Troubleshooting and Monitoring on the QFX Series

Release
13.2

Published: 2014-04-01

Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, California 94089
USA
408-745-2000
www.juniper.net

Troubleshooting and Monitoring on the QFX Series
13.2

The information in this document is current as of the date on the title page.

YEAR 2000 NOTICE

Juniper Networks hardware and software products are Year 2000 compliant. Junos OS has no known time-related limitations through the year 2038. However, the NTP application is known to have some difficulty in the year 2036.

END USER LICENSE AGREEMENT

The Juniper Networks product that is the subject of this technical documentation consists of (or is intended for use with) Juniper Networks software. Use of such software is subject to the terms and conditions of the End User License Agreement ("EULA") posted at http://www.juniper.net/support/eula.html. By downloading, installing or using such software, you agree to the terms and conditions of that EULA.

Table of Contents

List of Tables

About the Documentation

Documentation and Release Notes

To obtain the most current version of all Juniper Networks® technical documentation, see the product documentation page on the Juniper Networks website at http://www.juniper.net/techpubs/.

If the information in the latest release notes differs from the information in the documentation, follow the product Release Notes.

Juniper Networks Books publishes books by Juniper Networks engineers and subject matter experts. These books go beyond the technical documentation to explore the nuances of network architecture, deployment, and administration. The current list can be viewed at http://www.juniper.net/books.

Supported Platforms

For the features described in this document, the following platforms are supported:

- QFX Series standalone switches

Using the Examples in This Manual

If you want to use the examples in this manual, you can use the **load merge** or the **load merge relative** command. These commands cause the software to merge the incoming configuration into the current candidate configuration. The example does not become active until you commit the candidate configuration.

If the example configuration contains the top level of the hierarchy (or multiple hierarchies), the example is a *full example*. In this case, use the **load merge** command.

If the example configuration does not start at the top level of the hierarchy, the example is a *snippet*. In this case, use the **load merge relative** command. These procedures are described in the following sections.

Merging a Full Example

To merge a full example, follow these steps:

1. From the HTML or PDF version of the manual, copy a configuration example into a text file, save the file with a name, and copy the file to a directory on your routing platform.

   For example, copy the following configuration to a file and name the file **ex-script.conf**. Copy the **ex-script.conf** file to the **/var/tmp** directory on your routing platform.

   ```
   system {
     scripts {
       commit {
         file ex-script.xsl;
       }
     }
   }
   interfaces {
     fxp0 {
       disable;
       unit 0 {
         family inet {
           address 10.0.0.1/24;
         }
       }
     }
   }
   ```

2. Merge the contents of the file into your routing platform configuration by issuing the **load merge** configuration mode command:

   ```
   [edit]
   user@host# load merge /var/tmp/ex-script.conf
   load complete
   ```

Merging a Snippet

To merge a snippet, follow these steps:

1. From the HTML or PDF version of the manual, copy a configuration snippet into a text file, save the file with a name, and copy the file to a directory on your routing platform.

   For example, copy the following snippet to a file and name the file **ex-script-snippet.conf**. Copy the **ex-script-snippet.conf** file to the **/var/tmp** directory on your routing platform.

   ```
   commit {
     file ex-script-snippet.xsl; }
   ```

2. Move to the hierarchy level that is relevant for this snippet by issuing the following configuration mode command:

```
[edit]
user@host# edit system scripts
[edit system scripts]
```

3. Merge the contents of the file into your routing platform configuration by issuing the **load merge relative** configuration mode command:

```
[edit system scripts]
user@host# load merge relative /var/tmp/ex-script-snippet.conf
load complete
```

For more information about the **load** command, see the *CLI User Guide*.

Documentation Conventions

Table 1 on page xi defines notice icons used in this guide.

Table 1: Notice Icons

Icon	Meaning	Description
	Informational note	Indicates important features or instructions.
	Caution	Indicates a situation that might result in loss of data or hardware damage.
	Warning	Alerts you to the risk of personal injury or death.
	Laser warning	Alerts you to the risk of personal injury from a laser.

Table 2 on page xi defines the text and syntax conventions used in this guide.

Table 2: Text and Syntax Conventions

Convention	Description	Examples
Bold text like this	Represents text that you type.	To enter configuration mode, type the **configure** command: user@host> **configure**
`Fixed-width text like this`	Represents output that appears on the terminal screen.	`user@host>` **show chassis alarms** `No alarms currently active`
Italic text like this	• Introduces or emphasizes important new terms. • Identifies guide names. • Identifies RFC and Internet draft titles.	• A policy *term* is a named structure that defines match conditions and actions. • *Junos OS CLI User Guide* • RFC 1997, *BGP Communities Attribute*

Table 2: Text and Syntax Conventions *(continued)*

Convention	Description	Examples
Italic text like this	Represents variables (options for which you substitute a value) in commands or configuration statements.	Configure the machine's domain name: [edit] root@# **set system domain-name** *domain-name*
Text like this	Represents names of configuration statements, commands, files, and directories; configuration hierarchy levels; or labels on routing platform components.	• To configure a stub area, include the **stub** statement at the **[edit protocols ospf area area-id]** hierarchy level. • The console port is labeled **CONSOLE**.
< > (angle brackets)	Encloses optional keywords or variables.	**stub <default-metric** ***metric*****>;**
| (pipe symbol)	Indicates a choice between the mutually exclusive keywords or variables on either side of the symbol. The set of choices is often enclosed in parentheses for clarity.	**broadcast | multicast** **(*string1* | *string2* | *string3*)**
# (pound sign)	Indicates a comment specified on the same line as the configuration statement to which it applies.	**rsvp { # Required for dynamic MPLS only**
[ ] (square brackets)	Encloses a variable for which you can substitute one or more values.	**community name members [** ***community-ids*** **]**

Indention and braces ({ })	Identifies a level in the configuration hierarchy.	[edit]
		routing-options {
		  static {
		  route default {
		  nexthop *address*;
		  retain;
		  }
		  }
		}

; (semicolon)	Identifies a leaf statement at a configuration hierarchy level.	
GUI Conventions		
Bold text like this	Represents graphical user interface (GUI) items you click or select.	• In the Logical Interfaces box, select **All Interfaces**. • To cancel the configuration, click **Cancel**.
> (bold right angle bracket)	Separates levels in a hierarchy of menu selections.	In the configuration editor hierarchy, select **Protocols>Ospf**.

Documentation Feedback

We encourage you to provide feedback, comments, and suggestions so that we can improve the documentation. You can send your comments to techpubs-comments@juniper.net, or fill out the documentation feedback form at

https://www.juniper.net/cgi-bin/docbugreport/. If you are using e-mail, be sure to include the following information with your comments:

- Document or topic name
- URL or page number
- Software release version (if applicable)

Requesting Technical Support

Technical product support is available through the Juniper Networks Technical Assistance Center (JTAC). If you are a customer with an active J-Care or JNASC support contract, or are covered under warranty, and need post-sales technical support, you can access our tools and resources online or open a case with JTAC.

- JTAC policies—For a complete understanding of our JTAC procedures and policies, review the *JTAC User Guide* located at http://www.juniper.net/us/en/local/pdf/resource-guides/7100059-en.pdf.
- Product warranties—For product warranty information, visit http://www.juniper.net/support/warranty/.
- JTAC hours of operation—The JTAC centers have resources available 24 hours a day, 7 days a week, 365 days a year.

Self-Help Online Tools and Resources

For quick and easy problem resolution, Juniper Networks has designed an online self-service portal called the Customer Support Center (CSC) that provides you with the following features:

- Find CSC offerings: http://www.juniper.net/customers/support/
- Search for known bugs: http://www2.juniper.net/kb/
- Find product documentation: http://www.juniper.net/techpubs/
- Find solutions and answer questions using our Knowledge Base: http://kb.juniper.net/
- Download the latest versions of software and review release notes: http://www.juniper.net/customers/csc/software/
- Search technical bulletins for relevant hardware and software notifications: http://kb.juniper.net/InfoCenter/
- Join and participate in the Juniper Networks Community Forum: http://www.juniper.net/company/communities/
- Open a case online in the CSC Case Management tool: http://www.juniper.net/cm/

To verify service entitlement by product serial number, use our Serial Number Entitlement (SNE) Tool: https://tools.juniper.net/SerialNumberEntitlementSearch/

Opening a Case with JTAC

You can open a case with JTAC on the Web or by telephone.

- Use the Case Management tool in the CSC at http://www.juniper.net/cm/.
- Call 1-888-314-JTAC (1-888-314-5822 toll-free in the USA, Canada, and Mexico).

For international or direct-dial options in countries without toll-free numbers, see http://www.juniper.net/support/requesting-support.html.

PART 1

Overview

CHAPTER 1

General Troubleshooting

Understanding Troubleshooting Resources

This topic describes some of the troubleshooting resources available for the QFX Series. These resources include tools such as the Junos OS CLI, Junos Space applications, and the Advanced Insight Scripts (AI-Scripts).

Table 3 on page 3 provides a list of some of the troubleshooting resources.

Table 3: Troubleshooting Resources on the QFX Series

Troubleshooting Resource	Description	Documentation
Chassis alarms	Chassis alarms indicate a failure on the switch or one of its components. A chassis alarm count is displayed on the LCD panel on the front of the switch.	"Chassis Alarm Messages on a QFX3500 Device" on page 10
Chassis Status LEDs and Fan Tray LEDs	A blinking amber Power, Fan, or Fan Tray LED indicates a hardware component error. A blinking amber Status LED indicates a software error.	*Chassis Status LEDs on a QFX3500 Device*
Interface alarms	A predefined alarm (red or yellow) for an interface type is triggered when an interface of that type goes down.	"Interface Alarm Messages" on page 13
System alarms	A predefined alarm is triggered by a missing rescue configuration or problem with the software license.	"Understanding Alarms" on page 9
System log messages	The system log includes details of system and user events, including errors. Specify the severity and type of system log messages you wish to view or save, and configure the output to be sent to local or remote hosts.	• *Overview of Single-Chassis System Logging Configuration* • *Junos OS System Log Configuration Statements*

Table 3: Troubleshooting Resources on the QFX Series *(continued)*

Troubleshooting Resource	Description	Documentation
Junos OS operational mode commands	Operational mode commands can be used to monitor switch performance and current activity on the network. For example, use the **traceroute monitor** command to locate points of failure in a network.	• *Monitoring System Process Information* • *Monitoring System Properties* • *traceroute monitor*
Junos OS automation scripts (event scripts)	Event scripts can be used to automate network troubleshooting and management tasks.	*Junos OS Automation Library*
Junos OS XML operational tags	XML operational tags are equivalent in function to operational mode commands in the CLI, which you can use to retrieve status information for a device.	*Junos XML API Operational Developer Reference*
NETCONF XML management protocol	The NETCONF XML management protocol defines basic operations that are equivalent to Junos OS CLI configuration mode commands. Client applications use the protocol operations to display, edit, and commit configuration statements (among other operations), just as administrators use CLI configuration mode commands such as **show**, **set**, and **commit** to perform those operations.	*NETCONF XML Management Protocol Developer Guide*
SNMP MIBs and traps	MIBs enable the monitoring of network devices from a central location. For example, use the Traceroute MIB to monitor devices remotely.	• *SNMP MIBs Support* • *SNMP Traps Support* • *Using the Traceroute MIB for Remote Monitoring Devices Running Junos OS*
AI-Scripts and Advanced Insight Manager (AIM)	AI-Scripts installed on the switch can automatically detect and monitor faults on the switch, and depending on the configuration on the AIM application, send notifications of potential problems and submit problem reports to Juniper Support Systems.	*Advanced Insight Scripts (AI-Scripts) Release Notes*
Junos Space Service Now	This application enables you to display and manage information about problem events. When problems are detected on the switch by Advanced Insight Scripts (AI-Scripts) that are installed on the switch, the data is collected and sent to Service Now for your review and action.	*Service Automation*

Table 3: Troubleshooting Resources on the QFX Series *(continued)*

Troubleshooting Resource	Description	Documentation
Junos Space Service Insight	This application helps in accelerating operational analysis and managing the exposure to known issues. You can identify devices that are nearing their End Of Life (EOL) and also discover and prevent issues that could occur in your network. The functionality of Service Insight is dependent on the information sent from Service Now.	*Service Automation*
Juniper Networks Knowledge Base	You can search in this database for Juniper Networks product information, including alerts and troubleshooting tips.	http://kb.juniper.net

Troubleshooting Overview

This topic provides a general guide to troubleshooting some typical problems you may encounter on your QFX Series product.

Table 4 on page 5 provides a list of problem categories, summary of the symptom or problem, and recommended actions with links to the troubleshooting documentation.

Table 4: Troubleshooting on the QFX Series

Problem Category	Symptom or Problem	Recommended Action
Switch hardware components	LCD panel shows a chassis alarm count.	See "Chassis Alarm Messages on a QFX3500 Device" on page 10.
	Fan tray LED is blinking amber.	See *Fan Tray LED on a QFX3500 Device*.
	Chassis status LED for the power is blinking amber.	See *Chassis Status LEDs on a QFX3500 Device*.
	Chassis status LED for the fan (on the management board) is blinking amber.	Replace the management board as soon as possible. See *Chassis Status LEDs on a QFX3500 Device*.

Table 4: Troubleshooting on the QFX Series *(continued)*

Problem Category	Symptom or Problem	Recommended Action
Port configuration	Cannot configure a port as a Gigabit Ethernet port.	Check whether the port is a valid Gigabit Ethernet port (6 through 41). See *QFX3500 Device Overview*.
	Cannot configure a port as a Fibre Channel port.	Check whether the port is a valid Fibre Channel port (0 through 5 and 42 through 47). See *QFX3500 Device Overview*.
	Cannot configure a port as a 10-Gigabit Ethernet port.	If the port is not a 40-Gbps QSFP+ interface, check whether the port is in the range of 0 through 5 or 42 through 47. If one of the ports in that block (0 through 5 or 42 through 47) is configured as a Fibre Channel port, then all ports in that block must also be configured as Fibre Channel ports. If the port is a 40-Gbps QSFP+ interface. make sure the configuration does not exceed the interface limit. Each 40-Gbps QSFP+ interface can be split into four 10-Gigabit Ethernet interfaces, but because port 0 is reserved, so you can only configure an additional fifteen 10-Gigabit Ethernet interfaces. See *QFX3500 Device Overview*.
	Cannot configure a 40-Gbps QSFP+ interface.	The 40-Gbps QSFP+ interfaces can only be used as 10-Gigabit Ethernet interfaces. Each 40-Gbps QSFP+ interface can be split into four 10-Gigabit Ethernet interfaces using a breakout cable. However, port 0 is reserved, so you can only configure an additional fifteen 10-Gigabit Ethernet interfaces. See *QFX3500 Device Overview*.
External devices (USB devices)	Upgrading software from a USB device results in an upgrade failure, and the system enters an invalid state.	Unplug the USB device and reboot the switch.
Initial device configuration	Cannot configure management Ethernet ports.	Configure the management ports from the console port. You cannot configure the management ports by directly connecting to them. NOTE: The management ports are on the front panel of the QFX3500 switch. They are labeled **C0** and **C1** on the front panel. In the CLI they are referred to as **me0** and **me1**. See *Configuring a QFX3500 Device as a Standalone Switch*.

Table 4: Troubleshooting on the QFX Series *(continued)*

Problem Category	Symptom or Problem	Recommended Action
Software upgrade and configuration	Failed software upgrade.	See "Recovering from a Failed Software Installation" on page 48.
	Active partition becomes inactive after upgrade.	
	Problem with the active configuration file.	See the following topics: • Loading a Previous Configuration File on page 31 • Reverting to the Default Factory Configuration on page 32 • Reverting to the Rescue Configuration on page 32 • Performing a Recovery Installation on a QFX Series Device on page 52
	Root password is lost or forgotten.	Recover the root password. See "Recovering the Root Password" on page 49.
Network interfaces	An aggregated Ethernet interface is down.	See "Troubleshooting an Aggregated Ethernet Interface" on page 39.
	Interface on built-in network port is down.	See "Troubleshooting Network Interfaces" on page 39.
	Interface on port in which SFP or SFP+ transceiver is installed in an SFP+ uplink module is down.	
Ethernet switching	A MAC address entry in the Ethernet switching table is not updated after the device with that MAC address has been moved from one interface to another on the switch.	See "Troubleshooting Ethernet Switching" on page 35.
Firewall filter	Firewall configuration exceeded available Ternary Content Addressable Memory (TCAM) space.	See "Troubleshooting Firewall Filter Configuration" on page 57.

CHAPTER 2

Alarms

Understanding Alarms

The QFX Series support different alarm types and severity levels. Table 5 on page 9 provides a list of alarm terms and definitions that may help you in monitoring the device.

Table 5: Alarm Terms and Definitions

Term	Definition
Alarm	Signal alerting you to conditions that might prevent normal operation. On the device, alarm indicators might include the LCD panel and LEDs on the device. The LCD panel (if present on the device) displays the chassis alarm message count. Blinking amber LEDs indicate yellow alarm conditions for chassis components.
Alarm condition	Failure event that triggers an alarm.

Alarm severity levels	Seriousness of the alarm. The level of severity can be either major (red) or minor (yellow).
	• Major (red)—Indicates a critical situation on the device that has resulted from one of the following conditions. A red alarm condition requires immediate action.
	  • One or more hardware components have failed.
	  • One or more hardware components have exceeded temperature thresholds.
	  • An alarm condition configured on an interface has triggered a critical warning.
	• Minor (yellow or amber)—Indicates a noncritical condition on the device that, if left unchecked, might cause an interruption in service or degradation in performance. A yellow alarm condition requires monitoring or maintenance. For example, a missing rescue configuration generates a yellow system alarm.

Table 5: Alarm Terms and Definitions *(continued)*

Term	Definition
Alarm types	Alarms include the following types: • Chassis alarm—Predefined alarm triggered by a physical condition on the device such as a power supply failure or excessive component temperature. • Interface alarm—Alarm you configure to alert you when an interface link is down. Applies to **ethernet**, **fibre-channel**, and **management-ethernet** interfaces. You can configure a red (major) or yellow (minor) alarm for the link-down condition, or have the condition ignored. • System alarm—Predefined alarm that might be triggered by a missing rescue configuration, failure to install a license for a licensed software feature, or high disk usage.

Related Documentation

- *Chassis Alarm Messages on a QFX3008-I Interconnect Device*
- Chassis Alarm Messages on a QFX3500 Device on page 10
- Interface Alarm Messages on page 13
- *show chassis alarms*
- *show system alarms*

Chassis Alarm Messages on a QFX3500 Device

Chassis alarms indicate a failure on the device or one of its components. Chassis alarms are preset and cannot be modified.

The chassis alarm message count is displayed on the LCD panel on the front of the device. To view the chassis alarm message text remotely, use the **show chassis lcd** CLI command.

Chassis alarms on QFX3500 devices have two severity levels:

- Major (red)—Indicates a critical situation on the device that has resulted from one of the conditions described in Table 6 on page 11. A red alarm condition requires immediate action.
- Minor (yellow or amber)—Indicates a noncritical condition on the device that, if left unchecked, might cause an interruption in service or degradation in performance. A yellow alarm condition requires monitoring or maintenance.

Table 6 on page 11 describes the chassis alarm messages on QFX3500 devices.

Table 6: QFX3500 Chassis Alarm Messages

Component	Alarm Type	CLI Message	Recommended Action
Fans	Major (red)	**Fan/Blower Absent**	The fan is missing. Install a fan.
		Fan Failure	Replace the fan and report the failure to customer support.
		Fan I2C Failure	Check the system log for one of the following messages and report the error message to customer support: • **CM ENV Monitor: Get fan speed failed**. • **CM ENV Monitor: Get fan speed failed *Fan-number* is NOT spinning @ correct speed**, where ***fan-number*** may be 1, 2, or 3.
		***fan-number* Not Spinning** Fan	Remove and check the fan for obstructions, and then reinsert the fan. If the problem persists, replace the fan.
Power Supplies	Major (red)	**PEM *pem-number* Airflow not matching Chassis Airflow**	The power supply airflow direction is the opposite of the chassis airflow direction. Replace the power supply with a power supply that supports the same airflow direction as the chassis.
		PEM *pem-number* I2C Failure	Check the system log for one of the following messages and report the error message to customer support: • **I2C Read failed for device *number***, where ***number*** may be from 123 to 125. • **PS *number*: Transitioning from online to offline**, where power supply (PS) ***number***may be 1 or 2.
		PEM *pem-number* is not supported	Indicates a power supply problem, or the power supply is not supported on the device. Report the problem to customer support.
		PEM *pem-number* Not OK	Indicates a problem with the incoming AC or outgoing DC power. Replace the power supply.

Table 6: QFX3500 Chassis Alarm Messages *(continued)*

Component	Alarm Type	CLI Message	Recommended Action
	Minor (yellow)	**PEM *pem-number* Absent**	For information only. Indicates the device was powered on with two power supplies installed, but now one is missing. The device can continue to operate with a single power supply. If you wish to remove this alarm message, reboot the device with one power supply.
		PEM *pem-number* is not powered	For information only. Check the power cord connection and reconnect it if necessary.
		PEM *pem-number* Power Supply Type Mismatch	For information only. Indicates that an AC power supply and DC power supply have been installed in the same chassis. If you wish to remove this alarm message, reboot the device with two AC power supplies or two DC power supplies.
		PEM *pem-number* Removed	For information only. Indicates the device was powered on with two power supplies installed, but one has been removed. The device can continue to operate with a single power supply. If you wish to remove this alarm message, reboot the device with one power supply.
Temperature Sensors	Major (red)	***sensor-location* Temp Sensor Fail**	Check the system log for the following message and report it to customer support: **Temp sensor *sensor-number* failed**, where *sensor-number* may range from 1 through 10.
		***sensor-location* Temp Sensor Too Hot**	Check environmental conditions and alarms on other devices. Ensure that environmental factors (such as hot air blowing around the equipment) are not affecting the temperature sensor. If the condition persists, the device may shut down.
	Minor (yellow)	***sensor-location* Temp Sensor Too Warm**	For information only. Check environmental conditions and alarms on other devices. Ensure that environmental factors (such as hot air blowing around the equipment) are not affecting the temperature sensor.

Related Documentation

- *Front Panel of a QFX3500 Device*

- *Configuring the Junos OS to Determine Conditions That Trigger Alarms on Different Interface Types*
- *alarm*

Interface Alarm Messages

Interface alarms are alarms that you configure to alert you when an interface is down.

To configure an interface link-down condition to trigger a red or yellow alarm, or to configure the link-down condition to be ignored, use the **alarm** statement at the [**edit chassis**] hierarchy level. You can specify the **ethernet**, **fibre-channel**, or **management-ethernet** interface type.

NOTE: Fibre Channel alarms are only valid on QFX3500 devices.

By default, major alarms are configured for interface link-down conditions on the control plane and management network interfaces in a QFabric system. The link-down alarms indicate that connectivity to the control plane network is down. You can configure these alarms to be ignored using the **alarm** statement at the [**edit chassis**] hierarchy level.

NOTE: If you configure a yellow alarm on the QFX3008-I Interconnect device, it will be handled as a red alarm.

Related Documentation

- Understanding Alarms on page 9

System Utilization Alarms

QFX Series devices provide system alarms that alert you when disk usage in the **/var** partition exceeds acceptable levels.

You can display the messages for these alarms by issuing the **show system alarms** operational mode command if the **/var** partition usage exceeds 75 percent. A usage level between 76 and 90 percent indicates high usage and raises a minor alarm condition, whereas a usage level above 90 percent indicates that the partition is full and raises a major alarm condition.

The following sample output from the **show system alarms** command shows system alarm messages that are displayed when disk usage is exceeded on the switch.

```
user@host> show system alarms
4 alarms currently active
Alarm time               Class  Description
2013-10-08 20:08:20 UTC  Minor  RE 0 /var partition usage is high
2013-10-08 20:08:20 UTC  Major  RE 0 /var partition is full
2013-10-08 20:08:08 UTC  Minor  FPC 1 /var partition usage is high
2013-10-08 20:08:08 UTC  Major  FPC 1 /var partition is full
```

BEST PRACTICE: We recommend that you regularly request a system file storage cleanup to optimize the performance of the switch and prevent generating system alarms.

Related Documentation

- Cleaning Up the System File Storage Space on page 33
- Understanding Alarms on page 9
- *show system alarms*

PART 2

Administration

CHAPTER 3

Routine Monitoring Using the CLI

Monitoring SNMP

There are several commands that you can access in Junos OS operational mode to monitor SNMP information. Some of the commands are:

- **show snmp health-monitor**, which displays the health monitor log and alarm information.
- **show snmp mib**, which displays information from the MIBs, such as device and system information.
- **show snmp statistics**, which displays SNMP statistics such as the number of packets, silent drops, and invalid output values.
- **show snmp rmon**, which displays the RMON alarm, event, history, and log information

The following example provides sample output from the **show snmp health-monitor** command:

```
user@switch> show snmp health-monitor
Alarm
Index  Variable description                          Value State

 32768 Health Monitor: root file system utilization
       jnxHrStoragePercentUsed.1                       58 active

 32769 Health Monitor: /config file system utilization
       jnxHrStoragePercentUsed.2                        0 active

 32770 Health Monitor: RE 0 CPU utilization
       jnxOperatingCPU.9.1.0.0                          0 active

 32773 Health Monitor: RE 0 Memory utilization
```

```
      jnxOperatingBuffer.9.1.0.0                          35 active

 32775 Health Monitor: jkernel daemon CPU utilization
      Init daemon                                          0 active
      Chassis daemon                                      50 active
      Firewall daemon                                      0 active
      Interface daemon                                     5 active
      SNMP daemon                                         11 active
      MIB2 daemon                                         42 active
      ...
```

The following example provides sample output from the **show snmp mib** command:

```
user@switch> show snmp mib walk system

sysDescr.0    = Juniper Networks, Inc. qfx3500s internet router, kernel
JUNOS 11.1-20100926.0 #0: 2010-09-26 06:17:38 UTC builder@abc.juniper.net:
/volume/build/junos/11.1/production/20100926.0/obj-xlr/bsd/sys/compile/JUNIPER-xxxxx

Build date: 2010-09-26 06:00:10 U
sysObjectID.0 = jnxProductQFX3500
sysUpTime.0   = 24444184
sysContact.0  = J Smith
sysName.0     = Lab QFX3500
sysLocation.0 = Lab
sysServices.0 = 4
```

The following example provides sample output from the **show snmp statistics** command:

```
user@switch> show snmp statistics

SNMP statistics:
  Input:
    Packets: 0, Bad versions: 0, Bad community names: 0,
    Bad community uses: 0, ASN parse errors: 0,
    Too bigs: 0, No such names: 0, Bad values: 0,
    Read onlys: 0, General errors: 0,
    Total request varbinds: 0, Total set varbinds: 0,
    Get requests: 0, Get nexts: 0, Set requests: 0,
    Get responses: 0, Traps: 0,
    Silent drops: 0, Proxy drops: 0, Commit pending drops: 0,
    Throttle drops: 0, Duplicate request drops: 0
    Output:
    Packets: 0, Too bigs: 0, No such names: 0,
    Bad values: 0, General errors: 0,
    Get requests: 0, Get nexts: 0, Set requests: 0,
    Get responses: 0, Traps: 0
```

Related Documentation

- *health-monitor*
- *show snmp mib*
- *show snmp statistics*

Tracing SNMP Activity on a Device Running Junos OS

SNMP tracing operations track activity for SNMP agents and record the information in log files. The logged error descriptions provide detailed information to help you solve problems faster.

By default, Junos OS does not trace any SNMP activity. If you include the **traceoptions** statement at the **[edit snmp]** hierarchy level, the default tracing behavior is:

- Important activities are logged in files located in the **/var/log** directory. Each log is named after the SNMP agent that generates it. Currently, the following log files are created in the **/var/log** directory when the **traceoptions** statement is used:
  - chassisd
  - craftd
  - ilmid
  - mib2d
  - rmopd
  - serviced
  - snmpd
- When a trace file named ***filename*** reaches its maximum size, it is renamed ***filename*.0**, then ***filename*.1**, and so on, until the maximum number of trace files is reached. Then the oldest trace file is overwritten. (For more information about how log files are created, see the *Junos OS System Log Messages Reference*.)
- Log files can be accessed only by the user who configured the tracing operation.

You cannot change the directory (**/var/log**) in which trace files are located. However, you can customize the other trace file settings by including the following statements at the **[edit snmp]** hierarchy level:

```
[edit snmp]
traceoptions {
   file <files number> <match regular-expression> <size size> <world-readable |
      no-world-readable>;
   flag flag;
   no-remote-trace;
}
```

These statements are described in the following sections:

- Configuring the Number and Size of SNMP Log Files on page 20
- Configuring Access to the Log File on page 20
- Configuring a Regular Expression for Lines to Be Logged on page 20
- Configuring the Trace Operations on page 20

Configuring the Number and Size of SNMP Log Files

By default, when the trace file reaches 128 kilobytes (KB) in size, it is renamed ***filename*.0**, then ***filename*.1**, and so on, until there are three trace files. Then the oldest trace file (***filename*.2**) is overwritten.

You can configure the limits on the number and size of trace files by including the following statements at the **[edit snmp traceoptions]** hierarchy level:

```
[edit snmp traceoptions]
file files number size size;
```

For example, set the maximum file size to 2 MB, and the maximum number of files to 20. When the file that receives the output of the tracing operation (***filename***) reaches 2 MB, ***filename*** is renamed ***filename*.0**, and a new file called ***filename*** is created. When the new ***filename*** reaches 2 MB, ***filename*.0** is renamed ***filename*.1** and ***filename*** is renamed ***filename*.0**. This process repeats until there are 20 trace files. Then the oldest file (***filename*.19**) is overwritten by the newest file (***filename*.0**).

The number of files can be from 2 through 1000 files. The file size of each file can be from 10 KB through 1 gigabyte (GB).

Configuring Access to the Log File

By default, log files can be accessed only by the user who configured the tracing operation.

To specify that any user can read all log files, include the **file world-readable** statement at the **[edit snmp traceoptions]** hierarchy level:

```
[edit snmp traceoptions]
file world-readable;
```

To explicitly set the default behavior, include the **file no-world-readable** statement at the **[edit snmp traceoptions]** hierarchy level:

```
[edit snmp traceoptions]
file no-world-readable;
```

Configuring a Regular Expression for Lines to Be Logged

By default, the trace operation output includes all lines relevant to the logged activities.

You can refine the output by including the **match** statement at the **[edit snmp traceoptions file *filename*]** hierarchy level and specifying a regular expression (regex) to be matched:

```
[edit snmp traceoptions]
file filename match regular-expression;
```

Configuring the Trace Operations

By default, only important activities are logged. You can specify which trace operations are to be logged by including the following **flag** statement (with one or more tracing flags) at the **[edit snmp traceoptions]** hierarchy level:

```
[edit snmp traceoptions]
flag {
```

```
    all;
    configuration;
    database;
    events;
    general;
    interface-stats;
    nonvolatile-sets;
    pdu;
    policy;
    protocol-timeouts;
    routing-socket;
    server;
    subagent;
    timer;
    varbind-error;
}
```

Table 7 on page 21 describes the meaning of the SNMP tracing flags.

Table 7: SNMP Tracing Flags

Flag	Description	Default Setting
all	Log all operations.	Off
configuration	Log reading of the configuration at the **[edit snmp]** hierarchy level.	Off
database	Log events involving storage and retrieval in the events database.	Off
events	Log important events.	Off
general	Log general events.	Off
interface-stats	Log physical and logical interface statistics.	Off
nonvolatile-set	Log nonvolatile SNMP set request handling.	Off
pdu	Log SNMP request and response packets.	Off
policy	Log policy processing.	Off
protocol-timeouts	Log SNMP response timeouts.	Off
routing-socket	Log routing socket calls.	Off
server	Log communication with processes that are generating events.	Off
subagent	Log subagent restarts.	Off
timer	Log internal timer events.	Off

Table 7: SNMP Tracing Flags *(continued)*

Flag	Description	Default Setting
varbind-error	Log variable binding errors.	Off

To display the end of the log for an agent, issue the **show log *agent*d | last** operational mode command:

```
[edit]
user@host# run show log agentd | last
```

where ***agent*** is the name of an SNMP agent.

Related Documentation

- *Configuring SNMP on a Device Running Junos OS*
- *Configuration Statements at the [edit snmp] Hierarchy Level*
- *Example: Tracing SNMP Activity*
- *Configuring SNMP*

Monitoring RMON MIB Tables

Purpose Monitor remote monitoring (RMON) alarm, event, and log tables.

Action To display the RMON tables:

```
user@switch> show snmp rmon
Alarm
Index  Variable description                          Value State

     5 monitor
       jnxOperatingCPU.9.1.0.0                          5 falling threshold

Event
Index  Type                        Last Event
    1  log and trap                2010-07-10 11:34:17 PDT
Event Index: 1
    Description: Event 1 triggered by Alarm 5, rising threshold (90) crossed,
(variable: jnxOperatingCPU.9.1.0.0, value: 100)
    Time: 2010-07-10 11:34:07 PDT
    Description: Event 1 triggered by Alarm 5, falling threshold (75) crossed,
(variable: jnxOperatingCPU.9.1.0.0, value: 5)
    Time: 2010-07-10 11:34:17 PDT
```

Meaning The display shows that an alarm has been defined to monitor jnxRmon MIB object jnxOperatingCPU, which represents the CPU utilization of the Routing Engine. The alarm is configured to generate an event that sends an SNMP trap and adds an entry to the logTable in the RMON MIB. The log table shows that two occurrences of the event have been generated—one for rising above a threshold of 90 percent, and one for falling below a threshold of 75 percent.

Related Documentation

- *Configuring RMON Alarms and Events*
- *show snmp rmon*

- *show snmp rmon history*
- *clear snmp statistics*
- *clear snmp history*

Displaying a Log File from a Single-Chassis System

To display a log file stored on a single-chassis system such as the QFX3500 switch, enter Junos OS CLI operational mode and issue the following commands:

user@switch> **show log** *log-filename*
user@switch> **file show** *log-file-pathname*

By default, the commands display the file stored on the local Routing Engine.

The following example shows the output from the **show log messages** command:

```
user@switch1> show log messages
Nov  4 11:30:01 switch1 newsyslog[2283]: logfile turned over due to size>128K
Nov  4 11:30:01 switch1 newsyslog[2283]: logfile turned over due to size>128K
Nov  4 11:30:06  switch1 chassism[952]: CM ENV Monitor: set fan speed is 65 percent
 for Fan 1
Nov  4 11:30:06  switch1 chassism[952]: CM ENV Monitor: set fan speed is 65 percent
 for Fan 2
Nov  4 11:30:06  switch1 chassism[952]: CM ENV Monitor: set fan speed is 65 percent
 for Fan 3
...
Nov  4 11:52:53  switch1 snmpd[944]: SNMPD_HEALTH_MON_INSTANCE: Health Monitor:
jroute daemon memory usage (Management
process): new instance detected (variable: sysApplElmtRunMemory.5.6.2293)
Nov  4 11:52:53  switch1 snmpd[944]: SNMPD_HEALTH_MON_INSTANCE: Health Monitor:
jroute daemon memory usage (Command-line
interface): new instance detected (variable: sysApplElmtRunMemory.5.8.2292)
...
Nov  4 12:08:30  switch1 rpdf[957]: task_connect: task BGP_100.10.10.1.6+179 addr
 10.10.1.6+179: Can't assign requested
address
Nov  4 12:08:30  switch1 rpdf[957]: bgp_connect_start: connect 10.10.1.6 (Internal
 AS 100): Can't assign requested address
Nov  4 12:10:24  switch1 mgd[2293]: UI_CMDLINE_READ_LINE: User 'jsmith', command
 'exit '
Nov  4 12:10:27  switch1 mgd[2293]: UI_DBASE_LOGOUT_EVENT: User 'jsmith' exiting
 configuration mode
Nov  4 12:10:31  switch1 mgd[2293]: UI_CMDLINE_READ_LINE: User 'jsmith', command
 'show log messages
```

The following example shows the output from the **file show** command. The file in the pathname **/var/log/processes** has been previously configured to include messages from the daemon facility.

```
user@switch1> file show /var/log/processes
Feb 22 08:58:24 switch1 snmpd[359]: SNMPD_TRAP_WARM_START: trap_generate_warm:
SNMP trap: warm start
Feb 22 20:35:07 switch1 snmpd[359]: SNMPD_THROTTLE_QUEUE_DRAINED:
trap_throttle_timer_handler: cleared all throttled traps
Feb 23 07:34:56 switch1 snmpd[359]: SNMPD_TRAP_WARM_START: trap_generate_warm:
SNMP trap: warm start
Feb 23 07:38:19 switch1 snmpd[359]: SNMPD_TRAP_COLD_START: trap_generate_cold:
```

```
SNMP trap: cold start
...
```

Related Documentation

- *Interpreting Messages Generated in Standard Format*
- *Interpreting Messages Generated in Structured-Data Format*

Monitoring System Log Messages

Purpose Display system log messages about the QFX Series. By looking through a system log file for any entries pertaining to the interface that you are interested in, you can further investigate a problem with an interface on the switch.

Action To view system log messages:

```
user@switch1> show log messages
```

Sample Output

```
Nov  4 11:30:01 switch1 newsyslog[2283]: logfile turned over due to size>128K
Nov  4 11:30:01 switch1 newsyslog[2283]: logfile turned over due to size>128K
Nov  4 11:30:06  switch1 chassism[952]: CM ENV Monitor: set fan speed is 65 percent
 for Fan 1
Nov  4 11:30:06  switch1 chassism[952]: CM ENV Monitor: set fan speed is 65 percent
 for Fan 2
Nov  4 11:30:06  switch1 chassism[952]: CM ENV Monitor: set fan speed is 65 percent
 for Fan 3
...
Nov  4 11:52:53  switch1 snmpd[944]: SNMPD_HEALTH_MON_INSTANCE: Health Monitor:
jroute daemon
memory usage (Management process): new instance detected (variable:
sysApplElmtRunMemory.5.6.2293)
Nov  4 11:52:53  switch1 snmpd[944]: SNMPD_HEALTH_MON_INSTANCE: Health Monitor:
jroute daemon
memory usage (Command-line interface): new instance detected (variable:
sysApplElmtRunMemory.5.8.2292)
...
Nov  4 12:10:24  switch1 mgd[2293]: UI_CMDLINE_READ_LINE: User 'jsmith', command
 'exit '
Nov  4 12:10:27  switch1 mgd[2293]: UI_DBASE_LOGOUT_EVENT: User 'jsmith' exiting
 configuration mode
Nov  4 12:10:31  switch1 mgd[2293]: UI_CMDLINE_READ_LINE: User 'jsmith', command
 'show log messages
```

Meaning The sample output shows the following entries in the **messages** file:

- A new log file was created when the previous file reached the maximum size of 128 kilobytes (KB).
- The fan speed for Fan 1, 2, and 3 is set at 65 percent.
- Health monitoring activity is detected.
- CLI commands were entered by the user jsmith.

Related Documentation

- *Overview of Junos OS System Log Messages*
- *Understanding the Implementation of System Log Messages on the QFabric System*
- *Example: Configuring System Log Messages*
- *clear log*
- *show log*
- *syslog*

Monitoring Traffic Through the Router or Switch

To help with the diagnosis of a problem, display real-time statistics about the traffic passing through physical interfaces on the router or switch.

To display real-time statistics about physical interfaces, perform these tasks:

1. Displaying Real-Time Statistics About All Interfaces on the Router or Switch on page 25
2. Displaying Real-Time Statistics About an Interface on the Router or Switch on page 26

Displaying Real-Time Statistics About All Interfaces on the Router or Switch

Purpose Display real-time statistics about traffic passing through all interfaces on the router or switch.

Action To display real-time statistics about traffic passing through all interfaces on the router or switch:

```
user@host> monitor interface traffic
```

Sample Output

```
user@host> monitor interface traffic
host name                  Seconds: 15                 Time: 12:31:09
Interface    Link  Input packets        (pps)     Output packets        (pps)
 so-1/0/0    Down              0          (0)                  0          (0)
 so-1/1/0    Down              0          (0)                  0          (0)
 so-1/1/1    Down              0          (0)                  0          (0)
 so-1/1/2    Down              0          (0)                  0          (0)
 so-1/1/3    Down              0          (0)                  0          (0)
 t3-1/2/0    Down              0          (0)                  0          (0)
 t3-1/2/1    Down              0          (0)                  0          (0)
 t3-1/2/2    Down              0          (0)                  0          (0)
 t3-1/2/3    Down              0          (0)                  0          (0)
 so-2/0/0      Up         211035          (1)              36778          (0)
 so-2/0/1      Up         192753          (1)              36782          (0)
 so-2/0/2      Up         211020          (1)              36779          (0)
 so-2/0/3      Up         211029          (1)              36776          (0)
 so-2/1/0      Up         189378          (1)              36349          (0)
 so-2/1/1    Down              0          (0)              18747          (0)
 so-2/1/2    Down              0          (0)              16078          (0)
 so-2/1/3      Up              0          (0)              80338          (0)
 at-2/3/0      Up              0          (0)                  0          (0)
 at-2/3/1    Down              0          (0)                  0          (0)
Bytes=b, Clear=c, Delta=d, Packets=p, Quit=q or ESC, Rate=r, Up=^U, Down=^D
```

Meaning The sample output displays traffic data for active interfaces and the amount that each field has changed since the command started or since the counters were cleared by using the **C** key. In this example, the **monitor interface** command has been running for 15 seconds since the command was issued or since the counters last returned to zero.

Displaying Real-Time Statistics About an Interface on the Router or Switch

Purpose Display real-time statistics about traffic passing through an interface on the router or switch.

Action To display traffic passing through an interface on the router or switch, use the following Junos OS CLI operational mode command:

user@host> **monitor interface** ***interface-name***

Sample Output

```
user@host> monitor interface so-0/0/1
Next='n', Quit='q' or ESC, Freeze='f', Thaw='t', Clear='c', Interface='i'
R1
Interface: so-0/0/1, Enabled, Link is Up
Encapsulation: PPP, Keepalives, Speed: OC3 Traffic statistics:
  Input bytes:                  5856541 (88 bps)
  Output bytes:                 6271468 (96 bps)
  Input packets:                 157629 (0 pps)
  Output packets:                157024 (0 pps)
Encapsulation statistics:
  Input keepalives:               42353
  Output keepalives:              42320
  LCP state: Opened
Error statistics:
  Input errors:                       0
  Input drops:                        0
  Input framing errors:               0
  Input runts:                        0
  Input giants:                       0
  Policed discards:                   0
  L3 incompletes:                     0
  L2 channel errors:                  0
  L2 mismatch timeouts:               0
  Carrier transitions:                1
  Output errors:                      0
  Output drops:                       0
  Aged packets:                       0
Active alarms : None
Active defects: None
SONET error counts/seconds:
  LOS count                           1
  LOF count                           1
  SEF count                           1
  ES-S                               77
  SES-S                              77
SONET statistics:
  BIP-B1                              0
  BIP-B2                              0
  REI-L                               0
  BIP-B3                              0
```



```
  REI-P                                 0
Received SONET overhead:  F1        : 0x00  J0        : 0xZ
```

Meaning The sample output shows the input and output packets for a particular SONET interface (**so-0/0/1**). The information can include common interface failures, such as SONET/SDH and T3 alarms, loopbacks detected, and increases in framing errors. For more information, see *Checklist for Tracking Error Conditions*.

To control the output of the command while it is running, use the keys shown in Table 8 on page 27.

Table 8: Output Control Keys for the monitor interface Command

Action	Key
Display information about the next interface. The **monitor interface** command scrolls through the physical or logical interfaces in the same order that they are displayed by the **show interfaces terse** command.	N
Display information about a different interface. The command prompts you for the name of a specific interface.	I
Freeze the display, halting the display of updated statistics.	F
Thaw the display, resuming the display of updated statistics.	T
Clear (zero) the current delta counters since **monitor interface** was started. It does not clear the accumulative counter.	C
Stop the **monitor interface** command.	Q

See the CLI Explorer for details on using match conditions with the **monitor traffic** command.

Pinging Hosts

Purpose Use the CLI **ping** command to verify that a host can be reached over the network. This command is useful for diagnosing host and network connectivity problems. The switch sends a series of Internet Control Message Protocol (ICMP) echo (ping) requests to a specified host and receives ICMP echo responses.

Action To use the **ping** command to send four requests (ping count) to host3:

ping *host* count *number*

Sample Output

ping host3 count 4

```
user@switch> ping host3 count 4
PING host3.site.net (176.26.232.111): 56 data bytes
64 bytes from 176.26.232.111: icmp_seq=0 ttl=122 time=0.661 ms
64 bytes from 176.26.232.111: icmp_seq=1 ttl=122 time=0.619 ms
64 bytes from 176.26.232.111: icmp_seq=2 ttl=122 time=0.621 ms
64 bytes from 176.26.232.111: icmp_seq=3 ttl=122 time=0.634 ms
```

```
--- host3.site.net ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
round-trip min/avg/max/stddev = 0.619/0.634/0.661/0.017 ms
```

Meaning

- The **ping** results show the following information:
  - Size of the ping response packet (in bytes).
  - IP address of the host from which the response was sent.
  - Sequence number of the ping response packet. You can use this value to match the ping response to the corresponding ping request.
  - Time-to-live (ttl) hop-count value of the ping response packet.
  - Total time between the sending of the ping request packet and the receiving of the ping response packet, in milliseconds. This value is also called round-trip time.
  - Number of ping requests (probes) sent to the host.
  - Number of ping responses received from the host.
  - Packet loss percentage.
  - Round-trip time statistics: minimum, average, maximum, and standard deviation of the round-trip time.

Related Documentation

- Troubleshooting Overview on page 5
- Understanding Troubleshooting Resources on page 3

PART 3

Troubleshooting

CHAPTER 4

Configuration and File Management

Loading a Previous Configuration File

You can use the **rollback** <***number*** command to return to a previously committed configuration file. A switch saves the last 50 committed configurations, including the rollback number, date, time, and name of the user who issued the **commit** configuration command.

Syntax

rollback <***number***>

Options

- **none**— Return to the most recently saved configuration.
- ***number***—Configuration to return to.
  - **Range**: 0 through 49. The most recently saved configuration is number 0, and the oldest saved configuration is number 49.
  - **Default**: 0

To return to a configuration prior to the most recently committed one:

1. Specify the rollback number (here, 1 is entered and the configuration returns to the previously committed configuration):

   ```
   [edit]
   user@switch# rollback 1
   load complete
   ```

2. Activate the configuration you have loaded:

   ```
   [edit]
   user@switch# commit
   ```

Related Documentation

- *Configuration File Terms*

Reverting to the Default Factory Configuration

If for any reason the current active configuration fails, you can revert to the default factory configuration. The default factory configuration contains the basic configuration settings. This is the first configuration of the switch, and it is loaded when the switch is first installed and powered on.

The **load factory default** command is a standard Junos OS configuration command. This configuration command replaces the current active configuration with the default factory configuration.

To revert the switch to the rescue configuration:

1.
```
[edit]
user@switch# load factory-default
[edit]
user@switch# delete system commit factory-settings
[edit]
user@switch# commit
```

Related Documentation

- *Understanding Configuration Files*
- Loading a Previous Configuration File on page 31
- Reverting to the Rescue Configuration on page 32

Reverting to the Rescue Configuration

If someone inadvertently commits a configuration that denies management access to a QFX Series product and the console port is not accessible, you can overwrite the invalid configuration and replace it with the rescue configuration. The rescue configuration is a previously committed, valid configuration.

To revert the switch to the rescue configuration:

1. Enter the **load override** command.

```
[edit]
user@switch# load override filename
```

2. Commit your changes.

```
[edit]
user@switch# commit filename
```

Related Documentation

- *Setting or Deleting the Rescue Configuration*
- Reverting to the Default Factory Configuration on page 32
- *Configuration File Terms*

Cleaning Up the System File Storage Space

Problem The system file storage space on the switch is full. Rebooting the switch does not solve the problem.

The following error message is displayed during a typical operation on the switch after the file storage space is full.

```
user@switch% cli
user@switch> configure
/var: write failed, filesystem is full
```

Solution Clean up the file storage on the switch by deleting system files.

1. Request to delete system files on the switch.

   ```
   user@switch> request system storage cleanup
   ```

   The list of files to be deleted is displayed.

   ```
   List of files to delete:

            Size Date         Name
      11B Jul 26 20:55 /var/jail/tmp/alarmd.ts
     124B Aug  4 18:05 /var/log/default-log-messages.0.gz
    1301B Jul 26 20:42 /var/log/install.0.gz
     387B Jun  3 14:37 /var/log/install.1.gz
    4920B Aug  4 18:05 /var/log/messages.0.gz
    20.0K Jul 26 21:00 /var/log/messages.1.gz
    16.3K Jun 25 13:45 /var/log/messages.2.gz
     804B Aug  4 18:05 /var/log/security.0.gz
    16.8K Aug  3 11:15 /var/log/security.1.gz
     487B Aug  4 18:04 /var/log/wtmp.0.gz
     855B Jul 29 22:54 /var/log/wtmp.1.gz
     920B Jun 30 16:32 /var/log/wtmp.2.gz
      94B Jun  3 14:36 /var/log/wtmp.3.gz
   353.2K Jun  3 14:37 /var/sw/pkg/jloader-qfx-11.2I20110303_1117_dc-builder.tgz

   124.0K Jun  3 14:30 /var/tmp/gres-tp/env.dat
       0B Apr 14 16:20 /var/tmp/gres-tp/lock
       0B Apr 14 17:37 /var/tmp/if-rtsdb/env.lck
    12.0K Jul 26 20:55 /var/tmp/if-rtsdb/env.mem
  2688.0K Jul 26 20:55 /var/tmp/if-rtsdb/shm_usr1.mem
   132.0K Jul 26 20:55 /var/tmp/if-rtsdb/shm_usr2.mem
  2048.0K Jul 26 20:55 /var/tmp/if-rtsdb/trace.mem
     155B Jul 26 20:55 /var/tmp/krt_gencfg_filter.txt
       0B Jul 26 20:55 /var/tmp/rtsdb/if-rtsdb
  1400.6K Aug  3 10:13 /var/tmp/sfid.core.0.gz
  1398.9K Aug  3 17:01 /var/tmp/sfid.core.1.gz
  Delete these files ? [yes,no] (no)
   ```

2. Enter **yes** to delete the files.

3. Reboot the switch.

BEST PRACTICE: We recommend that you regularly request a system file storage cleanup to optimize the performance of the switch.

Related Documentation

- *request system storage cleanup*

CHAPTER 5

Ethernet Switching

Troubleshooting Ethernet Switching

Problem Sometimes a MAC address entry in the switch's Ethernet switching table is not updated after the device with that MAC address has been moved from one interface to another on the switch. Typically, the switch does not wait for a MAC address expiration when a MAC move operation occurs. As soon as the switch detects the MAC address on the new interface, it immediately updates the table. Many network devices send a gratuitous ARP packet when switching an IP address from one device to another. The switch updates its ARP cache table after receipt of such gratuitous ARP messages, and then it also updates its Ethernet switching table.

Sometimes silent devices, such as syslog servers or SNMP trap receivers that receive UDP traffic but do not return acknowledgment (ACK) messages to the traffic source, fail to send gratuitous ARP packets when a device moves. If such a move occurs when the system administrator is not available to explicitly clear the affected interfaces by issuing the **clear ethernet-switching table** command, the entry for the moved device in the Ethernet switching table is not updated.

Solution Set up the switch to handle unattended MAC address switchovers.

1. Reduce the system-wide ARP aging timer. (By default, the ARP aging timer is set at 20 minutes. The range of the ARP aging timer is from 1 through 240 minutes.)

   ```
   [edit system arp]
   user@switch# set aging-timer 3
   ```

2. Set the MAC aging timer to the same value as the ARP timer. (By default, the MAC aging timer is set to 300 seconds. The range is 15 to 1,000,000 seconds.)

   ```
   [edit vlans]
   user@switch# set vlans sales mac-table-aging-time 180
   ```

The ARP entry and the MAC address entry for the moved device expire within the times specified by the aging timer values. After the entries expire, the switch sends a new ARP message to the IP address of the device. The device responds to the ARP message, thereby refreshing the entries in the switch's ARP cache table and Ethernet switching table.

Related Documentation

- *arp*
- *mac-table-aging-time*

CHAPTER 6

High Availability

Troubleshooting VRRP

Problem If you configure multiple VRRP groups on an interface (using multiple VLANs), traffic for some of the groups might be briefly dropped if a failover occurs. This can happen because the new master must send gratuitous ARP replies for each VRRP group to update the ARP tables in the connected devices, and there is a short delay between each gratuitous ARP reply. Traffic sent by devices that have not yet received the gratuitous ARP reply is dropped (until the device receives the reply and learns the MAC address of the new master).

Solution Configure a failover delay so that the new master delays sending gratuitous ARP replies for the period that you set. This allows the new master to send the ARP replies for all of the VRRP groups simultaneously.

Related Documentation

- *failover-delay*

CHAPTER 7

Interfaces

Troubleshooting an Aggregated Ethernet Interface

Problem The **show interfaces terse** command shows that the LAG is down.

Solution Check the following:

- Verify that there is no configuration mismatch.
- Verify that all member ports are up.
- Verify that a LAG is part of family ethernet-switching (Layer 2 LAG) or family inet (Layer 3 LAG).
- Verify that the LAG member is connected to the correct LAG at the other end.
- Verify that the LAG members belong to the same switch.

Related Documentation

- *Verifying the Status of a LAG Interface*
- *Example: Configuring Link Aggregation Between a QFX Series Product and an Aggregation Switch*

Troubleshooting Network Interfaces

The interface on the port in which an SFP or SFP+ transceiver is installed in an SFP or SFP+ module is down

Problem The QFX Series has an SFP or SFP+ module installed. The interface on the port in which an SFP or SFP+ transceiver is installed is down.

When you check the status with the CLI command **show interfaces *interface-name*** , the disabled port is not listed.

Cause By default, the SFP or SFP+ module operates in the 10-Gigabit Ethernet mode and supports only SFP or SFP+ transceivers. The operating mode for the module is incorrectly set.

Solution Only SFP or SFP+ transceivers can be installed in SFP or SFP+ modules. You must configure the operating mode of the SFP or SFP+ module to match the type of transceiver you want to use. For SFP+ transceivers, configure 10-Gigabit Ethernet operating mode.

Troubleshooting Multichassis Link Aggregation

Use the following information to troubleshoot multichassis link aggregation configuration.

MAC Addresses Learned on MC-AE Interfaces Are Not Removed from the MAC Address Table

Problem When both of the mulitchassis aggregated Ethernet (MC-AE) interfaces on both connected multichassis link aggregation group (MC-LAG) peers are down, the MAC addresses learned on the MC-AE interfaces are not removed from the MAC address table.

For example, if you disable the MC-AE interface (ae0) on both MC-LAG peers by issuing the **set interfaces ae0 disable** command and commit the configuration, the MAC table still shows the MAC addresses as being learned on the MC-AE interfaces of both MC-LAG peers:

```
user@switchA> show ethernet-switching table
Ethernet-switching table: 6 entries, 2 learned, 0 persistent entries
  VLAN             MAC address       Type         Age Interfaces
  v10               *                 Flood          - All-members
  v10               00:10:94:00:00:01 Learn(L)     3:55 ae0.0 (MCAE)
  v10               00:10:94:00:00:02 Learn(R)        0 xe-0/0/9.0
  v20               *                 Flood          - All-members
  v30               *                 Flood          - All-members
  v30               84:18:88:de:b1:2e Static         - Router


user@switchB> show ethernet-switching table
Ethernet-switching table: 6 entries, 2 learned, 0 persistent entries
  VLAN             MAC address       Type         Age Interfaces
  v10               *                 Flood          - All-members
  v10               00:10:94:00:00:01 Learn(R)        0 ae0.0 (MCAE)
  v10               00:10:94:00:00:02 Learn         40 xe-0/0/10.0
  v20               *                 Flood          - All-members
  v30               *                 Flood          - All-members
  v30               84:18:88:df:83:0a Static         - Router
```

Solution This is expected behavior.

MC-LAG Peer Does Not Go into Standby Mode

Problem A multichassis link aggregation group (MC-LAG) peer does not go into standby mode if the MC-LAG peer IP address specified in the Interchassis Control Protocol (ICCP) configuration and the IP address specified in the multichassis protection configuration are different.

Solution To prevent failure to enter standby mode, make sure the peer IP address in the ICCP configurations and the IP address in multichassis protection configurations are the same.

Secondary MC-LAG Peer with Status Control Set to Standby Becomes Inactive

Problem When the interchassis control link-protection link (ICL-PL) and multichassis aggregated Ethernet (MC-AE) interfaces go down on the primary multichassis link aggregation group (MC-LAG) peer, the secondary MC-LAG peer's MC-AE interfaces with status control set to standby become inactive instead of active.

Solution This is expected behavior.

Redirect Filters Take Priority over User-Defined Filters

Problem Multichassis link aggregation group (MC-LAG) implicit failover redirection filters take precedence over user-configured explicit filters. This is expected behavior.

Solution This is expected behavior.

Operational Command Output Is Wrong

Problem After you deactivate the Interchassis Control Protocol (ICCP), the **show iccp** operational command output still shows registered client daemons, such as mcsnoopd, lacpd, and eswd.

For example:

```
user@switch> show iccp
Client Application: MCSNOOPD
  Redundancy Group IDs Joined: None

Client Application: lacpd
  Redundancy Group IDs Joined: 1

Client Application: eswd
  Redundancy Group IDs Joined: 1
```

The **show iccp** command output always shows registered modules regardless of whether or not ICCP peers are configured.

Solution This is expected behavior.

ICCP Connection Might Take Up to 60 Seconds to Become Active

Problem When the Interchassis Control Protocol (ICCP) configuration and the routed VLAN interface (RVI) configuration are committed together, the ICCP connection might take up to 60 seconds to become active.

Solution This is expected behavior.

MAC Address Age Learned on an MC-AE Interface Is Reset to Zero

Problem When you activate and then deactivate an interchassis control link-protection link (ICL-PL), the MAC address age learned on the multichassis aggregated Ethernet (MC-AE) interface is reset to zero. The next-hop interface changes trigger MAC address updates in the hardware, which then triggers aging updates in the Packet Forwarding Engine (PFE). The result is that the MAC address age is updated to zero.

For example, the ICL-PL has been deactivated, and the **show ethernet-switching table** command output shows that the MAC addresses have an age of 0.

```
user@switch> show ethernet-switching table
Ethernet-switching table: 3 entries, 2 learned, 0 persistent entries
  VLAN              MAC address       Type         Age Interfaces
  v100              *                 Flood          - All-members
  v100              00:10:00:00:00:01 Learn(L)     0 ae0.0 (MCAE)
  v100              00:10:00:00:00:02 Learn(L)     0 ae0.0 (MCAE)
```

Solution This is expected behavior.

MAC Address Is Not Learned Remotely in a Default VLAN

Problem If a multichassis link aggregation group (MC-LAG) peer learns a MAC address in the default VLAN, the Interchassis Control Protocol (ICCP) does not synchronize the MAC address with the MAC address of the other MC-LAG peer.

Solution This is expected behavior.

Snooping Entries Learned on MC-AE Interfaces Are Not Removed

Problem When multichassis aggregated Ethernet (MC-AE) interfaces are configured on a VLAN that is enabled for multicast snooping, the membership entries learned on the MC-AE interfaces on the VLAN are not cleared when the MC-AE interfaces go down. This is done to speed up convergence time when the interfaces come up, or come up and go down.

Solution This is expected behavior.

ICCP Does Not Come Up After You Add or Delete an Authentication Key

Problem The Interchassis Control Protocol (ICCP) connection is not established when you add an authentication key and then delete it only at the global ICCP level. However, authentication works correctly at the ICCP peer level.

Solution Delete the ICCP configuration , and then add the ICCP configuration.

Local Status Is Standby When It Should Be Active

Problem If the multichassis aggregated Ethernet (MC-AE) interface is down when the state machine is in a synchronized state, the multichassis link aggregation group (MC-LAG) peer local status is standby. If the MC-AE interface goes down after the state machine is in an active state, then the local status remains active, and the local state indicates that the interface is down.

Solution This is expected behavior.

Packets Loop on the Server When ICCP Fails

Problem When you enable backup liveness detection for a multichassis link aggregation group (MC-LAG), and the backup liveness detection packets are lost because of a temporary failure on the MC-LAG, then both of the peers in the MC-LAG remain active. If this happens, both of the MC-LAG peers send packets to the connected server.

Solution This is expected behavior.

Both MC-LAG Peers Use the Default System ID After a Reboot or an ICCP Configuration Change

Problem After a reboot or after a new Interchassis Control Protocol (ICCP) configuration has been committed, and the ICCP connection does not become active, the Link Aggregation

Control Protocol (LACP) messages transmitted over the multichassis aggregated Ethernet (MC-AE) interfaces use the default system ID. The configured system ID is used instead of the default system ID only after the MC-LAG peers synchronize with each other.

Solution This is expected behavior.

No Commit Checks Are Done for ICL-PL Interfaces

Problem There are no commit checks on the interface being configured as an interchassis control link-protection link (ICL-PL), so you must provide a valid interface name for the ICL-PL.

Solution This is expected behavior.

Double Failover Scenario

Problem If the following events happen in this exact order—the Interchassis Control Protocol (ICCP) goes down, and the multichassis aggregated Ethernet (MC-AE) interface on the multichassis link aggregation group (MC-LAG) peer in active mode goes down—a double failover occurs. In this scenario, the MC-LAG peer in standby mode does not detect what happens on the active MC-LAG peer. The MC-LAG peer in standby mode operates as if the MC-AE interface on the MC-LAG in active mode were up and blocks the interchassis control protocol-protection link (ICL-PL) traffic. The ICL-PL traffic is not forwarded.

Solution This is expected behavior.

Multicast Traffic Floods the VLAN When the ICL-PL Interface Goes Down and Up

Problem When the interchassis control link-protection link (ICL-PL) goes down and up, multicast traffic is flooded to all of the interfaces in the VLAN. The Packet Forwarding Engine (PFE) flag Ip4McastFloodMode for the VLAN is changed to MCAST_FLOOD_ALL. This problem only occurs when a multichassis link aggregation group (MC-LAG) is configured for Layer 2.

Solution This is expected behavior.

Layer 3 Traffic Sent to the Standby MC-LAG Peer Is Not Redirected to Active MC-LAG Peer

Problem When the Interchassis Control Protocol (ICCP) is down, the status of a remote MC-LAG peer is unknown. Even if the MC-LAG peer is configured as standby, the traffic is not redirected to this peer because it is assumed that this peer is down.

Solution This is expected behavior.

AE Interfaces Go Down

Problem When a multichassis aggregated Ethernet (MC-AE) interface is converted to an aggregated Ethernet (AE) interface, it retains some MC-AE properties. For example, the

AE interface might retain the administrative key of the MC-AE. When this happens, the AE interface goes down.

Solution Restart the Link Aggregation Control Protocol (LACP) on the multichassis link aggregation group (MC-LAG) peer hosting the AE interface to bring up the AE interface. Restarting LACP removes the MC-AE properties of the AE interface.

Flooding of Upstream Traffic

Problem When MAC synchronization is enabled, the multichassis link aggregation group (MC-LAG) peer can resolve Address Resolution Protocol (ARP) entries for the MC-LAG routed VLAN interface (RVI) with either of the MC-LAG peer MAC addresses. If the downstream traffic is sent with one MAC address (MAC1) but the peer has resolved the MAC address with a different MAC address (MAC2), the MAC2 address might not be learned by any of the access layer switches. Flooding of the upstream traffic for the MAC2 address might then occur.

Solution Make sure that downstream traffic is sent from the MC-LAG peers periodically to prevent the MAC addresses from aging out.

Related Documentation

- *Understanding Multichassis Link Aggregation*
- *Example: Configuring Multichassis Link Aggregation*
- *Configuring Multichassis Link Aggregation*

CHAPTER 8

Junos OS Basics

Rebooting and Halting a QFX Series Product

To reboot the switch, issue the **request system reboot** command.

```
user@switch> request system reboot ?
Possible completions:
  <[Enter]>            Execute this command
  at                   Time at which to perform the operation
  in                   Number of minutes to delay before operation
  media                Boot media for next boot
  message              Message to display to all users
  |                    Pipe through a command
```

```
user@switch> request system reboot
Reboot the system ? [yes,no] (no) yes
Rebooting switch
```

Similarly, to halt the switch, issue the **request system halt** command.

CAUTION: Before entering this command, you must have access to the switch's console port in order to bring up the Routing Engine.

```
user@switch> request system halt ?
Possible completions:
  <[Enter]>            Execute this command
  at                   Time at which to perform the operation
  in                   Number of minutes to delay before operation
  media                Boot media for next boot
  message              Message to display to all users
  |                    Pipe through a command
```

NOTE: When you issue this command on an individual component in a QFabric system, you will receive a warning that says "Hardware-based members will halt, Virtual Junos Routing Engines will reboot." If you want to halt only one member, use the member option. You cannot issue this command from the QFabric CLI.

Issuing the **request system halt** command on the switch halts the Routing Engine. To reboot a Routing Engine that has been halted, you must connect through the console.

Related Documentation

- *clear system reboot*
- *request system reboot*
- *request system halt*
- *request system power-off*
- *Connecting a QFX Series Device to a Management Console*

Recovering from a Failed Software Installation

Problem If the Junos OS appears to have been installed but the CLI does not work, or if the switch has no software installed, you can use this recovery installation procedure to install the Junos OS.

Solution If a Junos OS image already exists on the switch, you can either install the new Junos OS package in a separate partition, in which case both Junos OS images remain on the switch, or you can remove the existing Junos OS image before you start the new installation process.

To perform a recovery installation:

1. Power on the switch. The loader script starts.
2. After the message **Loading /boot/defaults/loader.conf** appears, you are prompted with the following message:

   Hit [Enter] to boot immediately, or space bar for command prompt.

   Press the Spacebar to enter the manual loader. The **loader>** prompt appears.
3. Enter the following command:

   loader> **install [– –format] [– –external]** ***source***

   where:

   - **format**—Enables you to erase the installation media before installing the installation package. If you do not include this option, the system installs the new Junos OS in a different partition from that of the most recently installed Junos OS.
   - **external**—Installs the installation package onto external media (a USB stick, for example).

- ***source***—Represents the name and location of the Junos OS package, either on a server on the network or as a file on an external media, as shown in the following two examples:
  - Network address of the server and the path on the server; for example, **tftp://192.17.1.28/junos/jinstall-qfx-11.1R1.5-domestic-signed.tgz**
  - Junos OS package on a USB device (commonly stored in the root drive as the only file), for example, **file:///jinstall-qfx-11.1R1.5-domestic-signed.tgz**).

The installation now proceeds normally and ends with a login prompt.

Recovering the Root Password

If you forget the root password for the QFX3500 switch, you can use the password recovery procedure to reset the root password.

NOTE: **The root password cannot be recovered on a QFabric system.**

NOTE: **You need console access to the switch to recover the root password.**

To recover the root password:

1. Power off the switch by switching off the AC power outlet of the device or, if necessary, by pulling the power cords out of the QFX3500 switch power supplies.
2. Turn off the power to the management device, such as a PC or laptop computer, that you want to use to access the CLI.
3. Plug one end of the Ethernet rollover cable supplied with the switch into the RJ-45–to–DB-9 serial port adapter supplied with the switch.
4. Plug the RJ-45–to–DB-9 serial port adapter into the serial port on the management device.
5. Connect the other end of the Ethernet rollover cable to the console port on the switch.
6. Turn on the power to the management device.
7. On the management device, start your asynchronous terminal emulation application (such as Microsoft Windows Hyperterminal) and select the appropriate **COM** port to use (for example, **COM1**).
8. Configure the port settings as follows:
   - Bits per second: 9600
   - Data bits: 8
   - Parity: None

- Stop bits: 1
- Flow control: None

9. Power on the switch by (if necessary) plugging the power cords into the QFX3500 switch power supply, or turning on the power to the device or switch by switching on the AC power outlet the device is plugged into

   The terminal emulation screen on your management device displays the switch's boot sequence.

10. When the following prompt appears, press the Spacebar to access the switch's bootstrap loader command prompt:

    ```
    Hit [Enter] to boot immediately, or space bar for command prompt.
    Booting [kernel] in 9 seconds...
    ```

11. At the following prompt, enter **boot -s** to start up the system in single-user mode.

    ```
    ok boot -s
    ```

12. At the following prompt, enter **recovery** to start the root password recovery procedure.

    ```
    Enter full pathname of shell or 'recovery' for root password recovery or RETURN
    for /bin/sh: recovery
    ```

13. Enter configuration mode in the CLI.
14. Set the root password. For example:

    ```
    user@switch# set system root-authentication plain-text-password
    ```

15. At the following prompt, enter the new root password. For example:

    ```
    New password: juniper1
    Retype new password:
    ```

16. At the second prompt, reenter the new root password.
17. After you have finished configuring the password, commit the configuration.

    ```
    root@host# commit
    commit complete
    ```

18. Exit configuration mode in the CLI.
19. Exit operational mode in the CLI.
20. At the prompt, enter **y** to reboot the switch.

    ```
    Reboot the system? [y/n] y
    ```

Related Documentation

- *Configuring the Root Password*

Creating an Emergency Boot Device for a QFX Series Device

If Junos OS on the QFX Series is damaged in some way that prevents the software from loading properly, you can use an emergency boot device to repartition the primary disk and load a fresh installation of Junos OS. Use the following procedure to create an emergency boot device.

Before you begin, you need to download the installation media image for your device and Junos OS release from http://www.juniper.net/customers/support/ .

NOTE: In the following procedure, we assume that you are creating the emergency boot device on a QFX device. You can create the emergency boot device on another Juniper Networks switch or router, or any PC or laptop that supports Linux. The steps you take to create the emergency boot device vary, depending on the device.

To create an emergency boot device from a QFX device:

1. Use FTP to copy the installation media image into the **/var/tmp** directory on the QFX device.
2. Insert a USB device into the USB port.
3. From the Junos OS command-line interface (CLI), start the shell:

   ```
   user@device> start shell
   %
   ```
4. Switch to the root account using the **su** command:

   ```
   % su
   Password: password
   ```

   NOTE: The password is the root password for the QFX device. If you logged in to the device as root, you do not need to perform this step.
5. Enter the following command on the QFX3500, QFX3600, and QFX3600-I devices:

   ```
   root@device% dd if=/var/tmp/filename of=/dev/da1 bs=16k
   ```

   The device writes the installation media image to the USB device:

   ```
   root@device% dd if=/var/tmp/install-media-qfx3500.junos_11.1 of=/dev/da1 bs=16k
   11006+1 records in
   11006+1 records out
   180332544 bytes transferred in 71.764266 secs (2512846 bytes/sec)
   ```
6. Enter the following command on the QFX5100 device:

   ```
   root@device% dd if=/var/tmp/filename of=/dev/da0 bs=1048576
   ```

   The device writes the installation media image to the USB device:

   ```
   root@device% dd if=/var/tmp/jinstall-vjunos-usb-13.2.img of=/dev/da0 bs=1048576
   11006+1 records in
   11006+1 records out
   180332544 bytes transferred in 71.764266 secs (2512846 bytes/sec)
   ```
7. Log out of the shell:

   ```
   root@device% exit
   % exit
   user@device>
   ```

Related Documentation

- *USB Port Specifications for the QFX Series*
- Performing a Recovery Installation on a QFX Series Device on page 52
- *Performing a QFabric System Recovery Installation on the Director Group*

- *Performing a Recovery Installation on a QFX5100 Switch*

Performing a Recovery Installation on a QFX Series Device

If Junos OS on your device is damaged in some way that prevents the software from loading correctly, you may need to perform a recovery installation using an emergency boot device (for example, a USB flash drive) to restore the default factory installation. Once you have recovered the software, you need to restore the device configuration. You can either create a new configuration as you did when the device was shipped from the factory, or if you saved the previous configuration, you can simply restore that file to the device.

If at all possible, you should try to perform the following steps before you perform the recovery installation:

1. Ensure that you have an emergency boot device to use during the installation. See "Creating an Emergency Boot Device for a QFX Series Device" on page 50 for information on how to create an emergency boot device.
2. Copy the existing configuration in the file **/config/juniper.conf.gz** from the device to a remote system, such as a server, or to an emergency boot device. For extra safety, you can also copy the backup configurations (the files named **/config/juniper.conf.*n***, where *n* is a number from 0 through 9) to a remote system or to an emergency boot device.

   WARNING: **The recovery installation process completely overwrites the entire contents of the internal flash storage.**

3. Copy any other stored files to a remote system as desired.

To reinstall Junos OS:

1. Insert the emergency boot device into the QFX Series device.
2. Reboot the QFX Series device.

   NOTE: **Do not power off the device if it is already on.**

   [edit system]
   user@device> request system reboot

   If you do not have access to the CLI, power cycle the QFX Series device.

   The emergency boot device (external USB install media) is detected. At this time, you can load the Junos OS from the emergency boot device onto the internal flash storage.

3. The software prompts you with the following options:

```
External USB install media detected.
You can load Junos from this media onto an internal drive.
```

```
Press 'y' to proceed, 'f' to format and install, or 'n' to abort.
Do you wish to continue ([y]/f/n)? f
```

4. Type **f** to format the internal flash storage and install the Junos OS on the emergency boot device onto the internal flash storage.

   If you do no want to format the internal flash storage, type **y**.

   The following messages are displayed:

```
Installing packages from external USB drive da1
Packages will be installed to da0, media size: 8G

Processing format options
Fri September  4 01:18:44 UTC 2012

-- IMPORTANT INFORMATION --
Installer has detected settings to format system boot media.
This operation will erase all data from your system.

Formatting installation disk .. this will take a while, please wait
Disabling platform watchdog - threshold 12 mins

Determining installation slice
Fri September  4 01:27:07 UTC 2012
```

5. The device copies the software from the emergency boot device, occasionally displaying status messages. Copying the software can take up to 12 minutes.

   When the device is finished copying the software, you are presented with the following prompt:

```
*** Fri September  4 01:19:00 UTC 2012***
Installation successful..
Please select one of the following options:
Reboot to installed Junos after removing install media (default) ... 1
Reboot to installed Junos by disabling install media ............... 2
Exit to installer debug shell ..................................... 3
Install Junos to alternate slice .................................. 4
Your choice: 4
NOTE: System installer will now install Junos to alternate slice
Do not power off or remove the external installer media or
interrupt the installation mechanism.
```

6. Select **4** to install Junos OS to the alternate slice of the partition, and then press Enter.
7. Remove the emergency boot device when prompted and then press Enter. The device then reboots from the internal flash storage on which the software was just installed. When the reboot is complete, the device displays the login prompt.
8. Create a new configuration as you did when the device was shipped from the factory, or restore the previously saved configuration file to the device.

Related Documentation

- Creating an Emergency Boot Device for a QFX Series Device on page 50

CHAPTER 9

Layer 3 Protocols

Troubleshooting Virtual Routing Instances

Direct Routes Not Leaked Between Routing Instances

Problem Direct routes are not exported (leaked) between virtual routing instances For example, consider the following scenario:

- QFX switch with two virtual routing instances:
  - Routing instance 1 connects to downstream device through interface xe-0/0/1.
  - Routing instance 2 connects to upstream device through interface xe-0/0/2.

If you enable route leaking between the routing instances (by using the **rib-group** statement, for example), the downstream device cannot connect to the upstream device because the QFX switch connects to the upstream device over a direct route and these routes are not leaked between instances.

NOTE: **You can see a route to the upstream device in the routing table of the downstream device, but this route is not functional.**

Indirect routes *are* leaked between routing instances, so the downstream device can connect to any upstream devices that are connected to the QFX switch over indirect routes.

Solution This is expected behavior.

Related Documentation

- *Understanding Virtual Router Routing Instances*
- *Configuring Virtual Router Routing Instances*
- *rib-group*

CHAPTER 10

Security

Troubleshooting Firewall Filter Configuration

Use the following information to troubleshoot your firewall filter configuration.

Firewall Filter Configuration Returns a No Space Available in TCAM Message

Problem When a firewall filter configuration exceeds the amount of available Ternary Content Addressable Memory (TCAM) space, the system returns the following **syslogd** message:

```
No space available in tcam.
Rules for filter filter-name will not be installed.
```

A switch returns this message during the commit operation if the firewall filter that has been applied to a port, VLAN, or Layer 3 interface exceeds the amount of space available in the TCAM table. The filter is not applied, but the commit operation for the firewall filter configuration is completed in the CLI module.

Solution When a firewall filter configuration exceeds the amount of available TCAM table space, you must configure a new firewall filter with fewer filter terms so that the space requirements for the filter do not exceed the available space in the TCAM table.

You can perform either of the following procedures to correct the problem:

To delete the filter and its binding and apply the new smaller firewall filter to the same binding:

1. Delete the filter and its binding to ports, VLANs, or Layer 3 interfaces. For example:

   ```
   [edit]
   user@switch# delete firewall family ethernet-switching filter ingress-vlan-rogue-block
   user@switch# delete vlans employee-vlan description "filter to block rogue devices on
   employee-vlan"
   user@switch# delete vlans employee-vlan filter input ingress-vlan-rogue-block
   ```
2. Commit the changes:

   ```
   [edit]
   user@switch# commit
   ```
3. Configure a smaller filter with fewer terms that does not exceed the amount of available TCAM space. For example:

   ```
   [edit]
   user@switch# set firewall family ethernet-switching filter new-ingress-vlan-rogue-block ...
   ```
4. Apply (bind) the new firewall filter to a port, VLAN , or Layer 3 interface. For example:

   ```
   [edit]
   user@switch# set vlans employee-vlan description "filter to block rogue devices on
   employee-vlan"
   user@switch# set vlans employee-vlan filter input new-ingress-vlan-rogue-block
   ```
5. Commit the changes:

   ```
   [edit]
   user@switch# commit
   ```

To apply a new firewall filter and overwrite the existing binding but not delete the original filter:

1. Configure a firewall filter with fewer terms than the original filter:

   ```
   [edit]
   user@switch# set firewall family ethernet-switching filter new-ingress-vlan-rogue-block...
   ```
2. Apply the firewall filter to the port, VLAN, or Layer 3 interfaces to overwrite the binding of the original filter—for example:

   ```
   [edit]
   user@switch# set vlans employee-vlan description "smaller filter to block rogue devices on
   employee-vlan"
   user@switch# set vlans employee-vlan filter input new-ingress-vlan-rogue-block
   ```

   Because you can apply no more than one firewall filter per VLAN per direction, the binding of the original firewall filter to the VLAN is overwritten with the new firewall filter **new-ingress-vlan-rogue-block**.
3. Commit the changes:

   ```
   [edit]
   user@switch# commit
   ```


NOTE: The original filter is not deleted and is still available in the configuration.

Filter Counts Previously Dropped Packet

Problem If you configure two or more filters in the same direction for a physical interface and one of the filters includes a counter, the counter will be incorrect if the following circumstances apply:

- You configure the filter that is applied to packets first to discard certain packets. For example, imagine that you have a VLAN filter that accepts packets sent to 10.10.1.0/24 addresses and implicitly discards packets sent to any other addresses. You apply the filter to the **admin** VLAN in the output direction, and interface xe-0/0/1 is a member of that VLAN.
- You configure a subsequent filter to accept and count packets that are dropped by the first filter. In this example, you have a port filter that accepts and counts packets sent to 192.168.1.0/24 addresses that is also applied to xe-0/0/1 in the output direction.

The egress VLAN filter is applied first and correctly discards packets sent to 192.168.1.0/24 addresses. The egress port filter is applied next and counts the discarded packets as matched packets. The packets are not forwarded, but the counter displayed by the egress port filter is incorrect.

Remember that the order in which filters are applied depends on the direction in which they are applied, as indicated here:

Ingress filters:

1. Port (Layer 2) filter
2. VLAN filter
3. Router (Layer 3) filter

Egress filters:

1. Router (Layer 3) filter
2. VLAN filter
3. Port (Layer 2) filter

Solution This is expected behavior.

Matching Packets Not Counted

Problem If you configure two egress filters with counters for a physical interface and a packet matches both of the filters, only one of the counters includes that packet.

For example:

- You configure an egress port filter with a counter for interface xe-0/0/1.
- You configure an egress VLAN filter with a counter for the **admin**VLAN, and interface xe-0/0/1 is a member of that VLAN.
- A packet matches both filters.

In this case, the packet is counted by only one of the counters even though it matched both filters.

Solution This is expected behavior.

Counter Reset When Editing Filter

Problem If you edit a firewall filter term, the value of any counter associated with any term in the same filter is set to 0, including the implicit counter for any policer referenced by the filter. Consider the following examples:

- Assume that your filter has **term1**, **term2**, and **term3**, and each term has a counter that has already counted matching packets. If you edit any of the terms in any way, the counters for all the terms are reset to 0.
- Assume that your filter has **term1** and **term2**. Also assume that **term2** has a **policer** action modifier and the implicit counter of the policer has already counted 1000 matching packets. If you edit **term1** or **term2** in any way, the counter for the policer referenced by **term2** is reset to 0.

Solution This is expected behavior.

Cannot Include loss-priority and policer Actions in Same Term

Problem You cannot include both of the following actions in the same firewall filter term in a QFX Series switch:

- **loss-priority**
- **policer**

If you do so, you see the following error message when you attempt to commit the configuration: “cannot support policer action if loss-priority is configured.”

Solution This is expected behavior.

Cannot Egress Filter Certain Traffic Originating on QFX Switch

Problem On a QFX Series switch, you cannot filter certain traffic with a firewall filter applied in the output direction if the traffic originates on the QFX switch. This limitation applies to control traffic for protocols such as ICMP (ping), STP, LACP, and so on.

Solution This is expected behavior.

Firewall Filter Match Condition Not Working with Q-in-Q Tunneling

Problem If you create a firewall filter that includes a match condition of **dot1q-tag** or **dot1q-user-priority** and apply the filter on input to a trunk port that participates in a service VLAN, the match condition does not work if the Q-in-Q EtherType is not 0x8100. (When Q-in-Q tunneling is enabled, trunk interfaces are assumed to be part of the service provider or data center network and therefore participate in service VLANs.)

Solution This is expected behavior. To set the Q-in-Q EtherType to 0x8100, enter the **set dot1q-tunneling ethertype 0x8100** statemen at the **[edit ethernet-switching-options]** hierarchy level. You must also configure the other end of the link to use the same Ethertype.

Egress Firewall Filters with Private VLANs

Problem If you apply a firewall filter in the output direction to a primary VLAN, the filter also applies to the secondary VLANs that are members of the primary VLAN when the traffic egresses with the primary VLAN tag or isolated VLAN tag, as listed below:

- Traffic forwarded from a secondary VLAN trunk port to a promiscuous port (trunk or access)
- Traffic forwarded from a secondary VLAN trunk port that carries an isolated VLAN to a PVLAN trunk port.
- Traffic forwarded from a promiscuous port (trunk or access) to a secondary VLAN trunk port
- Traffic forwarded from a PVLAN trunk port. to a secondary VLAN trunk port
- Traffic forwarded from a community port to a promiscuous port (trunk or access)

If you apply a firewall filter in the output direction to a primary VLAN, the filter does *not* apply to traffic that egresses with a community VLAN tag, as listed below:

- Traffic forwarded from a community trunk port to a PVLAN trunk port
- Traffic forwarded from a secondary VLAN trunk port that carries a community VLAN to a PVLAN trunk port
- Traffic forwarded from a promiscuous port (trunk or access) to a community trunk port
- Traffic forwarded from a PVLAN trunk port. to a community trunk port

If you apply a firewall filter in the output direction to a community VLAN, the following behaviors apply:

- The filter is applied to traffic forwarded from a promiscuous port (trunk or access) to a community trunk port (because the traffic egresses with the community VLAN tag).
- The filter is applied to traffic forwarded from a community port to a PVLAN trunk port (because the traffic egresses with the community VLAN tag).

- The filter is *not* applied to traffic forwarded from a community port to a promiscuous port (because the traffic egresses with the primary VLAN tag or untagged).

Solution These are expected behaviors. They occur only if you apply a firewall filter to a private VLAN in the output direction and do not occur if you apply a firewall filter to a private VLAN in the input direction.

Egress Filtering of L2PT Traffic Not Supported

Problem Egress filtering of L2PT traffic is not supported on the QFX3500 switch. That is, if you configure L2PT to tunnel a protocol on an interface, you cannot also use a firewall filter to filter traffic for that protocol on that interface in the output direction. If you commit a configuration for this purpose, the firewall filter is not applied to the L2PT-tunneled traffic.

Solution This is expected behavior.

Cannot Drop BGP Packets in Certain Circumstances

Problem BGP packets with a time-to-live (TTL) value greater than 1 cannot be discarded using a firewall filter applied to a loopback interface or applied on input to a Layer 3 interface. BGP packets with TTL value of 1 or 0 can be discarded using a firewall filter applied to a loopback interface or applied on input to a Layer 3 interface.

Solution This is expected behavior.

Invalid Statistics for Policer

Problem If you apply a single-rate two-color policer in more than 128 terms in a firewall filter, the output of the **show firewall** command displays incorrect data for the policer.

Solution This is expected behavior.

Policers can Limit Egress Filters

Problem The number of egress policers that you configure can affect the total number of allowed egress firewall filters. Every policer has two implicit counters that consume two entries in a 1024-entry TCAM that is used for counters, including counters that are configured as action modifiers in firewall filter terms. (Policers consume two entries because one is used for green packets and one is used for nongreen packets regardless of policer type.) If the TCAM becomes full, you cannot commit any more egress firewall filters that have terms with counters. For example, if you configure and commit 512 egress policers (two-color, three-color, or a combination of both policer types), all of the memory entries for counters are used up. If later in your configuration file you insert additional egress firewall filters with terms that also include counters, *none* of the terms in those filters are committed because there is no available memory space for the counters.

Here are some additional examples:

- Assume that you configure egress filters that include a total of 512 policers and no counters. Later in your configuration file you include another egress filter with 10 terms, 1 of which has a counter action modifier. None of the terms in this filter are committed because there is not enough TCAM space for the counter.
- Assume that you configure egress filters that include a total of 500 policers, so 1000 TCAM entries are occupied. Later in your configuration file you include the following two egress filters:
  - Filter A with 20 terms and 20 counters. All the terms in this filter are committed because there is enough TCAM space for all the counters.
  - Filter B comes after Filter A and has five terms and five counters. *None* of the terms in this filter are committed because there is not enough memory space for *all* the counters. (Five TCAM entries are required but only four are available.)

Solution You can prevent this problem by ensuring that egress firewall filter terms with counter actions are placed earlier in your configuration file than terms that include policers. In this circumstance, Junos OS commits policers even if there is not enough TCAM space for the implicit counters. For example, assume the following:

- You have 1024 egress firewall filter terms with counter actions.
- Later in your configuration file you have an egress filter with 10 terms. None of the terms have counters but one has a policer action modifier.

You can successfully commit the filter with 10 terms even though there is not enough TCAM space for the implicit counters of the policer. The policer is committed without the counters.

Related Documentation

- *Understanding FIP Snooping, FBF, and MVR Filter Scalability*
- *Configuring Firewall Filters*
- *Verifying That Firewall Filters Are Operational*

Troubleshooting Policer Configuration

Incomplete Count of Packet Drops

Problem Under certain circumstances, Junos OS might display a misleading number of packets dropped by an ingress policer.

If packets are dropped because of ingress admission control, policer statistics might not show the number of packet drops you would expect by calculating the difference between ingress and egress packet counts. This might happen if you apply an ingress policer to multiple interfaces, and the aggregate ingress rate of those interfaces exceeds the line rate of a common egress interface. In this case, packets might be dropped from the ingress buffer. These drops are not included in the count of packets dropped by the policer, which causes policer statistics to underreport the total number of drops.

Solution This is expected behavior.

Counter Reset When Editing Filter

Problem If you edit a firewall filter term, the value of any counter associated with any term in the same filter is set to 0, including the implicit counter for any policer referenced by the filter. Consider the following examples:

- Assume that your filter has **term1**, **term2**, and **term3**, and each term has a counter that has already counted matching packets. If you edit any of the terms in any way, the counters for all the terms are reset to 0.
- Assume that your filter has **term1** and **term2**. Also assume that **term2** has a **policer** action modifier and the implicit counter of the policer has already counted 1000 matching packets. If you edit **term1** or **term2** in any way, the counter for the policer referenced by **term2** is reset to 0.

Solution This is expected behavior.

Invalid Statistics for Policer

Problem If you apply a single-rate two-color policer in more than 128 terms in a firewall filter, the output of the **show firewall** command displays incorrect data for the policer.

Solution This is expected behavior.

Egress Policers on QFX3500 Devices Might Allow More Throughput Than Is Configured

Problem If you configure a policer to rate-limit throughput and apply it on egress to multiple interfaces on a QFX3500 switch or Node, the measured aggregate policed rate might be twice the configured rate, depending on which interfaces you apply the policer to. The doubling of the policed rate occurs if you apply a policer to multiple interfaces and *both* of the following are true:

- There is at least one policed interface in the range xe-0/0/0 to xe-0/0/23 or the range xe-0/1/1 to xe-0/1/7.

- There is at least one policed interface in the range xe-0/0/24 to xe-0/0/47 or the range xe-0/1/8 to xe-0/1/15.

For example, if you configure a policer to rate-limit traffic at 1 Gbps and apply the policer (by using a firewall filter) to xe-0/0/0 and xe-0/0/24 in the output direction, each interface is rate-limited at 1 Gbps, for a total allowed throughput of 2 Gbps. The same behavior occurs if you apply the policer to xe-0/1/1 and xe-0/0/24—each interface is rate-limited at 1 Gbps.

If you apply the same policer on egress to multiple interfaces in these groups, each *group* is rate-limited at 1 Gbps. For example, if you apply the policer to xe-0/0/0 through xe-0/0/4 (five interfaces) and xe-0/0/24 through xe-0/0/33 (ten interfaces), each group is rate-limited at 1 Gbps, for a total allowed throughput of 2 Gbps.

Here is another example: If you apply the policer to xe-0/0/0 through xe-0/0/4 and xe-0/1/1 through xe-0/1/5 (a total of ten interfaces), that group is rate-limited at 1 Gbps in aggregate. If you also apply the policer to xe-0/0/24, that one interface is rate-limited at 1 Gbps while the other ten are still rate-limited at 1 Gbps in aggregate.

Interfaces xe-0/1/1 through xe-0/1/15 are physically located on the QSFP+ uplink ports, according to the following scheme:

- xe-0/1/1 through xe-0/1/3 are on Q0.
- xe-0/1/4 through xe-0/1/7 are on Q1.
- xe-0/1/8 through xe-0/1/11 are on Q2.
- xe-0/1/2 through xe-0/1/15 are on Q3.

The doubling of the policed rate occurs only if the policer is applied in the output direction. If you configure a policer as described above but apply it in the input direction, the total allowed throughput for all interfaces is 1 Gbps.

Solution This is expected behavior.

Filter-Specific Egress Policers on QFX3500 Devices Might Allow More Throughput Than Is Configured

Problem You can configure policers to be filter-specific, which means that Junos OS creates only one policer instance regardless of how many times the policer is referenced. When you do this, rate limiting is applied in aggregate, so if you configure a policer to discard traffic that exceeds 1 Gbps and reference that policer in three different terms, the total bandwidth allowed by the filter is 1 Gbps. However, the behavior of a filter-specific policer is affected by how the firewall filter terms that reference the policer are stored in ternary content addressable memory (TCAM). If you create a filter-specific policer and reference it in multiple firewall filter terms, the policer allows more traffic than expected if the terms are stored in different TCAM slices. For example, if you configure a policer to discard traffic that exceeds 1 Gbps and reference that policer in three different terms that are stored in three separate memory slices, the total bandwidth allowed by the filter is 3 Gbps, not 1 Gbps.

Solution To prevent this unexpected behavior, use the information about TCAM slices presented in *Planning the Number of Firewall Filters to Create* to organize your configuration file so that all the firewall filter terms that reference a given filter-specific policer are stored in the same TCAM slice.

Policers Can Limit Egress Filters

Problem The number of egress policers that you configure can affect the total number of allowed egress firewall filters. Every policer has two implicit counters that consume two entries in a 1024-entry TCAM that is used for counters, including counters that are configured as action modifiers in firewall filter terms. (Policers consume two entries because one is used for green packets and one is used for nongreen packets regardless of policer type.) If the TCAM becomes full, you cannot commit any more egress firewall filters that have terms with counters. For example, if you configure and commit 512 egress policers (two-color, three-color, or a combination of both policer types), all of the memory entries for counters are used up. If later in your configuration file you insert additional egress firewall filters with terms that also include counters, *none* of the terms in those filters are committed because there is no available memory space for the counters.

Here are some additional examples:

- Assume that you configure egress filters that include a total of 512 policers and no counters. Later in your configuration file you include another egress filter with 10 terms, 1 of which has a counter action modifier. None of the terms in this filter are committed because there is not enough TCAM space for the counter.
- Assume that you configure egress filters that include a total of 500 policers, so 1000 TCAM entries are occupied. Later in your configuration file you include the following two egress filters:
  - Filter A with 20 terms and 20 counters. All the terms in this filter are committed because there is enough TCAM space for all the counters.
  - Filter B comes after Filter A and has five terms and five counters. *None* of the terms in this filter are committed because there is not enough memory space for *all* the counters. (Five TCAM entries are required but only four are available.)

Solution You can prevent this problem by ensuring that egress firewall filter terms with counter actions are placed earlier in your configuration file than terms that include policers. In this circumstance, Junos OS commits policers even if there is not enough TCAM space for the implicit counters. For example, assume the following:

- You have 1024 egress firewall filter terms with counter actions.
- Later in your configuration file you have an egress filter with 10 terms. None of the terms have counters but one has a policer action modifier.

You can successfully commit the filter with 10 terms even though there is not enough TCAM space for the implicit counters of the policer. The policer is committed without the counters.

CHAPTER 11

Services

Troubleshooting Port Mirroring

Port Mirroring Constraints and Limitations

Local and Remote Port Mirroring

The following constraints and limitations apply to local and remote port mirroring with the QFX Series:

- You can create a total of four port-mirroring configurations on a QFX Series standalone switch.
- You can create a total of four port-mirroring configurations on each Node group in a QFabric system, subject to the following constraints:
  - As many as four of the configurations can be for local port mirroring.
  - As many as three of the configurations can be for remote port mirroring.
- Regardless of whether you are configuring a standalone switch or a Node group, the following limits apply:
  - There can be no more than two configurations that mirror ingress traffic. (If you configure a firewall filter to send traffic to a port mirror—that is, you use the **analyzer** action modifier in a filter term—this counts as an ingress mirroring configuration for switch or Node group on which the filter is applied.)
  - There can be no more than two configurations that mirror egress traffic.

NOTE: On QFabric systems, there is no system-wide limit on the total number of mirror sessions.

- You can configure no more than one type of output in one port-mirroring configuration. That is, you can use no more than one of the following to complete a **set analyzer *name* output** statement:
  - **interface**
  - **ip-address**
  - **vlan**
- If you configure Junos OS to mirror egress packets, do not configure more than 2000 VLANs on a QFX3500 device or QFabric system. If you do so, some VLAN packets might contain incorrect VLAN IDs. This applies to any VLAN packets—not only the mirrored copies.
- The **ratio** and **loss-priority** options are not supported.
- Packets with physical layer errors are filtered out and are not sent to the output port or VLAN.
- If you use sFlow monitoring to sample traffic, it does not sample the mirror copies when they exit from the output interface.
- You cannot mirror packets exiting or entering the following ports:
  - Dedicated Virtual Chassis interfaces
  - Management interfaces (me0 or vme0)
  - Fibre Channel interfaces
  - Routed VLAN interfaces
- An aggregated Ethernet interface cannot be an output interface.
- Do not include an 802.1Q subinterface that has a unit number other than 0 in a port mirroring configuration. Port mirroring does not work with subinterfaces if their unit number is not 0. (You configure 802.1Q subinterfaces using the **vlan-tagging** statement.)
- When packet copies are sent out the output interface, they are not modified for any changes that are normally applied on egress, such as CoS rewriting.
- An interface can be the input interface for only one mirroring configuration. Do not use the same interface as the input interface for multiple mirroring configurations.
- CPU-generated packets (such as ARP, ICMP, BPDU, and LACP packets) cannot be mirrored on egress.
- VLAN-based mirroring is not supported for STP traffic.
- (QFabric systems only) If you configure a QFabric analyzer to mirror egress traffic and the input and output interfaces are on different Node devices, the mirrored copies have incorrect VLAN IDs. This limitation does not apply if you configure a QFabric analyzer

to mirror egress traffic and the input and output interfaces are on the *same* Node device. In this case the mirrored copies have the correct VLAN IDs (as long as you do not configure more than 2000 VLANs on the QFabric system).

Remote Port Mirroring Only

The following constraints and limitations apply to remote port mirroring with the QFX Series:

- If you configure an output IP address, the address cannot be in the same subnetwork as any of the switch's management interfaces.
- If you create virtual routing instances and also create an analyzer configuration that includes an output IP address, the output address belongs to the default virtual routing instance (inet.0 routing table).
- An output VLAN cannot be a private VLAN or VLAN range.
- An output VLAN cannot be shared by multiple **analyzer** statements.
- An output VLAN interface cannot be a member of any other VLAN.
- An output VLAN interface cannot be an aggregated Ethernet interface.
- On the source (monitored) switch, only one interface can be a member of the analyzer VLAN.

Egress Port Mirroring with VLAN Translation

Problem If you create a port-mirroring configuration that mirrors customer VLAN (CVLAN) traffic on egress and the traffic undergoes VLAN translation before being mirrored, the VLAN translation does not apply to the mirrored packets. That is, the mirrored packets retain the service VLAN (SVLAN) tag that should be replaced by the CVLAN tag on egress. The original packets are unaffected—on these packets VLAN translation works properly, and the SVLAN tag is replaced with the CVLAN tag on egress.

Solution This is expected behavior.

Egress Port Mirroring with Private VLANs

Problem If you create a port-mirroring configuration that mirrors private VLAN (PVLAN) traffic on egress, the mirrored traffic (the traffic that is sent to the analyzer system) has the VLAN tag of the ingress VLAN instead of the egress VLAN. For example, assume the following PVLAN configuration:

- Promiscuous trunk port that carries primary VLANs pvlan100 and pvlan400.
- Isolated access port that carries secondary VLAN isolated200. This VLAN is a member of primary VLAN pvlan100.

- Community port that carries secondary VLAN comm300. This VLAN is also a member of primary VLAN pvlan100.
- Output interface (monitor interface) that connects to the analyzer system. This interface forwards the mirrored traffic to the analyzer.

If a packet for pvlan100 enters on the promiscuous trunk port and exits on the isolated access port, the original packet is untagged on egress because it is exiting on an access port. However, the mirror copy retains the tag for pvlan100 when it is sent to the analyzer.

Here is another example: If a packet for comm300 ingresses on the community port and egresses on the promiscuous trunk port, the original packet carries the tag for pvlan100 on egress, as expected. However, the mirrored copy retains the tag for comm300 when it is sent to the analyzer.

Solution This is expected behavior.

Related Documentation

- *Understanding Port Mirroring*
- *Example: Configuring Port Mirroring for Local Analysis*
- *Example: Configuring Port Mirroring for Remote Analysis*

CHAPTER 12

Traffic Management

Troubleshooting Egress Bandwidth That Exceeds the Configured Maximum Bandwidth

Problem The maximum bandwidth of a queue when measured at the egress port exceeds the maximum bandwidth (shaping rate) configured for the queue.

Cause When you configure bandwidth for a queue or a priority group, the switch accounts for the configured bandwidth as data only. The switch does not rate-shape the preamble and the interframe gap (IFG) associated with frames, so the switch does not account for the bandwidth consumed by the preamble and the IFG in its maximum bandwidth calculations.

The measured egress bandwidth can exceed the configured maximum bandwidth when small packet sizes (64 or 128 bytes) are transmitted because the preamble and the IFG are a larger percentage of the total traffic. For larger packet sizes, the preamble and IFG overhead are a small portion of the total traffic, and the effect on egress bandwidth is minor.

Solution When you calculate the bandwidth requirements for queues on which you expect a significant amount of traffic with small packet sizes, consider the shaping rate as the maximum bandwidth for the data only. Add sufficient bandwidth to your calculations to account for the preamble and IFG so that the port bandwidth is sufficient to handle the combined maximum data rate (shaping rate) and the preamble and IFG.

If the maximum bandwidth measured at the egress port exceeds the amount of bandwidth that you want to allocate to the queue, reduce the shaping rate for that queue.

Related Documentation

- *shaping-rate*
- *Example: Configuring Maximum Output Bandwidth*

- *Example: Configuring Queue Schedulers*
- *Understanding CoS Output Queue Schedulers*

Troubleshooting Egress Bandwidth That Exceeds the Configured Minimum Bandwidth

Problem The minimum bandwidth of a queue or a priority group when measured at the egress port exceeds the minimum bandwidth configured for the queue (transmit-rate) or for the priority group (guaranteed-rate).

Cause When you configure bandwidth for a queue or a priority group, the switch accounts for the configured bandwidth as data only. The switch does not include the preamble and the interframe gap (IFG) associated with frames, so the switch does not account for the bandwidth consumed by the preamble and the IFG in its minimum bandwidth calculations.

The measured egress bandwidth can exceed the configured minimum bandwidth when small packet sizes (64 or 128 bytes) are transmitted because the preamble and the IFG are a larger percentage of the total traffic. For larger packet sizes, the preamble and IFG overhead are a small portion of the total traffic, and the effect on egress bandwidth is minor.

NOTE: The sum of the queue transmit rates in a priority group should not exceed the guaranteed rate for the priority group. (You cannot guarantee a minimum bandwidth for the queues that is greater than the minimum bandwidth guaranteed for the entire set of queues.)

Solution When you calculate the bandwidth requirements for queues and priority groups on which you expect a significant amount of traffic with small packet sizes, consider the transmit rate and the guaranteed rate as the minimum bandwidth for the data only. Add sufficient bandwidth to your calculations to account for the preamble and IFG so that the port bandwidth is sufficient to handle the combined minimum data rate and the preamble and IFG.

If the minimum bandwidth measured at the egress port exceeds the amount of bandwidth that you want to allocate to a queue or to a priority group, reduce the transmit rate for that queue and reduce the guaranteed rate of the priority group that contains the queue.

Related Documentation

- *guaranteed-rate*
- *transmit-rate*
- *Example: Configuring Minimum Guaranteed Output Bandwidth*
- *Example: Configuring Queue Schedulers*
- *Understanding CoS Output Queue Schedulers*

Troubleshooting Egress Queue Bandwidth Impacted by Congestion

Problem Congestion on an egress port causes egress queues to receive less bandwidth than expected. Egress port congestion can impact the amount of bandwidth allocated to queues on the congested port and, in some cases, on ports that are not congested.

Cause Egress queue congestion can cause the ingress port buffer to fill above a certain threshold and affect the flow to the queues on the egress port. One queue receives its configured bandwidth, but the other queues on the egress port are affected and do not receive their configured share of bandwidth.

Solution The solution is to configure a drop profile to apply weighted random early detection (WRED) to the queue or queues on the congested ports.

Configure a drop profile on the queue that is receiving its configured bandwidth. This queue is preventing the other queues from receiving their expected bandwidth. The drop profile prevents the queue from affecting the other queues on the port.

To configure a tail-drop profile using the CLI:

- Name the drop profile and set the drop start point, drop end point, minimum drop rate, and maximum drop rate for the drop profile:

```
[edit class-of-service]
user@switch# set drop-profile drop-profile-name interpolate fill-level percentage fill-level
percentage drop-probability 0 drop-probability percentage
```

Related Documentation

- *drop-profile*
- *Example: Configuring Tail-Drop Profiles*
- *Example: Configuring CoS Hierarchical Port Scheduling (ETS)*
- *Understanding CoS Tail-Drop Profiles*

Troubleshooting an Unexpected Rewrite Value

Problem Traffic from one or more forwarding classes on an egress port is assigned an unexpected rewrite value.

NOTE: For packets that carry both an inner VLAN tag and an outer VLAN tag, the rewrite rules rewrite only the outer VLAN tag.

Cause If you configure a rewrite rule for a forwarding class on an egress port but you do not configure a rewrite rule for every forwarding class on that egress port, then the forwarding classes that do not have a configured rewrite rule are assigned random rewrite values.

For example:

1. Configure forwarding classes **fc1**, **fc2**, and **fc3**.
2. Configure rewrite rules for forwarding classes **fc1** and **fc2**, but not for forwarding class **fc3**.
3. Assign forwarding classes **fc1**, **fc2**, and **fc3** to a port.

When traffic for these forwarding classes flows through the port, traffic for forwarding classes **fc1** and **fc2** is rewritten correctly. However, traffic for forwarding class **fc3** is assigned a random rewrite value.

Solution If any forwarding class on an egress port has a configured rewrite rule, then all forwarding classes on that egress port must have a configured rewrite rule. Configuring a rewrite rule for any forwarding class that is assigned a random rewrite value solves the problem.

TIP: If you want the forwarding class to use the same code point value assigned to it by the ingress classifier, specify that value as the rewrite rule value. For example, if a forwarding class has the IEEE 802.1 ingress classifier code point value 011, configure a rewrite rule for that forwarding class that uses the IEEE 802.1p code point value 011.

NOTE: There are no default rewrite rules. You can bind one rewrite rule for each type (DSCP and IEEE 802.1) to a given interface. A rewrite rule can contain multiple forwarding-class-to-rewrite-value associations.

1. Assign a rewrite value to a forwarding class. Add the new rewrite value to the same rewrite rule as the other forwarding classes on the port:

   ```
   [edit class-of-service rewrite-rules]
   user@switch# set (dscp | ieee-802.1) rewrite-name forwarding-class class-name loss-priority priority code-point (alias | bits)
   ```

   For example, if the other forwarding classes on the port use rewrite values defined in the rewrite rule **custom-rw**, the forwarding class **fcoe** is being randomly rewritten, and you want to use IEEE 802.1 code point **011** for the **fcoe** forwarding class:

   ```
   [edit class-of-service rewrite-rules]
   user@switch# set ieee-802.1 custom-rw forwarding-class fcoe loss-priority high code-point 011
   ```

2. Enable the rewrite rule on an interface if it is not already enabled on the desired interface:

   ```
   [edit]
   user@switch# set class-of-service interfaces interface-name unit unit rewrite-rules (dscp | ieee-802.1) rewrite-rule-name
   ```

   For example, to enable the rewrite rule **custom-rw** on interface **xe-0/0/24.0**:

```
[edit]
user@switch# set class-of-service interfaces xe-0/0/24 unit 0 rewrite-rules ieee-802.1
custom-rw
```

Related Documentation

- *interfaces*
- *rewrite-rules*
- *Defining CoS Rewrite Rules*
- *Monitoring CoS Rewrite Rules*