. The stack continues to use the master’s chassis MAC address even after a failover. The MAC address is not refreshed until the stack is reloaded and a different unit becomes the stack manager.

Note: If the removed management unit is brought up as a standalone unit or as part of a different stack, there is a possibility of MAC address collisions.

In Figure 40-3 and Figure 40-4, a standalone is added to a stack. The standalone and the stack master have the same priority, but the standalone has a lower MAC address, so the standalone reboots. In Figure 40-4 and Figure 40-5, a standalone is added to a stack. The standalone has a higher priority than the stack, so the stack (excluding the new unit) reloads.

Figure 40-3. Adding a Standalone S55 with a Lower MAC Address to a Stack— Before

```
-------------------------------STANDALONE BEFORE CONNECTION----------------------------------
Standalone#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Management   online          S55         S55         8.3.5.1     52
  1   Member       not present
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
Standalone#show system | grep priority
Master priority : 0
------------------------------------STACK BEFORE CONNECTION----------------------------------
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Standby      online         S55         S55         8.3.5.1     52
  1   Management   online         S55         S55         8.3.5.1     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
Stack#show system stack-unit 0 | grep priority
Master priority : 0
Stack#show system stack-unit 1 | grep priority
Master priority : 0
```

Figure 40-4. Adding a Standalone S55 with a Lower MAC Address and Equal Priority to a Stack—After

```
-------------------------------STANDALONE AFTER CONNECTION----------------------------------
Standalone#%STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
00:20:20: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
00:20:22: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present

Going for reboot. Reason is Stack merge
[bootup messages omitted]
------------------------------------STACK AFTER CONNECTION----------------------------------
Stack# 3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
3w1d14h: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 2 (type S55, 52 ports)
3w1d14h: %S55:2 %CHMGR-0-PS_UP: Power supply 0 in unit 2 is up
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 2 is up

Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-----------------------------------------------------------------------------
  0   Standby      online         S55         S55         8.3.5.1     52
  1   Management   online         S55         S55         8.3.5.1     52
  2   Member       online         S55         S55         8.3.5.1     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
  8   Member       not present
  9   Member       not present
  10   Member       not present
  11   Member       not present
```

Figure 40-5. Adding a Standalone S55 with a Lower MAC Address but Higher Priority to a Stack—Before

```
-------------------------------STANDALONE BEFORE CONNECTION-----------------------------------
Standalone#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       not present    S55
  1   Member       not present    S55
  2   Management   online         S55        S55        8.3.5.1       52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
  8   Member       not present
  9   Member       not present
  10   Member       not present
  11   Member       not present

[output omitted]
Stack#show system | grep priority
Master priority : 1
------------------------------------STACK BEFORE CONNECTION-----------------------------------
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Standby      online         S55          S55        8.3.5.1     52
  1   Management   online         S55          S55        8.3.5.1     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
  8   Member       not present
  9   Member       not present
  10   Member       not present
  11   Member       not present
Stack#show system stack-unit 0 | grep priority
Master priority : 0
Stack#show system stack-unit 1 | grep priority
Master priority : 0
```

Figure 40-6. Adding a Standalone with a Lower MAC Address but Higher Priority to a Stack—After

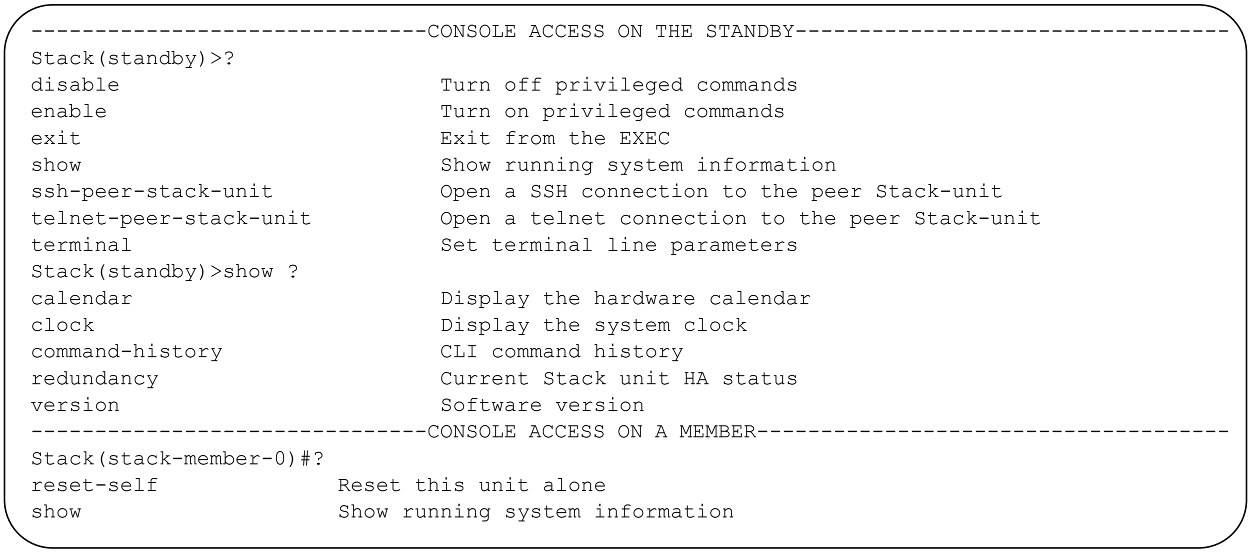
```
-------------------------------STANDALONE AFTER CONNECTION----------------------------------
Standalone#00:18:27: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
00:18:27: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
00:18:40: %STKUNIT2-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
00:18:40: %STKUNIT2-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 1 down - card removed
00:19:30: %STKUNIT2-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit i
s present
00:19:30: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
00:19:30: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
[remaining bootup messages omitted]
------------------------------------STACK AFTER CONNECTION----------------------------------
Stack#3w1d15h: %STKUNIT1-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is not
present
3w1d15h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present

Going for reboot. Reason is Stack merge
3w1d15h: %STKUNIT1-M:CP %CHMGR-2-STACK_UNIT_DOWN: Stack-unit 0 down - card removed
[bootup messages omitted]
Stack#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Standby      online         S55          S55        8.3.5.1     52
  1   Management   online         S55          S55        8.3.5.1     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
  8   Member       not present
  9   Member       not present
  10   Member       not present
  11   Member       not present
```

Management Access on S-Series Stacks

You can access the stack via the console port or VTY line.

- **Console access**: You may access the stack through the console port of the stack manager only. Like a standby RPM, the console port of the standby unit does not provide management capability; only a limited number of commands are available. Member units provide a severely limited set of commands, as shown in Figure 40-7.
- **Remote access**: You may access the stack with SNMP, SSH, or Telnet through any enabled, Layer 3 interface on any stack unit. The S60 and S55 have a dedicated Management port and support the routing table, similarly to the E-Series systems. No other S-Series has a dedicated management port or management route table on.

Figure 40-7. Accessing Non-Master Units on a Stack via the Console Port



```
-------------------------------CONSOLE ACCESS ON THE STANDBY----------------------------------
Stack(standby)>?
disable                         Turn off privileged commands
enable                          Turn on privileged commands
exit                            Exit from the EXEC
show                            Show running system information
ssh-peer-stack-unit             Open a SSH connection to the peer Stack-unit
telnet-peer-stack-unit          Open a telnet connection to the peer Stack-unit
terminal                        Set terminal line parameters
Stack(standby)>show ?
calendar                        Display the hardware calendar
clock                           Display the system clock
command-history                 CLI command history
redundancy                      Current Stack unit HA status
version                         Software version
-------------------------------CONSOLE ACCESS ON A MEMBER-------------------------------------
Stack(stack-member-0)#?
reset-self              Reset this unit alone
show                    Show running system information
```

Important Points to Remember

- You may stack up to eight S55s systems.
- You may not stack a combination of S-Series models. S55 systems can only stack with other S55 systems.
- All stack units must have the same version of FTOS.
- Insert S55 stacking modules in the lower optional module slot (slot 0) only. The upper optional module slot does not support stacking modules.

S-Series Stacking Installation Tasks

- Create an S-Series Stack
- Add a Unit to an S-Series Stack
- Remove a Unit from an S-Series Stack
- Merge Two S-Series Stacks
- Split an S-Series Stack

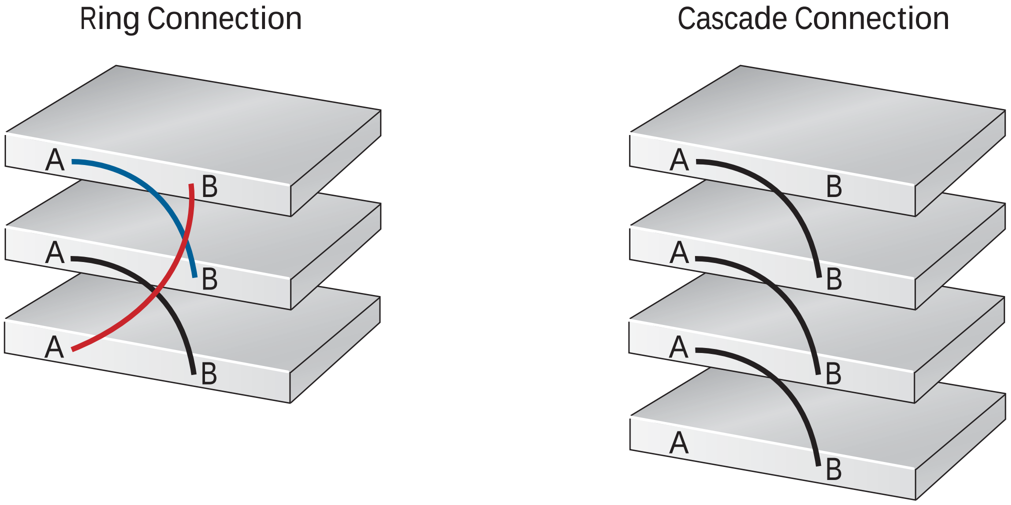
Create an S-Series Stack

Stacking modules are pluggable units in the back of the unit that switch traffic between units in a stack. Units are connected using bi-directional stacking cables; if you stacking modules have two ports, it does not matter if you connect port A to B, or A to A, or B to B. Install stacking modules before powering the unit. If you install a stacking module while the unit is online, FTOS does not register the new hardware; in this case, you must reload the unit.

The S50 and S25 systems support mixed stacking, as long as the units have the same FTOS version. The S55 and S60 systems DO NOT SUPPORT mixed stacking. Stack only S55 systems together or only S60 systems together; do not stack them with any other system type. Figure 40-8 shows two common stacking topologies, ring and cascade (also called daisy-chain). A ring topology provides some performance gains and stack integrity.

Note: The illustration below is an example to show ring and cascade topologies. Please refer to the Installation document for your system type to view a diagram for the topology supported by your system.

Figure 40-8. Common S-Series Stacking Topologies (S50-type)



Facing the rear of an S-Series unit, stack-port are numbered from left to right, beginning with the highest Ethernet port number (*n*) plus 1. For example, for a 48-port unit with two 12-Gigabyte stacking modules, the stack-ports are 49, 50, 51, and 52.

Note: The S55 stacking modules are installed in the lower optional module slot (slot 0) and use ports 48 and 49.

To add a unit to an existing stack:

Step	Task	Command Syntax	Command Mode
1	Verify that each unit has the same FTOS version prior to stacking them together.	**show version**	EXEC Privilege
2	Pre-configure unit numbers for each unit so that the stacking is deterministic upon boot up.	**stack-unit renumber**	EXEC Privilege
3	Configure the switch priority for each unit to make management unit selection deterministic.	**stack-unit priority**	CONFIGURATION
4	Connect the units using stacking cables.		

Step	Task	Command Syntax	Command Mode
5	Power the stack one unit at a time. Start with the management unit, then the standby, followed by each of the members in order of their assigned stack number (or the position in the stack you want each unit to take). Allow each unit to completely boot, and verify that the unit is detected by the stack manager, and then power the next unit.	**show system brief**	EXEC Privilege

To display the status of the stacking ports, including the topology:

Task	Command Syntax	Command Mode
Display the stacking ports.	**show system stack-ports**	EXEC Privilege

Note: The output samples below are example to show ring and cascade topologies. Please refer to the Installation document for your system type to view a diagram for the topology supported by your system.

Figure 40-9 shows an example of a daisy-chain topology. Figure 40-10 shows the same stack converted to a ring by connecting stack-port 2/51 to 0/51; you may rearrange the stacking cables without triggering a unit reset, so long as the stack manager is never disconnected from the stack.

Note: The S55 stacking modules are installed in the lower optional module slot (slot 0) and use ports 48 and 49 only.

Figure 40-9. Displaying the S-Series Stacking Topology

```
Stack#show system stack-ports
Topology: Daisy chain
Interface  Connection    Link Speed      Admin     Link     Trunk
                           (Gb/s)        Status    Status   Group
------------------------------------------------------------------
 0/51                      12             up        down
 0/52        1/50          12             up        up
 1/49        2/52          12             up        up
 1/50        0/52          12             up        up
 2/51                      12             up        down
 2/52        1/49          12             up        up
```

Figure 40-10. Displaying the S-Series Stacking Topology

```
Force10#show system stack-ports
Topology: Ring
Interface  Connection    Link Speed      Admin     Link     Trunk
                           (Gb/s)        Status    Status   Group
------------------------------------------------------------------
 0/51        2/51          12             up        up
 0/52        1/50          12             up        up
 1/49        2/52          12             up        up
 1/50        0/52          12             up        up
 2/51        0/51          12             up        up
 2/52        1/49          12             up        up
```

LED Status Indicators on an S-Series Stack

The stack unit is displayed in an LED panel on the front of each switch. Refer to the installation guide for your system type for a full discussion of all system display.

Add a Unit to an S-Series Stack

If you are adding units to a stack, you can either:

- allow FTOS to automatically assign the new unit a position in the stack, or
- manually determine each units position in the stack by configuring each unit to correspond with the stack before connecting it

Three configurable system variables affect how a new unit joins a stack: priority, stack number, and provision.

- Depending on which has the higher priority, either the standalone unit or the entire stack reloads (excluding the new unit). If the new unit has the higher priority, it becomes the new stack manager and the stack reloads, as shown in Figure 40-3, Figure 40-4, Figure 40-5, and Figure 40-6.
- If you add a unit that has a stack number that conflicts with the stack, the stack assigns the first available stack number, as shown in Figure 40-11 and Figure 40-12.
- If the stack has a provision for the stack-number that will be assigned to the new unit, the provision must match the unit type, or FTOS generates a type mismatch error, as show in Figure 40-13 and Figure 40-14.

After the new unit loads, it synchronizes its running and startup configurations with the stack.

To manually assign a new unit a position in the stack:

Step	Task	Command Syntax	Command Mode
1	While the unit is unpowered, install stacking modules in the new unit.		
2	On the stack, determine the next available stack-unit number, and the management prioritity of the management unit.	**show system brief** show system stack-unit	EXEC Privilege
3	Create a virtual unit and assign it the next available stack-unit number.	**stack-unit provision**	CONFIGURATION
4	On the new unit, number it the next available stack-unit number.	stack-unit renumber	EXEC Privilege
5	(OPTIONAL) On the new unit, assign a management priority based on whether you want the new unit to be the stack manager.	stack-unit priority	CONFIGURATION
6	Connect the new unit to the stack using stacking cables.		

Figure 40-11. Adding a Stack Unit with a Conflicting Stack Number—Before (S50 type)

```
------------------------STANDALONE BEFORE CONNECTION----------------------------------
Standalone#show system brief
Stack MAC : 00:01:e8:d5:ef:81
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       not present    S50V
  1   Management   online         S50V        S50V        7.8.1.0     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
-----------------------------STACK BEFORE CONNECTION----------------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       not present
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Figure 40-12. Adding a Stack Unit with a Conflicting Stack Number—After (S50 type)

```
------------------------STANDALONE AFTER CONNECTION----------------------------------
00:08:45: %STKUNIT1-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
00:08:45: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
00:08:47: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
Going for reboot. Reason is Stack merge
[bootup messages omitted]
Stack(stack-member-0)#
-----------------------------STACK AFTER CONNECTION----------------------------------
Stack#21:27:22: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
21:27:39: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
21:28:24: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
21:28:33: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52 ports)
21:28:33: %S50V:0 %CHMGR-0-PS_UP: Power supply 0 in unit 0 is up
21:28:34: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 0 is up
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Figure 40-13. Adding a Stack Unit with a Conflicting Stack Provision—Before (S50 type)

```
------------------------STANDALONE BEFORE CONNECTION----------------------------------
Standalone#show system brief
Stack MAC : 00:01:e8:d5:ef:81
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Management   online         S50V        S50V        7.8.1.0     52
  1   Member       not present    S50N
  2   Member       not present    S50V
  3   Member       not present    S50V
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
-----------------------------STACK BEFORE CONNECTION----------------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       not present    S25N
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Figure 40-14. Adding a Stack Unit with a Conflicting Stack Number—After (S50 type)

```
------------------------STANDALONE AFTER CONNECTION----------------------------------
01:38:34: %STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 presentGoing for reboot.
Reason is Stack merge
Going for reboot. Reason is Stack merge
[bootup messages omitted]
Stack(stack-member-0)#
-----------------------------STACK AFTER CONNECTION----------------------------------
23:11:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:11:40: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
23:12:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:12:34: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52 ports)
23:12:34: %STKUNIT1-M:CP %CHMGR-3-STACKUNIT_MISMATCH: Mismatch: Stack unit 0 is type S50V -
type S25N required
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       type mismatch  S25N        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Remove a Unit from an S-Series Stack

The running-configuration and startup-configuration are synchronized on all stack units. A stack member that is disconnected from the stack maintain this configuration.

To remove a stack member from the stack, disconnect the stacking cables from the unit. You may do this at any time, whether the unit is powered or unpowered, online or offline. Note that if you remove a unit in the middle of the stackin a cascade topology, the stack will be split into multiple parts, and each will form a new stack according to the stacking algorithm described throughout this chapter.

Figure 40-15. Removing a Stack Member—Before (S50 type)

```
----------------------------STANDALONE BEFORE DISCONNECTION----------------------------------
Standalone(stack-member-2)#?
reset-self                Reset this unit alone
show                      Show running system information
Standalone(stack-member-2)#show ?
version                   Software version
---------------------------------STACK BEFORE DISCONNECTION----------------------------------
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Figure 40-16. Removing a Stack Member—After (S50 type)

```
----------------------------STANDALONE AFTER DISCONNECTION----------------------------------
Standalone(stack-member-2)#
                           Going for reboot. Reason is Stack split
[bootup messages omitted]
Stack#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       not present    S50V
  1   Member       not present    S50N
  2   Management   online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
---------------------------------STACK AFTER DISCONNECTION----------------------------------
Stack#3w1d15h: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
3w1d15h: %STKUNIT1-M:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
3w1d15h: %STKUNIT0-S:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       not present    S50V
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Merge Two S-Series Stacks

You may merge two stacks while they are powered and online. To merge two stacks, connect one stack to the other using stacking cables. You may not connect 12G and 24G stack ports.

- FTOS selects a primary stack manager from the two existing mangers.
- FTOS resets all the units in the losing stack, and they all become stack members.
- If there is no unit numbering conflict, the stack members retain their previous unit numbers. Otherwise, the stack manager assigns new unit numbers, based on the order that they come online.
- The stack manager overwrites the startup and running config on the losing stack members with its own.

Split an S-Series Stack

To split a stack, unplug the desired stacking cables.You may do this at any time, whether the stack is powered or unpowered, and the units are online or offline. Each portion of the split stack retains the startup and running configuration of the original stack.

For a parent stack that is split into two child stacks, A and B, each with multiple units:

- If one of the new stacks receives the primary and the secondary management units, it is unaffected by the split.
- If one of the new stacks receives only the primary management unit, that units remains the stack manager, and FTOS elects a new secondary management unit.
- If one of the new stacks receives only the secondary management unit, it becomes the primary management, and FTOS elects a new secondary management unit.
- If one of the new stacks receives neither the primary nor the secondary management unit, the stack is reset so that a new election can take place.

S-Series Stacking Configuration Tasks

- Assign Unit Numbers to Units in an S-Series Stack on page 730
- Create a Virtual Stack Unit on an S-Series Stack
- Display Information about an S-Series Stack
- Influence Management Unit Selection on an S-Series Stack
- Manage Redundancy on an S-Series Stack
- Reset a Unit on an S-Series Stack
- Recover from Stack Link Flaps

Assign Unit Numbers to Units in an S-Series Stack

Each unit in the stack has a stack number that is either assigned by you or FTOS. S55 units are numbered from 0 to 11. Stack numbers are stored in NVRAM and are preserved upon reload.

Task	Command Syntax	Command Mode
Assign a stack-number to a unit.	**stack-unit renumber**	EXEC Privilege

Note: Renumbering the stack manager triggers a failover, as shown in Message 1.

Message 1 Renumbering the Stack Manager

```
Renumbering master unit will reload the stack. Proceed to renumber [confirm yes/no]: yes
```

Create a Virtual Stack Unit on an S-Series Stack

Use virtual stack units to configure ports on the stack before adding a new unit, or to prevent FTOS from assigning a particular stack-number.

Task	Command Syntax	Command Mode
Create a virtual stack unit.	**stack-unit provision**	CONFIGURATION

Display Information about an S-Series Stack

Task	Command Syntax	Command Mode
Display for stack-identity, status, and hardware information on every unit in a stack (Figure 40-17).	**show system**	EXEC Privilege
Display most of the information in **show system**, but in a more convenient tabular form (Figure 40-18).	**show system brief**	EXEC Privilege
Display the same information in **show system**, but only for the specified unit (Figure 40-18).	**show system stack-unit**	EXEC Privilege
Display topology and stack link status for the entire stack. The available options separate the **show system stack-port** output into topology information from link status information (Figure 40-18).	**show system stack-ports** [**status** | **topology**]	EXEC Privilege

Figure 40-17. Displaying Information about an S-Series Stack—show system (S50 type)

```
Force10#show system

Stack MAC : 00:01:e8:d5:f9:6f

--  Unit 0 --
Unit Type       : Member Unit
Status          : online
Next Boot       : online
Required Type   : S50V - 48-port E/FE/GE with POE (SB)
Current Type    : S50V - 48-port E/FE/GE with POE (SB)
Master priority : 0
Hardware Rev    : 2.0
Num Ports       : 52
Up Time         : 30 min, 7 sec
FTOS Version    : 7.8.1.0
Jumbo Capable   : yes
POE Capable     : yes
Burned In MAC   : 00:01:e8:d5:ef:81
No Of MACs      : 3

--  Module 0 --
Status          : not present

--  Module 1 --
Status          : online
Module Type     : S50-01-12G-2S     - 2-port 12G Stacking (SB)
Num Ports       : 2
Hot Pluggable   : no

--  Power Supplies  --
Unit   Bay   Status       Type
---------------------------------------------------------------------------
  0     0     up            AC
  0     1     absent

--  Fan  Status  --
Unit   TrayStatus  Speed    Fan0    Fan1    Fan2    Fan3    Fan4    Fan5
--------------------------------------------------------------------------------
 0     up          low      up      up      up      up      up      up
```

Figure 40-18. Displaying Information about a stack—show system brief (S50 type)

```
Force10#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present

--  Module Info  --
Unit  Module No   Status        Module Type        Ports
---------------------------------------------------------------------------
  0   0           not present   No Module           0
  0   1           online        S50-01-12G-2S       2
  1   0           online        S50-01-12G-2S       2
  1   1           not present   No Module           0
  2   0           not present   No Module           0
  2   1           online        S50-01-12G-2S       2

--  Power Supplies  --
Unit   Bay   Status       Type
---------------------------------------------------------------------------
  0     0     up            AC
  0     1     absent
  1     0     absent
  1     1     up            DC
  2     0     up            AC
  2     1     absent

--  Fan  Status  --
Unit   TrayStatus  Speed    Fan0    Fan1    Fan2    Fan3    Fan4    Fan5
--------------------------------------------------------------------------------
 0     up          low      up      up      up      up      up      up
 1     up          low      up      up      up      up      up      up
 2     up          low      up      up      up      up      up      up
```

Figure 40-19. Displaying Information about a Stack—show system stack-ports (S50 type)

```
Force10#show system stack-ports
Topology: Daisy chain
Interface  Connection    Link Speed      Admin     Link     Trunk
                           (Gb/s)        Status    Status   Group
------------------------------------------------------------------
 0/51                      12             up        down
 0/52        2/51          12             up        up
 1/49        2/52          12             up        up
 1/50                      12             up        down
 2/51        0/52          12             up        up
 2/52        1/49          12             up        up
```

Figure 40-20. Show information about a stack—show system brief (S55)

```
Force10#show system brief
Stack MAC : 00:01:e8:55:00:85
Reload Type : normal-reload

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Management   online         S55         S55         8-3-5-34    52
  1   Standby      online         S55         S55         8-3-5-34    52
  2   Member       online         S55         S55         8-3-5-34    52
  3   Member       online         S55         S55         8-3-5-34    52
  4   Member       online         S55         S55         8-3-5-34    52
  5   Member       online         S55         S55         8-3-5-34    52
  6   Member       online         S55         S55         8-3-5-34    52
  7   Member       online         S55         S55         8-3-5-34    52
  8   Member       online         S55         S55         8-3-5-34    52
  9   Member       online         S55         S55         8-3-5-34    52
 10   Member       online         S55         S55         8-3-5-34    52
 11   Member       online         S55         S55         8-3-5-34    52

--  Module Info  --
Unit  Module No   Status        Module Type        Ports
---------------------------------------------------------------------------
  0   0           online        S55-12G-2S          2
  0   1           online        S55-10GE-2SFP+      2
  1   0           online        S55-12G-2S          2
  1   1           not present   No Module           0
  2   0           online        S55-12G-2S          2
  2   1           not present   No Module           0
<<<<<<<<<<<<<<information truncated >>>>>>>>>>>>>>>>>>>>>>>>>>>>
 10   1           online        S55-10GE-2SFP+      2
 11   0           online        S55-12G-2S          2
 11   1           not present   No Module           0

--  Power Supplies  --
Unit   Bay   Status       Type    FanStatus
---------------------------------------------------------------------------
  0     0     absent              absent or down
  0     1     up            DC     good
  1     0     up            DC     good
  1     1     absent              absent or down
  2     0     absent              absent or down
  2     1     up            AC     good
<<<<<<<<<<<<<<information truncated >>>>>>>>>>>>>>>>>>>>>>>>>>>>
 11     0     up            AC     good
 11     1     absent              absent or down

--  Fan  Status  --
Unit Bay   TrayStatus  Fan0    Speed
--------------------------------------------------------------------------------
 0    0     up          up      9000
 0    1     up          up      8760
 1    0     up          up      8880
 1    1     up          up      8880
 2    0     up          up      8520
 2    1     up          up      8880
<<<<<<<<<<<<<information truncated >>>>>>>>>>>>>>>>>>>>>>>>>>>>
10    0     up          up      8880
10    1     up          up      8880
11    0     up          up      8880
11    1     up          up      8760
```

Figure 40-21. Displaying Information about an S-Series Stack—show system stack-ports (S55)

```
Force10#Force10#show system stack-ports status
Topology: Ring
Interface  Link Speed      Admin     Link     Trunk
             (Gb/s)        Status    Status   Group
----------------------------------------------------
 0/48        12             up        up
 0/49        12             up        up
 1/48        12             up        up
 1/49        12             up        up
 2/48        12             up        up
 2/49        12             up        up
 3/48        12             up        up
 3/49        12             up        up
 4/48        12             up        up
 4/49        12             up        up
 5/48        12             up        up
 5/49        12             up        up
 6/48        12             up        up
 6/49        12             up        up
 7/48        12             up        up
 7/49        12             up        up
 8/48        12             up        up
 8/49        12             up        up
 9/48        12             up        up
 9/49        12             up        up
 10/48        12             up        up
 10/49        12             up        up
 11/48        12             up        up
 11/49        12             up        up
Force10#
```

Influence Management Unit Selection on an S-Series Stack

Stack Priority is the system variable that FTOS uses to determine which units in the stack will be the primary and secondary management units. If multiple units tie for highest priority, then the unit with the highest MAC address prevails.

If management was determined by priority only, a change in management occurs when:

- you powered down, or offline the management unit, or a failover occurs
- you disconnect the management unit from the stack

Task	Command Syntax	Command Mode
Influence the selection of the stack management units. The unit with the numerically highest priority is elected the primary management unit, and the unit with the second highest priority is the secondary management unit. Default: 0 Range: 1-14	**stack-unit priority**	CONFIGURATION

Manage Redundancy on an S-Series Stack

Task	Command Syntax	Command Mode
Reset the current management unit, and make the secondary management unit the new primary. A new secondary is elected, and when the former stack manager comes back online, it becomes a member unit.	**redundancy force-failover stack-unit**	EXEC Privilege
Prevent the stack manager from rebooting after a failover. This command does not affect a forced failover, manual reset, or a stack-link disconnect.	**redundancy disable-auto-reboot stack-unit**	CONFIGURATION
Display redundancy information.	**show redundancy**	EXEC Privilege

Reset a Unit on an S-Series Stack

You may reset any stack unit except for the master (Message 2).

Message 2 Master Reset Disallowed

```
% Error: Reset of master unit is not allowed.
```

Task	Command Syntax	Command Mode
Reload a stack-unit	**reset stack-unit** *0-11*	EXEC Privilege
Reload a member unit, from the unit itself	**reset-self**	EXEC Privilege
Reset a stack-unit when the unit is in a problem state.	**reset stack-unit** *0-11* **hard**	EXEC Privilege

Monitor an S-Series Stack with SNMP

S-Series supports the following tables in *f10-ss-chassis.mib* for stack management through SNMP:

- chStackUnitTable
- chSysStackPortTable

Troubleshoot an S-Series Stack

- Recover from Stack Link Flaps
- Recover from a Card Problem State on an S-Series Stack on page 737
- Recover from a Card Mismatch State on an S-Series Stack

Recover from Stack Link Flaps

S-Series Stack Link Integrity Monitoring enables units to monitor their own stack ports, and disable any stack port that flaps five times within 10 seconds. FTOS displays console messages the local and remote members of a flapping link, and on the primary and secondary management units as KERN-2-INT messages if the flapping port belongs to either of these units.

In Figure 40-22, a stack-port on the manager flaps. The remote member, Member 2, displays a console message, and the manager and standby display KERN-2-INT messages.

To re-enable the downed stack-port, power cycle the offending unit.

Figure 40-22. Recovering from a Stack Link Flapping Error

```
------------------------------------MANAGMENT UNIT-----------------------------------------
Error: Stack Port 50 has flapped 5 times within 10 seconds.Shutting down this st
ack port now.
Error: Please check the stack cable/module and power-cycle the stack.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times w
ithin 10 seconds.Shutting down this stack port now.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Please check the stack cable/module
 and power-cycle the stack.
---------------------------------------STANDBY UNIT-----------------------------------------
10:55:18: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times within 10
seonds.Shutting down this stack port now.
10:55:18: %STKUNIT1-M:CP %KERN-2-INT: Error: Please check the stack cable/module
 and power-cycle the stack.
----------------------------------------MEMBER 2--------------------------------------------
Error: Stack Port 51 has flapped 5 times within 10 seconds.Shutting down this stack port now.
Error: Please check the stack cable/module and power-cycle the stack.
```

Recover from a Card Problem State on an S-Series Stack

If a unit added to a stack has has a different FTOS version, the unit does not come online, and FTOS cites a card problem error, as shown in Figure 40-23. To recover, disconnect the new unit from the stack, change the FTOS version to match the stack, and then reconnect it to the stack.

Figure 40-23. Recovering from a Card Problem Error on an S-Series Stack (S50 type)

```
F
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       card problem   S25N        unknown     7.7.1.1     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Recover from a Card Mismatch State on an S-Series Stack

A card mismatch occurs if the stack has a provision for the lowest available stack number which does not match the model of a newly added unit (Figure 40-24). To recover, disconnect the new unit. Then, either:

- remove the provision from the stack, then reconnect the standalone unit, or
- renumber the standalone unit with another available stack number on the stack.

Figure 40-24. Recovering from a Card Mismatch State on an S-Series Stack (S50 type)

```
-----------------------------------STANDALONE UNIT BEFORE-----------------------------------
Standalone#show system brief
Stack MAC : 00:01:e8:d5:ef:81
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Management   online         S50V        S50V        7.8.1.0     52
  1   Member       not present    S50N
  2   Member       not present    S50V
  3   Member       not present    S50V
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
---------------------------------------STACK BEFORE----------------------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       not present    S25N
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
-----------------------------------STANDALONE UNIT AFTER-----------------------------------
01:38:34: %STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
Going for reboot. Reason is Stack merge
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
----------------------------------------STACK AFTER----------------------------------------
23:11:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:11:40: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
23:12:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:12:34: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52 ports)
23:12:34: %STKUNIT1-M:CP %CHMGR-3-STACKUNIT_MISMATCH: Mismatch: Stack unit 0 is type S50V -
type S25N required

Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
  0   Member       type mismatch  S25N        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

41

Storm Control

Storm Control is supported on platforms: C E S

Storm Control for Multicast is supported on platforms: C S

The storm control feature enables you to control unknown-unicast and broadcast traffic on Layer 2 and Layer 3 physical interfaces.

FTOS Behavior: On the E-Series, FTOS supports broadcast control for Layer 3 traffic only. To control Layer 2 broadcast traffic use the command **storm-control unknown-unicast**. On the C-Series and S-Series, FTOS supports broadcast control (command **storm-control broadcast**) for Layer 2 *and* Layer 3 traffic.

FTOS Behavior: On E-Series, bi-directional traffic (unknown unicast and broadcast) along with egress storm control causes the configured traffic rates to be split between the involved ports. The percentage of traffic that each port receives after the split is not predictable. These ports can be in the same/different port-pipes or on the same/different line cards.

FTOS Behavior: The minimum number of packets per second (PPS) that storm control can limit on the S55 is 2.

Configure Storm Control

Storm control is supported in INTERFACE mode and CONFIGURATION mode

Configure storm control from INTERFACE mode

Configure storm control from INTERFACE mode using the command storm control. From INTERFACE mode:

- You can only on configure storm control for ingress traffic.
- If you configure storm control from both INTERFACE and CONFIGURATION mode, the INTERFACE mode configurations override the CONFIGURATION mode configurations.
- The percentage of storm control is calculated based on the advertised rate of the line card, not by the speed setting.

Configure storm control from CONFIGURATION mode

Configure storm control from CONFIGURATION mode using the command storm control. From CONFIGURATION mode you can configure storm control for ingress and egress traffic.

Do not apply per-VLAN QoS on an interface that has storm-control enabled (either on an interface or globally)

- On the E-Series, when broadcast storm-control is enabled on an interface or globally on the ingress and DSCP marking for a DSCP value 1 is configured for the data traffic, the traffic goes to queue 1 instead of queue 0. Similarly, if unicast storm-control is enabled on an interface or globally on the ingress, and DSCP marking for a DSCP value 2 is configured for the data traffic, the traffic goes to queue 2 instead of queue 0.

42

System Time and Date

Chapter 42, System Time and Date settings, and Network Time Protocol are supported on platforms: C E S

System times and dates can be set and maintained through the Network Time Protocol (NTP). They are also set through FTOS CLIs and hardware settings.

This chapter includes the following sections:

- Network Time Protocol
    - Protocol Overview
    - Implementation Information
    - Configuring Network Time Protocol
- FTOS Time and Date
    - Configuring time and date settings
    - Set daylight savings time

Network Time Protocol

Network Time Protocol (NTP) synchronizes timekeeping among a set of distributed time servers and clients. The protocol also coordinates time distribution in a large, diverse network with a variety of interfaces. In NTP, servers maintain the time and NTP clients synchronize with a time-serving host. NTP clients choose from among several NTP servers to determine which offers the best available source of time and the most reliable transmission of information.

NTP is a fault-tolerant protocol that will automatically select the best of several available time sources to synchronize to. Multiple candidates can be combined to minimize the accumulated error. Temporarily or permanently insane time sources will be detected and avoided.

Dell Force10 recommends configuring NTP for the most accurate time. In FTOS, other time sources can be configured (the hardware clock and the software clock).

NTP is designed to produce three products: clock offset, roundtrip delay, and dispersion, all of which are relative to a selected reference clock.

- **Clock offset** represents the amount to adjust the local clock to bring it into correspondence with the reference clock.

- **Roundtrip delay** provides the capability to launch a message to arrive at the reference clock at a specified time.
- **Dispersion** represents the maximum error of the local clock relative to the reference clock.

Since most host time servers will synchronize via another peer time server, there are two components in each of these three products, those determined by the peer relative to the primary reference source of standard time and those measured by the host relative to the peer.

Each of these components are maintained separately in the protocol in order to facilitate error control and management of the subnet itself. They provide not only precision measurements of offset and delay, but also definitive maximum error bounds, so that the user interface can determine not only the time, but the quality of the time as well.

In what may be the most common client/server model a client sends an NTP message to one or more servers and processes the replies as received. The server interchanges addresses and ports, overwrites certain fields in the message, recalculates the checksum and returns the message immediately. Information included in the NTP message allows the client to determine the server time with respect to local time and adjust the local clock accordingly. In addition, the message includes information to calculate the expected timekeeping accuracy and reliability, as well as select the best from possibly several servers.

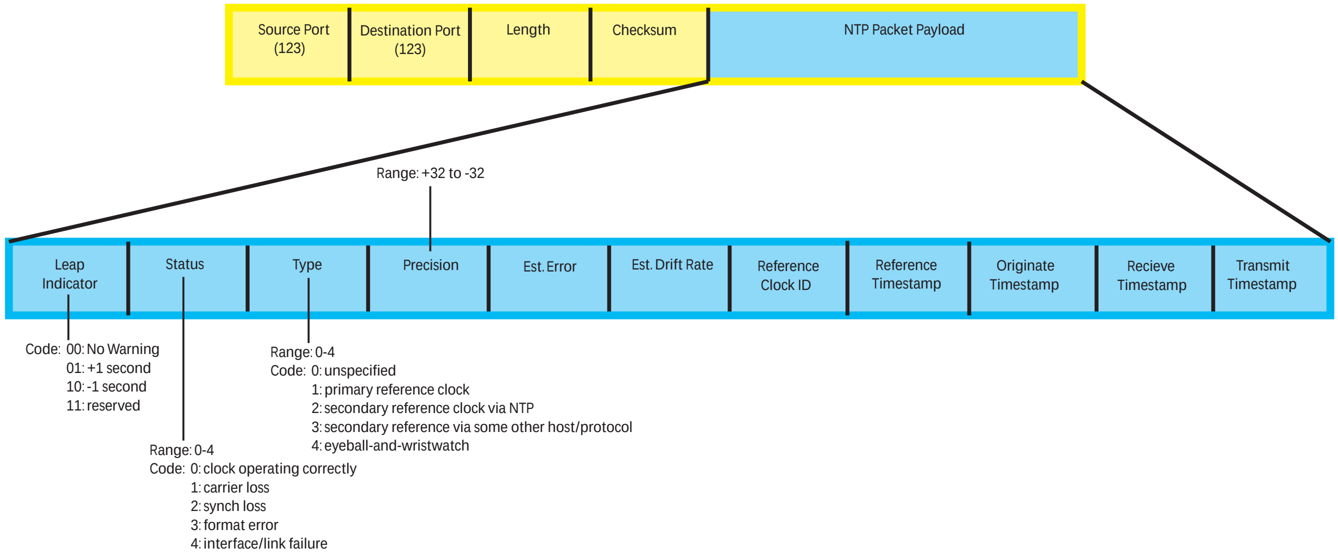
Following conventions established by the telephone industry [BEL86], the accuracy of each server is defined by a number called the stratum, with the topmost level (primary servers) assigned as one and each level downwards (secondary servers) in the hierarchy assigned as one greater than the preceding level.

FTOS synchronizes with a time-serving host to get the correct time. You can set FTOS to poll specific NTP time-serving hosts for the current time. From those time-serving hosts, the system chooses one NTP host with which to synchronize and serve as a client to the NTP host. As soon as a host-client relationship is established, the networking device propagates the time information throughout its local network.

Protocol Overview

NTP message to one or more servers and processes the replies as received. The server interchanges addresses and ports, fills in or overwrites certain fields in the message, recalculates the checksum and returns it immediately. Information included in the NTP message allows each client/server peer to determine the timekeeping characteristics of its other peers, including the expected accuracies of their clocks. Using this information each peer is able to select the best time from possibly several other clocks, update the local clock and estimate its accuracy.

Figure 42-1. NTP Fields



Implementation Information

- Dell Force10 systems can only be an NTP client.

Configuring Network Time Protocol

Configuring NTP is a one-step process:

1. Enable NTP. See page 746.

Related Configuration Tasks

- Configure NTP broadcasts on page 747
- Set the Hardware Clock with the Time Derived from NTP on page 747
- Set the Hardware Clock with the Time Derived from NTP on page 747
- Disable NTP on an interface on page 747
- Configure a source IP address for NTP packets on page 748 (optional)

Enable NTP

NTP is disabled by default. To enable it, specify an NTP server to which the Dell Force10 system will synchronize. Enter the command multiple times to specify multiple servers. You may specify an unlimited number of servers at the expense of CPU resources.

Task	Command	Command Mode
Specify the NTP server to which the Dell Force10 system will synchronize.	**ntp server** *ip-address*	CONFIGURATION

Display the system clock state with respect to NTP using the command **show ntp status** from EXEC Privilege mode, as shown in Figure 42-2.

Figure 42-2. Displaying the System Clock State with respect to NTP

```
R6_E300(conf)#do show ntp status
Clock is synchronized, stratum 2, reference is 192.168.1.1
frequency is -369.623 ppm, stability is 53.319 ppm, precision is 4294967279
reference time is CD63BCC2.0CBBD000 (16:54:26.049 UTC Thu Mar 12 2009)
clock offset is 997.529984 msec, root delay is 0.00098 sec
root dispersion is 10.04271 sec, peer dispersion is 10032.715 msec
peer mode is client
```

Display the calculated NTP synchronization variables received from the server that the system will use to synchronize its clock using the command **show ntp associations** from EXEC Privilege mode, as shown in Figure 42-3.

Figure 42-3. Displaying the Calculated NTP Synchronization Variables

```
R6_E300(conf)#do show ntp associations
   remote        ref clock     st when poll reach   delay   offset    disp
==========================================================================
#192.168.1.1     .LOCL.           1   16   16   76     0.98   -2.470  879.23
* master (synced), # master (unsynced), + selected, - candidate
```

Set the Hardware Clock with the Time Derived from NTP

Task	Command	Command Mode
Periodically update the system hardware clock with the time value derived from NTP.	**ntp update-calendar**	CONFIGURATION

Figure 42-4. Displaying the Calculated NTP Synchronization Variables

```
R5/R8(conf)#do show calendar
06:31:02 UTC Mon Mar 13 1989
R5/R8(conf)#ntp update-calendar 1
R5/R8(conf)#do show calendar
06:31:26 UTC Mon Mar 13 1989
R5/R8(conf)#do show calendar
12:24:11 UTC Thu Mar 12 2009
```

Configure NTP broadcasts

With FTOS, you can receive broadcasts of time information. You can set interfaces within the system to receive NTP information through broadcast.

To configure an interface to receive NTP broadcasts, use the following commands in the INTERFACE mode:

Task	Command	Command
Set the interface to receive NTP packets.	**ntp broadcast client**	INTERFACE

Table 42-1.

```
2w1d11h : NTP: Maximum Slew:-0.000470, Remainder = -0.496884
```

Disable NTP on an interface

By default, NTP is enabled on all active interfaces. If you disable NTP on an interface, FTOS drops any NTP packets sent to that interface.

To disable NTP on an interface, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ntp disable	INTERFACE	Disable NTP on the interface.

To view whether NTP is configured on the interface, use the **show config** command in the INTERFACE mode. If **ntp disable** is not listed in the **show config** command output, then NTP is enabled. (The **show config** command displays only non-default configuration information.)

Configure a source IP address for NTP packets

By default, the source address of NTP packets is the IP address of the interface used to reach the network. You can configure one interface's IP address to be included in all NTP packets.

To configure an IP address as the source address of NTP packets, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ntp source *interface*	CONFIGURATION	Enter the following keywords and slot/port or number information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by a number between 0 and 16383. • For a port channel interface, enter the keyword **lag** followed by a number from 1 to 255 for TeraScale and ExaScale. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

To view the configuration, use the **show running-config ntp** command (Figure 38) in the EXEC privilege mode.

Configure NTP authentication

NTP authentication and the corresponding trusted key provide a reliable means of exchanging NTP packets with trusted time sources. NTP authentication begins when the first NTP packet is created following the configuration of keys. NTP authentication in FTOS uses the MD5 algorithm and the key is embedded in the synchronization packet that is sent to an NTP time source.

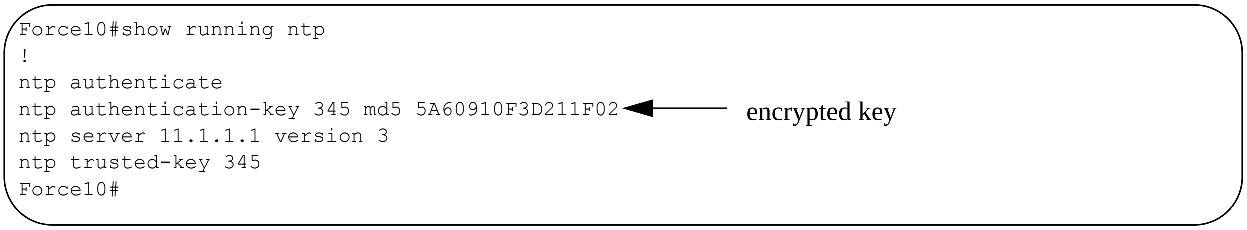
FTOS Behavior: FTOS versions 8.2.1.0 and later use an encryption algorithm to store the authentication key that is different from previous FTOS versions; beginning in version 8.2.1.0, FTOS uses DES encryption to store the key in the startup-config when you enter the command **ntp authentication-key**. Therefore, if your system boots with a startup-configuration from an FTOS versions prior to 8.2.1.0 in which you have configured **ntp authentication-key**, the system cannot correctly decrypt the key, and cannot authenticate NTP packets. In this case you must re-enter this command and save the running-config to the startup-config.

To configure NTP authentication, use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ntp authenticate**	CONFIGURATION	Enable NTP authentication.
2	**ntp authentication-key** *number* **md5** *key*	CONFIGURATION	Set an authentication key. Configure the following parameters: *number:* Range 1 to 4294967295. This *number* must be the same as the *number* in the **ntp trusted-key** command. *key:* Enter a text string. This text string is encrypted.
3	**ntp trusted-key** *number*	CONFIGURATION	Define a trusted key. Configure a number from 1 to 4294967295. The *number* must be the same as the *number* used in the **ntp authentication-key** command.

To view the NTP configuration, use the **show running-config ntp** command (Figure 40) in the EXEC privilege mode. Figure 42-5 shows an encrypted authentication key. All keys are encrypted.

Figure 42-5. show running-config ntp Command Example



```
Force10#show running ntp
!
ntp authenticate
ntp authentication-key 345 md5 5A60910F3D211F02
ntp server 11.1.1.1 version 3
ntp trusted-key 345
Force10#
```

Command Syntax	Command Mode	Purpose
ntp server *ip-address* [**key** *keyid*] [**prefer**] [**version** *number*]	CONFIGURATION	Configure an NTP server. Configure the IP address of a server and the following optional parameters: • **key** *keyid:* Configure a text string as the key exchanged between the NTP server and client. • **prefer:** Enter the keyword to set this NTP server as the preferred server. • **version** *number:* Enter a number 1 to 3 as the NTP version.

```
R6_E300(conf)#1w6d23h : NTP: xmit packet to 192.168.1.1:
 leap 0, mode 3, version 3, stratum 2, ppoll 1024
 rtdel 0219 (8.193970), rtdsp AF928 (10973.266602), refid C0A80101 (192.168.1.1)
 ref CD7F4F63.6BE8F000 (14:51:15.421 UTC Thu Apr 2 2009)
 org CD7F4F63.68000000 (14:51:15.406 UTC Thu Apr 2 2009)
 rec CD7F4F63.6BE8F000 (14:51:15.421 UTC Thu Apr 2 2009)
 xmt CD7F5368.D0535000 (15:8:24.813 UTC Thu Apr 2 2009)
1w6d23h : NTP: rcv packet from 192.168.1.1
 leap 0, mode 4, version 3, stratum 1, ppoll 1024
 rtdel 0000 (0.000000), rtdsp AF587 (10959.090820), refid 4C4F434C (76.79.67.76)
 ref CD7E14FD.43F7CED9 (16:29:49.265 UTC Wed Apr 1 2009)
 org CD7F5368.D0535000 (15:8:24.813 UTC Thu Apr 2 2009)
 rec CD7F5368.D0000000 (15:8:24.812 UTC Thu Apr 2 2009)
 xmt CD7F5368.D0000000 (15:8:24.812 UTC Thu Apr 2 2009)
 inp CD7F5368.D1974000 (15:8:24.818 UTC Thu Apr 2 2009)

rtdel-root delay
rtdsp - round trip dispersion
refid - reference id
org -
rec - (last?) receive timestamp
xmt - transmit timestamp

mode - 3 client, 4 server
stratum - 1 primary reference clock, 2 secondary reference clock (via NTP)
version - NTP version 3
leap -
```

- Leap Indicator (sys.leap, peer.leap, pkt.leap): This is a two-bit code warning of an impending leap second to be inserted in the NTP time scale. The bits are set before 23:59 on the day of insertion and reset after 00:00 on the following day. This causes the number of seconds (rollover interval) in the day of insertion to be increased or decreased by one. In the case of primary servers the bits are set by operator intervention, while in the case of secondary servers the bits are set by the protocol. The two bits, bit 0 and bit 1, respectively, are coded as follows:

- Poll Interval: integer indicating the minimum interval between transmitted messages, in seconds as a power of two. For instance, a value of six indicates a minimum interval of 64 seconds.
- Precision: integer indicating the precision of the various clocks, in seconds to the nearest power of two. The value must be rounded to the next larger power of two; for instance, a 50-Hz (20 ms) or 60-Hz (16.67ms) power-frequency clock would be assigned the value -5 (31.25 ms), while a 1000-Hz (1 ms) crystal-controlled clock would be assigned the value -9 (1.95 ms).
- Root Delay (sys.rootdelay, peer.rootdelay, pkt.rootdelay): This is a signed fixed-point number indicating the total roundtrip delay to the primary reference source at the root of the synchronization subnet, in seconds. Note that this variable can take on both positive and negative values, depending on clock precision and skew.
- Root Dispersion (sys.rootdispersion, peer.rootdispersion, pkt.rootdispersion): This is a signed fixed-point number indicating the maximum error relative to the primary reference source at the root of the synchronization subnet, in seconds. Only positive values greater than zero are possible.
- Reference Clock Identifier (sys.refid, peer.refid, pkt.refid): This is a 32-bit code identifying the particular reference clock. In the case of stratum 0 (unspecified) or stratum 1 (primary reference source), this is a four-octet, left-justified, zero-padded ASCII string, for example (see Appendix A for comprehensive list): the case of stratum 2 and greater (secondary reference) this is the four-octet Internet address of the peer selected for synchronization.
- Reference Timestamp (sys.reftime, peer.reftime, pkt.reftime): This is the local time, in timestamp format, when the local clock was last updated. If the local clock has never been synchronized, the value is zero.
- **Originate Timestamp**: The departure time on the server of its last NTP message. If the server becomes unreachable, the value is set to zero.
- **Receive Timestamp**: The arrival time on the client of the last NTP message from the server. If the server becomes unreachable, the value is set to zero.
- **Transmit Timestamp**: The departure time on the server of the current NTP message from the sender.
- Filter dispersion is the error in calculating the minimum delay from a set of sample data from a peer.

FTOS Time and Date

The time and date can be set using the FTOS CLI.

Configuring time and date settings

The following list includes the configuration tasks for setting the system time:

- Set the time and date for the switch hardware clock
- Set the time and date for the switch software clock
- Set the timezone
- Set daylight savings time
    - Set Daylight Saving Time Once
    - Set Recurring Daylight Saving Time

Set the time and date for the switch hardware clock

Command Syntax	Command Mode	Purpose
calendar set *time month day year*	EXEC Privilege	Set the hardware clock to the current time and date. *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. Range: 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. Range: 1993 to 2035.

```
Force10#calendar set 08:55:00 september 18 2009
Force10#
```

Set the time and date for the switch software clock

You can change the order of the *month* and *day* parameters to enter the time and date as *time day month year*. You cannot delete the software clock.

The software clock runs only when the software is up. The clock restarts, based on the hardware clock, when the switch reboots.

Command Syntax	Command Mode	Purpose
clock set *time month day year*	EXEC Privilege	Set the system software clock to the current time and date. *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. Range: 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. Range: 1993 to 2035.

```
Force10#clock set 16:20:00 19 september 2009
Force10#
```

Set the timezone

Coordinated Universal Time (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 UTC and your local timezone. For example, San Jose, CA is the Pacific Timezone with a UTC offset of -8.

Command Syntax	Command Mode	Purpose
clock timezone *timezone-name offset*	CONFIGURATION	Set the clock to the appropriate timezone. *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 (-) followed by a number from 1 to 23 as the number of hours

```
Force10#conf
Force10(conf)#clock timezone Pacific -8
Force10(conf)#01:40:19: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Timezone
configuration changed from "UTC 0 hrs 0 mins" to "Pacific -8 hrs
0 mins"
```

Set daylight savings time

FTOS supports setting the system to daylight savings time once or on a recurring basis every year.

Set Daylight Saving Time Once

Set a date (and time zone) on which to convert the switch to daylight savings time on a one-time basis.

Command Syntax	Command Mode	Purpose
clock summer-time *time-zone* **date** *start-month start-day start-year start-time end-month end-day end-year end-time* [*offset*]	CONFIGURATION	Set the clock to the appropriate timezone and daylight savings time. *time-zone: Enter* the three-letter name for the time zone. This name is displayed 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. Range: 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. Range: 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. Range: 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. Range: 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. Range: 1 to1440. Default: 60 minutes

Command Syntax	Command Mode	Purpose

```
Force10(conf)#clock summer-time pacific date Mar 14 2009 00:00 Nov 7 2009 00:00
Force10(conf)#02:02:13: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from
"none" to "Summer time starts 00:00:00 Pacific Sat Mar 14 2009;Summer time ends 00:00:00 pacific
Sat Nov 7 2009"
```

Set Recurring Daylight Saving Time

Set a date (and time zone) on which to convert the switch to daylight savings time on a specific day every year.

If you have already set daylight savings 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.

Command Syntax	Command Mode	Purpose
clock summer-time *time-zone* **recurring** *start-week start-day start-month start-time end-week end-day end-month end-time* [*offset*]	CONFIGURATION	Set the clock to the appropriate timezone and adjust to daylight savings time every year. *time-zone: Enter* the three-letter name for the time zone. This name is displayed in the show clock output. *start-week:* (OPTIONAL) Enter one of the following as the week that daylight savings begins and then enter values for *start-day* through *end-time*: • *week-number:* Enter a number from 1-4 as the number of the week in the month to start daylight savings time. • **first:** Enter this keyword to start daylight savings time in the first week of the month. • **last:** Enter this keyword to start daylight savings 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. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.

Command Syntax	Command Mode	Purpose
		start-year: Enter a four-digit number as the year. Range: 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 savings ends: • *week-number:* enter a number from 1-4 as the number of the week to end daylight savings time. • **first:** enter the keyword first to end daylight savings time in the first week of the month. • **last:** enter the keyword last to end daylight savings 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. Range: 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. Range: 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. Range: 1 to1440. Default: 60 minutes

```
Force10(conf)#clock summer-time pacific recurring Mar 14 2009 00:00 Nov 7 2009 00:00 ?

Force10(conf)#02:02:13: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from
"none" to "Summer time starts 00:00:00 Pacific Sat Mar 14 2009;Summer time ends 00:00:00 pacific
Sat Nov 7 2009"
```

Note: If you enter <CR> after entering the recurring command parameter, and you have already set a one-time daylight saving time/date, the system will use that time and date as the recurring setting.

Command Syntax	Command Mode	Purpose

```
Force10(conf)#clock summer-time pacific recurring ?
<1-4>                   Week number to start
first                   Week number to start
last                    Week number to start
<cr>
Force10(conf)#clock summer-time pacific recurring
Force10(conf)#02:10:57: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from
"Summer time starts 00:00:00 Pacific Sat Mar 14 2009 ; Summer time ends 00:00:00 pacific Sat Nov
7 2009" to "Summer time starts 02:00:00 Pacific Sun Mar 8 2009;Summer time ends 02:00:00 pacific
```

43

Upgrade Procedures

Find the upgrade procedures

Go to the *FTOS Release Notes* for your system type to see all the requirements to upgrade to the desired FTOS version. Follow the procedures in the *FTOS Release Notes* for the software version you wish to upgrade *to*.

Get Help with upgrades

Direct any questions or concerns about FTOS Upgrade Procedures to Dell Force10' Technical Support Center. You can reach Technical Support:

- On the Web: www.force10networks.com/support/
- By email: support@force10networks.com
- By phone: US and Canada: 866.965.5800, International: 408.965.5800

44

Virtual LANs (VLAN)

VLANs are supported on platforms C E S

This section contains the following subsections:

- Default VLAN on page 762
- Port-Based VLANs on page 763
- VLANs and Port Tagging on page 763
- Configuration Task List for VLANs on page 764
- Enable Null VLAN as the Default VLAN on page 769

Virtual LANs, or VLANs, are 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. When in Layer 2 mode, VLANs move traffic at wire speed and can span multiple devices. FTOS supports up to 4093 port-based VLANs and 1 Default VLAN, as specified in IEEE 802.1Q.

VLANs provide the following benefits:

- Improved security because you can isolate groups of users into different VLANs
- Ability to create one VLAN across multiple devices

For more information on VLANs, refer to IEEE Standard 802.1Q *Virtual Bridged Local Area Networks*. In this guide, see also:

- Bulk Configuration on page 306 in Chapter 15, “Interfaces,” on page 283
- VLAN Stacking on page 647

For a complete listing of all commands related to FTOS VLANs, see these *FTOS Command Reference* chapters:

- *Interfaces* chapter
- Port Authentication (802.1x) section in the *Security* chapter
- Chapter 13, GARP VLAN Registration Protocol.
- Chapter 36, Service Provider Bridging
- Chapter 30, Per-VLAN Spanning Tree Plus.
- For E-Series, see also the *ACL VLAN Group* and *Dell Force10 Resilient Ring Protocol* chapters.

Table 44-1 displays the defaults for VLANs in FTOS.

Table 44-1. VLAN Defaults on FTOS

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

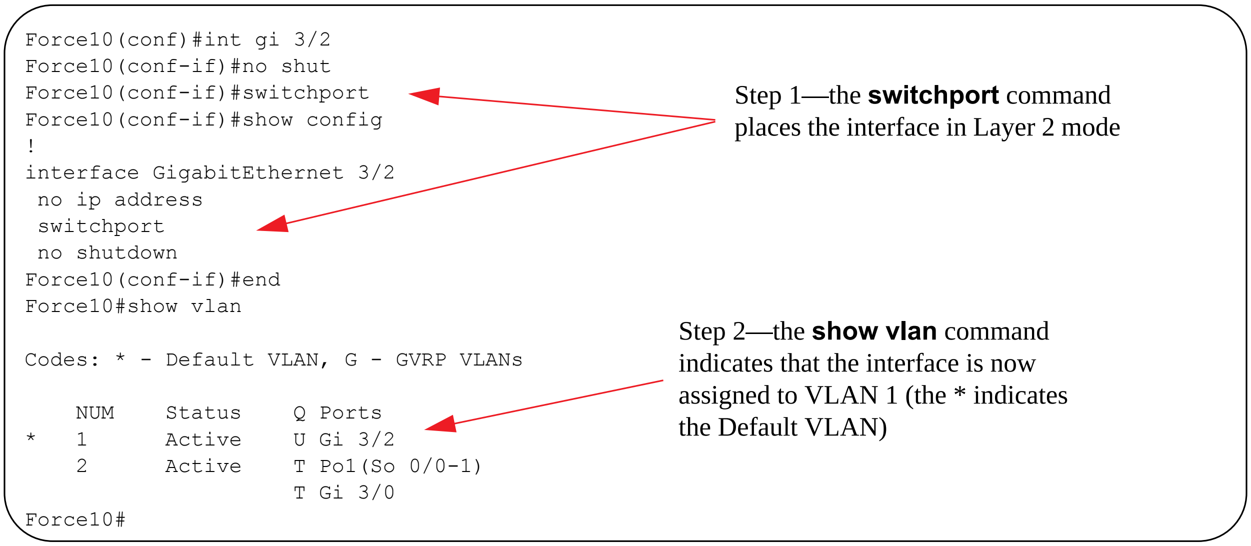
Default VLAN

When interfaces are configured for Layer 2 mode, they are automatically placed in the Default VLAN as untagged interfaces. Only untagged interfaces can belong to the Default VLAN.

Figure 44-1 displays the outcome of placing an interface in Layer 2 mode. To configure an interface for Layer 2 mode, use the **switchport** command. In Step 1, the **switchport** command places the interface in Layer 2 mode.

In Step 2, the **show vlan** command in EXEC privilege mode indicates that the interface is now part of the Default VLAN (VLAN 1).

Figure 44-1. Interfaces and the Default VLAN Example



```
Force10(conf)#int gi 3/2
Force10(conf-if)#no shut
Force10(conf-if)#switchport
Force10(conf-if)#show config
!
interface GigabitEthernet 3/2
 no ip address
 switchport
 no shutdown
Force10(conf-if)#end
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Active    U Gi 3/2
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
Force10#
```

Step 1—the **switchport** command places the interface in Layer 2 mode

Step 2—the **show vlan** command indicates that the interface is now assigned to VLAN 1 (the * indicates the Default VLAN)

By default, VLAN 1 is the Default VLAN. To change that designation, use the **default vlan-id** command in the CONFIGURATION mode. You cannot delete the Default VLAN.

Note: An IP address cannot be assigned 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 details on assigning IP addresses, see Assign an IP address to a VLAN on page 768.

Untagged interfaces must be part of a VLAN. To remove an untagged interface from the Default VLAN, you must create another VLAN and place the interface into that VLAN. Alternatively, enter the **no switchport** command, and FTOS removes the interface from the Default VLAN.

A tagged interface requires an additional step to remove it from Layer 2 mode. Since tagged interfaces can belong to multiple VLANs, you must 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, see VLANs and Port Tagging on page 763.

Port-Based VLANs

Port-based VLANs are a broadcast domain defined by different ports or interfaces. In FTOS, a port-based VLAN can contain interfaces from different line cards within the chassis. FTOS supports 4094 port-based VLANs.

Note: E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

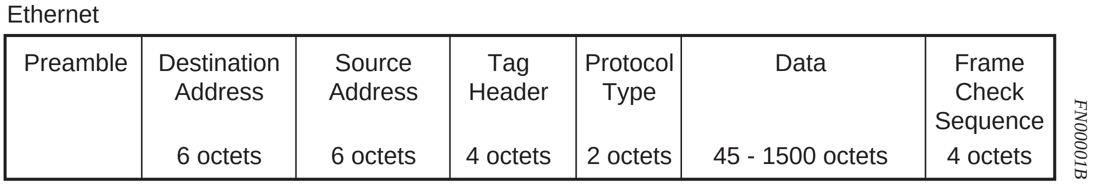
Port-based VLANs offer increased security for traffic, conserve bandwidth, and allow switch segmentation. Interfaces in different VLANs do not communicate with each other, adding some security to the traffic on those interfaces. Different VLANs can communicate between each other by means of IP routing. 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. FTOS supports IEEE 802.1Q tagging at the interface level to filter traffic. When tagging is enabled, 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. Figure 44-2 illustrates the structure of a frame with a tag header. The VLAN ID is inserted in the tag header.

Figure 44-2. Tagged Frame Format



Ethernet

Preamble	Destination Address	Source Address	Tag Header	Protocol Type	Data	Frame Check Sequence
	6 octets	6 octets	4 octets	2 octets	45 - 1500 octets	4 octets

The tag header contains some key information used by FTOS:

- 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 2 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.

Configuration Task List for VLANs

This section contains the following VLAN configuration tasks:

- Create a port-based VLAN on page 764 (mandatory)
- Assign interfaces to a VLAN on page 765 (optional)
- Assign an IP address to a VLAN on page 768 (optional)
- Enable Null VLAN as the Default VLAN on page 769

Create a port-based VLAN

The Default VLAN as VLAN 1 is part of the system startup configuration and does not require configuration. To configure a port-based VLAN, you must create the VLAN and then add physical interfaces or port channel (LAG) interfaces to the VLAN.

To create a port-based VLAN, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface vlan *vlan-id*	CONFIGURATION	Configure a port-based VLAN (if the *vlan-id* is different from the Default VLAN ID) and enter INTERFACE VLAN mode. After you create a VLAN, you must assign interfaces in Layer 2 mode to the VLAN to activate the VLAN.

Use the **show vlan** command (Figure 44-3) in the EXEC privilege mode to view the configured VLANs.

Figure 44-3. show vlan Command Example

```
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive  U So 7/4-11
    2      Active    U Gi 0/1,18
    3      Active    U Gi 0/2,19
    4      Active    T Gi 0/3,20
    5      Active    U Po 1
    6      Active    U Gi 0/12
                     U So 9/0
Force10#
```

A VLAN is active only if the VLAN contains interfaces and those interfaces are operationally up. In Figure 44-3, VLAN 1 is inactive because it contains the interfaces that are not active. The other VLANs listed in the Figure 44-3 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.

Assign interfaces to a VLAN

Only interfaces in Layer 2 mode can be assigned to a VLAN using the **tagged** and **untagged** commands. Use the **switchport** command to place an interface in Layer 2 mode.

These Layer 2 interfaces can further be designated as tagged or untagged. For more information, refer to the Interfaces chapter and Configure Layer 2 (Data Link) Mode on page 288. When an interface is placed 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. For example, Figure 44-3 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 on this command, see the command statement in the Layer 2 chapter of the *FTOS Command Reference*.

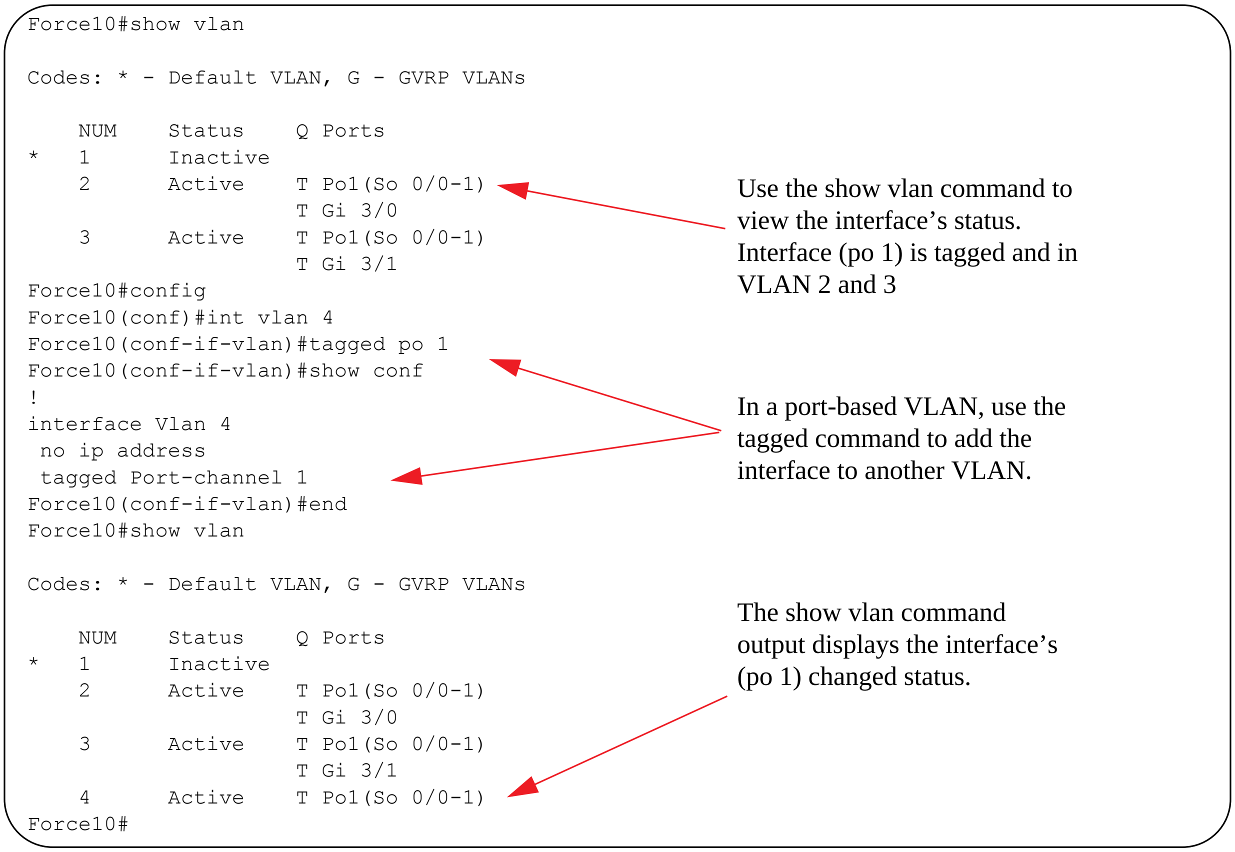
To view just the interfaces that are in Layer 2 mode, enter the **show interfaces switchport** command in the EXEC privilege mode or EXEC mode.

To tag frames leaving an interface in Layer 2 mode, you must assign that interface to a port-based VLAN to tag it with that VLAN ID. To tag interfaces, use these commands in the following sequence:

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Access the INTERFACE VLAN mode of the VLAN to which you want to assign the interface.
2	**tagged** *interface*	INTERFACE	Enable an interface to include the IEEE 802.1Q tag header.

Figure 44-4 shows the steps to add a tagged interface (in this case, port channel 1) to VLAN 4.

Figure 44-4. Example of Adding an Interface to Another VLAN



```
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
    3      Active    T Po1(So 0/0-1)
                     T Gi 3/1
Force10#config
Force10(conf)#int vlan 4
Force10(conf-if-vlan)#tagged po 1
Force10(conf-if-vlan)#show conf
!
interface Vlan 4
 no ip address
 tagged Port-channel 1
Force10(conf-if-vlan)#end
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
    3      Active    T Po1(So 0/0-1)
                     T Gi 3/1
    4      Active    T Po1(So 0/0-1)
Force10#
```

Use the show vlan command to view the interface’s status. Interface (po 1) is tagged and in VLAN 2 and 3

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. Hybrid ports can be assigned to two VLANs if the port is untagged in one VLAN and tagged in all others.

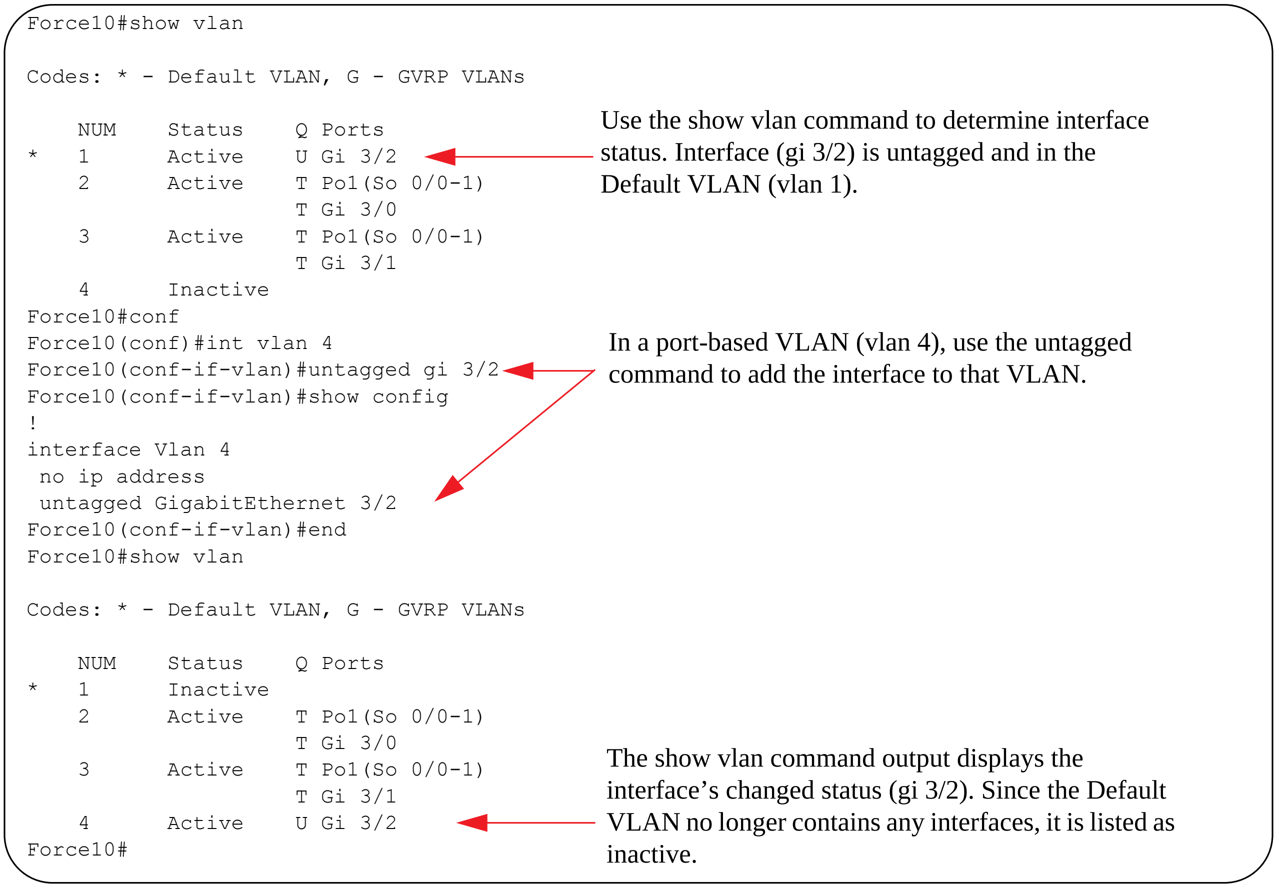
When you remove a tagged interface from a VLAN (using the **no tagged** *interface* command), it will remain tagged only if it is a tagged interface in another VLAN. If the tagged interface is removed from the only VLAN to which it belongs, the interface is placed in the Default VLAN as an untagged interface.

Use the **untagged** command to move untagged interfaces from the Default VLAN to another VLAN:

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Access the INTERFACE VLAN mode of the VLAN to which you want to assign the interface.
2	**untagged** *interface*	INTERFACE	Configure an interface as untagged. This command is available only in VLAN interfaces.

The **no untagged** *interface* command removes the untagged interface from a port-based VLAN and places the interface in the Default VLAN. You cannot use the **no untagged** *interface* command in the Default VLAN. Figure 44-5 illustrates the steps and commands to move an untagged interface from the Default VLAN to another VLAN.

Figure 44-5. Example of Moving an Untagged Interface to Another VLAN



```
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Active    U Gi 3/2
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
    3      Active    T Po1(So 0/0-1)
                     T Gi 3/1
    4      Inactive
Force10#conf
Force10(conf)#int vlan 4
Force10(conf-if-vlan)#untagged gi 3/2
Force10(conf-if-vlan)#show config
!
interface Vlan 4
 no ip address
 untagged GigabitEthernet 3/2
Force10(conf-if-vlan)#end
Force10#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
    3      Active    T Po1(So 0/0-1)
                     T Gi 3/1
    4      Active    U Gi 3/2
Force10#
```

Use the show vlan command to determine interface status. 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). Since the Default VLAN no longer contains any interfaces, it is listed as inactive.

The only way to remove an interface from the Default VLAN is to place the interface in Default mode by entering the **no switchport** command in the INTERFACE mode.

Assign 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: An IP address cannot be assigned to the Default VLAN, which, by default, is VLAN 1. To assign another VLAN ID to the Default VLAN, use the **default vlan-id** *vlan-id* command.

To assign an IP address, use the following command in INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure an IP address and mask on the interface. • *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.

In FTOS, VLANs and other logical interfaces can be placed in Layer 3 mode to receive and send routed traffic. For details, see Bulk Configuration on page 306.

VLAN Interface Counters

VLAN counters can be enabled for either Ingress packets, egress packets, or both. VLAN counters are disabled by default, and are supported on E-Series ExaScale EX only.

Command Syntax	Command Mode	Purpose
enable vlan-counter [ingress | egress | all]	CONFIGURATION	Configure ingress, egress or both counters for VLAN interfaces.

To return to the default without any VLAN counters, enter no **enable vlan-counter**.

Note: VLAN output counters may show higher than expected values because source-suppression drops are counted.

Native VLANs

Traditionally, ports can be either untagged for membership to one VLAN or tagged for membership to multiple VLANs. An untagged port must be connected to a VLAN-unaware station (one that does not understand VLAN tags), and a tagged port must be connected to a VLAN-aware station (one that generates and understands VLAN tags).

Native VLAN support breaks this barrier so that a port can be connected 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 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.

To configure a port so that it can be a member of an untagged and tagged VLANs:

Step	Task	Command	Command Mode
1	Remove any Layer 2 or Layer 3 configurations from the interface.		INTERFACE
2	Configure the interface for hybrid mode.	**portmode hybrid**	INTERFACE
3	Configure the interface for switchport mode.	**switchport**	INTERFACE
4	Add the interface to a tagged or untagged VLAN.	[**tagged** | **untagged**]	VLAN INTERFACE

Note: An existing switchport or port channel interface cannot be configured for Native VLAN. Interfaces must have no other Layer 2 or Layer 3 configurations when entering the command **portmode hybrid** or a message like Message 1 is displayed.

Message 1 Native VLAN Error

```
% Error: Port is in Layer-2 mode Gi 5/6.
```

Enable 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. FTOS has a Null VLAN to eliminate this vulnerability. When you enable the Null VLAN, all ports are placed into by it default, so that 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.

Task	Command Syntax	Command Mode
Disable the default VLAN, so that all ports belong to the Null VLAN until configured as a member of another VLAN.	**default-vlan disable** Default: the default VLAN is enabled (**no default-vlan disable**).	CONFIGURATION

45

Virtual Router Redundancy Protocol (VRRP)

Virtual Router Redundancy Protocol (VRRP) is supported on platforms C E S

This chapter covers the following information:

- VRRP Overview
- VRRP Benefits
- VRRP Implementation
- VRRP Configuration
- Sample Configurations

Virtual Router Redundancy Protocol (VRRP) is designed to eliminate a single point of failure in a statically routed network. This protocol is defined in RFC 2338 and RFC 3768.

VRRP Overview

VRRP specifies a MASTER router that owns the next hop IP and MAC address for end stations on a 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 represented by IP addresses are BACKUP routers.

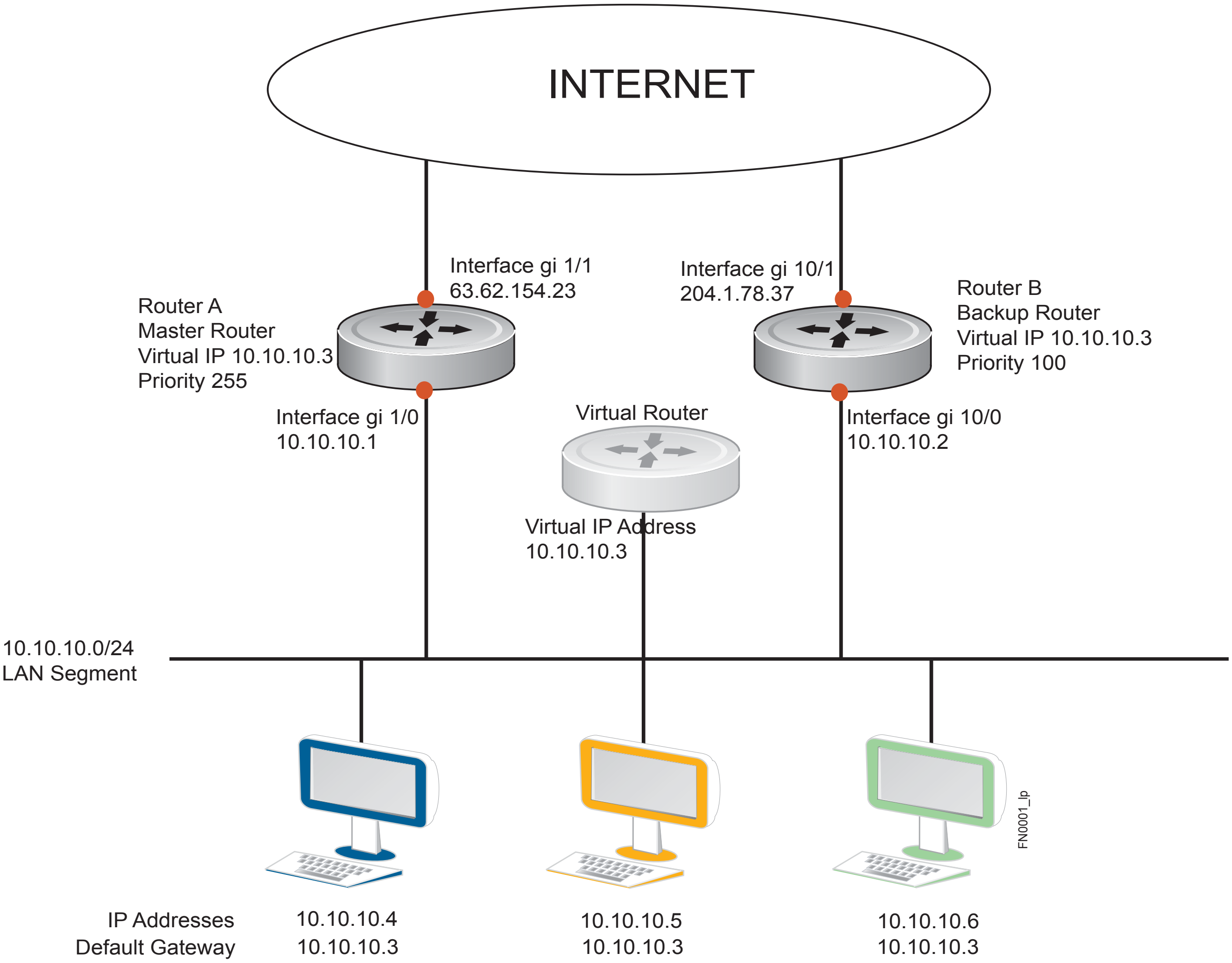
VRRP packets are transmitted with the virtual router MAC address as the source MAC address. The MAC address is in the following format: 00-00-5E-00-01-{VRID}. The first three octets are unchangeable. The next two octets (00-01) indicate the address block assigned to the VRRP protocol, and are unchangeable. The final octet changes depending on the VRRP Virtual Router Identifier and allows for up to 255 VRRP routers on a network.

Figure 45-1 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 Figure 45-1 below, 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.

Figure 45-1. Basic VRRP Configuration



For more detailed information on VRRP, refer to RFC 2338, *Virtual Router Redundancy Protocol*.

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 they are not dependent on IGP protocols to converge or update routing tables.

VRRP Implementation

E-Series supports an unlimited total number of VRRP groups on the router while supporting up to 255 VRRP groups on a single interface (Table 45-1).

C-Series supports a total of 128 VRRP groups on the switch with varying number of maximum VRRP groups per interface (Table 45-1).

S-Series supports a total of 120 VRRP groups on a switch with FTOS *or* a total of 20 VRRP groups when using SFTOS. The S-Series supports varying number of maximum VRRP groups per interface (Table 45-1).

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.

Though FTOS on E-Series supports unlimited VRRP groups, default VRRP settings may affect the maximum number of groups that can be configured and work efficiently, as a result of hardware throttling VRRP advertisement packets reaching the RP2 processor on the E-Series, the CP on the C-Series, or the FP on the S-Series. To avoid throttling VRRP advertisement packets, Dell Force10 recommends you to increase the VRRP advertisement interval to a value higher than the default value of 1 second. The recommendations are as follows:

Table 45-1. Recommended VRRP Advertise Intervals

	Recommended Advertise Interval			Groups/Interface			
Total VRRP Groups	**E-Series**	**C-Series**	**S-Series**	**E-Series ExaScale**	**E-Series TeraScale**	**C-Series**	**S-Series**
Less than 250	1 second	1 second	1 second	512	255	12	12
Between 250 and 450	2 seconds	2 - 3 seconds	2 - 3 seconds	512	255	24	24
Between 450 and 600	3 seconds	4 seconds	3 - 4 seconds	512	255	36	36
Between 600 and 800	4 seconds	5 seconds	4 seconds	512	255	48	48
Between 800 and 1000	5 seconds	5 seconds	5 seconds	512	255	84	84
Between 1000 and 1200	7 seconds	7 seconds	7 seconds	512	255	100	100
Between 1200 and 1500	8 seconds	8 seconds	8 seconds	512	255	120	120

Note: The 1500 VRRP groups are supported in FTOS Release 6.3.1.0 and later.

The recommendations in Table 45-1 may vary depending on various factors like ARP broadcasts, IP broadcasts, or STP before changing the advertisement interval. When the number of packets processed by RP2/CP/FP processor increases or decreases based on the dynamics of the network, the advertisement intervals in may increase or decrease accordingly.

△ **CAUTION:** Increasing the advertisement interval increases the VRRP Master dead interval, resulting in an increased failover time for Master/Backup election. Take extra caution when increasing the advertisement interval, as the increased dead interval may cause packets to be dropped during that switch-over time.

VRRP Configuration

By default, VRRP is not configured.

Configuration Task List for VRRP

The following list specifies the configuration tasks for VRRP:

- Create a Virtual Router on page 774 (mandatory)
- Assign Virtual IP addresses on page 775 (mandatory)
- Set VRRP Group (Virtual Router) Priority on page 777 (optional)
- Configure VRRP Authentication on page 778 (optional)
- Disable Preempt on page 779 (optional)
- Change the Advertisement interval on page 780 (optional)
- Track an Interface on page 781 (optional)

For a complete listing of all commands related to VRRP, refer to *FTOS Command Line Interface*.

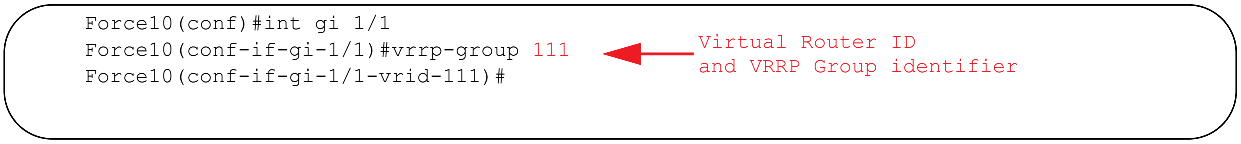
Create a Virtual Router

To enable VRRP, you must create a Virtual Router. In FTOS, a VRRP Group is identified by the Virtual Router Identifier (VRID).

To enable a Virtual Router, use the following command in the INTERFACE mode. To delete a VRRP group, use the **no vrrp-group** *vrid* command in the INTERFACE mode.

Task	Command Syntax	Command Mode
Create a virtual router for that interface with a VRID.	**vrrp-group** *vrid* VRID Range: 1-255	INTERFACE
	Note: The interface must already have a Primary IP Address defined, and be enabled.	

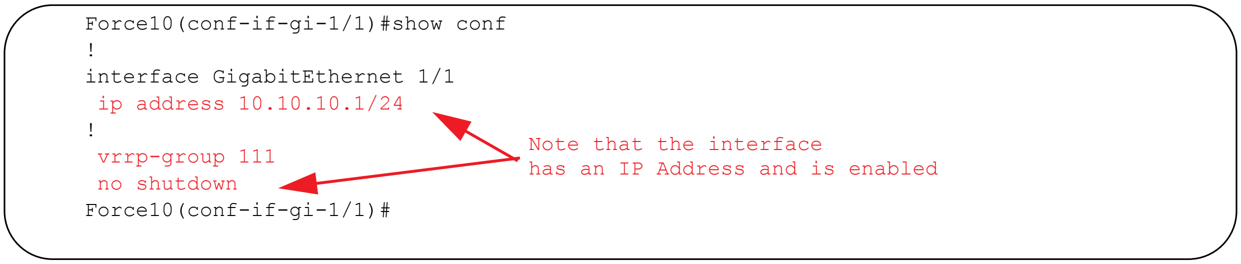
Figure 45-2. Command Example: vrrp-group



```
Force10(conf)#int gi 1/1
Force10(conf-if-gi-1/1)#vrrp-group 111
Force10(conf-if-gi-1/1-vrid-111)#
```

Virtual Router ID
and VRRP Group identifier

Figure 45-3. Command Example Display: show config for the Interface



```
Force10(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 1/1
 ip address 10.10.10.1/24
!
 vrrp-group 111
 no shutdown
Force10(conf-if-gi-1/1)#
```

Note that the interface
has an IP Address and is enabled

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.

E-Series supports an unlimited total number of VRRP Groups on the router while supporting up to 255 VRRP groups on a single interface (Table 45-1).

C-Series supports a total of 128 VRRP groups on the switch with varying number of maximum VRRP groups per interface (Table 45-1).

S-Series supports a total of 120 VRRP groups on a switch with FTOS *or* a total of 20 VRRP groups when using SFTOS. The S-Series supports varying number of maximum VRRP groups per interface (Table 45-1).

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 Force10 recommends you configure virtual IP addresses belonging to the *same* IP subnet for any one VRRP group.

For example, an interface (on which VRRP is to be enabled) 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 FTOS 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 multiple VRRP groups are configured on an interface, only one of the VRRP Groups can contain the interface primary or secondary IP address.

Configure a Virtual IP address with these commands in the following sequence in the INTERFACE mode.

Step	Task	Command Syntax	Command Mode
1	Configure a VRRP group.	**vrrp-group** *vrrp-id* VRID Range: 1-255	INTERFACE
2	Configure virtual IP addresses for this VRID.	**virtual-address** *ip-address1* [*...ip-address12*] Range: up to 12 addresses	INTERFACE -VRID

Figure 45-4. Command Example: virtual-address

```
Force10(conf-if-gi-1/1-vrid-111)#virtual-address 10.10.10.1
Force10(conf-if-gi-1/1-vrid-111)#virtual-address 10.10.10.2
Force10(conf-if-gi-1/1-vrid-111)#virtual-address 10.10.10.3
Force10(conf-if-gi-1/1-vrid-111)#
```

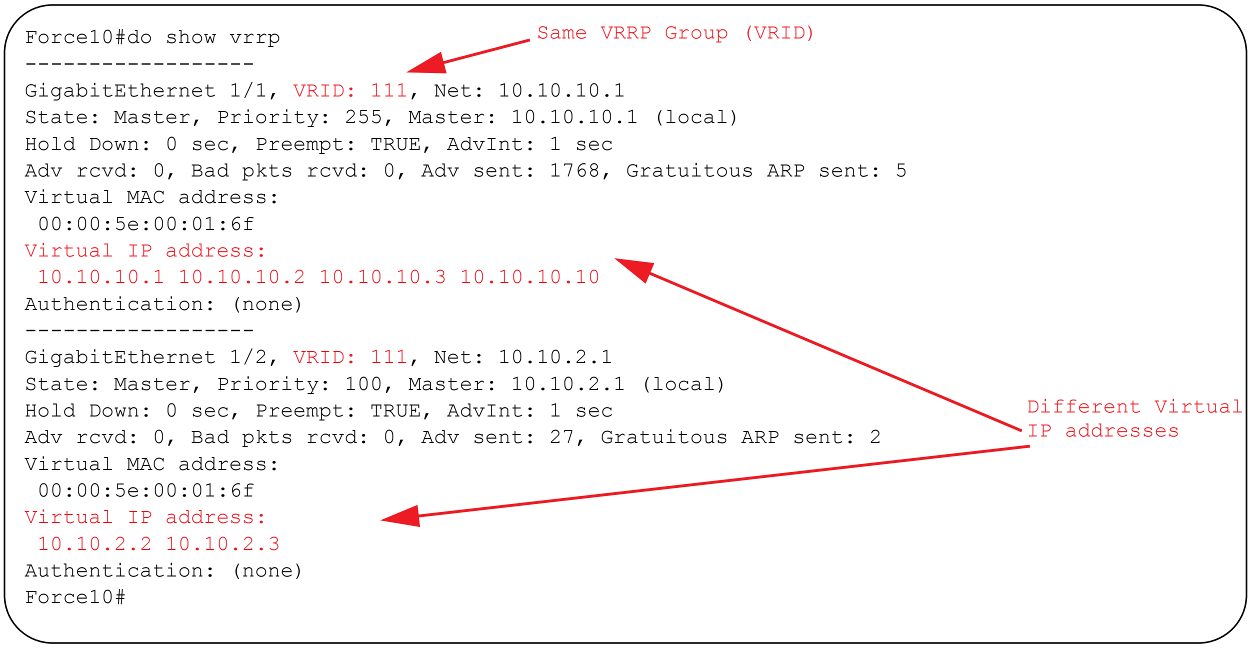
Figure 45-5. Command Example Display: show config for the Interface

```
Force10(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 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
Force10(conf-if-gi-1/1)#
```

Note that the Primary IP address and the Virtual IP addresses are on the same subnet

Figure 45-6 shows the same VRRP group configured on multiple interfaces on different subnets.

Figure 45-6. Command Example Display: show vrrp



```
Force10#do show vrrp
------------------
GigabitEthernet 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: 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)
------------------
GigabitEthernet 1/2, VRID: 111, 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)
Force10#
```

Same VRRP Group (VRID)

Different Virtual IP addresses

When the VRRP process completes its initialization, the State field contains either Master or Backup.

Set 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 will become MASTER.

Configure the VRRP Group's priority with the following command in the VRRP mode:

Task	Command Syntax	Command Mode
Configure the priority for the VRRP group.	INTERFACE -VRID	**priority** *priority* Range: 1-255 Default: 100

Figure 45-7. Command Example: priority in Interface VRRP mode

```
Force10(conf-if-gi-1/2)#vrrp-group 111
Force10(conf-if-gi-1/2-vrid-111)#priority 125
```

Figure 45-8. Command Example Display: show vrrp

```
Force10#show vrrp
------------------
GigabitEthernet 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)
------------------
GigabitEthernet 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)
Force10(conf)#
```

Configure VRRP Authentication

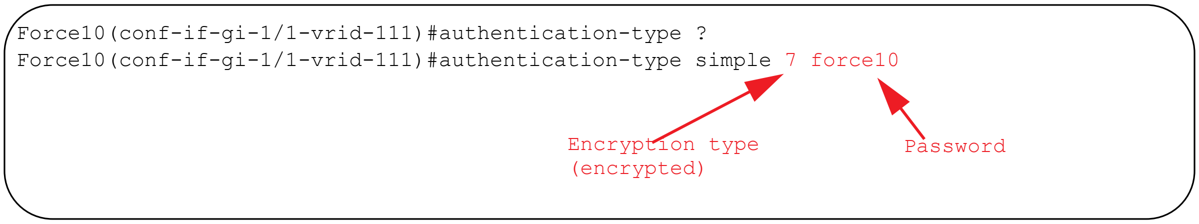
Simple authentication of VRRP packets ensures that only trusted routers participate in VRRP processes. When authentication is enabled, FTOS includes the password in its VRRP transmission, and the receiving router uses that password to verify the transmission.

Note: All virtual routers in the VRRP group must be configured the same: authentication must be enabled with the same password or authentication is disabled.

Configure simple authentication with the following command in the VRRP mode:

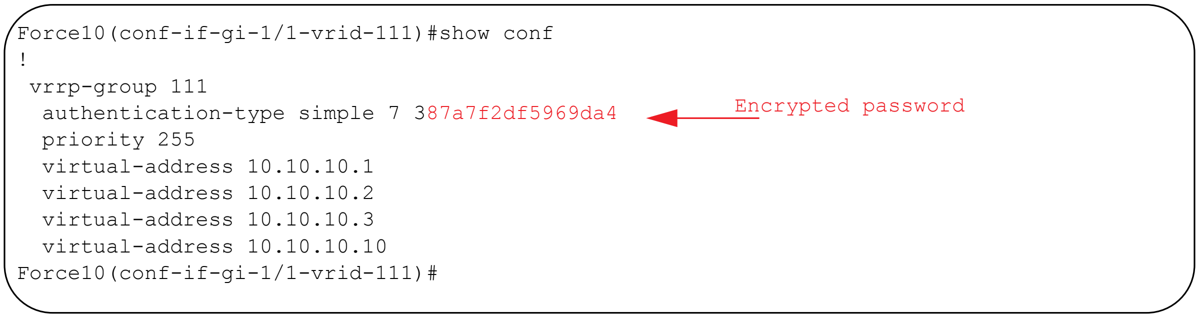
Task	Command Syntax	Command Mode
Configure a simple text password.	**authentication-type simple** [*encryption-type*] *password* Parameters: *encryption-type:* 0 indicates unencrypted; 7 indicates encrypted *password: plain text*	INTERFACE-VRID

Figure 45-9. Command Example: authentication-type



```
Force10(conf-if-gi-1/1-vrid-111)#authentication-type ?
Force10(conf-if-gi-1/1-vrid-111)#authentication-type simple 7 force10
```

Figure 45-10. Command Example: show config in VRID mode with a Simple Password Configured



```
Force10(conf-if-gi-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
Force10(conf-if-gi-1/1-vrid-111)#
```

Disable Preempt

The **preempt** command is enabled by default, and it 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: All virtual routers in the VRRP group must be configured the same: all configured with preempt enabled or configured with preempt disabled.

Since preempt is enabled by default, disable the preempt function with the following command in the VRRP mode. Re-enable preempt by entering the **preempt** command. When preempt is enabled, it does not display in the show commands, because it is a default setting.,

Task	Command Syntax	Command Mode
Prevent any BACKUP router with a higher priority from becoming the MASTER router.	**no preempt**	INTERFACE-VRID

Figure 45-11. Command Example: no preempt

```
Force10(conf-if-gi-1/1)#vrrp-group 111
Force10(conf-if-gi-1/1-vrid-111)#no preempt
Force10(conf-if-gi-1/1-vrid-111)#show conf
```

Figure 45-12. Command Example Display: show config in VRID mode

```
Force10(conf-if-gi-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
Force10(conf-if-gi-1/1-vrid-111)#
```

Change the Advertisement interval

By default, the MASTER router transmits a VRRP advertisement to all members of the VRRP group every 1 second, indicating it is operational and is the MASTER router. If the VRRP group misses 3 consecutive advertisements, then the election process begins and the BACKUP virtual router with the highest priority transitions to MASTER.

Note: Dell Force10 recommends you to increase the VRRP advertisement interval to a value higher than the default value of 1 second to avoid throttling VRRP advertisement packets . If you do change the time interval between VRRP advertisements on one router, you must change it on all participating routers.

Change that advertisement interval with the following command in the VRRP mode:

Task	Command Syntax	Command Mode
Change the advertisement interval setting.	**advertise-interval** *seconds* Range: 1-255 seconds Default: 1 second	INTERFACE-VRID

Figure 45-13. Command Example: advertise-interval

```
Force10(conf-if-gi-1/1)#vrrp-group 111
Force10(conf-if-gi-1/1-vrid-111)#advertise-interval 10
Force10(conf-if-gi-1/1-vrid-111)#
```

Figure 45-14. Command Example Display: advertise-interval in VRID mode

```
Force10(conf-if-gi-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
Force10(conf-if-gi-1/1-vrid-111)#
```

Track an Interface

Set FTOS to monitor the state of any interface according to the Virtual group. Each VRRP group can track up to 12 interfaces, which may affect the priority of the VRRP group. If the tracked interface goes down, the VRRP group's priority is decreased by a default value of 10 (also known as cost). If the tracked interface's state goes up, the VRRP group's priority is increased 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 should not exceed 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.

To track an interface, use the following command in the VRRP mode:

Task	Command Syntax	Command Mode
Monitor an interface and, optionally, set a value to be subtracted from the interface's VRRP group priority.	**track** *interface* [**priority-cost** *cost*] Cost Range: 1-254 Default: 10	INTERFACE-VRID

The sum of all the costs for all tracked interfaces must be less than or equal to the configured priority of the VRRP group.

Figure 45-15. Command Example: track

```
Force10(conf-if-gi-1/1)#vrrp-group 111
Force10(conf-if-gi-1/1-vrid-111)#track gigabitethernet 1/2
Force10(conf-if-gi-1/1-vrid-111)#
```

Figure 45-16. Command Example Display: track in VRID mode

```
Force10(conf-if-gi-1/1-vrid-111)#show conf
!
 vrrp-group 111
  advertise-interval 10
  authentication-type simple 7 387a7f2df5969da4
  no preempt
  priority 255
  track GigabitEthernet 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
Force10(conf-if-gi-1/1-vrid-111)#
```

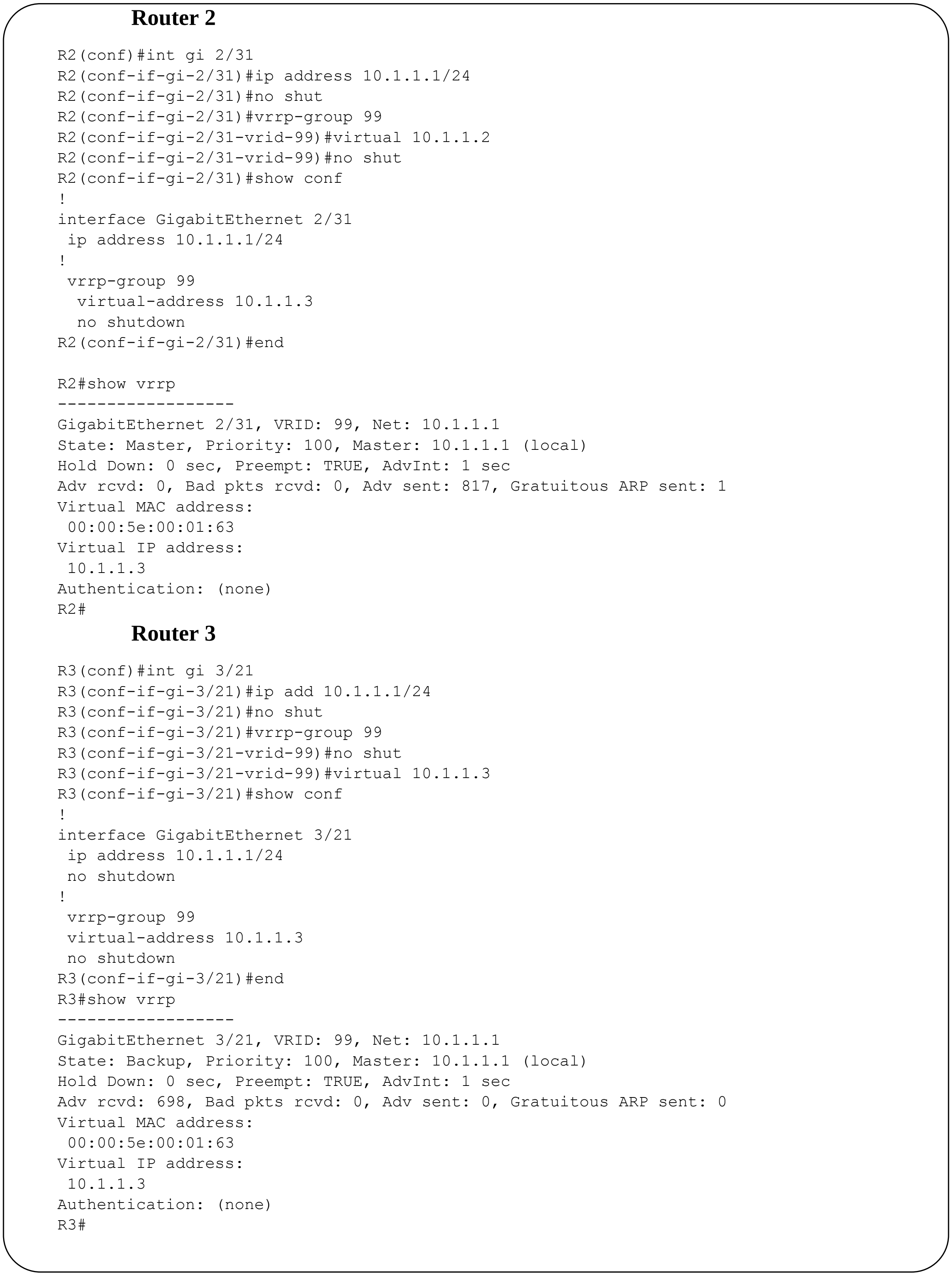
Sample Configurations

The following configurations are examples for enabling VRRP. 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.

Figure 45-17 is a sample configuration for enabling VRRP. Figure 45-18 illustrates the topology created with that CLI configuration.

Figure 45-17. Configure VRRP



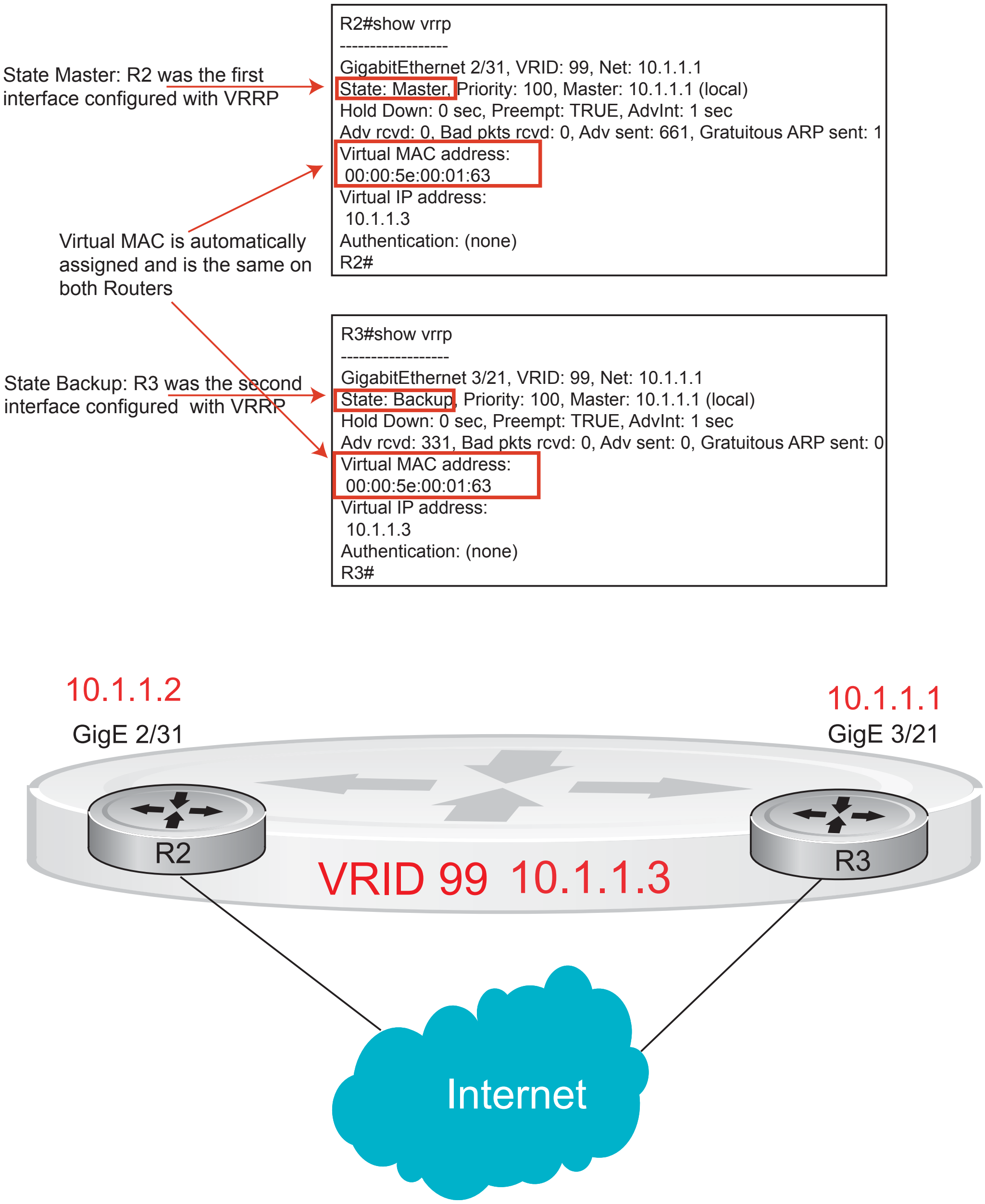
Router 2

```
R2(conf)#int gi 2/31
R2(conf-if-gi-2/31)#ip address 10.1.1.1/24
R2(conf-if-gi-2/31)#no shut
R2(conf-if-gi-2/31)#vrrp-group 99
R2(conf-if-gi-2/31-vrid-99)#virtual 10.1.1.2
R2(conf-if-gi-2/31-vrid-99)#no shut
R2(conf-if-gi-2/31)#show conf
!
interface GigabitEthernet 2/31
 ip address 10.1.1.1/24
!
 vrrp-group 99
  virtual-address 10.1.1.3
  no shutdown
R2(conf-if-gi-2/31)#end

R2#show vrrp
------------------
GigabitEthernet 2/31, VRID: 99, Net: 10.1.1.1
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 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 gi 3/21
R3(conf-if-gi-3/21)#ip add 10.1.1.1/24
R3(conf-if-gi-3/21)#no shut
R3(conf-if-gi-3/21)#vrrp-group 99
R3(conf-if-gi-3/21-vrid-99)#no shut
R3(conf-if-gi-3/21-vrid-99)#virtual 10.1.1.3
R3(conf-if-gi-3/21)#show conf
!
interface GigabitEthernet 3/21
 ip address 10.1.1.1/24
 no shutdown
!
 vrrp-group 99
 virtual-address 10.1.1.3
 no shutdown
R3(conf-if-gi-3/21)#end
R3#show vrrp
------------------
GigabitEthernet 3/21, VRID: 99, Net: 10.1.1.1
State: Backup, Priority: 100, Master: 10.1.1.1 (local)
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)
R3#
```

Figure 45-18. VRRP Topography Illustration



```
R2#show vrrp
------------------
GigabitEthernet 2/31, VRID: 99, Net: 10.1.1.1
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 661, Gratuitous ARP sent: 1
Virtual MAC address:
 00:00:5e:00:01:63
Virtual IP address:
 10.1.1.3
Authentication: (none)
R2#
```

```
R3#show vrrp
------------------
GigabitEthernet 3/21, VRID: 99, Net: 10.1.1.1
State: Backup, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 331, 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)
R3#
```

State Master: R2 was the first interface configured with VRRP

Virtual MAC is automatically assigned and is the same on both Routers

State Backup: R3 was the second interface configured with VRRP

46

S-Series Debugging and Diagnostics

The chapter contains the following major sections:

- Offline diagnostics
- Trace logs
- Last restart reason (S55)
- show hardware commands (S55)
- Troubleshooting packet loss
- Application core dumps
- Mini core dumps

Offline diagnostics

The offline diagnostics test suite is useful for isolating faults and debugging hardware.The diagnostics tests are grouped into three levels:

- **Level 0**—Level 0 diagnostics check for the presence of various components and perform essential path verifications. In addition, they verify the identification registers of the components on the board.
- **Level 1**—A smaller set of diagnostic tests. Level 1 diagnostics perform status/self-test for all the components on the board and test their registers for appropriate values. In addition, they perform extensive tests on memory devices (e.g., SDRAM, flash, NVRAM, EEPROM) wherever possible.
- **Level 2**—The full set of diagnostic tests. Level 2 diagnostics are used primarily for on-board loopback tests and more extensive component diagnostics. Various components on the board are put into loopback mode, and test packets are transmitted through those components. These diagnostics also perform snake tests using VLAN configurations.

Important Points to Remember

- You can only perform offline diagnostics on an offline standalone unit or offline member unit of a stack of three or more. You cannot perform diagnostics on the management or standby unit in a stack of two or more (Message 1).

  Message 1 Offline Diagnostics on Master/Standby Error

  ```
  Running Diagnostics on master/standby unit is not allowed on stack.
  ```

- Perform offline diagnostics on one stack member at a time.
- 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

1. Place the unit in the offline state using the **offline stack-unit** command from EXEC Privilege mode, as shown in Figure 46-1. YOu cannot enter the command on a Master or Standby stack unit.

The system reboots when the off-line diagnostics complete. This is an automatic process. A warning message appears when the **offline stack-unit** command is implemented.

```
Warning - Diagnostic execution will cause stack-unit to reboot after completion of
diags.
Proceed with Offline-Diags [confirm yes/no]:y
```

Figure 46-1. Taking an S-Series Stack Unit Offline

```
Force10#offline stack-unit 2
Warning - Diagnostic execution will cause stack-unit to reboot after completion of diags.
Proceed with Offline-Diags [confirm yes/no]:y
5w6d12h: %STKUNIT0-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - stack unit offline
5w6d12h: %STKUNIT0-M:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
Force10#5w6d12h: %STKUNIT1-S:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
```

2. Use the **show system brief** command from EXEC Privilege mode to confirm offline status, as shown in Figure 46-2.

Figure 46-2. Verifying the Offline/Online Status of an S-Series Stack Unit

```
Force10#show system brief | no-more

Stack MAC : 00:01:e8:d6:02:39

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
---------------------------------------------------------------------------
0 Standby      online         S25V        S25V        4.7.7.220   28
  1   Management   offline S50N         S50N        4.7.7.220   52
  2   Member       online         S25P        S25P        4.7.7.220   28
3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present

--  Module Info  --
Unit  Module No   Status        Module Type        Ports
---------------------------------------------------------------------------
  0   0           online        S50-01-10GE-2C      2
  0   1           online        S50-01-12G-2S       2
  1   0           online        S50-01-10GE-2P      2
  1   1           online        S50-01-12G-2S       2
  2   0           not present   No Module           0
  2   1           offline       S50-01-12G-2S       2

--  Power Supplies  --
Unit   Bay   Status       Type
---------------------------------------------------------------------------
  0     0     up            AC
  0     1     absent
  1     0     up            AC
  1     1     absent
  2     0     up            AC
  2     1     absent
```

3. Start diagnostics on the unit using the command **diag**, as shown in Figure 46-3. When the tests are complete, the system displays syslog Message 2, and automatically reboots the unit. Diagnostic results are printed to a file in the flash using the filename format *TestReport-SU-<stack-unit>.txt*.

Message 2 Offline Diagnostics Complete

```
Force10#00:09:32 : Diagnostic test results are stored on file: flash:/TestReport-SU-1.txt
00:09:37: %S50N:1 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on stack unit 1
Diags completed... Rebooting the system now!!!
```

As shown in Figure 46-3 and Figure 46-4, log messages differ somewhat when diagnostics are done on a standalone unit and on a stack member.

Figure 46-3. Running Offline Diagnostics on an S-Series Standalone Unit

```
Force10#diag stack-unit 1 alllevels
Warning - diagnostic execution will cause multiple link flaps on the peer side - advisable
to shut directly connected ports
Proceed with Diags [confirm yes/no]: yes
00:03:35: %S50N:1 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on stack unit 1
00:03:35 : Approximate time to complete these Diags ... 6 Min
S50N#00:09:32 : Diagnostic test results are stored on file: flash:/TestReport-SU-0.txt
00:09:37: %S50N:0 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on stack unit 0
Diags completed... Rebooting the system now!!!

[reboot output omitted]

S50N#00:01:35: %STKUNIT0-M:CP %SYS-5-CONFIG_I: Configured from console by  console
dir
Directory of flash:

  1  drw-      16384   Jan 01 1980 00:00:00 +00:00 .
  2  drwx       1536   Feb 29 1996 00:05:22 +00:00 ..
  3  drw-        512   Aug 15 1996 23:09:48 +00:00 TRACE_LOG_DIR
  4  d---        512   Aug 15 1996 23:09:52 +00:00 ADMIN_DIR
  5  -rw-       3854   Sep 24 1996 03:43:46 +00:00 startup-config
  6  -rw-      12632   Nov 05 2008 17:15:16 +00:00 TestReport-SU-1.txt

flash: 3104256 bytes total (3086336 bytes free)
```

Figure 46-4 shows the output of the master and member units when you run offline diagnostics on a member unit.

Figure 46-4. Running Offline Diagnostics on an S-Series Stack Member

```
[output from master unit]
Force10#diag stack-unit 2
Warning - the stack unit will be pulled out of the stack for diagnostic execution
Proceed with Diags [confirm yes/no]: yes
Warning - diagnostic execution will cause multiple link flaps on the peer side - advisable
to shut directly connected ports
Proceed with Diags [confirm yes/no]: yes
Force10#00:03:13: %S25P:2 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on stack unit 2
00:03:13 : Approximate time to complete these Diags ... 6 Min
00:03:13 : Diagnostic test results will be stored on stack unit 2 file: flash:/
TestReport-SU-2.txt
Force10#00:03:35: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
00:08:50: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
00:09:00: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 2 (type S25P, 28 ports)
00:09:00: %S25P:2 %CHMGR-0-PS_UP: Power supply 0 in unit 2 is up
00:09:00: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 2 is up
[output from the console of the unit in which diagnostics are performed]
Force10(stack-member-2)#
Diagnostic test results are stored on file: flash:/TestReport-SU-2.txt
Diags completed... Rebooting the system now!!!
```

4. View the results of the diagnostic tests using the command **show file flash://** from EXEC Privilege mode, as shown in Figure 46-5.

Figure 46-5. Viewing the Results of Offline Diagnostics on a Standalone Unit

```
Force10#show file flash://TestReport-SU-0.txt

**********************************S-Series Diagnostics********************
Stack Unit Board Serial Number : DL267160098
CPU Version : MPC8541, Version: 1.1
PLD Version : 5
Diag image based on build : E_MAIN4.7.7.206
Stack Unit Board Voltage levels - 3.300000 V, 2.500000 V, 1.800000 V, 1.250000 V, 1.200000
V, 2.000000 V
Stack Unit Board temperature : 26 Degree C
Stack Unit Number : 0

****************************Stack Unit EEPROM INFO*******************************

********MFG INFO*******************

Data in Chassis Eeprom Mfg Info is listed as...
Vendor Id: 07
Country Code: 01
Date Code: 12172007
Serial Number: DL267160098
Part Number: 7590003600
Product Revision: B
Product Order Number: ${

**************************** LEVEL 0 DIAGNOSTICS**************************

Test 0 - CPLD Presence Test ......................................... PASS
 Hardware PCB Revision is - Revision B
Test 1 - CPLD Hardware PCB Revision Test ............................ PASS
Test 2.000 - CPLD Fan-0 Presence Test ............................... PASS
Test 2.001 - CPLD Fan-1 Presence Test ............................... PASS
Test 2.002 - CPLD Fan-2 Presence Test ............................... PASS
Test 2.003 - CPLD Fan-3 Presence Test ............................... PASS
Test 2.004 - CPLD Fan-4 Presence Test ............................... PASS
Test 2.005 - CPLD Fan-5 Presence Test ............................... PASS
Test 3.000 - CPLD Power Bay-0 Presence Test ......................... PASS
Test 3.001 - CPLD Power Bay-1 Presence Test .........................     NOT PRESENT
Test 4 - SDRAM Access Test .......................................... PASS
Test 5 - CPU Access Test ............................................ PASS
Test 6 - I2C Temp Access Test CPU Board ............................. PASS
Test 7 - I2C Temp Access Test Main Board ............................ PASS
Test 8 - RTC Access Test ............................................ PASS
--More--
```

Trace logs

In addition to the syslog buffer, FTOS buffers trace messages which are continuously written by various FTOS software tasks to report hardware and software events and status information. Each trace message provides the date, time, and name of the FTOS process. All messages are stored in a ring buffer and can be saved to a file either manually or automatically upon failover.

Auto Save on Crash or Rollover

Exception information on 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, the **TRACE_LOG_DIR** files can be reached by FTP or by using the **show file** command from the **flash://TRACE_LOG_DIR** directory.

On a Standby unit, the **TRACE_LOG_DIR** files can be reached only by using the **show file** command from the **flash://TRACE_LOG_DIR** directory.

Note: Non-management member units do not support this functionality.

Last restart reason (S55)

If an S55 system restarted for some reason (automatically or manually), the **show system** command output includes the reason for the restart. The following table shows the reasons displayed in the output and their corresponding causes.

Table 46-1. Line card restart causes and reasons

Causes	Displayed Reasons
Remote power cycle of the chassis	push button reset
reload	soft reset
reboot after a crash	soft reset

show hardware commands (S55)

Note: The **show hardware** command tree is supported on the S55 only.

The **show hardware** command tree consists of EXEC Privilege commands used with the S55 system. These commands display information from a hardware sub-component and from hardware-based feature tables.

Table 46-2 lists the **show hardware** commands available as of the latest FTOS version on the S55.

Note: The **show hardware** commands should only be used under the guidance of Dell Force10 Technical Assistance Center.

Table 46-2. show hardware Commands

Command	Description
show hardware stack-unit *{0-11}* **cpu management statistics**	View internal interface status of the stack-unit CPU port which connects to the external management interface.
show hardware stack-unit *{0-11}* **cpu data-plane statistics**	View driver-level statistics for the data-plane port on the CPU for the specified stack-unit. It 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.
show hardware stack-unit *{0-11}* **cpu party-bus statistics**	View input and output statistics on the party bus, which carries inter-process communication traffic between CPUs.
show hardware stack-unit *{0-11}* **drops unit** *{0-1}* **port** *{0-1}*	View the ingress and egress internal packet-drop counters, MAC counters drop, and FP packet drops for the stack unit on per port basis. It assists in identifying the stack unit/port pipe/port that may experience internal drops.
show hardware stack-unit *{0-11}* **stack-port** *{0-47}*	View the input and output statistics for a stack-port interface.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **counters**	View the counters in the field processors of the stack unit.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **details**	View the details of the the FP Devices, and Hi gig ports on the stack-unit.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **execute-shell-cmd** *{command}*	Execute a specified bShell commands from the CLI without going into the bShell.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **ipmc-replication**	View the Multicast IPMC replication table from the bShell.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **port-stats [detail]**	View the internal statistics for each port-pipe (unit) on per port basis.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **register**	View the stack-unit internal registers for each port-pipe.
show hardware stack-unit *{0-11}* **unit** *{0-1}* **table-dump** *{table name}*	View the tables from the bShell through the CLI without going into the bShell.

Troubleshooting packet loss

The **show hardware stack-unit** command is intended primarily to troubleshoot packet loss.

- **show hardware stack-unit cpu data-plane statistics**
- **show hardware stack-unit cpu party-bus statistics**
- **show hardware stack-unit 0-11 drops unit 0-1 port 0-49**
- **show hardware stack-unit 0-11 stack-port 48-51**
- **show hardware stack-unit 0-11 unit 0-1 {counters | details | port-stats [detail] | register | execute-shell-cmd | ipmc-replication | table-dump}**:
- **show hardware {layer2| layer3} {e.g. acl |in acl} stack-unit 0-11 port-set 0-1**
- **show hardware layer3 qos stack-unit 0-7 port-set 0-1**
- **show hardware ipv6 {e.g.-acl |in-acl} stack-unit 0-11 port-set 0-1**
- **show hardware system-flow layer2 stack-unit 0-11 port-set 0-1 [counters]**
- **clear hardware stack-unit 0-11 counters**

- **clear hardware stack-unit 0-11 unit 0-1 counters**
- **clear hardware stack-unit 0-11 cpu data-plane statistics**
- **clear hardware stack-unit 0-11 cpu party-bus statistics**
- **clear hardware stack-unit 0-11 stack-port 48-51**

Displaying Drop Counters

The **show hardware stack-unit 0–11 drops [unit 0–1 [port 0–49]]** command assists in identifying which stack unit, port pipe, and port is experiencing internal drops, as shown in Figure 46-6 and Figure 46-7.

Figure 46-6. Displaying Drop Counter Statistics

```
Force10#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

Force10#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
```

Display drop counters with the **show hardware stack-unit drops unit port** command:

Figure 46-7. Displaying Drop Counters

```
Force10#show hardware stack-unit 0 drops unit 0 port 1
 --- Ingress Drops      ---
Ingress Drops                 : 30
IBP CBP Full Drops            : 0
PortSTPnotFwd Drops           : 0
IPv4 L3 Discards              : 0
Policy Discards               : 0
Packets dropped by FP         : 14
(L2+L3) Drops                 : 0
Port bitmap zero Drops        : 16
Rx VLAN Drops                 : 0

--- Ingress MAC counters---
Ingress FCSDrops              : 0
Ingress MTUExceeds            : 0

--- MMU Drops          ---
HOL DROPS                     : 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
```

Dataplane Statistics

The **show hardware stack-unit cpu data-plane statistics** command provides insight into the packet types coming to the CPU. As shown in Figure 46-8, the command output 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.

Figure 46-8. Displaying Dataplane Statistics

```
Force10#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
 txPkt(UNIT1)    :0
 txPkt(UNIT2)    :0
 txPkt(UNIT3)    :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, as shown in Figure 46-9.

Figure 46-9. Displaying Party Bus Statistics

```
Force10#sh 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, as shown in Figure 46-10.

Figure 46-10. Displaying Stack Unit Statistics

```
Force10#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
Force10#
```

Displaying Stack Member Counters

The **show hardware stack-unit 0–11 {counters | details | port-stats [detail] | register}** command displays internal receive and transmit statistics, based on the selected command option. A sample of the output is shown for the **counters** option in Figure 46-11.

Figure 46-11. Displaying Stack Unit Counters

```
RIPC4.ge0        :                1,202              +1,202
RUC.ge0          :                1,224              +1,217
RDBGC0.ge0       :                   34                 +24
RDBGC1.ge0       :                  366                +235
RDBGC5.ge0       :                   16                 +12
RDBGC7.ge0       :                   18                 +12
GR64.ge0         :                5,176                 +24
GR127.ge0        :                1,566              +1,433
GR255.ge0        :                    4                  +4
GRPKT.ge0        :                1,602              +1,461
GRBYT.ge0        :              117,600            +106,202
GRMCA.ge0        :                  366                +235
GRBCA.ge0        :                   12                  +9
GT64.ge0         :                    4                  +3
GT127.ge0        :                  964                +964
GT255.ge0        :                    4                  +4
GT511.ge0        :                    1                  +1
GTPKT.ge0        :                  973                +972
GTBCA.ge0        :                    1                  +1
GTBYT.ge0        :               71,531             +71,467
RUC.cpu0         :                  972                +971
TDBGC6.cpu0      :                1,584              +1,449=
```

Application core dumps

Application core dumps are *disabled* by default. A core dump file can be very large. Due to memory requirements the file can only be sent directly to an FTP server. It is not stored on the local flash. Enable full application core dumps with the following:

Task	Command Syntax	Command Mode
Enable RPM core dumps and specify the shutdown mode.	**logging coredump server**	CONFIGURATION

Undo this command using the **no logging coredump server.**

Mini core dumps

FTOS supports mini core dumps on the for application and kernel crashes. The mini core dump apply to Master, Standby and Member units.

Application and kernel mini core dumps are always enabled. The mini core dumps contain the stack space and some other very minimal information that can be used to debug a crash. These files are small files and are written into flash until space is exhausted. When the flash is full, the write process is stopped.

A mini core dump contains critical information in the event of a crash. Mini core dump files are located in **flash:/ (root dir)**. The application mini core file name format is **f10StkUnit<***Stack_unit_no***>.<***Application name***>.acore.mini.txt**. The kernel mini core file name format is **f10StkUnit<***Stack_unit_no***>.kcore.mini.txt**. Sample files names are shown in Figure 46-12 and sample file text is shown in Figure 46-13.

Figure 46-12. Mini application core file naming example

```
Force10#dir
Directory of flash:

  1  drw-      16384   Jan 01 1980 00:00:00 +00:00 .
  2  drwx       1536   Sep 03 2009 16:51:02 +00:00 ..
  3  drw-        512   Aug 07 2009 13:05:58 +00:00 TRACE_LOG_DIR
  4  d---        512   Aug 07 2009 13:06:00 +00:00 ADMIN_DIR
  5  -rw-       8693   Sep 03 2009 16:50:56 +00:00 startup-config
  6  -rw-       8693   Sep 03 2009 16:44:22 +00:00 startup-config.bak
  7  -rw-        156   Aug 28 2009 16:16:10 +00:00 f10StkUnit0.mrtm.acore.mini.txt
  8  -rw-        156   Aug 28 2009 17:17:24 +00:00 f10StkUnit0.vrrp.acore.mini.txt
  9  -rw-        156   Aug 28 2009 18:25:18 +00:00 f10StkUnit0.sysd.acore.mini.txt
 10  -rw-        156   Aug 28 2009 19:07:36 +00:00 f10StkUnit0.frrp.acore.mini.txt
 11  -rw-        156   Aug 31 2009 16:18:50 +00:00 f10StkUnit2.sysd.acore.mini.txt
 12  -rw-        156   Aug 29 2009 14:28:34 +00:00 f10StkUnit0.ipm1.acore.mini.txt
 13  -rw-        156   Aug 31 2009 16:14:56 +00:00 f10StkUnit0.acl.acore.mini.txt

flash: 3104256 bytes total (2959872 bytes free)
Force10#
```

When a member or standby unit crashes, the mini core file gets uploaded to master unit. When the master unit crashes, the mini core file is uploaded to new master.

Figure 46-13. Mini core text file example

```
                       VALID MAGIC
------------------------PANIC STRING -----------------
panic string is :<null>
----------------------STACK TRACE START---------------
0035d60c <f10_save_mmu+0x120>:
00274f8c <panic+0x144>:
0024e2b0 <db_fncall+0x134>:
0024dee8 <db_command+0x258>:
0024d9c4 <db_command_loop+0xc4>:
002522b0 <db_trap+0x158>:
0026a8d0 <mi_switch+0x1b0>:
0026a00c <bpendtsleep>:
------------------------STACK TRACE END----------------

---------------------------FREE MEMORY---------------
uvmexp.free = 0x2312
```

The panic string contains key information regarding the crash. Several panic string types exist, and they are displayed in regular english text to enable easier understanding of the crash cause.

47

Standards Compliance

This appendix contains the following sections:

- IEEE Compliance
- RFC and I-D Compliance
- MIB Location

Note: Unless noted, when a standard cited here is listed as supported by FTOS, FTOS also supports predecessor standards. One way to search for predecessor standards is to use the http://tools.ietf.org/ website. Click on "**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

- 802.1AB — LLDP
- 802.1D — Bridging, STP
- 802.1p — L2 Prioritization
- 802.1Q — VLAN Tagging, Double VLAN Tagging, GVRP
- 802.1s — MSTP
- 802.1w — RSTP
- 802.1X — Network Access Control (Port Authentication)
- 802.3ab — Gigabit Ethernet (1000BASE-T)
- 802.3ac — Frame Extensions for VLAN Tagging
- 802.3ad — Link Aggregation with LACP
- 802.3ae — 10 Gigabit Ethernet (10GBASE-W, 10GBASE-X)
- 802.3af — Power over Ethernet
- 802.3ak — 10 Gigabit Ethernet (10GBASE-CX4)
- 802.3i — Ethernet (10BASE-T)
- 802.3u — Fast Ethernet (100BASE-FX, 100BASE-TX)
- 802.3x — Flow Control
- 802.3z — Gigabit Ethernet (1000BASE-X)
- ANSI/TIA-1057— LLDP-MED
- Dell Force10 — FRRP (Dell Force10 Redundant Ring Protocol)
- Dell Force10 — PVST+
- SFF-8431 — SFP+ Direct Attach Cable (10GSFP+Cu)
- MTU — 9,252 bytes

RFC and I-D Compliance

The following standards are supported by FTOS, and are grouped by related protocol. The columns showing support by platform indicate which version of FTOS first supports the standard.

Note: Checkmarks (✓) in the E-Series column indicate that FTOS support was added before FTOS version 7.5.1.

General Internet Protocols

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
768	User Datagram Protocol	7.6.1	7.5.1	✓	8.1.1
793	Transmission Control Protocol	7.6.1	7.5.1	✓	8.1.1
854	Telnet Protocol Specification	7.6.1	7.5.1	✓	8.1.1
959	File Transfer Protocol (FTP)	7.6.1	7.5.1	✓	8.1.1
1321	The MD5 Message-Digest Algorithm	7.6.1	7.5.1	✓	8.1.1
1350	The TFTP Protocol (Revision 2)	7.6.1	7.5.1	✓	8.1.1
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	7.7.1	7.5.1	✓	8.1.1
2615	PPP over SONET/SDH			✓	
2698	A Two Rate Three Color Marker			✓	8.1.1
3164	The BSD syslog Protocol	7.6.1	7.5.1	✓	8.1.1
draft-ietf-bfd-base-03	Bidirectional Forwarding Detection		7.6.1	✓	8.1.1

General IPv4 Protocols

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
791	Internet Protocol	7.6.1	7.5.1	✓	8.1.1
792	Internet Control Message Protocol	7.6.1	7.5.1	✓	8.1.1
826	An Ethernet Address Resolution Protocol	7.6.1	7.5.1	✓	8.1.1
1027	Using ARP to Implement Transparent Subnet Gateways	7.6.1	7.5.1	✓	8.1.1
1035	DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION (client)	7.6.1	7.5.1	✓	8.1.1
1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks	7.6.1	7.5.1	✓	8.1.1
1191	Path MTU Discovery	7.6.1	7.5.1	✓	8.1.1
1305	Network Time Protocol (Version 3) Specification, Implementation and Analysis	7.6.1	7.5.1	✓	8.1.1
1519	Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy	7.6.1	7.5.1	✓	8.1.1
1542	Clarifications and Extensions for the Bootstrap Protocol	7.6.1	7.5.1	✓	8.1.1
1812	Requirements for IP Version 4 Routers	7.6.1	7.5.1	✓	8.1.1
2131	Dynamic Host Configuration Protocol	7.6.1	7.5.1	✓	8.1.1
2338	Virtual Router Redundancy Protocol (VRRP)	7.6.1	7.5.1	✓	8.1.1
3021	Using 31-Bit Prefixes on IPv4 Point-to-Point Links	7.7.1	7.7.1	7.7.1	8.1.1
3046	DHCP Relay Agent Information Option	7.8.1	7.8.1		
3069	VLAN Aggregation for Efficient IP Address Allocation	7.8.1	7.8.1		
3128	Protection Against a Variant of the Tiny Fragment Attack	7.6.1	7.5.1	✓	8.1.1

General IPv6 Protocols

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
1886	DNS Extensions to support IP version 6	7.8.1	7.8.1	✓	8.2.1
1981 (Partial)	Path MTU Discovery for IP version 6	7.8.1	7.8.1	✓	8.2.1
2460	Internet Protocol, Version 6 (IPv6) Specification	7.8.1	7.8.1	✓	8.2.1
2461 (Partial)	Neighbor Discovery for IP Version 6 (IPv6)	7.8.1	7.8.1	✓	8.2.1
2462 (Partial)	IPv6 Stateless Address Autoconfiguration	7.8.1	7.8.1	✓	8.2.1
2463	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification	7.8.1	7.8.1	✓	8.2.1
2464	Transmission of IPv6 Packets over Ethernet Networks	7.8.1	7.8.1	✓	8.2.1
2675	IPv6 Jumbograms	7.8.1	7.8.1	✓	8.2.1
3587	IPv6 Global Unicast Address Format	7.8.1	7.8.1	✓	8.2.1
4291	Internet Protocol Version 6 (IPv6) Addressing Architecture	7.8.1	7.8.1	✓	8.2.1

Border Gateway Protocol (BGP)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1997	BGP Communities Attribute	7.8.1	7.7.1	✓	8.1.1
2385	Protection of BGP Sessions via the TCP MD5 Signature Option	7.8.1	7.7.1	✓	8.1.1
2439	BGP Route Flap Damping	7.8.1	7.7.1	✓	8.1.1
2545	Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing		7.8.1	✓	8.2.1
2796	BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)	7.8.1	7.7.1	✓	8.1.1
2842	Capabilities Advertisement with BGP-4	7.8.1	7.7.1	✓	8.1.1
2858	Multiprotocol Extensions for BGP-4	7.8.1	7.7.1	✓	8.1.1
2918	Route Refresh Capability for BGP-4	7.8.1	7.7.1	✓	8.1.1
3065	Autonomous System Confederations for BGP	7.8.1	7.7.1	✓	8.1.1
4360	BGP Extended Communities Attribute	7.8.1	7.7.1	7.6.1	8.1.1
4893	BGP Support for Four-octet AS Number Space	7.8.1	7.7.1	7.7.1	8.1.1
5396	Textual Representation of Autonomous System (AS) Numbers	8.1.2	8.1.2	8.1.2	8.2.1
draft-ietf-idr-bgp4-20	A Border Gateway Protocol 4 (BGP-4)	7.8.1	7.7.1	✓	8.1.1
draft-ietf-idr-restart-06	Graceful Restart Mechanism for BGP	7.8.1	7.7.1	✓	8.1.1

Open Shortest Path First (OSPF)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1587	The OSPF Not-So-Stubby Area (NSSA) Option	7.6.1	7.5.1	✓	8.1.1
2154	OSPF with Digital Signatures	7.6.1	7.5.1	✓	8.1.1
2328	OSPF Version 2	7.6.1	7.5.1	✓	8.1.1
2370	The OSPF Opaque LSA Option	7.6.1	7.5.1	✓	8.1.1
2740	OSPF for IPv6		7.8.1	✓	8.2.1
3623	Graceful OSPF Restart	7.8.1	7.5.1	✓	8.1.1
4222	Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance	7.6.1	7.5.1	✓	8.1.1

Intermediate System to Intermediate System (IS-IS)

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
1142	OSI IS-IS Intra-Domain Routing Protocol (ISO DP 10589)			✓	8.1.1
1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments			✓	8.1.1
2763	Dynamic Hostname Exchange Mechanism for IS-IS			✓	8.1.1
2966	Domain-wide Prefix Distribution with Two-Level IS-IS			✓	8.1.1
3373	Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies			✓	8.1.1
3567	IS-IS Cryptographic Authentication			✓	8.1.1
3784	Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)			✓	8.1.1
5120	M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)			7.8.1	8.2.1
5306	Restart Signaling for IS-IS			8.3.1	8.3.1
draft-ietf-isis-igp-p2p-over-lan-06	Point-to-point operation over LAN in link-state routing protocols			✓	8.1.1
draft-ietf-isis -ipv6-06	Routing IPv6 with IS-IS			7.5.1	8.2.1
draft-kaplan-isis-ext-eth-02	Extended Ethernet Frame Size Support			✓	8.1.1

Routing Information Protocol (RIP)

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
1058	Routing Information Protocol	7.8.1	7.6.1	✓	8.1.1
2453	RIP Version 2	7.8.1	7.6.1	✓	8.1.1

Multiprotocol Label Switching (MPLS)

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
2702	Requirements for Traffic Engineering Over MPLS				8.3.1
3031	Multiprotocol Label Switching Architecture				8.3.1
3032	MPLS Label Stack Encoding				8.3.1
3209	RSVP-TE: Extensions to RSVP for LSP Tunnels				8.3.1
3630	Traffic Engineering (TE) Extensions to OSPF Version 2				8.3.1
3784	Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)				8.3.1
3812	Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)				8.3.1
3813	Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)				8.3.1
4090	Fast Reroute Extensions to RSVP-TE for LSP Tunnels				8.3.1
4379	Detecting Multi-Protocol Label Switched Data Plane Failures (MPLS TE/LDP Ping & Traceroute				8.3.1
5036	LDP Specification				8.3.1
5063	Extensions to GMPLS Resource Reservation Protocol (RSVP) Graceful Restart				8.3.1

Multicast

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1112	Host Extensions for IP Multicasting	7.8.1	7.7.1	✓	8.1.1
2236	Internet Group Management Protocol, Version 2	7.8.1	7.7.1	✓	8.1.1
2710	Multicast Listener Discovery (MLD) for IPv6			✓	8.2.1
3376	Internet Group Management Protocol, Version 3	7.8.1	7.7.1	✓	8.1.1
3569	An Overview of Source-Specific Multicast (SSM)	7.8.1 SSM for IPv4	7.7.1 SSM for IPv4	7.5.1 SSM for IPv4/ IPv6	8.2.1 SSM for IPv4
3618	Multicast Source Discovery Protocol (MSDP)			✓	8.1.1
3810	Multicast Listener Discovery Version 2 (MLDv2) for IPv6			✓	8.2.1
3973	Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)			✓	
4541	Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches	7.6.1 (IGMPv1/v2)	7.6.1 (IGMPv1/v2)	✓ IGMPv1/v2/v3, MLDv1 Snooping	8.2.1 IGMPv1/v2/ v3, MLDv1 Snooping
draft-ietf-pim-sm-v2-new-05	Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)	7.8.1 PIM-SM for IPv4	7.7.1	✓ IPv4/ IPv6	8.2.1 PIM-SM for IPv4/IPv6

Network Management

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
1155	Structure and Identification of Management Information for TCP/IP-based Internets	7.6.1	7.5.1	✓	8.1.1
1156	Management Information Base for Network Management of TCP/IP-based internets	7.6.1	7.5.1	✓	8.1.1
1157	A Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
1212	Concise MIB Definitions	7.6.1	7.5.1	✓	8.1.1
1215	A Convention for Defining Traps for use with the SNMP	7.6.1	7.5.1	✓	8.1.1
1493	Definitions of Managed Objects for Bridges [except for the dot1dTpLearnedEntryDiscards object]	7.6.1	7.5.1	✓	8.1.1
1724	RIP Version 2 MIB Extension		7.5.1	✓	8.1.1
1850	OSPF Version 2 Management Information Base	7.6.1	7.5.1	✓	8.1.1
1901	Introduction to Community-based SNMPv2	7.6.1	7.5.1	✓	8.1.1
2011	SNMPv2 Management Information Base for the Internet Protocol using SMIv2	7.6.1	7.5.1	✓	8.1.1
2012	SNMPv2 Management Information Base for the Transmission Control Protocol using SMIv2	7.6.1	7.5.1	✓	8.1.1
2013	SNMPv2 Management Information Base for the User Datagram Protocol using SMIv2	7.6.1	7.5.1	✓	8.1.1
2024	Definitions of Managed Objects for Data Link Switching using SMIv2	7.6.1	7.5.1	✓	8.1.1
2096	IP Forwarding Table MIB	7.6.1	7.5.1	✓	8.1.1
2558	Definitions of Managed Objects for the Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Interface Type			✓	
2570	Introduction and Applicability Statements for Internet Standard Management Framework	7.6.1	7.5.1	✓	8.1.1
2571	An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks	7.6.1	7.5.1	✓	8.1.1
2572	Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
2574	User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)	7.6.1	7.5.1	✓	8.1.1
2575	View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1

Network Management (continued)

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
2576	Coexistence Between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework	7.6.1	7.5.1	✓	8.1.1
2578	Structure of Management Information Version 2 (SMIv2)	7.6.1	7.5.1	✓	8.1.1
2579	Textual Conventions for SMIv2	7.6.1	7.5.1	✓	8.1.1
2580	Conformance Statements for SMIv2	7.6.1	7.5.1	✓	8.1.1
2618	RADIUS Authentication Client MIB, except the following four counters: radiusAuthClientInvalidServerAddresses radiusAuthClientMalformedAccessResponses radiusAuthClientUnknownTypes radiusAuthClientPacketsDropped	7.6.1	7.5.1	✓	8.1.1
2665	Definitions of Managed Objects for the Ethernet-like Interface Types	7.6.1	7.5.1	✓	8.1.1
2674	Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions	7.6.1	7.5.1	✓	8.1.1
2787	Definitions of Managed Objects for the Virtual Router Redundancy Protocol	7.6.1	7.5.1	✓	8.1.1
2819	Remote Network Monitoring Management Information Base: Ethernet Statistics Table, Ethernet History Control Table, Ethernet History Table, Alarm Table, Event Table, Log Table	7.6.1	7.5.1	✓	8.1.1
2863	The Interfaces Group MIB	7.6.1	7.5.1	✓	8.1.1
2865	Remote Authentication Dial In User Service (RADIUS)	7.6.1	7.5.1	✓	8.1.1
3273	Remote Network Monitoring Management Information Base for High Capacity Networks (64 bits): Ethernet Statistics High-Capacity Table, Ethernet History High-Capacity Table	7.6.1	7.5.1	✓	8.1.1
3416	Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
3418	Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
3434	Remote Monitoring MIB Extensions for High Capacity Alarms, High-Capacity Alarm Table (64 bits)	7.6.1	7.5.1	✓	8.1.1
3580	IEEE 802.1X Remote Authentication Dial In User Service (RADIUS) Usage Guidelines	7.6.1	7.5.1	✓	8.1.1

Network Management (continued)

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
3815	Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)				8.3.1
5060	Protocol Independent Multicast MIB	7.8.1	7.8.1	7.7.1	8.1.1
ANSI/TIA-1057	The LLDP Management Information Base extension module for TIA-TR41.4 Media Endpoint Discovery information	7.7.1	7.6.1	7.6.1	8.1.1
draft-grant-tacacs-02	The TACACS+ Protocol	7.6.1	7.5.1	✓	8.1.1
draft-ietf-idr-bgp4-mib-06	Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2	7.8.1	7.7.1	✓	8.1.1
draft-ietf-isis-wg-mib-16	Management Information Base for Intermediate System to Intermediate System (IS-IS): isisSysObject (top level scalar objects) isisISAdjTable isisISAdjAreaAddrTable isisISAdjIPAddrTable isisISAdjProtSuppTable			✓	8.1.1
IEEE 802.1AB	Management Information Base module for LLDP configuration, statistics, local system data and remote systems data components.	7.7.1	7.6.1	7.6.1	8.1.1
IEEE 802.1AB	The LLDP Management Information Base extension module for IEEE 802.1 organizationally defined discovery information. (LLDP DOT1 MIB and LLDP DOT3 MIB)	7.7.1	7.6.1	7.6.1	8.1.1
IEEE 802.1AB	The LLDP Management Information Base extension module for IEEE 802.3 organizationally defined discovery information. (LLDP DOT1 MIB and LLDP DOT3 MIB)	7.7.1	7.6.1	7.6.1	8.1.1
ruzin-mstp-mib-02 (Traps)	Definitions of Managed Objects for Bridges with Multiple Spanning Tree Protocol	7.6.1	7.6.1	7.6.1	8.1.1
sFlow.org	sFlow Version 5	7.7.1	7.6.1	✓	8.1.1
sFlow.org	sFlow Version 5 MIB	7.7.1	7.6.1	✓	8.1.1
FORCE10-BGP4-V2-MIB	Dell Force10 BGP MIB (draft-ietf-idr-bgp4-mibv2-05)	7.8.1	7.7.1	✓	8.1.1
FORCE10-FIB-MIB	Dell Force10 CIDR Multipath Routes MIB (The IP Forwarding Table provides information that you can use to determine the egress port of an IP packet and troubleshoot an IP reachability issue. It reports the autonomous system of the next hop, multiple next hop support, and policy routing support)			7.6.1	8.1.1

Network Management (continued)

		FTOS support, per platform			
RFC#	Full Name	S-Series	C-Series	E-Series TeraScale	E-Series ExaScale
FORCE10-CS-CHASSIS-MIB	Dell Force10 C-Series Enterprise Chassis MIB		7.5.1		
FORCE10-IF-EXTENSION-MIB	Dell Force10 Enterprise IF Extension MIB (extends the Interfaces portion of the MIB-2 (RFC 1213) by providing proprietary SNMP OIDs for other counters displayed in the “show interfaces” output)	7.6.1	7.6.1	7.6.1	8.1.1
FORCE10-LINKAGG-MIB	Dell Force10 Enterprise Link Aggregation MIB	7.6.1	7.5.1	✓	8.1.1
FORCE10-CHASSIS-MIB	Dell Force10 E-Series Enterprise Chassis MIB			✓	8.1.1
FORCE10-COPY-CONFIG-MIB	Dell Force10 File Copy MIB (supporting SNMP SET operation)	7.7.1	7.7.1	✓	8.1.1
FORCE10-MON-MIB	Dell Force10 Monitoring MIB	7.6.1	7.5.1	✓	8.1.1
FORCE10-PRODUCTS-MIB	Dell Force10 Product Object Identifier MIB	7.6.1	7.5.1	✓	8.1.1
FORCE10-SS-CHASSIS-MIB	Dell Force10 S-Series Enterprise Chassis MIB	7.6.1			
FORCE10-SMI	Dell Force10 Structure of Management Information	7.6.1	7.5.1	✓	8.1.1
FORCE10-SYSTEM-COMPONENT-MIB	Dell Force10 System Component MIB (enables the user to view CAM usage information)	7.6.1	7.5.1	✓	8.1.1
FORCE10-TC-MIB	Dell Force10 Textual Convention	7.6.1	7.5.1	✓	8.1.1
FORCE10-TRAP-ALARM-MIB	Dell Force10 Trap Alarm MIB	7.6.1	7.5.1	✓	8.1.1

MIB Location

Dell Force10 MIBs are under the **Dell 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 Force10 TAC for assistance.

Index