Traffic Management on the OCX Series

Modified: 2018-08-21

Juniper Networks, Inc.
1133 Innovation Way
Sunnyvale, California 94089
USA
408-745-2000
www.juniper.net

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 https://www.juniper.net/support/eula/. By downloading, installing or using such software, you agree to the terms and conditions of that EULA.

Table of Contents

List of Figures

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 https://www.juniper.net/documentation/.

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 https://www.juniper.net/books.

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 CLI Explorer.

Documentation Conventions

Table 1 on page xxi 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.
	Tip	Indicates helpful information.
	Best practice	Alerts you to a recommended use or implementation.

Table 2 on page xxi 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**

Table 2: Text and Syntax Conventions (continued)

Convention	Description	Examples
`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*
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**.

Table 2: Text and Syntax Conventions (continued)

Convention	Description	Examples
> (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 provide feedback by using either of the following methods:

- Online feedback rating system—On any page of the Juniper Networks TechLibrary site at https://www.juniper.net/documentation/index.html, simply click the stars to rate the content, and use the pop-up form to provide us with information about your experience. Alternately, you can use the online feedback form at https://www.juniper.net/documentation/feedback/.
- E-mail—Send your comments to **techpubs-comments@juniper.net**. Include the document or topic name, URL or page number, and software 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 Partner Support Service 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 https://www.juniper.net/us/en/local/pdf/resource-guides/7100059-en.pdf.
- Product warranties—For product warranty information, visit https://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: https://www.juniper.net/customers/support/
- Search for known bugs: https://prsearch.juniper.net/
- Find product documentation: https://www.juniper.net/documentation/
- Find solutions and answer questions using our Knowledge Base: https://kb.juniper.net/

- Download the latest versions of software and review release notes: https://www.juniper.net/customers/csc/software/
- Search technical bulletins for relevant hardware and software notifications: https://kb.juniper.net/InfoCenter/
- Join and participate in the Juniper Networks Community Forum: https://www.juniper.net/company/communities/
- Open a case online in the CSC Case Management tool: https://www.juniper.net/cm/

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

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 https://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 https://www.juniper.net/support/requesting-support.html.

PART 1

CoS Overview

CHAPTER 1

Basic Concepts

Overview of CoS on OCX Series Switches

Many Juniper Networks data center switching platforms are optimized for Layer 2 Ethernet transport. However, OCX Series switches are optimized for Layer 3 IP transport. This difference results in some differences in class-of-service (CoS) feature support on OCX Series switches compared with, for example, QFX Series switches. This topic describes those differences.

In addition, some Juniper Networks documentation is shared among different Juniper Networks platforms. Because of this, you might see references to lossless transport, data center bridging exchange (DCBX) protocol, priority-based flow control (PFC), and Fibre Channel over Ethernet (FCoE) in the documentation. These references do not apply to OCX Series switches.

Supported CoS Features

OCX Series switches support the following CoS features:

- Incoming packet classification on Layer 3 physical interfaces when at least one logical interface is defined on the physical interface:

  - DSCP, DSCP IPv6, and IEEE 802.1p behavior aggregate (BA) classifiers
  - Fixed classifiers
  - Multifield classifiers

  The default DSCP BA classifier is the default classifier. It maps incoming unicast traffic into the best-effort (queue 0) and network-control (queue 7) forwarding classes.

- Up to eight unicast forwarding classes and up to four multidestination (multicast, broadcast, destination lookup fail) forwarding classes.
- DSCP, DSCP IPv6, and IEEE 802.1p rewrite rules on Layer 3 physical interfaces when at least one logical interface is defined on the physical interface.
- Hierarchical, two-tier port scheduling, also known as enhanced transmission selection (ETS).
- Per output queue control of forwarding classes:
  - Guaranteed minimum bandwidth
  - Maximum bandwidth
  - Scheduling priority
  - Weighted random early detection (WRED) packet drop characteristics for congestion management
- Explicit congestion notification (ECN).
- Symmetric Ethernet PAUSE flow control.
- Shared and dedicated buffer pool configuration.

Unsupported Lossless Transport Features

OCX Series switches do not support lossless transport. Lossless transport does not refer to best-effort traffic on a link with Ethernet PAUSE enabled. Lossless transport refers to traffic classified into lossless forwarding classes on which you enable priority-based flow control (PFC) (defined in IEEE 802.1Qbb). OCX Series switches do not support lossless forwarding classes and do not support PFC.

Unsupported Default Lossless Unicast Forwarding Classes

Because the Junos OS software is common to several data center switching platforms, two of the five default forwarding classes are lossless forwarding classes (the unicast *fcoe* and *no-loss* forwarding classes, which are mapped by default to output queue 3 and output queue 4, respectively). On OCX Series switches, the default fcoe and no-loss forwarding classes are not supported.

The default fcoe and no-loss lossless forwarding classes carry the no-loss packet drop attribute. On OCX Series switches, the no-loss packet drop attribute is not supported. Do not classify traffic into the default fcoe or no-loss forwarding classes. Do not configure a forwarding class with the no-loss packet drop attribute.

You can use queues 3 and 4, but you must configure a forwarding class and map it to the desired queue, configure a classifier to map incoming traffic to the forwarding class, and then apply the classifier to the appropriate interfaces.

The forwarding class names fcoe and no-loss are just that—names. If you want to carry traffic on queue 3 or queue 4, you can remove the no-loss packet drop attribute from these forwarding classes without changing the names. Or, you can change the forwarding class names to something else, just as long as you do not configure the no-loss packet drop attribute on the forwarding class.

For example, to configure a new forwarding class named be2, map it to queue 3, map traffic identified by DSCP code point 001010 to forwarding class be2, and apply the configuration to interface xe-0/0/20:

1. Configure a new forwarding class named be2, without the no-loss attribute, and map it to queue 3:

   ```
   [edit class-of-service]
   user@switch# set forwarding-classes class be2 queue-num 3
   ```

2. Configure a classifier named be_classifier to classify incoming traffic with DSCP code point 001010 into forwarding class be2:

   ```
   [edit class-of-service]
   user@switch# set classifiers dscp be_classifier import default forwarding-class be2
   loss-priority low code-points 001010
   ```

   Importing the default classifier bases the new classifier on the default classifier. The default packet classification is retained, except for the changes you make to the classifier configuration. In this case, you are adding a new mapping to the classifier default mapping and saving it as classifier be_classifier.

3. Apply the classifier to interface xe-0/0/20:

   ```
   [edit class-of-service]
   user@switch# set interfaces xe-0/0/20 classifiers dscp be_classifier
   ```


NOTE: In addition to configuring the forwarding class and packet classification, you must also add the new forwarding class to a forwarding class set, and apply the forwarding class set to interfaces in order to associate the traffic in a forwarding class with interfaces. For completeness, here are example statements to accomplish this

This brief addition to the example configures a forwarding class set named be_fc_set, maps forwarding class be2 and default forwarding class best-effort to be_fc_set, and applies be_fc_set to interface xe-0/0/20:

```
[edit class-of-service]
user@switch# set forwarding-class-sets be_fc_set class be2
user@switch# set forwarding-class-sets be_fc_set class best-effort
user@switch# set interfaces xe-0/0/20 forwarding-class-set be_fc_set
```

Because this example classifies traffic into a forwarding class that is mapped to one of the queues that the default scheduler services, traffic receives the default scheduling (bandwidth and priority) for that queue. If you classify traffic into a forwarding class that is mapped to a queue that does not receive default scheduling, configure a queue scheduler for the traffic to ensure that the traffic receives a minimum amount of bandwidth during periods of congestion.

Effect of Unsupported Unicast Lossless Forwarding Classes on Bandwidth Scheduling

The default scheduler provides bandwidth and scheduling priority for unicast queues as shown in Table 3 on page 6:

Table 3: Default Scheduler Configuration

Default Scheduler and Queue Number	Transmit Rate (Minimum Guaranteed Bandwidth)	Priority
best-effort forwarding class scheduler (queue 0)	5%	low
fcoe forwarding class scheduler (queue 3)	35%	low
no-loss forwarding class scheduler (queue 4)	35%	low
network-control forwarding class scheduler (queue 7)	5%	low
mcast forwarding class scheduler (queue 8)	20%	low

Because the default DSCP classifier maps unicast traffic only into queue 0 and queue 7, the 35 percent of the bandwidth allocated to queue 3 and queue 4 are shared with the traffic-bearing queues. However, if you create a new forwarding class and map it to queue 3 or queue 4, the traffic classified into that forwarding class receives the scheduling resources shown in Table 3 on page 6. If you want to change the default scheduler bandwidth allocations, configure a hierarchical port scheduler.

If you want to map unicast traffic to queues other than the default queues, configure a hierarchical port scheduler to allocate port bandwidth resources to the queues.

Priority-Based Flow Control

OCX Series switches do not support PFC, and they do not support configuration of the [edit class-of-service congestion-notification-profile] hierarchy in the CLI or applying a congestion notification profile to an interface. In the rest of the CoS documentation set, information about PFC does not apply to OCX Series switches.

DCBX

OCX Series switches do not support DCBX, and they do not support configuration of the [edit class-of-service congestion-notification-profile] hierarchy in the CLI. In the rest of the CoS documentation set, information about DCBX does not apply to OCX Series switches.

Unsupported CoS Features

OCX Series switches do not support the following CoS features that some other data center switches support:

- CoS on Layer 2 interfaces (OCX Series switches do not support family ethernet-switching interfaces).
- Incoming MPLS EXP packet classification (OCX Series switches do not support MPLS and do not support the [edit class-of-service system-defaults classifiers] hierarchy).
- Lossless transport.
- Default lossless forwarding classes (fcoe and no-loss forwarding classes).
- No-loss packet drop attribute; do not apply the no-loss packet drop attribute to any forwarding class, and do not use the default fcoe and no-loss forwarding classes, which carry the no-loss packet drop attribute by default.
- Priority-based flow control (PFC).
- Data center bridging exchange (DCBX) protocol.
- Outgoing MPLS EXP rewrite rules (OCX Series switches do not support MPLS).
- Asymmetric Ethernet PAUSE flow control.

Related Documentation

- Overview of Junos OS CoS on page 8
- Understanding Default CoS Settings on page 25
- Understanding Default CoS Scheduling and Classification on page 114
- Understanding CoS Classifiers on page 123
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162
- Understanding CoS Forwarding Classes on page 137

- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Understanding CoS Buffer Configuration on page 546

Overview of Junos OS CoS

When a network experiences congestion and delay, some packets must be dropped. Junos OS class of service (CoS) enables you to divide traffic into classes and set various levels of throughput and packet loss when congestion occurs. You have greater control over packet loss because you can configure rules tailored to your needs.

You can configure CoS features to provide multiple classes of service for different applications. CoS also allows you to rewrite the Differentiated Services code point (DSCP) or IEEE 802.1p code-point bits of packets leaving an interface, thus allowing you to tailor packets for the network requirements of the remote peers.

CoS provides multiple classes of service for different applications. You can configure multiple forwarding classes for transmitting packets, define which packets are placed into each output queue, schedule the transmission service level for each queue, and manage congestion using a weighted random early detection (WRED) algorithm.

In designing CoS applications, you must carefully consider your service needs, and you must thoroughly plan and design your CoS configuration to ensure consistency and interoperability across all platforms in a CoS domain.

Because CoS is implemented in hardware rather than in software, you can experiment with and deploy CoS features without affecting packet forwarding and switching performance.

NOTE: CoS policies can be enabled or disabled on each switch interface. Also, each physical and logical interface on the switch can have associated custom CoS rules.

When you change or when you deactivate and then reactivate the class-of-service configuration, the system experiences packet drops because the system momentarily blocks traffic to change the mapping of incoming traffic to input queues.

This topic describes:

- CoS Standards on page 9
- How Junos OS CoS Works on page 9
- Default CoS Behavior on page 10

CoS Standards

The following RFCs define the standards for CoS capabilities:

- RFC 2474, *Definition of the Differentiated Services Field in the IPv4 and IPv6 Headers*
- RFC 2597, *Assured Forwarding PHB Group*
- RFC 2598, *An Expedited Forwarding PHB*
- RFC 2698, *A Two Rate Three Color Marker*
- RFC 3168, *The Addition of Explicit Congestion Notification (ECN) to IP*

The following data center bridging (DCB) standards are also supported to provide the CoS (and other characteristics) that Fibre Channel over Ethernet (FCoE) requires for transmitting storage traffic over an Ethernet network:

- IEEE 802.1Qbb, priority-based flow control (PFC)
- IEEE 802.1Qaz, enhanced transmission selection (ETS)
- IEEE 802.1AB (LLDP) extension called Data Center Bridging Capability Exchange Protocol (DCBX)

NOTE: **OCX Series switches and NFX250 Network Services platforms do not support PFC and DCBX.**

Juniper Networks QFX10000 switches support both enhanced transmission selection (ETS) hierarchical port scheduling and direct port scheduling.

How Junos OS CoS Works

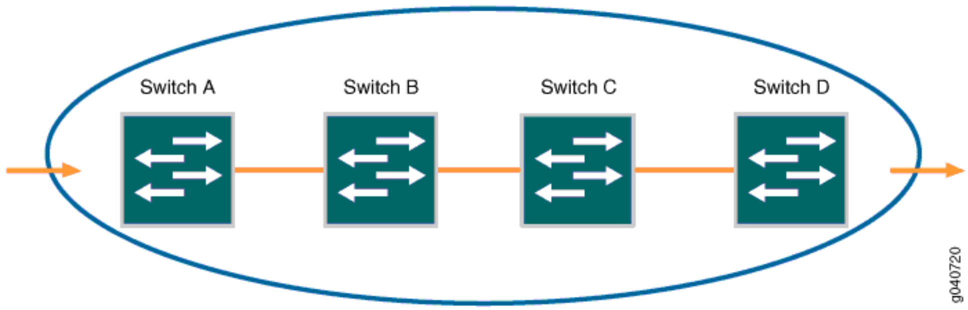
Junos OS CoS works by examining traffic entering the edge of your network. The switch classifies traffic into defined service groups to provide the special treatment of traffic across the network. For example, you can send voice traffic across certain links and data traffic across other links. In addition, the data traffic streams can be serviced differently along the network path to ensure that higher-paying customers receive better service. As the traffic leaves the network at the far edge, you can reclassify the traffic to meet the policies of the targeted peer by rewriting the DSCP or IEEE 802.1 code-point bits.

To support CoS, you must configure each switch in the network. Generally, each switch examines the packets that enter it to determine their CoS settings. These settings dictate which packets are transmitted first to the next downstream switch. Switches at the edges of the network might be required to alter the CoS settings of the packets that enter the network to classify the packets into the appropriate service groups.

In Figure 1 on page 10, Switch A is receiving traffic. As each packet enters, Switch A examines the packet's current CoS settings and classifies the traffic into one of the groupings defined on the switch. This definition allows Switch A to prioritize its resources for servicing the traffic streams it receives. Switch A might alter the CoS settings (forwarding class and loss priority) of the packets to better match the defined traffic groups.

When Switch B receives the packets, it examines the CoS settings, determines the appropriate traffic groups, and processes the packet according to those settings. It then transmits the packets to Switch C, which performs the same actions. Switch D also examines the packets and determines the appropriate groups. Because Switch D sits at the far end of the network, it can reclassify (rewrite) the CoS code-point bits of the packets before transmitting them.

Figure 1: Packet Flow Across the Network



Default CoS Behavior

If you do not configure CoS settings, the software performs some CoS functions to ensure that the system forwards traffic and protocol packets with minimum delay when the network is experiencing congestion. Some CoS settings, such as classifiers, are automatically applied to each logical interface that you configure. Other settings, such as rewrite rules, are applied only if you explicitly associate them with an interface.

Related Documentation

- Overview of Policers on page 40
- Understanding Junos CoS Components on page 15
- Understanding CoS Packet Flow on page 23
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315

Configuring CoS

The traffic management class-of-service topics describe how to configure the Junos OS class-of-service (CoS) components. Junos CoS provides a flexible set of tools that enable you to fine tune control over the traffic on your network.

- Define classifiers that classify incoming traffic into forwarding classes to place traffic in groups for transmission.
- Map forwarding classes to output queues to define the type of traffic on each output queue.
- Configure schedulers for each output queue to control the service level (priority, bandwidth characteristics) of each type of traffic.

- Provide different service levels for the same forwarding classes on different interfaces.
- On switches that support data center bridging standards, configure lossless transport across the Ethernet network using priority-based flow control (PFC), Data Center Bridging Exchange protocol (DCBX), and enhanced transmission selection (ETS) hierarchical scheduling (OCX Series switches and NFX250 Network Services platform do not support lossless transport, PFC, and DCBX).
- Configure various CoS components individually or in combination to define CoS services.

NOTE: When you change the CoS configuration or when you deactivate and then reactivate the CoS configuration, the system experiences packet drops because the system momentarily blocks traffic to change the mapping of incoming traffic to input queues.

Table 4 on page 12 lists the primary CoS configuration tasks by platform and provides links to those tasks.

NOTE: Links to features that are not supported on the platform for which you are looking up information might not be functional.

Table 4: CoS Configuration Tasks

<table>
<tr><th>CoS Configuration Task</th><th>Platforms Supported</th><th>Links</th></tr>
<tr><td>Basic CoS Configuration:
• Configure code-point aliases to assign a name to a pattern of code-point bits that you can use instead of the bit pattern when you configure CoS components such as classifiers and rewrite rules
• Configure classifiers and multidestination classifiers
  • Set the forwarding class and loss priority of a packet based on the incoming CoS value and assign packets to output queues based on the associated forwarding class
  • Change the host default output queue and mapping of DSCP bits used in the type of service (ToS) field
• Configure forwarding classes
• Configure rewrite rules to alter code point bit values in outgoing packets on the outbound interfaces of a switch so that the CoS treatment matches the policies of a targeted peer
• Configure Ethernet PAUSE flow control, a congestion relief feature that provides link-level flow control for all traffic on a full-duplex Ethernet link, including those that belong to Ethernet link aggregated (LAG) interfaces. On any particular interface, symmetric and asymmetric flow control are mutually exclusive.
• Assign the following CoS components to physical or logical interfaces:
  • Classifiers
  • Congestion notification profiles
  • Forwarding classes
  • Forwarding class sets
  • Output traffic control profiles
  • Port schedulers
  • Rewrite rules</td><td>• QFX3500
• QFX3600
• EX4600
• NFX250
• QFX5100
• QFX5200
• QFX5210
• QFX10000
• OCX1100 switches
• QFabric systems</td><td>• Defining CoS Code-Point Aliases on page 136
• (QFX10000 only) *Example: Configuring Classifiers*
• (Except QFX10000) “Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p)” on page 130
• (Except NFX250 and QFX10000) *Example: Configuring Multidestination (Multicast, Broadcast, DLF) Classifiers*
• Changing the Host Outbound Traffic Default Queue Mapping on page 156
• Example: Configuring Forwarding Classes on page 142
• Defining CoS Rewrite Rules on page 160
• (Except NFX250) “Enabling and Disabling CoS Symmetric Ethernet PAUSE Flow Control” on page 537
• (Except NFX250 and OCX1100) *Configuring CoS Asymmetric Ethernet PAUSE Flow Control*
• Assigning CoS Components to Interfaces on page 21</td></tr>
</table>

Table 4: CoS Configuration Tasks (continued)

CoS Configuration Task	Platforms Supported	Links
Configure Weighted random early detection (WRED) drop profiles that define the drop probability of packets of different packet loss probabilities (PLPs) as the output queue fills: • Configure WRED drop profiles where you associate WRED drop profiles with loss priorities in a scheduler. When you map the scheduler to a forwarding class (queue), you apply the interpolated drop profile to traffic of the specified loss priority on that queue. • Configure drop profile maps that map a drop profile to a packet loss priority, and associate the drop profile and packet loss priority with a scheduler • Configure explicit congestion notification (ECN) to enable end-to-end congestion notification between two endpoints on TCP/IP based networks. Apply WRED drop profiles to forwarding classes to control how the switch marks ECN-capable packets.	• QFX3500 • QFX3600 • EX4600 • QFX5100 • QFX5200 • QFX5210 • QFX10000 • OCX1100 switches • QFabric systems	• Example: Configuring WRED Drop Profiles on page 371 • Example: Configuring Drop Profile Maps on page 377 • Example: Configuring ECN on page 389
Configure queue schedulers and the bandwidth scheduling priority of individual queues. Schedulers define the CoS properties of output queues (output queues are mapped to forwarding classes, and classifiers map traffic into forwarding classes based on IEEE 802.1p or DSCP code points). Queue scheduling works with priority group scheduling to create a two-tier hierarchical scheduler. CoS scheduling properties include the amount of interface bandwidth assigned to the queue, the priority of the queue, whether explicit congestion notification (ECN) is enabled on the queue, and the WRED packet drop profiles associated with the queue.	• QFX3500 • QFX3600 • EX4600 • NFX250 • QFX5100 • QFX5200 • QFX5210 • QFX10000 • OCX1100 switches • QFabric systems	• (Except QFX10000) “Example: Configuring Queue Schedulers” on page 293 • Example: Configuring Queue Scheduling Priority on page 302 • (QFX10000 only) *Example: Configuring Queue Schedulers for Port Scheduling*
Configure traffic control profiles to define the output bandwidth and scheduling characteristics of forwarding class sets (priority groups). The forwarding classes (queues) mapped to a forwarding class set share the bandwidth resources that you configure in the traffic control profile.	• QFX3500 • QFX3600 • EX4600 • NFX250 • QFX5100 • QFX5200 • QFX5210 • QFX10000 • OCX1100 switches • QFabric systems	• (Except NFX250) “Defining CoS Traffic Control Profiles (Priority Group Scheduling)” on page 310 • (Except NFX250) “Example: Configuring Traffic Control Profiles (Priority Group Scheduling)” on page 311 • Example: Configuring Minimum Guaranteed Output Bandwidth on page 350 • (Except NFX250) “Example: Configuring Maximum Output Bandwidth” on page 357

Table 4: CoS Configuration Tasks (continued)

CoS Configuration Task	Platforms Supported	Links
Configure enhanced transmission selection (ETS) and forwarding class sets, and disable the ETS recommendation TLV. Hierarchical port scheduling, the Junos OS implementation of ETS, enables you to group priorities that require similar CoS treatment into priority groups. You define the port bandwidth resources for a priority group, and you define the amount of the priority group's resources that each priority in the group can use.	• QFX3500 • QFX3600 • EX4600 • QFX5100 • OCX1100 switches • QFX10000 • QFabric systems	• Example: Configuring Forwarding Class Sets on page 150 • Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321 • (Except OCX1100)*Disabling the ETS Recommendation TLV*
Configure Data Center Bridging Capability Exchange protocol (DCBX), which discovers the data center bridging (DCB) capabilities of peers by exchanging feature configuration information and is an extension of the Link Layer Discovery Protocol (LLDP) • Configure the DCBX mode that an interface uses to communicate with the connected peer • Configure DCBX autonegotiation on a per-interface basis for each supported feature or application • Define each application for which you want DCBX to exchange application protocol information • Map applications to IEEE 802.1p code points • Apply an application map to a DCBX interface	• QFX3500 • QFX3600 • EX4600 • QFX5100 • QFX5200 • QFX5210 • QFX10000 • QFabric systems	• *Example: Configuring DCBX Application Protocol TLV Exchange* • *Configuring the DCBX Mode* • *Configuring DCBX Autonegotiation* • *Defining an Application for DCBX Application Protocol TLV Exchange* • *Configuring an Application Map for DCBX Application Protocol TLV Exchange* • *Applying an Application Map to an Interface for DCBX Application Protocol TLV Exchange*
Configure CoS for FCoE: • Configure priority-based flow control (PFC) to divide traffic on one physical link into eight priorities • Configure a congestion notification profile (CNP) that enables priority-based flow control (PFC) on specified IEEE 802.1p priorities • Configure Multichassis link aggregation groups (MC-LAGs) to provide redundancy and load balancing between two switches • Configure two or more lossless forwarding classes and map them to different priorities • Configure lossless FCoE transport if your network uses a different priority than 3 • Configure multiple lossless FCoE priorities on a converged Ethernet network • If the FCoE network uses a different priority than priority 3 for FCoE traffic, configure a rewrite value to remap incoming traffic from the FC SAN to that priority after the interface encapsulates the FC packets in Ethernet • Configure lossless priorities for multiple types of traffic, such as FCoE and iSCSI	• QFX3500 • QFX3600 • EX4600 • QFX5100 • QFX5200 • QFX5210 • QFX10000 • QFabric systems	• *Example: Configuring CoS PFC for FCoE Traffic* • *Example: Configuring CoS for FCoE Transit Switch Traffic Across an MC-LAG* • *Configuring CoS PFC (Congestion Notification Profiles)* • (QFX3500 and QFabric only) *Example: Configuring IEEE 802.1p Priority Remapping on an FCoE-FC Gateway* • *Example: Configuring Two or More Lossless FCoE IEEE 802.1p Priorities on Different FCoE Transit Switch Interfaces* • *Example: Configuring Lossless FCoE Traffic When the Converged Ethernet Network Does Not Use IEEE 802.1p Priority 3 for FCoE Traffic (FCoE Transit Switch)* • *Example: Configuring Two or More Lossless FCoE Priorities on the Same FCoE Transit Switch Interface* • (QFX3500, NFX250, and QFabric only) *Configuring CoS Fixed Classifier Rewrite Values for Native FC Interfaces (NP_Ports)* • *Example: Configuring Lossless IEEE 802.1p Priorities on Ethernet Interfaces for Multiple Applications (FCoE and iSCSI)*

Understanding Junos CoS Components

This topic describes the Junos OS class-of-service (CoS) components:

Code-Point Aliases

A *code-point alias* assigns a name to a pattern of code-point bits. You can use this name instead of the bit pattern when you configure other CoS components such as classifiers and rewrite rules.

Policers

Policers limit traffic of a certain class to a specified bandwidth and burst size. Packets exceeding the policer limits can be discarded, or can be assigned to a different forwarding class, a different loss priority, or both. You define policers with filters that you can associate with input interfaces.

Classifiers

Packet classification associates incoming packets with a particular CoS servicing level. In Junos OS, *classifiers* associate packets with a forwarding class and loss priority and assign packets to output queues based on the associated forwarding class. Junos OS supports two general types of classifiers:

- Behavior aggregate (BA) or CoS value traffic classifiers—Examine the CoS value in the packet header. The value in this single field determines the CoS settings applied to the packet. BA classifiers allow you to set the forwarding class and loss priority of a packet based on the Differentiated Services code point (DSCP) value, IEEE 802.1p value, or MPLS EXP value.

  NOTE: OCX Series switches and NFX250 Network Services platform do not support MPLS.

- Multifield traffic classifiers—Examine multiple fields in the packet, such as source and destination addresses and source and destination port numbers of the packet. With multifield classifiers, you set the forwarding class and loss priority of a packet based on firewall filter rules.

On switches that require the separation of unicast and multidestination (multicast, broadcast, and destination lookup fail) traffic, you create separate unicast classifiers and multidestination classifiers. You cannot assign unicast traffic and multidestination traffic to the same classifier. You can apply unicast classifiers to one or more interfaces. Multidestination classifiers apply to all of the switch interfaces and cannot be applied to individual interfaces. Switches that require the separation of unicast and multidestination traffic have 12 output queues to provide 4 output queues reserved for multidestination traffic.

On switches that do not separate unicast and multidestination traffic, unicast and multidestination traffic use the same classifiers, and you do not create a separate special classifier for multidestination traffic. Switches that do not separate unicast and multidestination traffic have eight output queues because no extra queues are required to separate the traffic.

Forwarding Classes

Forwarding classes group packets for transmission and CoS. You assign each packet to an output queue based on the packet's forwarding class. Forwarding classes affect the forwarding, scheduling, and rewrite marking policies applied to packets as they transit the switch.

Switches provide up to five default forwarding classes:

- best-effort—Best-effort traffic
- fcoe—Fibre Channel over Ethernet traffic
- no-loss—Lossless traffic
- network-control—Network control traffic
- mcast—Multicast traffic

NOTE: The default **mcast** forwarding class applies only to switches that require the separation of unicast and multidestination (multicast, broadcast, and destination lookup fail) traffic. On these switches, you create separate forwarding classes for the two types of traffic. The default mcast forwarding class transports only multidestination traffic, and the default **best-effort**, **fcoe**, **no-loss**, and **network-control** forwarding classes transport only unicast traffic. Unicast forwarding classes map to unicast output queues, and multidestination forwarding classes map to multidestination output queues. You cannot assign unicast traffic and multidestination traffic to the same forwarding class or to the same output queue. Switches that require the separation of unicast and multidestination traffic have 12 output queues, 8 for unicast traffic and 4 for multidestination traffic.

On switches that do not separate unicast and multidestination traffic, unicast and multidestination traffic use the same forwarding classes and output queues, so the mcast forwarding class is not valid. You do not create separate forwarding classes for multidestination traffic. Switches that do not separate unicast and multidestination traffic have eight output queues because no extra queues are required to separate the traffic.

NOTE: On OCX Series switches only, do not map traffic to the default fcoe and no-loss forwarding classes. By default, the DSCP default classifier does not map traffic to the fcoe and no-loss forwarding classes, so by default, OCX Series switches do not classify traffic into those forwarding classes. (On other switches, the fcoe and no-loss forwarding classes provide lossless transport for Layer 2 traffic. OCX Series switches do not support lossless Layer 2 transport.)

Switches support a total of either 12 forwarding classes (8 unicast forwarding classes and 4 multicast forwarding classes), or 8 forwarding classes (unicast and multidestination traffic use the same forwarding classes), which provides flexibility in classifying traffic.

NFX250 Network Services platform provide the following forwarding classes:

- best-effort (be)—Provides no service profile. Loss priority is typically not carried in a CoS value.
- expedited-forwarding (ef)—Provides a low loss, low latency, low jitter, assured bandwidth, end-to-end service.
- assured-forwarding (af)—Provides a group of values you can define and includes four subclasses: AF1, AF2, AF3, and AF4, each with two drop probabilities: low and high.
- network-control (nc)—Supports protocol control and thus is typically high priority.

Forwarding Class Sets

You can group forwarding classes (output queues) into *forwarding class sets* to apply CoS to groups of traffic that require similar treatment. Forwarding class sets map traffic into priority groups to support enhanced transmission selection (ETS), which is described in IEEE 802.1Qaz.

You can configure up to three unicast forwarding class sets and one multicast forwarding class set. For example, you can configure different forwarding class sets to apply CoS to unicast groups of local area network (LAN) traffic, storage area network (SAN) traffic, and high-performance computing (HPC) traffic, and configure another group for multicast traffic.

Within each forwarding class set, you can configure special CoS treatment for the traffic mapped to each individual queue. This provides the ability to configure CoS in a two-tier hierarchical manner. At the forwarding class set tier, you configure CoS for groups of traffic using a *traffic control profile*. At the queue tier, you configure CoS for individual output queues within a forwarding class set using a *scheduler* that you map to a queue (forwarding class) using a *scheduler map*.

Flow Control (Ethernet PAUSE, PFC, and ECN)

Ethernet PAUSE (described in IEEE 802.3X) is a link-level flow control mechanism. During periods of network congestion, Ethernet PAUSE stops all traffic on a full-duplex Ethernet link for a period of time specified in the PAUSE message.

NOTE: QFX10000 switches do not support Ethernet PAUSE.

Priority-based flow control (PFC) is described in IEEE 802.1Qbb as part of the IEEE data center bridging (DCB) specifications for creating a lossless Ethernet environment to transport loss-sensitive flows such as Fibre Channel over Ethernet (FCoE) traffic.

NOTE: OCX Series switches do not support PFC.

PFC is a link-level flow control mechanism similar to Ethernet PAUSE. However, Ethernet PAUSE stops all traffic on a link for a period of time. PFC decouples the pause function from the physical link and divides the traffic on the link into eight priorities (3-bit IEEE 802.1p code points). You can think of the eight priorities as eight “lanes” of traffic. You can apply pause selectively to the traffic on any priority without pausing the traffic on other priorities on the same link.

The granularity that PFC provides allows you to configure different levels of CoS for different types of traffic on the link. You can create lossless lanes for traffic such as FCoE, LAN backup, or management, while using standard frame-drop methods of congestion management for IP traffic on the same link.

NOTE: If you transport FCoE traffic, you must enable PFC on the priority assigned to FCoE traffic (usually IEEE 802.1p code point 011 on interfaces that carry FCoE traffic).

Explicit congestion notification (ECN) enables end-to-end congestion notification between two endpoints on TCP/IP based networks. ECN must be enabled on both endpoints and on all of the intermediate devices between the endpoints for ECN to work properly. Any device in the transmission path that does not support ECN breaks the end-to-end ECN functionality. ECN notifies networks about congestion with the goal of reducing packet loss and delay by making the sending device decrease the transmission rate until the congestion clears, without dropping packets. RFC 3168, *The Addition of Explicit Congestion Notification (ECN) to IP*, defines ECN.

WRED Profiles and Tail Drop

A weighted random early detection (WRED) profile (drop profile) defines parameters that enable the network to drop packets during periods of congestion. A *drop profile* defines the conditions under which packets of different loss priorities drop, by determining the probability of dropping a packet for each loss priority when output queues become congested. Drop profiles essentially set a value for a level of queue fullness—when the queue fills to the level of the queue fullness value, packets drop. The combination of queue fill level, the probability of dropping a packet at that fill level, and loss priority of the packet, determine whether a packet is dropped or forwarded. Each pairing of a fill level with a drop probability creates a point on a drop profile curve.

You can associate different drop profiles with different loss priorities to set the probability of dropping packets. You can apply a drop profile for each loss priority to a forwarding class (output queue) by applying a drop profile to a scheduler, and then mapping the scheduler to a forwarding class using a scheduler map. When the queue mapped to the forwarding class experiences congestion, the drop profile determines the level of packet drop for traffic of each loss priority in that queue.

Loss priority affects the scheduling of a packet without affecting the packet's relative ordering. Typically you mark packets exceeding a particular service level with a high loss priority.

Tail drop is a simple drop mechanism that drops all packets indiscriminately during periods of congestion, without differentiating among the packet loss priorities of traffic flows. Tail drop requires only one curve point that corresponds to the maximum depth of the output queue, and drop probability when traffic exceeds the buffer depth is 100 percent (all packets that cannot be stored in the queue are dropped). WRED is superior to tail-drop because WRED enables you to treat traffic of different priorities in a differentiated manner, so that higher priority traffic receives preference, and because of the ability to set multiple points on the drop curve.

Schedulers

Each switch interface has multiple queues assigned to store packets. The switch determines which queue to service based on a particular method of scheduling. This process often involves determining the sequence in which different types of packets should be transmitted.

You can define the scheduling priority (**priority**), minimum guaranteed bandwidth (**transmit-rate**), maximum bandwidth (**shaping-rate**), and WRED profiles to be applied to a particular queue (forwarding class) for packet transmission. By default, extra bandwidth is shared among queues in proportion to the minimum guaranteed bandwidth of each queue. On switches that support the **excess-rate** statement, you can configure the percentage of shared extra bandwidth an output queue receives independently from the minimum guaranteed bandwidth transmit rate, or you can use default bandwidth sharing based on the transmit rate.

A scheduler map associates a specified forwarding class with a scheduler configuration. You can associate up to four user-defined scheduler maps with the interfaces.

Rewrite Rules

A *rewrite rule* sets the appropriate CoS bits in the outgoing packet. This allows the next downstream device to classify the packet into the appropriate service group. Rewriting (marking) outbound packets is useful when the switch is at the border of a network and must change the CoS values to meet the policies of the targeted peer.

NOTE: Ingress firewall filters can also rewrite forwarding class and loss priority values.

Related Documentation

- Understanding CoS Packet Flow on page 23

Assigning CoS Components to Interfaces

After you define the following CoS components, you assign them to physical or logical interfaces. Components that you assign to physical interfaces are valid for all of the logical interfaces configured on the physical interface. Components that you assign to a logical interface are valid only for that logical interface.

- Classifiers—Assign only to logical interfaces; on some switches, you can apply classifiers to physical Layer 3 interfaces and the classifiers are applied to all logical interfaces on the physical interface.
- Congestion notification profiles—Assign only to physical interfaces.

NOTE: **OCX Series switches and NFX250 Network Services platform do not support congestion notification profiles.**

- Forwarding classes—Assign to interfaces by mapping to forwarding class sets.
- Forwarding class sets—Assign only to physical interfaces.
- Output traffic control profiles—Assign only to physical interfaces (with a forwarding class set).
- Port schedulers—Assign only to physical interfaces on switches that support port scheduling. Associate the scheduler with a forwarding class in a scheduler map and apply the scheduler map to the physical interface.
- Rewrite rules—Assign only to logical interfaces; on some switches, you can apply classifiers to physical Layer 3 interfaces and the classifiers are applied to all logical interfaces on the physical interface.

You can assign a CoS component to a single interface or to multiple interfaces using wildcards. You can also assign a congestion notification profile or a forwarding class set globally to all interfaces.

To assign CoS components to interfaces:

Assign a CoS component to a physical interface by associating a CoS component (for example, a forwarding class set named **be-priority-group**) with an interface:

```
[edit class-of-service interfaces]
user@switch# set xe-0/0/7 forwarding-class-set be-priority-group
```

Assign a CoS component to a logical interface by associating a CoS component (for example, a classifier named **be_classifier**) with a logical interface:

```
[edit class-of-service interfaces]
user@switch# set xe-0/0/7 unit 0 classifiers dscp be_classifier
```

Assign a CoS component to multiple interfaces by associating a CoS component (for example, a rewrite rule named **customup-rw**) to all 10-Gigabit Ethernet interfaces on the switch, use wildcard characters for the interface name and logical interface (unit) number:

```
[edit class-of-service interfaces]
user@switch# set xe-* unit * rewrite-rules ieee-802.1 customup-rw
```

Assign a congestion notification profile or a forwarding class set globally to all interfaces using the **set class-of-service interfaces all** statement. For example, to assign a forwarding class set named **be-priority-group** to all interfaces:

```
[edit class-of-service interfaces]
user@switch# set all forwarding-class-set be-priority-group
```

NOTE: If there is an existing CoS configuration of any type on an interface, the global configuration is not applied to that particular interface. The global configuration is applied to all interfaces that do not have an existing CoS configuration.

For example, if you configure a rewrite rule, assign it to interfaces xe-0/0/20.0 and xe-0/0/22.0, and then configure a forwarding class set and apply it to all interfaces, the forwarding class set is applied to every interface except xe-0/0/20 and xe-0/0/22.

Related Documentation

- Monitoring Interfaces That Have CoS Components on page 55
- Understanding Junos CoS Components on page 15

Understanding CoS Packet Flow

When a packet traverses a switch, the switch provides the appropriate level of service to the packet using either default class-of-service (CoS) settings or CoS settings that you configure. On ingress ports, the switch classifies packets into appropriate forwarding classes and assigns a loss priority to the packets. On egress ports, the switch applies packet scheduling and (if you have configured them) rewrite rules to re-mark packets.

You can configure CoS on Layer 2 logical interfaces, and you can configure CoS on Layer 3 physical interfaces if you have defined at least one logical interface on the Layer 3 physical interface. You cannot configure CoS on Layer 2 physical interfaces and Layer 3 logical interfaces.

For Layer 2 traffic, either use the default CoS settings or configure CoS on each logical interface. You can apply different CoS settings to different Layer 2 logical interfaces.

NOTE: OCX Series switches do not support Layer 2 interfaces (family ethernet-switching).

For Layer 3 traffic, either use the default CoS settings or configure CoS on the physical interface (not on the logical unit). The switch uses the CoS applied on the physical Layer 3 interface for all logical Layer 3 interfaces configured on the physical Layer 3 interface.

The switch applies CoS to packets as they flow through the system:

- An interface has one or more classifiers of different types applied to it (configure this at the **[edit class-of-service interfaces]** hierarchy level). The classifier types are based on the portion of the incoming packet that the classifier examines (IEEE 802.1p code point bits or DSCP code point bits).
- When a packet enters an ingress port, the classifier assigns the packet to a forwarding class and a loss priority based on the code point bits of the packet (configure this at the **[edit class-of-service classifiers]** hierarchy level).
- The switch assigns each forwarding class to an output queue (configure this at the **[edit class-of-service forwarding-classes]** hierarchy level).
- Input (and output) policers meter traffic and can change the forwarding class and loss priority if a traffic flow exceeds its service level.
- A scheduler map is applied to each interface. When a packet exits an egress port, the scheduler map controls how it is treated (configure this at the **[edit class-of-service interfaces]** hierarchy level). A scheduler map assigns schedulers to forwarding classes (configure this at the **[edit class-of-service scheduler-maps]** hierarchy level).
- A scheduler defines how traffic is treated at the egress interface output queue (configure this at the **[edit class-of-service schedulers]** hierarchy level). You control the transmit rate, shaping rate, priority, and drop profile of each forwarding class by mapping schedulers to forwarding classes in scheduler maps, then applying scheduler maps to interfaces.

- A drop-profile defines how aggressively to drop packets that are mapped to a particular scheduler (configure this at the **[edit class-of-service drop-profiles]** hierarchy level).
- A rewrite rule takes effect as the packet leaves an interface that has a rewrite rule configured (configure this at the **[edit class-of-service rewrite-rules]** hierarchy level). The rewrite rule writes information to the packet (for example, a rewrite rule can re-mark the code point bits of outgoing traffic) according to the forwarding class and loss priority of the packet.

Figure 2 on page 24 is a high-level flow diagram of how packets from various sources enter switch interfaces, are classified at the ingress, and then scheduled (provided bandwidth) at the egress queues.

Figure 2: CoS Classifier, Queues, and Scheduler

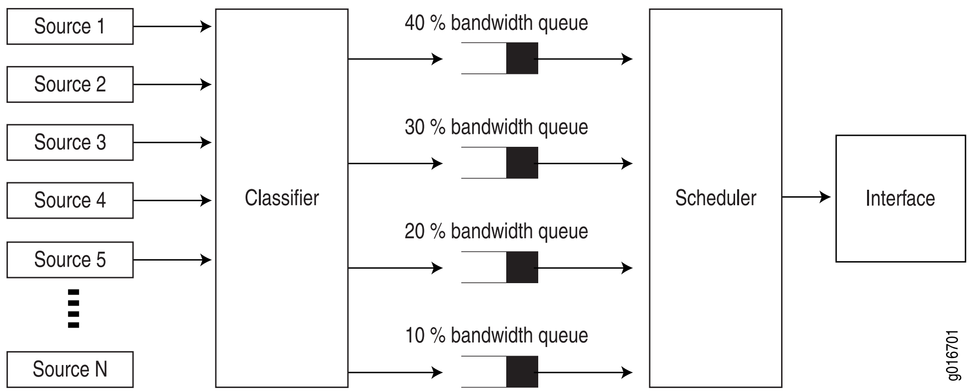
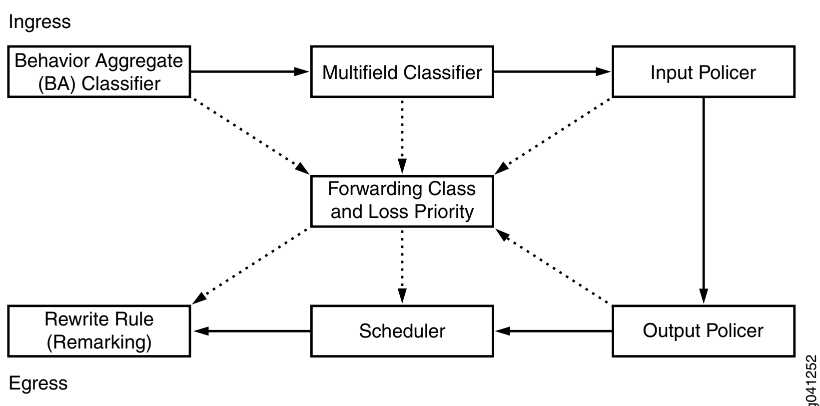


Figure 3 on page 24 shows the packet flow through the CoS components that you can configure.

Figure 3: Packet Flow Through Configurable CoS Components



The middle box (Forwarding Class and Loss Priority) represents two values that you can use on ingress and egress interfaces. The system uses these values for classifying traffic on ingress interfaces and for rewrite rule re-marking on egress interfaces. Each outer box represents a process component. The components in the top row apply to incoming packets. The components in the bottom row apply to outgoing packets.

The solid-line arrows show the direction of packet flow from ingress to egress. The dotted-line arrows that point to the forwarding class and loss priority box indicate processes that configure (set) the forwarding class and loss priority. The dotted-line arrows that point away from the forwarding class and loss priority box indicate processes that use forwarding class and loss priority as input values on which to base actions.

For example, the BA classifier sets the forwarding class and loss priority of incoming packets, so the forwarding class and loss priority are outputs of the classifier and the arrow points away from the classifier. The scheduler receives the forwarding class and loss priority settings, and queues the outgoing packets based on those settings, so the arrow points toward the scheduler.

Understanding Default CoS Settings

If you do not configure CoS settings, Junos OS performs some CoS functions to ensure that traffic and protocol packets are forwarded with minimum delay when the network experiences congestion. Some default mappings are automatically applied to each logical interface that you configure.

You can display default CoS settings by issuing the **show class-of-service** operational mode command.

This topic describes the default configurations for the following CoS components:

Default Forwarding Classes and Queue Mapping

Table 5 on page 25 shows the default mapping of the default forwarding classes to queues and packet drop attributes.

Table 5: Default Forwarding Classes and Queue Mapping

Default Forwarding Class	Description	Default Queue Mapping	Packet Drop Attribute
best-effort (be)	Best-effort traffic class (priority 0, IEEE 802.1p code point 000)	0	drop

Table 5: Default Forwarding Classes and Queue Mapping (continued)

Default Forwarding Class	Description	Default Queue Mapping	Packet Drop Attribute
fcoe	Guaranteed delivery for FCoE traffic (priority 3, IEEE 802.1p code point 011) NOTE: Do not map traffic to the default fcoe forwarding class. Other switches use the fcoe forwarding class for Fibre Channel over Ethernet (FCoE) lossless Layer 2 transport, but OCX Series switches do not support FCoE or lossless Layer 2 transport.	3	no-loss NOTE: The no-loss drop attribute is not supported on OCX Series switches. If you want to use queue 3, configure a new forwarding class *without* the no-loss packet drop attribute and map it to queue 3.
no-loss	Guaranteed delivery for TCP no-loss traffic (priority 4, IEEE 802.1p code point 100) NOTE: Do not map traffic to the default no-loss forwarding class. Other switches use the no-loss forwarding class for lossless Layer 2 transport, but OCX Series switches do not support lossless Layer 2 transport.	4	no-loss NOTE: The no-loss drop attribute is not supported on OCX Series switches. If you want to use queue 4, configure a new forwarding class *without* the no-loss packet drop attribute and map it to queue 4.
network-control (nc)	Network control traffic (priority 7, IEEE 802.1p code point 111)	7	drop
mcast	Multidestination traffic	8	drop

NOTE: On OCX Series switches only, do not map traffic to the default fcoe and no-loss forwarding classes. By default, the DSCP classifier does not map traffic to the fcoe and no-loss forwarding classes, so by default, the OCX Series does not classify traffic into those forwarding classes. (On other switches, the fcoe and no-loss forwarding classes provide lossless transport for Layer 2 traffic. OCX Series switches do not support lossless Layer 2 transport.)

Do not configure the no-loss packet drop attribute on a forwarding class, and do not map traffic to forwarding classes that have the no-loss attribute.

Default Forwarding Class Sets (Priority Groups)

If you do not explicitly configure forwarding class sets, the system automatically creates a default forwarding class set that contains all of the forwarding classes on the switch.

The system assigns 100 percent of the port output bandwidth to the default forwarding class set.

Ingress traffic is classified based on the default classifier settings. The forwarding classes (queues) in the default forwarding class set receive bandwidth based on the default scheduler settings. Forwarding classes that are not part of the default scheduler receive no bandwidth.

The default forwarding class set is transparent. It does not appear in the configuration and is used for Data Center Bridging Capability Exchange (DCBX) protocol advertisement.

Default Code-Point Aliases

Table 6 on page 27 shows the default mapping of code-point aliases to IEEE code points.

Table 6: Default IEEE 802.1 Code-Point Aliases

CoS Value Types	Mapping
be	000
be1	001
ef	010
ef1	011
af11	100
af12	101
nc1	110
nc2	111

Table 7 on page 27 shows the default mapping of code-point aliases to DSCP and DSCP IPv6 code points.

Table 7: Default DSCP and DCSP IPv6 Code-Point Aliases

CoS Value Types	Mapping
ef	101110
af11	001010
af12	001100
af13	001110
af21	010010

Table 7: Default DSCP and DCSP IPv6 Code-Point Aliases (continued)

CoS Value Types	Mapping
af22	010100
af23	010110
af31	011010
af32	011100
af33	011110
af41	100010
af42	100100
af43	100110
be	000000
cs1	001000
cs2	010000
cs3	011000
cs4	100000
cs5	101000
nc1	110000
nc2	111000

Default Classifiers

The switch applies default classifiers to each interface that does not have explicitly configured classifiers. If you explicitly configure one type of classifier but not other types of classifiers, the system uses only the configured classifier and does not use default classifiers for other types of traffic.

Table 8 on page 29 shows the default mapping of DSCP code-point values to unicast forwarding classes and loss priorities for DSCP IPv4 and DCSP IPv6. This is the default classifier used for Layer 3 IP traffic.

Table 8: Default DSCP IPv4 and IPv6 Unicast Classifiers

Code Point	Forwarding Class	Loss Priority
000000 (be)	best-effort	low
000001	best-effort	low
000010	best-effort	low
000011	best-effort	low
000100	best-effort	low
000101	best-effort	low
000110	best-effort	low
000111	best-effort	low
001000 (cs1)	best-effort	low
001001	best-effort	low
001010 (af11)	best-effort	low
001011	best-effort	low
001100 (af12)	best-effort	low
001101	best-effort	low
001110 (af13)	best-effort	low
001111	best-effort	low
010000 (cs2)	best-effort	low
010001	best-effort	low
010010 (af21)	best-effort	low
010011	best-effort	low
010100 (af22)	best-effort	low
010101	best-effort	low
010110 (af23)	best-effort	low
010111	best-effort	low

Table 8: Default DSCP IPv4 and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
011000 (cs3)	best-effort	low
011001	best-effort	low
011010 (af31)	best-effort	low
011011	best-effort	low
011100 (af32)	best-effort	low
011101	best-effort	low
011110 (af33)	best-effort	low
011111	best-effort	low
100000 (cs4)	best-effort	low
100001	best-effort	low
100010 (af41)	best-effort	low
100011	best-effort	low
100100 (af42)	best-effort	low
100101	best-effort	low
100110 (af43)	best-effort	low
100111	best-effort	low
101000 (cs5)	best-effort	low
101001	best-effort	low
101011	best-effort	low
101100	best-effort	low
101101	best-effort	low
101110 (ef)	best-effort	low
101111	best-effort	low
110000 (nc1)	network-control	low

Table 8: Default DSCP IPv4 and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
110001	network-control	low
110010	network-control	low
110011	network-control	low
110100	network-control	low
110101	network-control	low
110110	network-control	low
110111	network-control	low
111000 (nc2)	network-control	low
111001	network-control	low
111010	network-control	low
111011	network-control	low
111100	network-control	low
111101	network-control	low
111110	network-control	low
111111	network-control	low

NOTE: There are no default DSCP IPv4 or IPv6 classifiers for multidestination traffic. DSCP IPv6 classifiers are not supported for multidestination traffic.

Table 9 on page 31 shows the DSCP IPv6 compatibility default classifier mappings.

Table 9: Default DSCP IPv6 Compatibility Classifiers

Code Point	Forwarding Class	Loss Priority
000000 (be)	best-effort	low
000001	best-effort	low
000010	best-effort	low

Table 9: Default DSCP IPv6 Compatibility Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
000011	best-effort	low
000100	best-effort	low
000101	best-effort	low
000110	best-effort	low
000111	best-effort	low
001000 (cs1)	best-effort	low
001001	best-effort	low
001010 (af11)	best-effort	low
001011	best-effort	low
001100 (af12)	best-effort	low
001101	best-effort	low
001110 (af13)	best-effort	low
001111	best-effort	low
010000 (cs2)	best-effort	low
010001	best-effort	low
010010 (af21)	best-effort	low
010011	best-effort	low
010100 (af22)	best-effort	low
010101	best-effort	low
010110 (af23)	best-effort	low
010111	best-effort	low
011000 (cs3)	best-effort	low
011001	best-effort	low
011010 (af31)	best-effort	low

Table 9: Default DSCP IPv6 Compatibility Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
011011	best-effort	low
011100 (af32)	best-effort	low
011101	best-effort	low
011110 (af33)	best-effort	low
011111	best-effort	low
100000 (cs4)	best-effort	low
100001	best-effort	low
100010 (af41)	best-effort	low
100011	best-effort	low
100100 (af42)	best-effort	low
100101	best-effort	low
100110 (af43)	best-effort	low
100111	best-effort	low
101000 (cs5)	best-effort	low
101001	best-effort	low
101011	best-effort	low
101100	best-effort	low
101101	best-effort	low
101110 (ef)	best-effort	low
101111	best-effort	low
110000 (nc1)	network-control	low
110001	best-effort	low
110010	best-effort	low
110011	best-effort	low

Table 9: Default DSCP IPv6 Compatibility Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
110100	best-effort	low
110101	best-effort	low
110110	best-effort	low
110111	best-effort	low
111000 (nc2)	network-control	low
111001	best-effort	low
111010	best-effort	low
111011	best-effort	low
111100	best-effort	low
111101	best-effort	low
111110	best-effort	low
111111	best-effort	low

Table 10 on page 34 shows the default mapping of IEEE 802.1 code-point values to unicast forwarding classes and loss priorities.

Table 10: Default IEEE 802.1 Unicast Trusted Classifiers

Code Point	Forwarding Class	Loss Priority
be (000)	best-effort	low
be1 (001)	best-effort	low
ef (010)	best-effort	low
ef1 (011)	fcoe	low
af11 (100)	no-loss	low
af12 (101)	best-effort	low
nc1 (110)	network-control	low
nc2 (111)	network-control	low

Table 11 on page 35 shows the default mapping of IEEE 802.1p code-point values to unicast forwarding classes and loss priorities for untrusted ports.

Table 11: Default IEEE 802.1 Unicast Untrusted Classifiers

Code Point	Forwarding Class	Loss Priority
000	best-effort	low
001	best-effort	low
010	best-effort	low
011	best-effort	low
100	best-effort	low
101	best-effort	low
110	best-effort	low
111	best-effort	low

Table 12 on page 35 shows the default mapping of IEEE 802.1 code-point values to multidestination (multicast, broadcast, and destination lookup fail traffic) forwarding classes and loss priorities.

Table 12: Default IEEE 802.1 Multidestination Classifiers

Code Point	Forwarding Class	Loss Priority
be (000)	mcast	low
be1 (001)	mcast	low
ef (010)	mcast	low
ef1 (011)	mcast	low
af11 (100)	mcast	low
af12 (101)	mcast	low
nc1 (110)	mcast	low
nc2 (111)	mcast	low

Default Rewrite Rules

There are no default rewrite rules. If you do not explicitly configure rewrite rules, the switch does not reclassify egress traffic.

Default Drop Profile

Table 13 on page 36 shows the default drop profile configuration.

Table 13: Default Drop Profile

Fill Level	Drop Probability
100	100

Default Schedulers

Table 14 on page 36 shows the default scheduler configuration.

Table 14: Default Schedulers

Default Scheduler and Queue Number	Transmit Rate (Guaranteed Minimum Bandwidth)	Shaping Rate (Maximum Bandwidth)	Excess Bandwidth Sharing	Priority	Buffer Size
best-effort unicast forwarding class scheduler (queue 0)	5%	None	5%	low	5%
fcoe unicast forwarding class scheduler (queue 3)	35%	None	35%	low	35%
no-loss unicast forwarding class scheduler (queue 4)	35%	None	35%	low	35%
network-control unicast forwarding class scheduler (queue 7)	5%	None	5%	low	5%
mcast forwarding class scheduler (queue 8)	20%	None	20%	low	20%

NOTE: **The minimum guaranteed bandwidth (transmit rate) also determines the amount of excess (extra) bandwidth that the queue can share. Extra bandwidth is allocated to queues in proportion to the transmit rate of each queue.**

By default, only the best-effort and network-control unicast forwarding classes show in Table 14 on page 36 have traffic mapped to them. Even though the default schedulers allocate 35 percent of the port bandwidth to the fcoe and no-loss forwarding classes, by default, no traffic is mapped (classified) to the fcoe and no-loss forwarding classes. Because the fcoe and no-loss forwarding classes carry no traffic, the bandwidth that the default schedulers assign to them can be used for the other forwarding classes.

Only the queues associated with the default schedulers receive default bandwidth, which is based on the default scheduler transmit rate. (You can configure schedulers and forwarding classes to allocate bandwidth to other queues or to change the default bandwidth of a default queue.) In addition, multidestination queue 11 receives enough bandwidth from the default multidestination scheduler to handle CPU-generated multidestination traffic.

Default hierarchical scheduling divides the total port bandwidth between two groups of traffic: unicast traffic and multidestination traffic. Unicast traffic receives and shares a total of 80 percent of the port bandwidth. By default, multidestination traffic (**mcast** queue 8) receives a total of 20 percent of the port bandwidth.

NOTE: **Multidestination queue 11 also receives a small amount of default bandwidth from the multidestination scheduler. CPU-generated multidestination traffic uses queue 11, so you might see a small number of packets egress from queue 11. In addition, in the unlikely case that firewall filter match conditions map multidestination traffic to a unicast forwarding class, that traffic uses queue 11.**

Default scheduling uses weighted round-robin (WRR) scheduling. Each queue receives a portion (weight) of the total available interface bandwidth. The scheduling weight is based on the transmit rate of the default scheduler for that queue. For example, queue 7 receives a default scheduling weight of 5 percent of the available bandwidth, and queue 4 receives a default scheduling weight of 35 percent of the available bandwidth. Queues are mapped to forwarding classes, so forwarding classes receive the default bandwidth for the queues to which they are mapped. Unused bandwidth is shared with other default queues.

If you want non-default (unconfigured) queues to forward traffic, you should explicitly map traffic to those queues (configure the forwarding classes and queue mapping) and create schedulers to allocate bandwidth to those queues. By default, unicast queues 1, 2, 5, and 6 are unconfigured, and multidestination queues 9, 10, and 11 are unconfigured. Unconfigured queues have a default scheduling weight of 1 so that they can receive a small amount of bandwidth in case they need to forward traffic. (However, queue 11 can use more of the default multidestination scheduler bandwidth if necessary to handle CPU-generated multidestination traffic.)

NOTE: All four multidestination queues have a scheduling weight of 1. Because by default multidestination traffic goes to queue 8, queue 8 receives almost all of the multidestination bandwidth. (There is no default traffic on queue 9 and queue 10, and very little default traffic on queue 11, so there is almost no competition for multidestination bandwidth.)

However, if you explicitly configure queue 9, 10, or 11 (by mapping code points to the unconfigured multidestination forwarding classes using the multidestination classifier), the explicitly configured queues share the multidestination scheduler bandwidth equally with default queue 8, because all of the queues have the same scheduling weight (1). To ensure that multidestination bandwidth is allocated to each queue properly and that the bandwidth allocation to the default queue (8) is not reduced too much, we strongly recommend that you configure a scheduler if you explicitly classify traffic into queue 9, 10, or 11.

If you map traffic to an unconfigured queue, the queue receives only the amount of group bandwidth proportional to its default weight (1). The actual amount of bandwidth an unconfigured queue receives depends on how much bandwidth the other queues in the group are using.

If the other unicast queues use less than their allocated amount of bandwidth, the unconfigured queues can share the unused bandwidth. Sharing unused bandwidth is one of the key advantages of hierarchical port scheduling. Configured queues have higher priority for bandwidth than unconfigured queues, so if a configured queue needs more bandwidth, then less bandwidth is available for unconfigured queues. Unconfigured queues always receive a minimum amount of bandwidth based on their scheduling weight (1). If you map traffic to an unconfigured queue, to allocate bandwidth to that queue, configure a scheduler for the forwarding class that is mapped to the queue.

Default Scheduler Maps

Table 15 on page 38 shows the default mapping of forwarding classes to schedulers.

Table 15: Default Scheduler Maps

Forwarding Class	Scheduler
best-effort	Default BE scheduler
fcoe	Default FCoE scheduler NOTE: OCX Series switches do not support FCoE traffic. By default, no traffic is mapped to this forwarding class. Do not map traffic to this default forwarding class.
no-loss	No-loss scheduler NOTE: OCX Series switches do not support lossless Layer 2 traffic. By default, no traffic is mapped to this forwarding class. Do not map traffic to this default forwarding class.
network-control	Default network-control scheduler

Table 15: Default Scheduler Maps (continued)

Forwarding Class	Scheduler
mcast-be	Default multidestination scheduler

Default Shared Buffer Configuration

Table Table 16 on page 39 and Table 17 on page 39 show the default shared buffer allocations:

Table 16: Default Ingress Shared Buffer Configuration

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
100%	9%	45%	46%

Table 17: Default Egress Shared Buffer Configuration

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	50%	31%	19%

Related Documentation

- Overview of Junos OS CoS on page 8
- Understanding Junos CoS Components on page 15
- Understanding Default CoS Scheduling and Classification on page 114
- Understanding CoS Classifiers on page 123
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162
- Understanding CoS Code-Point Aliases on page 134
- Understanding CoS Forwarding Classes on page 137
- Understanding CoS Rewrite Rules on page 157
- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS WRED Drop Profiles on page 362

CoS Inputs and Outputs Overview

Some CoS components map one set of values to another set of values. Each mapping contains one or more inputs and one or more outputs. When you configure a mapping, you set the outputs for a given set of inputs, as shown in Table 18 on page 40.

Table 18: CoS Mappings—Inputs and Outputs

CoS Mappings	Inputs	Outputs	Comments
classifiers	code-points	forwarding-class, loss-priority	The map sets the forwarding class and packet loss priority (PLP) for a specific set of code points.
drop-profile-map	**loss-priority**, protocol	drop-profile	The map sets the drop profile for a specific PLP and protocol type.
rewrite-rules	loss-priority, forwarding-class	code-points	The map sets the code points for a specific forwarding class and PLP.
rewrite-value (Fibre Channel Interfaces)	*forwarding-class*	*code-point*	(Systems that support native Fibre Channel interfaces only) The map sets the code point for the forwarding class specified in the fixed classifier attached to the native Fibre Channel (NP_Port) interface.

Related Documentation

- Understanding CoS Packet Flow on page 23

Overview of Policers

A switch polices traffic by limiting the input or output transmission rate of a class of traffic according to user-defined criteria. Policing (or rate-limiting) traffic allows you to control the maximum rate of traffic sent or received on an interface and to provide multiple priority levels or classes of service.

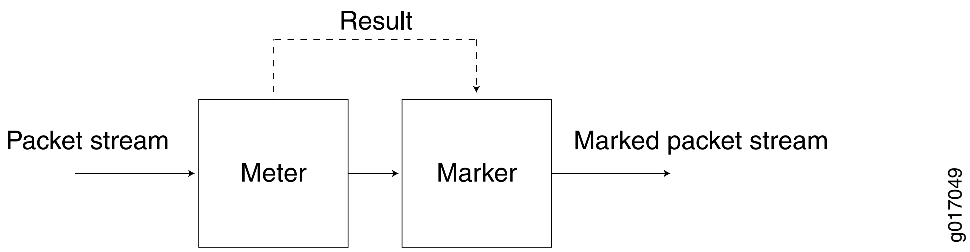
Policing is also an important component of firewall filters. You can achieve policing by including policers in firewall filter configurations.

Policer Overview

You use policers to apply limits to traffic flow and set consequences for packets that exceed these limits—usually applying a higher loss priority—so that if packets encounter downstream congestion, they can be discarded first. Policers apply only to unicast packets.

Policers provide two functions: metering and marking. A policer meters (measures) each packet against traffic rates and burst sizes that you configure. It then passes the packet and the metering result to the marker, which assigns a packet loss priority that corresponds to the metering result. Figure 4 on page 41 illustrates this process.

Figure 4: Flow of Tricolor Marking Policer Operation



After you name and configure a policer, you can use it by specifying it as an action in one or more firewall filters.

Policer Types

A switch supports three types of policers:

- Single-rate two-color marker—A two-color policer (or "policer" when used without qualification) meters the traffic stream and classifies packets into two categories of packet loss priority (PLP) according to a configured bandwidth and burst-size limit. You can mark packets that exceed the bandwidth and burst-size limit with a specified PLP or simply discard them.

  You can specify this type of policer in an ingress or egress firewall.

  NOTE: **A two-color policer is most useful for metering traffic at the port (physical interface) level.**

- Single-rate three-color marker—This type of policer is defined in RFC 2697, *A Single Rate Three Color Marker*, as part of an assured forwarding (AF) per-hop-behavior (PHB) classification system for a Differentiated Services (DiffServ) environment. This type of policer meters traffic based on one rate—the configured committed information rate (CIR) as well as the committed burst size (CBS) and the excess burst size (EBS). The CIR specifies the average rate at which bits are admitted to the switch. The CBS specifies the usual burst size in bytes and the EBS specifies the maximum burst size in bytes. The EBS must be greater than or equal to the CBS, and neither can be 0.

  You can specify this type of policer in an ingress or egress firewall.

  NOTE: **A single-rate three-color marker (TCM) is most useful when a service is structured according to packet length and not peak arrival rate.**

- Two-rate three-color marker—This type of policer is defined in RFC 2698, *A Two Rate Three Color Marker*, as part of an assured forwarding per-hop-behavior classification system for a Differentiated Services environment. This type of policer meters traffic based on two rates—the CIR and peak information rate (PIR) along with their associated burst sizes, the CBS and peak burst size (PBS). The PIR specifies the maximum rate at which bits are admitted to the network and must be greater than or equal to the CIR.

You can specify this type of policer in an ingress or egress firewall.

NOTE: A two-rate three-color policer is most useful when a service is structured according to arrival rates and not necessarily packet length.

See Table 19 on page 42 for information about how metering results are applied for each of these policer types.

Policer Actions

Policer actions are implicit or explicit and vary by policer type. *Implicit* means that Junos OS assigns the loss priority automatically. Table 19 on page 42 describes the policer actions.

Table 19: Policer Actions

Policer	Marking	Implicit Action	Configurable Action
Single-rate two-color	Green (conforming)	Assign low loss priority	None
	Red (nonconforming)	None	Discard
Single-rate three-color	Green (conforming)	Assign low loss priority	None
	Yellow (above the CIR and CBS)	Assign medium-high loss priority	None
	Red (above the EBS)	Assign high loss priority	Discard
Two-rate three-color	Green (conforming)	Assign low loss priority	None
	Yellow (above the CIR and CBS)	Assign medium-high loss priority	None
	Red (above the PIR and PBS)	Assign high loss priority	Discard

NOTE: If you specify a policer in an egress firewall filter, the only supported action is discard.

Policer Colors

Single-rate and two-rate three-color policers can operate in two modes:

- Color-blind—In color-blind mode, the three-color policer assumes that all packets examined have not been previously marked or metered. In other words, the three-color policer is “blind” to any previous coloring a packet might have had.
- Color-aware—In color-aware mode, the three-color policer assumes that all packets examined have been previously marked or metered. In other words, the three-color policer is “aware” of the previous coloring a packet might have had. In color-aware mode, the three-color policer can increase the PLP of a packet but cannot decrease it. For example, if a color-aware three-color policer meters a packet with a medium PLP marking, it can raise the PLP level to high but cannot reduce the PLP level to low.

Filter-Specific Policers

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 on some QFX switches, 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 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. (This behavior does not occur in QFX10000 switches.)

To prevent this unexpected behavior from occurring, use the information about TCAM slices presented in *Understanding How Many Firewall Filters Are Supported* 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.

Suggested Naming Convention for Policers

We recommend that you use the naming convention ***policertypeTCM#-color type*** when configuring three-color policers and ***policer#*** when configuring two-color policers. TCM stands for three-color marker. Because policers can be numerous and must be applied correctly to work, a simple naming convention makes it easier to apply the policers properly. For example, the first single-rate, color-aware three-color policer configured would be named **srTCM1-ca**. The second two-rate, color-blind three-color configured would be named **trTCM2-cb**. The elements of this naming convention are explained below:

- sr (single-rate)
- tr (two-rate)
- TCM (tricolor marking)

- 1 or 2 (number of marker)
- ca (color-aware)
- cb (color-blind)

Policer Counters

On some QFX switches, each policer that you configure includes an implicit counter that counts the number of packets that exceed the rate limits that are specified for the policer. If you use the same policer in multiple terms—either within the same filter or in different filters—the implicit counter counts all the packets that are policed in all of these terms and provides the total amount. (This does not apply to QFX10000 switches.) If you want to obtain separate packet counts for each term on an affected switch, use these options:

- Configure a unique policer for each term.
- Configure only one policer, but use a unique, explicit counter in each term.

Policer Algorithms

Policing uses the *token-bucket algorithm*, which enforces a limit on average bandwidth while allowing bursts up to a specified maximum value. It offers more flexibility than the *leaky bucket algorithm* in allowing a certain amount of bursty traffic before it starts discarding packets.

NOTE: In an environment of light bursty traffic, QFX5200 might not replicate all multicast packets to two or more downstream interfaces. This occurs only at a line rate burst—if traffic is consistent, the issue does not occur. In addition, the issue occurs only when packet size increases beyond 6k in a one gigabit traffic flow.

How Many Policers Are Supported?

QFX10000 switches support 8K policers (all policer types). QFX5100 and QFX5200 switches support 1535 ingress policers and 1024 egress policers (assuming one policer per firewall filter term). QFX5110 switches support 6144 ingress policers and 1024 egress policers (assuming one policer per firewall filter term).

QFX3500 and QFX3600 standalone switches and QFabric Node devices support the following numbers of policers (assuming one policer per firewall filter term):

- Two-color policers used in ingress firewall filters: 767
- Three-color policers used in ingress firewall filters: 767
- Two-color policers used in egress firewall filters: 1022
- Three-color policers used in egress firewall filters: 512

Policers Can Limit Egress Firewall Filters

On some switches, the number of egress policers that you configure can affect the total number of allowed egress firewall filters. (This issue does not affect QFX10000 switches.) On the affected switches, 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.)

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 Color-Blind Mode for Single-Rate Tricolor Marking*
- *Understanding Color-Blind Mode for Two-Rate Tricolor Marking*
- *Understanding Color-Aware Mode for Single-Rate Tricolor Marking*

- *Understanding Color-Aware Mode for Two-Rate Tricolor Marking*
- *Configuring Two-Color and Three-Color Policers to Control Traffic Rates*

CHAPTER 2

Configuration Statements for Basic Concepts

class-of-service

Syntax

```
class-of-service {
  classifiers {
    (dscp | dscp-ipv6 | ieee-802.1 | exp) classifier-name {
      import (classifier-name | default);
      forwarding-class class-name {
        loss-priority level {
          code-points [ aliases ] [ bit-patterns ];
        }
      }
    }
  }
  code-point-aliases {
    (dscp| dscp-ipv6 | ieee-802.1) {
      alias-name bits;
    }
  }
  congestion-notification-profile profile-name {
    input {
      (dscp | ieee-802.1) {
        code-point [code-point-bits] {
          pfc {
            mru mru-value;
          }
        }
      }
      cable-length cable-length-value;
    }
    output {
      ieee-802.1 {
        code-point [code-point-bits] {
          flow-control-queue [queue | list-of-queues];
        }
      }
    }
  }
  drop-profiles {
    profile-name {
      interpolate {
        fill-level low-value fill-level high-value drop-probability 0 drop-probability high-value;
      }
    }
  }
  forwarding-class class-name {
    loss-priority level {
      code-points [ aliases ] [ bit-patterns ];
    }
  }
  forwarding-class class-name {
    scheduler scheduler-name;
  }
  forwarding-class-sets forwarding-class-set-name {
    class class-name;
```

```
}
forwarding-classes {
  class {
    class-name {
      pfc-priority pfc-priority;
      queue-num queue-number <no-loss>;
    }
  }
}
host-outbound-traffic{
  forwarding-class class-name;
  dscp-code-point code-point;
}
interfaces {
  interface-name  {
    congestion-notification-profile profile-name {
    }
    forwarding-class lossless-forwarding-class-name;
    forwarding-class-set forwarding-class-set-name {
      output-traffic-control-profile profile-name;
    }
    rewrite-value {
      input {
        ieee-802.1{
          code-point code-point-bits;
        }
      }
    }
    scheduler-map scheduler-map-name
    unit logical-unit-number {
      classifiers {
        (dscp | dscp-ipv6 | ieee-802.1 | exp) (classifier-name | default);
      }
      forwarding-class class-name;
      rewrite-rules {
        (dscp | dscp-ipv6 | ieee-802.1 | exp) (classifier-name | default);
      }
    }
  }
}
multi-destination {
  classifiers {
    (dscp | ieee-802.1) classifier-name;
  }
}
rewrite-rules {
  (dscp | dscp-ipv6 | ieee-802.1 | exp) classifier-name {
    import (rewrite-name | default);
    forwarding-class class-name {
      loss-priority priority code-point (alias | bits);
    }
  }
}
scheduler-map-forwarding-class-sets {
  fabric-scheduler-map-name {
    forwarding-class-set fabric-forwarding-class-set-name scheduler scheduler-name;
```

```
    }
  }
  scheduler-maps {
    map-name {
      forwarding-class class-name scheduler scheduler-name;
    }
  }
  schedulers {
    scheduler-name  {
      buffer-size (percent percentage | remainder);
      drop-profile-map loss-priority (low | medium-high | high) protocol protocol drop-profile
        drop-profile-name;
      excess-rate percent percentage;
      explicit-congestion-notification;
      priority priority;
      shaping-rate (rate | percent percentage);
      transmit-rate (percent percentage) <exact>;
    }
  }
  shared-buffer {
    egress {
      percent percent;
      buffer-partition (lossless | lossy | multicast) {
        percent percent
      }
    }
    ingress {
      percent percent;
      buffer-partition (lossless | lossless-headroom | lossy) {
        percent percent
      }
    }
  }
  system-defaults {
    classifiers exp classifier-name;
  }
  traffic-control-profiles profile-name {
    guaranteed-rate(rate| percent percentage);
    scheduler-map map-name;
    shaping-rate (rate| percent percentage);
  }
}
```

Hierarchy Level [edit]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

> NOTE: Not all switches support all portions of the class of service hierarchy. For example, some switches use the same classifiers for unicast and multidestination traffic, and those switches do not support the **multi-destination** classifier hierarchy, and some switches do not support shared buffer configuration, and those switches do not support the **shared-buffer** hierarchy.

> NOTE: OCX Series switches do not support MPLS exp classifiers and rewrite rules (including MPLS system defaults), and they do not support congestion notification profiles.

Description

Configure class-of-service parameters on the switch.

The remaining statements are explained separately. Search for a statement in CLI Explorer or click a linked statement in the Syntax section for details.

Default

If you do not configure any CoS features, the default CoS settings are used.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Assigning CoS Components to Interfaces on page 21
- Overview of Junos OS CoS on page 8

traceoptions (Class of Service)

Syntax

```
traceoptions {
  file filename <size size> <files number>
  <world-readable | no-world-readable>;
  flag flag <flag-modifier>;
  no-remote-trace
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Set class-of-service (CoS) tracing options.

NOTE: **The traceoptions statement is not supported on the QFabric system.**

Default Traceoptions is disabled.

Options **file *filename***—Name of the file to receive the tracing operation output. Enclose the name in quotation marks. Traceoption output files are located in the **/var/log/** directory.

files *number*—(Optional) Maximum number of trace files. When a trace file named ***trace-file*** reaches its maximum size, it is renamed ***trace-file*.0**. The traceoption output continues in a second trace file named ***trace-file*.1**. When ***trace-file*.1** reaches its maximum size, output continues in a third file named ***trace-file*.2**, and so on. When the maximum number of trace files is reached, the oldest trace file is overwritten.

If you specify a maximum number of files, you must also specify a maximum file size with the size option.

Range: 2 through 1000 files

Default: 1 trace file

flag—Tracing operation to perform. To specify more than one tracing operation, include multiple **flag** statements:

- **all**—Trace all operations.
- **asynch**—Trace asynchronous configuration processing.
- **chassis-scheduler**—Trace chassis stream scheduler processing.
- **cos-adjustment**—Trace CoS rate adjustments.
- **dynamic**—Trace dynamic CoS functions.

- **hardware-database**—Trace the chassis hardware database related processing.
- **init**—Trace initialization events.
- **performance-monitor**—Trace performance monitor countors.
- **process**—Trace configuration processing.
- **restart**—Trace restart processing.
- **route-socket**—Trace route-socket events.
- **show**—Trace show command servicing.
- **snmp**—Trace SNMP-related processing.
- **util**—Trace utilities.

The following are the global tracing options:

- **all**—Perform all tracing operations
- **parse**—Trace parser processing.

no-remote-trace—(Optional) Disable remote tracing.

no-world-readable—(Optional) Prevent any user from reading the log file.

size *size*—(Optional) Maximum size of each trace file, in kilobytes (KB), megabytes (MB), or gigabytes (GB). When a trace file named ***trace-file*** reaches its maximum size, it is renamed ***trace-file*.0.** Incoming tracefile data is logged in the now empty ***trace-file***. When ***trace-file*** again reaches its maximum size, ***trace-file*.0** is renamed ***trace-file*.1** and ***trace-file*** is renamed ***trace-file*.0**. This renaming scheme continues until the maximum number of trace files is reached. Then the oldest trace file is overwritten.

If you specify a maximum file size, you must also specify a maximum number of trace files with the **files** option.
Syntax: ***x*k** to specify KB, ***x*m** to specify MB, or ***x*g** to specify GB
Range: 10 KB through the maximum file size of 4 GB (maximum is lower if 4 GB is not supported on your system)
Default: 1 MB

world-readable—(Optional) Allow any user to read the log file.

Required Privilege Level

routing—To view this statement in the configuration.
routing-control—To add this statement to the configuration.

CHAPTER 3

Monitoring Commands for Basic Concepts

Monitoring Interfaces That Have CoS Components

Purpose Use the monitoring functionality to display details about the physical and logical interfaces and the CoS components assigned to them.

Action To monitor interfaces that have CoS components in the CLI, enter the command:

user@switch> **show class-of-service interface**

To monitor a specific interface in the CLI, enter the command:

user@switch> **show class-of-service interface** ***interface-name***

Meaning Table 20 on page 55 summarizes key output fields for CoS interfaces.

Table 20: Summary of Key CoS Interfaces Output Fields

Field	Values
Physical interface	Name of a physical interface to which CoS components are assigned.
Index	Index of this interface or the internal index of a specific object.
Queues supported	Number of queues you can configure on the interface.
Queues in use	Number of queues currently configured.
Scheduler map	Name of the scheduler map associated with this interface.

Table 20: Summary of Key CoS Interfaces Output Fields (continued)

Field	Values
Congestion-notification	Status of congestion notification (enabled or disabled). NOTE: OCX Series switches do not support congestion notification profiles.
Rewrite Input IEEE Code-point	(Fibre Channel NP_Port interfaces only) IEEE 802.1p code point (priority) the interface assigns to incoming Fibre Channel (FC) traffic when the interface encapsulates the FC traffic in Ethernet before forwarding it onto the FCoE network.
Logical Interface	Name of a logical interface on the physical interface to which CoS components are assigned.
Object	Category of an object—for example, **classifier**, **scheduler-map**, or **rewrite**.
Name	Name of the object—for example, **ba-classifier**.
Type	Type of the object—for example, **ieee8021p** for a classifier.

Related Documentation

- Assigning CoS Components to Interfaces on page 21

show class-of-service

Syntax **show class-of-service**

Release Information Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Display the class-of-service (CoS) information.

Required Privilege Level view

Related Documentation

- Monitoring CoS Code-Point Value Aliases on page 206
- Monitoring CoS Classifiers on page 201
- Monitoring CoS Forwarding Classes on page 202
- Monitoring Interfaces That Have CoS Components on page 55
- Monitoring CoS Scheduler Maps on page 431
- Monitoring CoS Rewrite Rules on page 205

List of Sample Output show class-of- service on page 59

Output Fields Table 21 on page 57 lists the output fields for the **show class-of-service** command. Output fields are listed in the approximate order in which they appear.

Table 21: show class-of-service Output Fields

Field Name	Field Description	Level of Output
Forwarding class	The forwarding class configuration: • **Forwarding class**—Name of the forwarding class. • **ID**—Forwarding class ID. • **Queue**—Queue number.	All levels
Code point type	The type of code-point alias: • **dscp**—Aliases for DiffServ code point (DSCP) values. • **ieee–802.1**—Aliases for IEEE 802.1p values. • **exp**—Aliases for MPLS EXP values.	All levels
Alias	Names given to CoS values.	All levels
Bit pattern	Set of bits associated with an alias.	All levels
Classifier	Name of the classifier.	All levels

Table 21: show class-of-service Output Fields (continued)

Field Name	Field Description	Level of Output
Code point	Code-point values.	All levels
Loss priority	Loss priority assigned to specific CoS values and aliases of the classifier.	All levels
Rewrite rule	Name of the rewrite rule if one has been configured.	All levels
Drop profile	Name of the drop profile.	All levels
Type	Type of drop profile. QFX Series supports only the **discrete** type of drop-profile.	All levels
Fill level	Percentage of queue buffer fullness in a drop profile at which packets begin to drop during periods of congestion.	All levels
Scheduler map	Name of the scheduler map.	All levels
Scheduler	Name of the scheduler.	All levels
Transmit rate	Transmission rate of the scheduler.	All levels
Buffer size	Delay buffer size in the queue.	All levels
Drop profiles	Drop profiles configured for the specified scheduler.	All levels
Protocol	Transport protocol corresponding to the drop profile.	All levels
Name	Name of the drop profile.	All levels
Queues supported	Number of queues that can be configured on the interface.	All levels
Queues in use	Number of queues currently configured.	All levels
Physical interface	Name of the physical interface.	All levels
Scheduler map	Name of the scheduler map.	All levels
Congestion-notification	Enabled if a congestion notification profile is applied to the interface; disabled if no congestion notification profile is applied to the interface. NOTE: OCX Series switches do not support congestion notification profiles.	All levels
Forwarding class set	Name of the forwarding class set (priority group). NOTE: Only on systems that support enhanced transmission selection (ETS) hierarchical port scheduling.	
Index	Internal index of an object.	All levels

Sample Output

show class-of- service

```
user@switch> show class-of-service
Forwarding class                          ID      Queue
  best-effort                              0        0
  fcoe                                     1        3
  no-loss                                  2        4
  network-control                          3        7
  mcast                                    8        8

Code point type: dscp
  Alias                Bit pattern
  af11                 001010
  af12                 001100
  ...                  ...

Code point type: ieee-802.1
  Alias                Bit pattern
  af11                 100
  ...                  ...

Classifier: dscp-default, Code point type: dscp, Index: 7
  Code point           Forwarding class                    Loss priority
  000000               best-effort                         low
  000001               best-effort                         low
  ...                  ...                                 ...

Classifier: ieee8021p-default, Code point type: ieee-802.1, Index: 11
  Code point           Forwarding class                    Loss priority
  000                  best-effort                         low
  001                  best-effort                         low
  010                  best-effort                         low
  011                  fcoe                                low
  100                  no-loss                             low
  101                  best-effort                         low
  110                  network-control                     low
  111                  network-control                     low

Drop profile:<default-drop-profile>, Type: discrete, Index: 1
  Fill level
         100

Scheduler map: <default>, Index: 2

  Scheduler: <default-be>, Forwarding class: best-effort, Index: 21
    Transmit rate: 5 percent, Rate Limit: none, Buffer size: 5 percent, Buffer
Limit: none,
    Priority: low
    Excess Priority: low
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority    Protocol    Index    Name
      Low              any             1    <default-drop-profile>
      Medium high      any             1    <default-drop-profile>
      High             any             1    <default-drop-profile>

  Scheduler: <default-fcoe>, Forwarding class: fcoe, Index: 50
    Transmit rate: 35 percent, Rate Limit: none, Buffer size: 35 percent, Buffer
```

```
 Limit: none,
    Priority: low
    Excess Priority: low
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    <default-drop-profile>
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

  Scheduler: <default-noloss>, Forwarding class: no-loss, Index: 51
    Transmit rate: 35 percent, Rate Limit: none, Buffer size: 35 percent, Buffer
 Limit: none,
    Priority: low
    Excess Priority: low
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    <default-drop-profile>
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

  Scheduler: <default-nc>, Forwarding class: network-control, Index: 23
    Transmit rate: 5 percent, Rate Limit: none, Buffer size: 5 percent, Buffer
Limit: none,
    Priority: low
    Excess Priority: low
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    <default-drop-profile>
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

  Scheduler: <default-mcast>, Forwarding class: mcast, Index: 49
    Transmit rate: 20 percent, Rate Limit: none, Buffer size: 20 percent, Buffer
 Limit: none,
    Priority: low
    Excess Priority: low
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    <default-drop-profile>
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

Physical interface: xe-0/0/0, Index: 129
Queues supported: 12, Queues in use: 12
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

Physical interface: xe-0/0/1, Index: 130
Queues supported: 12, Queues in use: 12
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

...            ...              ...

Forwarding class set: lan-fcset, Type: normal-type, Forwarding class set index:
7
```



```
Forwarding class                                  Index
best-effort                                         0
```

show class-of-service interface

Syntax

show class-of-service interface
<comprehensive | detail> <*interface-name*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Forwarding class map information added in Junos OS Release 9.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 12.1 for the PTX Series Packet Transport routers.
Command introduced in Junos OS Release 12.2 for the ACX Series Universal Metro routers.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
Options **detail** and **comprehensive** introduced in Junos OS Release 11.4.
Command introduced in Junos OS Release 15.1R3 on MX Series routers for enhanced subscriber management.

Description

Display the logical and physical interface associations for the classifier, rewrite rules, and scheduler map objects.

NOTE: On routing platforms with dual Routing Engines, running this command on the backup Routing Engine, with or without any of the available options, is not supported and produces the following error message:

error: the class-of-service subsystem is not running

Options

none—Display CoS associations for all physical and logical interfaces.

comprehensive—(M Series, MX Series, and T Series routers) (Optional) Display comprehensive quality-of-service (QoS) information about all physical and logical interfaces.

detail—(M Series, MX Series, and T Series routers) (Optional) Display QoS and CoS information based on the interface.

If the **interface** *interface-name* is a physical interface, the output includes:

- Brief QoS information about the physical interface
- Brief QoS information about the logical interface
- CoS information about the physical interface
- Brief information about filters or policers of the logical interface
- Brief CoS information about the logical interface

If the **interface** *interface-name* is a logical interface, the output includes:

- Brief QoS information about the logical interface
- Information about filters or policers for the logical interface
- CoS information about the logical interface

interface-name—(Optional) Display class-of-service (CoS) associations for the specified interface.

none—Display CoS associations for all physical and logical interfaces.

NOTE: ACX5000 routers do not support classification on logical interfaces and therefore do not show CoS associations for logical interfaces with this command.

Required Privilege Level
view

Related Documentation
- *Verifying and Managing Junos OS Enhanced Subscriber Management*

List of Sample Output
show class-of-service interface (Physical) on page 75
show class-of-service interface (Logical) on page 75
show class-of-service interface (Gigabit Ethernet) on page 76
show class-of-service interface (ANCP) on page 76
show class-of-service interface (PPPoE Interface) on page 76
show class-of-service interface (DHCP Interface) on page 76
show class-of-service interface (T4000 Routers with Type 5 FPCs) on page 77
show class-of-service interface detail on page 77
show class-of-service interface comprehensive on page 78
show class-of-service interface (ACX Series Routers) on page 89
show class-of-service interface (PPPoE Subscriber Interface for Enhanced Subscriber Management) on page 91

Output Fields
Table 22 on page 63 describes the output fields for the **show class-of-service interface** command. Output fields are listed in the approximate order in which they appear.

Table 22: show class-of-service interface Output Fields

Field Name	Field Description
Physical interface	Name of a physical interface.
Index	Index of this interface or the internal index of this object. (Enhanced subscriber management for MX Series routers) Index values for dynamic CoS traffic control profiles and dynamic scheduler maps are larger for enhanced subscriber management than they are for legacy subscriber management.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Dedicated Queues	Status of dedicated queues configured on an interface. Supported only on Trio MPC/MIC interfaces on MX Series routers. (Enhanced subscriber management for MX-Series routers) This field is not displayed for enhanced subscriber management.
Maximum usable queues	Number of queues you can configure on the interface.
Maximum usable queues	Maximum number of queues you can use.
Total non-default queues created	Number of queues created in addition to the default queues. Supported only on Trio MPC/MIC interfaces on MX Series routers. (Enhanced subscriber management for MX Series routers) This field is not displayed for enhanced subscriber management.
Rewrite Input IEEE Code-point	(QFX3500 switches only) IEEE 802.1p code point (priority) rewrite value. Incoming traffic from the Fibre Channel (FC) SAN is classified into the forwarding class specified in the native FC interface (NP_Port) fixed classifier and uses the priority specified as the IEEE 802.1p rewrite value.
Shaping rate	Maximum transmission rate on the physical interface. You can configure the shaping rate on the physical interface, or on the logical interface, but not on both. Therefore, the **Shaping rate** field is displayed for either the physical interface or the logical interface.
Scheduler map	Name of the output scheduler map associated with this interface. (Enhanced subscriber management for MX Series routers) The name of the dynamic scheduler map object is associated with a generated UID (for example, **SMAP-1_UID1002**) instead of with a subscriber interface.
Scheduler map forwarding class sets	(QFX Series only) Name of the output fabric scheduler map associated with a QFabric system Interconnect device interface.
Input shaping rate	For Gigabit Ethernet IQ2 PICs, maximum transmission rate on the input interface.
Input scheduler map	For Gigabit Ethernet IQ2 PICs, name of the input scheduler map associated with this interface.
Chassis scheduler map	Name of the scheduler map associated with the packet forwarding component queues.
Rewrite	Name and type of the rewrite rules associated with this interface.
Traffic-control-profile	Name of the associated traffic control profile. (Enhanced subscriber management for MX Series routers) The name of the dynamic traffic control profile object is associated with a generated UID (for example, **TC_PROF_100_199_SERIES_UID1006**) instead of with a subscriber interface.
Classifier	Name and type of classifiers associated with this interface.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Forwarding-class-map	Name of the forwarding map associated with this interface.
Congestion-notification	(QFX Series and EX4600 switches only) Congestion notification state, **enabled** or **disabled**.
Logical interface	Name of a logical interface.
Object	Category of an object: **Classifier**, **Fragmentation-map** (for LSQ interfaces only), **Scheduler-map**, **Rewrite**, **Translation Table** (for IQE PICs only), or **traffic-class-map** (for T4000 routers with Type 5 FPCs).
Name	Name of an object.
Type	Type of an object: **dscp**, **dscp-ipv6**, **exp**, **ieee-802.1**, **ip**, **inet-precedence**, or **ieee-802.1ad** (for traffic class map on T4000 routers with Type 5 FPCs)..
Link-level type	Encapsulation on the physical interface.
MTU	MTU size on the physical interface.
Speed	Speed at which the interface is running.
Loopback	Whether loopback is enabled and the type of loopback.
Source filtering	Whether source filtering is enabled or disabled.
Flow control	Whether flow control is enabled or disabled.
Auto-negotiation	(Gigabit Ethernet interfaces) Whether autonegotiation is enabled or disabled.
Remote-fault	(Gigabit Ethernet interfaces) Remote fault status. • **Online**—Autonegotiation is manually configured as online. • **Offline**—Autonegotiation is manually configured as offline.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Device flags	The **Device flags** field provides information about the physical device and displays one or more of the following values: • **Down**—Device has been administratively disabled. • **Hear-Own-Xmit**—Device receives its own transmissions. • **Link-Layer-Down**—The link-layer protocol has failed to connect with the remote endpoint. • **Loopback**—Device is in physical loopback. • **Loop-Detected**—The link layer has received frames that it sent, thereby detecting a physical loopback. • **No-Carrier**—On media that support carrier recognition, no carrier is currently detected. • **No-Multicast**—Device does not support multicast traffic. • **Present**—Device is physically present and recognized. • **Promiscuous**—Device is in promiscuous mode and recognizes frames addressed to all physical addresses on the media. • **Quench**—Transmission on the device is quenched because the output buffer is overflowing. • **Recv-All-Multicasts**—Device is in multicast promiscuous mode and therefore provides no multicast filtering. • **Running**—Device is active and enabled.

Interface flags	The **Interface flags** field provides information about the physical interface and displays one or more of the following values:
	• **Admin-Test**—Interface is in test mode and some sanity checking, such as loop detection, is disabled.
	• **Disabled**—Interface is administratively disabled.
	• **Down**—A hardware failure has occurred.
	• **Hardware-Down**—Interface is nonfunctional or incorrectly connected.
	• **Link-Layer-Down**—Interface keepalives have indicated that the link is incomplete.
	• **No-Multicast**—Interface does not support multicast traffic.
	• **No-receive No-transmit**—Passive monitor mode is configured on the interface.
	• **Point-To-Point**—Interface is point-to-point.
	• **Pop all MPLS labels from packets of depth**—MPLS labels are removed as packets arrive on an interface that has the **pop-all-labels** statement configured. The depth value can be one of the following:
	  ◦ **1**—Takes effect for incoming packets with one label only.
	  ◦ **2**—Takes effect for incoming packets with two labels only.
	  ◦ **[1 2]**—Takes effect for incoming packets with either one or two labels.
	• **Promiscuous**—Interface is in promiscuous mode and recognizes frames addressed to all physical addresses.
	• **Recv-All-Multicasts**—Interface is in multicast promiscuous mode and provides no multicast filtering.
	• **SNMP-Traps**—SNMP trap notifications are enabled.
	• **Up**—Interface is enabled and operational.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Flags	The **Logical interface flags** field provides information about the logical interface and displays one or more of the following values: • **ACFC Encapsulation**—Address control field Compression (ACFC) encapsulation is enabled (negotiated successfully with a peer). • **Device-down**—Device has been administratively disabled. • **Disabled**—Interface is administratively disabled. • **Down**—A hardware failure has occurred. • **Clear-DF-Bit**—GRE tunnel or IPsec tunnel is configured to clear the Don't Fragment (DF) bit. • **Hardware-Down**—Interface protocol initialization failed to complete successfully. • **PFC**—Protocol field compression is enabled for the PPP session. • **Point-To-Point**—Interface is point-to-point. • **SNMP-Traps**—SNMP trap notifications are enabled. • **Up**—Interface is enabled and operational.
Encapsulation	Encapsulation on the logical interface.
Admin	Administrative state of the interface (**Up** or **Down**)
Link	Status of physical link (**Up** or **Down**).
Proto	Protocol configured on the interface.
Input Filter	Names of any firewall filters to be evaluated when packets are received on the interface, including any filters attached through activation of dynamic service.
Output Filter	Names of any firewall filters to be evaluated when packets are transmitted on the interface, including any filters attached through activation of dynamic service.
Link flags	Provides information about the physical link and displays one or more of the following values: • **ACFC**—Address control field compression is configured. The Point-to-Point Protocol (PPP) session negotiates the ACFC option. • **Give-Up**—Link protocol does not continue connection attempts after repeated failures. • **Loose-LCP**—PPP does not use the Link Control Protocol (LCP) to indicate whether the link protocol is operational. • **Loose-LMI**—Frame Relay does not use the Local Management Interface (LMI) to indicate whether the link protocol is operational. • **Loose-NCP**—PPP does not use the Network Control Protocol (NCP) to indicate whether the device is operational. • **Keepalives**—Link protocol keepalives are enabled. • **No-Keepalives**—Link protocol keepalives are disabled. • **PFC**—Protocol field compression is configured. The PPP session negotiates the PFC option.
Hold-times	Current interface hold-time up and hold-time down, in milliseconds.
CoS queues	Number of CoS queues configured.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Last flapped	Date, time, and how long ago the interface went from down to up. The format is **Last flapped:** ***year-month-day hour:minute:second:timezone*** **(*hour:minute:second* ago)**. For example, **Last flapped: 2002-04-26 10:52:40 PDT (04:33:20 ago).**
Statistics last cleared	Number and rate of bytes and packets received and transmitted on the physical interface. • **Input bytes**—Number of bytes received on the interface. • **Output bytes**—Number of bytes transmitted on the interface. • **Input packets**—Number of packets received on the interface. • **Output packets**—Number of packets transmitted on the interface.
Exclude Overhead Bytes	Exclude the counting of overhead bytes from aggregate queue statistics. • **Disabled**—Default configuration. Includes the counting of overhead bytes in aggregate queue statistics. • **Enabled**—Excludes the counting of overhead bytes from aggregate queue statistics for just the physical interface. • **Enabled for hierarchy**—Excludes the counting of overhead bytes from aggregate queue statistics for the physical interface as well as all child interfaces, including logical interfaces and interface sets.
IPv6 transit statistics	Number of IPv6 transit bytes and packets received and transmitted on the logical interface if IPv6 statistics tracking is enabled.
Input errors	Input errors on the interface. The labels are explained in the following list: • **Errors**—Sum of the incoming frame aborts and FCS errors. • **Drops**—Number of packets dropped by the input queue of the I/O Manager ASIC. If the interface is saturated, this number increments once for every packet that is dropped by the ASIC's RED mechanism. • **Framing errors**—Number of packets received with an invalid frame checksum (FCS). • **Runts**—Number of frames received that are smaller than the runt threshold. • **Giants**—Number of frames received that are larger than the giant threshold. • **Bucket Drops**—Drops resulting from the traffic load exceeding the interface transmit or receive leaky bucket configuration. • **Policed discards**—Number of frames that the incoming packet match code discarded because they were not recognized or not of interest. Usually, this field reports protocols that Junos OS does not handle. • **L3 incompletes**—Number of incoming packets discarded because they failed Layer 3 (usually IPv4) sanity checks of the header. For example, a frame with less than 20 bytes of available IP header is discarded. Layer 3 incomplete errors can be ignored by configuring the **ignore-l3-incompletes** statement. • **L2 channel errors**—Number of times the software did not find a valid logical interface for an incoming frame. • **L2 mismatch timeouts**—Number of malformed or short packets that caused the incoming packet handler to discard the frame as unreadable. • **HS link CRC errors**—Number of errors on the high-speed links between the ASICs responsible for handling the router interfaces. • **HS link FIFO overflows**—Number of FIFO overflows on the high-speed links between the ASICs responsible for handling the router interfaces.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Output errors	Output errors on the interface. The labels are explained in the following list: • **Carrier transitions**—Number of times the interface has gone from **down** to **up**. This number does not normally increment quickly, increasing only when the cable is unplugged, the far-end system is powered down and up, or another problem occurs. If the number of carrier transitions increments quickly (perhaps once every 10 seconds), the cable, the far-end system, or the PIC is malfunctioning. • **Errors**—Sum of the outgoing frame aborts and FCS errors. • **Drops**—Number of packets dropped by the output queue of the I/O Manager ASIC. If the interface is saturated, this number increments once for every packet that is dropped by the ASIC's RED mechanism. NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), the **Drops** field does not always use the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs. • **Aged packets**—Number of packets that remained in shared packet SDRAM so long that the system automatically purged them. The value in this field should never increment. If it does, it is most likely a software bug or possibly malfunctioning hardware. • **HS link FIFO underflows**—Number of FIFO underflows on the high-speed links between the ASICs responsible for handling the router interfaces. • **MTU errors**—Number of packets whose size exceeds the MTU of the interface.
Egress queues	Total number of egress Maximum usable queues on the specified interface.
Queue counters	CoS queue number and its associated user-configured forwarding class name. • **Queued packets**—Number of queued packets. • **Transmitted packets**—Number of transmitted packets. • **Dropped packets**—Number of packets dropped by the ASIC's RED mechanism. NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), the **Dropped packets** field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.
SONET alarms **SONET defects**	(SONET) SONET media-specific alarms and defects that prevent the interface from passing packets. When a defect persists for a certain period, it is promoted to an alarm. Based on the router configuration, an alarm can ring the red or yellow alarm bell on the router or light the red or yellow alarm LED on the craft interface. See these fields for possible alarms and defects: **SONET PHY**, **SONET section**, **SONET line**, and **SONET path**.
SONET PHY	Counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET PHY** field has the following subfields: • **PLL Lock**—Phase-locked loop • **PHY Light**—Loss of optical signal

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
SONET section	Counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET section** field has the following subfields: • **BIP-B1**—Bit interleaved parity for SONET section overhead • **SEF**—Severely errored framing • **LOS**—Loss of signal • **LOF**—Loss of frame • **ES-S**—Errored seconds (section) • **SES-S**—Severely errored seconds (section) • **SEFS-S**—Severely errored framing seconds (section)
SONET line	Active alarms and defects, plus counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET line** field has the following subfields: • **BIP-B2**—Bit interleaved parity for SONET line overhead • **REI-L**—Remote error indication (near-end line) • **RDI-L**—Remote defect indication (near-end line) • **AIS-L**—Alarm indication signal (near-end line) • **BERR-SF**—Bit error rate fault (signal failure) • **BERR-SD**—Bit error rate defect (signal degradation) • **ES-L**—Errored seconds (near-end line) • **SES-L**—Severely errored seconds (near-end line) • **UAS-L**—Unavailable seconds (near-end line) • **ES-LFE**—Errored seconds (far-end line) • **SES-LFE**—Severely errored seconds (far-end line) • **UAS-LFE**—Unavailable seconds (far-end line)

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
SONET path	Active alarms and defects, plus counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET path** field has the following subfields: • **BIP-B3**—Bit interleaved parity for SONET section overhead • **REI-P**—Remote error indication • **LOP-P**—Loss of pointer (path) • **AIS-P**—Path alarm indication signal • **RDI-P**—Path remote defect indication • **UNEQ-P**—Path unequipped • **PLM-P**—Path payload (signal) label mismatch • **ES-P**—Errored seconds (near-end STS path) • **SES-P**—Severely errored seconds (near-end STS path) • **UAS-P**—Unavailable seconds (near-end STS path) • **ES-PFE**—Errored seconds (far-end STS path) • **SES-PFE**—Severely errored seconds (far-end STS path) • **UAS-PFE**—Unavailable seconds (far-end STS path)
Received SONET overhead **Transmitted SONET overhead**	Values of the received and transmitted SONET overhead: • **C2**—Signal label. Allocated to identify the construction and content of the STS-level SPE and for PDI-P. • **F1**—Section user channel byte. This byte is set aside for the purposes of users. • **K1** and **K2**—These bytes are allocated for APS signaling for the protection of the multiplex section. • **J0**—Section trace. This byte is defined for STS-1 number 1 of an STS-*N* signal. Used to transmit a 1-byte fixed-length string or a 16-byte message so that a receiving terminal in a section can verify its continued connection to the intended transmitter. • **S1**—Synchronization status. The S1 byte is located in the first STS-1 number of an STS-*N* signal. • **Z3** and **Z4**—Allocated for future use.
Received path trace **Transmitted path trace**	SONET/SDH interfaces allow path trace bytes to be sent inband across the SONET/SDH link. Juniper Networks and other router manufacturers use these bytes to help diagnose misconfigurations and network errors by setting the transmitted path trace message so that it contains the system hostname and name of the physical interface. The received path trace value is the message received from the router at the other end of the fiber. The transmitted path trace value is the message that this router transmits.
HDLC configuration	Information about the HDLC configuration. • **Policing bucket**—Configured state of the receiving policer. • **Shaping bucket**—Configured state of the transmitting shaper. • **Giant threshold**—Giant threshold programmed into the hardware. • **Runt threshold**—Runt threshold programmed into the hardware.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Packet Forwarding Engine configuration	Information about the configuration of the Packet Forwarding Engine: • **Destination slot**—FPC slot number. • **PLP byte**—Packet Level Protocol byte.
CoS information	Information about the CoS queue for the physical interface. • **CoS transmit queue**—Queue number and its associated user-configured forwarding class name. • **Bandwidth %**—Percentage of bandwidth allocated to the queue. • **Bandwidth bps**—Bandwidth allocated to the queue (in bps). • **Buffer %**—Percentage of buffer space allocated to the queue. • **Buffer usec**—Amount of buffer space allocated to the queue, in microseconds. This value is nonzero only if the buffer size is configured in terms of time. • **Priority**—Queue priority: **low** or **high**. • **Limit**—Displayed if rate limiting is configured for the queue. Possible values are **none** and **exact**. If **exact** is configured, the queue transmits only up to the configured bandwidth, even if excess bandwidth is available. If **none** is configured, the queue transmits beyond the configured bandwidth if bandwidth is available.
Forwarding classes	Total number of forwarding classes supported on the specified interface.
Egress queues	Total number of egress Maximum usable queues on the specified interface.
Queue	Queue number.
Forwarding classes	Forwarding class name.
Queued Packets	Number of packets queued to this queue.
Queued Bytes	Number of bytes queued to this queue. The byte counts vary by PIC type.
Transmitted Packets	Number of packets transmitted by this queue. When fragmentation occurs on the egress interface, the first set of packet counters shows the postfragmentation values. The second set of packet counters (displayed under the **Packet Forwarding Engine Chassis Queues** field) shows the prefragmentation values.
Transmitted Bytes	Number of bytes transmitted by this queue. The byte counts vary by PIC type.
Tail-dropped packets	Number of packets dropped because of tail drop.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description

RED-dropped packets	Number of packets dropped because of random early detection (RED).
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, the total number of dropped packets is displayed. On all other M Series routers, the output classifies dropped packets into the following categories:
	  ◦ **Low, non-TCP**—Number of low-loss priority non-TCP packets dropped because of RED.
	  ◦ **Low, TCP**—Number of low-loss priority TCP packets dropped because of RED.
	  ◦ **High, non-TCP**—Number of high-loss priority non-TCP packets dropped because of RED.
	  ◦ **High, TCP**—Number of high-loss priority TCP packets dropped because of RED.
	• (MX Series routers with enhanced DPCs, and T Series routers with enhanced FPCs only) The output classifies dropped packets into the following categories:
	  ◦ **Low**—Number of low-loss priority packets dropped because of RED.
	  ◦ **Medium-low**—Number of medium-low loss priority packets dropped because of RED.
	  ◦ **Medium-high**—Number of medium-high loss priority packets dropped because of RED.
	  ◦ **High**—Number of high-loss priority packets dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

RED-dropped bytes	Number of bytes dropped because of RED. The byte counts vary by PIC type.
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, only the total number of dropped bytes is displayed. On all other M Series routers, the output classifies dropped bytes into the following categories:
	  ◦ **Low, non-TCP**—Number of low-loss priority non-TCP bytes dropped because of RED.
	  ◦ **Low, TCP**—Number of low-loss priority TCP bytes dropped because of RED.
	  ◦ **High, non-TCP**—Number of high-loss priority non-TCP bytes dropped because of RED.
	  ◦ **High, TCP**—Number of high-loss priority TCP bytes dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

Transmit rate	Configured transmit rate of the scheduler. The rate is a percentage of the total interface bandwidth.
Rate Limit	Rate limiting configuration of the queue. Possible values are : • **None**—No rate limit. • **exact**—Queue transmits at the configured rate.
Buffer size	Delay buffer size in the queue.
Priority	Scheduling priority configured as **low** or **high**.
Excess Priority	Priority of the excess bandwidth traffic on a scheduler: **low**, **medium-low**, **medium-high**, **high**, or **none**.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Drop profiles	Display the assignment of drop profiles. • **Loss priority**—Packet loss priority for drop profile assignment. • **Protocol**—Transport protocol for drop profile assignment. • **Index**—Index of the indicated object. Objects that have indexes in this output include schedulers and drop profiles. • **Name**—Name of the drop profile. • **Type**—Type of the drop profile: **discrete** or **interpolated**. • **Fill Level**—Percentage fullness of a queue. • **Drop probability**—Drop probability at this fill level.
Excess Priority	Priority of the excess bandwidth traffic on a scheduler.
Drop profiles	Display the assignment of drop profiles. • **Loss priority**—Packet loss priority for drop profile assignment. • **Protocol**—Transport protocol for drop profile assignment. • **Index**—Index of the indicated object. Objects that have indexes in this output include schedulers and drop profiles. • **Name**—Name of the drop profile. • **Type**—Type of the drop profile: **discrete** or **interpolated**. • **Fill Level**—Percentage fullness of a queue. • **Drop probability**—Drop probability at this fill level.

Table 22: show class-of-service interface Output Fields (continued)

Field Name	Field Description

Adjustment information	Display the assignment of shaping-rate adjustments on a scheduler node or queue.
	
	• **Adjusting application**—Application that is performing the shaping-rate adjustment.
	  ◦ The adjusting application can appear as **ancp LS-0**, which is the Junos OS Access Node Control Profile process (**ancpd**) that performs shaping-rate adjustments on schedule nodes.
	  ◦ The adjusting application can appear as **DHCP**, which adjusts the shaping-rate and overhead-accounting class-of-service attributes based on DHCP option 82, suboption 9 (Vendor Specific Information). The shaping rate is based on the actual-data-rate-downstream attribute. The overhead accounting value is based on the access-loop-encapsulation attribute and specifies whether the access loop uses Ethernet (frame mode) or ATM (cell mode).
	  ◦ The adjusting application can also appear as **pppoe**, which adjusts the shaping-rate and overhead-accounting class-of-service attributes on dynamic subscriber interfaces in a broadband access network based on access line parameters in Point-to-Point Protocol over Ethernet (PPPoE) Tags [TR-101]. This feature is supported on MPC/MIC interfaces on MX Series routers. The shaping rate is based on the actual-data-rate-downstream attribute. The overhead accounting value is based on the access-loop-encapsulation attribute and specifies whether the access loop uses Ethernet (frame mode) or ATM (cell mode).
	• **Adjustment type**—Type of adjustment: **absolute** or **delta**.
	• **Configured shaping rate**—Shaping rate configured for the scheduler node or queue.
	• **Adjustment value**—Value of adjusted shaping rate.
	• **Adjustment target**—Level of shaping-rate adjustment performed: **node** or **queue**.
	• **Adjustment overhead-accounting mode**—Configured shaping mode: **frame** or **cell**.
	• **Adjustment overhead bytes**—Number of bytes that the ANCP agent adds to or subtracts from the actual downstream frame overhead before reporting the adjusted values to CoS.
	• **Adjustment target**—Level of shaping-rate adjustment performed: **node** or **queue**.
	• **Adjustment multicast index**—

Sample Output

show class-of-service interface (Physical)

```
user@host> show class-of-service interface so-0/2/3
Physical interface: so-0/2/3, Index: 135
Maximum usable queues: 8, Queues in use: 4
Total non-default queues created: 4
  Scheduler map: <default>, Index: 2032638653

  Logical interface: fe-0/0/1.0, Index: 68, Dedicated Queues: no
    Shaping rate: 32000
    Object                   Name                   Type          Index
    Scheduler-map            <default>                             27
    Rewrite                  exp-default            exp            21
    Classifier               exp-default            exp            5
    Classifier               ipprec-compatibility   ip             8
    Forwarding-class-map     exp-default            exp            5
```

show class-of-service interface (Logical)

```
user@host> show class-of-service interface so-0/2/3.0
Logical interface: so-0/2/3.0, Index: 68, Dedicated Queues: no
    Shaping rate: 32000
```

```
Object                   Name                   Type             Index
Scheduler-map            <default>                                27
Rewrite                  exp-default            exp               21
Classifier               exp-default            exp               5
Classifier               ipprec-compatibility   ip                8
Forwarding–class–map     exp-default            exp               5
```

show class-of-service interface (Gigabit Ethernet)

```
user@host> show class-of-service interface ge-6/2/0
Physical interface: ge-6/2/0, Index: 175
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Input scheduler map: <default>, Index: 3
  Chassis scheduler map: <default-chassis>, Index: 4
```

show class-of-service interface (ANCP)

```
user@host> show class-of-service interface pp0.1073741842
  Logical interface: pp0.1073741842, Index: 341
Object                 Name                   Type                    Index
Traffic-control-profile TCP-CVLAN             Output                  12408
Classifier             dscp-ipv6-compatibility dscp-ipv6                  9
Classifier             ipprec-compatibility   ip                         13

    Adjusting application: ancp LS-0
      Adjustment type: absolute
      Configured shaping rate: 4000000
      Adjustment value: 11228000
      Adjustment overhead-accounting mode: Frame Mode
      Adjustment overhead bytes: 50
      Adjustment target: node
```

show class-of-service interface (PPPoE Interface)

```
user@host> show class-of-service interface pp0.1
Logical interface: pp0.1, Index: 85
    Object                 Name                   Type           Index
    Traffic-control-profile tcp-pppoe.o.pp0.1     Output         2726446535
    Classifier             ipprec-compatibility   ip             13

    Adjusting application: PPPoE
      Adjustment type: absolute
      Adjustment value: 5000000
      Adjustment overhead-accounting mode: cell
      Adjustment target: node
```

show class-of-service interface (DHCP Interface)

```
user@host> show class-of-service interface demux0.1
Logical interface: pp0.1, Index: 85
    Object                 Name                   Type           Index
    Traffic-control-profile tcp-dhcp.o.demux0.1   Output         2726446535
    Classifier             ipprec-compatibility   ip             13

    Adjusting application: DHCP
      Adjustment type: absolute
      Adjustment value: 5000000
      Adjustment overhead-accounting mode: cell
      Adjustment target: node
```


show class-of-service interface (T4000 Routers with Type 5 FPCs)

```
user@host> show class-of-service interface xe-4/0/0
Physical interface: xe-4/0/0, Index: 153
    Maximum usable queues: 8, Queues in use: 4
      Shaping rate: 5000000000 bps
      Scheduler map: <default>, Index: 2
      Congestion-notification: Disabled

      Logical interface: xe-4/0/0.0, Index: 77
        Object                  Name                    Type
Index
        Classifier              ipprec-compatibility    ip
  13
```

show class-of-service interface detail

```
user@host> show class-of-service interface ge-0/3/0 detail

Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Link-level type: Ethernet, MTU: 1518, Speed: 1000mbps, Loopback: Disabled,
Source filtering: Disabled, Flow control: Enabled, Auto-negotiation: Enabled,
Remote fault: Online
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000

  Physical interface: ge-0/3/0, Index: 138
  Maximum usable queues: 4, Queues in use: 5
  Shaping rate: 50000 bps
  Scheduler map: interface-schedular-map, Index: 58414
  Input shaping rate: 10000 bps
  Input scheduler map: schedular-map, Index: 15103
  Chassis scheduler map: <default-chassis>, Index: 4
  Congestion-notification: Disabled

  Logical interface ge-0/3/0.0
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    inet
    mpls
Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.0      up    up   inet
                           mpls
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.0      up    up   inet
                           mpls

  Logical interface: ge-0/3/0.0, Index: 68
    Object                  Name                    Type                Index
    Rewrite                 exp-default             exp (mpls-any)        33
    Classifier              exp-default             exp                   10
    Classifier              ipprec-compatibility    ip                    13


  Logical interface ge-0/3/0.1
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    inet
Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.1      up    up   inet
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.1      up    up   inet
```



```
  Logical interface: ge-0/3/0.1, Index: 69
    Object                  Name                   Type              Index
    Classifier              ipprec-compatibility   ip                 13
```

show class-of-service interface comprehensive

```
user@host> show class-of-service interface ge-0/3/0 comprehensive
Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Interface index: 138, SNMP ifIndex: 601, Generation: 141
  Link-level type: Ethernet, MTU: 1518, Speed: 1000mbps, BPDU Error: None,
MAC-REWRITE Error: None, Loopback: Disabled, Source filtering: Disabled, Flow
control: Enabled,
  Auto-negotiation: Enabled, Remote fault: Online
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000
  CoS queues     : 4 supported, 4 maximum usable queues
  Schedulers     : 256
  Hold-times     : Up 0 ms, Down 0 ms
  Current address: 00:14:f6:f4:b4:5d, Hardware address: 00:14:f6:f4:b4:5d
  Last flapped   : 2010-09-07 06:35:22 PDT (15:14:42 ago)
  Statistics last cleared: Never  Exclude Overhead Bytes: Disabled
  Traffic statistics:
   Input  bytes  :                    0                    0 bps
   Output bytes  :                    0                    0 bps
   Input  packets:                    0                    0 pps
   Output packets:                    0                    0 pps
   IPv6 total statistics:
    Input  bytes  :                   0
    Output bytes  :                   0
    Input  packets:                   0
    Output packets:                   0
  Ingress traffic statistics at Packet Forwarding Engine:
   Input  bytes  :                    0                    0 bps
   Input  packets:                    0                    0 pps
   Drop   bytes  :                    0                    0 bps
   Drop   packets:                    0                    0 pps
  Label-switched interface (LSI) traffic statistics:
   Input  bytes  :                    0                    0 bps
   Input  packets:                    0                    0 pps
  Input errors:
    Errors: 0, Drops: 0, Framing errors: 0, Runts: 0, Policed discards: 0, L3
incompletes: 0, L2 channel errors: 0, L2 mismatch timeouts: 0, FIFO errors: 0,
Resource errors: 0
  Output errors:
    Carrier transitions: 5, Errors: 0, Drops: 0, Collisions: 0, Aged packets: 0,
 FIFO errors: 0, HS link CRC errors: 0, MTU errors: 0, Resource errors: 0
  Ingress queues: 4 supported, 5 in use
  Queue counters:       Queued packets  Transmitted packets      Dropped packets

    0 af3                            0                    0                    0

    1 af2                            0                    0                    0

    2 ef2                            0                    0                    0

    3 ef1                            0                    0                    0

  Egress queues: 4 supported, 5 in use
  Queue counters:       Queued packets  Transmitted packets      Dropped packets
```



```
    0 af3                            0                    0                    0

    1 af2                            0                    0                    0

    2 ef2                            0                    0                    0

    3 ef1                            0                    0                    0

  Active alarms  : None
  Active defects : None
  MAC statistics:                      Receive         Transmit
    Total octets                             0                0
    Total packets                            0                0
    Unicast packets                          0                0
    Broadcast packets                        0                0
    Multicast packets                        0                0
    CRC/Align errors                         0                0
    FIFO errors                              0                0
    MAC control frames                       0                0
    MAC pause frames                         0                0
    Oversized frames                         0
    Jabber frames                            0
    Fragment frames                          0
    VLAN tagged frames                       0
    Code violations                          0
  Filter statistics:
    Input packet count                       0
    Input packet rejects                     0
    Input DA rejects                         0
    Input SA rejects                         0
    Output packet count                                       0
    Output packet pad count                                   0
    Output packet error count                                 0
    CAM destination filters: 0, CAM source filters: 0
  Autonegotiation information:
    Negotiation status: Complete
    Link partner:
       Link mode: Full-duplex, Flow control: Symmetric/Asymmetric, Remote fault:
 OK
    Local resolution:
      Flow control: Symmetric, Remote fault: Link OK
  Packet Forwarding Engine configuration:
    Destination slot: 0
  CoS information:
    Direction : Output
    CoS transmit queue               Bandwidth               Buffer Priority
Limit
                            %            bps     %           usec
    2 ef2                  39          19500     0            120     high
none
    Direction : Input
    CoS transmit queue               Bandwidth               Buffer Priority
Limit
                            %            bps     %           usec
    0 af3                  30           3000    45              0      low
none

Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Interface index: 138, SNMP ifIndex: 601
Forwarding classes: 16 supported, 5 in use
Ingress queues: 4 supported, 5 in use
```

```
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Forwarding classes: 16 supported, 5 in use
Egress queues: 4 supported, 5 in use
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                     0                     0 pps
    RL-dropped bytes     :                     0                     0 bps
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
```

```
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps

Packet Forwarding Engine Chassis Queues:
Queues: 4 supported, 5 in use
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
```

```
    RED-dropped bytes    : Not Available
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :               108546                     0 pps
    Bytes                :             12754752                   376 bps
  Transmitted:
    Packets              :               108546                      0 pps
    Bytes                :             12754752                   376 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available

Physical interface: ge-0/3/0, Index: 138
Maximum usable queues: 4, Queues in use: 5
  Shaping rate: 50000 bps

Scheduler map: interface-schedular-map, Index: 58414

  Scheduler: ef2, Forwarding class: ef2, Index: 39155
    Transmit rate: 39 percent, Rate Limit: none, Buffer size: 120 us, Buffer
Limit: none, Priority: high
    Excess Priority: unspecified
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
       100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
       100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
       100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
       100                100
  Input shaping rate: 10000 bps
  Input scheduler map: schedular-map

Scheduler map: schedular-map, Index: 15103

  Scheduler: af3, Forwarding class: af3, Index: 35058
    Transmit rate: 30 percent, Rate Limit: none, Buffer size: 45 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: unspecified
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         40582    green
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any         18928    yellow
Drop profile: green, Type: discrete, Index: 40582
  Fill level    Drop probability
        50                  0
       100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
```



```
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: yellow, Type: discrete, Index: 18928
  Fill level    Drop probability
          50                  0
         100                100
  Chassis scheduler map: < default-drop-profile>
Scheduler map: < default-drop-profile>, Index: 4

  Scheduler: < default-drop-profile>, Forwarding class: af3, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100

  Scheduler: < default-drop-profile>, Forwarding class: af2, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                100

  Scheduler: < default-drop-profile>, Forwarding class: ef2, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
```

```
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100

  Scheduler: < default-drop-profile>, Forwarding class: ef1, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: , Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
  Congestion-notification: Disabled
Forwarding class                       ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                   0       0          0              low
            normal
  af2                                   1       1          1              low
            normal
  ef2                                   2       2          2              high
            normal
  ef1                                   3       3          3              high
            normal
  af1                                   4       4          0              low
            normal

  Logical interface ge-0/3/0.0 (Index 68) (SNMP ifIndex 152) (Generation 159)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    Traffic statistics:
     Input  bytes  :                    0
     Output bytes  :                    0
     Input  packets:                    0
```



```
     Output packets:                   0
    Local statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
     Output packets:                   0
    Transit statistics:
     Input  bytes  :                   0                    0 bps
     Output bytes  :                   0                    0 bps
     Input  packets:                   0                    0 pps
     Output packets:                   0                    0 pps
    Protocol inet, MTU: 1500, Generation: 172, Route table: 0
      Flags: Sendbcast-pkt-to-re
      Input Filters: filter-in-ge-0/3/0.0-i,
      Policer: Input: p1-ge-0/3/0.0-inet-i
    Protocol mpls, MTU: 1488, Maximum labels: 3, Generation: 173, Route table: 0

      Flags: Is-Primary
      Output Filters: exp-filter,,,,,

Logical interface ge-1/2/0.0 (Index 347) (SNMP ifIndex 638) (Generation 156)

Forwarding class ID  Queue  Restricted queue  Fabric priority  Policing priority
   SPU priority
best-effort      0   0      0                 low              normal
    low

Aggregate Forwarding-class statistics per forwarding-class
  Aggregate Forwarding-class statistics:
   Forwarding-class statistics:

Forwarding-class best-effort statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

IPv4 protocol forwarding-class statistics:
  Forwarding-class statistics:
  Forwarding-class best-effort statistics:

    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0
```

```
    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

    IPv6  protocol forwarding-class statistics:
    Forwarding-class statistics:
     Forwarding-class best-effort statistics:

    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Logical interface ge-0/3/0.0 (Index 68) (SNMP ifIndex 152)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    Input packets : 0
    Output packets: 0

Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.0      up    up   inet  filter-in-ge-0/3/0.0-i
                           mpls                       exp-filter
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.0      up    up
                           inet  p1-ge-0/3/0.0-inet-i
                           mpls

Filter: filter-in-ge-0/3/0.0-i
Counters:
Name                                                Bytes              Packets
count-filter-in-ge-0/3/0.0-i                            0                    0

Filter: exp-filter
```



```
Counters:
Name                                                 Bytes              Packets
count-exp-seven-match                                    0                    0
count-exp-zero-match                                     0                    0
Policers:
Name                                               Packets
p1-ge-0/3/0.0-inet-i                                     0


  Logical interface: ge-0/3/0.0, Index: 68
    Object                  Name                   Type                    Index

    Rewrite                 exp-default            exp (mpls-any)             33

Rewrite rule: exp-default, Code point type: exp, Index: 33
  Forwarding class                    Loss priority       Code point
  af3                                 low                 000
  af3                                 high                001
  af2                                 low                 010
  af2                                 high                011
  ef2                                 low                 100
  ef2                                 high                101
  ef1                                 low                 110
  ef1                                 high                111
    Object                  Name                   Type                    Index

    Classifier              exp-default            exp                        10

Classifier: exp-default, Code point type: exp, Index: 10
  Code point         Forwarding class                    Loss priority
  000                af3                                 low
  001                af3                                 high
  010                af2                                 low
  011                af2                                 high
  100                ef2                                 low
  101                ef2                                 high
  110                ef1                                 low
  111                ef1                                 high
    Object                  Name                   Type                    Index

    Classifier              ipprec-compatibility   ip                         13

Classifier: ipprec-compatibility, Code point type: inet-precedence, Index: 13
  Code point         Forwarding class                    Loss priority
  000                af3                                 low
  001                af3                                 high
  010                af3                                 low
  011                af3                                 high
  100                af3                                 low
  101                af3                                 high
  110                ef1                                 low
  111                ef1                                 high
Forwarding class                    ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                0       0          0             low
          normal
  af2                                1       1          1             low
          normal
  ef2                                2       2          2             high
          normal
  ef1                                3       3          3             high
```

```
          normal
  af1                                 4       4          0             low
          normal

  Logical interface ge-0/3/0.1 (Index 69) (SNMP ifIndex 154) (Generation 160)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    Traffic statistics:
     Input  bytes  :                  0
     Output bytes  :                  0
     Input  packets:                  0
     Output packets:                  0
    Local statistics:
     Input  bytes  :                  0
     Output bytes  :                  0
     Input  packets:                  0
     Output packets:                  0
    Transit statistics:
     Input  bytes  :                  0                    0 bps
     Output bytes  :                  0                    0 bps
     Input  packets:                  0                    0 pps
     Output packets:                  0                    0 pps
    Protocol inet, MTU: 1500, Generation: 174, Route table: 0
      Flags: Sendbcast-pkt-to-re

  Logical interface ge-0/3/0.1 (Index 69) (SNMP ifIndex 154)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    Input packets : 0
    Output packets: 0


Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.1      up    up   mpls
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.1      up    up
                           mpls

  Logical interface: ge-0/3/0.1, Index: 69
    Object                Name                   Type                    Index

    Classifier            ipprec-compatibility   ip                         13

Classifier: ipprec-compatibility, Code point type: inet-precedence, Index: 13
  Code point         Forwarding class                   Loss priority
  000                af3                                low
  001                af3                                high
  010                af3                                low
  011                af3                                high
  100                af3                                low
  101                af3                                high
  110                ef1                                low
  111                ef1                                high
Forwarding class                       ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                 0       0          0             low
          normal
  af2                                 1       1          1             low
          normal
  ef2                                 2       2          2             high
          normal
  ef1                                 3       3          3             high
          normal
```



```
  af1                                  4       4          0              low
            normal
```

show class-of-service interface (ACX Series Routers)

```
user@host-g11# show class-of-service interface
Physical interface: at-0/0/0, Index: 130
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

  Logical interface: at-0/0/0.0, Index: 69

  Logical interface: at-0/0/0.32767, Index: 70

Physical interface: at-0/0/1, Index: 133
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

  Logical interface: at-0/0/1.0, Index: 71

  Logical interface: at-0/0/1.32767, Index: 72

Physical interface: ge-0/1/0, Index: 146
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Rewrite                 dscp-default          dscp                       31
Classifier              d1                    dscp                    11331
Classifier              ci                    ieee8021p                 583

  Logical interface: ge-0/1/0.0, Index: 73
Object                  Name                  Type                    Index
Rewrite                 custom-exp            exp (mpls-any)          46413

  Logical interface: ge-0/1/0.1, Index: 74

  Logical interface: ge-0/1/0.32767, Index: 75

Physical interface: ge-0/1/1, Index: 147
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Classifier              ipprec-compatibility  ip                         13

  Logical interface: ge-0/1/1.0, Index: 76

Physical interface: ge-0/1/2, Index: 148
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Rewrite                 ri                    ieee8021p (outer)       35392
Classifier              ci                    ieee8021p                 583

Physical interface: ge-0/1/3, Index: 149
```

```
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

  Logical interface: ge-0/1/3.0, Index: 77
Object                 Name                   Type                    Index
Rewrite                custom-exp2            exp (mpls-any)          53581

Physical interface: ge-0/1/4, Index: 150
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/5, Index: 151
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/6, Index: 152
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/7, Index: 153
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             d1                     dscp                    11331

Physical interface: ge-0/2/0, Index: 154
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/2/1, Index: 155
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

  Logical interface: ge-0/2/1.0, Index: 78

  Logical interface: ge-0/2/1.32767, Index: 79

Physical interface: xe-0/3/0, Index: 156
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
```

```
Classifier             ipprec-compatibility   ip                        13

  Logical interface: xe-0/3/0.0, Index: 80

Physical interface: xe-0/3/1, Index: 157
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                        13

  Logical interface: xe-0/3/1.0, Index: 81

[edit]
user@host-g11#
```

show class-of-service interface (PPPoE Subscriber Interface for Enhanced Subscriber Management)

```
user@host> show class-of-service interface pp0.3221225474
  Logical interface: pp0.3221225475, Index: 3221225475
Object                 Name                   Type                    Index
Traffic-control-profile TC_PROF_100_199_SERIES_UID1006 Output     4294967312
Scheduler-map          SMAP-1_UID1002         Output                4294967327
Rewrite-Output         ieee-rewrite           ieee8021p                60432
Rewrite-Output         rule1                  ip                       50463

    Adjusting application: PPPoE IA tags
      Adjustment type: absolute
      Configured shaping rate: 11000000
      Adjustment value: 5000000
      Adjustment target: node

    Adjusting application: ucac
      Adjustment type: delta
      Configured shaping rate: 5000000
      Adjustment value: 100000
      Adjustment target: node
```

show pfe next-hop

List of Syntax

Syntax on page 92
Syntax (TX Matrix and TX Matrix Plus Routers) on page 92

Syntax

show pfe next-hop
<interface *interface-name*>

Syntax (TX Matrix and TX Matrix Plus Routers)

show pfe next-hop
<fpc *slot*>
<interface *interface-name*>
<lcc *number*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display Packet Forwarding Engine next-hop information.

Options

none—Display all Packet Forwarding Engine next-hop information.

fpc *slot*—(TX Matrix and TX Matrix Plus routers only) (Optional) Show the next hops for a Flexible PIC Concentrator (FPC) slot.

- On a TX Matrix router, if you specify the number of a T640 router by using the **lcc** *number* option (the recommended method), replace ***slot*** with a value from **0** through **7**. Otherwise, replace ***slot*** with a value from **0** through **31**.
- On a TX Matrix Plus router, if you specify the number of a T1600 router by using the **lcc** *number* option (the recommended method), replace ***slot*** with a value from **0** through **7**. Otherwise, replace ***slot*** with a value from **0** through **31**.
- On a TX Matrix Plus router in the TXP-T1600-3D, TXP-T4000-3D, or TXP-Mixed-LCC-3D configuration, if you specify the number of a T1600 or T4000 router by using the **lcc** *number* option (the recommended method), replace ***slot*** with a value from **0** through **7**. Otherwise, replace ***slot*** with a value from **0** through **63**.

For example, the following commands have the same result:

user@host> **show pfe next-hop fpc 1 lcc 1**
user@host> **show pfe next-hop fpc 9**

interface *interface-name*—(Optional) Display the Packet Forwarding Engine next-hop interface.

lcc *number*—(TX Matrix and TX Matrix Plus routers only) (Optional) On a TX Matrix router, display Packet Forwarding Engine next-hop interface for a specific T640 router (or line-card chassis) that is connected to a TX Matrix router. On a TX Matrix Plus router, display Packet Forwarding Engine next-hop interface for the router (or line-card chassis) that is connected to a TX Matrix Plus router.

Replace *number* with the following values depending on the LCC configuration:

- 0 through 3, when T640 routers are connected to a TX Matrix router in a routing matrix.
- 0 through 3, when T1600 routers are connected to a TX Matrix Plus router in a routing matrix.
- 0 through 7, when T1600 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.
- 0, 2, 4, or 6, when T4000 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.

Required Privilege Level admin

Related Documentation

- *Routing Matrix with TXP-T1600 Configuration*
- *Routing Matrix with TXP-T1600-3D Configuration*
- *Routing Matrix with TXP-T4000-3D Configuration*
- *Routing Matrix with a TXP-Mixed-LCC-3D Configuration*

List of Sample Output

show pfe next-hop on page 94
show pfe next-hop fpc (TX Matrix Router) on page 94
show pfe next-hop fpc (TX Matrix Plus Router) on page 95

Output Fields Table 23 on page 94 lists the output fields for the **show pfe next hop** command. Output fields are listed in the approximate order in which they appear.

Table 23: show pfe next-hop Output Fields

Field Name	Field Description
ID	The next-hop ID for the entry.
Type	The next-hop type for the entry.
Interface	The interface to which the next-hop entry is assigned.
Protocol	The protocol type for the next-hop entry.
Encap	Encapsulation type for the next-hop entry.
Next Hop Addr	Next-hop address for the next-hop entry.
MTU	MTU value for the nexthop entry.

Sample Output

show pfe next-hop

```
user@host> show pfe next-hop
Nexthop Info:
   ID        Type     Interface     Protocol          Encap  Next Hop Addr    MTU
-----  --------  -------------  ---------  ------------  ---------------  -----
    4     Mcast  -                   IPv4              -  0.0.0.0              0
    5     Bcast  -                   IPv4              -  -                    0
    7   Discard  -                   IPv4              -  -                    0
    8  MDiscard  -                   IPv4              -  -                    0
    9    Reject  -                   IPv4              -  -                    0
   13     Local  -                   IPv4              -  192.168.4.60         0
   14   Resolve  fxp0.0              IPv4    Unspecified  -                    0
   17     Local  -                   IPv4              -  127.0.0.1            0
   18   Unicast  fxp0.0              IPv4    Unspecified  192.168.4.254        0
   21     Local  -                   IPv4              -  11.1.0.1             0
   22   Unicast  at-0/1/0.0          IPv4       ATM SNAP  11.1.0.2          4482
 ...
```

show pfe next-hop fpc (TX Matrix Router)

```
user@host> show pfe next-hop fpc 1
Slot 1
Nexthop Info:
   ID        Type     Interface     Next Hop Addr    Protocol        Encap     MTU
-----  --------  -------------  ---------------  ----------  ------------  ----
    5     Mcast  -              default                IPv4             -     0
    6     Bcast  -              -                      IPv4             -     0
    8   Discard  -              -                      IPv4             -     0
    9  MDiscard  -              -                      IPv4             -     0
   13     Mcast  -              default                IPV6             -     0
   17  MDiscard  -              -                      IPV6             -     0
   18    Reject  -              -                      IPV6             -     0
```



```
  24   Discard  -              -                 None           -     0
  68     Local  -              192.168.66.113    IPv4           -     0
  69   Resolve  fxp0.0         -                 IPv4  Unspecified    0
  70   Unicast  fxp0.0         192.168.71.254    IPv4  Unspecified    0
 256     Local  -              10.71.71.1        IPv4           -     0
 257     Local  -              127.0.0.1         IPv4           -     0
 258     Mcast.local..1      default           IPv4  Unspecified    0
 259     Bcast.local..1      -                 IPv4  Unspecified    0
 261   Discard.local..1      -                 IPv4  Unspecified    0
 262  MDiscard.local..1      -                 IPv4  Unspecified    0
 269     Mcast.local..1      default           IPV6  Unspecified    0
 271   Discard.local..1      -                 IPV6  Unspecified    0
...
```

show pfe next-hop fpc (TX Matrix Plus Router)

user@host> **show pfe next-hop fpc 0**

```
Slot 0

   ID      Type     Interface     Next Hop Addr    Protocol        Encap    MTU
-----  --------  -------------  ---------------  ----------  ------------  ----
   31     Mcast  -              default                IPv4             -     0
   32     Bcast  -              -                      IPv4             -     0
   34   Discard  -              -                      IPv4             -     0
   35  MDiscard  -              -                      IPv4             -     0
   36    Reject  -              -                      IPv4             -     0
   39     Mcast  -              default                IPv6             -     0
   42   Discard  -              -                      IPv6             -     0
   43  MDiscard  -              -                      IPv6             -     0
   44    Reject  -              -                      IPv6             -     0
   49   Receive  -              -                      MPLS             -     0
   50   Discard  -              -                      MPLS             -     0
  111     Mcast  .local..1      default                IPv4   Unspecified     0
  112     Bcast  .local..1      -                      IPv4   Unspecified     0
  114   Discard  .local..1      -                      IPv4   Unspecified     0
  115  MDiscard  .local..1      -                      IPv4   Unspecified     0
  116    Reject  .local..1      -                      IPv4   Unspecified     0
  119     Mcast  .local..1      default                IPv6   Unspecified     0
  122   Discard  .local..1      -                      IPv6   Unspecified     0
  123  MDiscard  .local..1      -                      IPv6   Unspecified     0
  124    Reject  .local..1      -                      IPv6   Unspecified     0
  191     Mcast  .local..2      default                IPv4   Unspecified     0
  192     Bcast  .local..2      -                      IPv4   Unspecified     0
  194   Discard  .local..2      -                      IPv4   Unspecified     0
  195  MDiscard  .local..2      -                      IPv4   Unspecified     0
  196    Reject  .local..2      -                      IPv4   Unspecified     0
  322     Local  -              10.1.0.5               IPv4             -     0
  323   Resolve  bcm0.0         -                      IPv4   Unspecified     0
  326     Local  -              129.0.0.5              IPv4             -     0
  327   Resolve  bcm0.0         -                      IPv4   Unspecified     0
  328     Local  -              fe80::201:ff:fe01:5
                                                       IPv6             -     0
  329   Receive  bcm0.0         ff02::1:ff01:5         IPv6   Unspecified     0
  330   Receive  bcm0.0         fe80::                 IPv6   Unspecified     0
  331   Resolve  bcm0.0         -                      IPv6   Unspecified     0
  332     Local  -              fec0::a:1:0:5          IPv6             -     0
  333   Receive  bcm0.0         ff02::1:ff00:5         IPv6   Unspecified     0
  334   Receive  bcm0.0         fec0::                 IPv6   Unspecified     0
  335   Resolve  bcm0.0         -                      IPv6   Unspecified     0
  348     Local  -              192.168.178.4          IPv4             -     0
```

```
349   Resolve  em0.0          -                      IPv4   Unspecified     0
350   Unicast  em0.0          192.168.178.126        IPv4   Unspecified     0
357     Local  -              fe80::201:1ff:fe01:5
                                                     IPv6             -     0
512     Local  -              10.255.178.11          IPv4             -     0
513     Local  -              127.0.0.1              IPv4             -     0
515     Local  -              abcd::10:255:178:11
                                                     IPv6             -     0
516     Local  -              fe80::200:ff:fe00:0
                                                     IPv6             -     0
517     Local  -              127.0.0.1              IPv4             -     0
518     Mcast  .local..3      default                IPv4   Unspecified     0
519     Bcast  .local..3      -                      IPv4   Unspecified     0
521   Discard  .local..3      -                      IPv4   Unspecified     0
522  MDiscard  .local..3      -                      IPv4   Unspecified     0
523    Reject  .local..3      -                      IPv4   Unspecified     0
531     Mcast  .local..3      default                IPv6   Unspecified     0
533   Discard  .local..3      -                      IPv6   Unspecified     0
534  MDiscard  .local..3      -                      IPv6   Unspecified     0
535    Reject  .local..3      -                      IPv6   Unspecified     0
539    Mgroup  -              -                      IPv4             -     0
540     Bcast  ge-15/0/3.0    -                      IPv4      Ethernet     0
541   Receive  ge-15/0/3.0    14.2.1.0               IPv4      Ethernet     0
542     Local  -              14.2.1.1               IPv4             -     0
543   Resolve  ge-15/0/3.0    -                      IPv4      Ethernet     0
544     Bcast  ge-31/0/4.0    -                      IPv4      Ethernet     0
545   Receive  ge-31/0/4.0    14.1.1.0               IPv4      Ethernet     0
546     Local  -              14.1.1.1               IPv4             -     0
547   Resolve  ge-31/0/4.0    -                      IPv4      Ethernet     0
548   Unicast  ge-31/0/4.0    14.1.1.2               IPv4      Ethernet     0
549   Unicast  ge-15/0/3.0    14.2.1.2               IPv4      Ethernet     0
550     Bcast  ae1.0          -                      IPv4      Ethernet     0
551   Receive  ae1.0          11.1.1.0               IPv4      Ethernet     0
552     Local  -              11.1.1.1               IPv4             -     0
553   Resolve  ae1.0          -                      IPv4      Ethernet     0
554   Aggreg.  ae1.0          -                      IPv4      Ethernet     0
555   Unicast  ge-23/0/8.0    11.1.1.2               IPv4      Ethernet     0
556   Unicast  ge-7/0/9.0     11.1.1.2               IPv4      Ethernet     0
557   Aggreg.  ae1.0          -                      MPLS      Ethernet     0
558   Unicast  ge-23/0/8.0    -                      MPLS      Ethernet     0
559   Unicast  ge-7/0/9.0     -                      MPLS      Ethernet     0
560   Aggreg.  ae1.0          -                      MPLS      Ethernet     0
561   Unicast  ge-23/0/8.0    -                      MPLS      Ethernet     0
562   Unicast  ge-7/0/9.0     -                      MPLS      Ethernet     0
```


show pfe route

List of Syntax

Syntax on page 97
Syntax (EX Series Switches) on page 97
Syntax (QFX Series) on page 97
Syntax (MX Series) on page 97
Syntax (TX Matrix and TX Matrix Plus Routers) on page 97

Syntax

show pfe route
<<inet6 | ip | iso> <prefix *prefix*> | <table <*table-name*> <index *index*> <prefix *prefix*>>>
<mpls>
<summary>

Syntax (EX Series Switches)

show pfe route
<<inet6 | ip> <prefix *prefix*> | <table <*table-name*> <index *index*> <prefix *prefix*>>>
<mpls>
<summary>

Syntax (QFX Series)

show pfe route
<<inet6 | ip> <prefix *prefix*> | <table <*table-name*> <index *index*> <prefix *prefix*>> <hw (host | lpm | multicast)>>
<<clnp> <prefix *prefix*> | <table <*table-name*> <index *index*> <prefix *prefix*>>>
<mpls>
<summary> <hw>

Syntax (MX Series)

show pfe route
<<inet6 | ip> <prefix *prefix*> | <table <*table-name*> <index *index*> <prefix *prefix*>>>
<dhcp>
<mpls>
<summary>

Syntax (TX Matrix and TX Matrix Plus Routers)

show pfe route
<fpc *slot*>
<<inet6 | ip | iso> <prefix *prefix*> | <table <*table-name*> <index *index*> <prefix *prefix*>>>
<lcc *number*>
<mpls>
<summary>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 13.3 for the MX Series.
Command option **hw** introduced in Junos OS Release 14.1X53-D10 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display the routes in the Packet Forwarding Engine forwarding table. The Packet Forwarding Engine forwards packets between input and output interfaces.

NOTE: The Routing Engine maintains a master copy of the forwarding table. It copies the forwarding table to the Packet Forwarding Engine, which is the part of the router or switch responsible for forwarding packets. To display the routes in the Routing Engine forwarding table, use the show route forwarding table command. For more information, see the CLI Explorer.

Options

none—Display all Packet Forwarding Engine forwarding table information.

clnp—(Optional) Show International Standards Organization (ISO) connectionless-mode network protocol (CLNP) route table information.

dhcp—(Optional) Display Packet Forwarding Engine DHCP-Snooping route table information.

fpc *slot*—(TX Matrix and TX Matrix Plus routers only) (Optional) Show the next hops for a Flexible PIC Concentrator (FPC) slot.

- On a TX Matrix router, if you specify the number of a T640 router by using the **lcc** *number* option (the recommended method), replace ***slot*** with a value from **0** through **7**. Otherwise, replace ***slot*** with a value from **0** through **31**.
- On a TX Matrix Plus router, if you specify the number of a T1600 router by using the **lcc** *number* option (the recommended method), replace ***slot*** with a value from **0** through **7**. Otherwise, replace ***slot*** with a value from **0** through **31**
- On a TX Matrix Plus router in the TXP-T1600-3D, TXP-T4000-3D, or TXP-Mixed-LCC-3D configuration, if you specify the number of a T1600 or T4000 router by using the **lcc** *number* option (the recommended method), replace ***slot*** with a value from **0** through **7**. Otherwise, replace ***slot*** with a value from **0** through **63**.

For example, the following commands have the same result:

```
user@host> show pfe route fpc 1 lcc 1
user@host> show pfe route fpc 9
```

host—(QFX standalone switches, pure mode QFX5100-only VCF and VC, and pure mode QFX3500-only VC) (Optional) Display host routes installed in the on-chip hardware table.

hw—(QFX standalone switches, pure mode QFX5100-only VCF and VC, and pure mode QFX3500-only VC) (Optional) Display routes installed in the on-chip hardware table (as opposed to displaying routes from the routing table and the PFE forwarding table before they are installed in the hardware).

index *index*—(Optional) Display table index.

inet6—(Optional) Display Packet Forwarding Engine IPv6 routes.

ip—(Optional) Display Packet Forwarding Engine IPv4 routes.

iso —(Optional) Display ISO version routing tables.

lcc ***number***—(TX Matrix and TX Matrix Plus routers only) (Optional) On a TX Matrix router, the slot number of the T640 router (or line-card chassis) that houses the FPC. On a TX Matrix Plus router, the slot number of the router (line-card chassis) that houses the FPC.

Replace *number* with the following values depending on the LCC configuration:

- 0 through 3, when T640 routers are connected to a TX Matrix router in a routing matrix.
- 0 through 3, when T1600 routers are connected to a TX Matrix Plus router in a routing matrix.
- 0 through 7, when T1600 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.
- 0, 2, 4, or 6, when T4000 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.

lpm—(QFX standalone switches, pure mode QFX5100-only VCF and VC, and pure mode QFX3500-only VC) (Optional) Display longest prefix match (LPM) routes installed in the on-chip hardware table.

mpls—(Optional) Display Packet Forwarding Engine MPLS information.

multicast—(QFX standalone switches, pure mode QFX5100-only VCF and VC, and pure mode QFX3500-only VC) (Optional) Display multicast routes installed in the on-chip hardware table.

prefix ***prefix***—(Optional) IPv4 or IPv6 prefix for which to show table entries.

summary—(Optional) Display summary of Packet Forwarding Engine information.

table <*table-name*>—(Optional) Display table information.

Required Privilege Level

admin

Related Documentation

- *Routing Matrix with TXP-T1600 Configuration*
- *Routing Matrix with TXP-T1600-3D Configuration*
- *Routing Matrix with TXP-T4000-3D Configuration*
- *Routing Matrix with a TXP-Mixed-LCC-3D Configuration*

List of Sample Output

show pfe route ip on page 101
show pfe route iso on page 101
show pfe route lcc summary (TX Matrix Router) on page 102

show pfe route lcc summary (TX Matrix Plus Router) on page 103
show pfe route summary (MX Series Router) on page 104
show pfe route summary hw (QFX Series, EX4600 Switches, OCX Series) on page 105
show pfe route ip hw host (QFX Series) on page 105

Output Fields Table 24 on page 100 lists the output fields for the **show pfe route** command. Output fields are listed in the approximate order in which they appear.

Table 24: show pfe route Output Fields

Field Name	Field Description
Destination	Destination address for the entry.
NH IP Addr	Next-hop IP address for the entry.
Type	Next-hop type for the entry
NH ID	Next-hop ID for the entry
Encap	Encapsulation type for the next-hop entry.
Interface	Interface to which the next-hop entry is assigned.

Table 25 on page 100 lists the output fields for the QFX Series **show pfe route** hardware table (**hw**) commands. Output fields are listed in the approximate order in which they appear.

Table 25: QFX Series, EX4600 switches, and OCX Series show pfe route Hardware Table Output Fields

Field Name	Field Description
Max	Maximum routing entries per route type.
Used	Number of routing entries consumed per route type.
Free	Number of unused routing entries per route type.
% Free	Percentage of unused routing entries per route type.
Rtt	Internal routing engine index number of the route table.
VRF	Internal hardware index number for the corresponding route table.
Destination	Destination address for the entry.
Type	(**show pfe route summary hw**)—Route type for the entry: IPv4 or IPv6 route, and host, LPM, or multicast route. (**show pfe route (ip | inet6) hw**)—Next-hop type for the entry.
NH ID	Next-hop ID for the entry

Table 25: QFX Series, EX4600 switches, and OCX Series show pfe route Hardware Table Output Fields (continued)

Field Name	Field Description
Interface	Interface to which the next-hop entry is assigned.
HW NH-ID	Internal hardware index number of the next-hop.
Src-MAC-Address	Source MAC address.
Port	Port number.
Dst-MAC-Address	Destination MAC address.
VLAN	ID of the multicast group VLAN.
GROUP	Internal hardware index number of the multicast group next-hop.
CLASS	Internal class number of the multicast group.

Sample Output

show pfe route ip

```
user@host> show pfe route ip

IPv4 Route Table 0, default.0, 0x0:
Destination                       NH IP Addr      Type     NH ID Interface
--------------------------------- --------------- -------- ----- ---------
default                                            Discard     8
127.0.0.1                         127.0.0.1          Local   256
172.16/12                         192.168.71.254   Unicast    68 fxp0.0
192.168.0/18                      192.168.71.254   Unicast    68 fxp0.0
192.168.40/22                     192.168.71.254   Unicast    68 fxp0.0
192.168.64/18                     192.168.71.254   Unicast    68 fxp0.0
192.168.64/21                                      Resolve    67 fxp0.0
192.168.71.249                    192.168.71.249     Local    66
192.168.220.0/30                                   Resolve   303 fe-0/0/0.0
192.168.220.0                     192.168.220.0    Receive   301 fe-0/0/0.0
224.0.0.1                                            Mcast     5
255.255.255.255                                      Bcast     6

...
```

show pfe route iso

```
user@host# show pfe route iso

CLNS Route Table 0, CLNP.0, 0x0:
Destination                              Type     NH ID Interface
---------------------------------------- -------- ----- ---------
default                                    Reject    60
47.0005.80ff.f800.0000.0108.0001.0102.5508.2159/152     Local    514
49.0001.00a0.c96b.c491/72                   Local    536
```


show pfe route lcc summary (TX Matrix Router)

```
user@host> show pfe route lcc 2 summary

Slot 0

IPv4 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           43        3081
1                  4         281

MPLS Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default            1          68

IPV6 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default            9         717
1                  5         389

Slot 1


IPv4 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           43        3081
1                  4         281

MPLS Route Tables:

Index          Routes     Size(b)
--------  ----------  ----------
Default            1          68

IPV6 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default            9         717
1                  5         389

Slot 16

IPv4 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           41        2938
1                  4         281

MPLS Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default            1          68

IPV6 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default            9         717
```

```
1                 5         389

Slot 17

IPv4 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default          41        2938
1                 4         281


MPLS Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           1          68

IPV6 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           9         717
1                 5         389
```

show pfe route lcc summary (TX Matrix Plus Router)

user@host> **show pfe route lcc 2 summary**

```
Slot 0


IPv4 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default          25        2266
1                 9         815
2                 6         545
3                 5         453
4                15        1371
5                 5         453
6                13        1187

MPLS Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           1          88
4                 5         452

IPv6 Route Tables:
Index          Routes     Size(b)
--------  ----------  ----------
Default           7         697
1                13        1305
3                 4         385
4                 4         385
5                 4         385
6                18        1833

Slot 6


IPv4 Route Tables:
Index          Routes     Size(b)
```

```
--------  ----------  ----------
Default           25        2266
1                  9         815
2                  6         545
3                  5         453
4                 15        1371
5                  5         453
6                 13        1187

MPLS Route Tables:
Index         Routes     Size(b)
--------  ----------  ----------
Default            1          88
4                  5         452

IPv6 Route Tables:
Index         Routes     Size(b)
--------  ----------  ----------
Default            7         697
1                 13        1305
3                  4         385
4                  4         385
5                  4         385
6                 18        1833
...
```

show pfe route summary (MX Series Router)

```
user@host> show pfe route summary

Slot 0


DHCP-Snooping Route Tables:
Index         Routes     Size(b)
--------  ----------  ----------
Default            1         144

IPv4 Route Tables:
Index         Routes     Size(b)
--------  ----------  ----------
Default           25        2266
1                  9         815
2                  6         545
3                  5         453
4                 15        1371
5                  5         453
6                 13        1187

MPLS Route Tables:
Index         Routes     Size(b)
--------  ----------  ----------
Default            1          88
4                  5         452

IPv6 Route Tables:
Index         Routes     Size(b)
--------  ----------  ----------
Default            7         697
1                 13        1305
3                  4         385
```

```
4                  4         385
5                  4         385
6                 18        1833

...
```

show pfe route summary hw (QFX Series, EX4600 Switches, OCX Series)

```
user@switch> show pfe route summary hw
Slot 0
Unit: 0
Profile active: l2-profile-three
Type            Max       Used      Free      % free
----------------------------------------------------
IPv4 Host       8192      103       8073      98.55
IPv4 LPM        16384     9         16369     99.91
IPv4 Mcast      4096      2         4037      98.56

IPv6 Host       4096      6         4037      98.56
IPv6 LPM(< 64)  8192      3         8185      99.91
IPv6 LPM(> 64)  256       1     255        99.61
IPv6 Mcast      2048      0         2019      98.58
```

show pfe route ip hw host (QFX Series)

```
user@switch> show pfe route ip hw host
Slot 0
Unit: 0
IPv4 Host entries present: 103
Rtt  VRF  Destination                      Type     NH-ID     Interface
         HW NH-ID  Src-MAC-Address   Port Dst-MAC-Address
_________________________________________________________________________________________
4    3    255.255.255.255                  Bcast    1695      .local.    .4
ifl 550   100003    00:00:00:01:02:03 127  00:00:00:01:02:03
0    1    200.1.1.42                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.56                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.61                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    11.1.1.2                         Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.73                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.76                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.18                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.5                        Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.23                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    101.1.1.255                      Bcast    1664      ae0        .0
ifl 544   100003    00:00:00:01:02:03 127  00:00:00:01:02:03
0    1    200.1.1.40                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23
0    1    200.1.1.58                       Unicast  1743      et-0/1/1   .0
ifl 559   100268    84:18:88:de:96:fd 53   00:00:00:21:12:23. . .
. . .
```


show pfe terse

List of Syntax

Syntax on page 107
Syntax (TX Matrix and TX Matrix Plus Router) on page 107
Syntax (MX Series Router) on page 107

Syntax

```
show pfe terse
```

Syntax (TX Matrix and TX Matrix Plus Router)

```
show pfe terse
<lcc number | scc>
<sfc number>
```

Syntax (MX Series Router)

```
show pfe terse
<all-members>
<local>
<member member-id>
```

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display Packet Forwarding Engine status information.

Options

none—Display brief information about the Packet Forwarding Engine.

all-members—(MX Series routers only) (Optional) Display Packet Forwarding Engine status information for all members in the Virtual Chassis configuration.

lcc ***number***—(TX Matrix and TX Matrix Plus routers only) (Optional) On a TX Matrix router, display Packet Forwarding Engine information for a T640 router (or line-card chassis) that is connected to a TX Matrix router. On a TX Matrix Plus router, display Packet Forwarding Engine information for the router (or line-card chassis) that is connected to a TX Matrix Plus router.

Replace *number* with the following values depending on the LCC configuration:

- 0 through 3, when T640 routers are connected to a TX Matrix router in a routing matrix.
- 0 through 3, when T1600 routers are connected to a TX Matrix Plus router in a routing matrix.
- 0 through 7, when T1600 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.
- 0, 2, 4, or 6, when T4000 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.

local—(MX Series routers only) (Optional) Display Packet Forwarding Engine status information for the local Virtual Chassis member.

member *member-id*—(MX Series routers only) (Optional) Display Packet Forwarding Engine status information for the specified member of the Virtual Chassis configuration. Replace ***member-id*** with a value of 0 or 1.

scc—(TX Matrix routers only) (Optional) Display Packet Forwarding Engine information for the TX Matrix router (or switch-card chassis).

sfc—(TX Matrix Plus routers only) (Optional) Display Packet Forwarding Engine information for the TX Matrix Plus router (or switch-fabric chassis).

Required Privilege Level admin

List of Sample Output
show pfe terse (TX Matrix Router) on page 108
show pfe terse (TX Matrix Plus Router) on page 108
show pfe terse sfc (TX Matrix Plus Router) on page 108

Sample Output

show pfe terse (TX Matrix Router)

```
user@host> show pfe terse
Slot Type Slot    State   Flags Uptime
  0  SFM  Present Online  0x0bf 01:25:42
  2  SFM  Present Online  0x0bf 01:25:40
  0  FPC  Present Online  0x102 01:25:57
  1  FPC  Present Online  0x102 01:25:55
  2  FPC  Present Online  0x102 01:25:53
```

show pfe terse (TX Matrix Plus Router)

```
user@host> show pfe terse
sfc0-re0:
--------------------------------------------------------------------------
Slot Type Slot      State     Uptime
 0   LCC  Present   Online    2d 05:26

lcc0-re0:
--------------------------------------------------------------------------
Slot Type Slot      State     Uptime
 0   GFPC Present   Online    2d 05:25
 1   GFPC Present   Online    2d 05:25
```

show pfe terse sfc (TX Matrix Plus Router)

```
user@host> show pfe terse sfc 0
sfc0-re0:
--------------------------------------------------------------------------
Slot Type Slot      State     Uptime
 0   LCC  Present   Online    2d 05:25
```


show pfe version

Syntax show pfe version <brief | detail>

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Display Packet Forwarding Engine version information.

Options **brief | detail**—Display the specified level of output.

Required Privilege Level admin

List of Sample Output show pfe version brief on page 109
show pfe version detail on page 109

Sample Output

show pfe version brief

```
user@host> show pfe version brief
PFED release 11.1D0 built by builder on 2010-11-11 05:16:11 UTC
```

show pfe version detail

```
user@host> show pfe version detail
PFED release 11.1D0 built by builder on 2010-11-11 05:16:11 UTC

device01.example.com:/volume/build/junos/rpd_feb11/11.1/development/20101111.0/obj-i386/
junos/usr.sbin/pfed
```

PART 2

Classifying and Rewriting Traffic

CHAPTER 4

Using Classifiers, Forwarding Classes, and Rewrite Rules

Understanding Default CoS Scheduling and Classification

If you do not explicitly configure classifiers and apply them to interfaces, the switch uses the default classifier for ingress traffic. If you do not configure hierarchical scheduling (also known as enhanced transmission selection (ETS)) on an interface, the switch uses the default schedulers for egress traffic. Default classification maps all traffic into default forwarding classes (best-effort, fcoe, no-loss, network-control, and mcast).

Hierarchical scheduling groups IEEE 802.1p priorities (IEEE 802.1p code points, which classifiers map to forwarding classes, which in turn are mapped to output queues) into priority groups (forwarding class sets). If you use only the default traffic scheduling and classification, the switch automatically creates a default priority group that contains all of the priorities (which are mapped to forwarding classes and output queues), and assigns 100 percent of the port output bandwidth to that priority group. The forwarding classes (queues) in the default forwarding class set receive bandwidth based on the default classifier settings. The default priority group is transparent. It does not appear in the configuration (it is used for Data Center Bridging Capability Exchange (DCBX) protocol advertisement on QFX Series switches).

NOTE: If you explicitly configure one or more priority groups on an interface, any forwarding class that is not assigned to a priority group on that interface receives *no bandwidth*. This means that if you configure hierarchical scheduling on an interface, every forwarding class (priority) that you want to forward traffic on that interface must belong to a forwarding class set (priority group).

The following sections describe:

Default Classification

The default classifiers assign unicast and multicast best-effort and network-control ingress traffic to forwarding classes and loss priorities. The switch applies default unicast DSCP, unicast IEEE 802.1, and multidestination classifiers to each interface that does not have explicitly configured classifiers. If you explicitly configure one type of classifier but not other types of classifiers, the system uses only the configured classifier and does not use default classifiers for other types of traffic. There are two different default unicast IEEE 802.1 classifiers, a trusted classifier and an untrusted classifier. On QFX Series switches, the trusted classifier is the default IEEE classifier.

Table 26 on page 115 shows the default mapping of DSCP code-point values to unicast forwarding classes and loss priorities for DSCP IP and DCSP IPv6.

Table 26: Default DSCP IP and IPv6 Unicast Classifiers

Code Point	Forwarding Class	Loss Priority
000000 (be)	best-effort	low
000001	best-effort	low
000010	best-effort	low
000011	best-effort	low
000100	best-effort	low
000101	best-effort	low
000110	best-effort	low
000111	best-effort	low
001000 (cs1)	best-effort	low
001001	best-effort	low
001010 (af11)	best-effort	low
001011	best-effort	low
001100 (af12)	best-effort	low
001101	best-effort	low
001110 (af13)	best-effort	low
001111	best-effort	low
010000 (cs2)	best-effort	low
010001	best-effort	low
010010 (af21)	best-effort	low
010011	best-effort	low
010100 (af22)	best-effort	low
010101	best-effort	low
010110 (af23)	best-effort	low
010111	best-effort	low

Table 26: Default DSCP IP and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
011000 (cs3)	best-effort	low
011001	best-effort	low
011010 (af31)	best-effort	low
011011	best-effort	low
011100 (af32)	best-effort	low
011101	best-effort	low
011110 (af33)	best-effort	low
011111	best-effort	low
100000 (cs4)	best-effort	low
100001	best-effort	low
100010 (af41)	best-effort	low
100011	best-effort	low
100100 (af42)	best-effort	low
100101	best-effort	low
100110 (af43)	best-effort	low
100111	best-effort	low
101000 (cs5)	best-effort	low
101001	best-effort	low
101011	best-effort	low
101100	best-effort	low
101101	best-effort	low
101110 (ef)	best-effort	low
101111	best-effort	low
110000 (nc1)	network-control	low

Table 26: Default DSCP IP and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
110001	network-control	low
110010	network-control	low
110011	network-control	low
110100	network-control	low
110101	network-control	low
110110	network-control	low
110111	network-control	low
111000 (nc2)	network-control	low
111001	network-control	low
111010	network-control	low
111011	network-control	low
111100	network-control	low
111101	network-control	low
111110	network-control	low
111111	network-control	low

NOTE: There are no default DSCP IP or IPv6 multidestination classifiers for multidestination traffic. DSCP IPv6 multidestination classifiers are not supported for multidestination traffic.

Table 27 on page 117 shows the default trusted classifier mapping of IEEE 802.1 code-point values to unicast forwarding classes and loss priorities .

Table 27: Default IEEE 802.1 Unicast Classifiers (Trusted)

Code Point	Forwarding Class	Loss Priority
be (000)	best-effort	low
be1 (001)	best-effort	low

Table 27: Default IEEE 802.1 Unicast Classifiers (Trusted) (continued)

Code Point	Forwarding Class	Loss Priority
ef (010)	best-effort	low
ef1 (011)	fcoe	low
af11 (100)	no-loss	low
af12 (101)	best-effort	low
nc1 (110)	network-control	low
nc2 (111)	network-control	low

Table 28 on page 118 shows the default untrusted mapping of IEEE 802.1p code-point values to unicast forwarding classes and loss priorities.

Table 28: Default IEEE 802.1 Unicast Classifiers (Untrusted)

Code Point	Forwarding Class	Loss Priority
000	best-effort	low
001	best-effort	low
010	best-effort	low
011	best-effort	low
100	best-effort	low
101	best-effort	low
110	best-effort	low
111	best-effort	low

Table 29 on page 118 shows the default mapping of IEEE 802.1 code-point values to multidestination (multicast, broadcast, and destination lookup fail traffic) forwarding classes and loss priorities.

Table 29: Default IEEE 802.1 Multidestination Classifiers

Code Point	Forwarding Class	Loss Priority
be (000)	mcast	low
be1 (001)	mcast	low

Table 29: Default IEEE 802.1 Multidestination Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
ef (010)	mcast	low
ef1 (011)	mcast	low
af11 (100)	mcast	low
af12 (101)	mcast	low
nc1 (110)	mcast	low
nc2 (111)	mcast	low

Default Scheduling

The default schedulers allocate egress bandwidth resources to unicast and multicast egress traffic as shown in Table 30 on page 119:

Table 30: Default Scheduler Configuration

Default Scheduler and Queue Number	Transmit Rate (Minimum Guaranteed Bandwidth)	Shaping Rate (Maximum Bandwidth)	Excess Bandwidth Sharing	Priority	Buffer Size
best-effort forwarding class scheduler (queue 0)	5%	None	5%	low	5%
fcoe forwarding class scheduler (queue 3) NOTE: Do not map traffic to the default fcoe forwarding class. The default fcoe forwarding class carries the no-loss packet drop attribute, which is not supported on OCX Series switches. Other switches use the fcoe forwarding class for Fibre Channel over Ethernet (FCoE) lossless Layer 2 transport, but OCX Series switches do not support FCoE or lossless Layer 2 transport. If you want to use queue 3, configure a new forwarding class *without* the no-loss packet drop attribute and map it to queue 3.	35%	None	35%	low	35%

Table 30: Default Scheduler Configuration (continued)

Default Scheduler and Queue Number	Transmit Rate (Minimum Guaranteed Bandwidth)	Shaping Rate (Maximum Bandwidth)	Excess Bandwidth Sharing	Priority	Buffer Size
no-loss forwarding class scheduler (queue 4) NOTE: Do not map traffic to the default no-loss forwarding class. The default no-loss forwarding class carries the no-loss packet drop attribute, which is not supported on OCX Series switches. Other switches use the no-loss forwarding class for lossless Layer 2 transport, but OCX Series switches do not support lossless Layer 2 transport. If you want to use queue 4, configure a new forwarding class *without* the no-loss packet drop attribute and map it to queue 4.	35%	None	35%	low	35%
network-control forwarding class scheduler (queue 7)	5%	None	5%	low	5%
mcast forwarding class scheduler (queue 8)	20%	None	20%	low	20%

NOTE: **The minimum guaranteed bandwidth rate also determines the amount of excess (extra) bandwidth that the queue can share. Extra bandwidth is allocated to queues in proportion to the minimum guaranteed bandwidth (transmit rate) of each queue.**

The default DSCP classifier maps traffic only to the best-effort (queue 0), network-control (queue 7), and mcast (queue 8) forwarding classes. Only the five default schedulers shown in Table 30 on page 119 have port resources (for example, bandwidth) mapped to them by default, but only the queues that are forwarding traffic use port resources. So even though 35 percent of port bandwidth is allocated to the fcoe and no-loss schedulers, that bandwidth is available to the best-effort, network-control, and mcast schedulers because no traffic is mapped by default to the fcoe and no-loss forwarding classes. The amount of default bandwidth each forwarding class receives on a port is proportional to the default scheduler transmit rate. Unused bandwidth is shared among queues that need more bandwidth. (You can configure schedulers and forwarding classes to allocate bandwidth to other queues or to change the default bandwidth of a default queue.) In addition, multidestination queue 11 receives enough bandwidth from the default multidestination scheduler to handle CPU-generated multidestination traffic.

Default hierarchical scheduling divides the total port bandwidth between two groups of traffic: unicast traffic and multidestination traffic. By default, unicast traffic consists of queue 0 (**best-effort** forwarding class) and queue 7 (**network-control** forwarding class) because no traffic is mapped by default to queue 3 (**fcoe** forwarding class) or queue 4 (**no-loss** forwarding class). Unicast traffic receives and shares a total of 80 percent of the port bandwidth. By default, multidestination traffic (**mcast** queue 8) receives a total of 20 percent of the port bandwidth. So on a 10-Gigabit port, unicast traffic receives 8-Gbps of bandwidth and multidestination traffic receives 2-Gbps of bandwidth.

NOTE: **Multidestination queue 11 also receives a small amount of default bandwidth from the multidestination scheduler. CPU-generated multidestination traffic uses queue 11, so you might see a small number of packets egress from queue 11. In addition, in the unlikely case that firewall filter match conditions map multidestination traffic to a unicast forwarding class, that traffic uses queue 11.**

Default scheduling uses weighted round-robin (WRR) scheduling. Each queue receives a portion (weight) of the total available interface bandwidth. The scheduling weight is based on the transmit rate of the default scheduler for that queue. For example, queue 7 receives a default scheduling weight of 5 percent of the available bandwidth, and queue 8 receives a default scheduling weight of 20 percent of the available bandwidth. Queues are mapped to forwarding classes, so forwarding classes receive the default bandwidth for the queues to which they are mapped.

You should explicitly map traffic to non-default (unconfigured) queues and create schedulers to allocate bandwidth to those queues if you want to use them to forward traffic. By default, unicast queues 1, 2, 5, and 6 are unconfigured, and multidestination queues 9, 10, and 11 are unconfigured.

NOTE: **If you want to map traffic to queue 3 or to queue 4 by configuring a DSCP classifier and mapping DSCP code points to forwarding classes , then you must explicitly configure forwarding classes that do not have the no-loss packet attribute and map them to those queues. Do not map traffic to the default fcoe forwarding class (queue 3) or to the default no-loss forwarding class (queue 4). If you map traffic to queue 3 or queue 4, unless you configure a scheduler, that traffic uses the default scheduler and receives 35 percent of the port bandwidth by default.**

Unconfigured queues have a default scheduling weight of 1 so that they can receive a small amount of bandwidth in case they need to forward traffic. (However, queue 11 can use more of the default multidestination scheduler bandwidth if necessary to handle CPU-generated multidestination traffic.)

NOTE: All four multidestination queues have a scheduling weight of 1. Because by default multidestination traffic goes to queue 8, queue 8 receives almost all of the multidestination bandwidth. (There is no traffic on queue 9 and queue 10, and very little traffic on queue 11, so there is almost no competition for multidestination bandwidth.)

However, if you explicitly configure queue 9, 10, or 11 (by mapping code points to the unconfigured multidestination forwarding classes using the multidestination classifier), the explicitly configured queues share the multidestination scheduler bandwidth equally with default queue 8, because all of the queues have the same scheduling weight (1). To ensure that multidestination bandwidth is allocated to each queue properly and that the bandwidth allocation to the default queue (8) is not reduced too much, we strongly recommend that you configure a scheduler if you explicitly classify traffic into queue 9, 10, or 11.

If you map traffic to an unconfigured queue, the queue receives only the amount of group bandwidth proportional to its default weight (1). The actual amount of bandwidth an unconfigured queue receives depends on how much bandwidth the other queues in the group are using.

If the other unicast queues use less than their allocated amount of bandwidth, the unconfigured queues can share the unused bandwidth. Sharing unused bandwidth is one of the key advantages of hierarchical port scheduling. Configured queues have higher priority for bandwidth than unconfigured queues, so if a configured queue needs more bandwidth, then less bandwidth is available for unconfigured queues. Unconfigured queues always receive a minimum amount of bandwidth based on their scheduling weight (1). If you map traffic to an unconfigured queue, to allocate bandwidth to that queue, configure a scheduler for the forwarding class that is mapped to the queue.

Default Scheduling and Classification Summary

If you do not configure hierarchical scheduling on an interface:

- Default classifiers classify ingress traffic. The default DSCP classifier classifies unicast traffic into two queues, queue 0 (best-effort) and queue 7 (network-control). By default, multidestination traffic is classified into queue 8 (mcast).
- Default schedulers schedule egress traffic.
- A single default priority group (fc-set) receives 100 percent of the port bandwidth. All priorities (forwarding classes) are assigned to the default priority group and receive bandwidth based on their default schedulers. The default priority group is generated automatically and is not user-configurable.

Related Documentation

- Understanding CoS Packet Flow on page 23
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Understanding Default CoS Settings on page 25

- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- Example: Configuring Queue Schedulers on page 293

Understanding CoS Classifiers

Packet classification maps incoming packets to a particular class-of-service (CoS) servicing level. Classifiers map packets to a forwarding class and a loss priority, and assign packets to output queues based on the forwarding class. There are three general types of classifiers:

- Behavior aggregate (BA) classifiers—DSCP and DSCP IPv6 classify IP and IPv6 traffic, and IEEE 802.1p classifiers classify all other traffic.
- Fixed classifiers—Fixed classifiers classify all ingress traffic on a physical interface into one forwarding class, regardless of the CoS bits in the packet header.
- Multifield (MF) classifiers—MF classifiers classify traffic based on more than one field in the packet header and take precedence over BA and fixed classifiers.

By default, OCX Series switches use the default DSCP classifier. The default DSCP classifier maps incoming traffic to the best-effort (queue 0) and network-control (queue 7) forwarding classes.

NOTE: If you configure a classifier, do not map traffic to the default fcoe (queue 3) or no-loss (queue 4) forwarding classes. (Do not map DSCP code points to the fcoe or no-loss default forwarding classes. On other switches, the fcoe and no-loss default forwarding classes provide lossless transport for Layer 2 traffic. OCX Series switches do not support lossless Layer 2 transport. You can map traffic to queue 3 and queue 4 only if you configure forwarding classes that do not have the no-loss packet drop attribute, and map them to those queues.)

- Interfaces and Output Queues on page 123
- Behavior Aggregate Classifiers on page 124
- Fixed Classifiers on page 129
- Multifield Classifiers on page 129

Interfaces and Output Queues

On Gigabit interfaces, 10-Gigabit interfaces, and link aggregation (LAG) interfaces, you can apply classifiers to Layer 3 physical interfaces if the Layer 3 physical interface has at least one defined logical interface. Classifiers applied to Layer 3 physical interfaces are used on all logical interfaces on that physical interface. *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces* describes the interaction between classifiers and interfaces in greater detail (see "Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces" on page 162 for OCX Series switch information).

You can configure both a BA classifier and an MF classifier on an interface. If you do this, the BA classification is performed first, and then the MF classification is performed. If the two classification results conflict, the MF classification result overrides the BA classification result.

You cannot configure a fixed classifier and a BA classifier on the same interface.

You can configure both a DSCP or a DSCP IPv6 classifier and an IEEE 802.1p classifier on the same interface. IP traffic uses the DSCP or DSCP IPv6 classifier. All other traffic uses the IEEE classifier. You can configure only one DSCP classifier on a physical interface (either one DSCP classifier or one DSCP IPv6 classifier, but not both).

You can create unicast BA classifiers for unicast traffic and multicast BA classifiers for multidestination traffic, which includes multicast, broadcast, and destination lookup fail (DLF) traffic. You cannot assign unicast traffic and multidestination traffic to the same BA classifier.

On each interface, the switch has separate output queues for unicast traffic and for multidestination traffic:

- The switch supports 12 output queues, with 8 queues dedicated to unicast traffic and 4 queues dedicated to multidestination traffic.
- Queues 0 through 7 are unicast traffic queues. You can apply only unicast BA classifiers to unicast queues. A unicast BA classifier should contain only forwarding classes that are mapped to unicast queues.
- Queues 8 through 11 are multidestination traffic queues. You can apply only multidestination BA classifiers to multidestination queues. A multidestination BA classifier should contain only forwarding classes that are mapped to multidestination queues.

You can apply unicast classifiers to one or more interfaces. Multidestination classifiers apply to all of the switch interfaces and cannot be applied to individual interfaces. Use the DSCP multidestination classifier for both IP and IPv6 multidestination traffic. The DSCP IPv6 classifier is not supported for multidestination traffic.

Behavior Aggregate Classifiers

The behavior aggregate classifier maps a class-of-service (CoS) value to a forwarding class and loss priority. The forwarding class determines the output queue. A scheduler uses the loss priority to control packet discard during periods of congestion by associating different drop profiles with different loss priorities.

The switch supports two types of BA classifiers:

- Differentiated Services Code Point (DSCP) for IP DiffServ (IP and IPv6)
- IEEE 802.1p CoS bits

BA classifiers are based on fixed-length fields, which makes them computationally more efficient than MF classifiers. Therefore, core devices, which handle high traffic volumes, are normally configured to perform BA classification.

Unless you explicitly configure a classifier and apply it to interfaces, OCX Series switch interfaces use the default DSCP classifier shown in Table 31 on page 125.

Table 31: Default DSCP IP and IPv6 Unicast Classifiers

Code Point	Forwarding Class	Loss Priority
000000 (be)	best-effort	low
000001	best-effort	low
000010	best-effort	low
000011	best-effort	low
000100	best-effort	low
000101	best-effort	low
000110	best-effort	low
000111	best-effort	low
001000 (cs1)	best-effort	low
001001	best-effort	low
001010 (af11)	best-effort	low
001011	best-effort	low
001100 (af12)	best-effort	low
001101	best-effort	low
001110 (af13)	best-effort	low
001111	best-effort	low
010000 (cs2)	best-effort	low
010001	best-effort	low
010010 (af21)	best-effort	low
010011	best-effort	low

Table 31: Default DSCP IP and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
010100 (af22)	best-effort	low
010101	best-effort	low
010110 (af23)	best-effort	low
010111	best-effort	low
011000 (cs3)	best-effort	low
011001	best-effort	low
011010 (af31)	best-effort	low
011011	best-effort	low
011100 (af32)	best-effort	low
011101	best-effort	low
011110 (af33)	best-effort	low
011111	best-effort	low
100000 (cs4)	best-effort	low
100001	best-effort	low
100010 (af41)	best-effort	low
100011	best-effort	low
100100 (af42)	best-effort	low
100101	best-effort	low
100110 (af43)	best-effort	low
100111	best-effort	low
101000 (cs5)	best-effort	low
101001	best-effort	low
101011	best-effort	low
101100	best-effort	low

Table 31: Default DSCP IP and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
101101	best-effort	low
101110 (ef)	best-effort	low
101111	best-effort	low
110000 (nc1)	network-control	low
110001	network-control	low
110010	network-control	low
110011	network-control	low
110100	network-control	low
110101	network-control	low
110110	network-control	low
110111	network-control	low
111000 (nc2)	network-control	low
111001	network-control	low
111010	network-control	low
111011	network-control	low
111100	network-control	low
111101	network-control	low
111110	network-control	low
111111	network-control	low

Unicast and multicast traffic cannot share the same classifier. You can map unicast traffic and multicast traffic to the same classifier CoS value, but the unicast traffic must belong to a unicast classifier and the multicast traffic must belong to a multidestination classifier.

Default Behavior Aggregate Classification

Juniper Networks Junos OS automatically assigns implicit default classifiers to all logical interfaces based on the type of interface. Table 32 on page 128 lists different types of interfaces and the corresponding implicit default BA classifiers.

Table 32: Default BA Classification

Type of Interface	Default BA Classification
Layer 3 interface	**dscp-default**

NOTE: There are default BA classifiers for the **best-effort**, **fcoe**, **no-loss**, **network-control**, and **mcast** forwarding classes.

When you explicitly associate a unicast classifier with a logical interface, you override the default unicast classifier with the explicit unicast classifier.

NOTE: You can apply only one classifier of each type, DSCP and IEEE 802.1p, to an interface. If both types of classifiers are present, DSCP classifiers take precedence over IEEE 802.1p classifiers.

Importing a Classifier

You can use any existing classifier, including the default classifiers, as the basis for defining a new classifier. You accomplish this using the **import** statement.

The imported classifier is used as a template and is not modified. The modifications you make become part of a new classifier (and a new template) identified by the name of the new classifier. Whenever you commit a configuration that assigns a new forwarding class-name and loss-priority value to a code-point alias or set of bits, it replaces that entry in the new classifier template. As a result, you must explicitly specify every CoS value in every designation that requires modification.

Multidestination Classifiers

Multidestination classifiers are applied to all interfaces and cannot be applied to individual interfaces. You can configure both a DSCP multidestination classifier and an IEEE multidestination classifer. IP and IPv6 traffic use the DSCP classifier, and all other traffic uses the IEEE classifier.

DSCP IPv6 multidestination classifiers are not supported, so IPv6 traffic uses the DSCP multidestination classifier.

The default multidestination classifier is the IEEE 802.1p multidestination classifier.

Fixed Classifiers

Fixed classifiers map all traffic on a physical interface to a forwarding class and a loss priority. (As opposed to BA classifiers, which map traffic into multiple different forwarding classes based on the IEEE 802.1p CoS bits field value in the VLAN header or the DSCP field value in the type-of-service bits in the packet IP header.) Each forwarding class maps to an output queue. However, when you use a fixed classifier, regardless of the CoS or DSCP bits, all Incoming traffic is classified into the forwarding class specified in the fixed classifier. A scheduler uses the loss priority to control packet discard during periods of congestion by associating different drop profiles with different loss priorities.

You cannot configure a fixed classifier and a DSCP or IEEE 802.1p BA classifier on the same interface. If you configure a fixed classifier on an interface, you cannot configure a DSCP or an IEEE classifier on that interface. If you configure a DSCP classifier, an IEEE classifier, or both classifiers on an interface, you cannot configure a fixed classifier on that interface.

To switch from a fixed classifier to a BA classifier or to switch from a BA classifier to a fixed classifier, deactivate the existing classifier attachment on the interface, and then attach the new classifier to the interface.

Multifield Classifiers

Multifield classifiers examine multiple fields in a packet such as source and destination addresses and source and destination port numbers of the packet. With MF classifiers, you set the forwarding class and loss priority of a packet based on firewall filter rules.

MF classification is normally performed at the network edge because of the general lack of DiffServ code point (DSCP) support in end-user applications. On a switch at the edge of a network, an MF classifier provides the filtering functionality that scans through a variety of packet fields to determine the forwarding class for a packet. Typically, a classifier performs matching operations on the selected fields against a configured value.

Related Documentation

- Understanding CoS Packet Flow on page 23
- Understanding Default CoS Settings on page 25
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Example: Configuring Unicast Classifiers*
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Example: Configuring Multidestination (Multicast, Broadcast, DLF) Classifiers*

Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p)

Packet classification associates incoming packets with a particular CoS servicing level. Behavior aggregate (BA) classifiers examine the Differentiated Services code point (DSCP or DSCP IPv6) value, the IEEE 802.1p CoS value, or the MPLS EXP value in the packet header to determine the CoS settings applied to the packet. (See *Configuring a Global MPLS EXP Classifier* to learn how to define EXP classifiers for MPLS traffic.) BA classifiers allow you to set the forwarding class and loss priority of a packet based on the incoming CoS value.

NOTE: OCX Series switches do not support MPLS EXP classifiers.

On switches except QFX10000 and NFX Series devices, unicast traffic must use different classifiers than multidestination (mulitcast, broadcast, and destination lookup fail) traffic. You use the **multi-destination** statement at the **[edit class-of-service]** hierarhcy level to configure a multidestination BA classifier.

On QFX10000 switches and NFX Series devices, unicast and multidestination traffic use the same classifiers and forwarding classes.

Multidestination classifiers apply to all of the switch interfaces and handle multicast, broadcast, and destination lookup fail (DLF) traffic. You cannot apply a multidestination classifier to a single interface or to a range of interfaces.

To configure a DSCP, DSCP IPv6, or IEEE 802.1p BA classifier using the CLI:

1. Create a BA classifier:

   - To create a DSCP, DSCP IPv6, or IEEE 802.1p BA classifier based on the default classifier, import the default DSCP, DSCP IPv6, or IEEE 802.1p classifier and associate it with a forwarding class, a loss priority, and a code point:

     ```
     [edit class-of-service classifiers]
     user@switch# set (dscp | dscp-ipv6 | ieee-802.1) classifier-name import default
     forwarding-class forwarding-class-name loss-priority level code-points [aliases]
     [bit-patterns]
     ```

   - To create a BA classifier that is not based on the default classifier, create a DSCP, DSCP IPv6, or IEEE 802.1p classifier and associate it with a forwarding class, a loss priority, and a code point:

     ```
     [edit class-of-service classifiers]
     user@switch# set (dscp | dscp-ipv6 | ieee-802.1) classifier-name forwarding-class
     forwarding-class-name loss-priority level code-points [aliases] [bit-patterns]
     ```

2. For multidestination traffic, except on QFX10000 switches or NFX Series devices, configure the classifier as a multidestination classifier:

   ```
   [edit class-of-service]
   user@switch# set multi-destination classifiers (dscp | dscp-ipv6 | ieee-802.1 | inet-precedence)
   classifier-name
   ```


3. Apply the classifier to a specific Ethernet interface or to all Ethernet interfaces, or to all Fibre Channel interfaces on the device.

   - To apply the classifier to a specific interface:

     ```
     [edit class-of-service interfaces]
     user@switch# set interface-name unit unit classifiers (dscp | dscp-ipv6 | ieee-802.1) classifier-name
     ```

   - To apply the classifier to all Ethernet interfaces on the switch, use wildcards for the interface name and the logical interface (unit) number:

     ```
     [edit class-of-service interfaces]
     user@switch# set xe-* unit * classifiers (dscp | dscp-ipv6 | ieee-802.1) classifier-name
     ```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- *Example: Configuring Unicast Classifiers*
- *Configuring a Global MPLS EXP Classifier*
- *Configuring Rewrite Rules for MPLS EXP Classifiers*
- Monitoring CoS Classifiers on page 201
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- *Understanding CoS MPLS EXP Classifiers and Rewrite Rules*
- *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces*
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162

Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p)

Packet classification associates incoming packets with a particular CoS servicing level. Behavior aggregate (BA) classifiers examine the Differentiated Services code point (DSCP or DSCP IPv6) value, the IEEE 802.1p CoS value, or the MPLS EXP value in the packet header to determine the CoS settings applied to the packet. (See *Configuring a Global MPLS EXP Classifier* to learn how to define EXP classifiers for MPLS traffic.) BA classifiers allow you to set the forwarding class and loss priority of a packet based on the incoming CoS value.

NOTE: OCX Series switches do not support MPLS EXP classifiers.

On switches except QFX10000 and NFX Series devices, unicast traffic must use different classifiers than multidestination (mulitcast, broadcast, and destination lookup fail) traffic. You use the **multi-destination** statement at the **[edit class-of-service]** hierarhcy level to configure a multidestination BA classifier.

On QFX10000 switches and NFX Series devices, unicast and multidestination traffic use the same classifiers and forwarding classes.

Multidestination classifiers apply to all of the switch interfaces and handle multicast, broadcast, and destination lookup fail (DLF) traffic. You cannot apply a multidestination classifier to a single interface or to a range of interfaces.

To configure a DSCP, DSCP IPv6, or IEEE 802.1p BA classifier using the CLI:

1. Create a BA classifier:

   - To create a DSCP, DSCP IPv6, or IEEE 802.1p BA classifier based on the default classifier, import the default DSCP, DSCP IPv6, or IEEE 802.1p classifier and associate it with a forwarding class, a loss priority, and a code point:

     ```
     [edit class-of-service classifiers]
     user@switch# set (dscp | dscp-ipv6 | ieee-802.1) classifier-name import default
     forwarding-class forwarding-class-name loss-priority level code-points [aliases]
     [bit-patterns]
     ```

   - To create a BA classifier that is not based on the default classifier, create a DSCP, DSCP IPv6, or IEEE 802.1p classifier and associate it with a forwarding class, a loss priority, and a code point:

     ```
     [edit class-of-service classifiers]
     user@switch# set (dscp | dscp-ipv6 | ieee-802.1) classifier-name forwarding-class
     forwarding-class-name loss-priority level code-points [aliases] [bit-patterns]
     ```

2. For multidestination traffic, except on QFX10000 switches or NFX Series devices, configure the classifier as a multidestination classifier:

   ```
   [edit class-of-service]
   user@switch# set multi-destination classifiers (dscp | dscp-ipv6 | ieee-802.1 | inet-precedence)
   classifier-name
   ```

3. Apply the classifier to a specific Ethernet interface or to all Ethernet interfaces, or to all Fibre Channel interfaces on the device.

   - To apply the classifier to a specific interface:

     ```
     [edit class-of-service interfaces]
     user@switch# set interface-name unit unit classifiers (dscp | dscp-ipv6 | ieee-802.1)
     classifier-name
     ```

   - To apply the classifier to all Ethernet interfaces on the switch, use wildcards for the interface name and the logical interface (unit) number:

     ```
     [edit class-of-service interfaces]
     user@switch# set xe-* unit * classifiers (dscp | dscp-ipv6 | ieee-802.1) classifier-name
     ```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- *Example: Configuring Unicast Classifiers*
- *Configuring a Global MPLS EXP Classifier*
- *Configuring Rewrite Rules for MPLS EXP Classifiers*
- Monitoring CoS Classifiers on page 201
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- *Understanding CoS MPLS EXP Classifiers and Rewrite Rules*
- *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces*
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162

Understanding Host Inbound Traffic Classification

The destination address of traffic that enters the switch can be an external device such as another switch, a router, or a server, or the destination can be the host (the switch Routing Engine or CPU). When the destination is an external device, the DSCP and IEEE 802.1p code-point bits of incoming traffic are preserved as the traffic travels through the switch to the egress port. At the egress port, the code-point bits are either preserved when the packets are sent to the next hop or they are rewritten according to the rewrite rule attached to the egress interface.

When the destination of incoming traffic is the host, DSCP bits are preserved. However, IEEE 802.1p bits are not preserved. The IEEE 802.1p bits of traffic destined for the host are set to zero (0). This does not affect system behavior because the switch prioritizes traffic destined for the host based on the protocol type. For example, the switch gives a higher priority to BPDU traffic than to ping traffic.

Understanding CoS Code-Point Aliases

A code-point alias assigns a name to a pattern of code-point bits. You can use this name instead of the bit pattern when you configure other CoS components such as classifiers and rewrite rules.

NOTE: This topic applies to all EX Series switches except the EX4600. Because the EX4600 uses a different chipset than other EX Series switches, the code-point aliases on EX4600 match those on QFX Series switches. For EX4600 code-point aliases, see "Understanding CoS Code-Point Aliases" on page 134.

Behavior aggregate classifiers use class-of-service (CoS) values such as Differentiated Services Code Points (DSCPs) or IEEE 802.1 bits to associate incoming packets with a particular forwarding class and the CoS servicing level associated with that forwarding class. You can assign a meaningful name or alias to the CoS values and use that alias instead of bits when configuring CoS components. These aliases are not part of the specifications but are well known through usage. For example, the alias for DSCP 101110 is widely accepted as ef (expedited forwarding).

When you configure forwarding classes and define classifiers, you can refer to the markers by alias names. You can configure code point alias names for user-defined classifiers. If the value of an alias changes, it alters the behavior of any classifier that references it.

You can configure code-point aliases for the following type of CoS markers:

- dscp or dscp-ipv6—Handles incoming IP and IPv6 packets.
- ieee-802.1—Handles Layer 2 frames.

Table 33 on page 134 shows the default mapping of code-point aliases to IEEE code points.

Table 33: Default IEEE 802.1 Code-Point Aliases

CoS Value Types	Mapping
be	000
be1	001
ef	010
ef1	011
af11	100
af12	101

Table 33: Default IEEE 802.1 Code-Point Aliases (continued)

CoS Value Types	Mapping
nc1	110
nc2	111

Table 34 on page 135 shows the default mapping of code-point aliases to DSCP and DSCP IPv6 code points.

Table 34: Default DSCP and DSCP IPv6 Code-Point Aliases

CoS Value Types	Mapping
ef	101110
af11	001010
af12	001100
af13	001110
af21	010010
af22	010100
af23	010110
af31	011010
af32	011100
af33	011110
af41	100010
af42	100100
af43	100110
be	000000
cs1	001000
cs2	010000
cs3	011000
cs4	100000

Table 34: Default DSCP and DSCP IPv6 Code-Point Aliases (continued)

CoS Value Types	Mapping
cs5	101000
nc1	110000
nc2	111000

Related Documentation

- Understanding Junos CoS Components on page 15
- Defining CoS Code-Point Aliases on page 136

Defining CoS Code-Point Aliases

You can use code-point aliases to streamline the process of configuring CoS features on your switch. A code-point alias assigns a name to a pattern of code-point bits. You can use this name instead of the bit pattern when you configure other CoS components such as classifiers and rewrite rules.

You can configure code-point aliases for the following CoS marker types:

- DSCP or DSCP IPv6—Handles incoming IPv4 or IPv6 packets.
- IEEE 802.1p—Handles Layer 2 frames.

To configure a code-point alias:

1. Specify a CoS marker type (IEEE 802.1 or DSCP).
2. Assign an alias.
3. Specify the code point that corresponds to the alias.

```
[edit class-of-service code-point-aliases]
user@switch# set (dscp | dscp-ipv6 | ieee-802.1) alias-name code-point-bits
```

For example, to configure a code-point alias for an IEEE 802.1 CoS marker type that has the alias name be2 and maps to the code-point bits 001:

```
[edit class-of-service code-point-aliases]
user@switch# set ieee-802.1 be2 001
```

Related Documentation

- Monitoring CoS Code-Point Value Aliases on page 206
- Understanding CoS Code-Point Aliases on page 134

Understanding CoS Forwarding Classes

Forwarding classes group traffic and assign the traffic to output queues. Each forwarding class is mapped to an output queue. Classification maps incoming traffic to forwarding classes based on the code point bits in the packet or frame header. Forwarding class to queue mapping defines the output queue used for the traffic classified into a forwarding class.

A classifier must associate each packet with one of the following default forwarding classes or with a user-configured forwarding class in order to assign an output queue to the packet:

- fcoe—Do not use.
- no-loss—Do not use.
- best-effort—Provides best-effort delivery without a service profile. Loss priority is typically not carried in a class-of-service (CoS) value.
- network-control—Supports protocol control and is typically high priority.
- mcast—Delivery of multidestination (multicast, broadcast, and destination lookup fail) packets.

NOTE: Do not map traffic to the default fcoe and no-loss forwarding classes. By default, the DSCP default classifier does not map traffic to the fcoe and no-loss forwarding classes, so by default, the OCX Series does not classify traffic into those forwarding classes. (On other switches, the fcoe and no-loss forwarding classes provide lossless transport for Layer 2 traffic. OCX Series switches do not support lossless Layer 2 transport.)

The default fcoe and no-loss forwarding classes have the no-loss drop attribute. You can map traffic to the fcoe and no-loss forwarding classes only if you explicitly configure them and do not configure the no-loss drop attribute. The OCX Series does not support the no-loss drop attribute.

The switch supports up to 12 forwarding classes, 8 for unicast traffic and 4 for multidestination traffic, thus enabling flexible, differentiated, packet classification. For example, you can configure multiple classes of best-effort traffic such as **best-effort**, **best-effort1**, and **best-effort2**.

The switch supports up to 12 output queues: 8 output queues for unicast traffic (queues 0 through 7) and 4 output queues for multidestination traffic (queues 8 through 11). Forwarding classes mapped to unicast queues are associated with unicast traffic, and forwarding classes mapped to multidestination queues are associated with multidestination traffic. You cannot map unicast and multidestination traffic to the same queue. You cannot map a strict-high priority queue to a multidestination forwarding class (queues 8 through 11 do not support strict-high priority configuration).

- Default Forwarding Classes on page 138

• Forwarding Class Configuration Rules on page 139

Default Forwarding Classes

Table 35 on page 138 shows the four default forwarding classes defined for unicast traffic, and Table 36 on page 139 shows the default forwarding class defined for multicast traffic.

If desired, you can rename the forwarding classes. Assigning a new class name to an output queue does not alter the default classification or scheduling that is applicable to that queue. CoS configurations can be complicated, so unless it is required by your scenario, we recommend that you not alter the default class names or queue number associations.

Table 35: Default Forwarding Classes for Unicast Packets

Forwarding Class Name	Default Queue Mapping	Comments
best-effort	0	The software does not apply any special CoS handling to packets with 000000 in the DiffServ field. These packets are usually dropped under congested network conditions. By default, this is a lossy forwarding class with a packet drop attribute of **drop**.
fcoe	3	Do not map traffic to the default fcoe forwarding class. (On other switches, the fcoe forwarding class provides lossless transport for Layer 2 Fibre Channel over Ethernet (FCoE) traffic. OCX Series switches do not support FCoE or lossless Layer 2 transport.) The default fcoe forwarding class has the no-loss drop attribute. You can map traffic to the fcoe forwarding class only if you explicitly configure it and do not configure the **no-loss** drop attribute. The OCX Series does not support the **no-loss** drop attribute.
no-loss	4	Do not map traffic to the default no-loss forwarding class. OCX Series switches do not support lossless Layer 2 transport. The default no-loss forwarding class has the no-loss drop attribute. You can map traffic to the no-loss forwarding class only if you explicitly configure it and do not configure the **no-loss** drop attribute. The OCX Series does not support the **no-loss** drop attribute.
network-control	7	The software delivers packets in this service class with a high priority. (These packets are not delay-sensitive.) Typically, these packets represent routing protocol hello or keepalive messages. Because loss of these packets jeopardizes proper network operation, packet delay is preferable to packet discard. By default, this is a lossy forwarding class with a packet drop attribute of **drop**.

Table 36: Default Forwarding Classes for Multicast Packets

Forwarding Class Name	Default Queue Mapping	Comments
mcast	8	The software does not apply any special CoS handling to the multidestination packets. These packets are usually dropped under congested network conditions. By default, this is a lossy forwarding class with a packet drop attribute of **drop**.

NOTE: **Mirrored traffic is always sent to the queue that corresponds to the multidestination forwarding class. The switched copy of the mirrored traffic is forwarded with the priority determined by the behavior aggregate classification process.**

Forwarding Class Configuration Rules

Take the following rules into account when you configure forwarding classes:

- Queue Assignment Rules on page 139
- Scheduling Rules on page 140
- Rewrite Rules on page 140

Queue Assignment Rules

The following rules govern queue assignment:

- CoS configurations that specify more queues than the switch can support are not accepted. The commit operation fails with a detailed message that states the total number of queues available.
- All default CoS configurations are based on queue number. The name of the forwarding class that appears in the default configuration is the forwarding class currently mapped to that queue.
- Only unicast forwarding classes can be mapped to unicast queues (0 through 7), and only multidestination forwarding classes can be mapped to multidestination queues (8 through 11).
- Strict-high priority queues cannot be mapped to multidestination forwarding classes. (Strict-high priority traffic cannot be mapped to queues 8 through 11).
- If you map more than one forwarding class to a queue, all of the forwarding classes mapped to the same queue must have the same packet drop attribute (all of the forwarding classes must be lossy, or all of the forwarding classes mapped to a queue must be lossless). This is important because the default fcoe and no-loss forwarding classes have the no-loss drop attribute, which is not supported on OCX Series switches. Do not map traffic to the default fcoe and no-loss forwarding classes.

In addition, if you configure a strict-high priority queue, we recommend that you always apply a shaping rate to prevent the strict-high priority queue from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

Scheduling Rules

When you configure a forwarding class and map traffic to it on the switch (that is, you are not using a default classifier and forwarding class), you must also define a scheduling policy for the forwarding class. Defining a scheduling policy means:

- Mapping a scheduler to the forwarding class in a scheduler map
- Including the forwarding class in a forwarding class set
- Associating the scheduler map with a traffic control profile
- Attaching the traffic control profile to a forwarding class set and applying the traffic control profile to an interface

Rewrite Rules

On each physical interface, either all forwarding classes that are being used on the interface must have rewrite rules configured, or no forwarding classes that are being used on the interface can have rewrite rules configured. On any physical port, do not mix forwarding classes with rewrite rules and forwarding classes without rewrite rules.

Related Documentation

- Understanding Junos CoS Components on page 15
- Understanding CoS Packet Flow on page 23
- Understanding CoS Flow Control (Ethernet PAUSE and PFC) on page 525
- Example: Configuring Forwarding Classes on page 142
- Defining CoS Forwarding Classes on page 140

Defining CoS Forwarding Classes

Forwarding classes allow you to group packets for transmission. The switch supports a total of eight (QFX10000 and NFX Series devices), 10 (QFX5200 switches), or 12 (other switches) forwarding classes. To forward traffic, you map (assign) the forwarding classes to output queues.

The QFX10000 switches and NFX Series devices have eight output queues, queues 0 through 7. These queues support both unicast and multidestination traffic.

Except on QFX10000 and NFX Series devices, the switch has 10 output queues (QFX5200) or 12 output queues (other switches). Queues 0 through 7 are for unicast traffic and queues 8 through 11 are for multicast traffic. Forwarding classes mapped to unicast queues must carry unicast traffic, and forwarding classes mapped to multidestination queues must carry multidestination traffic. There are four default unicast forwarding classes and one default multidestination forwarding class.

The default forwarding classes, except on NFX Series devices, are:

NOTE: **Except on QFX10000, these are the default unicast forwarding classes.**

- **best-effort**—Best-effort traffic
- **fcoe**—Guaranteed delivery for Fibre Channel over Ethernet traffic (do not use on OCX Series switches)
- **no-loss**—Guaranteed delivery for TCP no-loss traffic (do not use on OCX Series switches)
- **network-control**—Network control traffic

The default multidestination forwarding class, except on QFX10000 switches and NFX Series devices, is:

- **mcast**—Multidestination traffic

The NFX Series devices have the following default forwarding classes:

- best-effort (be)—Provides no service profile. Loss priority is typically not carried in a CoS value.
- expedited-forwarding (ef)—Provides a low loss, low latency, low jitter, assured bandwidth, end-to-end service.
- assured-forwarding (af)—Provides a group of values you can define and includes four subclasses: AF1, AF2, AF3, and AF4, each with two drop probabilities: low and high.
- network-control (nc)—Supports protocol control and thus is typically high priority.

You can map forwarding classes to queues using the **class** statement. You can map more than one forwarding class to a single queue. Except on QFX10000 or NFX Series devices, all forwarding classes mapped to a particular queue must be of the same type, either unicast or multicast. You cannot mix unicast and multicast forwarding classes on the same queue.

All of the forwarding classes mapped to the same queue must have the same packet drop attribute: either all of the forwarding classes must be lossy or all of the forwarding classes must be lossless. This is important because the default fcoe and no-loss forwarding classes have the **no-loss** drop attribute, which is not supported on OCX Series switches. On OCX Series switches, do not map traffic to the default fcoe and no-loss forwarding classes.

```
[edit class-of-service forwarding-classes]
user@switch# set class class-name queue-num queue-number <no-loss>
```

One example is to create a forwarding class named **be2** and map it to queue 1:

```
[edit class-of-service forwarding-classes]
user@switch# set class be2 queue-num 1
```

Another example is to create a lossless forwarding class named **fcoe2** and map it to queue 5:

```
[edit class-of-service forwarding-classes]
user@switch# set class fcoe2 queue-num 5 no-loss
```

NOTE: On switches that do not run ELS software, if you are using Junos OS Release 12.2 or later, use the default forwarding-class-to-queue mapping for the lossless **fcoe** and **no-loss** forwarding classes. If you explicitly configure the lossless forwarding classes, the traffic mapped to those forwarding classes is treated as lossy (**best-effort**) traffic and does *not* receive lossless treatment unless you include the optional **no-loss** packet drop attribute introduced in Junos OS Release 12.3 in the forwarding class configuration..

NOTE: On switches that do not run ELS software, Junos OS Release 11.3R1 and earlier supported an alternate method of mapping forwarding classes to queues that allowed you to map only one forwarding class to a queue using the statement:

```
[edit class-of-service forwarding-classes]
user@switch# set queue queue-number class-name
```

The **queue** statement has been deprecated and is no longer valid in Junos OS Release 11.3R2 and later. If you have a configuration that uses the **queue** statement to map forwarding classes to queues, edit the configuration to replace the **queue** statement with the **class** statement.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Forwarding Classes on page 142
- Monitoring CoS Forwarding Classes on page 202

- *Understanding CoS Forwarding Classes*
- *Understanding CoS Port Schedulers on QFX Switches*

Example: Configuring Forwarding Classes

Forwarding classes group packets for transmission. Forwarding classes map to output queues, so the packets assigned to a forwarding class use the output queue mapped to that forwarding class. Except on QFX10000, unicast traffic and multidestination (multicast, broadcast, and destination lookup fail) traffic use separate forwarding classes and output queues.

Requirements

This example uses the following hardware and software components for two configuration examples:

Configuring forwarding classes for switches except QFX10000

- One switch except QFX10000 (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Configuring forwarding classes for QFX10000 switches

- One QFX10000 switch
- Junos OS Release 15.1X53-D10 or later for the QFX Series

Overview

The QFX10000 switch supports eight forwarding classes. Other switches support up to 12 forwarding classes. To forward traffic, you must map (assign) the forwarding classes to output queues. On the QFX10000 switch, queues 0 through 7 are for both unicast and multidestination traffic. On other switches, queues 0 through 7 are for unicast traffic, and queues 8 through 9 (QFX5200 switch) or 8 through 11 (other switches) are for multidestination traffic. Except for OCX Series switches, switches support up to six lossless forwarding classes. (OCX Series switches do not support lossless Layer 2 transport.)

The switch provides four default forwarding classes, and except on QFX10000 switches, these four forwarding classes are unicast, plus one default multidestination forwarding class. You can define the remaining forwarding classes and configure them as unicast or multidestination forwarding classes by mapping them to unicast or multidestination queues. The type of queue, unicast or multidestination, determines the type of forwarding class.

The four default forwarding classes (unicast except on QFX10000) are:

- **be**—Best-effort traffic
- **fcoe**—Guaranteed delivery for Fibre Channel over Ethernet traffic (do not use on OCX Series switches)

- **no-loss**—Guaranteed delivery for TCP no-loss traffic (do not use on OCX Series switches)
- **nc**—Network control traffic

Except on QFX10000 switches, the default multidestination forwarding class is:

- **mcast**—Multidestination traffic

Map forwarding classes to queues using the **class** statement. You can map more than one forwarding class to a single queue, but all forwarding classes mapped to a particular queue must be of the same type:

- Except on QFX10000 switches, all forwarding classes mapped to a particular queue must be either unicast or multicast. You cannot mix unicast and multicast forwarding classes on the same queue.
- On QFX10000 switches, all forwarding classes mapped to a particular queue must have the same packet drop attribute: all of the forwarding classes must be lossy, or all of the forwarding classes mapped to a queue must be lossless.

```
[edit class-of-service forwarding-classes]
user@switch# set class class-name queue-num queue-number;
```

NOTE: On switches that do not run ELS software, if you are using Junos OS Release 12.2, use the default forwarding-class-to-queue mapping for the lossless **fcoe** and **no-loss** forwarding classes. If you explicitly configure the lossless forwarding classes, the traffic mapped to those forwarding classes is treated as lossy (**best-effort**) traffic and does *not* receive lossless treatment.

In Junos OS Release 12.3 and later, you can include the *no-loss* packet drop attribute in explicit forwarding class configurations to configure a lossless forwarding class.

NOTE: On switches that do not run ELS software, Junos OS Release 11.3R1 and earlier supported an alternate method of mapping forwarding classes to queues that allowed you to map only one forwarding class to a queue using the statement:

```
[edit class-of-service forwarding-classes]
user@switch# set queue queue-number class-name
```

The **queue** statement has been deprecated and is no longer valid in Junos OS Release 11.3R2 and later. If you have a configuration that uses the **queue** statement to map forwarding classes to queues, edit the configuration to replace the **queue** statement with the **class** statement.

NOTE: Hierarchical scheduling controls output queue forwarding. When you define a forwarding class and classify traffic into it, you must also define a scheduling policy for the forwarding class. Defining a scheduling policy means:

- Mapping a scheduler to the forwarding class in a scheduler map
- Including the forwarding class in a forwarding class set
- Associating the scheduler map with a traffic control profile
- Attaching the traffic control profile to a forwarding class set and applying the traffic control profile to an interface

On QFX10000 switches, you can define a scheduling policy using port scheduling:

- Mapping a scheduler to the forwarding class in a scheduler map.
- Applying the scheduler map to one or more interfaces.

Example 1: Configuring Forwarding Classes for Switches Except QFX10000

Configuration

Step-by-Step Procedure

Table 37 on page 145 shows the configuration forwarding-class-to-queue mapping for this example:

Table 37: Forwarding-Class-to-Queue Example Configuration Except on QFX10000

Forwarding Class	Queue
best-effort	0
nc	7
mcast	8

To configure CoS forwarding classes for switches except QFX10000:

1. Map the **best-effort** forwarding class to queue **0**:

```
[edit class-of-service forwarding-classes]
user@switch# set class best-effort queue-num 0
```

2. Map the **nc** forwarding class to queue **7**:

```
[edit class-of-service forwarding-classes]
user@switch# set class nc queue-num 7
```

3. Map the **mcast-be** forwarding class to queue **8**:

```
[edit class-of-service forwarding-classes]
user@switch# set class mcast-be queue-num 8
```

Verification

Verifying the Forwarding-Class-to-Queue Mapping

Purpose Verify the forwarding-class-to-queue mapping. (The system shows only the explicitly configured forwarding classes; it does not show default forwarding classes such as **fcoe** and **no-loss**.)

Action Verify the results of the forwarding class configuration using the operational mode command **show configuration class-of-service forwarding-classes**:

```
user@switch> show configuration class-of-service forwarding-classes
class best-effort queue-num 0;
class network-control queue-num 7;
class mcast queue-num 8;
```

Example 2: Configuring Forwarding Classes for QFX10000 Switches

Configuration

Step-by-Step Procedure Table 38 on page 146 shows the configuration forwarding-class-to-queue mapping for this example:

Table 38: Forwarding-Class-to-Queue Example Configuration on QFX10000

Forwarding Class	Queue
best-effort	0
be1	1
nc	7

To configure CoS forwarding classes for QFX10000 switches:

1. Map the **best-effort** forwarding class to queue **0**:

```
[edit class-of-service forwarding-classes]
user@switch# set class best-effort queue-num 0
```

2. Map the **be1** forwarding class to queue **1**:

```
[edit class-of-service forwarding-classes]
user@switch# set class be1 queue-num 1
```


3. Map the **nc** forwarding class to queue **7**:

```
[edit class-of-service forwarding-classes]
user@switch# set class nc queue-num 7
```

Verification

Verifying the Forwarding-Class-to-Queue Mapping

Purpose Verify the forwarding-class-to-queue mapping. (The system shows only the explicitly configured forwarding classes; it does not show default forwarding classes such as **fcoe** and **no-loss**.)

Action Verify the results of the forwarding class configuration using the operational mode command **show configuration class-of-service forwarding-classes**:

```
user@switch> show configuration class-of-service forwarding-classes
class best-effort queue-num 0;
class be1 queue-num 1;
class network-control queue-num 7;
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Defining CoS Forwarding Classes on page 140
- Monitoring CoS Forwarding Classes on page 202
- *Overview of CoS Changes Introduced in Junos OS Release 11.3*
- *Overview of CoS Changes Introduced in Junos OS Release 12.2*
- *Understanding CoS Forwarding Classes*
- Understanding CoS Forwarding Classes on page 137

Understanding CoS Forwarding Class Sets (Priority Groups)

A forwarding class set is the Junos OS configuration construct that equates to a priority group in enhanced transmission selection (ETS, described in IEEE 802.1Qaz). The switch implements ETS using a two-tier hierarchical scheduler.

A priority group is a group of forwarding classes. Each forwarding class is mapped to an output queue and an IEEE 802.1p priority (code points). Classifying traffic into a forwarding class based on its code points, and mapping the forwarding class to a queue, defines the traffic assigned to that queue. The forwarding classes that belong to a priority group share the port bandwidth allocated to that priority group. The traffic mapped to forwarding classes in one priority group usually shares similar traffic-handling requirements.

You can configure up to three unicast forwarding class sets and one multicast forwarding class set. Only unicast forwarding classes can belong to unicast forwarding class sets. Only multicast forwarding classes can belong to the multicast forwarding class set.

If you configure a strict-high priority forwarding class (you can configure only one strict-high priority forwarding class), you must observe the following rules when configuring forwarding class sets:

- You must create a separate forwarding class set for the strict-high priority forwarding class.
- Only one forwarding class set can contain the strict-high priority forwarding class.
- A strict-high priority forwarding class cannot belong to the same forwarding class set as forwarding classes that are not strict-high priority.
- A strict-high priority forwarding class cannot belong to a multidestination forwarding class set.
- You cannot configure a guaranteed minimum bandwidth (guaranteed rate) for a forwarding class set that includes a strict-high priority forwarding class. (You also cannot configure a guaranteed minimum bandwidth for a strict-high forwarding class.)
- We recommend that you always apply a shaping rate to a strict-high priority forwarding class to prevent it from starving the queues mapped to other forwarding classes. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority forwarding class can use, then the strict-high priority forwarding class can use all of the available port bandwidth and starve other forwarding classes on the port.

You must use hierarchical scheduling if you explicitly configure CoS. The two-tier hierarchical scheduler defines bandwidth resources for the forwarding class set (priority group), and then allocates those resources among the forwarding classes (priorities) that belong to the forwarding class set.

If you do not explicitly configure forwarding class sets, the system automatically creates a default forwarding class set that contains all of the forwarding classes on the switch. The system assigns 100 percent of the port output bandwidth to the default forwarding class set. Ingress traffic is classified based on the default classifier settings. The forwarding classes in the default forwarding class set receive bandwidth based on the default scheduler settings. Forwarding classes that are not part of the default scheduler receive no bandwidth. The default priority group is transparent. It does not appear in the configuration and is used for Data Center Bridging Capability Exchange Protocol (DCBX) advertisement (except on OCX Series switches, which do not support DCBX).

When you explicitly configure forwarding class sets and apply them to interfaces, on those interfaces, forwarding classes that you do not map to a forwarding class set receive no guaranteed bandwidth. Forwarding classes that belong to the default forwarding class set might receive bandwidth if the other forwarding class sets are not using all of the port bandwidth. However, the amount of bandwidth received by forwarding classes that are not members of a forwarding class set is not guaranteed. In this case, the bandwidth a forwarding class receives if it is not a member of a forwarding class set depends on whether unused port bandwidth is available and therefore is not deterministic.

To guarantee bandwidth for forwarding classes in a predictable manner, be sure to map all forwarding classes that you expect to carry traffic on an interface to a forwarding class set, and apply the forwarding class set to the interface.

Related Documentation

- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Forwarding Class Sets on page 150
- Defining CoS Forwarding Class Sets on page 149

Defining CoS Forwarding Class Sets

A forwarding class set is a priority group for enhanced transmission selection (ETS) traffic control. Each forwarding class set consists of one or more forwarding classes. Classifiers map traffic into forwarding classes based on code points (priority), and forwarding classes are mapped to output queues.

You can configure up to three unicast forwarding class sets and one multicast forwarding class set.

To configure a forwarding class set using the CLI:

1. Assign one or more forwarding classes to the forwarding class set:

```
[edit class-of-service]
user@switch# set forwarding-class-sets forwarding-class-set-name class
forwarding-class-name
```

2. Map the forwarding class set to an interface:

```
[edit class-of-service]
user@switch# set interfaces interface-name forwarding-class-set forwarding-class-set-name
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Forwarding Class Sets on page 150
- Defining CoS Queue Schedulers on page 290
- Defining CoS Traffic Control Profiles (Priority Group Scheduling) on page 310
- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148

Example: Configuring Forwarding Class Sets

A forwarding class set (fc-set) is a priority group for enhanced transmission selection (ETS) traffic control. Each fc-set consists of one or more forwarding classes (priorities). Classifiers map traffic to forwarding classes based on code points, and forwarding classes are mapped to output queues.

ETS enables you to configure link resources (bandwidth and bandwidth sharing characteristics) for an fc-set, and then allocate the fc-set's resources among the forwarding classes that belong to the fc-set. This is called two-tier, or hierarchical, scheduling. Traffic control profiles control the scheduling for the fc-set (priority group), and schedulers control the scheduling for individual forwarding classes (priorities).

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series.

Overview

You can configure up to three unicast fc-sets and one multicast fc-set. A common way to configure unicast priority groups is to configure separate fc-sets for local area network (LAN) traffic, storage area network (SAN) traffic, and high-performance computing (HPC) traffic, and then assign the appropriate forwarding classes to each fc-set.

NOTE: If you configure a **strict-high** priority forwarding class, you must create an fc-set that is dedicated only to **strict-high** priority traffic. You can only configure one strict-high priority forwarding class, and only one fc-set can contain a strict-high priority queue. Queues that are not strict-high priority cannot belong to the same fc-set as a strict-high priority queue. The multidestination fc-set cannot contain a strict-high priority queue.

To apply ETS, you use a traffic control profile to map one or more fc-sets to a physical egress port. You can map up to three unicast forwarding class sets and one multidestination forwarding class set to each port. When you map an fc-set to a port, the port uses hierarchical scheduling to allocate port resources to the priority group (fc-set) and to allocate the priority group resources to the queues (forwarding classes) that belong to the priority group.

This example describes how to:

- Configure three fc-sets called **lan-pg**, **san-pg**, and **hpc-pg**.
- Assign forwarding classes to each of the fc-sets.
- Apply the fc-sets and their output traffic control profiles to an egress interface.

This example does not describe how to configure the forwarding classes assigned to the fc-sets or how to configure traffic control profiles (scheduling). “Example: Configuring CoS Hierarchical Port Scheduling (ETS)” on page 321 provides a complete example of how to configure ETS, including forwarding class and scheduling configuration. Table 39 on page 151 shows the configuration components for this example:

Table 39: Components of the Forwarding Class Sets Configuration Example

Component	Settings
Hardware	QFX3500 switch
LAN traffic priority group	Forwarding class set: **lan-pg** Forwarding classes: **best-effort-1**, **best-effort-2**
SAN traffic priority group	Forwarding class set: **san-pg** Forwarding classes: **fcoe**, **fcoe-2** NOTE: OCX Series switches do not support FCoE traffic or lossless Layer 2 transport. If you were configuring this example on an OCX Series switch, you could omit this priority group, or rename it and map different forwarding classes to it.
HPC traffic priority group	Forwarding class set: **hpc-pg** Forwarding classes: **nc**, **high-perf**
Egress interface	**xe-0/0/7**

Configuring Forwarding Class Sets

1. Define the **lan-pg** priority group (fc-set) and assign to it the forwarding classes **best-effort-1** and **best-effort-2**:

```
[edit class-of-service]
user@switch# set forwarding-class-sets lan-pg class best-effort-1
user@switch# set forwarding-class-sets lan-pg class best-effort-2
```

2. Define the **san-pg** priority group and assign to it the forwarding classes **fcoe** and **fcoe-2**:

```
[edit class-of-service]
user@switch# set forwarding-class-sets san-pg class fcoe
user@switch# set forwarding-class-sets san-pg class fcoe-2
```

3. Define the **hpc-pg** priority group and assign to it the forwarding classes **nc** and **high-perf**:

```
[edit class-of-service]
user@switch# set forwarding-class-sets hpc-pg class nc
user@switch# set forwarding-class-sets hpc-pg class high-perf
```

4. Map the three forwarding class sets to an interface (the output traffic control profiles associated with the forwarding class sets determine the class of service scheduling for the priority groups):

```
[edit class-of-service]
user@switch# set interfaces xe-0/0/7 forwarding-class-set lan-pg
output-traffic-control-profile lan-tcp
user@switch# set interfaces xe-0/0/7 forwarding-class-set san-pg
output-traffic-control-profile san-tcp
user@switch# set interfaces xe-0/0/7 forwarding-class-set hpc-pg
output-traffic-control-profile hpc-tcp
```

Verification

To verify the priority group configuration, perform these tasks:

- Verifying Forwarding Class Set Membership on page 152
- Verifying the Egress Interface Configuration on page 153

Verifying Forwarding Class Set Membership

Purpose Verify that you configured the **lan-pg**, **san-pg**, and **hpc-pg** priority groups with the correct forwarding classes.

Action List the forwarding class set member configuration using the operational mode command **show configuration class-of-service forwarding-class-sets**:

```
user@switch> show configuration class-of-service forwarding-class-sets
lan-pg {
    class best-effort-1;
    class best-effort-2;
}
san-pg {
    class fcoe;
    class fcoe-2;
}
hpc-pg {
    class high-perf;
    class nc;
}
```


Verifying the Egress Interface Configuration

Purpose Verify that egress interface **xe-0/0/7** is associated with the **lan-pg**, **san-pg**, and **hpc-pg** priority groups and with the correct output traffic control profiles.

Action Display the egress interface using the operational mode command **show configuration class-of-service interfaces xe-0/0/7**:

```
user@switch> show configuration class-of-service interfaces xe-0/0/7
forwarding-class-set {
    lan-pg {
        output-traffic-control-profile lan-tcp;
    }
    san-pg {
        output-traffic-control-profile san-tcp;
    }
    hpc-pg {
        output-traffic-control-profile hpc-tcp;
    }
}
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Defining CoS Forwarding Class Sets on page 149
- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148

Understanding Host Routing Engine Outbound Traffic Queues and Defaults

The host Routing Engine and CPU generate outbound traffic that is transmitted using different protocols. You cannot configure a classifier to map different types of outbound traffic that the host generates to forwarding classes (queues). The traffic that the host generates is assigned to forwarding classes by default as shown in Table 40 on page 154.

If you want to separate host outbound traffic from other traffic or if you want to assign that traffic to a particular queue, you can configure a single forwarding class for all traffic that the host generates. If you configure a forwarding class for outbound host traffic, that forwarding class is used globally for all traffic generated by the host. (That is, the host outbound traffic is mapped to the selected queue on all egress interfaces.) Configuring a forwarding class for host outbound traffic does not affect transit or incoming traffic.

Whether you use the default host outbound traffic forwarding class configuration or configure a forwarding class for all host outbound traffic, the configuration applies to all Layer 2 and Layer 3 protocols and to all application-level traffic such as FTP and ping operations.

If you configure a queue for host outbound traffic, the queue must be properly configured on all interfaces.

NOTE: Fibre Channel over Ethernet (FCoE) Initialization Protocol (FIP) packets generated by the CPU are always transmitted on the **fcoe** queue (queue 3), even if you configure a queue for host outbound traffic. This helps to ensure lossless behavior for FCoE traffic. QFabric systems classify FIP control packets into the same traffic class (**fcoe**) across the Interconnect device (fabric) and the egress Node device.

This does not apply to OCX Series switches, which do not support FCoE.

By default, traffic generated by the host is sent to the best effort queue (queue 0) or to the network control queue (queue 7). Table 40 on page 154 lists the default host traffic to output queue mapping.

Table 40: Routing Engine Protocol Default Queue Mapping

Routing Engine Protocol	Default Queue Mapping
Address Resolution Protocol (ARP) reply	Queue 0
ARP request	Queue 0
Border Gateway Protocol (BGP)	Queue 0
BGP TCP Retransmission	Queue 7

Table 40: Routing Engine Protocol Default Queue Mapping (continued)

Routing Engine Protocol	Default Queue Mapping
Fibre Channel over Ethernet (FCoE) Initialization Protocol (FIP)	Queue 3
File Transfer Protocol (FTP)	Queue 0
Internet Control Message Protocol (ICMP) reply	Queue 0
ICMP request	Queue 0
Internet Group Management Protocol (IGMP) query	Queue 7
IGMP report	Queue 0
Link Aggregation Control Protocol (LACP)	Queue 7
Open Shortest Path First (OSPF) hello	Queue 7
OSPF protocol data unit (PDU)	Queue 7
OSPF link state advertisements (LSAs)	Queue 7
Protocol Independent Multicast (PIM)	Queue 7
PIM hello	Queue 7
Simple Network Management Protocol (SNMP)	Queue 0
Secure Shell (SSH)	Queue 0
Telnet	Queue 0
Virtual Router Redundancy Protocol (VRRP)	Queue 7
VLAN Spanning Tree Protocol (VSTP)	Queue 7
xnm-clear-text	Queue 0
xnm-ssl	Queue 0

Related Documentation

- *Understanding CoS Forwarding Classes*
- Understanding CoS Forwarding Classes on page 137
- Changing the Host Outbound Traffic Default Queue Mapping on page 156
- Example: Configuring Forwarding Classes on page 142

Changing the Host Outbound Traffic Default Queue Mapping

If you do not want to use the default mapping of host Routing Engine and CPU outbound traffic to queues, you can change the default output queue. You can also change the default DSCP bits used in the type of service (ToS) field of packets generated by the Routing Engine.

Configuring a queue for host outbound traffic maps all traffic that the host generates to one forwarding class (queue). The configuration is global and applies to all host-generated traffic on the switch. Configuring a forwarding class for host outbound traffic does not affect transit or incoming traffic.

NOTE: Fibre Channel over Ethernet (FCoE) Initialization Protocol (FIP) packets generated by the CPU are always transmitted on the **fcoe** queue (queue 3), even if you configure a queue for host outbound traffic. This helps to ensure lossless behavior for FCoE traffic. QFabric systems classify FIP control packets into the same traffic class (**fcoe**) across the Interconnect device (fabric) and the egress Node device.

This does not apply to OCX Series switches, which do not support FCoE.

To change the host outbound traffic egress queue by including the **host-outbound-traffic** statement at the **[edit class-of-service]** hierarchy level:

```
[edit class-of-service]
host-outbound-traffic {
  forwarding-class class-name;
  dscp-code-point code-point;
}
```

For example, to map host outbound traffic to queue 7 (the network control forwarding class) and set the DSCP code point value to 101010:

```
[edit class-of-service]
host-outbound-traffic {
  forwarding-class network-control;
  dscp-code-point 101010
}
```

Related Documentation

- Understanding Host Routing Engine Outbound Traffic Queues and Defaults on page 154

Understanding CoS Rewrite Rules

As packets enter or exit a network, edge switches might be required to alter the class-of-service (CoS) settings of the packets. Rewrite rules set the value of the code point bits (Layer 3 DSCP bits, Layer 2 CoS bits, or MPLS EXP bits) within the header of the outgoing packet. Each rewrite rule:

1. Reads the current forwarding class and loss priority associated with the packet.
2. Locates the new (rewrite) code point value from a table.
3. Writes that code point value into the packet header, replacing the old code point value.

Rewrite rules must be assigned to an interface for rewrites to take effect.

You can apply (bind) one DSCP or DSCP IPv6 rewrite rule and one IEEE 802.1p rewrite rule to each interface. You can also bind EXP rewrite rules to **family mpls** logical interfaces to rewrite the CoS bits of MPLS traffic.

NOTE: **OCX Series switches do not support MPLS and do not support EXP rewrite rules.**

You cannot apply both a DSCP and a DSCP IPv6 rewrite rule to the same physical interface. Each physical interface supports only one DSCP rewrite rule. Both IP and IPv6 packets use the same DSCP rewrite rule, regardless if the configured rewrite rule is DSCP or DSCP IPv6. You can apply an EXP rewrite rule on an interface that has DSCP or IEEE rewrite rules. Only MPLS traffic on **family mpls** interfaces uses the EXP rewrite rule.

You *can* apply both a DSCP rewrite rule and a DSCP IPv6 rewrite rule to a logical interface. IPv6 packets are rewritten with DSCP-IPv6 rewrite-rules and IPv4 packets are remarked with DSCP rewrite-rules.

NOTE: **There are no default rewrite rules. If you want to apply a rewrite rule to outgoing packets, you must explicitly configure the rewrite rule.**

You can look at behavior aggregate (BA) classifiers and rewrite rules as two sides of the same coin. A BA classifier reads the code point bits of incoming packets and classifies the packets into forwarding classes, then the system applies the CoS configured for the forwarding class to those packets. Rewrite rules change (rewrite) the code point bits just before the packets leave the system so that the next switch or router can apply the appropriate level of CoS to the packets. When you apply a rewrite rule to an interface, the rewrite rule is the last CoS action performed on the packet before it is forwarded.

Rewrite rules alter CoS values in outgoing packets on the outbound interfaces of an edge switch to accommodate the policies of a targeted peer. This allows the downstream switch in a neighboring network to classify each packet into the appropriate service group.

NOTE: On each physical interface, either all forwarding classes that are being used on the interface must have rewrite rules configured or no forwarding classes that are being used on the interface can have rewrite rules configured. On any physical port, do not mix forwarding classes with rewrite rules and forwarding classes without rewrite rules.

NOTE: Rewrite rules are applied *before* the egress filter is matched to traffic. Because the code point rewrite occurs before the egress filter is matched to traffic, the egress filter match is based on the rewrite value, not on the original code point value in the packet.

For packets that carry both an inner VLAN tag and an outer VLAN tag, the rewrite rule rewrites only the outer VLAN tag.

MPLS EXP rewrite rules apply only to **family mpls** logical interfaces. You cannot apply to an EXP rewrite rule to a physical interface. You can configure up to 64 EXP rewrite rules, but you can only use 16 EXP rewrite rules at any time on the switch. On a given logical interface, all pushed MPLS labels have the same EXP rewrite rule applied to them. You can apply different EXP rewrite rules to different logical interfaces on the same physical interface.

NOTE: If the switch is performing penultimate hop popping (PHP), EXP rewrite rules do not take effect. If both an EXP classifier and an EXP rewrite rule are configured on the switch, then the EXP value from the last popped label is copied into the inner label. If either an EXP classifier or an EXP rewrite rule (but not both) is configured on the switch, then the inner label EXP value is sent unchanged.

You can configure enough rewrite rules to handle most, if not all, network scenarios. Table 41 on page 158 shows how many of each type of rewrite rules you can configure, and how many entries you can configure per rewrite rule.

Table 41: Configuring Rewrite Rules

Rewrite Rule Type	Maximum Number of Rewrite Rules	Maximum Number of Entries per Rewrite Rule
IEEE 802.1p	64	128
DSCP	32	128
DSCP IPv6	32	128
MPLS EXP	64	128

You cannot apply rewrite rules directly to integrated routing and bridging (IRB), also known as routed VLAN interfaces (RVIs), because the members of IRBs/RVIs are VLANs, not ports. However, you can apply rewrite rules to the VLAN port members of an IRB/RVI.

NOTE: OCX Series witches do not support IRBs/RVIs.

Related Documentation

- Understanding Junos CoS Components on page 15
- Defining CoS Rewrite Rules on page 160
- *Configuring Rewrite Rules for MPLS EXP Classifiers*

Defining CoS Rewrite Rules

Edge switches might need to change the class-of-service (CoS) settings of the packets. You can configure rewrite rules to alter code point bit values in outgoing packets on the outbound interfaces of a switch so that the CoS treatment matches the policies of a targeted peer. Policy matching allows the downstream routing platform or switch in a neighboring network to classify each packet into the appropriate service group.

To configure a CoS rewrite rule, create the rule by giving it a name and associating it with a forwarding class, loss priority, and code point. This creates a rewrite table. After the rewrite rule is created, enable it on an interface (EXP rewrite rules can only be enabled on **family mpls** logical interfaces, not on physical interfaces). You can also apply an existing rewrite rule on an interface.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP rewrite rules.

NOTE: On each physical interface, either all forwarding classes that are being used on the interface must have rewrite rules configured, or no forwarding classes that are being used on the interface can have rewrite rules configured. On any physical port, do not mix forwarding classes with rewrite rules and forwarding classes without rewrite rules.

NOTE: To replace an existing rewrite rule on the interface with a new rewrite rule of the same type, first explicitly remove the existing rewrite rule and then apply the new rule.

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

To create rewrite rules and enable them on interfaces:

- To create an 802.1p rewrite rule named **customup-rw** in the rewrite table for all Layer 2 interfaces:

```
[edit class-of-service rewrite-rules]
user@switch# set ieee-802.1 customup-rw forwarding-class be loss-priority low code-point 000
user@switch# set ieee-802.1 customup-rw forwarding-class be loss-priority high code-point 001
user@switch# set ieee-802.1 customup-rw forwarding-class be loss-priority low code-point 010
user@switch# set ieee-802.1 customup-rw forwarding-class fcoe loss-priority low code-point 011
```


`user@switch#` **set ieee-802.1 customup-rw forwarding-class ef-no-loss loss-priority low code-point 100**
`user@switch#` **set ieee-802.1 customup-rw forwarding-class ef-no-loss loss-priority high code-point 101**
`user@switch#` **set ieee-802.1 customup-rw forwarding-class nc loss-priority low code-point 110**
`user@switch#` **set ieee-802.1 customup-rw forwarding-class nc loss-priority high code-point 111**

- To enable an 802.1p rewrite rule named **customup-rw** on a Layer 2 interface:

  `[edit]`
  `user@switch#` **set class-of-service interfaces xe-0/0/7 unit 0 rewrite-rules ieee-802.1 customup-rw**

  NOTE: All forwarding classes assigned to port xe-0/0/7 must have rewrite rules. Do not mix forwarding classes that have rewrite rules with forwarding classes that do not have rewrite rules on the same physical interface.

- To enable an 802.1p rewrite rule named **customup-rw** on all 10-Gigabit Ethernet interfaces on the switch, use wildcards for the interface name and logical interface (unit) number:

  `[edit]`
  `user@switch#` **set class-of-service interfaces xe-* unit * rewrite-rules customup-rw**

  NOTE: In this case, *all* forwarding classes assigned to *all* 10-Gigabit Ethernet ports must have rewrite rules. Do not mix forwarding classes that have rewrite rules with forwarding classes that do not have rewrite rules on the same physical interface.

Related Documentation

- Monitoring CoS Rewrite Rules on page 205
- *Configuring Rewrite Rules for MPLS EXP Classifiers*
- Understanding CoS Rewrite Rules on page 157
- *Understanding CoS MPLS EXP Classifiers and Rewrite Rules*

Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces

At ingress interfaces, classifiers group incoming traffic into classes based on the DSCP or IEEE 802.1p class of service (CoS) code point bits in the packet header. At egress interfaces, you can use rewrite rules to change (re-mark) the code point bits before the interface forwards the packets. At ingress interfaces, classifiers group incoming traffic into classes based on the DSCP or IEEE 802.1p code point bits in the packet header. At egress interfaces, rewrite rules can change (re-mark) the code point bits before the interface forwards the packets.

You can apply classifiers and rewrite rules to interfaces to control the level of CoS applied to each packet as it traverses the system and the network. This topic describes:

Supported Classifier and Rewrite Rule Types

Table 42 on page 162 shows the supported types of classifiers and rewrite rules:

Table 42: Supported Classifiers and Rewrite Rules

Classifier or Rewrite Rule Type	Description
Fixed classifier	Classifies all ingress traffic on a physical interface into one fixed forwarding class, regardless of the CoS bits in the packet header.
DSCP and DSCP IPv6 unicast classifiers	Classifies IP and IPv6 traffic into forwarding classes and assigns loss priorities to the traffic based on DSCP code point bits.
IEEE 802.1p unicast classifier	Classifies Ethernet traffic into forwarding classes and assigns loss priorities to the traffic based on IEEE 802.1p code point bits.
DSCP multidestination classifier (also used for IPv6 multidestination traffic)	Classifies IP and IPv6 multicast, broadcast, and destination lookup fail (DLF) traffic into multidestination forwarding classes. Multidestination classifiers are applied to all interfaces and cannot be applied to individual interfaces.
IEEE 802.1p multidestination classifier	Classifies Ethernet multicast, broadcast, and destination lookup fail (DLF) traffic into multidestination forwarding classes. Multidestination classifiers are applied to all interfaces and cannot be applied to individual interfaces.

Table 42: Supported Classifiers and Rewrite Rules (continued)

Classifier or Rewrite Rule Type	Description
DSCP and DSCP IPv6 rewrite rules	Re-marks the DSCP code points of IP and IPv6 packets before forwarding the packets.
IEEE 802.1p rewrite rule	Re-marks the IEEE 802.1p code points of Ethernet packets before forwarding the packets.

DSCP and IEEE 802.1p classifiers are behavior aggregate (BA) classifiers. Multidestination classifiers are global and apply to all interfaces; you cannot apply a multidestination classifier to individual interfaces.

Classifying packets into forwarding classes assigns packets to the output queues associated with the forwarding classes. Classifying traffic into a forwarding class associates the CoS scheduling for the forwarding class with that traffic.

NOTE: **In addition to BA classifiers and fixed classifiers, which classify traffic based on the CoS field in the packet header, you can use firewall filters to configure multifield (MF) classifiers. MF classifiers classify traffic based on more than one field in the packet header and take precedence over BA and fixed classifiers.**

Interfaces Supported for Classifier and Rewrite Rule Configuration

To apply a classifier to incoming traffic or a rewrite rule to outgoing traffic, you need to apply the classifier or rewrite rule to one or more interfaces. When you apply a classifier or rewrite rule to an interface, the interface uses the classifier to group incoming traffic into forwarding classes and uses the rewrite rule to re-mark the CoS code point value of each packet before it leaves the system.

Not all interface types support all types of CoS configuration. This section describes:

Classifier and Rewrite Rule Physical and Logical Interface Support

The ports can function as:

- Layer 3 physical interfaces (family inet/inet6)
- Layer 3 logical interfaces (family inet/inet6)

You can apply CoS classifiers and rewrite rules to Layer 3 physical interfaces if at least one logical Layer 3 interface is configured on the physical interface.

NOTE: The CoS you configure on a Layer 3 physical interface is applied to all of the Layer 3 logical interfaces on that physical interface. This means that each Layer 3 interface uses the same classifiers and rewrite rules for all of the Layer 3 traffic on that interface.

Table 43 on page 164 shows on which interfaces you can configure and apply classifiers and rewrite rules. For Layer 3 LAGs, configure BA or fixed classifiers on the LAG (ae) interface. The classifier configured on the LAG is valid on all of the LAG member interfaces.

Table 43: Interface Support for Classifier and Rewrite Rule Configuration

CoS Classifiers and Rewrite Rules	Layer 3 Physical Interfaces (If at Least One Logical Layer 3 Interface Is Defined)	Layer 3 Logical Interfaces
Fixed classifier	Yes	Yes
DSCP classifier	Yes	Yes
DSCP IPv6 classifier	Yes	Yes
IEEE 802.1p classifier	Yes	Yes
DSCP rewrite rule	Yes	Yes
DSCP IPv6 rewrite rule	Yes	Yes
IEEE 802.1p rewrite rule	Yes	Yes

NOTE: DSCP mutidestination and IEEE 802.1p multidestination classifiers are applied to all interfaces and cannot be applied to individual interfaces. No DSCP IPv6 multidestination classifier is supported. IPv6 multidestination traffic uses the DSCP multidestination classifier.

Default Classifiers

If you do not explicitly configure classifiers on an interface, the switch applies default classifiers (see “Understanding Default CoS Settings” on page 25) so that the traffic receives basic CoS treatment. The system applies a default classifier using the following rules:

- If the physical interface has at least one Layer 3 logical interface configured, it uses the default DSCP classifier.

- If the physical interface has no logical interface configured, no default classifier is applied.
- The default multidestination classifier is the IEEE 802.1p multidestination classifier.

Default Rewrite Rules

No default rewrite rules are applied to interfaces. If you want to re-mark packets at the egress interface, you must explicitly configure a rewrite rule.

Classifier Precedence

You can apply multiple unicast classifiers (MF, fixed, DSCP, or IEEE 802.1p) to a physical interface to handle different types of traffic. When you apply more than one classifier to an interface, the system uses an order of precedence to determine which classifier to use on interfaces:

The precedence of unicast classifiers on physical interfaces, from the highest-priority classifier to the lowest-priority classifier, is:

- MF classifier (no classifier has a higher priority than MF classifiers)
- Fixed classifier
- DSCP or DSCP IPv6 classifier
- IEEE 802.1p classifier

You can apply a DSCP classifier and an IEEE 802.1p classifier to a physical interface. When both classifiers are on an interface, IP traffic uses the DSCP classifier, and all other traffic uses the IEEE classifier.

NOTE: **You cannot apply a fixed classifier and a DSCP or IEEE classifier to the same interface. If a DSCP classifier, an IEEE classifier, or both are on an interface, you cannot apply a fixed classifier to that interface unless you first delete the DSCP and IEEE classifiers. If a fixed classifier is on an interface, you cannot apply a DSCP classifier or an IEEE classifier unless you first delete the fixed classifier.**

Classifier Behavior and Limitations

Consider the following behaviors and constraints when you apply classifiers to interfaces:

- You can configure only one DSCP classifier (IP or IPv6) on a physical interface. You cannot configure both types of DSCP classifier on one physical interface. Both IP and IPv6 traffic use whichever DSCP classifier is configured on the interface.
- When you configure a DSCP or a DSCP IPv6 classifier on a physical interface and the physical interface has at least one logical Layer 3 interface, all packets (IP, IPv6, and non-IP) use that classifier.

- An interface with both a DSCP classifier (IP or IPv6) and an IEEE 802.1p classifier uses the DSCP classifier for IP and IPv6 packets, and uses the IEEE classifier for all other packets.
- Fixed classifiers and BA classifiers (DSCP and IEEE classifiers) are not permitted simultaneously on an interface. If you configure a fixed classifier on an interface, you cannot configure a DSCP or an IEEE classifier on that interface. If you configure a DSCP classifier, an IEEE classifier, or both classifiers on an interface, you cannot configure a fixed classifier on that interface.
- When you configure an IEEE 802.1p classifier on a physical interface and a DSCP classifier is not explicitly configured on that interface, the interface uses the IEEE classifier for all types of packets. No default DSCP classifier is applied to the interface. (In this case, if you want a DSCP classifier on the interface, you must explicitly configure it and apply it to the interface.)
- The system does not apply a default classifier to a physical interface until you create a logical interface on that physical interface. If you configure a Layer 3 logical interface, the system uses the default DSCP classifier.
- MF classifiers take precedence over BA and fixed classifiers. (Use firewall filters to configure MF classifiers.) When BA or fixed classifiers are present on an interface, you can still configure an MF classifier on that interface.
- If no explicit classifier (IEEE-802.1 or DSCP or DSCP-IPv6) is configured on an L3 logical interface, a DSCP classifier default (*dscp-default*) and a DSCP-IPv6 classifier default (*dscp-ipv6-default*) are configured. All IPv4 traffic uses *dscp-default*, while IPv6 traffic uses *dscp-ipv6-default* classifiers for classification. Remaining traffic is classified to be fc=0 and dp=0.
- When only IEEE802.1p is configured on an L3 logical interface, no *dscp-default* or *dscp-ipv6-default* classifiers are configured. All traffic uses IEEE 802.1p classifier for classification.
- When only one of either a DSCP or DSCP-IPv6 classifier is configured, the default classifier for the other type (either *dscp-default* or *dscp-ipv6-default*) is automatically configured.
- On an L3 logical interface, you cannot configure both a fixed classifier and a BA classifier. If a fixed classifier is configured first, then no BA classifier is allowed on that interface and vice versa.
- L3 logical interface MF classification is configured with firewall filters and has highest precedence over both BA and fixed classifiers. MF classifiers can be configured even when a BA classifier or fixed classifier is already present on the L3 logical interface.
- An interface can be simultaneously configured with an IEEE 802.1p classifier, a DSCP classifier, and a DSCP-IPv6 classifier. IP, IPv6, and any remaining traffic are all classified using DSCP, DSCP-IPv6 and 802.1p classifiers respectively.
- When a classifier is bound to a LAG logical interface, the classifier is internally applied to all member interfaces on both L2 and L3 interfaces.
- Multicast and unicast traffic are classified using the same classifier,
- RVI interfaces do not support classifiers.

Rewrite Rule Precedence and Behavior

The following rules apply on interfaces for rewrite rules:

- If you configure one DSCP (or DSCP IPv6) rewrite rule and one IEEE 802.1p rewrite rule on an interface, both rewrite rules take effect. Traffic with IP and IPv6 headers use the DSCP rewrite rule, and traffic with a VLAN tag uses the IEEE rewrite rule.
- If you do not explicitly configure a rewrite rule, there is no default rewrite rule, so the system does not apply any rewrite rule to either the physical or logical interface.
- You can apply either a DSCP rewrite rule or a DSCP IPv6 rewrite rule to any physical interface, but you *cannot* apply both a DSCP and a DSCP IPv6 rewrite rule to the same physical interface. Both IP and IPv6 packets use the same DSCP rewrite rule, regardless of whether the configured rewrite rule is DSCP or DSCP IPv6.
- You *can* apply both a DSCP rewrite rule and a DSCP IPv6 rewrite rule to a logical interface. IPv6 packets are rewritten with DSCP-IPv6 rewrite-rules and IPv4 packets are remarked with DSCP rewrite-rules.
- Rewrite-rules on RVI are not supported.
- When rewrite-rules are bound to LAG logical interfaces, the rewrite-rules are internally applied on all the member interfaces. The same rules apply to LAGs in L2 logical interfaces and L3 logical interfaces.

Related Documentation

- Understanding CoS Packet Flow on page 23
- Understanding Default CoS Settings on page 25
- Understanding Default CoS Scheduling and Classification on page 114
- Understanding CoS Classifiers on page 123
- Understanding CoS Rewrite Rules on page 157
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- Defining CoS Rewrite Rules on page 160

Troubleshooting an Unexpected Rewrite Value

Problem

Description: 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 DSCP traffic and one rewrite rule for IEEE 802.1p traffic to an interface. A rewrite rule can contain multiple forwarding-class-to-rewrite-value mappings.

1. To 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 **be2** is being randomly rewritten, and you want to use IEEE 802.1 code point **002** for the **be2** forwarding class:

   ```
   [edit class-of-service rewrite-rules]
   user@switch# set ieee-802.1 custom-rw forwarding-class be2 loss-priority low code-point 002
   ```


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 on page 192**
- **rewrite-rules on page 199**
- Defining CoS Rewrite Rules on page 160
- Monitoring CoS Rewrite Rules on page 205

CHAPTER 5

Configuration Statements for Classifiers and Rewrite Rules

class (Forwarding Classes)

Syntax

```
class {
  class-name {
    pfc-priority pfc-priority;
    queue-num queue-number <no-loss>;
  }
}
```

Hierarchy Level [edit class-of-service forwarding-classes]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
No-loss option introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
PFC-priority statement introduced in Junos OS Release 17.4R1 for the QFX Series.

Description Map one or more forwarding classes to a single queue. Also, when configuring DSCP-based PFC, map a forwarding class to a PFC priority value to use in pause frames when traffic on a DSCP value becomes congested (see *Configuring DSCP-based PFC for Layer 3 Untagged Traffic* for details).

You can map unicast forwarding classes to a unicast queue (0 through 7) and multidestination forwarding classes to a multicast queue (8 through 11). The queue to which you map a forwarding class determines if the forwarding class is a unicast or multicast forwarding class.

NOTE: On systems that do not use the ELS CLI, if you are using Junos OS Release 12.2, use the default forwarding-class-to-queue mapping for the lossless **fcoe** and **no-loss** forwarding classes. If you explicitly configure the lossless forwarding classes, the traffic mapped to those forwarding classes is treated as lossy (best effort) traffic and does *not* receive lossless treatment.

NOTE: On systems that do not use the ELS CLI, if you are using Junos OS Release 12.3 or later, the default configuration is the same as the default configuration for Junos OS Release 12.2, and the default behavior is the same (the **fcoe** and **no-loss** forwarding classes receive lossless treatment). However, if you explicitly configure lossless forwarding classes, you can configure up to six lossless forwarding classes by specifying the **no-loss** option. If you do not specify the **no-loss** option in an explicit forwarding class configuration, the forwarding class is lossy. For example, if you explicitly configure the **fcoe** forwarding class and you do not include the **no-loss** option, the **fcoe** forwarding class is lossy, not lossless.

Options

class-name —Name of the forwarding class.

The remaining statements are explained separately. Click a linked statement in the Syntax section or search for a statement in CLI Explorer for details.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring Forwarding Classes on page 142
- *Understanding CoS Forwarding Classes*
- Understanding CoS Forwarding Classes on page 137

class (Forwarding Class Sets)

Syntax

class *class-name*;

Hierarchy Level

[edit class-of-service forwarding-class-sets *forwarding-class-set-name*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Group forwarding classes into sets of forwarding classes (priority groups). You can group some or all of the configured forwarding classes into up to three unicast forwarding class sets and one multidestination forwarding class set.

Options

class-name —Name of the forwarding class.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Forwarding Class Sets on page 150
- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148

classifiers

List of Syntax

Syntax (BA Classifiers) on page 174
Syntax (Multidestination BA Classifiers) on page 174
Syntax (Interface Classifier Association: DSCP, DSCP IPv6, IEEE) on page 174
Syntax (Global EXP Interface Classifier Association with Interfaces) on page 174

Syntax (BA Classifiers)

```
classifiers {
  (dscp | dscp-ipv6 | ieee-802.1) classifier-name {
    import (classifier-name | default);
    forwarding-class class-name {
      loss-priority level {
        code-points [ aliases ] [ bit-patterns ];
      }
    }
  }
}
```

Syntax (Multidestination BA Classifiers)

```
classifiers {
  (dscp | ieee-802.1) classifier-name;
}
```

Syntax (Interface Classifier Association: DSCP, DSCP IPv6, IEEE)

```
classifiers {
  (dscp | dscp-ipv6 | ieee-802.1) (default | classifier-name);
}
```

Syntax (Global EXP Interface Classifier Association with Interfaces)

```
classifiers {
  exp classifier-name;
}
```

Hierarchy Level (BA Classifiers)

[edit class-of-service],

Hierarchy Level (Multidestination BA Classifiers)

[edit class-of-service multi-destination],

Hierarchy Level (Interface Classifier Association: DSCP, DSCP IPv6, IEEE)

[edit class-of-service interfaces *interface-name* unit *logical-unit-number*]

Hierarchy Level (Global EXP Classifier)

[edit class-of-service system-defaults]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
EXP statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Define a unicast or multidestination CoS behavior aggregate (BA) classifier.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifier configuration.

Options The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Configuring a Global MPLS EXP Classifier*
- *Example: Configuring Unicast Classifiers*
- *Example: Configuring Multidestination (Multicast, Broadcast, DLF) Classifiers*
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- *Understanding CoS MPLS EXP Classifiers and Rewrite Rules*

code-point (Rewrite Rules)

Syntax code-point [*alias*] [*bit-pattern*];

Hierarchy Level [edit class-of-service rewrite-rules (dscp | ieee-802.1) forwarding-class *class-name* loss-priority *level*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure a code-point alias or bit set to apply to a forwarding class for a rewrite rule.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP rewrite rules.

Options *alias*—Name of the alias.

bit-pattern—Value of the code-point bits, in decimal form.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS Rewrite Rules on page 160
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123

code-point-aliases

Syntax

```
code-point-aliases {
  (dscp| dscp-ipv6 | ieee-802.1 | exp) {
    alias-name bits;
  }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Define an alias for a CoS marker. You can use the alias instead of the bit pattern when you specify the code point during configuration.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP code-point aliases.

Options **(dcsp | dscp-ipv6 | ieee-802.1 | exp)**—Set the type of classifier for which you are creating an alias.

alias-name—Name of the code-point alias.

bits —Value of the code-point bits, in decimal form.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS Code-Point Aliases on page 136
- Understanding CoS Code-Point Aliases on page 134

code-points (CoS)

Syntax code-points ([*aliases*] | [*bit-patterns*]);

Hierarchy Level [edit class-of-service classifiers *type classifier-name* forwarding-class *class-name* loss-priority *level*]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced in Junos OS Release 9.2 for SRX Series devices.
Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 12.1X44 for the SRX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
Statement introduced in Junos OS Release 14.2 for PTX Series Packet Transport Routers.

Description Specify one or more DSCP code-point aliases or bit sets to apply to a forwarding class..

NOTE: OCX Series switches do not support MPLS, and therefore, do not support EXP code points or code point aliases.

Options ***aliases***—Name of the DSCP alias.

bit-patterns—Value of the code-point bits, in six-bit binary form.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- *Understanding Interfaces*
- *Understanding How Behavior Aggregate Classifiers Prioritize Trusted Traffic*
- *Example: Configuring Behavior Aggregate Classifiers*
- *Example: Configuring BA Classifiers on Transparent Mode Security Devices*

dscp

List of Syntax

Syntax (Classifier) on page 179
Syntax (Code-Point Alias) on page 179
Syntax (Multidestination Classifier) on page 179
Syntax (Interface Classifier Association) on page 179
Syntax (Rewrite Rule) on page 179

Syntax (Classifier)

```
dscp classifier-name {
  import (classifier-name | default);
  forwarding-class class-name {
    loss-priority level {
      code-points [ aliases ] [ bit-patterns ];
    }
  }
}
```

Syntax (Code-Point Alias) / **Code-Point Alias Configuration**

```
dscp alias-name bit-pattern;
```

Syntax (Multidestination Classifier) / **Multidestination Classifier Configuration**

```
dscp classifier-name;
```

Syntax (Interface Classifier Association) / **Interface Classifier Association**

```
dscp (classifier-name | default);
```

Syntax (Rewrite Rule) / **Rewrite Rule Configuration**

```
dscp rewrite-name {
  import (rewrite-name | default);
  forwarding-class class-name {
    loss-priority level {
      code-point [ aliases ] [ bit-patterns ];
    }
  }
}
```

Hierarchy Level (Classifier)

[edit class-of-service classifiers],

Hierarchy Level (Code-Point Aliases)

[edit class-of-service code-point-aliases],

Hierarchy Level (Multidestination Classifier)

[edit class-of-service multi-destination classifiers],

Hierarchy Level (Interface Classifier Association)

[edit class-of-service interfaces *interface-name* unit *logical-unit-number* classifiers],
[edit class-of-service interfaces *interface-name* unit *logical-unit-number* rewrite-rules],

Hierarchy Level (Rewrite Rule)

[edit class-of-service rewrite-rules]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Define the Differentiated Services code point (DSCP) mapping that is applied to the packets.

Options

classifier-name—Name of the classifier.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- *Example: Configuring Unicast Classifiers*
- *Example: Configuring Classifiers*
- Defining CoS Code-Point Aliases on page 136
- Defining CoS Rewrite Rules on page 160
- Assigning CoS Components to Interfaces on page 21
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- Understanding CoS Rewrite Rules on page 157
- *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces*
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162

dscp-ipv6

List of Syntax

Syntax (Classifier) on page 181
Syntax (Code-Point Alias) on page 181
Syntax (Interface Classifier Association) on page 181
Syntax (Rewrite Rule) on page 181

Syntax (Classifier)

```
dscp-ipv6 classifier-name {
  import (classifier-name | default);
  forwarding-class class-name {
    loss-priority level {
      code-points [ aliases ] [ bit-patterns ];
    }
  }
}
```

Syntax (Code-Point Alias)

dscp-ipv6 *alias-name bit-pattern*;

Syntax (Interface Classifier Association)

dscp-ipv6 (*classifier-name* | default);

Syntax (Rewrite Rule)

```
dscp-ipv6 rewrite-name {
  import (rewrite-name | default);
  forwarding-class class-name {
    loss-priority level {
      code-point [ aliases ] [ bit-patterns ];
    }
  }
}
```

Hierarchy (Classifier)

[edit class-of-service classifiers],

Hierarchy (Code-Point Alias)

[edit class-of-service code-point-aliases],

Hierarchy (Interface Classifier Association)

[edit class-of-service interfaces *interface-name* unit *logical-unit-number* classifiers],
[edit class-of-service interfaces *interface-name* unit *logical-unit-number* rewrite-rules],

Hierarchy (Rewrite Rule)

[edit class-of-service rewrite-rules]

Release Information

Statement introduced in Junos OS Release 12.2 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series

Description

Define the Differentiated Services code point (DSCP) IPv6 mapping that is applied to the packets.

NOTE: On switches that use different classifiers for unicast and multidestination (multicast, broadcast, and destination lookup fail) traffic, there is no DSCP IPv6 classifier for multidestination (multicast, broadcast, and destination lookup fail) traffic. Multidestination IPv6 traffic uses the multidestination DSCP classifier.

Options The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS Code-Point Aliases on page 136
- *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces*
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162

dscp-code-point

Syntax dscp-code-point *code-point*;

Hierarchy Level [edit class-of-service host-outbound-traffic]

Release Information Statement introduced in Junos OS Release 11.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Set the value of the DSCP code point in the type of service (ToS) field of the packet generated by the Routing Engine (host).

Options **code-point**—Six-bit DSCP code point value.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Changing the Host Outbound Traffic Default Queue Mapping on page 156
- Understanding Host Routing Engine Outbound Traffic Queues and Defaults on page 154

forwarding-class

List of Syntax

Classifier on page 183
Rewrite Rule on page 183
Scheduler Map on page 183
Interface on page 183

Classifier

```
forwarding-class class-name {
  loss-priority level {
    code-points [ aliases ] [ bit-patterns ];
  }
}
```

Rewrite Rule

```
forwarding-class class-name {
  loss-priority level {
    code-point [ aliases ] [ bit-patterns ];
  }
}
```

Scheduler Map

```
forwarding-class class-name {
  scheduler scheduler-name;
}
```

Interface

```
forwarding-class class-name;
```

Classifier Hierarchy Level

[edit class-of-service classifiers (dscp | dscp-ipv6 | ieee-802.1 | exp) *classifier-name*],

Rewrite Rule Hierarchy Level

[edit class-of-service rewrite-rules] (dscp | dscp-ipv6 | ieee-802.1) *rewrite-name* | exp],

Scheduler Map Hierarchy Level

[edit class-of-service scheduler-maps *map-name*],

Interface Hierarchy Level

[edit class-of-service interfaces *interface-name* unit *logical-unit-number*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

- Classifiers—Assign incoming traffic to the specified forwarding class based on the specified code point values and assign that traffic the specified loss priority
- Rewrite rules—At the egress interface, change (rewrite) the value of the code point bits and the loss priority to specified new values for traffic assigned to the specified forwarding class, before forwarding the traffic to the next hop.
- Scheduler maps—Apply the specified scheduler to the specified forwarding class.

- Interfaces—Assign the specified forwarding class to the interface to use as a fixed classifier (all incoming traffic on the interface is classified into that forwarding class).

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifiers or rewrite rules.

Options *class-name*—Name of the forwarding class.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

forwarding-class (Host Outbound Traffic)

Syntax forwarding-class *class-name*;

Hierarchy Level [edit class-of-service host-outbound-traffic]

Release Information Statement introduced in Junos OS Release 11.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Define forwarding class name for outbound host traffic (traffic generated by the Routing Engine).

Options *class-name*—Name of the forwarding class.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Changing the Host Outbound Traffic Default Queue Mapping on page 156
- Understanding Host Routing Engine Outbound Traffic Queues and Defaults on page 154

forwarding-class-sets

Syntax

forwarding-class-sets *forwarding-class-set-name* {
  class *class-name*;
}

Hierarchy Level

[edit class-of-service]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series

Description

Assign forwarding classes to forwarding class sets (priority groups).

Options

fowarding-class-set-name—Name of the forwarding class set.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Forwarding Class Sets on page 150
- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148

forwarding-classes

Syntax

```
forwarding-classes {
  class {
    class-name {
      pfc-priority pfc-priority;
      queue-num queue-number <no-loss>;
    }
  }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
No-loss option introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series
PFC-priority statement introduced in Junos OS Release 17.4R1 for the QFX Series.

Description Map one or more forwarding classes to a single output queue. Also, when configuring DSCP-based PFC, map a forwarding class to a PFC priority value to use in pause frames when traffic on a DSCP value becomes congested (see *Configuring DSCP-based PFC for Layer 3 Untagged Traffic* for details).

Switches that use different forwarding classes for unicast and multidestination (multicast, broadcast, and destination lookup fail) traffic support 12 forwarding classes and 12 output queues (0 through 11). You map unicast forwarding classes to a unicast queue (0 through 7) and multidestination forwarding classes to a multidestination queue (8 through 11). The queue to which you map a forwarding class determines if the forwarding class is a unicast or multidestination forwarding class.

Switches that use the same forwarding classes for unicast and multidestination traffic support eight forwarding classes and eight output queues (0 through 7). You map forwarding classes to output queues. All traffic classified into one forwarding class (unicast and multidestination) uses the same output queue.

You cannot configure weighted random early detection (WRED) packet drop on forwarding classes configured with the no-loss packet drop attribute. Do not associate a drop profile with lossless forwarding classes.

NOTE: If you map more than one forwarding class to a queue, all of the forwarding classes mapped to the queue must to the same queue must have the same packet drop attribute (all of the forwarding classes must be lossy, or all of the forwarding classes mapped to a queue must be lossless).

OCX Series switches do not support the no-loss packet drop attribute and do not support lossless forwarding classes. On OCX Series switches, do not configure the no-loss packet drop attribute on forwarding classes, and do not map traffic to the default **fcoe** and **no-loss** forwarding classes (both of these default forwarding classes carry the no-loss packet drop attribute).

NOTE: On switches that do not use the ELS CLI, if you are using Junos OS Release 12.2, use the default forwarding-class-to-queue mapping for the lossless **fcoe** and **no-loss** forwarding classes. If you explicitly configure the lossless forwarding classes, the traffic mapped to those forwarding classes is treated as lossy (best effort) traffic and does *not* receive lossless treatment.

NOTE: On switches that do not use the ELS CLI, if you are using Junos OS Release 12.3 or later, the default configuration is the same as the default configuration for Junos OS Release 12.2, and the default behavior is the same (the **fcoe** and **no-loss** forwarding classes receive lossless treatment). However, if you explicitly configure lossless forwarding classes, you can configure up to six lossless forwarding classes by specifying the **no-loss** option. If you do not specify the **no-loss** option in an explicit forwarding class configuration, the forwarding class is lossy. For example, if you explicitly configure the **fcoe** forwarding class and you do not include the **no-loss** option, the **fcoe** forwarding class is lossy, not lossless.

Options The remaining statements are explained separately. Click a linked statement in the Syntax section or search for a statement in CLI Explorer for details.

Required Privilege Level
interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

host-outbound-traffic

Syntax

```
host-outbound-traffic {
  forwarding-class class-name;
  dscp-code-point code-point;
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Allow queue selection for traffic generated by the Routing Engine (host). The selected queue must be configured properly. You can also configure specific DSCP code point bits for the type of service (ToS) field of the generated packets. This configuration does not affect transit packets or incoming packets. This is a global configuration that only affects packets originating on the Routing Engine. If you do not configure an output queue for host outbound traffic, the switch uses the default queue mapping.

Options The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Changing the Host Outbound Traffic Default Queue Mapping on page 156
- Understanding Host Routing Engine Outbound Traffic Queues and Defaults on page 154

ieee-802.1

List of Syntax

Syntax (Classifier)

```
ieee-802.1 classifier-name {
  import (classifier-name | default);
  forwarding-class class-name {
    loss-priority level {
      code-points [ aliases ] [ bit-patterns ];
    }
  }
}
```

Syntax (Code-Point Alias) / Code-Point Alias Configuration

```
ieee-802.1 alias-name bit-pattern;
```

Syntax (Multidestination Classifier) / Multidestination Classifier Configuration

```
ieee-802.1 classifier-name;
```

Syntax (Interface Classifier Association) / Interface Classifier Association

```
ieee-802.1 (classifier-name | default);
```

Syntax (Rewrite Rule) / Rewrite Rule Configuration

```
ieee-802.1 rewrite-name {
  import (rewrite-name | default);
  forwarding-class class-name {
    loss-priority level {
      code-point [ aliases ] [ bit-patterns ];
    }
  }
}
```

Hierarchy Level (Classifier)

[edit class-of-service classifiers],

Hierarchy Level (Code-Point Alias)

[edit class-of-service code-point-aliases],

Hierarchy Level (Multidestination Classifier)

[edit class-of-service multi-destination classifiers],

Hierarchy Level (Interface Classifier Association)

[edit class-of-service interfaces *interface-name* unit *logical-unit-number* classifiers],
[edit class-of-service interfaces *interface-name* unit *logical-unit-number* rewrite-rules],

Hierarchy Level (Rewrite Rule) [edit class-of-service rewrite-rules]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure an IEEE 802.1 classifier, configure an IEEE 802.1 code-point alias, apply a fixed IEEE 802.1 classifier to an interface, or apply an IEEE-802.1 rewrite rule.

Options ***classifier-name***—Name of the classifier.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- *Example: Configuring Unicast Classifiers*
- Defining CoS Code-Point Aliases on page 136
- Defining CoS Rewrite Rules on page 160
- Assigning CoS Components to Interfaces on page 21
- *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces*
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- Understanding CoS Rewrite Rules on page 157

import

Syntax import (*import* | default);

Hierarchy Level [edit class-of-service classifiers (dscp | dscp-ipv6 | ieee-802.1 | exp) *classifier-name*],
[edit class-of-service rewrite-rules (dscp | dscp-ipv6 | ieee-802.1 | exp) *classifier-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Specify a default or previously defined classifier.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifiers and rewrite rules.

Options *import* —Name of the classifier mapping configured at the **[edit class-of-service classifiers]** hierarchy level.

default—Default classifier mapping.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- *Example: Configuring Unicast Classifiers*
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- Defining CoS Rewrite Rules on page 160
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- *Understanding CoS Classifiers*
- Understanding CoS Rewrite Rules on page 157

interfaces (Class of Service)

Syntax

```
interfaces {
interface-name {
  congestion-notification-profile profile-name {
  }
  forwarding-class forwarding-class-name;
  forwarding-class-set forwarding-class-set-name {
    output-traffic-control-profile profile-name;
  }
  rewrite-value {
    input {
      ieee-802.1{
        code-point code-point-bits;
      }
    }
  }
  scheduler-map scheduler-map-name
  unit logical-unit-number {
    classifiers {
      (dscp | dscp-ipv6 | ieee-802.1 | exp) (classifier-name | default);
    }
    forwarding-class class-name;
    rewrite-rules {
      (dscp | dscp-ipv6 | ieee-802.1 | exp) (classifier-name | default);
    }
  }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series

Description Configure interface-specific CoS properties for incoming packets.

NOTE: Only switches that support direct port scheduling also support applying a scheduler map directly to an interface. When using enhanced transmission selection (ETS) hierarchical port scheduling, you cannot apply a scheduler map directly to an interface; instead, you associate the scheduler map with a traffic control profile and apply the traffic control profile to the interface.

NOTE: Only switches that support native Fibre Channel interfaces support the **rewrite-value** statement, which enables you to rewrite the IEEE 802.1p code points on native Fibre Channel interfaces.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifiers or rewrite rules. OCX Series switches do not support the congestion-notification-profile configuration statement, which applies priority-based flow control (PFC) to interface output queues.

Options *interface-name*—Name of the interface.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Assigning CoS Components to Interfaces on page 21
- *Interfaces Overview for Switches*

loss-priority (Classifiers)

Syntax

```
loss-priority level {
  code-points [ aliases ] [ bit-patterns ];
}
```

Hierarchy Level

[edit class-of-service classifiers (dscp | dscp-ipv6 | ieee-802.1) *classifier-name* forwarding-class *class-name*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Configure packet loss priority value for a specific set of code-point aliases and bit patterns.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifiers.

Options

level—Can be one of the following:

- **low**—Packet has low loss priority.
- **medium-high**—Packet has medium-high loss priority.
- **high**—Packet has high loss priority.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- *Example: Configuring Unicast Classifiers*
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123

loss-priority (Rewrite Rules)

Syntax

```
loss-priority level {
  code-point (alias | bit-pattern);
}
```

Hierarchy Level

[edit class-of-service rewrite-rules (dscp | ieee-802.1) *rewrite-name* forwarding-class *class-name*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Specify a loss priority to which to apply a rewrite rule. The rewrite rule sets the code-point aliases and bit patterns for a specific forwarding class and loss priority. Packets that match the forwarding class and loss priority are rewritten with the rewrite code-point alias or bit pattern.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP rewrite rules.

Options

level—Can be one of the following:

- **low**—Packet has low loss priority.
- **medium-high**—Packet has medium-high loss priority.
- **high**—Packet has high loss priority.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS Rewrite Rules on page 160
- Understanding CoS Rewrite Rules on page 157

multi-destination

Syntax

```
multi-destination {
  classifiers {
    (dscp | dscp-ipv6 | ieee-802.1 | inet-precedence) classifier-name;
  }
}
```

Hierarchy Level [edit class-of-service],

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Define a multicast CoS behavior aggregate (BA) classifier.

Options The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Example: Configuring Multidestination (Multicast, Broadcast, DLF) Classifiers*
- Assigning CoS Components to Interfaces on page 21
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123

queue-num

Syntax queue-num *queue-number* <no-loss>;

Hierarchy Level [edit class-of-service forwarding-classes class *class-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
No-loss option introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Map a forwarding class to an output queue number. Optionally, configure the forwarding class as a lossless forwarding class. Each switch provides enough output queues so that you can map forwarding classes to queues on a one-to-one basis, so each forwarding class can have a dedicated output queue.

On switches that use different forwarding classes and output queues for unicast and multidestination (multicast, broadcast, destination lookup fail) traffic, the switch supports 12 forwarding classes and 12 output queues, eight of each for unicast traffic and four of each for multidestination traffic. You can map some or all of the eight unicast forwarding classes to a unicast queue (0 through 7) and some or all of the four multidestination forwarding classes to the a multidestination queue (8 through 11). You cannot map a forwarding class to more than one queue (each forwarding class maps to one and only one queue), but you can map multiple forwarding classes to one queue. The queue to which you map a forwarding class determines if the forwarding class is a unicast or multidestination forwarding class.

On switches that use the same forwarding classes and output queues for unicast and multidestination traffic, the switch supports eight forwarding classes and eight output queues. You can map some or all of the eight of the forwarding classes to queues (0 through 7). You cannot map a forwarding class to more than one queue (each forwarding class maps to one and only one queue), but you can map multiple forwarding classes to one queue.

You cannot configure weighted random early detection (WRED) packet drop on forwarding classes configured with the no-loss packet drop attribute. Do not associate a drop profile with lossless forwarding classes. Instead, use priority-based flow control (PFC) to prevent frame drop on lossless forwarding classes.

NOTE: If you map more than one forwarding class to a queue, all of the forwarding classes mapped to the same queue must have the same packet drop attribute (all of the forwarding classes must be lossy, or all of the forwarding classes mapped to a queue must be lossless).

OCX Series switches do not support the no-loss packet drop attribute and do not support lossless forwarding classes. On OCX Series switches, do not configure the no-loss packet drop attribute on forwarding classes, and do not map traffic to the default **fcoe** and **no-loss** forwarding classes (both of these default forwarding classes carry the no-loss packet drop attribute).

NOTE: On systems that do not use the ELS CLI, if you are using Junos OS Release 12.2, use the default forwarding-class-to-queue mapping for the lossless **fcoe** and **no-loss** forwarding classes. If you explicitly configure lossless forwarding classes, the traffic mapped to those forwarding classes is treated as lossy (best effort) traffic and does *not* receive lossless treatment.

NOTE: On systems that do not use the ELS CLI, if you are using Junos OS Release 12.3 or later, the default configuration is the same as the default configuration for Junos OS Release 12.2, and the default behavior is the same (the **fcoe** and **no-loss** forwarding classes receive lossless treatment). However, if you explicitly configure lossless forwarding classes, you can configure up to six lossless forwarding classes by specifying the **no-loss** option. If you do not specify the **no-loss** option in an explicit forwarding class configuration, the forwarding class is lossy. For example, if you explicitly configure the **fcoe** forwarding class and you do not include the **no-loss** option, the **fcoe** forwarding class is lossy, not lossless.

Options

queue-number—(Switches that use different output queues for unicast and multidestination traffic) Number of the CoS unicast queue (0 through 7) or the CoS multidestination queue (8 through 11).

queue-number—(Switches that use the same output queues for unicast and multidestination traffic) Number of the CoS queue (0 through 7).

no-loss—Optional packet drop attribute keyword to configure the forwarding class as lossless.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

rewrite-rules

List of Syntax

Syntax (Rewrite Rule Configuration) on page 199
Syntax (Rewrite Rule Association with Interface) on page 199

Syntax (Rewrite Rule Configuration)

```
rewrite-rules {
  (dscp | dscp-ipv6 | ieee-802.1 | exp) rewrite-name {
    import (rewrite-name | default);
    forwarding-class class-name {
      loss-priority priority code-point (alias | bits);
    }
  }
}
```

Syntax (Rewrite Rule Association with Interface)

```
rewrite-rules {
  (dscp | dscp-ipv6 | ieee-802.1 | exp) rewrite-name;
}
```

Hierarchy Level (Rewrite Rule Configuration)

[edit class-of-service],

Hierarchy Level (Rewrite Rule Association with Interface)

[edit class-of-service interfaces *interface-name* unit *logical-unit-number*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
EXP statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Configure rewrite rules that map traffic to code points when traffic exits the system, and apply the rewrite rules to a specific interface.

MPLS EXP rewrite rules can only be bound to logical interfaces, not to physical interfaces. You can configure up to 64 EXP rewrite rules, but you can use only 16 EXP rewrite rules on switch interfaces at any given time.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP rewrite rules.

Options

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level

interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Defining CoS Rewrite Rules on page 160
- *Configuring Rewrite Rules for MPLS EXP Classifiers*
- Understanding CoS Rewrite Rules on page 157
- *Understanding CoS MPLS EXP Classifiers and Rewrite Rules*

unit

Syntax

```
unit logical-unit-number {
  classifiers {
    (dscp | dscp-ipv6 | ieee-802.1 | exp) (classifier-name | default);
  }
  forwarding-class class-name;
  rewrite-rules {
    (dscp | dscp-ipv6 | ieee-802.1 | exp) (classifier-name | default);
  }
}
```

Hierarchy Level

[edit class-of-service interfaces *interface-name*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Configure a logical interface on the physical device. You must configure a logical interface to use the physical device.

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifiers and rewrite rules.

Options

logical-unit-number—Number of the logical unit.

Range: 0 through 16,385

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Assigning CoS Components to Interfaces on page 21

CHAPTER 6

Monitoring Commands for Classifiers and Rewrite Rules

Monitoring CoS Classifiers

Purpose Display the mapping of incoming CoS values to forwarding class and loss priority for each classifier.

Action To monitor CoS classifiers in the CLI, enter the CLI command:

user@switch> **show class-of-service classifier**

To monitor a particular classifier in the CLI, enter the CLI command:

user@switch> **show class-of-service classifier name *classifier-name***

To monitor a particular type of classifier in the CLI, enter the CLI command:

user@switch> **show class-of-service classifier type *classifier-type***

Meaning Table 44 on page 202 summarizes key output fields for CoS classifiers.

Table 44: Summary of Key CoS Classifier Output Fields

Field	Values
Classifier	Name of a classifier.
Code point type	Type of classifier: • **dscp**—All classifiers of the DSCP type. • **ieee-802.1**—All classifiers of the IEEE 802.1 type. • **ieee-mcast**—All classifiers of the IEEE 802.1 multicast type. NOTE: QFX10000 switches do not use different classifiers for unicast and multidestination (multicast, broadcast, destination lookup fail) traffic, so multicast-specific classifiers are not supported. • **exp**—All classifiers of the MPLS exp type. NOTE: OCX Series switches do not support MPLS.
Index	Internal index of the classifier.
Code point	DSCP or IEEE 802.1 code point value of the incoming packets, in bits. These values are used for classification.
Forwarding Class	Name of the forwarding class that the classifier assigns to an incoming packet. This class affects the forwarding and scheduling policies that are applied to the packet as it transits the switch.
Loss Priority	Loss priority value that the classifier assigns to the incoming packet based on its code point value.

Monitoring CoS Forwarding Classes

Purpose Use the monitoring functionality to view the current assignment of CoS forwarding classes to queue numbers on the system.

Action To monitor CoS forwarding classes in the CLI, enter the following CLI command:

user@switch> **show class-of-service forwarding-class**

Meaning Some switches use different forwarding classes, output queues, and classifiers for unicast and multidestination (multicast, broadcast, destination lookup fail) traffic. These switches support 12 forwarding classes and output queues, eight for unicast traffic and four for multidestination traffic.

Some switches use the same forwarding classes, output queues, and classifiers for unicast and multidestination traffic. These switches support eight forwarding classes and eight output queues.

Table 45 on page 203 summarizes key output fields on switches that use different forwarding classes and output queues for unicast and multidestination traffic.

Table 45: Summary of Key CoS Forwarding Class Output Fields on Switches that Separate Unicast and Multidestination Traffic

Field	Values
Forwarding Class	Names of forwarding classes assigned to queue numbers. By default, the following unicast forwarding classes are assigned to queues 0, 3, 4, and 7, respectively: • **best-effort**—Provides no special CoS handling of packets. Loss priority is typically not carried in a CoS value. • **fcoe**—Provides guaranteed delivery for Fibre Channel over Ethernet (FCoE) traffic. • **no-loss**—Provides guaranteed delivery for TCP lossless traffic • **network-control**—Packets can be delayed but not dropped. By default, the following multidestination forwarding class is assigned to queue 8: • **mcast**—Provides no special CoS handling of packets.
Queue	Queue number corresponding to (mapped to) the forwarding class name. By default, four queues (0, 3, 4, and 7) are assigned to unicast forwarding classes and one queue (8) is assigned to a multidestination forwarding class: • Queue 0—**best-effort** • Queue 3—**fcoe** • Queue 4—**no-loss** • Queue 7—**network-control** • Queue 8—**mcast**

Table 45: Summary of Key CoS Forwarding Class Output Fields on Switches that Separate Unicast and Multidestination Traffic (continued)

Field	Values
No-Loss	Packet drop attribute associated with each forwarding class: • Disabled—The forwarding class is configured for lossy transport (packets might drop during periods of congestion) • Enabled—The forwarding class is configured for lossless transport NOTE: To achieve lossless transport, you must ensure that priority-based flow control (PFC) and DCBX are properly configured on the lossless priority (IEEE 802.1p code point), and that sufficient port bandwidth is reserved for the lossless traffic flows. OCX Series switches do not support lossless transport.

NOTE: OCX Series switches do not support the default lossless forwarding classes **fcoe** and **no-loss**, and do not support the no-loss packet drop attribute used to configure lossless forwarding classes. On OCX Series switches, do not map traffic to the default **fcoe** and **no-loss** forwarding classes (both of these default forwarding classes carry the no-loss packet drop attribute), and do not configure the no-loss packet drop attribute on forwarding classes.

Table 46 on page 204 summarizes key output fields on switches that use the same forwarding classes and output queues for unicast and multidestination traffic.

Table 46: Summary of Key CoS Forwarding Class Output Fields on Switches That Do Not Separate Unicast and Multidestination Traffic

Field	Values
Forwarding Class	Names of forwarding classes assigned to queue numbers. By default, the following forwarding classes are assigned to queues 0, 3, 4, and 7, respectively: • **best-effort**—Provides no special CoS handling of packets. Loss priority is typically not carried in a CoS value. • **fcoe**—Provides guaranteed delivery for Fibre Channel over Ethernet (FCoE) traffic. • **no-loss**—Provides guaranteed delivery for TCP lossless traffic • **network-control**—Packets can be delayed but not dropped.

Table 46: Summary of Key CoS Forwarding Class Output Fields on Switches That Do Not Separate Unicast and Multidestination Traffic (continued)

Field	Values
Queue	Queue number corresponding to (mapped to) the forwarding class name. By default, four queues (0, 3, 4, and 7) are assigned to forwarding classes: • Queue 0—**best-effort** • Queue 3—**fcoe** • Queue 4—**no-loss** • Queue 7—**network-control**
No-Loss	Packet drop attribute associated with each forwarding class: • Disabled—The forwarding class is configured for lossy transport (packets might drop during periods of congestion). • Enabled—The forwarding class is configured for lossless transport. NOTE: To achieve lossless transport, you must ensure that priority-based flow control (PFC) and DCBX are properly configured on the lossless priority (IEEE 802.1p code point), and that sufficient port bandwidth is reserved for the lossless traffic flows. OCX Series switches do not support lossless transport.

Monitoring CoS Rewrite Rules

Purpose Use the monitoring functionality to display information about CoS value rewrite rules, which are based on the forwarding class and loss priority.

Action To monitor CoS rewrite rules in the CLI, enter the CLI command:

user@switch> **show class-of-service rewrite-rule**

To monitor a particular rewrite rule in the CLI, enter the CLI command:

user@switch> **show class-of-service rewrite-rule name *rewrite-rule-name***

To monitor a particular type of rewrite rule (for example, DSCP, DSCP IPv6, IEEE-802.1, or MPLS EXP) in the CLI, enter the CLI command:

user@switch> **show class-of-service rewrite-rule type *rewrite-rule-type***

Meaning Table 47 on page 206 summarizes key output fields for CoS rewrite rules.

Table 47: Summary of Key CoS Rewrite Rule Output Fields

Field	Values
Rewrite rule	Name of the rewrite rule.
Code point type	Rewrite rule type: • **dscp**—For IPv4 DiffServ traffic. • **dscp-ipv6**—For IPv6 Diffserv traffic. • **ieee-802.1**—For Layer 2 traffic. • **exp**—For MPLS traffic. NOTE: OCX Series switches do not support MPLS.
Index	Internal index for the rewrite rule.
Forwarding class	Name of the forwarding class that is used to determine CoS values for rewriting in combination with loss priority. Rewrite rules are applied to CoS values in outgoing packets based on forwarding class and loss priority setting.
Loss priority	Level of loss priority that is used to determine CoS values for rewriting in combination with forwarding class.
Code point	Rewrite code point value.

Related Documentation

- Defining CoS Rewrite Rules on page 160

Monitoring CoS Code-Point Value Aliases

Purpose Use the monitoring functionality to display information about the CoS code-point value aliases that the system is currently using to represent DSCP and IEEE 802.1p code point bits.

Action To monitor CoS value aliases in the CLI, enter the CLI command:

user@switch> **show class-of-service code-point-aliases**

To monitor a specific type of code-point alias (DSCP, DSCP IPv6, IEEE 802.1, or MPLS EXP) in the CLI, enter the CLI command:

user@switch> **show class-of-service code-point-aliases ieee-802.1**

Meaning Table 48 on page 207 summarizes key output fields for CoS value aliases.

Table 48: Summary of Key CoS Value Alias Output Fields

Field	Values
Code point type	Type of the CoS value: • **dscp**—Examines Layer 3 packet headers for IP packet classification. • **dscp-ipv6**—Examines Layer 3 packet headers for IPv6 packet classification. • **ieee-802.1**—Examines Layer 2 packet headers for packet classification. • **exp**—Examines MPLS packet headers for packet classification. NOTE: OCX Series switches do not support MPLS.
Alias	Name given to a set of bits—for example, **af11** is a name for bits **001010**.
Bit pattern	Set of bits associated with the alias.

Related Documentation

- Defining CoS Code-Point Aliases on page 136

show class-of-service classifier

Syntax

show class-of-service classifier
<name *name*>
<type dscp | type dscp-ipv6 | type exp | type ieee-802.1 | type inet-precedence>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

For each class-of-service (CoS) classifier, display the mapping of code point value to forwarding class and loss priority.

Options

none—Display all classifiers.

name ***name***—(Optional) Display named classifier.

type dscp—(Optional) Display all classifiers of the Differentiated Services code point (DSCP) type.

type dscp-ipv6—(Optional) Display all classifiers of the DSCP for IPv6 type.

type exp—(Optional) Display all classifiers of the MPLS experimental (EXP) type.

type ieee-802.1—(Optional) Display all classifiers of the ieee-802.1 type.

type inet-precedence—(Optional) Display all classifiers of the inet-precedence type.

Required Privilege Level

view

List of Sample Output

show class-of-service classifier type ieee-802.1 on page 209
show class-of-service classifier type ieee-802.1 (QFX Series) on page 209

Output Fields

Table 49 on page 208 describes the output fields for the **show class-of-service classifier** command. Output fields are listed in the approximate order in which they appear.

Table 49: show class-of-service classifier Output Fields

Field Name	Field Description
Classifier	Name of the classifier.
Code point type	Type of the classifier: **exp** (not on EX Series switch), **dscp**, **dscp-ipv6** (not on EX Series switch), **ieee-802.1**, or **inet-precedence**.
Index	Internal index of the classifier.
Code point	Code point value used for classification

Table 49: show class-of-service classifier Output Fields (continued)

Field Name	Field Description
Forwarding class	Classification of a packet affecting the forwarding, scheduling, and marking policies applied as the packet transits the router.
Loss priority	Loss priority value used for classification. For most platforms, the value is **high** or **low**. For some platforms, the value is **high**, **medium-high**, **medium-low**, or **low**.

Sample Output

show class-of-service classifier type ieee-802.1

user@host> show class-of-service classifier type ieee-802.1

```
Classifier: ieee802.1-default, Code point type: ieee-802.1, Index: 3
Code Point         Forwarding Class                  Loss priority
  000              best-effort                         low
  001              best-effort                         high
  010              expedited-forwarding                low
  011              expedited-forwarding                high
  100              assured-forwarding                  low
  101              assured-forwarding                  medium-high
  110              network-control                     low
  111              network-control                     high

Classifier: users-ieee802.1, Code point type: ieee-802.1
  Code point         Forwarding class                    Loss priority
  100                expedited-forwarding                low
```

show class-of-service classifier type ieee-802.1 (QFX Series)

user@switch> show class-of-service classifier type ieee-802.1

```
Classifier: ieee8021p-default, Code point type: ieee-802.1, Index: 11
  Code point         Forwarding class                    Loss priority
  000                best-effort                         low
  001                best-effort                         low
  010                best-effort                         low
  011                fcoe                                low
  100                no-loss                             low
  101                best-effort                         low
  110                network-control                     low
  111                network-control                     low

Classifier: ieee8021p-untrust, Code point type: ieee-802.1, Index: 16
  Code point         Forwarding class                    Loss priority
  000                best-effort                         low
  001                best-effort                         low
  010                best-effort                         low
  011                best-effort                         low
  100                best-effort                         low
  101                best-effort                         low
  110                best-effort                         low
  111                best-effort                         low

Classifier: ieee-mcast, Code point type: ieee-802.1, Index: 46
  Code point         Forwarding class                    Loss priority
```



```
000                 mcast                        low
001                 mcast                        low
010                 mcast                        low
011                 mcast                        low
100                 mcast                        low
101                 mcast                        low
110                 mcast                        low
111                 mcast                        low
```

show class-of-service code-point-aliases

Syntax

show class-of-service code-point-aliases
<dscp | dscp-ipv6 | exp | ieee-802.1 | inet-precedence>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display the mapping of class-of-service (CoS) code point aliases to corresponding bit patterns.

Options

none—Display code point aliases of all code point types.

dscp—(Optional) Display Differentiated Services code point (DSCP) aliases.

dscp-ipv6—(Optional) Display IPv6 DSCP aliases.

exp—(Optional) Display MPLS EXP code point aliases.

ieee-802.1—(Optional) Display IEEE-802.1 code point aliases.

inet-precedence—(Optional) Display IPv4 precedence code point aliases.

Required Privilege Level

view

List of Sample Output

show class-of-service code-point-aliases exp on page 212

Output Fields

Table 50 on page 211 describes the output fields for the **show class-of-service code-point-aliases** command. Output fields are listed in the approximate order in which they appear.

Table 50: show class-of-service code-point-aliases Output Fields

Field Name	Field Description
Code point type	Type of the code points displayed: **dscp**, **dscp-ipv6** (not on EX Series switch), **exp** (not on EX Series switch or the QFX Series), **ieee-802.1**, or **inet-precedence** (not on the QFX Series).
Alias	Alias for a bit pattern.
Bit pattern	Bit pattern for which the alias is displayed.

Sample Output

show class-of-service code-point-aliases exp

```
user@host> show class-of-service code-point-aliases exp
Code point type: exp
  Alias              Bit pattern
  af11               100
  af12               101
  be                 000
  be1                001
  cs6                110
  cs7                111
  ef                 010
  ef1                011
  nc1                110
  nc2                111
```

show class-of-service forwarding-class

Syntax

show class-of-service forwarding-class

Release Information

Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
PFC priority output field introduced for DSCP-based PFC in Junos OS Release 17.4R1 for the QFX Series.

Description

Display information about forwarding classes, including the mapping of forwarding classes to queue numbers.

Required Privilege Level

view

Related Documentation

- *Monitoring CoS Forwarding Classes*
- Monitoring CoS Forwarding Classes on page 202
- *Understanding PFC Using DSCP at Layer 3 for Untagged Traffic*

List of Sample Output

show class-of-service forwarding-class on page 214
show class-of-service forwarding-class (EX8200 Switch) on page 214
show class-of-service forwarding-class (QFX Series) on page 214
show class-of-service forwarding-class (QFX Series with DSCP-based PFC) on page 215

Output Fields

Table 51 on page 213 describes the output fields for the **show class-of-service forwarding-class** command. Output fields are listed in the approximate order in which they appear.

Table 51: show class-of-service forwarding-class Output Fields

Field Name	Field Description
Forwarding class	Name of the forwarding class.
ID	Forwarding class identifier. (QFX5110, QFX5200, and QFX5210 switches only) For DSCP-based PFC, the forwarding class ID is assigned from (and should be the same as) the configured PFC priority for the forwarding class. See *Configuring DSCP-based PFC for Layer 3 Untagged Traffic* for details.
Queue	CoS output queue mapped to the forwarding class.
Policing priority	Not supported on EX Series switches or the QFX Series and can be ignored.

Table 51: show class-of-service forwarding-class Output Fields (continued)

Field Name	Field Description
Fabric priority	(EX8200 switches only) Fabric priority for the forwarding class, either **high** or **low**. Determines the priority of packets entering the switch fabric.
No-Loss	(QFX Series only) Packet loss attribute to differentiate lossless forwarding classes from lossy forwarding classes: • Disabled—Lossless transport is not configured on the forwarding class (packet drop attribute is **drop**). • Enabled—Lossless transport is configured on the forwarding class (packet drop attribute is **no-loss**).
PFC Priority	(QFX5110, QFX5200, and QFX5210 switches only) For DSCP-based PFC, the explicitly configured PFC priority configured for the forwarding class. The DSCP value on which PFC is enabled maps to this priority, and this priority is used in PFC pause frames sent to the peer to request to pause traffic on the mapped DSCP value when the link becomes congested. The forwarding class ID is assigned from and should match this value in the output of this command. See *Configuring DSCP-based PFC for Layer 3 Untagged Traffic* for details.

Sample Output

show class-of-service forwarding-class

```
user@switch> show class-of-service forwarding-class
Forwarding class               ID                     Queue  Policing priority
  best-effort                       0                       0            normal
  expedited-forwarding              1                       5            normal
  assured-forwarding                2                       1            normal
  network-control                   3                       7            normal
```

Sample Output

show class-of-service forwarding-class (EX8200 Switch)

```
user@switch> show class-of-service forwarding-class
Forwarding class                       ID       Queue  Fabric priority
  best-effort                           0        0      low
  expedited-forwarding                  1        5      low
  assured-forwarding                    2        1      low
  network-control                       3        7      low
  mcast-be                              4        2      low
  mcast-ef                              5        4      low
  mcast-af                              6        6      low
```

Sample Output

show class-of-service forwarding-class (QFX Series)

```
user@switch> show class-of-service forwarding-class
```



```
Forwarding class                    ID      Queue   Policing priority  No-Loss

  best-effort                       0        0          normal         Disabled

  fcoe                              1         3          normal         Enabled

  no-loss                           2         4          normal         Enabled

  network-control                   3        7          normal         Disabled

  mcast                             8        8          normal         Disabled
```

show class-of-service forwarding-class (QFX Series with DSCP-based PFC)

```
user@switch> show class-of-service forwarding-class
Forwarding class        ID      Queue  Policing priority  No-Loss     PFC priority

  best-effort            0         0        normal          Disabled
  fcoe                   1         3        normal          Enabled
  no-loss                2         4        normal          Enabled
  fc2                    3         2        normal          Enabled          3
  network-control        5         7        normal          Disabled
  fc1                    7         1        normal          Enabled          7
  mcast                  8         8        normal          Disabled
```

On switches that do not use different forwarding classes and output queues for unicast and multidestination (multicast, broadcast, destination lookup fail) traffic, there is no **mcast** forwarding class and there is no queue 8. (Switches that use different forwarding classes and output queues for unicast and multidestination traffic support 12 forwarding classes and output queues, of which four of each are dedicated to multidestination traffic. Switches that use the same forwarding classes and output queues for unicast and multidestination traffic support eight forwarding classes and eight output queues.)

show class-of-service forwarding-class-set

Syntax show class-of-service forwarding-class-set
<*forwarding-class-set-name*>

Release Information Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Display the forwarding classes associated with each forwarding class set.

Options **none**—Display all forwarding class sets.

forwarding-class-set-name—(Optional) Display the forwarding classes associated with the specified forwarding class set.

Required Privilege Level view

Related Documentation

- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148
- Defining CoS Forwarding Class Sets on page 149
- Example: Configuring Forwarding Class Sets on page 150

Output Fields Table 52 on page 216 describes the output fields for the **show class-of-service forwarding-class-set** command. Output fields are listed in the approximate order in which they appear.

Table 52: show class-of-service forwarding-class-set Output Fields

Field Name	Field Description
Forwarding class set	Name of the forwarding class set.
Type	Internal Junos OS type.
Forwarding class set index	Index of this forwarding class set.
Forwarding class	Name of a forwarding class.
Index	Index of this forwarding class.

Sample Output

show class-of-service forwarding-class-set

```
user@switch> show class-of-service forwarding-class-set
```



```
Forwarding class set: san_fcset, Type: normal-type, Forwarding class set index:
37839
  Forwarding class                       Index
  fcoe                                     1

Forwarding class set: lan_fcset, Type: normal-type, Forwarding class set index:
37840
  Forwarding class                       Index
  best-effort                              0

Forwarding class set: multicast_fcset, Type: normal-type, Forwarding class set
index: 37841
  Forwarding class                       Index
  mcast                                    8
```

show class-of-service forwarding-table

List of Syntax

Syntax on page 218
Syntax (TX Matrix and TX Matrix Plus Router) on page 218

Syntax

show class-of-service forwarding-table

Syntax (TX Matrix and TX Matrix Plus Router)

show class-of-service forwarding-table
<lcc *number*> | <sfc *number*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display the entire class-of-service (CoS) configuration as it exists in the forwarding table. Executing this command is equivalent to executing all **show class-of-service forwarding-table** commands in succession.

Options

lcc *number*—(TX Matrix and TX Matrix Plus router only) (Optional) On a TX Matrix router, display the forwarding table configuration for a specific T640 router (or line-card chassis) configured in a routing matrix. On a TX Matrix Plus router, display the forwarding table configuration for a specific router (or line-card chassis) configured in the routing matrix.

Replace *number* with the following values depending on the LCC configuration:

- 0 through 3, when T640 routers are connected to a TX Matrix router in a routing matrix.
- 0 through 3, when T1600 routers are connected to a TX Matrix Plus router in a routing matrix.
- 0 through 7, when T1600 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.
- 0, 2, 4, or 6, when T4000 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.

sfc *number*—(TX Matrix Plus routers only) (Optional) Display the forwarding table configuration for the TX Matrix Plus router. Replace ***number*** with **0**.

Required Privilege Level

view

List of Sample Output

show class-of-service forwarding-table on page 219
show class-of-service forwarding-table lcc (TX Matrix Plus Router) on page 220

Output Fields

See the output field descriptions for **show class-of-service forwarding-table** commands:

- show class-of-service forwarding-table classifier
- show class-of-service forwarding-table classifier mapping
- show class-of-service forwarding-table drop-profile
- *show class-of-service forwarding-table fabric scheduler-map*
- show class-of-service forwarding-table rewrite-rule
- show class-of-service forwarding-table rewrite-rule mapping
- show class-of-service forwarding-table scheduler-map

Sample Output

show class-of-service forwarding-table

```
user@host> show class-of-service forwarding-table
Classifier table index: 9, # entries: 8, Table type: EXP
Entry #    Code point   Forwarding-class #    PLP
   0          000          0                   0
   1          001          0                   1
   2          010          1                   0
   3          011          1                   1
   4          100          2                   0
   5          101          2                   1
   6          110          3                   0
   7          111          3                   1

                       Table Index/
Interface      Index     Q num      Table type
sp-0/0/0.1001     66         11      IPv4 precedence
sp-0/0/0.2001     67         11      IPv4 precedence
sp-0/0/0.16383    68         11      IPv4 precedence
fe-0/0/0.0        69         11      IPv4 precedence

Interface: sp-0/0/0 (Index: 129, Map index: 2, Map type: FINAL,
 Num of queues: 2):
  Entry 0 (Scheduler index: 16, Forwarding-class #: 0):
    Tx rate: 0 Kb (95%), Buffer size: 95 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
  Entry 1 (Scheduler index: 18, Forwarding-class #: 3):
    Tx rate: 0 Kb (5%), Buffer size: 5 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1

Interface: fe-0/0/0 (Index: 137, Map index: 2, Map type: FINAL,
 Num of queues: 2):
  Entry 0 (Scheduler index: 16, Forwarding-class #: 0):
    Tx rate: 0 Kb (95%), Buffer size: 95 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
  Entry 1 (Scheduler index: 18, Forwarding-class #: 3):
    Tx rate: 0 Kb (5%), Buffer size: 5 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
Interface: fe-0/0/1 (Index: 138, Map index: 2, Map type: FINAL,
 Num of queues: 2):
  Entry 0 (Scheduler index: 16, Forwarding-class #: 0):
```

```
    Tx rate: 0 Kb (95%), Buffer size: 95 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
  Entry 1 (Scheduler index: 18, Forwarding-class #: 3):
    Tx rate: 0 Kb (5%), Buffer size: 5 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1

...

RED drop profile index: 1, # entries: 1
                             Drop
Entry      Fullness(%)   Probability(%)
    0             100              100
```

show class-of-service forwarding-table lcc (TX Matrix Plus Router)

```
user@host> show class-of-service forwarding-table lcc 0
lcc0-re0:
--------------------------------------------------------------------------

Classifier table index: 9, # entries: 64, Table type: IPv6 DSCP
Entry #   Code point   Forwarding-class #   PLP
   0        000000        0        0
   1        000001        0        0
   2        000010        0        0
   3        000011        0        0
   4        000100        0        0
   5        000101        0        0
   6        000110        0        0
   7        000111        0        0
   8        001000        0        0
   9        001001        0        0
  10        001010        0        0
  11        001011        0        0
  12        001100        0        0
  13        001101        0        0
  14        001110        0        0
  15        001111        0        0
  16        010000        0        0
  17        010001        0        0
  18        010010        0        0
  19        010011        0        0
  20        010100        0        0
  21        010101        0        0
  22        010110        0        0
  23        010111        0        0
  24        011000        0        0
  25        011001        0        0
  26        011010        0        0
  27        011011        0        0
  28        011100        0        0
  29        011101        0        0
  30        011110        0        0
  31        011111        0        0
  32        100000        0        0
  33        100001        0        0
  34        100010        0        0
  35        100011        0        0
  36        100100        0        0
  37        100101        0        0
```

```
    38          100110          0         0
    39          100111          0         0
    40          101000          0         0
    41          101001          0         0
    42          101010          0         0
    43          101011          0         0
    44          101100          0         0
    45          101101          0         0
    46          101110          0         0
...
```

show class-of-service forwarding-table classifier

Syntax show class-of-service forwarding-table classifier

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Display the mapping of code point value to queue number and loss priority for each classifier as it exists in the forwarding table.

Options This command has no options.

Required Privilege Level view

List of Sample Output show class-of-service forwarding-table classifier on page 223

Output Fields Table 53 on page 222 describes the output fields for the **show class-of-service forwarding-table classifier** command. Output fields are listed in the approximate order in which they appear.

Table 53: show class-of-service forwarding-table classifier Output Fields

Field Name	Field Description
Classifier table index	Index of the classifier table.
entries	Total number of entries.
Table type	Type of code points in the table: **DSCP**, **EXP** (not on the QFX Series), **IEEE 802.1**, **IPv4 precedence** (not on the QFX Series), or **IPv6 DSCP**.
Entry #	Entry number.
Code point	Code point value used for classification.
Forwarding-class #	Forwarding class to which the code point is assigned.
PLP	Packet loss priority value set by classification. For most platforms, the value can be **0** or **1**. For some platforms, the value is **0**, **1**, **2**, or **3**. The value **0** represents low PLP. The value **1** represents **high** PLP. The value **2** represents medium-low PLP. The value **3** represents medium-high PLP.

Sample Output

show class-of-service forwarding-table classifier

```
user@host> show class-of-service forwarding-table classifier
Classifier table index: 62436, # entries: 64, Table type: DSCP

Entry #   Code point   Forwarding-class #   PLP
   0        000000       0                   0
   1        000001       0                   0
   2        000010       0                   0
   3        000011       0                   0
   4        000100       0                   0
   5        000101       0                   0
   6        000110       0                   0
   7        000111       0                   0
   8        001000       0                   0
   9        001001       0                   0
  10        001010       1                   1
  11        001011       0                   0
...
  60        111100       0                   0
  61        111101       0                   0
  62        111110       0                   0
  63        111111       0                   0
```

show class-of-service forwarding-table classifier mapping

Syntax show class-of-service forwarding-table classifier mapping

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description For each logical interface, display either the table index of the classifier for a given code point type or the queue number (if it is a fixed classification) in the forwarding table.

Options This command has no options.

Required Privilege Level view

List of Sample Output show class-of-service forwarding-table classifier mapping on page 224

Output Fields Table 54 on page 224 describes the output fields for the **show class-of-service forwarding-table classifier mapping** command. Output fields are listed in the approximate order in which they appear.

Table 54: show class-of-service forwarding-table classifier mapping Output Fields

Field Name	Field Description
Table index/ Q num	If the table type is **Fixed**, the number of the queue to which the interface is mapped. For all other types, this value is the classifier index number.
Interface	Name of the logical interface. This field can also show the physical interface (QFX Series).
Index	Logical interface index.
Table type	Type of code points in the table: **DSCP**, **EXP** (not on the QFX Series), **Fixed**, **IEEE 802.1**, **IPv4 precedence** (not on the QFX Series),or **IPv6 DSCP**. **none** if **no-default** option set.

Sample Output

show class-of-service forwarding-table classifier mapping

```
user@host> show class-of-service forwarding-table classifier mapping
                       Table index/
Interface      Index     Q num      Table type
so-5/0/0.0        10     62436      DSCP
so-0/1/0.0        11     62436      DSCP
so-0/2/0.0        12         1      Fixed
so-0/2/1.0        13     62436      DSCP
```



```
so-0/2/1.0        13     62437       IEEE 802.1
so-0/2/2.0        14     62436       DSCP
so-0/2/2.0        14     62438       IPv4 precedence
```

show class-of-service forwarding-table rewrite-rule

Syntax show class-of-service forwarding-table rewrite-rule

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Display mapping of queue number and loss priority to code point value for each rewrite rule as it exists in the forwarding table.

Options This command has no options.

Required Privilege Level view

List of Sample Output show class-of-service forwarding-table rewrite-rule on page 226

Output Fields Table 55 on page 226 describes the output fields for the **show class-of-service forwarding-table rewrite-rule** command. Output fields are listed in the approximate order in which they appear.

Table 55: show class-of-service forwarding-table rewrite-rule Output Fields

Field Name	Field Description
Rewrite table index	Index for this rewrite rule.
# entries	Number of entries in this rewrite rule.
Table type	Type of table: **DSCP**, **EXP** (not on the QFX Series), **EXP-PUSH-3** (not on the QFX Series), **IEEE 802.1**,**IPv4 precedence** (not on the QFX Series), **IPv6 DSCP**, or **Fixed**.
Q#	Queue number to which this entry is assigned.
Low bits	Code point value for low-priority loss profile.
State	State of this code point: **enabled**, **rewritten**, or **disabled**.
High bits	Code point value for high-priority loss profile.

Sample Output

show class-of-service forwarding-table rewrite-rule

```
user@host> show class-of-service forwarding-table rewrite-rule
Rewrite table index: 3753, # entries: 4, Table type: DSCP
Q#     Low bits  State      High bits  State
```

```
0       000111  Enabled       001010   Enabled
2       000000  Disabled      001100   Enabled
1       101110  Enabled       110111   Enabled
3       110000  Enabled       111000   Enabled
```

show class-of-service forwarding-table rewrite-rule mapping

Syntax show class-of-service forwarding-table rewrite-rule mapping

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description For each logical interface, display the table identifier of the rewrite rule map for each code point type.

Options This command has no options.

Required Privilege Level view

List of Sample Output show class-of-service forwarding-table rewrite-rule mapping on page 228

Output Fields Table 56 on page 228 describes the output fields for the **show class-of-service forwarding-table rewrite-rule mapping** command. Output fields are listed in the approximate order in which they appear.

Table 56: show class-of-service forwarding-table rewrite-rule mapping Output Fields

Field Name	Field Description
Interface	Name of the logical interface. This field can also show the physical interface (QFX Series).
Index	Logical interface index.
Table index	Rewrite table index.
Type	Type of classifier: **DSCP**, **EXP** (not on the QFX Series), **EXP-PUSH-3** (not on the QFX Series), **EXP-SWAP-PUSH-2** (not on the QFX Series),**IEEE 802.1**, **IPv4 precedence** (not on the QFX Series), **IPv6 DSCP**, or **Fixed**.

Sample Output

show class-of-service forwarding-table rewrite-rule mapping

```
user@host> show class-of-service forwarding-table rewrite-rule mapping
Interface       Index    Table index   Type
so-5/0/0.0         10          3753   DSCP
so-0/1/0.0         11          3753   DSCP
so-0/2/0.0         12          3753   DSCP
so-0/2/1.0         13          3753   DSCP
so-0/2/2.0         14          3753   DSCP
so-0/2/3.0         15          3753   DSCP
```


show class-of-service interface

Syntax

show class-of-service interface
<comprehensive | detail> <*interface-name*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Forwarding class map information added in Junos OS Release 9.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 12.1 for the PTX Series Packet Transport routers.
Command introduced in Junos OS Release 12.2 for the ACX Series Universal Metro routers.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
Options **detail** and **comprehensive** introduced in Junos OS Release 11.4.
Command introduced in Junos OS Release 15.1R3 on MX Series routers for enhanced subscriber management.

Description

Display the logical and physical interface associations for the classifier, rewrite rules, and scheduler map objects.

NOTE: **On routing platforms with dual Routing Engines, running this command on the backup Routing Engine, with or without any of the available options, is not supported and produces the following error message:**

error: the class-of-service subsystem is not running

Options

none—Display CoS associations for all physical and logical interfaces.

comprehensive—(M Series, MX Series, and T Series routers) (Optional) Display comprehensive quality-of-service (QoS) information about all physical and logical interfaces.

detail—(M Series, MX Series, and T Series routers) (Optional) Display QoS and CoS information based on the interface.

If the **interface** *interface-name* is a physical interface, the output includes:

- Brief QoS information about the physical interface
- Brief QoS information about the logical interface
- CoS information about the physical interface
- Brief information about filters or policers of the logical interface
- Brief CoS information about the logical interface

If the **interface** *interface-name* is a logical interface, the output includes:

- Brief QoS information about the logical interface
- Information about filters or policers for the logical interface
- CoS information about the logical interface

interface-name—(Optional) Display class-of-service (CoS) associations for the specified interface.

none—Display CoS associations for all physical and logical interfaces.

NOTE: ACX5000 routers do not support classification on logical interfaces and therefore do not show CoS associations for logical interfaces with this command.

Required Privilege Level view

Related Documentation

- *Verifying and Managing Junos OS Enhanced Subscriber Management*

List of Sample Output

Output Fields Table 22 on page 63 describes the output fields for the **show class-of-service interface** command. Output fields are listed in the approximate order in which they appear.

Table 57: show class-of-service interface Output Fields

Field Name	Field Description
Physical interface	Name of a physical interface.
Index	Index of this interface or the internal index of this object. (Enhanced subscriber management for MX Series routers) Index values for dynamic CoS traffic control profiles and dynamic scheduler maps are larger for enhanced subscriber management than they are for legacy subscriber management.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Dedicated Queues	Status of dedicated queues configured on an interface. Supported only on Trio MPC/MIC interfaces on MX Series routers. (Enhanced subscriber management for MX-Series routers) This field is not displayed for enhanced subscriber management.
Maximum usable queues	Number of queues you can configure on the interface.
Maximum usable queues	Maximum number of queues you can use.
Total non-default queues created	Number of queues created in addition to the default queues. Supported only on Trio MPC/MIC interfaces on MX Series routers. (Enhanced subscriber management for MX Series routers) This field is not displayed for enhanced subscriber management.
Rewrite Input IEEE Code-point	(QFX3500 switches only) IEEE 802.1p code point (priority) rewrite value. Incoming traffic from the Fibre Channel (FC) SAN is classified into the forwarding class specified in the native FC interface (NP_Port) fixed classifier and uses the priority specified as the IEEE 802.1p rewrite value.
Shaping rate	Maximum transmission rate on the physical interface. You can configure the shaping rate on the physical interface, or on the logical interface, but not on both. Therefore, the **Shaping rate** field is displayed for either the physical interface or the logical interface.
Scheduler map	Name of the output scheduler map associated with this interface. (Enhanced subscriber management for MX Series routers) The name of the dynamic scheduler map object is associated with a generated UID (for example, **SMAP-1_UID1002**) instead of with a subscriber interface.
Scheduler map forwarding class sets	(QFX Series only) Name of the output fabric scheduler map associated with a QFabric system Interconnect device interface.
Input shaping rate	For Gigabit Ethernet IQ2 PICs, maximum transmission rate on the input interface.
Input scheduler map	For Gigabit Ethernet IQ2 PICs, name of the input scheduler map associated with this interface.
Chassis scheduler map	Name of the scheduler map associated with the packet forwarding component queues.
Rewrite	Name and type of the rewrite rules associated with this interface.
Traffic-control-profile	Name of the associated traffic control profile. (Enhanced subscriber management for MX Series routers) The name of the dynamic traffic control profile object is associated with a generated UID (for example, **TC_PROF_100_199_SERIES_UID1006**) instead of with a subscriber interface.
Classifier	Name and type of classifiers associated with this interface.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Forwarding-class-map	Name of the forwarding map associated with this interface.
Congestion-notification	(QFX Series and EX4600 switches only) Congestion notification state, **enabled** or **disabled**.
Logical interface	Name of a logical interface.
Object	Category of an object: **Classifier**, **Fragmentation-map** (for LSQ interfaces only), **Scheduler-map**, **Rewrite**, **Translation Table** (for IQE PICs only), or **traffic-class-map** (for T4000 routers with Type 5 FPCs).
Name	Name of an object.
Type	Type of an object: **dscp**, **dscp-ipv6**, **exp**, **ieee-802.1**, **ip**, **inet-precedence**, or **ieee-802.1ad** (for traffic class map on T4000 routers with Type 5 FPCs)..
Link-level type	Encapsulation on the physical interface.
MTU	MTU size on the physical interface.
Speed	Speed at which the interface is running.
Loopback	Whether loopback is enabled and the type of loopback.
Source filtering	Whether source filtering is enabled or disabled.
Flow control	Whether flow control is enabled or disabled.
Auto-negotiation	(Gigabit Ethernet interfaces) Whether autonegotiation is enabled or disabled.
Remote-fault	(Gigabit Ethernet interfaces) Remote fault status. • **Online**—Autonegotiation is manually configured as online. • **Offline**—Autonegotiation is manually configured as offline.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Device flags	The **Device flags** field provides information about the physical device and displays one or more of the following values: • **Down**—Device has been administratively disabled. • **Hear-Own-Xmit**—Device receives its own transmissions. • **Link-Layer-Down**—The link-layer protocol has failed to connect with the remote endpoint. • **Loopback**—Device is in physical loopback. • **Loop-Detected**—The link layer has received frames that it sent, thereby detecting a physical loopback. • **No-Carrier**—On media that support carrier recognition, no carrier is currently detected. • **No-Multicast**—Device does not support multicast traffic. • **Present**—Device is physically present and recognized. • **Promiscuous**—Device is in promiscuous mode and recognizes frames addressed to all physical addresses on the media. • **Quench**—Transmission on the device is quenched because the output buffer is overflowing. • **Recv-All-Multicasts**—Device is in multicast promiscuous mode and therefore provides no multicast filtering. • **Running**—Device is active and enabled.

Interface flags	The **Interface flags** field provides information about the physical interface and displays one or more of the following values:
	• **Admin-Test**—Interface is in test mode and some sanity checking, such as loop detection, is disabled.
	• **Disabled**—Interface is administratively disabled.
	• **Down**—A hardware failure has occurred.
	• **Hardware-Down**—Interface is nonfunctional or incorrectly connected.
	• **Link-Layer-Down**—Interface keepalives have indicated that the link is incomplete.
	• **No-Multicast**—Interface does not support multicast traffic.
	• **No-receive No-transmit**—Passive monitor mode is configured on the interface.
	• **Point-To-Point**—Interface is point-to-point.
	• **Pop all MPLS labels from packets of depth**—MPLS labels are removed as packets arrive on an interface that has the **pop-all-labels** statement configured. The depth value can be one of the following:
	  ◦ **1**—Takes effect for incoming packets with one label only.
	  ◦ **2**—Takes effect for incoming packets with two labels only.
	  ◦ **[1 2]**—Takes effect for incoming packets with either one or two labels.
	• **Promiscuous**—Interface is in promiscuous mode and recognizes frames addressed to all physical addresses.
	• **Recv-All-Multicasts**—Interface is in multicast promiscuous mode and provides no multicast filtering.
	• **SNMP-Traps**—SNMP trap notifications are enabled.
	• **Up**—Interface is enabled and operational.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Flags	The **Logical interface flags** field provides information about the logical interface and displays one or more of the following values: • **ACFC Encapsulation**—Address control field Compression (ACFC) encapsulation is enabled (negotiated successfully with a peer). • **Device-down**—Device has been administratively disabled. • **Disabled**—Interface is administratively disabled. • **Down**—A hardware failure has occurred. • **Clear-DF-Bit**—GRE tunnel or IPsec tunnel is configured to clear the Don't Fragment (DF) bit. • **Hardware-Down**—Interface protocol initialization failed to complete successfully. • **PFC**—Protocol field compression is enabled for the PPP session. • **Point-To-Point**—Interface is point-to-point. • **SNMP-Traps**—SNMP trap notifications are enabled. • **Up**—Interface is enabled and operational.
Encapsulation	Encapsulation on the logical interface.
Admin	Administrative state of the interface (**Up** or **Down**)
Link	Status of physical link (**Up** or **Down**).
Proto	Protocol configured on the interface.
Input Filter	Names of any firewall filters to be evaluated when packets are received on the interface, including any filters attached through activation of dynamic service.
Output Filter	Names of any firewall filters to be evaluated when packets are transmitted on the interface, including any filters attached through activation of dynamic service.
Link flags	Provides information about the physical link and displays one or more of the following values: • **ACFC**—Address control field compression is configured. The Point-to-Point Protocol (PPP) session negotiates the ACFC option. • **Give-Up**—Link protocol does not continue connection attempts after repeated failures. • **Loose-LCP**—PPP does not use the Link Control Protocol (LCP) to indicate whether the link protocol is operational. • **Loose-LMI**—Frame Relay does not use the Local Management Interface (LMI) to indicate whether the link protocol is operational. • **Loose-NCP**—PPP does not use the Network Control Protocol (NCP) to indicate whether the device is operational. • **Keepalives**—Link protocol keepalives are enabled. • **No-Keepalives**—Link protocol keepalives are disabled. • **PFC**—Protocol field compression is configured. The PPP session negotiates the PFC option.
Hold-times	Current interface hold-time up and hold-time down, in milliseconds.
CoS queues	Number of CoS queues configured.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Last flapped	Date, time, and how long ago the interface went from down to up. The format is **Last flapped:** ***year-month-day hour:minute:second:timezone*** **(*hour:minute:second* ago)**. For example, **Last flapped: 2002-04-26 10:52:40 PDT (04:33:20 ago).**
Statistics last cleared	Number and rate of bytes and packets received and transmitted on the physical interface. • **Input bytes**—Number of bytes received on the interface. • **Output bytes**—Number of bytes transmitted on the interface. • **Input packets**—Number of packets received on the interface. • **Output packets**—Number of packets transmitted on the interface.
Exclude Overhead Bytes	Exclude the counting of overhead bytes from aggregate queue statistics. • **Disabled**—Default configuration. Includes the counting of overhead bytes in aggregate queue statistics. • **Enabled**—Excludes the counting of overhead bytes from aggregate queue statistics for just the physical interface. • **Enabled for hierarchy**—Excludes the counting of overhead bytes from aggregate queue statistics for the physical interface as well as all child interfaces, including logical interfaces and interface sets.
IPv6 transit statistics	Number of IPv6 transit bytes and packets received and transmitted on the logical interface if IPv6 statistics tracking is enabled.
Input errors	Input errors on the interface. The labels are explained in the following list: • **Errors**—Sum of the incoming frame aborts and FCS errors. • **Drops**—Number of packets dropped by the input queue of the I/O Manager ASIC. If the interface is saturated, this number increments once for every packet that is dropped by the ASIC's RED mechanism. • **Framing errors**—Number of packets received with an invalid frame checksum (FCS). • **Runts**—Number of frames received that are smaller than the runt threshold. • **Giants**—Number of frames received that are larger than the giant threshold. • **Bucket Drops**—Drops resulting from the traffic load exceeding the interface transmit or receive leaky bucket configuration. • **Policed discards**—Number of frames that the incoming packet match code discarded because they were not recognized or not of interest. Usually, this field reports protocols that Junos OS does not handle. • **L3 incompletes**—Number of incoming packets discarded because they failed Layer 3 (usually IPv4) sanity checks of the header. For example, a frame with less than 20 bytes of available IP header is discarded. Layer 3 incomplete errors can be ignored by configuring the **ignore-l3-incompletes** statement. • **L2 channel errors**—Number of times the software did not find a valid logical interface for an incoming frame. • **L2 mismatch timeouts**—Number of malformed or short packets that caused the incoming packet handler to discard the frame as unreadable. • **HS link CRC errors**—Number of errors on the high-speed links between the ASICs responsible for handling the router interfaces. • **HS link FIFO overflows**—Number of FIFO overflows on the high-speed links between the ASICs responsible for handling the router interfaces.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Output errors	Output errors on the interface. The labels are explained in the following list: • **Carrier transitions**—Number of times the interface has gone from **down** to **up**. This number does not normally increment quickly, increasing only when the cable is unplugged, the far-end system is powered down and up, or another problem occurs. If the number of carrier transitions increments quickly (perhaps once every 10 seconds), the cable, the far-end system, or the PIC is malfunctioning. • **Errors**—Sum of the outgoing frame aborts and FCS errors. • **Drops**—Number of packets dropped by the output queue of the I/O Manager ASIC. If the interface is saturated, this number increments once for every packet that is dropped by the ASIC's RED mechanism. NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), the **Drops** field does not always use the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs. • **Aged packets**—Number of packets that remained in shared packet SDRAM so long that the system automatically purged them. The value in this field should never increment. If it does, it is most likely a software bug or possibly malfunctioning hardware. • **HS link FIFO underflows**—Number of FIFO underflows on the high-speed links between the ASICs responsible for handling the router interfaces. • **MTU errors**—Number of packets whose size exceeds the MTU of the interface.
Egress queues	Total number of egress Maximum usable queues on the specified interface.
Queue counters	CoS queue number and its associated user-configured forwarding class name. • **Queued packets**—Number of queued packets. • **Transmitted packets**—Number of transmitted packets. • **Dropped packets**—Number of packets dropped by the ASIC's RED mechanism. NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), the **Dropped packets** field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.
SONET alarms **SONET defects**	(SONET) SONET media-specific alarms and defects that prevent the interface from passing packets. When a defect persists for a certain period, it is promoted to an alarm. Based on the router configuration, an alarm can ring the red or yellow alarm bell on the router or light the red or yellow alarm LED on the craft interface. See these fields for possible alarms and defects: **SONET PHY**, **SONET section**, **SONET line**, and **SONET path**.
SONET PHY	Counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET PHY** field has the following subfields: • **PLL Lock**—Phase-locked loop • **PHY Light**—Loss of optical signal

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
SONET section	Counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET section** field has the following subfields: • **BIP-B1**—Bit interleaved parity for SONET section overhead • **SEF**—Severely errored framing • **LOS**—Loss of signal • **LOF**—Loss of frame • **ES-S**—Errored seconds (section) • **SES-S**—Severely errored seconds (section) • **SEFS-S**—Severely errored framing seconds (section)
SONET line	Active alarms and defects, plus counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET line** field has the following subfields: • **BIP-B2**—Bit interleaved parity for SONET line overhead • **REI-L**—Remote error indication (near-end line) • **RDI-L**—Remote defect indication (near-end line) • **AIS-L**—Alarm indication signal (near-end line) • **BERR-SF**—Bit error rate fault (signal failure) • **BERR-SD**—Bit error rate defect (signal degradation) • **ES-L**—Errored seconds (near-end line) • **SES-L**—Severely errored seconds (near-end line) • **UAS-L**—Unavailable seconds (near-end line) • **ES-LFE**—Errored seconds (far-end line) • **SES-LFE**—Severely errored seconds (far-end line) • **UAS-LFE**—Unavailable seconds (far-end line)

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
SONET path	Active alarms and defects, plus counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET path** field has the following subfields: • **BIP-B3**—Bit interleaved parity for SONET section overhead • **REI-P**—Remote error indication • **LOP-P**—Loss of pointer (path) • **AIS-P**—Path alarm indication signal • **RDI-P**—Path remote defect indication • **UNEQ-P**—Path unequipped • **PLM-P**—Path payload (signal) label mismatch • **ES-P**—Errored seconds (near-end STS path) • **SES-P**—Severely errored seconds (near-end STS path) • **UAS-P**—Unavailable seconds (near-end STS path) • **ES-PFE**—Errored seconds (far-end STS path) • **SES-PFE**—Severely errored seconds (far-end STS path) • **UAS-PFE**—Unavailable seconds (far-end STS path)
Received SONET overhead **Transmitted SONET overhead**	Values of the received and transmitted SONET overhead: • **C2**—Signal label. Allocated to identify the construction and content of the STS-level SPE and for PDI-P. • **F1**—Section user channel byte. This byte is set aside for the purposes of users. • **K1** and **K2**—These bytes are allocated for APS signaling for the protection of the multiplex section. • **J0**—Section trace. This byte is defined for STS-1 number 1 of an STS-*N* signal. Used to transmit a 1-byte fixed-length string or a 16-byte message so that a receiving terminal in a section can verify its continued connection to the intended transmitter. • **S1**—Synchronization status. The S1 byte is located in the first STS-1 number of an STS-*N* signal. • **Z3** and **Z4**—Allocated for future use.
Received path trace **Transmitted path trace**	SONET/SDH interfaces allow path trace bytes to be sent inband across the SONET/SDH link. Juniper Networks and other router manufacturers use these bytes to help diagnose misconfigurations and network errors by setting the transmitted path trace message so that it contains the system hostname and name of the physical interface. The received path trace value is the message received from the router at the other end of the fiber. The transmitted path trace value is the message that this router transmits.
HDLC configuration	Information about the HDLC configuration. • **Policing bucket**—Configured state of the receiving policer. • **Shaping bucket**—Configured state of the transmitting shaper. • **Giant threshold**—Giant threshold programmed into the hardware. • **Runt threshold**—Runt threshold programmed into the hardware.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Packet Forwarding Engine configuration	Information about the configuration of the Packet Forwarding Engine: • **Destination slot**—FPC slot number. • **PLP byte**—Packet Level Protocol byte.
CoS information	Information about the CoS queue for the physical interface. • **CoS transmit queue**—Queue number and its associated user-configured forwarding class name. • **Bandwidth %**—Percentage of bandwidth allocated to the queue. • **Bandwidth bps**—Bandwidth allocated to the queue (in bps). • **Buffer %**—Percentage of buffer space allocated to the queue. • **Buffer usec**—Amount of buffer space allocated to the queue, in microseconds. This value is nonzero only if the buffer size is configured in terms of time. • **Priority**—Queue priority: **low** or **high**. • **Limit**—Displayed if rate limiting is configured for the queue. Possible values are **none** and **exact**. If **exact** is configured, the queue transmits only up to the configured bandwidth, even if excess bandwidth is available. If **none** is configured, the queue transmits beyond the configured bandwidth if bandwidth is available.
Forwarding classes	Total number of forwarding classes supported on the specified interface.
Egress queues	Total number of egress Maximum usable queues on the specified interface.
Queue	Queue number.
Forwarding classes	Forwarding class name.
Queued Packets	Number of packets queued to this queue.
Queued Bytes	Number of bytes queued to this queue. The byte counts vary by PIC type.
Transmitted Packets	Number of packets transmitted by this queue. When fragmentation occurs on the egress interface, the first set of packet counters shows the postfragmentation values. The second set of packet counters (displayed under the **Packet Forwarding Engine Chassis Queues** field) shows the prefragmentation values.
Transmitted Bytes	Number of bytes transmitted by this queue. The byte counts vary by PIC type.
Tail-dropped packets	Number of packets dropped because of tail drop.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description

RED-dropped packets	Number of packets dropped because of random early detection (RED).
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, the total number of dropped packets is displayed. On all other M Series routers, the output classifies dropped packets into the following categories:
	  • **Low, non-TCP**—Number of low-loss priority non-TCP packets dropped because of RED.
	  • **Low, TCP**—Number of low-loss priority TCP packets dropped because of RED.
	  • **High, non-TCP**—Number of high-loss priority non-TCP packets dropped because of RED.
	  • **High, TCP**—Number of high-loss priority TCP packets dropped because of RED.
	• (MX Series routers with enhanced DPCs, and T Series routers with enhanced FPCs only) The output classifies dropped packets into the following categories:
	  • **Low**—Number of low-loss priority packets dropped because of RED.
	  • **Medium-low**—Number of medium-low loss priority packets dropped because of RED.
	  • **Medium-high**—Number of medium-high loss priority packets dropped because of RED.
	  • **High**—Number of high-loss priority packets dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

RED-dropped bytes	Number of bytes dropped because of RED. The byte counts vary by PIC type.
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, only the total number of dropped bytes is displayed. On all other M Series routers, the output classifies dropped bytes into the following categories:
	  • **Low, non-TCP**—Number of low-loss priority non-TCP bytes dropped because of RED.
	  • **Low, TCP**—Number of low-loss priority TCP bytes dropped because of RED.
	  • **High, non-TCP**—Number of high-loss priority non-TCP bytes dropped because of RED.
	  • **High, TCP**—Number of high-loss priority TCP bytes dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

Transmit rate	Configured transmit rate of the scheduler. The rate is a percentage of the total interface bandwidth.
Rate Limit	Rate limiting configuration of the queue. Possible values are : • **None**—No rate limit. • **exact**—Queue transmits at the configured rate.
Buffer size	Delay buffer size in the queue.
Priority	Scheduling priority configured as **low** or **high**.
Excess Priority	Priority of the excess bandwidth traffic on a scheduler: **low**, **medium-low**, **medium-high**, **high**, or **none**.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Drop profiles	Display the assignment of drop profiles. • **Loss priority**—Packet loss priority for drop profile assignment. • **Protocol**—Transport protocol for drop profile assignment. • **Index**—Index of the indicated object. Objects that have indexes in this output include schedulers and drop profiles. • **Name**—Name of the drop profile. • **Type**—Type of the drop profile: **discrete** or **interpolated**. • **Fill Level**—Percentage fullness of a queue. • **Drop probability**—Drop probability at this fill level.
Excess Priority	Priority of the excess bandwidth traffic on a scheduler.
Drop profiles	Display the assignment of drop profiles. • **Loss priority**—Packet loss priority for drop profile assignment. • **Protocol**—Transport protocol for drop profile assignment. • **Index**—Index of the indicated object. Objects that have indexes in this output include schedulers and drop profiles. • **Name**—Name of the drop profile. • **Type**—Type of the drop profile: **discrete** or **interpolated**. • **Fill Level**—Percentage fullness of a queue. • **Drop probability**—Drop probability at this fill level.

Table 57: show class-of-service interface Output Fields (continued)

Field Name	Field Description

Adjustment information	Display the assignment of shaping-rate adjustments on a scheduler node or queue.
	• **Adjusting application**—Application that is performing the shaping-rate adjustment.
	  ◦ The adjusting application can appear as **ancp LS-0**, which is the Junos OS Access Node Control Profile process (**ancpd**) that performs shaping-rate adjustments on schedule nodes.
	  ◦ The adjusting application can appear as **DHCP**, which adjusts the shaping-rate and overhead-accounting class-of-service attributes based on DHCP option 82, suboption 9 (Vendor Specific Information). The shaping rate is based on the actual-data-rate-downstream attribute. The overhead accounting value is based on the access-loop-encapsulation attribute and specifies whether the access loop uses Ethernet (frame mode) or ATM (cell mode).
	  ◦ The adjusting application can also appear as **pppoe**, which adjusts the shaping-rate and overhead-accounting class-of-service attributes on dynamic subscriber interfaces in a broadband access network based on access line parameters in Point-to-Point Protocol over Ethernet (PPPoE) Tags [TR-101]. This feature is supported on MPC/MIC interfaces on MX Series routers. The shaping rate is based on the actual-data-rate-downstream attribute. The overhead accounting value is based on the access-loop-encapsulation attribute and specifies whether the access loop uses Ethernet (frame mode) or ATM (cell mode).
	• **Adjustment type**—Type of adjustment: **absolute** or **delta**.
	• **Configured shaping rate**—Shaping rate configured for the scheduler node or queue.
	• **Adjustment value**—Value of adjusted shaping rate.
	• **Adjustment target**—Level of shaping-rate adjustment performed: **node** or **queue**.
	• **Adjustment overhead-accounting mode**—Configured shaping mode: **frame** or **cell**.
	• **Adjustment overhead bytes**—Number of bytes that the ANCP agent adds to or subtracts from the actual downstream frame overhead before reporting the adjusted values to CoS.
	• **Adjustment target**—Level of shaping-rate adjustment performed: **node** or **queue**.
	• **Adjustment multicast index**—

Sample Output

show class-of-service interface (Physical)

```
user@host> show class-of-service interface so-0/2/3
Physical interface: so-0/2/3, Index: 135
Maximum usable queues: 8, Queues in use: 4
Total non-default queues created: 4
  Scheduler map: <default>, Index: 2032638653

  Logical interface: fe-0/0/1.0, Index: 68, Dedicated Queues: no
    Shaping rate: 32000
    Object                  Name                  Type          Index
    Scheduler-map           <default>                            27
    Rewrite                 exp-default           exp            21
    Classifier              exp-default           exp            5
    Classifier              ipprec-compatibility  ip             8
    Forwarding-class-map    exp-default           exp            5
```

show class-of-service interface (Logical)

```
user@host> show class-of-service interface so-0/2/3.0
Logical interface: so-0/2/3.0, Index: 68, Dedicated Queues: no
    Shaping rate: 32000
```

```
    Object                 Name                  Type             Index
    Scheduler-map          <default>                               27
    Rewrite                exp-default           exp               21
    Classifier             exp-default           exp               5
    Classifier             ipprec-compatibility  ip                8
    Forwarding-class-map   exp-default           exp               5
```

show class-of-service interface (Gigabit Ethernet)

```
user@host> show class-of-service interface ge-6/2/0
Physical interface: ge-6/2/0, Index: 175
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Input scheduler map: <default>, Index: 3
  Chassis scheduler map: <default-chassis>, Index: 4
```

show class-of-service interface (ANCP)

```
user@host> show class-of-service interface pp0.1073741842
  Logical interface: pp0.1073741842, Index: 341
Object                 Name                  Type                   Index
Traffic-control-profile TCP-CVLAN            Output                 12408
Classifier             dscp-ipv6-compatibility dscp-ipv6                9
Classifier             ipprec-compatibility   ip                       13

    Adjusting application: ancp LS-0
      Adjustment type: absolute
      Configured shaping rate: 4000000
      Adjustment value: 11228000
      Adjustment overhead-accounting mode: Frame Mode
      Adjustment overhead bytes: 50
      Adjustment target: node
```

show class-of-service interface (PPPoE Interface)

```
user@host> show class-of-service interface pp0.1
Logical interface: pp0.1, Index: 85
    Object                 Name                  Type           Index
    Traffic-control-profile tcp-pppoe.o.pp0.1    Output         2726446535
    Classifier             ipprec-compatibility  ip             13

    Adjusting application: PPPoE
      Adjustment type: absolute
      Adjustment value: 5000000
      Adjustment overhead-accounting mode: cell
      Adjustment target: node
```

show class-of-service interface (DHCP Interface)

```
user@host> show class-of-service interface demux0.1
Logical interface: pp0.1, Index: 85
    Object                 Name                  Type           Index
    Traffic-control-profile tcp-dhcp.o.demux0.1  Output         2726446535
    Classifier             ipprec-compatibility  ip             13

    Adjusting application: DHCP
      Adjustment type: absolute
      Adjustment value: 5000000
      Adjustment overhead-accounting mode: cell
      Adjustment target: node
```

show class-of-service interface (T4000 Routers with Type 5 FPCs)

```
user@host> show class-of-service interface xe-4/0/0
Physical interface: xe-4/0/0, Index: 153
    Maximum usable queues: 8, Queues in use: 4
      Shaping rate: 5000000000 bps
      Scheduler map: <default>, Index: 2
      Congestion-notification: Disabled

      Logical interface: xe-4/0/0.0, Index: 77
        Object                  Name                   Type
Index
        Classifier              ipprec-compatibility   ip
  13
```

show class-of-service interface detail

```
user@host> show class-of-service interface ge-0/3/0 detail

Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Link-level type: Ethernet, MTU: 1518, Speed: 1000mbps, Loopback: Disabled,
Source filtering: Disabled, Flow control: Enabled, Auto-negotiation: Enabled,
Remote fault: Online
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000

  Physical interface: ge-0/3/0, Index: 138
  Maximum usable queues: 4, Queues in use: 5
  Shaping rate: 50000 bps
  Scheduler map: interface-schedular-map, Index: 58414
  Input shaping rate: 10000 bps
  Input scheduler map: schedular-map, Index: 15103
  Chassis scheduler map: <default-chassis>, Index: 4
  Congestion-notification: Disabled

  Logical interface ge-0/3/0.0
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    inet
    mpls
Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.0      up    up   inet
                           mpls
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.0      up    up   inet
                           mpls

  Logical interface: ge-0/3/0.0, Index: 68
    Object                  Name                   Type                Index
    Rewrite                 exp-default            exp (mpls-any)        33
    Classifier              exp-default            exp                   10
    Classifier              ipprec-compatibility   ip                    13


  Logical interface ge-0/3/0.1
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    inet
Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.1      up    up   inet
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.1      up    up   inet
```

```
  Logical interface: ge-0/3/0.1, Index: 69
    Object                  Name                   Type              Index
    Classifier              ipprec-compatibility   ip                 13
```

show class-of-service interface comprehensive

```
user@host> show class-of-service interface ge-0/3/0 comprehensive
Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Interface index: 138, SNMP ifIndex: 601, Generation: 141
  Link-level type: Ethernet, MTU: 1518, Speed: 1000mbps, BPDU Error: None,
MAC-REWRITE Error: None, Loopback: Disabled, Source filtering: Disabled, Flow
control: Enabled,
  Auto-negotiation: Enabled, Remote fault: Online
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000
  CoS queues     : 4 supported, 4 maximum usable queues
  Schedulers     : 256
  Hold-times     : Up 0 ms, Down 0 ms
  Current address: 00:14:f6:f4:b4:5d, Hardware address: 00:14:f6:f4:b4:5d
  Last flapped   : 2010-09-07 06:35:22 PDT (15:14:42 ago)
  Statistics last cleared: Never  Exclude Overhead Bytes: Disabled
  Traffic statistics:
   Input  bytes  :                    0                    0 bps
   Output bytes  :                    0                    0 bps
   Input  packets:                    0                    0 pps
   Output packets:                    0                    0 pps
   IPv6 total statistics:
    Input  bytes  :                   0
    Output bytes  :                   0
    Input  packets:                   0
    Output packets:                   0
  Ingress traffic statistics at Packet Forwarding Engine:
   Input  bytes  :                    0                    0 bps
   Input  packets:                    0                    0 pps
   Drop   bytes  :                    0                    0 bps
   Drop   packets:                    0                    0 pps
  Label-switched interface (LSI) traffic statistics:
   Input  bytes  :                    0                    0 bps
   Input  packets:                    0                    0 pps
  Input errors:
    Errors: 0, Drops: 0, Framing errors: 0, Runts: 0, Policed discards: 0, L3
incompletes: 0, L2 channel errors: 0, L2 mismatch timeouts: 0, FIFO errors: 0,
Resource errors: 0
  Output errors:
    Carrier transitions: 5, Errors: 0, Drops: 0, Collisions: 0, Aged packets: 0,
 FIFO errors: 0, HS link CRC errors: 0, MTU errors: 0, Resource errors: 0
  Ingress queues: 4 supported, 5 in use
  Queue counters:       Queued packets  Transmitted packets      Dropped packets

    0 af3                            0                    0                    0

    1 af2                            0                    0                    0

    2 ef2                            0                    0                    0

    3 ef1                            0                    0                    0

  Egress queues: 4 supported, 5 in use
  Queue counters:       Queued packets  Transmitted packets      Dropped packets
```



```
    0 af3                            0                    0                    0

    1 af2                            0                    0                    0

    2 ef2                            0                    0                    0

    3 ef1                            0                    0                    0

  Active alarms  : None
  Active defects : None
  MAC statistics:                      Receive         Transmit
    Total octets                             0                0
    Total packets                            0                0
    Unicast packets                          0                0
    Broadcast packets                        0                0
    Multicast packets                        0                0
    CRC/Align errors                         0                0
    FIFO errors                              0                0
    MAC control frames                       0                0
    MAC pause frames                         0                0
    Oversized frames                         0
    Jabber frames                            0
    Fragment frames                          0
    VLAN tagged frames                       0
    Code violations                          0
  Filter statistics:
    Input packet count                       0
    Input packet rejects                     0
    Input DA rejects                         0
    Input SA rejects                         0
    Output packet count                                       0
    Output packet pad count                                   0
    Output packet error count                                 0
    CAM destination filters: 0, CAM source filters: 0
  Autonegotiation information:
    Negotiation status: Complete
    Link partner:
       Link mode: Full-duplex, Flow control: Symmetric/Asymmetric, Remote fault:
 OK
    Local resolution:
      Flow control: Symmetric, Remote fault: Link OK
  Packet Forwarding Engine configuration:
    Destination slot: 0
  CoS information:
    Direction : Output
    CoS transmit queue               Bandwidth               Buffer Priority
Limit
                              %            bps     %            usec
    2 ef2                    39          19500     0             120     high
none
    Direction : Input
    CoS transmit queue               Bandwidth               Buffer Priority
Limit
                              %            bps     %            usec
    0 af3                    30           3000    45               0      low
none

Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Interface index: 138, SNMP ifIndex: 601
Forwarding classes: 16 supported, 5 in use
Ingress queues: 4 supported, 5 in use
```



```
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Forwarding classes: 16 supported, 5 in use
Egress queues: 4 supported, 5 in use
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                     0                     0 pps
    RL-dropped bytes     :                     0                     0 bps
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
```



```
      Tail-dropped packets : Not Available
      RL-dropped packets   :                    0                    0 pps
      RL-dropped bytes     :                    0                    0 bps
      RED-dropped packets  :                    0                    0 pps
      RED-dropped bytes    :                    0                    0 bps
  Queue: 2, Forwarding classes: ef2
    Queued:
      Packets              :                    0                    0 pps
      Bytes                :                    0                    0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                    0 bps
      Tail-dropped packets : Not Available
      RL-dropped packets   :                    0                    0 pps
      RL-dropped bytes     :                    0                    0 bps
      RED-dropped packets  :                    0                    0 pps
      RED-dropped bytes    :                    0                    0 bps
  Queue: 3, Forwarding classes: ef1
    Queued:
      Packets              :                    0                    0 pps
      Bytes                :                    0                    0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                    0 bps
      Tail-dropped packets : Not Available
      RL-dropped packets   :                    0                    0 pps
      RL-dropped bytes     :                    0                    0 bps
      RED-dropped packets  :                    0                    0 pps
      RED-dropped bytes    :                    0                    0 bps

  Packet Forwarding Engine Chassis Queues:
  Queues: 4 supported, 5 in use
  Queue: 0, Forwarding classes: af3
    Queued:
      Packets              :                    0                    0 pps
      Bytes                :                    0                    0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                    0 bps
      Tail-dropped packets :                    0                    0 pps
      RED-dropped packets  : Not Available
      RED-dropped bytes    : Not Available
  Queue: 1, Forwarding classes: af2
    Queued:
      Packets              :                    0                    0 pps
      Bytes                :                    0                    0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                    0 bps
      Tail-dropped packets :                    0                    0 pps
      RED-dropped packets  : Not Available
      RED-dropped bytes    : Not Available
  Queue: 2, Forwarding classes: ef2
    Queued:
      Packets              :                    0                    0 pps
      Bytes                :                    0                    0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                    0 bps
      Tail-dropped packets :                    0                    0 pps
      RED-dropped packets  : Not Available
```

```
    RED-dropped bytes    : Not Available
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :               108546                     0 pps
    Bytes                :             12754752                   376 bps
  Transmitted:
    Packets              :               108546                      0 pps
    Bytes                :             12754752                   376 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available

Physical interface: ge-0/3/0, Index: 138
Maximum usable queues: 4, Queues in use: 5
  Shaping rate: 50000 bps

Scheduler map: interface-schedular-map, Index: 58414

  Scheduler: ef2, Forwarding class: ef2, Index: 39155
    Transmit rate: 39 percent, Rate Limit: none, Buffer size: 120 us, Buffer
Limit: none, Priority: high
    Excess Priority: unspecified
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
  Input shaping rate: 10000 bps
  Input scheduler map: schedular-map

Scheduler map: schedular-map, Index: 15103

  Scheduler: af3, Forwarding class: af3, Index: 35058
    Transmit rate: 30 percent, Rate Limit: none, Buffer size: 45 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: unspecified
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         40582    green
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any         18928    yellow
Drop profile: green, Type: discrete, Index: 40582
  Fill level    Drop probability
          50                   0
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
```



```
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: yellow, Type: discrete, Index: 18928
  Fill level    Drop probability
         50                   0
        100                 100
  Chassis scheduler map: < default-drop-profile>
Scheduler map: < default-drop-profile>, Index: 4

  Scheduler: < default-drop-profile>, Forwarding class: af3, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100

  Scheduler: < default-drop-profile>, Forwarding class: af2, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                 100

  Scheduler: < default-drop-profile>, Forwarding class: ef2, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
```

```
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100

  Scheduler: < default-drop-profile>, Forwarding class: ef1, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: , Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
  Congestion-notification: Disabled
Forwarding class                       ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                   0       0          0             low
            normal
  af2                                   1       1          1             low
            normal
  ef2                                   2       2          2             high
            normal
  ef1                                   3       3          3             high
            normal
  af1                                   4       4          0             low
            normal

  Logical interface ge-0/3/0.0 (Index 68) (SNMP ifIndex 152) (Generation 159)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    Traffic statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
```



```
     Output packets:                   0
    Local statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
     Output packets:                   0
    Transit statistics:
     Input  bytes  :                   0                    0 bps
     Output bytes  :                   0                    0 bps
     Input  packets:                   0                    0 pps
     Output packets:                   0                    0 pps
    Protocol inet, MTU: 1500, Generation: 172, Route table: 0
      Flags: Sendbcast-pkt-to-re
      Input Filters: filter-in-ge-0/3/0.0-i,
      Policer: Input: p1-ge-0/3/0.0-inet-i
    Protocol mpls, MTU: 1488, Maximum labels: 3, Generation: 173, Route table: 0

      Flags: Is-Primary
      Output Filters: exp-filter,,,,,

Logical interface ge-1/2/0.0 (Index 347) (SNMP ifIndex 638) (Generation 156)

Forwarding class ID  Queue  Restricted queue  Fabric priority  Policing priority
   SPU priority
best-effort      0   0      0                 low              normal
    low

Aggregate Forwarding-class statistics per forwarding-class
  Aggregate Forwarding-class statistics:
   Forwarding-class statistics:

Forwarding-class best-effort statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

IPv4 protocol forwarding-class statistics:
  Forwarding-class statistics:
  Forwarding-class best-effort statistics:

    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0
```

```
    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

    IPv6  protocol forwarding-class statistics:
    Forwarding-class statistics:
     Forwarding-class best-effort statistics:

    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Logical interface ge-0/3/0.0 (Index 68) (SNMP ifIndex 152)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    Input packets : 0
    Output packets: 0

Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.0      up    up   inet  filter-in-ge-0/3/0.0-i
                           mpls                       exp-filter
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.0      up    up
                           inet  p1-ge-0/3/0.0-inet-i
                           mpls

Filter: filter-in-ge-0/3/0.0-i
Counters:
Name                                                Bytes              Packets
count-filter-in-ge-0/3/0.0-i                            0                    0

Filter: exp-filter
```



```
Counters:
Name                                                Bytes              Packets
count-exp-seven-match                                   0                    0
count-exp-zero-match                                    0                    0
Policers:
Name                                              Packets
p1-ge-0/3/0.0-inet-i                                    0


  Logical interface: ge-0/3/0.0, Index: 68
    Object                 Name                   Type                     Index

    Rewrite                exp-default            exp (mpls-any)              33

Rewrite rule: exp-default, Code point type: exp, Index: 33
  Forwarding class                    Loss priority       Code point
  af3                                 low                 000
  af3                                 high                001
  af2                                 low                 010
  af2                                 high                011
  ef2                                 low                 100
  ef2                                 high                101
  ef1                                 low                 110
  ef1                                 high                111
    Object                 Name                   Type                     Index

    Classifier             exp-default            exp                         10

Classifier: exp-default, Code point type: exp, Index: 10
  Code point         Forwarding class                    Loss priority
  000                af3                                 low
  001                af3                                 high
  010                af2                                 low
  011                af2                                 high
  100                ef2                                 low
  101                ef2                                 high
  110                ef1                                 low
  111                ef1                                 high
    Object                 Name                   Type                     Index

    Classifier             ipprec-compatibility   ip                          13

Classifier: ipprec-compatibility, Code point type: inet-precedence, Index: 13
  Code point         Forwarding class                    Loss priority
  000                af3                                 low
  001                af3                                 high
  010                af3                                 low
  011                af3                                 high
  100                af3                                 low
  101                af3                                 high
  110                ef1                                 low
  111                ef1                                 high
Forwarding class                    ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                0       0          0             low
          normal
  af2                                1       1          1             low
          normal
  ef2                                2       2          2             high
          normal
  ef1                                3       3          3             high
```

```
          normal
  af1                                  4       4          0             low
          normal

  Logical interface ge-0/3/0.1 (Index 69) (SNMP ifIndex 154) (Generation 160)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    Traffic statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
     Output packets:                   0
    Local statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
     Output packets:                   0
    Transit statistics:
     Input  bytes  :                   0                    0 bps
     Output bytes  :                   0                    0 bps
     Input  packets:                   0                    0 pps
     Output packets:                   0                    0 pps
    Protocol inet, MTU: 1500, Generation: 174, Route table: 0
      Flags: Sendbcast-pkt-to-re

  Logical interface ge-0/3/0.1 (Index 69) (SNMP ifIndex 154)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    Input packets : 0
    Output packets: 0


Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.1      up    up   mpls
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.1      up    up
                           mpls

  Logical interface: ge-0/3/0.1, Index: 69
    Object                 Name                   Type                    Index

    Classifier             ipprec-compatibility   ip                         13

Classifier: ipprec-compatibility, Code point type: inet-precedence, Index: 13
  Code point         Forwarding class                    Loss priority
  000                af3                                 low
  001                af3                                 high
  010                af3                                 low
  011                af3                                 high
  100                af3                                 low
  101                af3                                 high
  110                ef1                                 low
  111                ef1                                 high
Forwarding class                       ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                  0       0          0             low
          normal
  af2                                  1       1          1             low
          normal
  ef2                                  2       2          2             high
          normal
  ef1                                  3       3          3             high
          normal
```



```
  af1                                  4       4          0              low
            normal
```

show class-of-service interface (ACX Series Routers)

```
user@host-g11# show class-of-service interface
Physical interface: at-0/0/0, Index: 130
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

  Logical interface: at-0/0/0.0, Index: 69

  Logical interface: at-0/0/0.32767, Index: 70

Physical interface: at-0/0/1, Index: 133
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

  Logical interface: at-0/0/1.0, Index: 71

  Logical interface: at-0/0/1.32767, Index: 72

Physical interface: ge-0/1/0, Index: 146
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Rewrite                 dscp-default          dscp                       31
Classifier              d1                    dscp                    11331
Classifier              ci                    ieee8021p                 583

  Logical interface: ge-0/1/0.0, Index: 73
Object                  Name                  Type                    Index
Rewrite                 custom-exp            exp (mpls-any)          46413

  Logical interface: ge-0/1/0.1, Index: 74

  Logical interface: ge-0/1/0.32767, Index: 75

Physical interface: ge-0/1/1, Index: 147
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Classifier              ipprec-compatibility  ip                         13

  Logical interface: ge-0/1/1.0, Index: 76

Physical interface: ge-0/1/2, Index: 148
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Rewrite                 ri                    ieee8021p (outer)       35392
Classifier              ci                    ieee8021p                 583

Physical interface: ge-0/1/3, Index: 149
```



```
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

  Logical interface: ge-0/1/3.0, Index: 77
Object                 Name                   Type                    Index
Rewrite                custom-exp2            exp (mpls-any)          53581

Physical interface: ge-0/1/4, Index: 150
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/5, Index: 151
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/6, Index: 152
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/7, Index: 153
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             d1                     dscp                    11331

Physical interface: ge-0/2/0, Index: 154
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/2/1, Index: 155
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

  Logical interface: ge-0/2/1.0, Index: 78

  Logical interface: ge-0/2/1.32767, Index: 79

Physical interface: xe-0/3/0, Index: 156
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
```



```
Classifier             ipprec-compatibility   ip                      13

  Logical interface: xe-0/3/0.0, Index: 80

Physical interface: xe-0/3/1, Index: 157
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                      13

  Logical interface: xe-0/3/1.0, Index: 81

[edit]
user@host-g11#
```

show class-of-service interface (PPPoE Subscriber Interface for Enhanced Subscriber Management)

```
user@host> show class-of-service interface pp0.3221225474
  Logical interface: pp0.3221225475, Index: 3221225475
Object                 Name                   Type                    Index
Traffic-control-profile TC_PROF_100_199_SERIES_UID1006 Output     4294967312
Scheduler-map          SMAP-1_UID1002         Output                4294967327
Rewrite-Output         ieee-rewrite           ieee8021p               60432
Rewrite-Output         rule1                  ip                      50463

    Adjusting application: PPPoE IA tags
      Adjustment type: absolute
      Configured shaping rate: 11000000
      Adjustment value: 5000000
      Adjustment target: node

    Adjusting application: ucac
      Adjustment type: delta
      Configured shaping rate: 5000000
      Adjustment value: 100000
      Adjustment target: node
```


show class-of-service multi-destination

Syntax show class-of-service multi-destination

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
Command introduced in Junos OS Release 17.1 for the EX4300 switch.

Description For each class-of-service (CoS) multidestination classifier, display the classifier type.

Options **none**—Display all multidestination classifiers.

Required Privilege Level view

Related Documentation

- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Example: Configuring Multidestination (Multicast, Broadcast, DLF) Classifiers*
- *Understanding CoS Classifiers*
- Understanding CoS Classifiers on page 123
- *Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces*
- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162

Output Fields Table 58 on page 260 describes the output fields for the **show class-of-service multi-destination** command. Output fields are listed in the approximate order in which they appear.

Table 58: show class-of-service multi-destination Output Fields

Field Name	Field Description
Family ethernet	Family to which the classifier belongs.
Classifier Name	Name of the classifier.
Classifier Type	Type of the classifier: **dscp** or **ieee-802.1**.
Classifier Index	Internal index of the classifier.

Sample Output

show class-of-service multi-destination

```
user@switch> show class-of-service multi-destination
```

```
Family ethernet:
Classifier Name               Classifier Type               Classifier Index
  ba-mcast-classifier           ieee-802.1                    62376
```

show class-of-service rewrite-rule

Syntax

show class-of-service rewrite-rule
<name *name*>
<type *type*>

Release Information Command introduced before Junos OS Release 7.4.

Description Display the mapping of forwarding classes and loss priority to code point values.

Options

none—Display all rewrite rules.

name ***name***—(Optional) Display the specified rewrite rule.

type ***type***—(Optional) Display the rewrite rule of the specified type. The rewrite rule type can be one of the following:

- **dscp**—For IPv4 traffic.
- **dscp-ipv6**—For IPv6 traffic.
- **exp**—For MPLS traffic.
- **frame-relay-de**—(SRX Series only) For Frame Relay traffic.
- **ieee-802.1**—For Layer 2 traffic.
- **inet-precedence**—For IPv4 traffic.

Required Privilege Level view

Related Documentation

- *Rewrite Rules Overview*

List of Sample Output show class-of-service rewrite-rule type dscp on page 263

Output Fields Table 59 on page 262 describes the output fields for the **show class-of-service rewrite-rule** command. Output fields are listed in the approximate order in which they appear.

Table 59: show class-of-service rewrite-rule Output Fields

Field Name	Field Description
Rewrite rule	Name of the rewrite rule.
Code point type	Type of rewrite rule: **dscp**, **dscp-ipv6**, **exp**, **frame-relay-de**, or **inet-precedence**.
Forwarding class	Classification of a packet affecting the forwarding, scheduling, and marking policies applied as the packet transits the router or switch.

Table 59: show class-of-service rewrite-rule Output Fields (continued)

Field Name	Field Description
Index	Internal index for this particular rewrite rule.
Loss priority	Loss priority for rewriting.
Code point	Code point value to rewrite.

Sample Output

show class-of-service rewrite-rule type dscp

```
user@host> show class-of-service rewrite-rule type dscp
Rewrite rule: dscp-default, Code point type: dscp
  Forwarding class                    Loss priority       Code point
  gold                                high                000000
  silver                              low                 110000
  silver                              high                111000
  bronze                              low                 001010
  bronze                              high                001100
  lead                                high                101110

Rewrite rule: abc-dscp-rewrite, Code point type: dscp, Index: 3245
Forwarding class                    Loss priority       Code point
  gold                                low                 000111
  gold                                high                001010
  silver                              low                 110000
  silver                              high                111000
  bronze                              high                001100
  lead                                low                 101110
  lead                                high                110111
```

PART 3

Scheduling Traffic

CHAPTER 7

Using Schedulers

- Troubleshooting Egress Bandwidth That Exceeds the Configured Maximum Bandwidth on page 396
- Troubleshooting Egress Queue Bandwidth Impacted by Congestion on page 397

Understanding Default CoS Scheduling and Classification

If you do not explicitly configure classifiers and apply them to interfaces, the switch uses the default classifier for ingress traffic. If you do not configure hierarchical scheduling (also known as enhanced transmission selection (ETS)) on an interface, the switch uses the default schedulers for egress traffic. Default classification maps all traffic into default forwarding classes (best-effort, fcoe, no-loss, network-control, and mcast).

Hierarchical scheduling groups IEEE 802.1p priorities (IEEE 802.1p code points, which classifiers map to forwarding classes, which in turn are mapped to output queues) into priority groups (forwarding class sets). If you use only the default traffic scheduling and classification, the switch automatically creates a default priority group that contains all of the priorities (which are mapped to forwarding classes and output queues), and assigns 100 percent of the port output bandwidth to that priority group. The forwarding classes (queues) in the default forwarding class set receive bandwidth based on the default classifier settings. The default priority group is transparent. It does not appear in the configuration (it is used for Data Center Bridging Capability Exchange (DCBX) protocol advertisement on QFX Series switches).

NOTE: If you explicitly configure one or more priority groups on an interface, any forwarding class that is not assigned to a priority group on that interface receives *no bandwidth*. This means that if you configure hierarchical scheduling on an interface, every forwarding class (priority) that you want to forward traffic on that interface must belong to a forwarding class set (priority group).

The following sections describe:

- Default Classification on page 268
- Default Scheduling on page 273
- Default Scheduling and Classification Summary on page 276

Default Classification

The default classifiers assign unicast and multicast best-effort and network-control ingress traffic to forwarding classes and loss priorities. The switch applies default unicast DSCP, unicast IEEE 802.1, and multidestination classifiers to each interface that does not have explicitly configured classifiers. If you explicitly configure one type of classifier but not other types of classifiers, the system uses only the configured classifier and does not use default classifiers for other types of traffic. There are two different default unicast IEEE 802.1 classifiers, a trusted classifier and an untrusted classifier. On QFX Series switches, the trusted classifier is the default IEEE classifier.

Table 26 on page 115 shows the default mapping of DSCP code-point values to unicast forwarding classes and loss priorities for DSCP IP and DCSP IPv6.

Table 60: Default DSCP IP and IPv6 Unicast Classifiers

Code Point	Forwarding Class	Loss Priority
000000 (be)	best-effort	low
000001	best-effort	low
000010	best-effort	low
000011	best-effort	low
000100	best-effort	low
000101	best-effort	low
000110	best-effort	low
000111	best-effort	low
001000 (cs1)	best-effort	low
001001	best-effort	low
001010 (af11)	best-effort	low
001011	best-effort	low
001100 (af12)	best-effort	low
001101	best-effort	low
001110 (af13)	best-effort	low
001111	best-effort	low
010000 (cs2)	best-effort	low
010001	best-effort	low
010010 (af21)	best-effort	low
010011	best-effort	low
010100 (af22)	best-effort	low
010101	best-effort	low
010110 (af23)	best-effort	low
010111	best-effort	low

Table 60: Default DSCP IP and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
011000 (cs3)	best-effort	low
011001	best-effort	low
011010 (af31)	best-effort	low
011011	best-effort	low
011100 (af32)	best-effort	low
011101	best-effort	low
011110 (af33)	best-effort	low
011111	best-effort	low
100000 (cs4)	best-effort	low
100001	best-effort	low
100010 (af41)	best-effort	low
100011	best-effort	low
100100 (af42)	best-effort	low
100101	best-effort	low
100110 (af43)	best-effort	low
100111	best-effort	low
101000 (cs5)	best-effort	low
101001	best-effort	low
101011	best-effort	low
101100	best-effort	low
101101	best-effort	low
101110 (ef)	best-effort	low
101111	best-effort	low
110000 (nc1)	network-control	low

Table 60: Default DSCP IP and IPv6 Unicast Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
110001	network-control	low
110010	network-control	low
110011	network-control	low
110100	network-control	low
110101	network-control	low
110110	network-control	low
110111	network-control	low
111000 (nc2)	network-control	low
111001	network-control	low
111010	network-control	low
111011	network-control	low
111100	network-control	low
111101	network-control	low
111110	network-control	low
111111	network-control	low

NOTE: There are no default DSCP IP or IPv6 multidestination classifiers for multidestination traffic. DSCP IPv6 multidestination classifiers are not supported for multidestination traffic.

Table 27 on page 117 shows the default trusted classifier mapping of IEEE 802.1 code-point values to unicast forwarding classes and loss priorities .

Table 61: Default IEEE 802.1 Unicast Classifiers (Trusted)

Code Point	Forwarding Class	Loss Priority
be (000)	best-effort	low
be1 (001)	best-effort	low

Table 61: Default IEEE 802.1 Unicast Classifiers (Trusted) (continued)

Code Point	Forwarding Class	Loss Priority
ef (010)	best-effort	low
ef1 (011)	fcoe	low
af11 (100)	no-loss	low
af12 (101)	best-effort	low
nc1 (110)	network-control	low
nc2 (111)	network-control	low

Table 28 on page 118 shows the default untrusted mapping of IEEE 802.1p code-point values to unicast forwarding classes and loss priorities.

Table 62: Default IEEE 802.1 Unicast Classifiers (Untrusted)

Code Point	Forwarding Class	Loss Priority
000	best-effort	low
001	best-effort	low
010	best-effort	low
011	best-effort	low
100	best-effort	low
101	best-effort	low
110	best-effort	low
111	best-effort	low

Table 29 on page 118 shows the default mapping of IEEE 802.1 code-point values to multidestination (multicast, broadcast, and destination lookup fail traffic) forwarding classes and loss priorities.

Table 63: Default IEEE 802.1 Multidestination Classifiers

Code Point	Forwarding Class	Loss Priority
be (000)	mcast	low
be1 (001)	mcast	low

Table 63: Default IEEE 802.1 Multidestination Classifiers (continued)

Code Point	Forwarding Class	Loss Priority
ef (010)	mcast	low
ef1 (011)	mcast	low
af11 (100)	mcast	low
af12 (101)	mcast	low
nc1 (110)	mcast	low
nc2 (111)	mcast	low

Default Scheduling

The default schedulers allocate egress bandwidth resources to unicast and multicast egress traffic as shown in Table 30 on page 119:

Table 64: Default Scheduler Configuration

Default Scheduler and Queue Number	Transmit Rate (Minimum Guaranteed Bandwidth)	Shaping Rate (Maximum Bandwidth)	Excess Bandwidth Sharing	Priority	Buffer Size
best-effort forwarding class scheduler (queue 0)	5%	None	5%	low	5%
fcoe forwarding class scheduler (queue 3) NOTE: Do not map traffic to the default fcoe forwarding class. The default fcoe forwarding class carries the no-loss packet drop attribute, which is not supported on OCX Series switches. Other switches use the fcoe forwarding class for Fibre Channel over Ethernet (FCoE) lossless Layer 2 transport, but OCX Series switches do not support FCoE or lossless Layer 2 transport. If you want to use queue 3, configure a new forwarding class *without* the no-loss packet drop attribute and map it to queue 3.	35%	None	35%	low	35%

Table 64: Default Scheduler Configuration (continued)

Default Scheduler and Queue Number	Transmit Rate (Minimum Guaranteed Bandwidth)	Shaping Rate (Maximum Bandwidth)	Excess Bandwidth Sharing	Priority	Buffer Size
no-loss forwarding class scheduler (queue 4) NOTE: Do not map traffic to the default no-loss forwarding class. The default no-loss forwarding class carries the no-loss packet drop attribute, which is not supported on OCX Series switches. Other switches use the no-loss forwarding class for lossless Layer 2 transport, but OCX Series switches do not support lossless Layer 2 transport. If you want to use queue 4, configure a new forwarding class *without* the no-loss packet drop attribute and map it to queue 4.	35%	None	35%	low	35%
network-control forwarding class scheduler (queue 7)	5%	None	5%	low	5%
mcast forwarding class scheduler (queue 8)	20%	None	20%	low	20%

NOTE: **The minimum guaranteed bandwidth rate also determines the amount of excess (extra) bandwidth that the queue can share. Extra bandwidth is allocated to queues in proportion to the minimum guaranteed bandwidth (transmit rate) of each queue.**

The default DSCP classifier maps traffic only to the best-effort (queue 0), network-control (queue 7), and mcast (queue 8) forwarding classes. Only the five default schedulers shown in Table 30 on page 119 have port resources (for example, bandwidth) mapped to them by default, but only the queues that are forwarding traffic use port resources. So even though 35 percent of port bandwidth is allocated to the fcoe and no-loss schedulers, that bandwidth is available to the best-effort, network-control, and mcast schedulers because no traffic is mapped by default to the fcoe and no-loss forwarding classes. The amount of default bandwidth each forwarding class receives on a port is proportional to the default scheduler transmit rate. Unused bandwidth is shared among queues that need more bandwidth. (You can configure schedulers and forwarding classes to allocate bandwidth to other queues or to change the default bandwidth of a default queue.) In addition, multidestination queue 11 receives enough bandwidth from the default multidestination scheduler to handle CPU-generated multidestination traffic.

Default hierarchical scheduling divides the total port bandwidth between two groups of traffic: unicast traffic and multidestination traffic. By default, unicast traffic consists of queue 0 (**best-effort** forwarding class) and queue 7 (**network-control** forwarding class) because no traffic is mapped by default to queue 3 (**fcoe** forwarding class) or queue 4 (**no-loss** forwarding class). Unicast traffic receives and shares a total of 80 percent of the port bandwidth. By default, multidestination traffic (**mcast** queue 8) receives a total of 20 percent of the port bandwidth. So on a 10-Gigabit port, unicast traffic receives 8-Gbps of bandwidth and multidestination traffic receives 2-Gbps of bandwidth.

NOTE: Multidestination queue 11 also receives a small amount of default bandwidth from the multidestination scheduler. CPU-generated multidestination traffic uses queue 11, so you might see a small number of packets egress from queue 11. In addition, in the unlikely case that firewall filter match conditions map multidestination traffic to a unicast forwarding class, that traffic uses queue 11.

Default scheduling uses weighted round-robin (WRR) scheduling. Each queue receives a portion (weight) of the total available interface bandwidth. The scheduling weight is based on the transmit rate of the default scheduler for that queue. For example, queue 7 receives a default scheduling weight of 5 percent of the available bandwidth, and queue 8 receives a default scheduling weight of 20 percent of the available bandwidth. Queues are mapped to forwarding classes, so forwarding classes receive the default bandwidth for the queues to which they are mapped.

You should explicitly map traffic to non-default (unconfigured) queues and create schedulers to allocate bandwidth to those queues if you want to use them to forward traffic. By default, unicast queues 1, 2, 5, and 6 are unconfigured, and multidestination queues 9, 10, and 11 are unconfigured.

NOTE: If you want to map traffic to queue 3 or to queue 4 by configuring a DSCP classifier and mapping DSCP code points to forwarding classes , then you must explicitly configure forwarding classes that do not have the no-loss packet attribute and map them to those queues. Do not map traffic to the default fcoe forwarding class (queue 3) or to the default no-loss forwarding class (queue 4). If you map traffic to queue 3 or queue 4, unless you configure a scheduler, that traffic uses the default scheduler and receives 35 percent of the port bandwidth by default.

Unconfigured queues have a default scheduling weight of 1 so that they can receive a small amount of bandwidth in case they need to forward traffic. (However, queue 11 can use more of the default multidestination scheduler bandwidth if necessary to handle CPU-generated multidestination traffic.)

NOTE: All four multidestination queues have a scheduling weight of 1. Because by default multidestination traffic goes to queue 8, queue 8 receives almost all of the multidestination bandwidth. (There is no traffic on queue 9 and queue 10, and very little traffic on queue 11, so there is almost no competition for multidestination bandwidth.)

However, if you explicitly configure queue 9, 10, or 11 (by mapping code points to the unconfigured multidestination forwarding classes using the multidestination classifier), the explicitly configured queues share the multidestination scheduler bandwidth equally with default queue 8, because all of the queues have the same scheduling weight (1). To ensure that multidestination bandwidth is allocated to each queue properly and that the bandwidth allocation to the default queue (8) is not reduced too much, we strongly recommend that you configure a scheduler if you explicitly classify traffic into queue 9, 10, or 11.

If you map traffic to an unconfigured queue, the queue receives only the amount of group bandwidth proportional to its default weight (1). The actual amount of bandwidth an unconfigured queue receives depends on how much bandwidth the other queues in the group are using.

If the other unicast queues use less than their allocated amount of bandwidth, the unconfigured queues can share the unused bandwidth. Sharing unused bandwidth is one of the key advantages of hierarchical port scheduling. Configured queues have higher priority for bandwidth than unconfigured queues, so if a configured queue needs more bandwidth, then less bandwidth is available for unconfigured queues. Unconfigured queues always receive a minimum amount of bandwidth based on their scheduling weight (1). If you map traffic to an unconfigured queue, to allocate bandwidth to that queue, configure a scheduler for the forwarding class that is mapped to the queue.

Default Scheduling and Classification Summary

If you do not configure hierarchical scheduling on an interface:

- Default classifiers classify ingress traffic. The default DSCP classifier classifies unicast traffic into two queues, queue 0 (best-effort) and queue 7 (network-control). By default, multidestination traffic is classified into queue 8 (mcast).
- Default schedulers schedule egress traffic.
- A single default priority group (fc-set) receives 100 percent of the port bandwidth. All priorities (forwarding classes) are assigned to the default priority group and receive bandwidth based on their default schedulers. The default priority group is generated automatically and is not user-configurable.

Related Documentation

- Understanding CoS Packet Flow on page 23
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Understanding Default CoS Settings on page 25

- Understanding Applying CoS Classifiers and Rewrite Rules to Interfaces on page 162
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- Example: Configuring Queue Schedulers on page 293

Understanding CoS Scheduling Behavior and Configuration Considerations

Many factors affect scheduling configuration and bandwidth requirements, including:

- When you configure bandwidth for a forwarding class (each forwarding class is mapped to a queue) or a forwarding class set (priority group), the switch considers only the data as the configured bandwidth. The switch does not account for the bandwidth consumed by the preamble and the interframe gap (IFG). Therefore, when you calculate and configure the bandwidth requirements for a forwarding class or for a forwarding class set, consider the preamble and the IFG as well as the data in the calculations.
- When you configure a forwarding class to carry traffic on the switch (instead of using only default forwarding classes), you must also define a scheduling policy for the user-configured forwarding class. Some switches support enhanced transmission selection (ETS) hierarchical port scheduling, some switches support direct port scheduling, and some switches support both methods of scheduling.

  For ETS hierarchical port scheduling, defining a hierarchical scheduling policy using ETS means:

  - Mapping a scheduler to the forwarding class in a scheduler map
  - Including the forwarding class in a forwarding class set
  - Associating the scheduler map with a traffic control profile
  - Attaching the traffic control profile to a forwarding class set and an interface

  On switches that support port scheduling, defining a scheduling policy means:

  - Mapping a scheduler to the forwarding class in a scheduler map.
  - Applying the scheduler map to one or more interfaces.
- On each physical interface, either all forwarding classes that are being used on the interface must have rewrite rules configured, or no forwarding classes that are being used on the interface can have rewrite rules configured. On any physical port, do not mix forwarding classes with rewrite rules and forwarding classes without rewrite rules.
- For packets that carry both an inner VLAN tag and an outer VLAN tag, rewrite rules rewrite only the outer VLAN tag.
- For ETS hierarchical port scheduling, configuring the minimum guaranteed bandwidth (**transmit-rate**) for a forwarding class does not work unless you also configure the minimum guaranteed bandwidth (**guaranteed-rate**) for the forwarding class set in the traffic control profile.

Additionally, the sum of the transmit rates of the forwarding classes in a forwarding class set should not exceed the guaranteed rate for the forwarding class set. (You cannot guarantee a minimum bandwidth for the queues that is greater than the minimum bandwidth guaranteed for the entire set of queues.) If you configure transmit rates whose sum exceeds the guaranteed rate of the forwarding class set, the commit check fails and the system rejects the configuration.

- For ETS hierarchical port scheduling, the sum of the forwarding class set guaranteed rates cannot exceed the total port bandwidth. If you configure guaranteed rates whose sum exceeds the port bandwidth, the system sends a syslog message to notify you that the configuration is not valid. However, the system does not perform a commit check. If you commit a configuration in which the sum of the guaranteed rates exceeds the port bandwidth, the hierarchical scheduler behaves unpredictably.
- For ETS hierarchical port scheduling, if you configure the **guaranteed-rate** of a forwarding class set as a percentage, configure all of the transmit rates associated with that forwarding class set as percentages. In this case, if any of the transmit rates are configured as absolute values instead of percentages, the configuration is not valid and the system sends a syslog message.
- There are several factors to consider if you want to configure a strict-high priority queue (forwarding class):
  - On QFX5200, QFX3500, and QFX3600 switches and on QFabric systems, you can configure only one strict-high priority queue (forwarding class).

    On QFX5100 and EX4600 switches, you can configure only one forwarding-class-set (priority group) as strict-high priority. All queues which are part of that strict-high forwarding class set then act as strict-high queues.

    On QFX10000 switches, there is no limit to the number of strict-high priority queues you can configure.
  - You cannot configure a minimum guaranteed bandwidth (**transmit-rate**) for a strict-high priority queue on QFX5200, QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems.

    On QFX5200 and QFX10000 switches, you can set the **transmit-rate** on strict-high priority queues to set a limit on the amount of traffic that the queue treats as strict-high priority traffic. Traffic in excess of the **transmit-rate** is treated as best-effort traffic, and receives an excess bandwidth sharing weight of “1”, which is the proportion of extra bandwidth the strict-high priority queue can share on the port. Queues that are not strict-high priority queues use the transmit rate (default) or the configured excess rate to determine the proportion (weight) of extra port bandwidth the queue can share. However, you cannot configure an excess rate on a strict-high priority queue, and you cannot change the excess bandwidth sharing weight of “1” on a strict-high priority queue.

    For ETS hierarchical port scheduling, you cannot configure a minimum guaranteed bandwidth (**guaranteed-rate**) for a forwarding class set that includes a strict-high priority queue.
  - Except on QFX10000 switches, for ETS hierarchical port scheduling only, you must create a separate forwarding class set for a strict-high priority queue. On QFX10000

switches, you can mix strict-high priority and low priority queues in the same forwarding class set.

- Except on QFX10000 switches, for ETS hierarchical port scheduling, only one forwarding class set can contain a strict-high priority queue. On QFX10000 switches, this restriction does not apply.
- Except on QFX10000 switches, for ETS hierarchical port scheduling, a strict-high priority queue cannot belong to the same forwarding class set as queues that are not strict-high priority. (You cannot mix a strict-high priority forwarding class with forwarding classes that are not strict-high priority in one forwarding class set.) On QFX10000 switches, you can mix strict-high priority and low priority queues in the same forwarding class set.
- For ETS hierarchical port scheduling on switches that use different forwarding class sets for unicast and multidestination (multicast, broadcast, and destination lookup fail) traffic, a strict-high priority queue cannot belong to a multidestination forwarding class set.
- On QFX10000 systems, we recommend that you always configure a transmit rate on strict-high priority queues to prevent them from starving other queues. If you do not apply a transmit rate to limit the amount of bandwidth strict-high priority queues can use, then strict-high priority queues can use all of the available port bandwidth and starve other queues on the port.

  On QFX5200, QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, we recommend that you always apply a shaping rate to the strict-high priority queue to prevent it from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

- On QFabric systems, if any queue that contains outgoing packets does not transmit packets for 12 consecutive seconds, the port automatically resets. Failure of a queue to transmit packets for 12 consecutive seconds might be due to:
  - A strict-high priority queue consuming all of the port bandwidth
  - Several queues consuming all of the port bandwidth
  - Any queue or port receiving continuous priority-based flow control (PFC) or 802.3x Ethernet PAUSE messages (received PFC and PAUSE messages prevent a queue or a port, respectively, from transmitting packets because of network congestion)
  - Other conditions that prevent a queue from obtaining port bandwidth for 12 consecutive seconds

  If the cause is a strict-high priority queue consuming all of the port bandwidth, use rate shaping to configure a maximum rate for the strict-high priority queue and prevent it from using all of the port bandwidth. To configure rate shaping, include the **shaping-rate (*rate* | percent *percentage*)** statement at the **[edit class-of-service schedulers *scheduler-name*]** hierarchy level and apply the shaping rate to the strict-high priority scheduler. We recommend that you always apply a shaping rate to strict-high priority traffic to prevent the strict-high priority queue from starving other queues.

If several queues consume all of the port bandwidth, you can use a scheduler to rate shape those queues and prevent them from using all of the port bandwidth.

- For transmit rates below 1 Gbps, we recommend that you configure the transmit rate as a percentage instead of as a fixed rate. This is because the system converts fixed rates into percentages and might round small fixed rates to a lower percentage. For example, a fixed rate of 350 Mbps is rounded down to 3 percent instead of 3.5 percent.
- When you set the maximum bandwidth for a queue or for a priority group (**shaping-rate**) at 100 Kbps or lower, the traffic shaping behavior is accurate only within +/– 20 percent of the configured **shaping-rate**.
- On QFX10000 switches, configuring rate shaping (**[set class-of-service schedulers *scheduler-name* transmit-rate *(rate | percentage)* exact**) on a LAG interface using the **[edit class-of-service interfaces *lag-interface-name* scheduler-map *scheduler-map-name***] statement can result in scheduled traffic streams receiving more LAG link bandwidth than expected.

  You configure rate shaping in a scheduler to set the maximum bandwidth for traffic assigned to a forwarding class on a particular output queue on a port. For example, you can use a scheduler to configure rate shaping on traffic assigned to the best-effort forwarding class mapped to queue 0, and then apply the scheduler to an interface using a scheduler map, to set the maximum bandwidth for best-effort traffic mapped to queue 0 on that port. Traffic in the best-effort forwarding can use no more than the amount of port bandwidth specified by the transmit rate when you use the **exact** option.

  LAG interfaces are composed of two or more Ethernet links bundled together to function as a single interface. The switch can hash traffic entering a LAG interface onto any member link in the LAG interface. When you configure rate shaping and apply it to a LAG interface, the way that the switch applies the rate shaping to traffic depends on how the switch hashes the traffic onto the LAG links.

  To illustrate how link hashing affects the way the switch applies a shaping rate to LAG traffic, let's look at a LAG interface (**ae0**) that has two member links (**xe-0/0/20** and **xe-0/0/21**). On LAG **ae0**, we configure rate shaping of **2g** for traffic assigned to the **best-effort** forwarding class, which is mapped to output queue **0**. When traffic in the **best-effort** forwarding class reaches the LAG interface, the switch hashes the traffic onto one of the two member links.

  If the switch hashes all of the **best-effort** traffic onto the same LAG link, the traffic receives a maximum of 2g bandwidth on that link. In this case, the intended cumulative limit of 2g for best-effort traffic on the LAG is enforced.

  However, if the switch hashes the **best-effort** traffic onto both of the LAG links, the traffic receives a maximum of 2g bandwidth on *each* LAG link, not 2g as a cumulative total for the entire LAG, so the best-effort traffic receives a maximum of 4g on the LAG, not the 2g set by the rate shaping configuration. When hashing spreads the traffic assigned to an output queue (which is mapped to a forwarding class) across multiple LAG links, the effective rate shaping (cumulative maximum bandwidth) on the LAG is:

  (number of LAG member interfaces) x (rate shaping for the output queue) = cumulative LAG rate shaping

- On switches that do not use virtual output queues (VOQs), ingress port congestion can occur during periods of egress port congestion if an ingress port forwards traffic to more than one egress port, and at least one of those egress ports experiences congestion. If this occurs, the congested egress port can cause the ingress port to exceed its fair allocation of ingress buffer resources. When the ingress port exceeds its buffer resource allocation, frames are dropped at the ingress. Ingress port frame drop affects not only the congested egress ports, but also all of the egress ports to which the congested ingress port forwards traffic.

  If a congested ingress port drops traffic that is destined for one or more uncongested egress ports, configure a weighted random early detection (WRED) drop profile and apply it to the egress queue that is causing the congestion. The drop profile prevents the congested egress queue from affecting egress queues on other ports by dropping frames at the egress instead of causing congestion at the ingress port.

  NOTE: **On systems that support lossless transport, do not configure drop profiles for lossless forwarding classes such as the default fcoe and no-loss forwarding classes. FCoE and other lossless traffic queues require lossless behavior. Use priority-based flow control (PFC) to prevent frame drop on lossless priorities.**

- On systems that use different classifiers for unicast and multidestination traffic and that support lossless transport, on an ingress port, do not configure classifiers that map the same IEEE 802.1p code point to both a multidestination traffic flow and a lossless unicast traffic flow (such as the default lossless **fcoe** or **no-loss** forwarding classes). Any code point used for multidestination traffic on a port should not be used to classify unicast traffic into a lossless forwarding class on the same port.

  If a multidestination traffic flow and a lossless unicast traffic flow use the same code point on a port, the multidestination traffic is treated the same way as the lossless traffic. For example, if priority-based flow control (PFC) is applied to the lossless traffic, the multidestination traffic of the same code point is also paused. During periods of congestion, treating multidestination traffic the same as lossless unicast traffic can create ingress port congestion for the multidestination traffic and affect the multidestination traffic on all of the egress ports the multidestination traffic uses.

  For example, the following configuration can cause ingress port congestion for the multidestination flow:

  1. For unicast traffic, IEEE 802.1p code point 011 is classified into the **fcoe** forwarding class:

     `user@switch#` **set class-of-service classifiers ieee-802.1 ucast_cl forwarding-class fcoe loss-priority low code-points 011**

  2. For multidestination traffic, IEEE 802.1p code point **011** is classified into the **mcast** forwarding class:

     `user@switch#` **set class-of-service classifiers ieee-802.1 mcast-cl forwarding-class mcast loss-priority low code-points 011**

3. The unicast classifier that maps traffic with code point **011** to the **fcoe** forwarding class is mapped to interface **xe-0/0/1**:

   `user@switch#` **set class-of-service interfaces xe-0/0/1 unit 0 classifiers ieee-802.1 ucast_cl**

4. The multidestination classifier that maps traffic with code point **011** to the **mcast** forwarding class is mapped to all interfaces (multidestination traffic maps to all interfaces and cannot be mapped to individual interfaces):

   `user@switch#` **set class-of-service multi-destination classifiers ieee-802.1 mcast-cl**

   Because the same code point (**011**) maps unicast traffic to a lossless traffic flow and also maps multidestination traffic to a multidestination traffic flow, the multidestination traffic flow might experience ingress port congestion during periods of congestion.

To avoid ingress port congestion, do not map the code point used by the multidestination traffic to lossless unicast traffic. For example:

1. Instead of classifying code point **011** into the **fcoe** forwarding class, classify code point **011** into the **best-effort** forwarding class:

   `user@switch#` **set class-of-service classifiers ieee-802.1 ucast_cl forwarding-class best-effort loss-priority low code-points 011**

2. `user@switch#` **set class-of-service classifiers ieee-802.1 mcast-cl forwarding-class mcast loss-priority low code-points 011**

3. `user@switch#` **set class-of-service interfaces xe-0/0/1 unit 0 classifiers ieee-802.1 ucast_cl**

4. `user@switch#` **set class-of-service multi-destination classifiers ieee-802.1 mcast-cl**

   Because the code point **011** does not map unicast traffic to a lossless traffic flow, the multidestination traffic flow does not experience ingress port congestion during periods of congestion.

The best practice is to classify unicast traffic with IEEE 802.1p code points that are also used for multidestination traffic into best-effort forwarding classes.

Understanding CoS Output Queue Schedulers

Output queue scheduling defines the class-of-service (CoS) properties of output queues. Output queues are mapped to forwarding classes, and classifiers map incoming traffic into forwarding classes based on IEEE 802.1p or DSCP code points. Output queue properties include the amount of interface bandwidth assigned to the queue, the size of the memory buffer allocated for storing packets, the priority of the queue, and the weighted random early detection (WRED) drop profiles associated with the queue. Queue scheduling works with priority group scheduling to create a two-tier hierarchical scheduler.

The hierarchical scheduler allocates port bandwidth to a group of queues (forwarding classes) called a priority group (forwarding class set), and queue scheduling determines the portion of the priority group's bandwidth that a particular queue can use. So the first scheduling tier is allocating port bandwidth to a forwarding class set, and the second scheduling tier is allocating forwarding class set bandwidth to forwarding classes (queues).

Scheduler maps associate queue schedulers with forwarding classes. The queue mapped to a forwarding class receives the scheduling resources assigned to that forwarding class. You associate a scheduler map with a traffic control profile, and then associate the traffic control profile with a forwarding class set (priority group) and a port interface to apply scheduling to a port. In conjunction with the priority group scheduling configured in the traffic control profile, queue scheduling configures the packet schedulers and weighted random early detection (WRED) packet drop processes for queues.

NOTE: When you configure bandwidth for a queue or a priority group, the switch considers only the data as the configured bandwidth. The switch does not account for the bandwidth consumed by the preamble and the interframe gap (IFG). Therefore, when you calculate and configure the bandwidth requirements for a queue or for a priority group, consider the preamble and the IFG as well as the data in the calculations.

Output Queue Scheduling Components

Table 65 on page 284 provides a quick reference to the scheduler components you can configure to determine the bandwidth properties of output queues (forwarding classes), and Table 66 on page 285 provides a quick reference to some related scheduling configuration components.

Table 65: Output Queue Scheduler Components

Output Queue Scheduler Component	Description
Buffer size	Sets the size of the queue buffer. See "Understanding CoS Buffer Configuration" on page 546.
Drop profile map	Maps a drop profile to a loss priority. Drop profile map components include: • Drop profile—Sets the probability of dropping packets as the queue fills up. • Loss priority—Sets the traffic loss priority to which a drop profile applies. See "Configuring CoS Drop Profile Maps" on page 377.
Explicit congestion notification	Enables explicit congestion notification (ECN) on the queue. See "Understanding CoS Explicit Congestion Notification" on page 380.
Priority	Sets the scheduling priority applied to the queue. See "Defining CoS Queue Scheduling Priority" on page 300.
Shaping rate	Sets the maximum bandwidth the queue can consume. TIP: On QFX5200 Series switches, a granularity of 64kbps is supported for the shaping rate. See "Understanding CoS Priority Group Shaping and Queue Shaping (Maximum Bandwidth)" on page 355.
Transmit rate	Sets the minimum guaranteed bandwidth for the queue. Extra bandwidth is shared among queues in proportion to the minimum guaranteed bandwidth of each queue. See "Understanding CoS Priority Group and Queue Guaranteed Minimum Bandwidth" on page 347.

Table 66: Other Scheduling Components

Other Scheduling Components	Description
Forwarding class	Maps traffic to an output queue. Classifiers map forwarding classes to IEEE 802.1p, DSCP, or EXP code points. A forwarding class, an output queue, and code point bits are mapped to each other and identify the same traffic. (The code point bits identify incoming traffic. Classifiers assign traffic to forwarding classes based on the code point bits. Forwarding classes are mapped to output queues. This mapping determines the output queue each class of traffic uses on the switch egress interfaces.)
Output queue	Buffers traffic before the switch forwards the traffic out the egress interface. Output queues are mapped to forwarding classes. The switch applies CoS properties defined in schedulers to output queues, by mapping forwarding classes to schedulers in scheduler maps. The queue mapped to the forwarding class has the CoS properties defined in the scheduler mapped to that forwarding class.
Scheduler map	Maps schedulers to forwarding classes (forwarding classes are mapped to queues, so a forwarding class represents a queue, and the scheduler mapped to a forwarding class determines the CoS properties of the output queue mapped to that forwarding class).
Traffic control profile	Configures scheduling for the forwarding class set (priority group), and associates a scheduler map with the forwarding class set to apply queue scheduling to the forwarding classes in the forwarding class set. Extra port bandwidth is shared among forwarding class sets in proportion to the minimum guaranteed bandwidth of each forwarding class set.
Forwarding class set	Name of a priority group. You map forwarding classes to forwarding class sets. A forwarding class set consists of one or more forwarding classes.

Default Schedulers

Each forwarding class requires a scheduler to set the CoS properties of the forwarding class and its output queue. You can use the default schedulers or you can define new schedulers for the associated forwarding classes. For any other forwarding class, you must explicitly configure a scheduler. For more information, see *Default Scheduling*.

Transmit Rate (Minimum Guaranteed Bandwidth)

The transmit rate determines the minimum guaranteed bandwidth for each forwarding class. The switch applies the minimum bandwidth guarantee to the output queue mapped to the forwarding class. The transmit rate also determines how much excess (extra) bandwidth each low-priority queue can share; each queue shares extra bandwidth in proportion to its transmit rate. You specify the rate in bits per second as a fixed value such as 1 Mbps or as a percentage of the total forwarding class set minimum guaranteed bandwidth (the guaranteed rate set in the traffic control profile). Either the default scheduler or a scheduler you configure allocates a portion of the outgoing interface bandwidth to each forwarding class in proportion to the transmit rate.

NOTE: For transmit rates below 1 Gbps, we recommend that you configure the transmit rate as a percentage instead of as a fixed rate. This is because the system converts fixed rates into percentages and may round small fixed rates to a lower percentage. For example, a fixed rate of 350 Mbps is rounded down to 3 percent.

You cannot configure a transmit rate for a strict-high priority queue. Queues with a configured transmit rate cannot be included in a forwarding class set that has a strict-high priority queue (you cannot mix strict-high priority queues and queues that are not strict-high priority in the same forwarding class set).

The allocated bandwidth can exceed the configured minimum rate if additional bandwidth is available from other queues in the forwarding class set that are not using all of their allocated bandwidth. During periods of congestion, the configured transmit rate is the guaranteed bandwidth minimum for the queue. This behavior enables you to ensure that each queue receives the amount of bandwidth appropriate to its level of service and is also able to share unused bandwidth.

NOTE: Configuring the minimum guaranteed bandwidth (transmit rate) for a forwarding class does not work unless you also configure the minimum guaranteed bandwidth (guaranteed rate) for the forwarding class set in the traffic control profile.

Additionally, the sum of the transmit rates of the queues in a forwarding class set should not exceed the guaranteed rate for the forwarding class set. (You cannot guarantee a combined minimum bandwidth for the queues that is greater than the minimum bandwidth guaranteed for the entire set of queues.)

For more information, see "Understanding CoS Priority Group and Queue Guaranteed Minimum Bandwidth" on page 347.

Sharing Extra Bandwidth

Extra bandwidth is available to low-priority queues when a forwarding class set does not use its full amount of minimum guaranteed bandwidth (guaranteed-rate). Extra bandwidth is shared among the forwarding classes in a forwarding class set in proportion to the minimum guaranteed bandwidth (transmit-rate) of each queue.

For example, in a forwarding class set, Queue A has a transmit rate of 1 Gbps, Queue B has a transmit rate of 1 Gbps, and Queue C has a transmit rate of 2 Gbps. After servicing the minimum guaranteed bandwidth of these queues, the forwarding class set has an extra 2 Gbps of bandwidth available, and all three queues still have packets to forward. The queues receive the extra bandwidth in proportion to their transmit rates, so Queue A receives an extra 500 Mbps, Queue B receives an extra 500 Mbps, and Queue C receives an extra 1 Gbps.

Shaping Rate (Maximum Bandwidth)

The shaping rate sets the maximum bandwidth that a forwarding class can consume. You specify the rate in bits per second as a fixed value, such as 3 Mbps or as a percentage of the total forwarding class set maximum bandwidth (the shaping rate set in the traffic control profile).

The maximum bandwidth for a queue depends on the total bandwidth available to the forwarding class set to which the queue belongs, and on how much bandwidth the other queues in the forwarding class set consume.

NOTE: On QFabric systems, if any queue that contains outgoing packets does not transmit packets for 12 consecutive seconds, the port automatically resets. A strict-high priority queue (or several queues with higher priorities than the starved queue) can consume all of the port bandwidth and prevent another queue from transmitting packets. To prevent a queue from being starved for bandwidth, you can configure a shaping rate on the queue or queues to prevent them from consuming all of the port bandwidth.

NOTE: We recommend that you always configure a shaping rate in the scheduler for strict-high priority queues to prevent them from starving other queues.

For more information, see “Understanding CoS Priority Group Shaping and Queue Shaping (Maximum Bandwidth)” on page 355.

Scheduling Priority

Scheduling priority determines the order in which an interface transmits traffic from its output queues. This ensures that queues containing important traffic receive prioritized access to the outgoing interface bandwidth. The priority setting in the scheduler determines the priority for the queue.

For more information, see “Defining CoS Queue Scheduling Priority” on page 300.

Scheduler Drop-Profile Maps

Drop-profile maps associate drop profiles with queue schedulers and packet loss priorities (PLPs). Drop profiles set thresholds for dropping packets during periods of congestion, based on the queue fill level and a percentage probability of dropping packets at the specified queue fill level. At different fill levels, a drop profile sets different probabilities of dropping a packet during periods of congestion.

Classifiers assign incoming traffic to forwarding classes (which are mapped to output queues), and also assign a PLP to the incoming traffic. The PLP can be low, medium-high, or high. You can classify traffic with different PLPs into the same forwarding class to differentiate treatment of traffic within the forwarding class.

In a drop profile map, you can configure a different drop profile for each PLP and associate (map) the drop profiles to a queue scheduler. A scheduler map maps the queue scheduler to a forwarding class (output queue). Traffic classified into the forwarding class uses the drop characteristics defined in the drop profiles that the drop profile map associates with the queue scheduler. The drop profile the traffic uses depends on the PLP that the classifier assigns to the traffic. (You can map different drop profiles to the forwarding class for different PLPs.)

In summary:

- Classifiers assign one of three PLPs (low, medium-high, high) to incoming traffic when classifiers assign traffic to a forwarding class.
- Drop profiles set thresholds for packet drop at different queue fill levels.
- Drop profile maps associate a drop profile with each PLP, and map the drop profiles to schedulers.
- Scheduler maps map schedulers to forwarding classes, and forwarding classes are mapped to output queues. The scheduler mapped to a forwarding class determines the CoS characteristics of the output queue mapped to the forwarding class, including the drop profile mapping.

Buffer Size

Most of the total system buffer space is divided into two buffer pools, shared buffers and dedicated buffers. Shared buffers are a global pool that the ports share dynamically as needed. Dedicated buffers are a reserved portion of the buffer pool that is distributed evenly to all of the ports. Each port receives an equal allocation of dedicated buffer space. The dedicated buffer allocation to ports is not configurable because it is reserved for the ports.

The queue buffers are allocated from the dedicated buffer pool assigned to the port. By default, ports divide their allocation of dedicated buffers among the egress queues in the same proportion as the default scheduler sets the minimum guaranteed transmission rates (**transmit-rate**) for traffic. Only the queues included in the default scheduler receive dedicated buffers.

If you do not use the default configuration, you can explicitly configure the queue buffer size in either of two ways:

- As a percentage—The queue receives the specified percentage of dedicated port buffers when the queue is mapped to the scheduler and the scheduler is mapped to a port.
- As a remainder—After the port services the queues that have an explicit percentage buffer size configuration, the remaining port dedicated buffer space is divided equally among the other queues to which a scheduler is attached. (No default or explicit scheduler means no dedicated buffer allocation for the queue.) If you configure a

scheduler and you do not specify a buffer size as a percentage, *remainder* is the default setting.

NOTE: **The total of all of the explicitly configured buffer size percentages for all of the queues on a port cannot exceed 100 percent.**

For a complete discussion about queue buffer configuration in the context of ingress and egress port buffer configuration, see "Understanding CoS Buffer Configuration" on page 546.

Explicit Congestion Notification

Explicit congestion notification (ECN) notifies networks about congestion with the goal of reducing packet loss and delay by making the sending device decrease the transmission rate until the congestion clears, without dropping packets. ECN enables end-to-end congestion notification between two endpoints on TCP/IP based networks. ECN is disabled by default.

For more information, see "Understanding CoS Explicit Congestion Notification" on page 380.

Scheduler Maps

A scheduler map associates a forwarding class with a scheduler configuration. After configuring a scheduler, you must include it in a scheduler map, associate the scheduler map with a traffic control profile, and then associate the traffic control profile with an interface and a forwarding class set to implement the configured queue scheduling.

You can associate up to four user-defined scheduler maps with traffic control profiles. For more information, see *Default Schedulers Overview*.

Related Documentation

- Understanding Junos CoS Components on page 15
- Understanding CoS Priority Group Scheduling on page 307
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Understanding CoS Buffer Configuration on page 546
- Understanding CoS Explicit Congestion Notification on page 380
- Understanding CoS Scheduling Behavior and Configuration Considerations on page 277
- *Understanding CoS Scheduling on QFabric System Node Device Fabric (fte) Ports*
- *Understanding Default CoS Scheduling on QFabric System Interconnect Devices (Junos OS Release 13.1 and Later Releases)*
- Configuring CoS Drop Profile Maps on page 377
- Defining CoS Queue Scheduling Priority on page 300
- Understanding CoS Priority Group Shaping and Queue Shaping (Maximum Bandwidth) on page 355

• Understanding CoS Priority Group and Queue Guaranteed Minimum Bandwidth on page 347

Defining CoS Queue Schedulers

Schedulers define the CoS properties of output queues (output queues are mapped to forwarding classes, and classifiers map traffic into forwarding classes based on IEEE 802.1p, DSCP, or MPLS EXP code points). Queue scheduling works with priority group scheduling to create a two-tier hierarchical scheduler. CoS scheduling properties include the amount of interface bandwidth assigned to the queue, the priority of the queue, whether explicit congestion notification (ECN) is enabled on the queue, and the WRED packet drop profiles associated with the queue.

The parameters you configure in a scheduler define the following characteristics for the queues mapped to the scheduler:

- **transmit-rate**—Minimum bandwidth, also known as the *committed information rate (CIR)*, set as a percentage rate or as an absolute value in bits per second. The transmit rate also determines the amount of excess (extra) priority group bandwidth that the queue can share. Extra priority group bandwidth is allocated among the queues in the priority group in proportion to the transmit rate of each queue.

  NOTE: **Include the preamble bytes and interframe gap (IFG) bytes as well as the data bytes in your bandwidth calculations.**

  NOTE: **You cannot configure a transmit rate for strict-high priority queues. Queues (forwarding classes) with a configured transmit rate cannot be included in a forwarding class set that has strict-high priority queues.**

- **shaping-rate**—Maximum bandwidth, also known as the *peak information rate (PIR)*, set as a percentage rate or as an absolute value in bits per second.

  NOTE: **Include the preamble bytes and interframe gap (IFG) bytes as well as the data bytes in your bandwidth calculations.**

- **priority**—One of two bandwidth priorities that queues associated with a scheduler can receive:
  - **low**—The scheduler has low priority.
  - **strict-high**—The scheduler has strict-high priority. You can configure only one queue as a strict-high priority queue. Strict-high priority allocates the scheduled bandwidth to the queue before any other queue receives bandwidth. Other queues receive the bandwidth that remains after the strict-high queue has been serviced.

We recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

- **drop-profile-map**—Drop profile mapping to a loss priority and protocol, to apply WRED to the scheduler and control packet drop for different packet loss priorities during periods of congestion.
- **buffer-size**—Size of the queue buffer as a percentage of the dedicated buffer space on the port, or as a proportional share of the dedicated buffer space on the port that remains after the explicitly configured queues are served.
- **explicit-congestion-notification**—Enables ECN on a best-effort queue. ECN enables end-to-end congestion notification between two ECN-enabled endpoints on TCP/IP based networks. ECN must be enabled on both endpoints and on all of the intermediate devices between the endpoints for ECN to work properly. ECN is disabled by default.

NOTE: Ingress port congestion can occur during periods of egress port congestion if an ingress port forwards traffic to more than one egress port, and at least one of those egress ports experiences congestion. If this occurs, the congested egress port can cause the ingress port to exceed its fair allocation of ingress buffer resources. When the ingress port exceeds its buffer resource allocation, frames are dropped at the ingress. Ingress port frame drop affects not only the congested egress ports, but also all of the egress ports to which the congested ingress port forwards traffic.

If a congested ingress port drops traffic that is destined for one or more uncongested egress ports, configure a weighted random early detection (WRED) drop profile and apply it to the egress queue that is causing the congestion. The drop profile prevents the congested egress queue from affecting egress queues on other ports by dropping frames at the egress instead of causing congestion at the ingress port.

NOTE: Do not configure drop profiles for the fcoe and no-loss forwarding classes. FCoE and other lossless traffic queues require lossless behavior. Use priority-based flow control (PFC) to prevent frame drop on lossless priorities.

OCX Series switches do not support lossless transport or PFC. On OCX Series switches, do not map traffic to the default lossless fcoe and no-loss forwarding classes.

To apply scheduling properties to traffic, map schedulers to forwarding classes using a scheduler map, and then associate the scheduler map with interfaces. (You associate a scheduler map with an interface using a traffic control profile; see “Example: Configuring CoS Hierarchical Port Scheduling (ETS)” on page 321 for an example of the complete hierarchical scheduling process.) Using different scheduler maps, you can map different schedulers to the same traffic (the same forwarding class) on different interfaces, to apply different scheduling to that traffic on different interfaces.

To configure a scheduler using the CLI:

1. Name the scheduler and set the minimum guaranteed bandwidth for the queue:

   ```
   [edit class-of-service]
   user@switch# set schedulers scheduler-name transmit-rate (rate | percent percentage)
   ```

2. Set the maximum bandwidth for the queue:

   ```
   [edit class-of-service schedulers scheduler-name]
   user@switch# set shaping-rate (rate | percent percentage)
   ```

3. Set the queue priority:

   ```
   [edit class-of-service schedulers scheduler-name]
   user@switch# set priority level
   ```

4. Specify drop profiles for packet loss priorities using a drop profile map:

   ```
   [edit class-of-service schedulers scheduler-name]
   user@switch# set drop-profile-map loss-priority (low | medium-high | high) protocol protocol
   drop-profile drop-profile-name
   ```

5. Configure the size of the port dedicated buffer space for the queue:

   ```
   [edit class-of-service schedulers scheduler-name]
   user@switch# set buffer-size (percent percent | remainder)
   ```

6. Enable ECN, if desired (on best-effort traffic only):

   ```
   [edit class-of-service schedulers scheduler-name]
   user@switch# set explicit-congestion-notification
   ```

7. Configure a scheduler map to map the scheduler to a forwarding class, which applies the scheduler's properties to the traffic in that forwarding class:

   ```
   [edit class-of-service]
   user@switch# set scheduler-maps scheduler-map-name forwarding-class
   forwarding-class-name scheduler scheduler-name
   ```

8. Assign the scheduler map and its associated schedulers to one or more interfaces using hierarchical scheduling. See "Example: Configuring CoS Hierarchical Port Scheduling (ETS)" on page 321 for a detailed example of hierarchical scheduling.

   ```
   [edit class-of-service]
   ```



```
user@switch# set traffic-control-profiles tcp-name scheduler-map scheduler-map-name
user@switch# set interfaces interface-name forwarding-class-set fc-set-name
output-traffic-control-profile tcp-name
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357
- Example: Configuring ECN on page 389
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Defining CoS Queue Scheduling Priority on page 300
- Configuring CoS WRED Drop Profiles on page 369
- Monitoring CoS Scheduler Maps on page 431
- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS Priority Group Scheduling on page 307
- Understanding CoS Buffer Configuration on page 546
- Understanding CoS Explicit Congestion Notification on page 380

Example: Configuring Queue Schedulers

Schedulers define the CoS properties of output queues (output queues are mapped to forwarding classes, and classifiers map traffic into forwarding classes based on IEEE 802.1p or DSCP code points). Queue scheduling works with priority group scheduling to create a two-tier hierarchical scheduler. CoS scheduling properties include the amount of interface bandwidth assigned to the queue, the priority of the queue, whether explicit congestion notification (ECN) is enabled on the queue, and the WRED packet drop profiles associated with the queue.

- Requirements on page 293
- Overview on page 294
- Configuring a CoS Scheduler on page 296
- Verification on page 298

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

Scheduler parameters define the following characteristics for the queues mapped to the scheduler:

- **transmit-rate**—Minimum bandwidth, also known as the *committed information rate (CIR)*. Each queue mapped to the scheduler receives a minimum of either the configured amount of absolute bandwidth or the configured percentage of bandwidth. The transmit rate also determines the amount of excess (extra) priority group bandwidth that the queue can share. Extra priority group bandwidth is allocated among the queues in the priority group in proportion to the transmit rate of each queue. You cannot configure a transmit rate for strict-high priority queues. Queues (forwarding classes) with a configured transmit rate cannot be included in a forwarding class set that has strict-high priority queues.

  NOTE: The **transmit-rate** setting works only if you also configure the **guaranteed-rate** in the traffic control profile that is attached to the forwarding class set to which the queue belongs. If you do not configure the **guaranteed-rate**, the **transmit-rate** does not work. The sum of all queue transmit rates in a forwarding class set should not exceed the traffic control profile guaranteed rate. If you configure transmit rates whose sum exceeds the forwarding class set guaranteed rate, the commit check fails, and the system rejects the configuration.

  NOTE: Include the preamble bytes and interframe gap bytes as well as the data bytes in your bandwidth calculations.

  NOTE: You cannot configure a transmit rate for strict-high priority queues. Queues (forwarding classes) with a configured transmit rate cannot be included in a forwarding class set that has strict-high priority queues.

- **shaping-rate**—Maximum bandwidth, also known as the *peak information rate (PIR)*. Each queue receives a maximum of the configured amount of absolute bandwidth or the configured percentage of bandwidth, even if more bandwidth is available.

  NOTE: Include the preamble bytes and interframe gap bytes as well as the data bytes in your bandwidth calculations.

- **priority**—One of two bandwidth priorities that queues associated with a scheduler can receive:
  - **low**—The scheduler has low priority.

- **strict-high**—The scheduler has strict-high priority. You can configure only one queue as a strict-high priority queue. Strict-high priority allocates the scheduled bandwidth to the queue before any other queue receives bandwidth. Other queues receive the bandwidth that remains after the strict-high queue has been serviced.

  We recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

- **drop-profile-map**—Mapping of a drop profile to a loss priority and protocol to apply WRED to the scheduler.
- **buffer-size**—Size of the queue buffer as a percentage of the dedicated buffer space on the port, or as a proportional share of the dedicated buffer space on the port that remains after the explicitly configured queues are served.
- **explicit-congestion-notification**—Enables ECN on a best-effort queue. ECN enables end-to-end congestion notification between two ECN-enabled endpoints on TCP/IP based networks. ECN must be enabled on both endpoints and on all of the intermediate devices between the endpoints for ECN to work properly. ECN is disabled by default.

NOTE: Ingress port congestion can occur during periods of egress port congestion if an ingress port forwards traffic to more than one egress port, and at least one of those egress ports experiences congestion. If this occurs, the congested egress port can cause the ingress port to exceed its fair allocation of ingress buffer resources. When the ingress port exceeds its buffer resource allocation, frames are dropped at the ingress. Ingress port frame drop affects not only the congested egress ports, but also all of the egress ports to which the congested ingress port forwards traffic.

If a congested ingress port drops traffic that is destined for one or more uncongested egress ports, configure a weighted random early detection (WRED) drop profile and apply it to the egress queue that is causing the congestion. The drop profile prevents the congested egress queue from affecting egress queues on other ports by dropping frames at the egress instead of causing congestion at the ingress port.

NOTE: Do not configure drop profiles for the fcoe and no-loss forwarding classes. FCoE and other lossless traffic queues require lossless behavior. Use priority-based flow control (PFC) to prevent frame drop on lossless priorities.

OCX Series switches do not support lossless transport or PFC. On OCX Series switches, do not map traffic to the default lossless fcoe and no-loss forwarding classes.

Scheduler maps associate schedulers with forwarding classes (queues). After defining schedulers and mapping them to queues in a scheduler map, to configure hardware queue scheduling (hierarchical port scheduling) you:

1. Associate a scheduler map with a traffic control profile (a traffic control profile schedules resources for a group of forwarding classes, called a *forwarding class set* or *priority group*).

2. Attach a forwarding class and a traffic control profile to an interface.

   "Example: Configuring CoS Hierarchical Port Scheduling (ETS)" on page 321 provides a complete example of hierarchical scheduling.

You can associate up to four user-defined scheduler maps with forwarding class sets.

This process configures the bandwidth properties and WRED characteristics that you map to forwarding classes (and thus to output queues) in a scheduler map. The traffic control profile uses the scheduler CoS properties to determine the resources that should be allocated to the individual output queues from the total resources available to the priority group.

Table 67 on page 296 shows the configuration components for this example.

Table 67: Components of the Queue Scheduler Configuration Example

Component	Settings
Hardware	QFX3500 switch
Scheduler	Name: **be-sched** Transmit rate: **20%** Shaping rate: **40%** Buffer size: **20%** Priority: **low** Drop profile: **be-dp** ECN: **disable** (default)
Scheduler map	Name: **be-map** Forwarding class to associate with the **be-sched** scheduler: **best-effort**
Traffic control profile	Name: **be-tcp** NOTE: This topic does not describe how to define a traffic control profile.
Forwarding class set	Name: **lan-pg**

Configuring a CoS Scheduler

CLI Quick Configuration

To quickly configure a queue scheduler, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit] hierarchy level:

```
[edit class-of-service]
set schedulers be-sched transmit-rate percent 20
set schedulers be-sched shaping-rate percent 40
set schedulers be-sched buffer-size percent 20
set schedulers be-sched priority low
set schedulers be-sched drop-profile-map loss-priority low protocol any drop-profile be-dp
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set traffic-control-profiles be-tcp scheduler-map be-map
set interfaces xe-0/0/7 forwarding-class-set lan-pg output-traffic-control-profile be-tcp
```

To configure a CoS scheduler:

1. Create scheduler (**be-sched**) with a minimum guaranteed bandwidth of 2 Gbps, a maximum bandwidth of 4 Gbps, and low priority, and map it to the drop profile **be-dp**:

   ```
   [edit class-of-service schedulers]
   user@switch# set be-sched transmit-rate percent 20
   user@switch# set be-sched shaping-rate percent 40
   user@switch# set be-sched buffer-size percent 20
   user@switch# set be-sched priority low
   user@switch# set be-sched drop-profile-map loss-priority low protocol any drop-profile
   be-dp
   ```

   NOTE: Because ECN is disabled by default, no ECN configuration is shown.

2. Configure scheduler map (**be-map**) to associate the scheduler (**be-sched**) with the forwarding class (**best-effort**):

   ```
   [edit class-of-service scheduler-maps]
   user@switch# set be-map forwarding-class best-effort scheduler be-sched
   ```

3. Associate the scheduler map **be-map** with a traffic control profile (**be-tcp**):

   ```
   [edit class-of-service traffic-control-profiles]
   user@switch# set be-tcp scheduler-map be-map
   ```

4. Associate the traffic control profile **be-tcp** with a forwarding class set (**lan-pg**) and a 10-Gigabit Ethernet interface (**xe-0/0/7**):

   ```
   [edit class-of-service]
   user@switch# set interfaces xe-0/0/7 forwarding-class-set lan-pg
   output-traffic-control-profile be-tcp
   ```

Verification

To verify that the queue scheduler has been created and is mapped to the correct interfaces, perform these tasks:

- Verifying the Scheduler Configuration on page 298
- Verifying the Scheduler Map Configuration on page 298
- Verifying That the Scheduler Is Associated with the Interface on page 299

Verifying the Scheduler Configuration

Purpose Verify that the queue scheduler **be-sched** has been created with a minimum guaranteed bandwidth of 2 Gbps, a maximum bandwidth of 4 Gbps, the priority set to **low**, and the drop profile **be-dp**.

Action Display the scheduler using the operational mode command **show configuration class-of-service schedulers be-sched**:

```
user@switch> show configuration class-of-service schedulers be-sched
transmit-rate percent 20;
shaping-rate percent 40;
buffer-size percent 20;
priority low;
drop-profile-map loss-priority low protocol any drop-profile be-dp;
```

Verifying the Scheduler Map Configuration

Purpose Verify that the scheduler map **be-map** has been created and associates the forwarding class **best-effort** with the scheduler **be-sched**, and also that the scheduler map is attached to the traffic control profile **be-tcp**.

Action Display the scheduler map using the operational mode command **show configuration class-of-service scheduler-maps be-map**:

```
user@switch> show configuration class-of-service scheduler-maps be-map
forwarding-class best-effort scheduler be-sched;
```

Display the traffic control profile to verify that the scheduler map **be-map** is attached using the operational mode command **show configuration class-of-service traffic-control-profiles be-tcp scheduler-map**:

```
user@switch> show configuration class-of-service traffic-control-profiles be-tcp scheduler-map
scheduler-map be-map;
```


NOTE: This topic does not describe how to configure a traffic control profile or its allocation of port bandwidth. Using a traffic control profile to configure the port resource allocation to the priority group is necessary to implement hierarchical scheduling.

Verifying That the Scheduler Is Associated with the Interface

Purpose

Verify that the forwarding class set (**lan-pg**) and the traffic control profile (**be-tcp**) that are associated with the queue scheduler are attached to the interface **xe-0/0/7**.

Action

List the interface using the operational mode command **show configuration class-of-service interfaces xe-0/0/7**:

```
user@switch> show configuration class-of-service interfaces xe-0/0/7
forwarding-class-set {
    lan-pg {
        output-traffic-control-profile be-tcp;
    }
}
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring WRED Drop Profiles on page 371
- Example: Configuring ECN on page 389
- Defining CoS Queue Schedulers on page 290
- Monitoring CoS Scheduler Maps on page 431
- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- *Understanding CoS Scheduling on QFabric System Node Device Fabric (fte) Ports*
- *Understanding Default CoS Scheduling on QFabric System Interconnect Devices (Junos OS Release 13.1 and Later Releases)*
- Understanding CoS Buffer Configuration on page 546

Defining CoS Queue Scheduling Priority

You can configure the scheduling priority of individual queues by specifying the priority in a scheduler, and then associating the scheduler with a queue by using a scheduler map. On QFX5100, QFX5200, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, queues can have one of two bandwidth scheduling priorities, **strict-high** priority or **low** priority. On QFX10000 Series switches, queues can also be configured as **high** priority.

NOTE: By default, all queues are low priority queues.

The switch services low priority queues after servicing any queue that has strict-high priority traffic or high priority traffic. Strict-high priority queues receive preferential treatment over all other queues and receive all of their configured bandwidth before other queues are serviced. Low-priority queues do not transmit traffic until strict-high priority queues are empty, and receive the bandwidth that remains after the strict-high queues have been serviced. High priority queues receive preference over low priority queues.

Different switches handle traffic configured as **strict-high** priority traffic in different ways:

- QFX5100, QFX5200, QFX3500, QFX3600, and EX4600 switches, and QFabric systems—You can configure only one queue as a strict-high priority queue.

  On these switches, we recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

- QFX10000 switches—You can configure as many queues as you want as strict-high priority. However, keep in mind that too much strict-high priority traffic can starve low priority queues on the port.

  NOTE: We strongly recommend that you configure a transmit rate on all strict-high priority queues to limit the amount of traffic the switch treats as strict-high priority traffic and prevent strict-high priority queues from starving other queues on the port. This is especially important if you configure more than one strict-high priority queue on a port. If you do not configure a transmit rate to limit the amount of bandwidth strict-high priority queues can use, then the strict-high priority queues can use all of the available port bandwidth and starve other queues on the port.

  The switch treats traffic in excess of the transmit rate as best-effort traffic that receives bandwidth from the leftover (excess) port bandwidth pool. On strict-high priority queues, all traffic that exceeds the transmit rate shares in the port excess bandwidth pool based on the strict-high priority excess bandwidth sharing weight of “1”, which is not configurable. The actual amount of extra bandwidth that traffic exceeding the transmit rate receives depends on how many other queues consume excess bandwidth and the excess rates of those queues.

- To configure queue priority using the CLI:

```
[edit class-of-service]
user@switch# set schedulers scheduler-name priority level
```

Related Documentation

- Example: Configuring Queue Scheduling Priority on page 302
- Monitoring CoS Scheduler Maps on page 431

Example: Configuring Queue Scheduling Priority

You can configure the bandwidth scheduling priority of individual queues by specifying the priority in a scheduler, and then using a scheduler map to associate the scheduler with a queue.

Requirements

This example uses the following hardware and software components:

- One switch.
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series.

Overview

Queues can have one of several bandwidth priorities:

- **strict-high**—Strict-high priority allocates bandwidth to the queue before any other queue receives bandwidth. Other queues receive the bandwidth that remains after the strict-high queue has been serviced. On QFX10000 switches, you can configure as many queues as you want as strict-high priority queues. On QFX5200, QFX3500, and QFX3600 switches and on QFabric systems, you can configure only one queue as a strict-high queue. On QFX5100 and EX4600 switches, you can configure only one forwarding-class-set (priority group) as strict-high priority. All queues which are part of that strict-high forwarding class set then act as strict-high queues.

NOTE: On QFX5200 switches, it is not possible to support multiple queues with strict-high priority because QFX5200 doesn't support flexible hierarchical scheduling. When multiple strict-high priority queues are configured, all of those queues are treated as strict-high priority but the higher number queue among them is given highest priority.

On QFX10000 switches, if you configure strict-high priority queues on a port, we strongly recommend that you configure a transmit rate on those queues. The transmit rate sets the amount of traffic that the switch forwards as strict-high priority; traffic in excess of the transmit rate is treated as best-effort traffic that receives the queue excess rate. Even if you configure only one strict-high priority queue, we strongly recommend that you configure a transmit rate the queue to prevent it from starving other queues. If you do not configure a transmit rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

On QFX5200, QFX5100, QFX3500, QFX3600, and EX4600 switches and on QFabric systems, we recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

NOTE: On switches that support enhanced transmission selection (ETS) hierarchical scheduling, if you use ETS and you configure a **strict-high** priority queue, you must create a forwarding class set that is dedicated only to **strict-high** priority traffic. Only one forwarding class set can contain a strict-high priority queue. Queues that are not strict-high priority cannot belong to the same forwarding class set as strict-high priority queues.

On switches that use different output queues for unicast and multidestination traffic, the multidestination forwarding class set cannot contain strict-high priority queues.

- **high** (QFX10000 Series switches only)—High priority. Traffic with high priority is serviced after any queue that has a **strict-high** priority, and before queues with low priority.
- **low**—Low priority. Traffic with low priority is serviced after any queue that has a **strict-high** priority.

NOTE: By default, all queues are low priority queues.

Table 68 on page 304 shows the configuration components for this example.

This example describes how to set the queue priority for two forwarding classes (queues) named **fcoe** and **no-loss**. Both queues have a priority of **low**. The scheduler for the **fcoe** queue is named **fcoe-sched** and the scheduler for the **no-loss** queue is named **nl-sched**. One scheduler map, **schedmap1**, associates the schedulers to the queues.

Table 68: Components of the Queue Scheduler Priority Configuration Example

Component	Settings
Hardware	One switch
Schedulers	**fcoe-sched** for FCoE traffic **nl-sched** for no-loss traffic
Priority	**low** for FCoE traffic **low** for no-loss traffic
Scheduler map	**schedmap1**: FCoE mapping: scheduler **fcoe-sched** to forwarding class **fcoe** No-loss mapping: scheduler **nl-sched** to forwarding class **no-loss**

NOTE: OCX Series switches do not support lossless transport. On OCX Series switches, the default DSCP classifier does not map traffic to the default fcoe and no-loss forwarding classes. On an OCX Series switch, you could use this example by substituting other forwarding classes (for example, best-effort or network-control) for the fcoe and no-loss forwarding classes, and naming the schedulers appropriately. The active forwarding classes (**best-effort**, **network-control**, and **mcast**) share the unused bandwidth assigned to the **fcoe** and **no-loss** forwarding classes.

Configuring Queue Scheduling Priority

CLI Quick Configuration

To quickly configure queue scheduling priority, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit] hierarchy level:

```
[edit class-of-service]
set schedulers fcoe-sched priority low
set schedulers nl-sched priority low
set scheduler-maps schedmap1 forwarding-class fcoe scheduler fcoe-sched
set scheduler-maps schedmap1 forwarding-class no-loss scheduler nl-sched
```

To configure queue priority using the CLI:

1. Create the FCoE scheduler with **low** priority:

```
[edit class-of-service]
user@switch# set schedulers fcoe-sched priority low
```


2. Create the no-loss scheduler with **low** priority:

```
[edit class-of-service]
user@switch# set schedulers nl-sched priority low
```

3. Associate the schedulers with the desired queues in the scheduler map:

```
[edit class-of-service]
user@switch# set scheduler-maps schedmap1 forwarding-class fcoe scheduler fcoe-sched
user@switch# set scheduler-maps schedmap1 forwarding-class no-loss scheduler nl-sched
```

Verification

To verify that you configured the queue scheduling priority for bandwidth and mapped the schedulers to the correct forwarding classes, perform these tasks:

- Verifying the Queue Scheduling Priority on page 305
- Verifying the Scheduler-to-Forwarding-Class Mapping on page 305

Verifying the Queue Scheduling Priority

Purpose Verify that you configured the queue schedulers **fcoe-sched** and **nl-sched** with **low** queue scheduling priority.

Action Display the **fcoe-sched** scheduler priority configuration using the operational mode command **show configuration class-of-service schedulers fcoe-sched priority**:

```
user@switch> show configuration class-of-service schedulers fcoe-sched priority
priority low;
```

Display the **nl-sched** scheduler priority configuration using the operational mode command **show configuration class-of-service schedulers nl-sched priority**:

```
user@switch> show configuration class-of-service schedulers nl-sched priority
priority low;
```

Verifying the Scheduler-to-Forwarding-Class Mapping

Purpose Verify that you configured the scheduler map **schedmap1** to map scheduler **fcoe-sched** to forwarding class **fcoe** and schedule **nl-sched** to forwarding class **no-loss**.

Action Display the scheduler map **schedmap1** using the operational mode command **show configuration class-of-service scheduler-maps schedmap1**:

```
user@switch> show configuration class-of-service scheduler-maps schedmap1
forwarding-class fcoe scheduler fcoe-sched;
forwarding-class no-loss scheduler nl-sched;
```


Related Documentation

- Defining CoS Queue Scheduling Priority on page 300
- Monitoring CoS Scheduler Maps on page 431

Understanding CoS Traffic Control Profiles

A traffic control profile defines the output bandwidth and scheduling characteristics of forwarding class sets (priority groups). The forwarding classes (which are mapped to output queues) that belong to a forwarding class set (fc-set) share the bandwidth that you assign to the fc-set in the traffic control profile.

This two-tier hierarchical scheduling architecture provides flexibility in allocating resources among forwarding classes, and also:

- Assigns a portion of port bandwidth to an fc-set. You define the port resources for the fc-set in a traffic control profile.
- Allocates fc-set bandwidth among the forwarding classes (queues) that belong to the fc-set. A scheduler map attached to the traffic control profile defines the amount of the fc-set's resources that each forwarding class can use.

Attaching an fc-set and a traffic control profile to a port defines the hierarchical scheduling properties of the group and the forwarding classes that belong to the group.

The ability to create fc-sets supports enhanced transmission selection (ETS), which is described in IEEE 802.1Qaz. When an fc-set does not use its allocated port bandwidth, ETS shares the excess port bandwidth among other fc-sets on the port in proportion to their guaranteed minimum bandwidth (guaranteed rate). This utilizes the port bandwidth better than scheduling schemes that reserve bandwidth for groups even if that bandwidth is not used. ETS shares unused port bandwidth, so traffic groups that need extra bandwidth can use it if the bandwidth is available, while preserving the ability to specify the minimum guaranteed bandwidth for traffic groups.

Traffic control profiles define the following CoS properties for fc-sets:

- Minimum guaranteed bandwidth—Also known as the *committed information rate (CIR)*. This is the minimum amount of port bandwidth the priority group receives. Priorities in the priority group receive their minimum guaranteed bandwidth as a portion of the priority group's minimum guaranteed bandwidth. The **guaranteed-rate** statement defines the minimum guaranteed bandwidth.

  NOTE: **You cannot apply a traffic control profile with a minimum guaranteed bandwidth to a priority group that includes strict-high priority queues.**

- Shared excess (extra) bandwidth—When the priority groups on a port do not consume the full amount of bandwidth allocated to them or there is unallocated link bandwidth available, priority groups can contend for that extra bandwidth if they need it. Priorities in the priority group contend for extra bandwidth as a portion of the priority group's extra bandwidth. The amount of extra bandwidth for which a priority group can contend is proportional to the priority group's guaranteed minimum bandwidth (guaranteed rate).

- Maximum bandwidth—Also known as *peak information rate (PIR)*. This is the maximum amount of port bandwidth the priority group receives. Priorities in the priority group receive their maximum bandwidth as a portion of the priority group's maximum bandwidth. The **shaping-rate** statement defines the maximum bandwidth.
- Queue scheduling—Each traffic control profile includes a scheduler map. The scheduler map maps forwarding classes (priorities) to schedulers to define the scheduling characteristics of the individual forwarding classes in the fc-set. The resources scheduled for each forwarding class represent portions of the resources that the traffic control profile schedules for the entire fc-set, not portions of the total link bandwidth. The **scheduler-maps** statement defines the mapping of forwarding classes to schedulers.

Related Documentation

- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Defining CoS Traffic Control Profiles (Priority Group Scheduling) on page 310

Understanding CoS Priority Group Scheduling

Priority group scheduling defines the class-of-service (CoS) properties of a group of output queues (priorities). Priority group scheduling works with output queue scheduling to create a two-tier hierarchical scheduler. The hierarchical scheduler allocates bandwidth to a group of queues (a priority group, called a forwarding class set in Junos OS configuration). Queue scheduling determines the portion of the priority group bandwidth that the particular queue can use.

You configure priority group scheduling in a traffic control profile and then associate the traffic control profile with a forwarding class set and an interface. You attach a scheduler map to the traffic control profile to specify the queue scheduling characteristics.

NOTE: When you configure bandwidth for a queue or a priority group, the switch considers only the data as the configured bandwidth. The switch does not account for the bandwidth consumed by the preamble and the interframe gap (IFG). Therefore, when you calculate and configure the bandwidth requirements for a queue or for a priority group, consider the preamble and the IFG as well as the data in the calculations.

Priority Group Scheduling Components

Table 69 on page 308 provides a quick reference to the traffic control profile components you can configure to determine the bandwidth properties of priority groups, and Table 70 on page 308 provides a quick reference to some related scheduling configuration components.

Table 69: Priority Group Scheduler Components

Traffic Control Profile Component	Description
Guaranteed rate	Sets the minimum guaranteed port bandwidth for the priority group. Extra port bandwidth is shared among priority groups in proportion to the guaranteed rate of each priority group on the port.
Shaping rate	Sets the maximum port bandwidth the priority group can consume.
Scheduler map	Maps schedulers to queues (forwarding classes, also called priorities). This determines the portion of the priority group bandwidth that a queue receives.

Table 70: Other Scheduling Components

Other Scheduling Components	Description
Forwarding class	Maps traffic to a queue (priority).
Forwarding class set	Name of a priority group. You map forwarding classes to priority groups. A forwarding class set consists of one or more forwarding classes.
Scheduler	Sets the bandwidth and scheduling priority of individual queues (forwarding classes).

Default Traffic Control Profile

There is no default traffic control profile.

Guaranteed Rate (Minimum Guaranteed Bandwidth)

The guaranteed rate determines the minimum guaranteed bandwidth for each priority group. It also determines how much excess (extra) port bandwidth the priority group can share; each priority group shares extra port bandwidth in proportion to its guaranteed rate. You specify the rate in bits per second as a fixed value such as 3 Mbps or as a percentage of the total port bandwidth.

The minimum transmission bandwidth can exceed the configured rate if additional bandwidth is available from other priority groups on the port. In case of congestion, the configured guaranteed rate is guaranteed for the priority group. This property enables you to ensure that each priority group receives the amount of bandwidth appropriate to its level of service.

NOTE: Configuring the minimum guaranteed bandwidth (transmit rate) for a forwarding class does not work unless you also configure the minimum guaranteed bandwidth (guaranteed rate) for the forwarding class set in the traffic control profile.

Additionally, the sum of the transmit rates of the queues in a forwarding class set should not exceed the guaranteed rate for the forwarding class set. (You cannot guarantee a minimum bandwidth for the queues that is greater than the minimum bandwidth guaranteed for the entire set of queues.)

You cannot configure a guaranteed rate for forwarding class sets that include strict-high priority queues.

Sharing Extra Bandwidth

Extra bandwidth is available to priority groups when the priority groups do not use the full amount of available port bandwidth. This extra port bandwidth is shared among the priority groups based on the minimum guaranteed bandwidth of each priority group.

For example, Port A has three priority groups: fc-set-1, fc-set-2, and fc-set-3. Fc-set-1 has a guaranteed rate of 2 Gbps, fc-set-2 has a guaranteed rate of 2 Gbps, and fc-set-3 has a guaranteed rate of 4 Gbps. After servicing the minimum guaranteed bandwidth of these priority groups, the port has an extra 2 Gbps of available bandwidth, and all three priority groups have still have packets to forward. The priority groups receive the extra bandwidth in proportion to their guaranteed rates, so fc-set-1 receives an extra 500 Mbps, fc-set-2 receives an extra 500 Mbps, and fc-set-3 receives an extra 1 Gbps.

Shaping Rate (Maximum Bandwidth)

The shaping rate determines the maximum bandwidth the priority group can consume. You specify the rate in bits per second as a fixed value such as 5 Mbps or as a percentage of the total port bandwidth.

The maximum bandwidth for a priority group depends on the total bandwidth available on the port and how much bandwidth the other priority groups on the port consume.

Scheduler Maps

A scheduler map maps schedulers to queues. When you associate a scheduler map with a traffic control profile, then associate the traffic control profile with an interface and a forwarding class set, the scheduling defined by the scheduler map determines the portion of the priority group resources that each individual queue can use.

You can associate up to four user-defined scheduler maps with traffic control profiles.

Related Documentation

- Understanding Junos CoS Components on page 15
- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315

- Understanding CoS Scheduling Behavior and Configuration Considerations on page 277
- *Understanding CoS Scheduling on QFabric System Node Device Fabric (fte) Ports*
- *Understanding Default CoS Scheduling on QFabric System Interconnect Devices (Junos OS Release 13.1 and Later Releases)*
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring WRED Drop Profiles on page 371
- Example: Configuring Drop Profile Maps on page 377

Defining CoS Traffic Control Profiles (Priority Group Scheduling)

A traffic control profile defines the output bandwidth and scheduling characteristics of forwarding class sets (priority groups). The forwarding classes (which are mapped to output queues) contained in a forwarding class set (fc-set) share the bandwidth resources that you configure in the traffic control profile. A scheduler map associates forwarding classes with schedulers to define how the individual forwarding classes that belong to an fc-set share the bandwidth allocated to that fc-set.

The parameters you configure in a traffic control profile define the following characteristics for the fc-set:

- **guaranteed-rate**—Minimum bandwidth, also known as the *committed information rate (CIR)*. The guaranteed rate also determines the amount of excess (extra) port bandwidth that the fc-set can share. Extra port bandwidth is allocated among the fc-sets on a port in proportion to the guaranteed rate of each fc-set.

  NOTE: **You cannot configure a guaranteed rate for a, fc-set that includes strict-high priority queues. If the traffic control profile is for an fc-set that contains strict-high priority queues, do not configure a guaranteed rate.**

- **shaping-rate**—Maximum bandwidth, also known as the *peak information rate (PIR)*.
- **scheduler-map**—Bandwidth and scheduling characteristics for the queues, defined by mapping forwarding classes to schedulers. (The queue scheduling characteristics represent amounts or percentages of the fc-set bandwidth, not the amounts or percentages of total link bandwidth.)

NOTE: Because a port can have more than one fc-set, when you assign resources to an fc-set, keep in mind that the total port bandwidth must serve all of the queues associated with that port.

To configure a traffic control profile using the CLI:

1. Name the traffic control profile and define the minimum guaranteed bandwidth for the fc-set:

```
[edit class-of-service ]
user@switch# set traffic-control-profiles traffic-control-profile-name guaranteed-rate (rate | percent percentage)
```

2. Define the maximum bandwidth for the fc-set:

```
[edit class-of-service traffic-control-profiles traffic-control-profile-name]
user@switch# set shaping-rate (rate | percent percentage)
```

3. Attach a scheduler map to the traffic control profile:

```
[edit class-of-service traffic-control-profiles traffic-control-profile-name]
user@switch# set scheduler-map scheduler-map-name
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357
- Defining CoS Queue Schedulers on page 290
- Understanding CoS Traffic Control Profiles on page 306

Example: Configuring Traffic Control Profiles (Priority Group Scheduling)

A traffic control profile defines the output bandwidth and scheduling characteristics of forwarding class sets (priority groups). The forwarding classes (queues) mapped to a forwarding class set share the bandwidth resources that you configure in the traffic control profile. A scheduler map associates forwarding classes with schedulers to define how the individual queues in a forwarding class set share the bandwidth allocated to that forwarding class set.

Requirements

This example uses the following hardware and software components:

- A Juniper Networks QFX3500 Switch
- Junos OS Release 11.1 or later for the QFX Series

Overview

The parameters you configure in a traffic control profile define the following characteristics for the priority group:

- **guaranteed-rate**—Minimum bandwidth, also known as the *committed information rate (CIR)*. Each fc-set receives a minimum of either the configured amount of absolute bandwidth or the configured percentage of bandwidth. The guaranteed rate also determines the amount of excess (extra) port bandwidth that the fc-set can share. Extra port bandwidth is allocated among the fc-sets on a port in proportion to the guaranteed rate of each fc-set.

  NOTE: In order for the transmit-rate option (minimum bandwidth for a queue that you set using scheduler configuration) to work properly, you must configure the **guaranteed-rate** for the fc-set. If an fc-set does not have a guaranteed minimum bandwidth, the forwarding classes that belong to the fc-set cannot have a guaranteed minimum bandwidth.

  NOTE: Include the preamble bytes and interframe gap bytes as well as the data bytes in your bandwidth calculations.

- **shaping-rate**—Maximum bandwidth, also known as the *peak information rate (PIR)*. Each fc-set receives a maximum of the configured amount of absolute bandwidth or the configured percentage of bandwidth, even if more bandwidth is available.

  NOTE: Include the preamble bytes and interframe gap bytes as well as the data bytes in your bandwidth calculations.

- **scheduler-map**—Bandwidth and scheduling characteristics for the queues, defined by mapping forwarding classes to schedulers. (The queue scheduling characteristics represent amounts or percentages of the fc-set bandwidth, not the amounts or percentages of total link bandwidth.)

NOTE: Because a port can have more than one fc-set, when you assign resources to an fc-set, keep in mind that the total port bandwidth must serve all of the queues associated with that port.

For example, if you map three fc-sets to a 10-Gigabit Ethernet port, the queues associated with all three of the fc-sets share the 10-Gbps bandwidth as defined by the traffic control profiles. Therefore, the total combined **guaranteed-rate** value of the three fc-sets should not exceed 10 Gbps. If you configure guaranteed rates whose sum exceeds the port bandwidth, the system sends a syslog message to notify you that the configuration is not valid. However, the system does not perform a commit check. If you commit a configuration in which the sum of the guaranteed rates exceeds the port bandwidth, the hierarchical scheduler behaves unpredictably.

The sum of the forwarding class (queue) transmit rates cannot exceed the total **guaranteed-rate** of the fc-set to which the forwarding classes belong. If you configure transmit rates whose sum exceeds the fc-set guaranteed rate, the commit check fails and the system rejects the configuration.

If you configure the **guaranteed-rate** of an fc-set as a percentage, configure all of the transmit rates associated with that fc-set as percentages. In this case, if any of the transmit rates are configured as absolute values instead of percentages, the configuration is not valid and the system sends a syslog message.

Configuring a Traffic Control Profile

This example describes how to configure a traffic control profile named **san-tcp** with a scheduler map named **san-map1** and allocate to it a minimum bandwidth of 4 Gbps and a maximum bandwidth of 8 Gbps:

1. Create the traffic control profile and set the **guaranteed-rate** (minimum guaranteed bandwidth) to **4g**:

```
[edit class-of-service]
user@switch# set traffic-control-profiles san-tcp guaranteed-rate 4g
```

2. Set the **shaping-rate** (maximum guaranteed bandwidth) to **8g**:

```
[edit class-of-service]
user@switch# set traffic-control-profiles san-tcp shaping-rate 8g
```

3. Associate the scheduler map **san-map1** with the traffic control profile:

```
[edit class-of-service]
user@switch# set traffic-control-profiles san-tcp scheduler-map san-map1
```

Verification

Verifying the Traffic Control Profile Configuration

Purpose Verify that you created the traffic control profile **san-tcp** with a minimum guaranteed bandwidth of 4 Gbps, a maximum bandwidth of 8 Gbps, and the scheduler map **san-map1**.

Action List the traffic control profile using the operational mode command **show configuration class-of-service traffic-control-profiles san-tcp**:

```
user@switch> show configuration class-of-service traffic-control-profiles san-tcp
scheduler-map san-map1;
shaping-rate percent 8g;
guaranteed-rate 4g;
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357
- Example: Configuring Queue Schedulers on page 293
- Defining CoS Traffic Control Profiles (Priority Group Scheduling) on page 310
- Understanding CoS Traffic Control Profiles on page 306
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315

Understanding CoS Hierarchical Port Scheduling (ETS)

Scheduling defines the class-of-service (CoS) properties of output queues. Output queues are mapped to forwarding classes. CoS scheduler properties include the amount of interface bandwidth assigned to the queue, the queue priority, and the drop profiles associated with the queue.

Hierarchical port scheduling is a two-tier process that provides better port bandwidth utilization and greater flexibility to allocate resources to queues (forwarding classes) and to groups of queues (forwarding class sets). Hierarchical scheduling includes the Junos OS implementation of enhanced transmission selection (ETS), as described in IEEE 802.1Qaz.

NOTE: All QFX Series devices use ETS scheduling, except for QFX5120, QFX5200, and QFX5210 switches.

Starting with Junos OS 17.3, QFX10000 devices support ETS scheduling.

EX4600 switches use ETS scheduling while EX4650 switches do not.

Video: What is Enhanced Transmission Selection?

This topic describes:

Hierarchical Scheduling Tiers

The two tiers used in hierarchical scheduling are priorities and priority groups, as shown in Table 71 on page 316.

Table 71: Hierarchical Scheduling Tiers

Junos OS Configuration Construct	Equivalent ETS Construct	Description
Forwarding class	Priority	Think about priorities (forwarding classes) as output queues. You map forwarding classes to queues, so each forwarding class represents an output queue. When you use a classifier to map a forwarding class to an IEEE 802.1p code point, the code point identifies that traffic's priority for priority-based flow control (PFC). Thus the forwarding class, the queue mapped to the forwarding class, and the priority (code point) mapped to the forwarding class all identify the same traffic. NOTE: OCX Series switches do not support lossless transport or PFC.
Forwarding class set	Priority group	Priority groups (forwarding class sets) are groups of priorities (forwarding classes). Forwarding class membership in a forwarding class set defines the priority group to which each priority belongs. You can configure up to three unicast priority groups and one multicast priority group.

You apply scheduling properties to each hierarchical scheduling tier as descried in the next section.

NOTE: If you explicitly configure one or more priority groups on an interface, any priority (forwarding class) that is not assigned to a priority group (forwarding class set) on that interface is assigned to an automatically generated default priority group and receives *no bandwidth*. This means that if you configure hierarchical scheduling on an interface, every forwarding class that you want to forward traffic on that interface must belong to a forwarding class set.

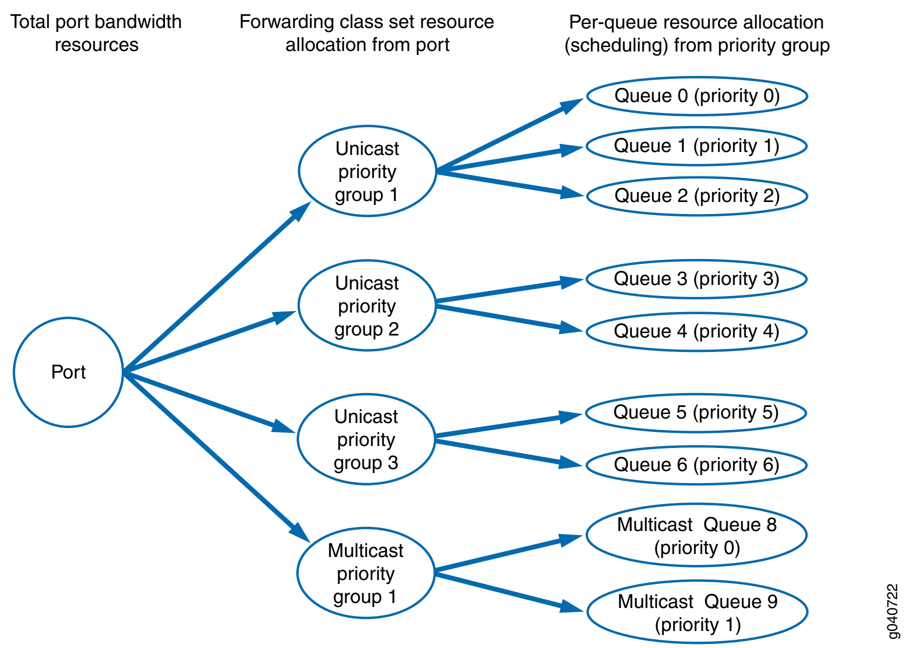
NOTE: On OCX Series switches, by default, classifiers use DSCP code points to map traffic to forwarding classes. However, hierarchical scheduling works in the same manner as when you use IEEE 802.1p code points to classify traffic. The OCX Series classifies traffic into forwarding classes based on DSCP code points, the forwarding classes are mapped to forwarding class sets, and you apply scheduling properties to each of the two tiers.

Hierarchical Scheduling and ETS

Two-tier hierarchical scheduling manages bandwidth efficiently by enabling you to define the CoS properties for each priority group and for each priority. The first tier of the hierarchical scheduler allocates port bandwidth to a priority group. The second tier of the hierarchical scheduler determines the portion of the priority group bandwidth that a priority (queue) can use.

The CoS properties of a priority group define the amount of port bandwidth resources available to the queues in that priority group. The CoS properties you configure for each queue specify the amount of the bandwidth available to the queue from the bandwidth allocated to the priority group. Figure 5 on page 317 shows the relationship of port resource allocation to priority groups, and priority group resource allocation to queues (priorities).

Figure 5: Hierarchical Scheduling Tiers



If a queue (priority) does not use its allocated bandwidth, ETS shares the unused bandwidth among the other queues in the priority group in proportion to the minimum guaranteed rate (transmit rate) scheduled for each queue. If a priority group does not use its allocated bandwidth, ETS shares the unused bandwidth among the priority groups on the port in proportion to the minimum guaranteed rate (guaranteed rate) scheduled for each priority group.

In this way, ETS improves link bandwidth utilization, and it provides each queue and each priority group with the maximum available bandwidth. For example, priorities that consist of bursty traffic can share bandwidth during periods of low traffic transmission, instead of reserving their entire bandwidth allocation when traffic loads are light.

NOTE: The available link bandwidth is the bandwidth remaining after servicing strict-high priority flows. Strict-high priority takes precedence over all other traffic. We recommend that you configure a shaping-rate (transmit-rate on QFX10000 switches) to limit the maximum amount of bandwidth that a strict-high priority forwarding class can use to prevent starving other queues.

ETS Advertisement in DCBX

When you configure hierarchical scheduling on a port, Data Center Bridging Capability Exchange protocol (DCBX) advertises:

- Each priority group
- The priorities in each priority group
- The bandwidth properties of each priority group and priority

When you configure hierarchical scheduling on a port, any priority that is not part of an explicitly configured priority group is assigned to the automatically generated default priority group and receives no bandwidth. The default priority group is transparent. It does not appear in the configuration.

NOTE: **OCX Series switches do support DCBX, so hierarchical scheduling information is not exchanged with connected peers on OCX Series switches.**

Hierarchical Scheduling Process

Hierarchical scheduling consists of multiple configuration steps that create the priorities and the priority groups, schedule their resources, and assign them to interfaces. The steps below correspond to the six blocks in the packet flow diagram shown in Figure 6 on page 319:

1. Packet classification:
   - Configure classification of incoming traffic into forwarding classes (priorities). This consists of either using the default classifiers or configuring classifiers to map code points and loss priorities to the forwarding classes.
   - Apply the classifiers to ingress interfaces or use the default classifiers. Applying a classifier to an interface groups incoming traffic on the interface into forwarding classes and loss priorities, by applying the classifier code point mapping to the incoming traffic.
2. Configure the output queues for the forwarding classes (priorities). This consists of either using the default forwarding classes and forwarding-class-to-queue mapping, or creating your own forwarding classes and mapping them to output queues.
3. Allocate resources to the forwarding classes:
   - Define resources for the priorities. This consists of configuring schedulers to set minimum guaranteed bandwidth, maximum bandwidth, drop profiles for Weighted Random Early Detection (WRED), and bandwidth priority to apply to a forwarding class. Extra bandwidth is shared among queues in proportion to the minimum guaranteed bandwidth (transmit rate) of each queue.
   - Map resources to priorities. This consists of mapping forwarding classes to schedulers, using a scheduler map.

4. Configure priority groups. This consists of mapping forwarding classes (priorities) to forwarding class sets (priority groups) to define the priorities that belong to each priority group.

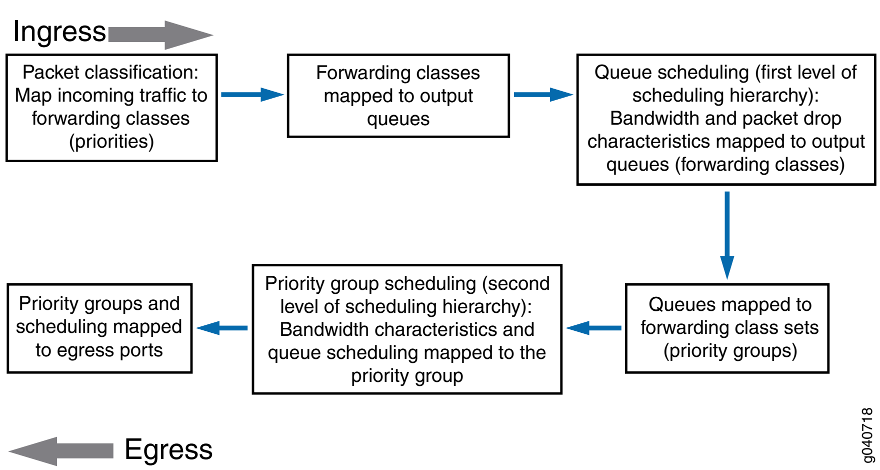
5. Define resources for the priority groups. This consists of configuring traffic control profiles to set minimum guaranteed bandwidth (guaranteed-rate) and maximum bandwidth (shaping-rate on switches other than QFX10000 switches, transmit-rate on QFX10000 switches) for a priority group. Traffic control profiles also specify a scheduler map, which defines the resources (schedulers) mapped to the priorities in the priority group. Extra port bandwidth is shared among priority groups in proportion to the minimum guaranteed bandwidth of each priority group.

   The traffic control profile bandwidth settings determine the port resources available to the priority group. The schedulers specified in the scheduler map determine the amount of priority group resources that each priority receives.

   NOTE: **QFX10000 switches do not support defining a shaping rate for priority groups. Instead, set the maximum bandwidth for a priority group by defining a transmit rate. See transmit-rate.**

6. Apply hierarchical scheduling to a port. This consists of attaching one or more priority groups (forwarding class sets) to an interface. For each priority group, you also attach a traffic control profile, which contains the scheduling properties of the priority group and the priorities in the priority group. Different priority groups on the same port can use different traffic control profiles, which provides fine tuned control of scheduling for each queue on each interface.

Figure 6: Hierarchical Scheduling Packet Flow



Strict-High Priority Queues and Hierarchical Scheduling

If you configure a strict-high priority queue, you must observe the following rules:

- You must create a separate forwarding class set (priority group) for the strict-high priority queue.
- Only one forwarding class set can contain strict-high priority queues.
- Strict-high priority queues cannot belong to the same forwarding class set as queues that are not strict-high priority.
- A strict-high priority queue cannot belong to a multidestination forwarding class set.
- We recommend that you always apply a **shaping-rate** (**transmit-rate** on QFX10000 switches) to strict-high priority queues to limit the amount of bandwidth a strict-high priority queue can use. If you do not limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

NOTE: **On a QFabric system, if a fabric (fte) interface handles strict-high priority traffic, you must define a separate forwarding class set (priority group) for strict-high priority traffic. Strict-high priority traffic cannot be mixed with traffic of other priorities in a forwarding class set. For example, you might choose to create different forwarding class sets for best effort, lossless, strict-high priority, and multidestination traffic.**

Default Hierarchical Scheduling

NOTE: **There is no default hierarchical scheduling on QFX10000 switches. QFX10000 switches use port scheduling by default, and you must explicitly configure hierarchical scheduling to enable ETS.**

If you do not explicitly configure hierarchical scheduling, the switch uses the default settings:

- The switch automatically creates a default forwarding class set that contains all of the forwarding classes on the switch. The switch assigns 100 percent of the port output bandwidth to the default forwarding class set. The default forwarding class set is transparent. It does not appear in the configuration and is used for Data Center Bridging Capability Exchange protocol (DCBX) advertisement.

  NOTE: **OCX Series switches do not support DCBX, so the ETS configuration is not advertised to connected peers.**

- Ingress traffic is classified based on the default classifier settings.

- The forwarding classes (queues) in the default forwarding class set receive bandwidth based on the default scheduler settings.

Release History Table

Release	Description
17.3	Starting with Junos OS 17.3, QFX10000 devices support ETS scheduling.

Related Documentation

- Understanding CoS Packet Flow on page 23
- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS Priority Group Scheduling on page 307
- *Benefits of Configuring CoS Hierarchical Port Scheduling*
- Understanding CoS Flow Control (Ethernet PAUSE and PFC) on page 525
- *Understanding CoS Classifiers*
- *Understanding Default CoS Scheduling and Classification*
- *Understanding CoS Scheduling on QFabric System Node Device Fabric (fte) Ports*
- *Understanding Default CoS Scheduling on QFabric System Interconnect Devices (Junos OS Release 13.1 and Later Releases)*
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357

Example: Configuring CoS Hierarchical Port Scheduling (ETS)

Hierarchical port scheduling defines the class-of-service (CoS) properties of output queues, which are mapped to forwarding classes. Traffic is classified into forwarding classes based on code point (priority), so mapping queues to forwarding classes also maps queues to priorities). Hierarchical port scheduling enables you to group priorities that require similar CoS treatment into priority groups. You define the port bandwidth resources for a priority group, and you define the amount of the priority group's resources that each priority in the group can use.

Hierarchical port scheduling is the Junos OS implementation of enhanced transmission selection (ETS), as described in IEEE 802.1Qaz. One major benefit of hierarchical port scheduling is greater port bandwidth utilization. If a priority group on a port does not use all of its allocated bandwidth, other priority groups on that port can use that bandwidth. Also, if a priority within a priority group does not use its allocated bandwidth, other priorities within that priority group can use that bandwidth.

Configuring hierarchical scheduling is a multistep procedure that includes:

- Mapping forwarding classes to queues
- Defining forwarding class sets (priority groups)
- Defining behavior aggregate classifiers
- Configuring priority-based flow control (PFC) for lossless priorities (queues)
- Applying classifiers and PFC configuration to ingress interfaces
- Defining drop profiles
- Defining schedulers
- Mapping forwarding classes to schedulers
- Defining traffic control profiles
- Assigning priority groups and traffic control profiles to egress ports

NOTE: OCX Series switches do not support lossless transport and do not support PFC. Although this example includes configuring lossless transport with PFC, the portions of the example that do not pertain to lossless transport still apply to OCX Series switches. (You can configure hierarchical scheduling on OCX Series switches, but you cannot configure lossless transport or lossless forwarding classes.)

This example describes how to configure hierarchical scheduling:

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

Keep the following considerations in mind when you plan the port bandwidth allocation for priority groups and for individual priorities:

- How much traffic and what types of traffic you expect to traverse the system.
- How you want to divide different types of traffic into priorities (forwarding classes) to apply different CoS treatment to different types of traffic. Dividing traffic into priorities includes:
  - Mapping the code points of ingress traffic to forwarding classes using behavior aggregate (BA) classifiers. This classifies incoming traffic into the appropriate forwarding class based on code point.
  - Mapping forwarding classes to output queues. This defines the output queue for each type of traffic.
  - Attaching the BA classifier to the desired ingress interfaces so that incoming traffic maps to the desired forwarding classes and queues.
- How you want to organize priorities into priority groups (forwarding class sets).

  Traffic that requires similar treatment usually belongs in the same priority group. To do this, place forwarding classes that require similar bandwidth, loss, and other characteristics in the same forwarding class set. For example, you can map all types of best-effort traffic forwarding classes into one forwarding class set.
- How much of the port bandwidth you want to allocate to each priority group and to each of the priorities in each priority group. The following considerations apply to bandwidth allocation:
  - Estimate how much traffic you expect in each forwarding class, and how much traffic you expect in each forwarding class set (the amount of traffic you expect in a forwarding class set is the aggregate amount of traffic in the forwarding classes that belong to the forwarding class set).
  - The combined minimum guaranteed bandwidth of the priorities (forwarding classes) in a priority group should not exceed the minimum guaranteed bandwidth of the priority group (forwarding class set). The transmit rate scheduler parameter defines the minimum guaranteed bandwidth for forwarding classes. Scheduler maps associate schedulers with forwarding classes.
  - The combined minimum guaranteed bandwidth of the priority groups (forwarding class sets) on a port should not exceed the port's total bandwidth. The guaranteed rate parameter in the traffic control profile defines the minimum bandwidth for a forwarding class set. Associating a scheduler map with a traffic control profile sets the scheduling for the individual forwarding classes in the forwarding class set.

This example creates hierarchical port scheduling by defining priority groups for best effort, guaranteed delivery, and high-performance computing (HPC) traffic. Each priority group includes priorities that need to receive similar CoS treatment. Each priority group and each priority within each priority group receive the CoS resources needed to service their flows. Lossless priorities use PFC to prevent packet loss when the network experiences congestion.

Topology

Table 72 on page 324 shows the configuration components for this example.

NOTE: OCX Series switches do not support lossless transport and do not support PFC. If you eliminate the configuration elements for the default lossless **fcoe** and **no-loss** forwarding classes (including classifier, forwarding class set, scheduler, and traffic control profile configuration for those forwarding classes) and for PFC, this example works for OCX Series switches. However, because the default **fcoe** and **no-loss** forwarding classes do not carry traffic on OCX Series switches, you can apply the bandwidth allocated to those forwarding classes to other forwarding classes. By default, the active forwarding classes (**best-effort**, **network-control**, and **mcast**) share the unused bandwidth assigned to the **fcoe** and **no-loss** forwarding classes.

Table 72: Components of the Hierarchical Port Scheduling (ETS) Configuration Topology

Property	Settings
Hardware	QFX3500 switch
Mapping of forwarding classes (priorities) to queues	**best-effort** to queue **0** **be2** to queue **1** **fcoe** (Fibre Channel over Ethernet) to queue **3** **no-loss** to queue **4** **hpc** (high-performance computing) to queue **5** **network-control** to queue **7** NOTE: On switches that do not support the ELS CLI, if you are using Junos OS Release 12.2 or later, use the default forwarding-class-to-queue mapping for the lossless **fcoe** and **no-loss** forwarding classes. If you explicitly configure the default lossless forwarding classes, the traffic mapped to those forwarding classes is treated as lossy (**best-effort**) traffic and does *not* receive lossless treatment. On switches that do not support the ELS CLI, in Junos OS Release 12.3 and later, you can include the *no-loss* packet drop attribute in the explicit forwarding class configuration to configure a lossless forwarding class.
Forwarding class sets (priority groups)	**best-effort-pg**: contains forwarding classes **best-effort**, **be2**, and **network control** **guar-delivery-pg**: contains forwarding classes **fcoe** and **no-loss** **hpc-pg**: contains forwarding class **hpc**

Table 72: Components of the Hierarchical Port Scheduling (ETS) Configuration Topology (continued)

Property	Settings
Behavior aggregate classifier (maps forwarding classes and loss priorities to incoming packets by IEEE 802.1 code point)	Name—**hsclassifier1** Code point mapping: • **000** to forwarding class **best-effort** and loss priority **low** • **001** to forwarding class **be2** and loss priority **high** • **011** to forwarding class **fcoe** and loss priority **low** • **100** to forwarding class **no-loss** and loss priority **low** • **101** to forwarding class **hpc** and loss priority **low** • **110** to forwarding class **network-control** and loss priority **low**
PFC	Congestion notification profile name—**gd-cnp** PFC enabled on code points: **011** (**fcoe** priority), **010** (**no-loss** priority)
Drop profiles NOTE: The **fcoe** and **no-loss** priorities (queues) do not use drop profiles because they are lossless traffic classes.	**dp-be-low**: drop start point **25**, drop end point **50**, maximum drop rate **80** **dp-be-high**: drop start point **10**, drop end point **40**, maximum drop rate **100** **dp-hpc**: drop start point **75**, drop end point **90**, maximum drop rate **75** **dp-nc**: drop start point **80**, drop end point **100**, maximum drop rate **100**
Queue schedulers	**be-sched**: minimum bandwidth **3g**, maximum bandwidth **100%**, priority **low**, drop profiles **dp-be-low** and **dp-be-high** **fcoe-sched**: minimum bandwidth **2.5g**, maximum bandwidth **100%**, priority **low** **hpc-sched**: minimum bandwidth **2g**, maximum bandwidth **100%**, priority **low**, drop profile **dp-hpc** **nc-sched**: minimum bandwidth **500m**, maximum bandwidth **100%**, priority **low**, drop profile **dp-nc** **nl-sched**: minimum bandwidth **2g**, maximum bandwidth **100%**, priority **low**
Forwarding class-to-scheduler mapping	Scheduler map **be-map**: Forwarding class **best-effort**, scheduler **be-sched** Forwarding class **be2**, scheduler **be-sched** Forwarding class **network-control**, scheduler **nc-sched** Scheduler map **gd-map**: Forwarding class **fcoe**, scheduler **fcoe-sched** Forwarding class **no-loss**, scheduler **nl-sched** Scheduler map **hpc-map**: Forwarding class **hpc**, scheduler **hpc-sched**
Traffic control profiles	**be-tcp**: scheduler map **be-map**, minimum bandwidth **3.5g**, maximum bandwidth **100%** **gd-tcp**: scheduler map **gd-map**, minimum bandwidth **4.5g**, maximum bandwidth **100%** **hpc-tcp**: scheduler map **hpc-map**, minimum bandwidth **2g**, maximum bandwidth **100%**

Table 72: Components of the Hierarchical Port Scheduling (ETS) Configuration Topology (continued)

Property	Settings
Interfaces	This example configures hierarchical port scheduling on interfaces **xe-0/0/20** and **xe-0/0/21**. Because traffic is bidirectional, you apply the ingress and egress configuration components to both interfaces: • Classifier Name—**hsclassifier1** • Forwarding class sets—**best-effort-pg**, **guar-deliver-pg**, **hpc-pg** • Congestion notification profile—**gd-cnp**

Figure 7 on page 326 shows a block diagram of the configuration components and the configuration flow of the CLI statements used in the example. You can perform the configuration steps in a different sequence if you want.

Figure 7: Hierarchical Port Scheduling Components Block Diagram

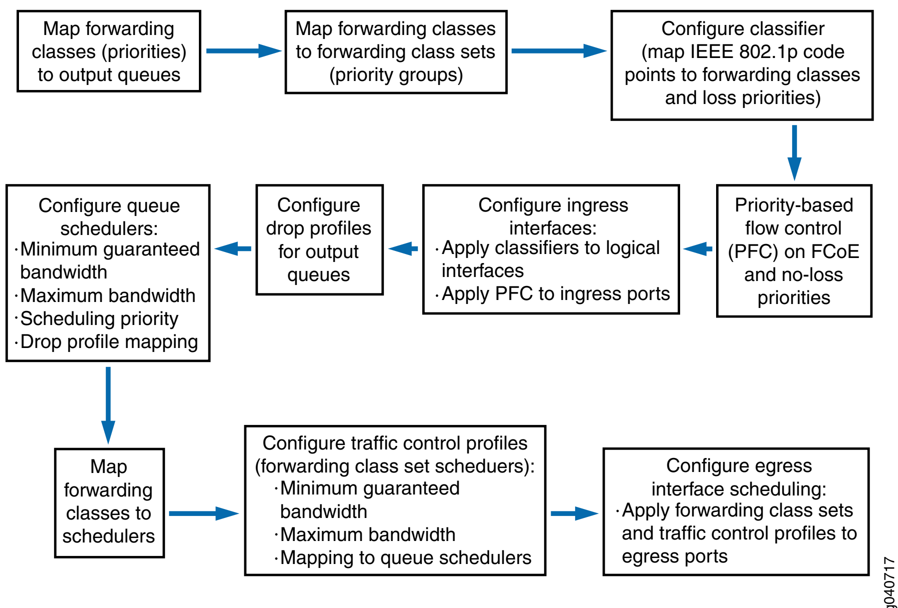
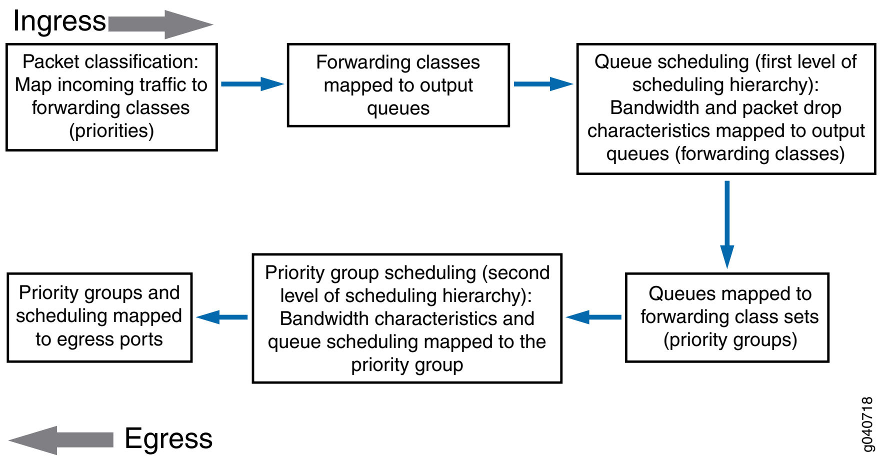


Figure 8 on page 327 shows a block diagram of the hierarchical scheduling packet flow from ingress to egress.

Figure 8: Hierarchical Port Scheduling Packet Flow Block Diagram



Configuration

CLI Quick Configuration

To quickly configure hierarchical port scheduling on systems that support lossless transport, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit class-of-service] hierarchy level:

```
[edit class-of-service]
set forwarding-classes class best-effort queue-num 0
set forwarding-classes class be2 queue-num 1
set forwarding-classes class hpc queue-num 5
set forwarding-classes class network-control queue-num 7
set forwarding-class-sets best-effort-pg class best-effort
set forwarding-class-sets best-effort-pg class be2
set forwarding-class-sets best-effort-pg class network-control
set forwarding-class-sets guar-delivery-pg class fcoe
set forwarding-class-sets guar-delivery-pg class no-loss
set forwarding-class-sets hpc-pg class hpc
set classifiers ieee-802.1 hsclassifier1 forwarding-class best-effort loss-priority low code-points
000
set classifiers ieee-802.1 hsclassifier1 forwarding-class be2 loss-priority high code-points 001
set classifiers ieee-802.1 hsclassifier1 forwarding-class fcoe loss-priority low code-points 011
set classifiers ieee-802.1 hsclassifier1 forwarding-class no-loss loss-priority low code-points 100
set classifiers ieee-802.1 hsclassifier1 forwarding-class hpc loss-priority low code-points 101
set classifiers ieee-802.1 hsclassifier1 forwarding-class network-control loss-priority low
code-points 110
set congestion-notification-profile gd-cnp input ieee-802.1 code-point 011 pfc
set congestion-notification-profile gd-cnp input ieee-802.1 code-point 100 pfc
set interfaces xe-0/0/20 unit 0 classifiers ieee-802.1 hsclassifier1
set interfaces xe-0/0/21 unit 0 classifiers ieee-802.1 hsclassifier1
set interfaces xe-0/0/20 congestion-notification-profile gd-cnp
set interfaces xe-0/0/21 congestion-notification-profile gd-cnp
set drop-profiles dp-be-low interpolate fill-level 25 fill-level 50 drop-probability 0 drop-probability
80
set drop-profiles dp-be-high interpolate fill-level 10 fill-level 40 drop-probability 0 drop-probability
100
set drop-profiles dp-nc interpolate fill-level 80 fill-level 100 drop-probability 0 drop-probability
100
```

```
set drop-profiles dp-hpc interpolate fill-level 75 fill-level 90 drop-probability 0 drop-probability 75
set schedulers be-sched priority low transmit-rate 3g
set schedulers be-sched shaping-rate percent 100
set schedulers be-sched drop-profile-map loss-priority low protocol any drop-profile dp-be-low
set schedulers be-sched drop-profile-map loss-priority high protocol any drop-profile dp-be-high
set schedulers fcoe-sched priority low transmit-rate 2500m
set schedulers fcoe-sched shaping-rate percent 100
set schedulers hpc-sched priority low transmit-rate 2g
set schedulers hpc-sched shaping-rate percent 100
set schedulers hpc-sched drop-profile-map loss-priority low protocol any drop-profile dp-hpc
set schedulers nc-sched priority low transmit-rate 500m
set schedulers nc-sched shaping-rate percent 100
set schedulers nc-sched drop-profile-map loss-priority low protocol any drop-profile dp-nc
set schedulers nl-sched priority low transmit-rate 2g
set schedulers nl-sched shaping-rate percent 100
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set scheduler-maps be-map forwarding-class be2 scheduler be-sched
set scheduler-maps be-map forwarding-class network-control scheduler nc-sched
set scheduler-maps gd-map forwarding-class fcoe scheduler fcoe-sched
set scheduler-maps gd-map forwarding-class no-loss scheduler nl-sched
set scheduler-maps hpc-map forwarding-class hpc scheduler hpc-sched
set traffic-control-profiles be-tcp scheduler-map be-map guaranteed-rate 3500m
set traffic-control-profiles be-tcp shaping-rate percent 100
set traffic-control-profiles gd-tcp scheduler-map gd-map guaranteed-rate 4500m
set traffic-control-profiles gd-tcp shaping-rate percent 100
set traffic-control-profiles hpc-tcp scheduler-map hpc-map guaranteed-rate 2g
set traffic-control-profiles hpc-tcp shaping-rate percent 100
set interfaces xe-0/0/20 forwarding-class-set best-effort-pg output-traffic-control-profile be-tcp
set interfaces xe-0/0/20 forwarding-class-set guar-delivery-pg output-traffic-control-profile gd-tcp
set interfaces xe-0/0/20 forwarding-class-set hpc-pg output-traffic-control-profile hpc-tcp
set interfaces xe-0/0/21 forwarding-class-set best-effort-pg output-traffic-control-profile be-tcp
set interfaces xe-0/0/21 forwarding-class-set guar-delivery-pg output-traffic-control-profile gd-tcp
set interfaces xe-0/0/21 forwarding-class-set hpc-pg output-traffic-control-profile hpc-tcp
```

OCX Series Switches

Because OCX Series switches do not support lossless transport, the following subset of the configuration eliminates the lossless configuration elements and provides hierarchical port scheduling for the best-effort, be2, hpc, and network-control forwarding classes. In addition, on OCX Series switches, you would probably use DSCP classifiers and code points instead of IEEE classifiers and code points. To quickly configure hierarchical port scheduling on an OCX Series switch, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit class-of-service] hierarchy level:

```
[edit class-of-service]
set forwarding-classes class best-effort queue-num 0
set forwarding-classes class be2 queue-num 1
set forwarding-classes class hpc queue-num 5
set forwarding-classes class network-control queue-num 7
set forwarding-class-sets best-effort-pg class best-effort
set forwarding-class-sets best-effort-pg class be2
set forwarding-class-sets best-effort-pg class network-control
set forwarding-class-sets hpc-pg class hpc
```



```
set classifiers ieee-802.1 hsclassifier1 forwarding-class best-effort loss-priority low code-points 000
set classifiers ieee-802.1 hsclassifier1 forwarding-class be2 loss-priority high code-points 001
set classifiers ieee-802.1 hsclassifier1 forwarding-class hpc loss-priority low code-points 101
set classifiers ieee-802.1 hsclassifier1 forwarding-class network-control loss-priority low code-points 110
set interfaces xe-0/0/20 unit 0 classifiers ieee-802.1 hsclassifier1
set interfaces xe-0/0/21 unit 0 classifiers ieee-802.1 hsclassifier1
set drop-profiles dp-be-low interpolate fill-level 25 fill-level 50 drop-probability 0 drop-probability 80
set drop-profiles dp-be-high interpolate fill-level 10 fill-level 40 drop-probability 0 drop-probability 100
set drop-profiles dp-nc interpolate fill-level 80 fill-level 100 drop-probability 0 drop-probability 100
set drop-profiles dp-hpc interpolate fill-level 75 fill-level 90 drop-probability 0 drop-probability 75
set schedulers be-sched priority low transmit-rate 3g
set schedulers be-sched shaping-rate percent 100
set schedulers be-sched drop-profile-map loss-priority low protocol any drop-profile dp-be-low
set schedulers be-sched drop-profile-map loss-priority high protocol any drop-profile dp-be-high
set schedulers hpc-sched priority low transmit-rate 2g
set schedulers hpc-sched shaping-rate percent 100
set schedulers hpc-sched drop-profile-map loss-priority low protocol any drop-profile dp-hpc
set schedulers nc-sched priority low transmit-rate 500m
set schedulers nc-sched shaping-rate percent 100
set schedulers nc-sched drop-profile-map loss-priority low protocol any drop-profile dp-nc
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set scheduler-maps be-map forwarding-class be2 scheduler be-sched
set scheduler-maps be-map forwarding-class network-control scheduler nc-sched
set scheduler-maps hpc-map forwarding-class hpc scheduler hpc-sched
set traffic-control-profiles be-tcp scheduler-map be-map guaranteed-rate 3500m
set traffic-control-profiles be-tcp shaping-rate percent 100
set traffic-control-profiles hpc-tcp scheduler-map hpc-map guaranteed-rate 2g
set traffic-control-profiles hpc-tcp shaping-rate percent 100
set interfaces xe-0/0/20 forwarding-class-set best-effort-pg output-traffic-control-profile be-tcp
set interfaces xe-0/0/20 forwarding-class-set hpc-pg output-traffic-control-profile hpc-tcp
set interfaces xe-0/0/21 forwarding-class-set best-effort-pg output-traffic-control-profile be-tcp
set interfaces xe-0/0/21 forwarding-class-set hpc-pg output-traffic-control-profile hpc-tcp
```

Step-by-Step Procedure

To perform a step-by-step configuration of the forwarding classes (priorities), forwarding class sets (priority groups), classifiers, queue schedulers, PFC, traffic control profiles, and interfaces to set up hierarchical port scheduling (ETS):

1. Configure the forwarding classes (priorities) and map them to unicast output queues (do not explicitly map the **fcoe** and **no-loss** forwarding classes to output queues; use the default configuration):

   ```
   [edit class-of-service]
   user@switch# set forwarding-classes class best-effort queue-num 0
   user@switch# set forwarding-classes class be2 queue-num 1
   user@switch# set forwarding-classes class hpc queue-num 5
   user@switch# set forwarding-classes class network-control queue-num 7
   ```

2. Configure forwarding class sets (priority groups) to group forwarding classes (priorities) that require similar CoS treatment:

```
[edit class-of-service]
user@switch# set forwarding-class-sets best-effort-pg class best-effort
user@switch# set forwarding-class-sets best-effort-pg class be2
user@switch# set forwarding-class-sets best-effort-pg class network-control
user@switch# set forwarding-class-sets guar-delivery-pg class fcoe
user@switch# set forwarding-class-sets guar-delivery-pg class no-loss
user@switch# set forwarding-class-sets hpc-pg class hpc
```

NOTE: On OCX Series switches, you would not configure the **guar-delivery-pg** forwarding class set for lossless traffic.

3. Configure a classifier to set the loss priority and IEEE 802.1 code points assigned to each forwarding class at the ingress:

```
[edit class-of-service]
user@switch# set classifiers ieee-802.1 hsclassifier1 forwarding-class best-effort loss-priority low code-points 000
user@switch# set classifiers ieee-802.1 hsclassifier1 forwarding-class be2 loss-priority high code-points 001
user@switch# set classifiers ieee-802.1 hsclassifier1 forwarding-class fcoe loss-priority low code-points 011
user@switch# set classifiers ieee-802.1 hsclassifier1 forwarding-class no-loss loss-priority low code-points 100
user@switch# set classifiers ieee-802.1 hsclassifier1 forwarding-class hpc loss-priority low code-points 101
user@switch# set classifiers ieee-802.1 hsclassifier1 forwarding-class network-control loss-priority low code-points 110
```

NOTE: On OCX Series switches, you would not configure the **fcoe** and **no-loss** portions of the classifier.

4. Configure a congestion notification profile to enable PFC on the FCoE and no-loss queue IEEE 802.1 code points:

```
[edit class-of-service]
user@switch# set congestion-notification-profile gd-cnp input ieee-802.1 code-point 011 pfc
user@switch# set congestion-notification-profile gd-cnp input ieee-802.1 code-point 100 pfc
```

NOTE: This step does not apply to OCX Series switches, which do not support PFC.

5. Assign the classifier to the interfaces:


```
[edit class-of-service]
user@switch# set interfaces xe-0/0/20 unit 0 classifiers ieee-802.1 hsclassifier1
user@switch# set interfaces xe-0/0/21 unit 0 classifiers ieee-802.1 hsclassifier1
```

6. Apply the PFC configuration to the interfaces:

```
[edit class-of-service]
user@switch# set interfaces xe-0/0/20 congestion-notification-profile gd-cnp
user@switch# set interfaces xe-0/0/21 congestion-notification-profile gd-cnp
```

NOTE: This step does not apply to OCX Series switches, which do not support PFC.

7. Configure the drop profile for the best-effort low loss-priority queue:

```
[edit class-of-service]
user@switch# set drop-profiles dp-be-low interpolate fill-level 25 fill-level 50
drop-probability 0 drop-probability 80
```

8. Configure the drop profile for the best-effort high loss-priority queue:

```
[edit class-of-service]
user@switch# set drop-profiles dp-be-high interpolate fill-level 10 fill-level 40
drop-probability 0 drop-probability 100
```

9. Configure the drop profile for the network-control queue:

```
[edit class-of-service]
user@switch# set drop-profiles dp-nc interpolate fill-level 80 fill-level 100 drop-probability
0 drop-probability 100
```

10. Configure the drop profile for the high-performance computing queue:

```
[edit class-of-service]
user@switch# set drop-profiles dp-hpc interpolate fill-level 75 fill-level 90 drop-probability
0 drop-probability 75
```

11. Define the minimum guaranteed bandwidth, priority, maximum bandwidth, and drop profiles for the best-effort queue:

```
[edit class-of-service]
user@switch# set schedulers be-sched priority low transmit-rate 3g
user@switch# set schedulers be-sched shaping-rate percent 100
user@switch# set schedulers be-sched drop-profile-map loss-priority low protocol any
drop-profile dp-be-low
user@switch# set schedulers be-sched drop-profile-map loss-priority high protocol any
drop-profile dp-be-high
```

12. Define the minimum guaranteed bandwidth, priority, and maximum bandwidth for the FCoE queue:

    [edit class-of-service]
    user@switch# **set schedulers fcoe-sched priority low transmit-rate 2500m**
    user@switch# **set schedulers fcoe-sched shaping-rate percent 100**

    NOTE: This step does not apply to OCX Series switches, which do not support lossless transport.

13. Define the minimum guaranteed bandwidth, priority, maximum bandwidth, and drop profile for the high-performance computing queue:

    [edit class-of-service]
    user@switch# **set schedulers hpc-sched priority low transmit-rate 2g**
    user@switch# **set schedulers hpc-sched shaping-rate percent 100**
    user@switch# **set schedulers hpc-sched drop-profile-map loss-priority low protocol any drop-profile dp-hpc**

14. Define the minimum guaranteed bandwidth, priority, maximum bandwidth, and drop profile for the network-control queue:

    [edit class-of-service]
    user@switch# **set schedulers nc-sched priority low transmit-rate 500m**
    user@switch# **set schedulers nc-sched shaping-rate percent 100**
    user@switch# **set schedulers nc-sched drop-profile-map loss-priority low protocol any drop-profile dp-nc**

15. Define the minimum guaranteed bandwidth, priority, and maximum bandwidth for the no-loss queue:

    [edit class-of-service]
    user@switch# **set schedulers nl-sched priority low transmit-rate 2g**
    user@switch# **set schedulers nl-sched shaping-rate percent 100**

    NOTE: This step does not apply to OCX Series switches, which do not support lossless transport.

16. Map the schedulers to the appropriate forwarding classes (queues):

    [edit class-of-service]
    user@switch# **set scheduler-maps be-map forwarding-class best-effort scheduler be-sched**
    user@switch# **set scheduler-maps be-map forwarding-class be2 scheduler be-sched**
    user@switch# **set scheduler-maps be-map forwarding-class network-control scheduler nc-sched**
    user@switch# **set scheduler-maps gd-map forwarding-class fcoe scheduler fcoe-sched**


```
user@switch# set scheduler-maps gd-map forwarding-class no-loss scheduler nl-sched
user@switch# set scheduler-maps hpc-map forwarding-class hpc scheduler hpc-sched
```

NOTE: On OCX Series switches, because lossless transport is not supported, you would not configure the **gd-map** scheduler map.

17. Define the traffic control profile for the best-effort priority group (queue scheduler to mapping, minimum guaranteed bandwidth, and maximum bandwidth):

```
[edit class-of-service]
user@switch# set traffic-control-profiles be-tcp scheduler-map be-map guaranteed-rate 3500m
user@switch# set traffic-control-profiles be-tcp shaping-rate percent 100
```

18. Define the traffic control profile for the guaranteed delivery priority group (queue to scheduler mapping, minimum guaranteed bandwidth, and maximum bandwidth):

```
[edit class-of-service]
user@switch# set traffic-control-profiles gd-tcp scheduler-map gd-map guaranteed-rate 4500m
user@switch# set traffic-control-profiles gd-tcp shaping-rate percent 100
```

NOTE: This step does not apply to OCX Series switches, which do not support lossless transport.

19. Define the traffic control profile for the high-performance computing priority group (queue to scheduler mapping, minimum guaranteed bandwidth, and maximum bandwidth):

```
[edit class-of-service]
user@switch# set traffic-control-profiles hpc-tcp scheduler-map hpc-map guaranteed-rate 2g
user@switch# set traffic-control-profiles hpc-tcp shaping-rate percent 100
```

20. Apply the three priority groups (forwarding class sets) and the appropriate traffic control profiles to the egress ports:

```
[edit class-of-service]
user@switch# set interfaces xe-0/0/20 forwarding-class-set best-effort-pg output-traffic-control-profile be-tcp
user@switch# set interfaces xe-0/0/20 forwarding-class-set guar-delivery-pg output-traffic-control-profile gd-tcp
user@switch# set interfaces xe-0/0/20 forwarding-class-set hpc-pg output-traffic-control-profile hpc-tcp
user@switch# set interfaces xe-0/0/21 forwarding-class-set best-effort-pg output-traffic-control-profile be-tcp
```


user@switch# **set interfaces xe-0/0/21 forwarding-class-set guar-delivery-pg output-traffic-control-profile gd-tcp**
user@switch# **set interfaces xe-0/0/21 forwarding-class-set hpc-pg output-traffic-control-profile hpc-tcp**

NOTE: Because OCX Series switches do not support lossless transport, on OCX Series switches, you would not apply the **guar-deliver-pg** forwarding class set and the **gd-tcp** traffic control profile to interfaces.

Results

Display the results of the configuration (the system shows only the explicitly configured parameters; it does not show default parameters such as the **fcoe** and **no-loss** lossless forwarding classes). On OCX Series switches, you would not see the lossless configuration components in the output:

```
user@switch> show configuration class-of-service
classifiers {
  ieee-802.1 hsclassifier1 {
    forwarding-class best-effort {
      loss-priority low code-points 000;
    }
    forwarding-class be2 {
      loss-priority high code-points 001;
    }
    forwarding-class fcoe {
      loss-priority low code-points 011;
    }
    forwarding-class no-loss {
      loss-priority low code-points 100;
    }
    forwarding-class hpc {
      loss-priority low code-points 101;
    }
    forwarding-class network-control {
      loss-priority low code-points 110;
    }
  }
  drop-profiles {
    dp-be-low {
      interpolate {
        fill-level [ 25 50 ];
        drop-probability [ 0 80 ];
      }
    }
    dp-be-high {
      interpolate {
        fill-level [ 10 40 ];
        drop-probability [ 0 100 ];
      }
```

```
    }
    dp-hpc {
        interpolate {
            fill-level [ 75 90 ];
            drop-probability [ 0 75 ];
        }
    }
    dp-nc {
        interpolate {
            fill-level [ 80 100 ];
            drop-probability [ 0 100 ];
        }
    }
}
forwarding-classes {
    class best-effort queue-num 0;
    class be2 queue-num 1;
    class hpc queue-num 5;
    class network-control queue-num 7;
}
traffic-control-profiles {
    be-tcp {
        scheduler-map be-map;
        shaping-rate percent 100;
        guaranteed-rate 3500000000;
    }
    gd-tcp {
        scheduler-map gd-map;
        shaping-rate percent 100;
        guaranteed-rate 4500000000;
    }
    hpc-tcp {
        scheduler-map hpc-map;
        shaping-rate percent 100;
        guaranteed-rate 2g;
    }
}
forwarding-class-sets {
    guar-delivery-pg {
        class fcoe;
        class no-loss;
    }
    best-effort-pg {
        class best-effort;
        class be2;
        class network-control;
    }
    hpc-pg {
        class hpc;
    }
}
congestion-notification-profile {
    gd-cnp {
        input {
            ieee-802.1 {
                code-point 011 {
```

```
                pfc;
            }
            code-point 100 {
                pfc;
            }
        }
    }
}
}
interfaces {
    xe-0/0/20 {
        forwarding-class-set {
            best-effort-pg {
                output-traffic-control-profile be-tcp;
            }
            guar-delivery-pg {
                output-traffic-control-profile gd-tcp;
            }
            hpc-pg {
                output-traffic-control-profile hpc-tcp;
            }
        }
        congestion-notification-profile gd-cnp;
        unit 0 {
            classifiers {
                ieee-802.1 hsclassifier1;
            }
        }
    }
    xe-0/0/21 {
        forwarding-class-set {
            best-effort-pg {
                output-traffic-control-profile be-tcp;
            }
            guar-delivery-pg {
                output-traffic-control-profile gd-tcp;
            }
            hpc-pg {
                output-traffic-control-profile hpc-tcp;
            }
        }
        congestion-notification-profile gd-cnp;
        unit 0 {
            classifiers {
                ieee-802.1 hsclassifier1;
            }
        }
    }
}
scheduler-maps {
    be-map {
        forwarding-class best-effort scheduler be-sched;
        forwarding-class network-control scheduler nc-sched;
        forwarding-class be2 scheduler be-sched;
    }
    gd-map {
```



```
        forwarding-class fcoe scheduler fcoe-sched;
        forwarding-class no-loss scheduler nl-sched;
    }
    hpc-map {
        forwarding-class hpc scheduler hpc-sched;
    }
}
schedulers {
    be-sched {
        transmit-rate 3g;
        shaping-rate percent 100;
        priority low;
        drop-profile-map loss-priority low protocol any drop-profile dp-be-low;
        drop-profile-map loss-priority high protocol any drop-profile dp-be-high;
    }
    fcoe-sched {
        transmit-rate 2500000000;
        shaping-rate percent 100;
        priority low;
    }
    hpc-sched {
        transmit-rate 2g;
        shaping-rate percent 100;
        priority low;
        drop-profile-map loss-priority low protocol any drop-profile dp-hpc;
    }
    nc-sched {
        transmit-rate 500m;
        shaping-rate percent 100;
        priority low;
        drop-profile-map loss-priority low protocol any drop-profile dp-nc;
    }
    nl-sched {
        transmit-rate 2g;
        shaping-rate percent 100;
        priority low;
    }
}
```

TIP: To quickly configure the interfaces, issue the **load merge terminal** command, and then copy the hierarchy and paste it into the switch terminal window.

Verification

NOTE: The verification output is based on the full example configuration. On OCX Series switches, you do not see lossless configuration components in the output. Comments about lossless configuration components do not apply to OCX Series switches.

To verify that you created the hierarchical port scheduling components and they are operating properly, perform these tasks:

Verifying the Forwarding Classes (Priorities)

Purpose Verify that you created the forwarding classes and mapped them to the correct queues. (The system shows only the explicitly configured forwarding classes. It does not show default forwarding classes such as **fcoe** and **no-loss**.)

Action List the forwarding classes using the operational mode command **show class-of-service forwarding-class**:

```
user@switch> show class-of-service forwarding-class
Forwarding class                    ID      Queue   Policing priority   No-Loss

  best-effort                       0        0          normal         Disabled

  be2                               1        3          normal         Disabled

  hpc                               2        4          normal         Disabled

  network-control                   3        7          normal         Disabled

  mcast                             8        8          normal         Disabled
```

Meaning The **show class-of-service forwarding-class** command lists all of the configured forwarding classes, the internal identification number of each forwarding class, the queues that are mapped to the forwarding classes, the policing priority, and whether the forwarding class is lossless (no-loss packet drop attribute enabled) or lossy forwarding class (no-loss packet drop attribute disabled). The command output shows that:

- Forwarding class **best-effort** maps to queue **0** and is lossy
- Forwarding class **be2** maps to queue **1** and is lossy
- Forwarding class **hpc** maps to queue **5** and is lossy
- Forwarding class **network-control** maps to queue **7** and is lossy

In addition, the command lists the default multicast (multidestination) forwarding class and the default queue to which it is mapped.

Verifying the Forwarding Class Sets (Priority Groups)

Purpose Verify that you created the priority groups and that the correct priorities (forwarding classes) belong to the appropriate priority group.

Action List the forwarding class sets using the operational mode command **show class-of-service forwarding-class-set**:

```
user@switch> show class-of-service forwarding-class-set
Forwarding class set: best-effort-pg, Type: normal-type, Forwarding class set
index: 19907
  Forwarding class                       Index
  best-effort                              0
  be2                                      1
  network-control                          5

Forwarding class set: guar-delivery-pg, Type: normal-type, Forwarding class set
index: 43700
  Forwarding class                       Index
  fcoe                                     2
  no-loss                                  3

Forwarding class set: hpc-pg, Type: normal-type, Forwarding class set index: 60758

  Forwarding class                       Index
  hpc                                      4
```

Meaning The **show class-of-service forwarding-class-set** command lists all of the configured forwarding class sets (priority groups), the forwarding classes (priorities) that belong to each priority group, and the internal index number of each priority group. The command output shows that:

- The forwarding class set **best-effort-pg** includes the forwarding classes **best-effort**, **be2**, and **network-control**.
- The forwarding class set **guar-delivery-pg** includes the forwarding classes **fcoe** and **no-loss**.
- The forwarding class set **hpc-pg** includes the forwarding class **hpc**.

Verifying the Classifier

Purpose Verify that the classifier maps forwarding classes to the correct IEEE 802.1p code points and packet loss priorities.

Action List the classifier configured for hierarchical port scheduling using the operational mode command **show class-of-service classifier name hsclassifier1**:

```
user@switch> show class-of-service classifier name hsclassifier1
Classifier: hsclassifier1, Code point type: ieee-802.1, Index: 43607
  Code point         Forwarding class                    Loss priority
```



```
  000            best-effort              low
  001            be2                      high
  011            fcoe                     low
  100            no-loss                  low
  101            hpc                      low
  110            network-control          low
```

Meaning The **show class-of-service classifier name hsclassifier1** command lists all of the IEEE 802.1p code points and the loss priorities mapped to all of the forwarding classes in the classifier. The command output shows that the forwarding classes **best-effort**, **be2**, **no-loss**, **fcoe**, **hpc**, and **network-control** have been created and mapped to IEEE 802.1p code points and loss priorities.

Verifying Priority-Based Flow Control

Purpose Verify that PFC is enabled on the correct priorities for lossless transport.

Action List the congestion notification profiles using the operational mode command **show class-of-service congestion-notification**:

```
user@switch> show class-of-service congestion-notification
Type: Input, Name: gd-cnp, Index: 51687
Cable Length: 100 m
  Priority     PFC          MRU
  000          Disabled
  001          Disabled
  010          Disabled
  011          Enabled      2500
  100          Enabled      2500
  101          Disabled
  110          Disabled
  111          Disabled
Type: Output
  Priority     Flow-Control-Queues
  000
               0
  001
               1
  010
               2
  011
               3
  100
               4
  101
               5
  110
               6
  111
               7
```

Meaning The **show class-of-service congestion-notification** command lists all of the congestion notification profiles and the IEEE 802.1p code points with PFC enabled. The command output shows that PFC is enabled for code points **011** (**fcoe** priority and queue) and **100** (**no-loss** priority and queue) for the **gd-cnp** congestion notification profile.

The command also shows the default cable length (100 meters), the default maximum receive unit (2500 bytes), and the default mapping of priorities to output queues because this example does not include configuring these options.

Verifying the Output Queue Schedulers

Purpose Verify that you created the output queue schedulers with the correct bandwidth parameters and priorities, mapped to the correct queues, and mapped to the correct drop profiles.

Action List the scheduler maps using the operational mode command **show class-of-service scheduler-map**:

```
user@switch> show class-of-service scheduler-map
Scheduler map: be-map, Index: 64023

  Scheduler: be-sched, Forwarding class: best-effort, Index: 13005
    Transmit rate: 3000000000 bps, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified
    Shaping rate: 100 percent,
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         55387    dp-be-low
      Medium high     any             1    <default-drop-profile>
      High            any          4369    dp-be-high

  Scheduler: be-sched, Forwarding class: be2, Index: 13005
    Transmit rate: 3000000000 bps, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified
    Shaping rate: 100 percent,
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         55387    dp-be-low
      Medium high     any             1    <default-drop-profile>
      High            any          4369    dp-be-high

  Scheduler: nc-sched, Forwarding class: network-control, Index: 45740
    Transmit rate: 500000000 bps, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified
    Shaping rate: 100 percent,
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         44207    dp-nc
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

Scheduler map: gd-map, Index: 61447
```



```
  Scheduler: fcoe-sched, Forwarding class: fcoe, Index: 37289
    Transmit rate: 2500000000 bps, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified
    Shaping rate: 100 percent,
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         44207    <default-drop-profile>
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

  Scheduler: nl-sched, Forwarding class: no-loss, Index: 29359
    Transmit rate: 2000000000 bps, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified
    Shaping rate: 100 percent,
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         44207    <default-drop-profile>
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>

Scheduler map: hpc-map, Index: 56941

  Scheduler: hpc-sched, Forwarding class: hpc, Index: 55900
    Transmit rate: 2000000000 bps, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified
    Shaping rate: 100 percent,
    drop-profile-map-set-type: mark
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         57716    dp-hpc
      Medium high     any             1    <default-drop-profile>
      High            any             1    <default-drop-profile>
```

Meaning The **show class-of-service scheduler-map** command lists all of the configured scheduler maps. For each scheduler map, the command output includes:

- The name of the scheduler map (**scheduler-map** field)
- The name of the scheduler (**scheduler** field)
- The forwarding classes mapped to the scheduler (**forwarding-class** field)
- The minimum guaranteed queue bandwidth (**transmit-rate** field)
- The scheduling priority (**priority** field)
- The maximum bandwidth in the priority group the queue can consume (**shaping-rate** field)
- The drop profile loss priority (**loss priority** field) for each drop profile name (**name** field)

The command output shows that:

- The scheduler map **be-map** was created and has these properties:
  - There are two schedulers, **be-sched** and **nc-sched**.
  - The scheduler **be-sched** has two forwarding classes, **best-effort** and **be2**.
  - Scheduler **be-sched** forwarding classes **best-effort** and **be2** share a minimum guaranteed bandwidth of **3,000,000,000 bps**, can consume a maximum of **100 percent** of the priority group bandwidth, and use the drop profile **dp-be-low** for low loss-priority traffic, the default drop profile for medium-high loss-priority traffic, and the drop profile **dp-be-high** for high loss-priority traffic.
  - The scheduler **nc-sched** has one forwarding class, **network-control**.
  - The **network-control** forwarding class has a minimum guaranteed bandwidth of **500,000,000 bps**, can consume a maximum of **100 percent** of the priority group bandwidth, and uses the drop profile **dp-nc** for low loss-priority traffic and the default drop profile for medium-high and high loss priority traffic.
- The scheduler map **gd-map** was created and has these properties:
  - There are two schedulers, **fcoe-sched** and **nl-sched**.
  - The scheduler **fcoe-sched** has one forwarding class, **fcoe**.
  - The **fcoe** forwarding class has a minimum guaranteed bandwidth of **2,500,000,000 bps**, and can consume a maximum of **100 percent** of the priority group bandwidth.
  - The scheduler **nl-sched** has one forwarding class, **no-loss**.
  - The **no-loss** forwarding class has a minimum guaranteed bandwidth of **2,000,000,000 bps**, and can consume a maximum of **100 percent** of the priority group bandwidth.
- The scheduler map **hpc-map** was created and has these properties:
  - There is one scheduler, **hpc-sched**.
  - The scheduler **hpc-sched** has one forwarding class, **hpc**.
  - The **hpc** forwarding class has a minimum guaranteed bandwidth of **2,000,000,000 bps**, can consume a maximum of **100 percent** of the priority group bandwidth, and uses the drop profile **dp-hpc** for low loss-priority traffic and the default drop profile for medium-high and high loss-priority traffic.

Verifying the Drop Profiles

Purpose Verify that you created the drop profiles **dp-be-high**, **dp-be-low**, **dp-hpc**, and **dp-nc** with the correct fill levels and drop probabilities.

Action List the drop profiles using the operational mode command **show configuration class-of-service drop-profiles**:

```
user@switch> show configuration class-of-service drop-profiles
dp-be-low {
    interpolate {
        fill-level [ 25 50 ];
        drop-probability [ 0 80 ];
    }
}
dp-be-high {
    interpolate {
        fill-level [ 10 40 ];
        drop-probability [ 0 100 ];
    }
}
dp-hpc {
    interpolate {
        fill-level [ 75 90 ];
        drop-probability [ 0 75 ];
    }
}
dp-nc {
    interpolate {
        fill-level [ 80 100 ];
        drop-probability [ 0 100 ];
    }
```

Meaning The **show configuration class-of-service drop-profiles** command lists the drop profiles and their properties. The command output shows that there are four drop profiles configured, **dp-be-high**, **dp-be-low**, **dp-hpc**, and **dp-nc**. The output also shows that:

- For **dp-be-low**, the drop start point (the first fill level) is when the queue is 25 percent filled, the drop end point (the second fill level) occurs when the queue is 50 percent filled, and the drop probability at the drop end point is 80 percent.
- For **dp-be-high**, the drop start point (the first fill level) is when the queue is 10 percent filled, the drop end point (the second fill level) occurs when the queue is 40 percent filled, and the drop probability at the drop end point is 100 percent.
- For **dp-hpc**, the drop start point (the first fill level) is when the queue is 75 percent filled, the drop end point (the second fill level) occurs when the queue is 90 percent filled, and the drop probability at the drop end point is 75 percent.
- For **dp-nc**, the drop start point (the first fill level) is when the queue is 80 percent filled, the drop end point (the second fill level) occurs when the queue is 100 percent filled, and the drop probability at the drop end point is 100 percent.

Verifying the Priority Group Output Schedulers (Traffic Control Profiles)

Purpose Verify that you created the traffic control profiles **be-tcp**, **gd-tcp**, and **hpc-tcp** with the correct bandwidth parameters and scheduler mapping.

Action List the traffic control profiles using the operational mode command **show class-of-service traffic-control-profile**:

```
user@switch> show class-of-service traffic-control-profile
Traffic control profile: be-tcp, Index: 40535
  Shaping rate: 100 percent
  Scheduler map: be-map
  Guaranteed rate: 3500000000

Traffic control profile: gd-tcp, Index: 37959
  Shaping rate: 100 percent
  Scheduler map: gd-map
  Guaranteed rate: 4500000000

Traffic control profile: hpc-tcp, Index: 47661
  Shaping rate: 100 percent
  Scheduler map: hpc-map
  Guaranteed rate: 2000000000
```

Meaning The **show class-of-service traffic-control-profile** command lists all of the configured traffic control profiles. For each traffic control profile, the command output includes:

- The name of the traffic control profile (**traffic-control-profile**)
- The maximum port bandwidth the priority group can consume (**shaping-rate**)
- The scheduler map associated with the traffic control profile (**scheduler-map**)
- The minimum guaranteed priority group port bandwidth (**guaranteed-rate**)

The command output shows that:

- The traffic control profile **be-tcp** can consume a maximum of **100 percent** of the port bandwidth, is associated with the scheduler map **be-map**, and has a minimum guaranteed bandwidth of **3,500,000,000 bps**.
- The traffic control profile **gd-tcp** can consume a maximum of **100 percent** of the port bandwidth, is associated with the scheduler map **gd-map**, and has a minimum guaranteed bandwidth of **4,500,000,000 bps**.
- The traffic control profile **hpc-tcp** can consume a maximum of **100 percent** of the port bandwidth, is associated with the scheduler map **hpc-map**, and has a minimum guaranteed bandwidth of **2,000,000,000 bps**.

Verifying the Interface Configuration

Purpose Verify that the classifier, the congestion notification profile, and the forwarding class sets are configured on interfaces **xe-0/0/20** and **xe-0/0/21**.

Action List the interfaces using the operational mode commands **show configuration class-of-service interfaces xe-0/0/20** and **show configuration class-of-service interfaces xe-0/0/21**:

```
user@switch> show configuration class-of-service interfaces xe-0/0/20
forwarding-class-set {
    best-effort-gp {
        output-traffic-control-profile be-tcp;
    }
    guar-delivery-pg {
        output-traffic-control-profile gd-tcp;
    }
    hpc-pg {
        output-traffic-control-profile hpc-tcp;
    }
}
congestion-notification-profile gd_cnp;
unit 0 {
    classifiers {
        ieee-802.1 hsclassifier1;
    }
}
```

```
user@switch> show configuration class-of-service interfaces xe-0/0/21
forwarding-class-set {
    best-effort-gp {
        output-traffic-control-profile be-tcp;
    }
    guar-delivery-pg {
        output-traffic-control-profile gd-tcp;
    }
    hpc-pg {
        output-traffic-control-profile hpc-tcp;
    }
}
congestion-notification-profile gd_cnp;
unit 0 {
    classifiers {
        ieee-802.1 hsclassifier1;
    }
}
```

Meaning The **show configuration class-of-service interfaces** ***interface-name*** command shows that each interface includes the forwarding class sets **best-effort-pg**, **guar-delivery-pg**, and **hpc-pg**, congestion notification profile **gd-cnp**, and the IEEE 802.1p classifier **hsclassifier1**.

Related Documentation

- Defining CoS BA Classifiers (DSCP, DSCP IPv6, IEEE 802.1p) on page 130
- *Benefits of Configuring CoS Hierarchical Port Scheduling*
- Assigning CoS Components to Interfaces on page 21
- Example: Configuring WRED Drop Profiles on page 371
- Example: Configuring Drop Profile Maps on page 377

- Example: Configuring Forwarding Classes on page 142
- Example: Configuring Forwarding Class Sets on page 150
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Queue Scheduling Priority on page 302
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Example: Configuring Maximum Output Bandwidth on page 357
- *Configuring CoS PFC (Congestion Notification Profiles)*
- *Overview of CoS Changes Introduced in Junos OS Release 12.2*
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315
- Understanding CoS Scheduling Behavior and Configuration Considerations on page 277
- *Understanding CoS Scheduling on QFabric System Node Device Fabric (fte) Ports*
- *Understanding Default CoS Scheduling on QFabric System Interconnect Devices (Junos OS Release 13.1 and Later Releases)*

Understanding CoS Priority Group and Queue Guaranteed Minimum Bandwidth

You can set a guaranteed minimum bandwidth for individual forwarding classes (queues) and for groups of forwarding classes called *forwarding class sets* (priority groups). Setting a minimum guaranteed bandwidth ensures that priority groups and queues receive the bandwidth required to support the expected traffic.

This topic covers:

- Guaranteeing Bandwidth Using Hierarchical Scheduling on page 347
- Priority Group Guaranteed Rate (Guaranteed Minimum Bandwidth) on page 349
- Queue Transmit Rate (Guaranteed Minimum Bandwidth) on page 349

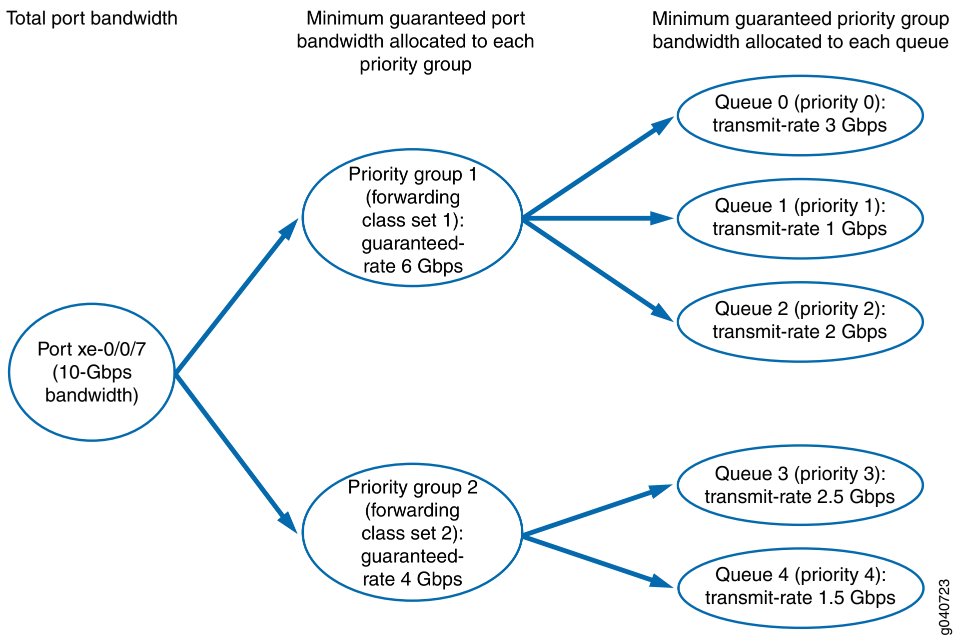
Guaranteeing Bandwidth Using Hierarchical Scheduling

The **guaranteed-rate** value for the priority group (configured in a traffic control profile) defines the minimum amount of bandwidth allocated to a forwarding class set on a port, whereas the **transmit-rate** value of the queue (configured in a scheduler) defines the minimum amount of bandwidth allocated to a particular queue in a priority group. The queue bandwidth is a portion of the priority group bandwidth.

NOTE: You cannot configure a minimum guaranteed bandwidth (transmit rate) for a forwarding class that is mapped to a strict-high priority queue, and you cannot configure a minimum guaranteed bandwidth (guaranteed rate) for a priority group that includes strict-high priority queues.

Figure 9 on page 348 shows how the total port bandwidth is allocated to priority groups (forwarding class sets) based on the guaranteed rate of each priority group. It also shows how the guaranteed bandwidth of each priority group is allocated to the queues in the priority group based on the transmit rate of each queue.

Figure 9: Allocating Guaranteed Bandwidth Using Hierarchical Scheduling



The sum of the priority group guaranteed rates cannot exceed the total port bandwidth. If you configure guaranteed rates whose sum exceeds the port bandwidth, the system sends a syslog message to notify you that the configuration is not valid. However, the system does not perform a commit check. If you commit a configuration in which the sum of the guaranteed rates exceeds the port bandwidth, the hierarchical scheduler behaves unpredictably.

The sum of the queue transmit rates cannot exceed the total guaranteed rate of the priority group to which the queues belong. If you configure transmit rates whose sum exceeds the priority group guaranteed rate, the commit check fails and the system rejects the configuration.

NOTE: You must set both the priority group **guaranteed-rate** value and the queue **transmit-rate** value in order to configure the minimum bandwidth for individual queues. If you set the **transmit-rate** value but do not set the **guaranteed-rate** value, the configuration fails.

You can set the **guaranteed-rate** value for a priority group without setting the **transmit-rate** value for individual queues in the priority group. However, queues that do not have a configured **transmit-rate** value can become starved for bandwidth if other higher-priority queues need the priority group's bandwidth. To avoid starving a queue, it is a good practice to configure a **transmit-rate** value for most queues.

If you configure the guaranteed rate of a priority group as a percentage, configure all of the transmit rates associated with that priority group as percentages. In this case, if any of the transmit rates are configured as absolute values instead of percentages, the configuration is not valid and the system sends a syslog message.

Priority Group Guaranteed Rate (Guaranteed Minimum Bandwidth)

Setting a priority group (forwarding class set) **guaranteed-rate** enables you to reserve a portion of the port bandwidth for the forwarding classes (queues) in that forwarding class set. The minimum bandwidth (**guaranteed-rate**) that you configure for a priority group sets the minimum bandwidth available to all of the forwarding classes in the forwarding class set.

The combined **guaranteed-rate** value of all of the forwarding class sets associated with an interface cannot exceed the amount of bandwidth available on that interface.

You configure the priority group **guaranteed-rate** in the traffic control profile. You cannot apply a traffic control profile that has a guaranteed rate to a priority group that includes a strict-high priority queue.

Queue Transmit Rate (Guaranteed Minimum Bandwidth)

Setting a queue (forwarding class) **transmit-rate** enables you to reserve a portion of the priority group bandwidth for the individual queue. For example, a queue that handles Fibre Channel over Ethernet (FCoE) traffic might require a minimum rate of 4 Gbps to ensure the class of service that storage area network (SAN) traffic requires.

The priority group **guaranteed-rate** sets the aggregate minimum amount of bandwidth available to the queues that belong to the priority group. The cumulative total minimum bandwidth the queues consume cannot exceed the minimum bandwidth allocated to the priority group to which they belong. (The combined transmit rates of the queues in a priority group cannot exceed the priority group's guaranteed rate.)

You must configure the **guaranteed-rate** value of the priority group in order to set a **transmit-rate** value for individual queues that belong to the priority group. The reason is that if there is no guaranteed bandwidth for a priority group, there is no way to guarantee bandwidth for queues in that priority group.

You configure the queue **transmit-rate** in the scheduler configuration. You cannot configure a transmit rate for a strict-high priority queue.

Related Documentation

- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS Traffic Control Profiles on page 306
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Defining CoS Queue Schedulers on page 290
- Defining CoS Traffic Control Profiles (Priority Group Scheduling) on page 310

Example: Configuring Minimum Guaranteed Output Bandwidth

Scheduling the minimum guaranteed output bandwidth for a queue (forwarding class) requires configuring both tiers of the two-tier hierarchical scheduler. One tier is scheduling the resources for the individual queue. The other tier is scheduling the resources for the priority group (forwarding class set) to which the queue belongs. You set a minimum guaranteed bandwidth to ensure than priority groups and queues receive the bandwidth required to support the expected traffic.

Requirements

This example uses the following hardware and software components:

- A Juniper Networks QFX3500 Switch
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

The priority group minimum guaranteed bandwidth defines the minimum total amount of bandwidth available for all of the queues in the priority group to meet their minimum bandwidth requirements.

The **transmit-rate** setting in the scheduler configuration determines the minimum guaranteed bandwidth for an individual queue. The transmit rate also determines the amount of excess (extra) priority group bandwidth that the queue can share. Extra priority group bandwidth is allocated among the queues in the priority group in proportion to the transmit rate of each queue.

The **guaranteed-rate** setting in the traffic control profile configuration determines the minimum guaranteed bandwidth for a priority group. The guaranteed rate also determines the amount of excess (extra) port bandwidth that the priority group can share. Extra port bandwidth is allocated among the priority groups on a port in proportion to the guaranteed rate of each priority group.

NOTE: You must configure both the **transmit-rate** value for the queue and the **guaranteed-rate** value for the priority group to set a valid minimum bandwidth guarantee for a queue. (If the priority group does not have a guaranteed minimum bandwidth, there is no guaranteed bandwidth pool from which the queue can take its guaranteed minimum bandwidth.)

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.)

NOTE: When you configure bandwidth for a queue or a priority group, the switch considers only the data as the configured bandwidth. The switch does not account for the bandwidth consumed by the preamble and the interframe gap (IFG). Therefore, when you calculate and configure the bandwidth requirements for a queue or for a priority group, consider the preamble and the IFG as well as the data in the calculations.

NOTE: You cannot configure minimum guaranteed bandwidth on strict-high priority queues or on a priority group that contains strict-high priority queues.

This example describes how to:

- Configure a transmit rate (minimum guaranteed queue bandwidth) of 2 Gbps for queues in a scheduler named **be-sched**.
- Configure a guaranteed rate (minimum guaranteed priority group bandwidth) of 4 Gbps for a priority group in a traffic control profile named **be-tcp**.
- Assign the scheduler to a queue named **best-effort** by using a scheduler map named **be-map**.
- Associate the scheduler map **be-map** with the traffic control profile **be-tcp**.
- Assign the queue **best-effort** to a priority group named **be-pg**.
- Assign the priority group and the minimum guaranteed bandwidth scheduling to the egress interface **xe-0/0/7**.

Table 73 on page 352 shows the configuration components for this example:

Table 73: Components of the Minimum Guaranteed Output Bandwidth Configuration Example

Component	Settings
Hardware	QFX3500 switch
Minimum guaranteed queue bandwidth	Transmit rate: **2g**
Minimum guaranteed priority group bandwidth	Guaranteed rate: **4g**
Scheduler	**be-sched**
Scheduler map	**be-map**
Traffic control profile	**be-tcp**
Forwarding class set (priority group)	**be-pg**
Queue (forwarding class)	**best-effort**
Egress interface	**xe-0/0/7**

Configuring Guaranteed Minimum Bandwidth

CLI Quick Configuration

To quickly configure the minimum guaranteed bandwidth for a priority group and a queue, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit] hierarchy level:

```
[edit class-of-service]
set schedulers be-sched transmit-rate percent 2g
set traffic-control-profiles be-tcp guaranteed-rate 4g
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set traffic-control-profiles be-tcp scheduler-map be-map
set forwarding-class-sets be-pg class best-effort
set interfaces xe-0/0/7 forwarding-class-set be-pg output-traffic-control-profile be-tcp
```

To configure the minimum guaranteed bandwidth hierarchical scheduling for a queue and a priority group:

1. Configure the minimum guaranteed queue bandwidth of 2 Gbps for scheduler **be-sched**:

   ```
   [edit class-of-service schedulers]
   user@switch# set be-sched transmit-rate 2g
   ```

2. Configure the minimum guaranteed priority group bandwidth of 4 Gbps for traffic control profile **be-tcp**:

   ```
   [edit class-of-service traffic-control-profiles]
   user@switch# set be-tcp guaranteed-rate 4g
   ```


3. Associate the scheduler **be-sched** with the **best-effort** queue in the scheduler map **be-map**:

```
[edit class-of-service scheduler-maps]
user@switch# set be-map forwarding-class best-effort scheduler be-sched
```

4. Associate the scheduler map with the traffic control profile:

```
[edit class-of-service traffic-control-profiles]
user@switch# set be-tcp scheduler-map be-map
```

5. Assign the **best-effort** queue to the priority group **be-pg**:

```
[edit class-of-service forwarding-class-sets]
user@switch# set be-pg class best-effort
```

6. Apply the configuration to interface **xe-0/0/7**:

```
[edit class-of-service interfaces]
user@switch# set xe-0/0/7 forwarding-class-set be-pg output-traffic-control-profile be-tcp
```

Verification

To verify the minimum guaranteed output bandwidth configuration, perform these tasks:

- Verifying the Minimum Guaranteed Queue Bandwidth on page 353
- Verifying the Priority Group Minimum Guaranteed Bandwidth and Scheduler Map Association on page 353
- Verifying the Scheduler Map Configuration on page 354
- Verifying Queue (Forwarding Class) Membership in the Priority Group on page 354
- Verifying the Egress Interface Configuration on page 354

Verifying the Minimum Guaranteed Queue Bandwidth

Purpose Verify that you configured the minimum guaranteed queue bandwidth as **2g** in the scheduler **be-sched**.

Action Display the minimum guaranteed bandwidth in the **be-sched** scheduler configuration using the operational mode command **show configuration class-of-service schedulers be-sched transmit-rate**:

```
user@switch> show configuration class-of-service schedulers be-sched transmit-rate
2g;
```

Verifying the Priority Group Minimum Guaranteed Bandwidth and Scheduler Map Association

Purpose Verify that the minimum guaranteed priority group bandwidth is **4g** and the attached scheduler map is **be-map** in the traffic control profile **be-tcp**.

Action Display the minimum guaranteed bandwidth in the **be-tcp** traffic control profile configuration using the operational mode command **show configuration class-of-service traffic-control-profiles be-tcp guaranteed-rate**:

```
user@switch> show configuration class-of-service traffic-control-profiles be-tcp guaranteed-rate
4g;
```

Display the scheduler map in the **be-tcp** traffic control profile configuration using the operational mode command **show configuration class-of-service traffic-control-profiles be-tcp scheduler-map**:

```
user@switch> show configuration class-of-service traffic-control-profiles be-tcp scheduler-map
scheduler-map be-map;
```

Verifying the Scheduler Map Configuration

Purpose Verify that the scheduler map **be-map** maps the forwarding class **best-effort** to the scheduler **be-sched**.

Action Display the **be-map** scheduler map configuration using the operational mode command **show configuration class-of-service schedulers maps be-map**:

```
user@switch> show configuration class-of-service scheduler-maps be-map
forwarding-class best-effort scheduler be-sched;
```

Verifying Queue (Forwarding Class) Membership in the Priority Group

Purpose Verify that the forwarding class set **be-pg** includes the forwarding class **best-effort**.

Action Display the **be-pg** forwarding class set configuration using the operational mode command **show configuration class-of-service forwarding-class-sets be-pg**:

```
user@switch> show configuration class-of-service forwarding-class-sets be-pg
class best-effort;
```

Verifying the Egress Interface Configuration

Purpose Verify that the forwarding class set **be-pg** and the traffic control profile **be-tcp** are attached to egress interface **xe-0/0/7**.

Action Display the egress interface using the operational mode command **show configuration class-of-service interfaces xe-0/0/7**:

```
user@switch> show configuration class-of-service interfaces xe-0/0/7
forwarding-class-set {
    be-pg {
```



```
            output-traffic-control-profile be-tcp;
        }
    }
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Queue Scheduling Priority on page 302
- Example: Configuring Forwarding Class Sets on page 150
- Understanding CoS Traffic Control Profiles on page 306
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315

Understanding CoS Priority Group Shaping and Queue Shaping (Maximum Bandwidth)

If the amount of traffic on an interface exceeds the maximum bandwidth available on the interface, it leads to congestion. You can use priority group (forwarding class set) shaping and queue (forwarding class) shaping to manage traffic and avoid congestion.

Configuring a maximum bandwidth sets the most bandwidth a priority group or a queue can use after all of the priority group and queue minimum bandwidth requirements are met, even if more bandwidth is available.

This topic covers:

Priority Group Shaping

Priority group shaping enables you to shape the aggregate traffic of a forwarding class set on a port to a maximum rate that is less than the line or port rate. The maximum bandwidth (**shaping-rate**) that you configure for a priority group sets the maximum bandwidth available to all of the forwarding classes (queues) in the forwarding class set.

If a port has more than one priority group and the combined **shaping-rate** value of the priority groups is greater than the amount of port bandwidth available, the bandwidth is shared proportionally among the priority groups.

You configure the priority group **shaping-rate** in the traffic control profile.

Queue Shaping

Queue shaping throttles the rate at which queues transmit packets. For example, using queue shaping, you can rate-limit a strict-high priority queue so that the strict-priority queue does not lock out (or starve) low-priority queues.

NOTE: We recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

Similarly, for any queue, you can configure queue shaping (**shaping-rate**) to set the maximum bandwidth for a particular queue.

The **shaping-rate** value of the priority group sets the aggregate maximum amount of bandwidth available to the queues that belong to the priority group. On a port, the cumulative total bandwidth the queues consume cannot exceed the maximum bandwidth of the priority group to which they belong.

If a priority group has more than one queue, and the combined **shaping-rate** of the queues is greater than the amount of bandwidth available to the priority group, the bandwidth is shared proportionally among the queues.

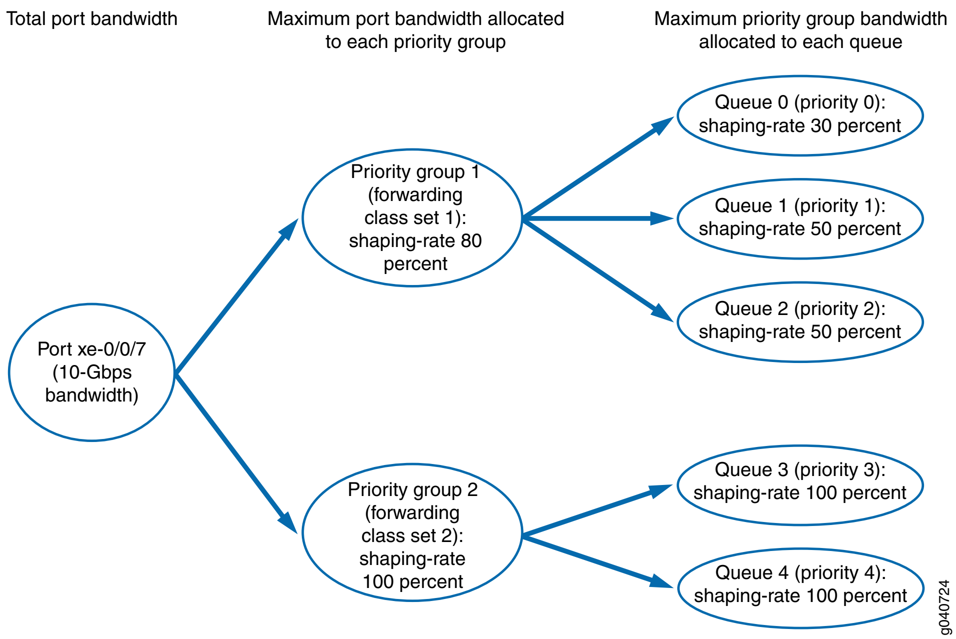
You configure the queue **shaping-rate** in the scheduler configuration, and you set the **shaping-rate** for priority groups in the traffic control profile configuration.

Shaping Maximum Bandwidth Using Hierarchical Scheduling

Priority group shaping defines the maximum bandwidth allocated to a forwarding class set on a port, whereas queue shaping defines a limit on maximum bandwidth usage per queue. The queue bandwidth is a portion of the priority group bandwidth.

Figure 10 on page 357 shows how the port bandwidth is allocated to priority groups (forwarding class sets) based on the shaping rate of each priority group, and how the bandwidth of each priority group is allocated to the queues in the priority group based on the shaping rate of each queue.

Figure 10: Setting Maximum Bandwidth Using Hierarchical Scheduling



Related Documentation

- Understanding CoS Output Queue Schedulers on page 283
- Understanding CoS Traffic Control Profiles on page 306
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Defining CoS Queue Schedulers on page 290
- Defining CoS Traffic Control Profiles (Priority Group Scheduling) on page 310

Example: Configuring Maximum Output Bandwidth

Scheduling the maximum output bandwidth for a queue (forwarding class) requires configuring both tiers of the hierarchical scheduler. One tier is scheduling the resources for the individual queue. The other tier is scheduling the resources for the priority group (forwarding class set) to which the queue belongs. You can use priority group and queue shaping to prevent traffic from using more bandwidth than you want the traffic to receive.

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

The priority group maximum bandwidth defines the maximum total amount of bandwidth available for all of the queues in the priority group.

The **shaping-rate** setting in the scheduler configuration determines the maximum bandwidth for an individual queue.

The **shaping-rate** setting in the traffic control profile configuration determines the maximum bandwidth for a priority group.

NOTE: When you configure bandwidth for a queue or a priority group, the switch considers only the data as the configured bandwidth. The switch does not account for the bandwidth consumed by the preamble and the interframe gap (IFG). Therefore, when you calculate and configure the bandwidth requirements for a queue or for a priority group, consider the preamble and the IFG as well as the data in the calculations.

NOTE: When you set the maximum bandwidth (**shaping-rate**) for a queue or for a priority group at 100 Kbps or less, the traffic shaping behavior is accurate only within +/– 20 percent of the configured **shaping-rate** value.

This example describes how to:

- Configure a maximum rate of 4 Gbps for queues in a scheduler named **be-sched**.
- Configure a maximum rate of 6 Gbps for a priority group in a traffic control profile named **be-tcp**.
- Assign the scheduler to a queue named **best-effort** by using a scheduler map named **be-map**.
- Associate the scheduler map **be-map** with the traffic control profile **be-tcp**.
- Assign the queue **best-effort** to a priority group named **be-pg**.
- Assign the priority group and the bandwidth scheduling to the interface **xe-0/0/7**.

Table 74 on page 359 shows the configuration components for this example:

Table 74: Components of the Maximum Output Bandwidth Configuration Example

Component	Settings
Hardware	QFX3500 switch
Maximum queue bandwidth	Shaping rate: **4g**
Maximum priority group bandwidth	Shaping rate: **6g**
Scheduler	**be-sched**
Scheduler map	**be-map**
Traffic control profile	**be-tcp**
Forwarding class set (priority group)	**be-pg**
Queue (forwarding class)	**best-effort**
Egress interface	**xe-0/0/7**

Configuring Maximum Bandwidth

CLI Quick Configuration

To quickly configure the maximum bandwidth for a priority group and a queue, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit] hierarchy level:

```
[edit class-of-service]
set schedulers be-sched shaping-rate percent 4g
set traffic-control-profiles be-tcp shaping-rate 6g
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set traffic-control-profiles be-tcp scheduler-map be-map
set forwarding-class-sets be-pg class best-effort
set interfaces xe-0/0/7 forwarding-class-set be-pg output-traffic-control-profile be-tcp
```

To configure the maximum bandwidth hierarchical scheduling for a queue and a priority group:

1. Configure the maximum queue bandwidth of 4 Gbps for scheduler **be-sched**:

   ```
   [edit class-of-service schedulers]
   user@switch# set be-sched shaping-rate 4g
   ```

2. Configure the maximum priority group bandwidth of 6 Gbps for traffic control profile **be-tcp**:

   ```
   [edit class-of-service traffic-control-profiles]
   user@switch# set be-tcp shaping-rate 6g
   ```

3. Associate the scheduler **be-sched** with the **best-effort** queue in the scheduler map **be-map**:

```
[edit class-of-service scheduler-maps]
user@switch# set be-map forwarding-class best-effort scheduler be-sched
```

4. Associate the scheduler map with the traffic control profile:

```
[edit class-of-service traffic-control-profiles]
user@switch# set be-tcp scheduler-map be-map
```

5. Assign the **best-effort** queue to the priority group **be-pg**:

```
[edit class-of-service forwarding-class-sets]
user@switch# set be-pg class best-effort
```

6. Apply the configuration to interface **xe-0/0/7**:

```
[edit class-of-service interfaces]
user@switch# set xe-0/0/7 forwarding-class-set be-pg output-traffic-control-profile be-tcp
```

Verification

To verify the maximum output bandwidth configuration, perform these tasks:

- Verifying the Maximum Queue Bandwidth on page 360
- Verifying the Priority Group Maximum Bandwidth and Scheduler Map Association on page 360
- Verifying the Scheduler Map Configuration on page 361
- Verifying Queue (Forwarding Class) Membership in the Priority Group on page 361
- Verifying the Egress Interface Configuration on page 361

Verifying the Maximum Queue Bandwidth

Purpose Verify that you configured the maximum queue bandwidth as **4g** in the scheduler **be-sched**.

Action List the maximum bandwidth in the **be-sched** scheduler configuration using the operational mode command **show configuration class-of-service schedulers be-sched shaping-rate**:

```
user@switch> show configuration class-of-service schedulers be-sched shaping-rate
4g;
```

Verifying the Priority Group Maximum Bandwidth and Scheduler Map Association

Purpose Verify that the maximum priority group bandwidth is **6g** and the attached scheduler map is **be-map** in the traffic control profile **be-tcp**.

Action List the maximum bandwidth in the **be-tcp** traffic control profile configuration using the operational mode command **show configuration class-of-service traffic-control-profiles be-tcp shaping-rate**:

```
user@switch> show configuration class-of-service traffic-control-profiles be-tcp shaping-rate
4g;
```

List the scheduler map in the **be-tcp** traffic control profile configuration using the operational mode command **show configuration class-of-service traffic-control-profiles be-tcp scheduler-map**:

```
user@switch> show configuration class-of-service traffic-control-profiles be-tcp scheduler-map
scheduler-map be-map;
```

Verifying the Scheduler Map Configuration

Purpose Verify that the scheduler map **be-map** maps the forwarding class **best-effort** to the scheduler **be-sched**.

Action List the **be-map** scheduler map configuration using the operational mode command **show configuration class-of-service schedulers maps be-map**:

```
user@switch> show configuration class-of-service scheduler-maps be-map
forwarding-class best-effort scheduler be-sched;
```

Verifying Queue (Forwarding Class) Membership in the Priority Group

Purpose Verify that the forwarding class set **be-pg** includes the forwarding class **best-effort**.

Action List the **be-pg** forwarding class set configuration using the operational mode command **show configuration class-of-service forwarding-class-sets be-pg**:

```
user@switch> show configuration class-of-service forwarding-class-sets be-pg
class best-effort;
```

Verifying the Egress Interface Configuration

Purpose Verify that the forwarding class set **be-pg** and the traffic control profile **be-tcp** are attached to egress interface **xe-0/0/7**.

Action List the egress interface using the operational mode command **show configuration class-of-service interfaces xe-0/0/7**:

```
user@switch> show configuration class-of-service interfaces xe-0/0/7
forwarding-class-set {
    be-pg {
```

```
            output-traffic-control-profile be-tcp;
        }
    }
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Forwarding Class Sets on page 150
- Understanding CoS Traffic Control Profiles on page 306
- Understanding CoS Hierarchical Port Scheduling (ETS) on page 315

Understanding CoS WRED Drop Profiles

When the number of packets queued is greater than the ability of the switch to empty an output queue, the queue requires a method for determining which packets to drop to relieve the congestion. Weighted random early detection (WRED) drop profiles define the drop probability of packets of different packet loss probabilities (PLPs) as the output queue fills. During periods of congestion, as the output queue fills, the switch drops incoming packets as determined by a drop profile, until the output queue becomes less congested.

Depending on the drop probabilities, a drop profile can drop many packets long before the buffer becomes full, or it can drop only a few packets even if the buffer is almost full.

You configure drop profiles in the drop profile section of the class-of-service (CoS) configuration hierarchy. You apply drop profiles using a drop profile map in queue scheduler configuration. For each queue scheduler, you can configure separate drop profiles for each PLP using the **loss-priority** attribute (low, medium-high, and high). This enables you to treat traffic of different PLPs in different ways during periods of congestion.

NOTE: Do not apply drop profiles to lossless traffic (traffic that belongs to a forwarding class that has the no-loss drop attribute.). Lossless traffic uses priority-based flow control (PFC) to control congestion.

OCX Series switches do not support lossless transport and do not support PFC.

NOTE: You cannot apply drop profiles to multidestination queues on switches that support them.

- Drop Profile Parameters on page 363
- Defining Drop Profiles on Switches Except QFX10000 on page 363

Drop Profile Parameters

Drop profiles specify two values, which work as pairs:

- Fill level—The queue fullness value, which represents a percentage of the memory used to store packets in relation to the total amount of memory allocated to the queue.
- Drop probability—The percentage value that corresponds to the likelihood that an individual packet is dropped.

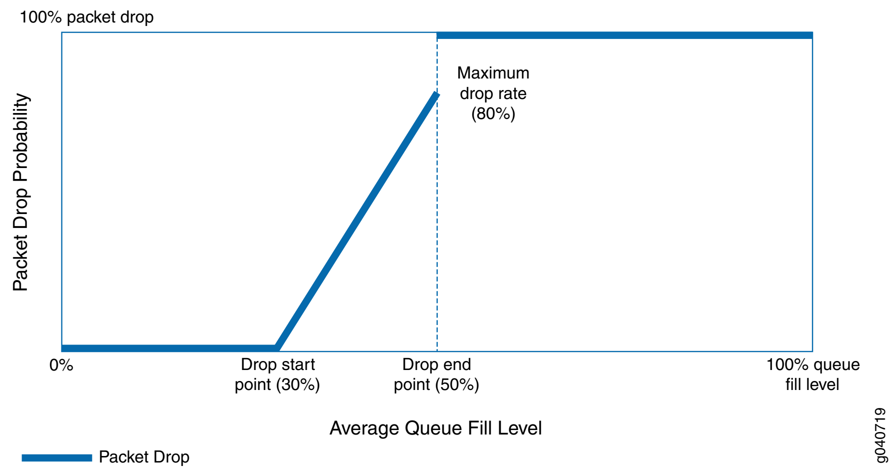
Defining Drop Profiles on Switches Except QFX10000

You set two queue fill levels and two drop probabilities in each drop profile. The first fill level and the first drop probability create one value pair and the second fill level and the second drop probability create a second value pair.

The first fill level value specifies the percentage of queue fullness at which packets begin to drop, known as the drop start point. Until the queue reaches this level of fullness, no packets are dropped. The second fill level value specifies the percentage of queue fullness at which all packets are dropped, known as the drop end point.

The first drop probability value is always **0** (zero). This pairs with the drop start point and specifies that until the queue fullness level reaches the first fill level, no packets drop. When the queue fullness exceeds the drop start point, packets begin to drop until the queue exceeds the second fill level, when all packets drop. The second drop probability value, known as the maximum drop rate, specifies the likelihood of dropping packets when the queue fullness reaches the drop end point. As the queue fills from the drop start point to the drop end point, packets drop in a smooth, linear pattern (called an interpolated graph) as shown in Figure 11 on page 364. After the drop end point, all packets drop.

Figure 11: WRED-Drop Profile Packet Drop Pattern



The thick line in Figure 11 on page 364 shows the packet drop characteristics for a sample WRED profile. At the drop start point, the queue reaches a fill level of 30 percent. At the drop end point, the queue fill level reaches 50 percent, and the maximum drop rate is 80 percent.

No packets drop until the queue fill level reaches the drop start point of 30 percent. When the queue reaches the 30 percent fill level, packets begin to drop. As the queue fills, the percentage of packets dropped increases in a linear fashion. When the queue fills to the drop end point of 50 percent, the rate of packet drop has increased to the maximum drop rate of 80 percent. When the queue fill level exceeds the drop end point of 50 percent, all of the packets drop until the queue fill level drops below 50 percent.

Defining Drop Profiles on QFX10000 Switches

Each queue fill level pairs with a drop probability. As the queue fills to different levels, every time it reaches a fill level configured in a drop profile, the queue applies the drop probability paired with that fill level to the traffic in the queue that exceeds the fill level. You can configure up to 32 pairs of fill levels and drop probabilities to create a customized packet drop probability curve with up to 32 points of differentiation.

Packets are not dropped until they reach the first configured queue fill level. When the queue reaches the first fill level, packets begin to drop at the configured drop probability rate paired with the first fill level. When the queue reaches the second fill level, packets begin to drop at the configured drop probability rate paired with the second fill level. This process continues for the number of fill level/drop probability pairs that you configure in the drop profile.

Drop profiles are interpolated, not segmented. An interpolated drop profile gradually increases the drop probability along a curve between each configured fill level. When the queue reaches the next fill level, the drop probability reaches the drop probability paired with that fill level. A segmented drop profile “jumps” from one fill level and drop probability setting to another in a stepped fashion. The drop probability of traffic does not change as the queue fills until the next fill level is reached.

An example of interpolation is a drop profile with three fill level/drop probability pairs:

- 25 percent queue fill level paired with a 30 percent drop probability
- 50 percent queue fill level paired with a 60 percent drop probability
- 75 percent queue fill level paired with a 100 percent drop probability (all packets that exceed the 75 percent queue fill level are dropped)

The queue drops no packets until its fill level reaches 25 percent. During periods of congestion, when the queue fills above 25 percent full, the queue begins to drop packets at a rate of 30 percent of the packets above the fill level.

However, as the queue continues to fill, it does not continue to drop packets at the 30 percent drop probability. Instead, the drop probability gradually increases as the queue fills to the 50 percent fullness level. When the queue reaches the 50 percent fill level, the drop probability has increased to the configured drop probability pair for the fill level, which is 60 percent.

As the queue continues to fill, the drop probability does not remain at 60 percent, but continues to rise as the queue fills. When the queue reaches the final fill level at 75 percent full, the drop probability has risen to 100 percent and all packets that exceed the 75 percent fill level are dropped.

Default Drop Profile

If you do not configure drop profiles and apply them to queue schedulers, the switch uses the default drop profile for lossy traffic classes. In the default drop profile, when the fill level is 0 percent, the drop probability is 0 percent. When the fill level is 100 percent, the drop probability is 100 percent. During periods of congestion, as soon as packets arrive on a queue, the default profile might begin to drop packets.

Packet Drop Method

When a packet reaches the head of a queue, the switch calculates a random number between 0 and 100. The switch plots the random number against the drop profile using the current fill level of the queue. When the random number falls above the graph line, the queue transmits the packet out the egress interface. When the number falls below graph the line, the switch drops the packet.

Packet Drop Example for Switches Except QFX10000

To create the linear drop pattern from the drop start point to the drop end point, the drop probabilities are derived using a linear approximation with eight sections, or steps, from the minimum queue fill level to the maximum queue fill level. The fill levels are divided into the eight sections equally, starting at the minimum fill level and ending at the maximum fill level. As the queue fills, the percentage of dropped packets increases. The percentage of packets dropped is based on the maximum drop rate.

For example, the default drop profile (which specifies a maximum drop rate of 100 percent) has the following drop probabilities at each section, or step, in the eight-section linear drop pattern:

- First section—The minimum drop probability is 6.25 percent of the maximum drop rate. The maximum drop probability is 12.5 percent of the maximum drop rate.
- Second section—The minimum drop probability is 18.75 percent of the maximum drop rate. The maximum drop probability is 25 percent of the maximum drop rate.
- Third section—The minimum drop probability is 30.25 percent of the maximum drop rate. The maximum drop probability is 37.5 percent of the maximum drop rate.
- Fourth section—The minimum drop probability is 43.75 percent of the maximum drop rate. The maximum drop probability is 50 percent of the maximum drop rate.
- Fifth section—The minimum drop probability is 56.25 percent of the maximum drop rate. The maximum drop probability is 62 percent of the maximum drop rate.
- Sixth section—The minimum drop probability is 68.75 percent of the maximum drop rate. The maximum drop probability is 75.5 percent of the maximum drop rate.
- Seventh section—The minimum drop probability is 81.25 percent of the maximum drop rate. The maximum drop probability is 87.5 percent of the maximum drop rate.
- Eighth section—The minimum drop probability is 92.75 percent of the maximum drop rate. The maximum drop probability is 100 percent of the maximum drop rate.

Packets drop even when there is no congestion, because packet drops begin at the drop start point regardless of whether congestion exists on the port. The default drop profile example represents the worst-case scenario, because the drop start point fill level is 0 percent, so packet drop begins when the queue starts to receive packets.

You can specify when packets begin to drop by configuring a drop start point at a fill level greater than 0 percent. For example, if you configure a drop profile that has a drop start point of 30 percent, packets do not drop until the queue is 30 percent full. We recommend that you configure drop profiles that are appropriate to your network traffic conditions.

The smaller the gap between the minimum drop rate (which is always 0) and the maximum drop rate, the smaller the gap between the minimum drop probability and the maximum drop probability at each section (step) of the linear drop pattern. The default drop profile, which has the maximum gap between the minimum drop rate (0 percent) and the maximum drop rate (100 percent), has the highest gap between the minimum drop probability and the maximum drop probability at each step. Configuring a lower maximum drop rate for a drop profile reduces the gap between the minimum drop probability and the maximum drop probability.

Drop Profile Maps

Drop profile maps are part of scheduler configuration. A drop profile map maps drop profiles to packet loss priorities. Specifying the drop profile map in a scheduler associates the drop profile with the forwarding classes (queues) that you map to the scheduler in a scheduler map.

You configure loss priority for a queue in the classifier section of the CoS configuration hierarchy, and the loss priority is applied to the traffic assigned to the forwarding class at the ingress interface.

Each scheduler can have multiple drop profile maps.

Congestion Prevention

Configuring drop profiles on output queues enables you to control how congestion affects other queues on a port. If you do not configure drop profiles and map them to output queues, the switch uses the default drop profile on queues that forward lossy traffic.

For example, if an ingress port forwards traffic to more than one egress port, and at least one of the egress ports experiences congestion, that can cause ingress port congestion. Ingress port congestion (ingress buffer exceeds its resource allocation) can cause frames to drop at the ingress port instead of at the egress port. Ingress port frame drop affects all of the egress ports to which the congested ingress port forwards traffic, not just the congested egress port.

NOTE: Do not configure drop profiles for the **fcoe** and **no-loss** forwarding classes. FCoE and other lossless traffic queues require lossless behavior (traffic queues that are configured with the **no-loss** packet drop attribute). Use priority-based flow control (PFC) to prevent frame drop on lossless priorities.

OCX Series switches do not support lossless transport and do not support PFC.

Configuring a WRED Drop Profile and Applying it to an Output Queue

To configure a WRED packet drop profile and apply it to an output queue:

1. Configure a drop profile:
   - On switches except QFX10000 use the statement **set class-of-service drop-profiles** ***profile-name*** **interpolate fill-level** ***drop-start-point*** **fill-level** ***drop-end-point*** **drop-probability 0 drop-probability** ***percentage***.
   - On QFX10000 switches use the statement **set class-of-service drop-profiles** ***profile-name*** **interpolate fill-level** ***level1 level2 ... level32*** **drop-probability** ***probability1 probability2 ... probability32***. You can specify as few as two fill level/drop probability pairs or as many as 32 pairs.
2. Map the drop profile to a queue scheduler using the statement **set class-of-service schedulers** ***scheduler-name*** **drop-profile-map loss-priority (low | medium-high | high) protocol any drop-profile** ***profile-name***. The name of the drop-profile is the name of the WRED profile configured in Step 1.
3. Map the scheduler, which Step 2 associates with the drop profile, to the output queue using the statement **set class-of-service scheduler-maps** ***map-name*** **forwarding-class** ***forwarding-class-name*** **scheduler** ***scheduler-name***. The forwarding class identifies the output queue. Forwarding classes are mapped to output queues by default, and can be remapped to different queues by explicit user configuration. The scheduler name is the scheduler configured in Step 2.

4. On switches except QFX10000, associate the scheduler map with a traffic control profile using the statement **set class-of-service traffic-control-profiles *tcp-name* scheduler-map *map-name***. The scheduler map name is the name configured in Step 3.

5. On switches except QFX10000, associate the traffic control profile with an interface using the statement **set class-of-service interfaces *interface-name* forwarding-class-set *forwarding-class-set-name* output-traffic-control-profile *tcp-name***. The output traffic control profile name is the name of the traffic control profile configured in Step 4.

   The interface uses the scheduler map in the traffic control profile to apply the drop profile (and other attributes) to the output queue (forwarding class) on that interface. Because you can use different traffic control profiles to map different schedulers to different interfaces, the same queue number on different interfaces can handle traffic in different ways.

6. On QFX10000 switches, associate the scheduler map with an interface using the statement **set class-of-service interfaces *interface-name* scheduler-map *scheduler-map-name*** .

   The interface uses the scheduler map to apply the drop profile (and other attributes) to the output queue mapped to the forwarding class on that interface. Because you can use different scheduler maps on different interfaces, the same queue number on different interfaces can handle traffic in different ways.

Drop Profiles on Explicit Congestion Notification Enabled Queues

You must configure a WRED drop profile on queues that you enable for explicit congestion notification (ECN). On ECN-enabled queues, the drop profile sets the threshold for when the queue should mark a packet as experiencing congestion (see "Understanding CoS Explicit Congestion Notification" on page 380). When a queue fills to the level at which the WRED drop profile has a packet drop probability greater than zero (0), the switch might mark a packet as experiencing congestion. Whether or not a switch marks a packet as experiencing congestion is the same probability as the drop probability of the queue at that fill level.

On ECN-enabled queues, the switch does not use the drop profile to control dropping packets that are not ECN-capable packets (packets marked non-ECT, ECN code bits 00) during periods of congestion. Instead, the switch uses the tail-drop algorithm to drop non-ECN-capable packets during periods of congestion. When a queue fills to its maximum level of fullness, tail-drop simply drops all subsequently arriving packets until there is space in the queue to buffer more packets. All non-ECN-capable packets are treated the same way.

To apply a WRED drop profile to non-ECT traffic, configure a multifield (MF) classifier to assign non-ECT traffic to a different output queue that is not ECN-enabled, and then apply the WRED drop profile to that queue.

Related Documentation

• Understanding Junos CoS Components on page 15

Configuring CoS WRED Drop Profiles

You can configure an interpolated weighted random early detection (WRED) profile to control traffic congestion by controlling packet drop characteristics for different packet loss priorities.

Drop profiles specify two values, which work as pairs:

- Fill level—The queue fullness value, which represents a percentage of the memory used to store packets in relation to the total amount of memory allocated to the queue.
- Drop probability—The percentage value that corresponds to the likelihood that an individual packet is dropped.

NOTE: Do not enable WRED on lossless traffic flows (forwarding classes configured with the no-loss packet drop attribute). Use priority-based flow control (PFC) to prevent packet loss on lossless forwarding classes.

Except on QFX10000, you cannot enable WRED on multidestination (multicast) queues on. You can enable WRED only on unicast queues.

OCX Series switches do not support lossless flows or PFC.

NOTE: On ECN-enabled queues, the drop profile sets the threshold for when the queue should mark a packet as experiencing congestion (see "Understanding CoS Explicit Congestion Notification" on page 380). On ECN-enabled queues, the switch does not use the drop profile to control dropping packets that are not ECN-capable packets during periods of congestion. Instead, the switch uses the tail-drop algorithm to drop non-ECN-capable packets during periods of congestion. When a queue fills to its maximum level of fullness, tail-drop simply drops all subsequently arriving packets until there is space in the queue to buffer more packets. All non-ECN-capable packets are treated the same way.

Drop Profiles on Switches Except QFX10000

Interpolated means that the switch creates a smooth drop curve from a drop start point to a drop end point, with a maximum drop rate that is reached at the drop end point.

The dropstart point is the average queue fill level when the WRED algorithm starts to drop packets. Before the drop start point, no packets are scheduled to drop. Specify the drop start point using the first of two **fill-level** statements.

The drop end point is the average queue fill level at which all subsequently arriving packets are dropped. When the queue fill levels falls below the drop end point, packets begin to be forwarded again. (At the drop end point, the packet drop probability becomes 100 percent.) Specify the drop end point using the second of two **fill-level** statements.

The minimum drop rate is always **0**. Specify the minimum drop rate using the first of two **drop-probability** statements. The maximum drop rate is the drop probability when the average queue fill level reaches the drop end point. Specify the maximum drop rate using the second of two **drop-probability** statements.

The drop rate is zero until the queue fill level reaches the drop start point. As the queue continues to fill, packets drop in smooth linear curveuntil the queue reaches the drop end point, when packets drop at the maximum drop rate. If the queue fills beyond the drop end point, all packets that match the drop profile are dropped.

To configure a WRED profile using the CLI on switches except QFX10000:

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
```


Drop Profiles on QFX 10000 Switches

Each queue fill level pairs with a drop probability. As the queue fills to different levels, every time it reaches afill level configured in a drop profile, the queue applies thedrop probability paired with that fill level to the traffic in the queue that exceeds the fill level. You can configure up to 32 pairs of fill levels and drop probabilities to create a customized packet drop probability curve with up to 32 points of differentiation.

Packets are not dropped until they reach the first configured queue fill level. When the queue reaches the firstfill level, packets begin to drop at the configured drop probability rate paired with the first fill level. When the queue reaches the second fill level, packets begin to drop at the configured drop probability rate paired with the second fill level. This process continues for the number of fill level/drop probability pairs that you configure in the drop profile.

Drop profiles are *interpolated*. An interpolated drop profile gradually increases the drop probability along a curve between each configured fill level. When the queue reaches the next fill level, the drop probability reaches the drop probability paired with that fill level.

To configure a WRED profile using the CLI on QFX10000 switches:

Name the drop profile and set the fill levels and their associated drop probabilities as percentages. For every fill level, there must be a paired drop probability (you must configure the same number of fill levels and drop probabilities).

```
[edit class-of-service]
user@switch# set drop-profile drop-profile-name interpolate fill-level level1 level2 ... level32
drop-probability probability1 probability2 ... probability32
```

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring WRED Drop Profiles on page 371
- Defining CoS Queue Schedulers on page 290
- *Defining CoS Queue Schedulers for Port Scheduling*
- Configuring CoS Drop Profile Maps on page 377
- Understanding CoS WRED Drop Profiles on page 362

Example: Configuring WRED Drop Profiles

You can configure interpolated weighted random early detection (WRED) profiles to control traffic congestion by controlling packet drop characteristics for different packet loss priorities.

NOTE: Do not enable WRED on lossless traffic flows. Use priority-based flow control (PFC) to prevent packet loss on lossless forwarding classes. (OCX Series switches do not support lossless flows or PFC.)

Except on QFX10000 switches, you cannot enable WRED on multidestination (multicast) queues. You can enable WRED only on unicast queues.

Requirements

This example uses the following hardware and software components:

- One switch
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series or Junos OS Release 15.1X53-D10 or later for the QFX10000.

Overview

You associate WRED drop profiles with loss priorities in a scheduler. When you map the scheduler to a forwarding class (queue), you apply the interpolated drop profile to traffic of the specified loss priority on that queue. Drop profiles specify two values, which work as pairs:

- Fill level—The queue fullness value, which represents a percentage of the memory used to store packets in relation to the total amount of memory allocated to the queue.
- Drop probability—The percentage value that corresponds to the likelihood that an individual packet is dropped.

NOTE: On ECN-enabled queues, the drop profile sets the threshold for when the queue should mark a packet as experiencing congestion (see "Understanding CoS Explicit Congestion Notification" on page 380). On ECN-enabled queues, the switch does not use the drop profile to control dropping packets that are not ECN-capable packets during periods of congestion. Instead, the switch uses the tail-drop algorithm to drop non-ECN-capable packets during periods of congestion. When a queue fills to its maximum level of fullness, tail-drop simply drops all subsequently arriving packets until there is space in the queue to buffer more packets. All non-ECN-capable packets are treated the same way.

Configuring WRED Drop Profiles on Switches Except QFX10000

Configuration

Step-by-Step Procedure

Interpolated means that the switch creates a smooth drop curve from a drop start point to a drop end point, with a maximum drop rate that is reached at the drop end point:

- Drop start point—Percentage of average queue fill level when the WRED algorithm starts to drop packets. Before the drop start point, no packets are scheduled to drop.
- Drop end point—Average queue fill level at which all subsequently arriving packets are dropped. When the queue fill levels falls below the drop end point, packets begin to be forwarded again. (At the drop end point, the packet drop probability becomes 100 percent.)
- Maximum drop rate—Drop probability when the average queue fill level reaches the drop end point.

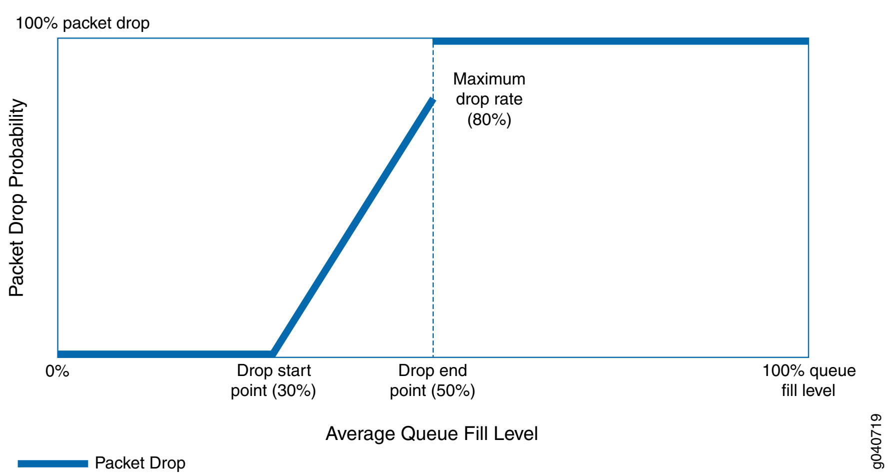
You set the drop start point and the drop end point by specifying two queue fill level percentage values. The first value is the drop start point and the second value is the drop end point.

You set the maximum drop rate by specifying two drop probability percentage values. The first value is always zero (0), which is the minimum drop rate, the probability of dropping a packet at the drop start point. The second value is the maximum drop rate at the drop end point.

The drop rate is zero until the queue fill level reaches the drop start point. As the queue continues to fill, packets drop in smooth linear curve until the queue reaches the drop end point, when packets drop at the maximum drop rate. If the queue fills beyond the drop end point, all packets that match the drop profile are dropped.

Figure 12 on page 374 shows the graph for a drop profile with a drop start point of 30 percent, a drop end point of 50 percent, and a maximum drop rate of 80 percent.

Figure 12: WRED Drop Profile Packet Drop Example



The graph shows that when the queue fill level is less than 30 percent, the packet drop rate is zero. When the queue fill level reaches 30 percent, packets begin to drop. As the queue fills, a higher percentage of packets drop. When the queue fill level reaches 50 percent, the packet drop rate has climbed to 80 percent. When the queue fill level exceeds 50 percent, all packets drop.

This example describes how to configure the drop profile shown in Figure 12 on page 374. The drop profile will have:

- The name **be-dp1**
- 30 percent for the drop start point (first **fill-level** setting)
- 50 percent for the drop end point (second **fill-level** setting)
- 0 percent for the minimum drop rate (first **drop-probability** setting)
- 80 percent for the maximum drop rate (second **drop-probability** setting)

You apply a drop profile by configuring a drop profile map that maps the drop profile to a packet loss priority, and associate the drop profile and packet loss priority with a scheduler. When you map the scheduler to a forwarding class (queue), the switch applies the drop profile to the packets in the forwarding class that have a matching packet loss priority.

1. Set the drop start point at **30** percent, the drop end point at **50** percent, the minimum drop rate at **0** percent, and the maximum drop rate at **80** percent for the drop profile **be-dp1**:

```
[edit class-of-service]
user@switch# set drop-profile be-dp1 interpolate fill-level 30 fill-level 50 drop-probability
0 drop-probability 80
```

Verification

Verifying the Drop Profile Configuration

Purpose Verify that you configured the drop profile **be-dp1** with the correct drop start and end points and with the correct drop rates.

Action Verify the results of the drop profile configuration using the operational mode command **show configuration class-of-service drop-profiles be-dp1**:

```
user@switch> show configuration class-of-service drop-profiles be-dp1
interpolate {
    fill-level [ 30 50 ];
    drop-probability [ 0 80 ];
}
```


Configuring WRED Drop Profiles on QFX10000 Switches

Configuration

Step-by-Step Procedure

Each queue fill level pairs with a drop probability. As the queue fills to different levels, every time it reaches a fill level configured in a drop profile, the queue applies the drop probability paired with that fill level to the traffic in the queue that exceeds the fill level. You can configure up to 32 pairs of fill levels and drop probabilities to create a customized packet drop probability curve with up to 32 points of differentiation.

Packets are not dropped until they reach the first configured queue fill level. When the queue reaches the first fill level, packets begin to drop at the configured drop probability rate paired with the first fill level. When the queue reaches the second fill level, packets begin to drop at the configured drop probability rate paired with the second fill level. This process continues for the number of fill level/drop probability pairs that you configure in the drop profile.

Drop profiles are *interpolated*. An interpolated drop profile gradually increases the drop probability along a curve between each configured fill level. When the queue reaches the next fill level, the drop probability reaches the drop probability paired with that fill level.

This example describes how to configure a drop profile with three fill level/drop probability pairs:

- Drop profile name—**be-dp1**
- Queue fill levels—25 percent, 50 percent, 75 percent
- Drop probabilities—30 percent, 60 percent, 100 percent

Each of the three fill levels pairs with a drop probability to program the interpolated drop profile curve.

You apply a drop profile by configuring a drop profile map that maps the drop profile to a packet loss priority, and associate the drop profile and packet loss priority with a scheduler. When you map the scheduler to a forwarding class (queue), the switch applies the drop profile to the packets in the forwarding class that have a matching packet loss priority.

To configure a drop profile:

1. Set the drop start point at a **25** percent fill level, an intermediate fill level of **50** percent, and a drop end point of **75** percent. Set the paired drop probabilities to **30** percent, **60** percent, and **100** percent, respectively, for drop profile **be-dp1**:

```
[edit class-of-service]
user@switch# set drop-profile be-dp1 interpolate fill-level [ 25 50 75 ] drop-probability [
30 60 100 ]
```

Verification

Verifying the Drop Profile Configuration

Purpose Verify that you configured the drop profile **be-dp1** with the correct fill levels and drop probabilities.

Action Verify the results of the drop profile configuration using the operational mode command **show configuration class-of-service drop-profiles be-dp1**:

```
user@switch> show configuration class-of-service drop-profiles be-dp1
interpolate {
    fill-level [ 25 50 75 ];
    drop-probability [ 30 60 100 ];
}
```

Configuring CoS Drop Profile Maps

A drop-profile map associates weighted random early detection (WRED) profiles for traffic of specified packet loss priorities with a scheduler. When you use a scheduler map to map a scheduler to a forwarding class, the drop profile map associated with the scheduler applies the specified WRED drop profile to traffic in the forwarding class that matches the specified packet loss priority.

Drop profile maps enable you to configure different drop profiles for traffic of different packet loss priorities within the same scheduler. You can associate different drop profiles with low-priority, medium-high priority, and high-priority traffic within a single scheduler, and then map that scheduler to a forwarding class. This applies the appropriate drop profile to traffic of each loss priority in a forwarding class. Drop profile maps apply to all traffic protocols.

To configure a drop-profile map:

- For the desired scheduler, configure the traffic loss priority and specify the drop profile you want to use to control the drop characteristics for traffic of that loss priority:

  ```
  [edit class-of-service]
  user@switch# set schedulers scheduler-name drop-profile-map loss-priority level protocol any drop-profile drop-profile-name
  ```

> NOTE: QFX10000 switches do not support the **protocol any** portion of the configuration. Drop profiles apply to all protocols.

Example: Configuring Drop Profile Maps

A drop-profile map associates weighted random early detection (WRED) profiles for traffic of specified packet loss priorities with a scheduler. When you use a scheduler map to map a scheduler to a forwarding class, the drop profile map associated with the scheduler applies the specified WRED drop profile to traffic in the forwarding class that matches the specified packet loss priority.

- Requirements on page 378
- Overview on page 378
- Configuring a Drop Profile Map on page 378
- Verification on page 379

Requirements

This example uses the following hardware and software components:

- A Juniper Networks QFX3500 Switch
- Junos OS Release 11.1 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series.

Overview

Drop profile maps enable you to configure different drop profiles for traffic of different packet loss priorities within the same scheduler. You can associate different drop profiles with low-priority, medium-high priority, and high-priority traffic within a single scheduler, and then map that scheduler to a forwarding class. This applies the appropriate drop profile to traffic of each loss priority in a forwarding class. Drop profile maps apply to all traffic protocols.

The following example describes how to configure a drop profile map for a scheduler named **mylan** that includes:

- A drop profile called **lp-profile** for low-priority traffic
- A drop profile called **mh-profile** for medium-high priority traffic
- A drop profile called **h-profile** for high-priority traffic

You apply the drop profiles in the drop profile map to a forwarding class by associating the scheduler **mylan** with a forwarding class in a scheduler map.

Configuring a Drop Profile Map

CLI Quick Configuration

To quickly configure a drop profile map, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the [edit] hierarchy level.

```
[edit class-of-service]
set schedulers mylan drop-profile-map loss-priority low protocol any drop-profile lp-profile
set schedulers mylan drop-profile-map loss-priority medium-high protocol any drop-profile
mh-profile
set schedulers mylan drop-profile-map loss-priority high protocol any drop-profile h-profile
```

To configure a drop profile map:

1. Configure the drop profile for low-priority traffic:

```
[edit class-of-service]
user@switch# set schedulers mylan drop-profile-map loss-priority low protocol any
drop-profile lp-profile
```

2. Configure the drop profile for medium-high priority traffic:

```
[edit class-of-service]
user@switch# set schedulers mylan drop-profile-map loss-priority medium-high protocol
any drop-profile mh-profile
```

3. Configure the drop profile for high-priority traffic:

```
[edit class-of-service]
user@switch# set schedulers mylan drop-profile-map loss-priority high protocol any
drop-profile h-profile
```

Verification

Verifying the Drop Profile Map Configuration

Purpose Verify that you configured the drop profile map for the scheduler **mylan** with the correct loss priorities and drop profiles.

Action Verify the results of the drop profile map configuration using the operational mode command **show configuration class-of-service schedulers mylan**:

```
user@switch> show configuration class-of-service schedulers mylan
transmit-rate 3g;
shaping-rate percent 100;
priority low;
drop-profile-map loss-priority low protocol any drop-profile lp-profile;
drop-profile-map loss-priority medium-high protocol any drop-profile mh-profile;
drop-profile-map loss-priority high protocol any drop-profile h-profile;
```

NOTE: This example does not include configuring scheduler bandwidth and priority. This information (transmit rate, shaping rate, and priority) is shown for completeness.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- *Example: Configuring Queue Schedulers for Port Scheduling*
- Example: Configuring WRED Drop Profiles on page 371

- Configuring CoS Drop Profile Maps on page 377
- Understanding CoS WRED Drop Profiles on page 362

Understanding CoS Explicit Congestion Notification

Explicit congestion notification (ECN) enables end-to-end congestion notification between two endpoints on TCP/IP based networks. The two endpoints are an ECN-enabled sender and an ECN-enabled receiver. ECN must be enabled on both endpoints and on all of the intermediate devices between the endpoints for ECN to work properly. Any device in the transmission path that does not support ECN breaks the end-to-end ECN functionality.

ECN notifies networks about congestion with the goal of reducing packet loss and delay by making the sending device decrease the transmission rate until the congestion clears, without dropping packets. RFC 3168, *The Addition of Explicit Congestion Notification (ECN) to IP*, defines ECN.

ECN is disabled by default. Normally, you enable ECN only on queues that handle best-effort traffic because other traffic types use different methods of congestion notification—lossless traffic uses priority-based flow control (PFC) and strict-high priority traffic receives all of the port bandwidth it requires up to the point of a configured maximum rate.

NOTE: OCX Series switches do not support lossless transport and do not support PFC.

You enable ECN on individual output queues (as represented by forwarding classes) by enabling ECN in the queue scheduler configuration, mapping the scheduler to forwarding classes (queues), and then applying the scheduler to interfaces.

NOTE: For ECN to work on a queue, you must also apply a weighted random early detection (WRED) packet drop profile to the queue.

- How ECN Works on page 380
- WRED Drop Profile Control of ECN Thresholds on page 385
- Support, Limitations, and Notes on page 388

How ECN Works

Without ECN, switches respond to network congestion by dropping TCP/IP packets. Dropped packets signal the network that congestion is occurring. Devices on the IP network respond to TCP packet drops by reducing the packet transmission rate to allow the congestion to clear. However, the packet drop method of congestion notification and management has some disadvantages. For example, packets are dropped and must be retransmitted. Also, bursty traffic can cause the network to reduce the transmission rate too much, resulting in inefficient bandwidth utilization.

Instead of dropping packets to signal network congestion, ECN marks packets to signal network congestion, without dropping the packets. For ECN to work, all of the switches in the path between two ECN-enabled endpoints must have ECN enabled. ECN is negotiated during the establishment of the TCP connection between the endpoints.

ECN-enabled switches determine the queue congestion state based on the WRED packet drop profile configuration applied to the queue, so each ECN-enabled queue must also have a WRED drop profile. If a queue fills to the level at which the WRED drop profile has a packet drop probability greater than zero (0), the switch might mark a packet as experiencing congestion. Whether or not a switch marks a packet as experiencing congestion is the same probability as the drop probability of the queue at that fill level.

ECN communicates whether or not congestion is experienced by marking the two least-significant bits in the differentiated services (DiffServ) field in the IP header. The most significant six bits in the DiffServ field contain the Differentiated Services Code Point (DSCP) bits. The state of the two ECN bits signals whether or not the packet is an ECN-capable packet and whether or not congestion has been experienced.

ECN-capable senders mark packets as ECN-capable. If a sender is not ECN-capable, it marks packets as not ECN-capable. If an ECN-capable packet experiences congestion at the egress queue of a switch, the switch marks the packet as experiencing congestion. When the packet reaches the ECN-capable receiver (destination endpoint), the receiver echoes the congestion indicator to the sender (source endpoint) by sending a packet marked to indicate congestion.

After receiving the congestion indicator from the receiver, the source endpoint reduces the transmission rate to relieve the congestion. This is similar to the result of TCP congestion notification and management, but instead of dropping the packet to signal network congestion, ECN marks the packet and the receiver echoes the congestion notification to the sender. Because the packet is not dropped, the packet does not need to be retransmitted.

ECN Bits in the DiffServ Field

The two ECN bits in the DiffServ field provide four codes that determine if a packet is marked as an ECN-capable transport (ECT) packet, meaning that both endpoints of the transport protocol are ECN-capable, and if there is congestion experienced (CE), as shown in Table 75 on page 381:

Table 75: ECN Bit Codes

ECN Bits (Code)	Meaning
00	Non-ECT—Packet is marked as not ECN-capable
01	ECT(1)—Endpoints of the transport protocol are ECN-capable

Table 75: ECN Bit Codes (continued)

ECN Bits (Code)	Meaning
10	ECT(0)—Endpoints of the transport protocol are ECN-capable
11	CE—Congestion experienced

Codes 01 and 10 have the same meaning: the sending and receiving endpoints of the transport protocol are ECN-capable. There is no difference between these codes.

End-to-End ECN Behavior

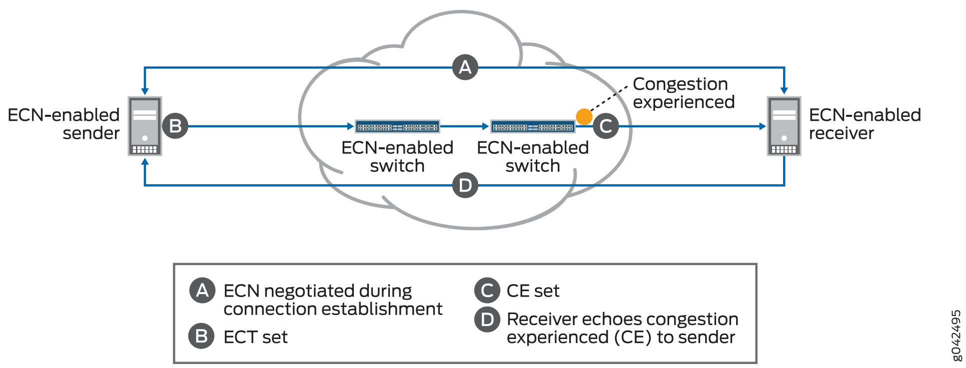
After the sending and receiving endpoints negotiate ECN, the sending endpoint marks packets as ECN-capable by setting the DiffServ ECN field to ECT(1) (01) or ECT(0) (10). Every intermediate switch between the endpoints must have ECN enabled or it does not work.

When a packet traverses a switch and experiences congestion at an output queue that uses the WRED packet drop mechanism, the switch marks the packet as experiencing congestion by setting the DiffServ ECN field to CE (11). Instead of dropping the packet (as with TCP congestion notification), the switch forwards the packet.

NOTE: **At the egress queue, the WRED algorithm determines whether or not a packet is drop eligible based on the queue fill level (how full the queue is). If a packet is drop eligible and marked as ECN-capable, the packet can be marked CE and forwarded. If a packet is drop eligible and is not marked as ECN-capable, it might be dropped. See** “WRED Drop Profile Control of ECN Thresholds” on page 385 **for more information about the WRED algorithm.**

When the packet reaches the receiver endpoint, the CE mark tells the receiver that there is network congestion. The receiver then sends (echoes) a message to the sender that indicates there is congestion on the network. The sender acknowledges the congestion notification message and reduces its transmission rate. Figure 13 on page 383 summarizes how ECN works to mitigate network congestion:

Figure 13: Explicit Congestion Notification



End-to-end ECN behavior includes:

1. The ECN-capable sender and receiver negotiate ECN capability during the establishment of their connection.

2. After successful negotiation of ECN capability, the ECN-capable sender sends IP packets with the ECT field set to the receiver.

   NOTE: **All of the intermediate devices in the path between the sender and the receiver must be ECN-enabled.**

3. If the WRED algorithm on a switch egress queue determines that the queue is experiencing congestion and the packet is drop eligible, the switch can mark the packet as “congestion experienced” (CE) to indicate to the receiver that there is congestion on the network. If the packet has already been marked CE (congestion has already been experienced at the egress of another switch), the switch forwards the packet with CE marked.

   If there is no congestion at the switch egress queue, the switch forwards the packet and does not change the ECT-enabled marking of the ECN bits, so the packet is still marked as ECN-capable but not as experiencing congestion.

   On QFX5210, QFX5200, QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, packets that are not marked as ECN-capable (ECT, 00) are treated according to the WRED drop profile configuration and might be dropped during periods of congestion.

   On QFX10000 switches, the switch uses the tail-drop algorithm to drop packets that are marked ECT (00) during periods of congestion. (When a queue fills to its maximum level of fullness, tail-drop simply drops all subsequently arriving packets until there is space in the queue to buffer more packets. All non-ECN-capable packets are treated the same.)

4. The receiver receives a packet marked CE to indicate that congestion was experienced along the congestion path.

5. The receiver echoes (sends) a packet back to the sender with the ECE bit (bit 9) marked in the flag field of the TCP header. The ECE bit is the ECN echo flag bit, which notifies the sender that there is congestion on the network.

6. The sender reduces the data transmission rate and sends a packet to the receiver with the CWR bit (bit 8) marked in the flag field of the TCP header. The CWR bit is the congestion window reduced flag bit, which acknowledges to the receiver that the congestion experienced notification was received.

7. When the receiver receives the CWR flag, the receiver stops setting the ECE bit in replies to the sender.

Table 76 on page 384 summarizes the behavior of traffic on ECN-enabled queues.

Table 76: Traffic Behavior on ECN-Enabled Queues

Incoming IP Packet Marking of ECN Bits	ECN Configuration on the Output Queue	Action if WRED Algorithm Determines Packet is Drop Eligible	Outgoing Packet Marking of ECN Bits
Non-ECT (00)	Does not matter	Drop (QFX5210, QFX5200, QFX5100, EX4600, QFX3500, QFX3600, QFabric systems). Tail drop occurs when queue reaches maximum fullness because no WRED drop probability is applied (QFX10000 switches).	No ECN bits marked
ECT (10 or 01)	ECN disabled	Drop	Packet dropped—no ECN bits marked
ECT (10 or 01)	ECN enabled	Do not drop. Mark packet as experiencing congestion (CE, bits 11).	Packet marked ECT (11) to indicate congestion
CE (11)	ECN disabled	Drop	Packet dropped—no ECN bits marked
CE (11)	ECN enabled	Do not drop. Packet is already marked as experiencing congestion, forward packet without changing the ECN marking.	Packet marked ECT (11) to indicate congestion

When an output queue is not experiencing congestion as defined by the WRED drop profile mapped to the queue, all packets are forwarded, and no packets are dropped.

ECN Compared to PFC and Ethernet PAUSE

ECN is an end-to-end network congestion notification mechanism for IP traffic. Priority-based flow control (PFC) (IEEE 802.1Qbb) and Ethernet PAUSE (IEEE 802.3X) are different types of congestion management mechanisms.

NOTE: QFX10000 switches do not support Ethernet PAUSE.

OCX Series switches do not support PFC. OCX Series switches support Ethernet PAUSE on tagged Layer 3 interfaces.

ECN requires that an output queue must also have an associated WRED packet drop profile. Output queues used for traffic on which PFC is enabled should not have an associated WRED drop profile. Interfaces on which Ethernet PAUSE is enabled should not have an associated WRED drop profile.

PFC is a peer-to-peer flow control mechanism to support lossless traffic. PFC enables connected peer devices to pause flow transmission during periods of congestion. PFC enables you to pause traffic on a specified type of flow on a link instead of on all traffic on a link. For example, you can (and should) enable PFC on lossless traffic classes such as the **fcoe** forwarding class. Ethernet PAUSE is also a peer-to-peer flow control mechanism, but instead of pausing only specified traffic flows, Ethernet PAUSE pauses all traffic on a physical link.

With PFC and Ethernet PAUSE, the sending and receiving endpoints of a flow do not communicate congestion information to each other across the intermediate switches. Instead, PFC controls flows between two PFC-enabled peer devices (for example, switches) that support data center bridging (DCB) standards. PFC works by sending a pause message to the connected peer when the flow output queue becomes congested. Ethernet PAUSE simply pauses all traffic on a link during periods of congestion and does not require DCB.

PFC works this way: if a switch output queue fills to a certain threshold, the switch sends a PFC pause message to the connected peer device that is transmitting data. The pause message tells the transmitting switch to pause transmission of the flow. When the congestion clears, the switch sends another PFC message to tell the connected peer to resume transmission. (If the output queue of the transmitting switch also reaches a certain threshold, that switch can in turn send a PFC pause message to the connected peer that is transmitting to it. In this way, PFC can propagate a transmission pause back through the network.)

See “Understanding CoS Flow Control (Ethernet PAUSE and PFC)” on page 525 for more information. For QFX5100 and EX4600 switches only, you can also refer to *Understanding PFC Functionality Across Layer 3 Interfaces*.

WRED Drop Profile Control of ECN Thresholds

You apply WRED drop profiles to forwarding classes (which are mapped to output queues) to control how the switch marks ECN-capable packets. A scheduler map associates a drop profile with a scheduler and a forwarding class, and then you apply the scheduler map to interfaces to implement the scheduling properties for the forwarding class on those interfaces.

Drop profiles define queue fill level (the percentage of queue fullness) and drop probability (the percentage probability that a packet is dropped) pairs. When a queue fills to a specified level, traffic that matches the drop profile has the drop probability paired with that fill level. When you configure a drop profile, you configure pairs of fill levels and drop probabilities to control how packets drop at different levels of queue fullness.

The first fill level and drop probability pair is the drop start point. Until the queue reaches the first fill level, packets are not dropped. When the queue reaches the first fill level, packets that exceed the fill level have a probability of being dropped that equals the drop probability paired with the fill level.

The last fill level and drop probability pair is the drop end point. When the queue reaches the last fill level, all packets are dropped unless they are configured for ECN.

NOTE: Lossless queues (forwarding class configured with the **no-loss** packet drop attribute) and strict-high priority queues do not use drop profiles. Lossless queues use PFC to control the flow of traffic. Strict-high priority queues receive all of the port bandwidth they require up to the configured maximum bandwidth limit (scheduler **transmit-rate** on QFX10000 switches, and **shaping-rate** on QFX5210, QFX5200, QFX5100, QFX3500, QFX3600, and EX4600 switches, and QFabric systems).

Different switches support different amounts of fill level/drop probability pairs in drop profiles. For example, QFX10000 switches support 32 fill level/drop probability pairs, so there can be as many as 30 intermediate fill level/drop probability pairs between the drop start and drop endpoints. QFX5210, QFX5200, QFX5100, QFX3500, QFX3600, and EX4600 switches, and QFabric systems support two fill level/drop probability pairs—by definition, the two pairs you configure on these switches are the drop start and drop end points.

NOTE: Do not configure the last fill level as 100 percent.

The drop profile configuration affects ECN packets as follows:

- Drop start point—ECN-capable packets might be marked as congestion experienced (CE).
- Drop end point—ECN-capable packets are always marked CE.

As a queue fills from the drop start point to the drop end point, the probability that an ECN packet is marked CE is the same as the probability that a non-ECN packet is dropped if you apply the drop profile to best-effort traffic. As the queue fills, the probability of an ECN packet being marked CE increases, just as the probability of a non-ECN packet being dropped increases when you apply the drop profile to best-effort traffic.

At the drop end point, all ECN packets are marked CE, but the ECN packets are not dropped. When the queue fill level exceeds the drop end point, all ECN packets are marked CE. (At this point on QFX5210, QFX5200, QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, all non-ECN packets are dropped.) ECN packets (and all other packets) are tail-dropped if the queue fills completely.

To configure a WRED packet drop profile and apply it to an output queue (using hierarchical scheduling on switches that support ETS):

1. Configure a drop profile using the statement **set class-of-service drop-profiles** ***profile-name*** **interpolate fill-level** ***drop-start-point*** **fill-level** ***drop-end-point*** **drop-probability 0 drop-probability** ***percentage***.

2. Map the drop profile to a queue scheduler using the statement **set class-of-service schedulers** ***scheduler-name*** **drop-profile-map loss-priority (low | medium-high | high) protocol any drop-profile** ***profile-name***. The name of the drop-profile is the name of the WRED profile configured in Step 1.

3. Map the scheduler, which Step 2 associates with the drop profile, to the output queue using the statement **set class-of-service scheduler-maps** ***map-name*** **forwarding-class** ***forwarding-class-name*** **scheduler** ***scheduler-name***. The forwarding class identifies the output queue. Forwarding classes are mapped to output queues by default, and can be remapped to different queues by explicit user configuration. The scheduler name is the scheduler configured in Step 2.

4. Associate the scheduler map with a traffic control profile using the statement **set class-of-service traffic-control-profiles** ***tcp-name*** **scheduler-map** ***map-name***. The scheduler map name is the name configured in Step 3.

5. Associate the traffic control profile with an interface using the statement **set class-of-service interface** ***interface-name*** **forwarding-class-set** ***forwarding-class-set-name*** **output-traffic-control-profile** ***tcp-name***. The output traffic control profile name is the name of the traffic control profile configured in Step 4.

   The interface uses the scheduler map in the traffic control profile to apply the drop profile (and other attributes, including the enable ECN attribute) to the output queue (forwarding class) on that interface. Because you can use different traffic control profiles to map different schedulers to different interfaces, the same queue number on different interfaces can handle traffic in different ways.

Starting in Release 15.1, you can configure a WRED packet drop profile and apply it to an output queue on switches that support port scheduling (ETS hierarchical scheduling is either not supported or not used). To configure a WRED packet drop profile and apply it to an output queue on switches that support port scheduling (ETS hierarchical scheduling is either not supported or not used):

1. Configure a drop profile using the statement **set class-of-service drop-profiles *profile-name* interpolate fill-level *level1 level2 ... level32* drop-probability *probability1 probability2 ... probability32***. You can specify as few as two fill level/drop probability pairs or as many as 32 pairs.

2. Map the drop profile to a queue scheduler using the statement **set class-of-service schedulers *scheduler-name* drop-profile-map loss-priority (low | medium-high | high) drop-profile *profile-name***. The name of the drop-profile is the name of the WRED profile configured in Step 1.

3. Map the scheduler, which Step 2 associates with the drop profile, to the output queue using the statement **set class-of-service scheduler-maps *map-name* forwarding-class *forwarding-class-name* scheduler *scheduler-name***. The forwarding class identifies the output queue. Forwarding classes are mapped to output queues by default, and can be remapped to different queues by explicit user configuration. The scheduler name is the scheduler configured in Step 2.

4. Associate the scheduler map with an interface using the statement **set class-of-service interfaces *interface-name* scheduler-map *scheduler-map-name***.

   The interface uses the scheduler map to apply the drop profile (and other attributes) to the output queue mapped to the forwarding class on that interface. Because you can use different scheduler maps on different interfaces, the same queue number on different interfaces can handle traffic in different ways.

Support, Limitations, and Notes

If the WRED algorithm that is mapped to a queue does not find a packet drop eligible, then the ECN configuration and ECN bits marking does not matter. The packet transport behavior is the same as when ECN is not enabled.

ECN is disabled by default. Normally, you enable ECN only on queues that handle best-effort traffic, and you do not enable ECN on queues that handle lossless traffic or strict-high priority traffic.

ECN supports the following:

- IPv4 and IPv6 packets
- Untagged, single-tagged, and double-tagged packets
- The outer IP header of IP tunneled packets (but not the inner IP header)

ECN does not support the following:

- IP packets with MPLS encapsulation
- The inner IP header of IP tunneled packets (however, ECN works on the outer IP header)
- Multicast, broadcast, and destination lookup fail (DLF) traffic
- Non-IP traffic

NOTE: On QFX10000 switches, when you enable a queue for ECN and apply a WRED drop profile to the queue, the WRED drop profile only sets the thresholds for marking ECN traffic as experiencing congestion (CE, 11). On ECN-enabled queues, the WRED drop profile does not set drop thresholds for non-ECT (00) traffic (traffic that is not ECN-capable). Instead, the switch uses the tail-drop algorithm on traffic is that is marked non-ECT on ECN-enabled queues during periods of congestion.

To apply a WRED drop profile to non-ECT traffic, configure a multifield (MF) classifier to assign non-ECT traffic to a different output queue that is not ECN-enabled, and then apply the WRED drop profile to that queue.

Release History Table

Release	Description
15.1	Starting in Release 15.1, you can configure a WRED packet drop profile and apply it to an output queue on switches that support port scheduling (ETS hierarchical scheduling is either not supported or not used).

Related Documentation

- Example: Configuring ECN on page 389

Example: Configuring ECN

This example shows how to enable explicit congestion notification (ECN) on an output queue.

Requirements

This example uses the following hardware and software components:

- One switch.
- Junos OS Release 13.2X51-D25 or later for the QFX Series or Junos OS Release 14.1X53-D20 for the OCX Series

Overview

ECN enables end-to-end congestion notification between two endpoints on TCP/IP based networks. The two endpoints are an ECN-enabled sender and an ECN-enabled receiver. ECN must be enabled on both endpoints and on all of the intermediate devices between the endpoints for ECN to work properly. Any device in the transmission path that does not support ECN breaks the end-to-end ECN functionality

A weighted random early detection (WRED) packet drop profile must be applied to the output queues on which ECN is enabled. ECN uses the WRED drop profile thresholds to mark packets when the output queue experiences congestion.

ECN reduces packet loss by forwarding ECN-capable packets during periods of network congestion instead of dropping those packets. (TCP notifies the network about congestion by dropping packets.) During periods of congestion, ECN marks ECN-capable packets that egress from congested queues. When the receiver receives an ECN packet that is marked as experiencing congestion, the receiver echoes the congestion state back to the sender. The sender then reduces its transmission rate to clear the congestion.

ECN is disabled by default. You can enable ECN on best-effort traffic. ECN should not be enabled on lossless traffic queues, which uses priority-based flow control (PFC) for congestion notification, and ECN should not be enabled on strict-high priority traffic queues.

To enable ECN on an output queue, you not only need to enable ECN in the queue scheduler, you also need to:

- Configure a WRED packet drop profile.
- Configure a queue scheduler that includes the WRED drop profile and enables ECN. (This example shows only ECN and drop profile configuration; you can also configure bandwidth, priority, and buffer settings in a scheduler.)
- Map the queue scheduler to a forwarding class (output queue) in a scheduler map.
- Starting in Junos OS 15.1, enhanced transmission selection (ETS) hierarchical scheduling is supported. If you are using enhanced transmission selection (ETS) hierarchical scheduling, add the forwarding class to a forwarding class set (priority group).
- If you are using ETS, associate the queue scheduler map with a traffic control profile (priority group scheduler for hierarchical scheduling).
- If you are using ETS, apply the traffic control profile and the forwarding class set to an interface. On that interface, the output queue uses the scheduler mapped to the forwarding class, as specified by the scheduler map attached to the traffic control profile. This enables ECN on the queue and applies the WRED drop profile to the queue.

  If you are using port scheduling, apply the scheduler map to an interface. On that interface, the output queue uses the scheduler mapped to the forwarding class in the scheduler map, which enables ECN on the queue and applies the WRED drop profile to the queue.

Table 77 on page 391 shows the configuration components for this example.

Table 77: Components of the ECN Configuration Example

Component	Settings
Hardware	QFX Series switch
Drop profile (with two fill level/drop probability pairs)	Name: **be-dp** Drop start fill level: **30** percent Drop end fill level: **75** percent Drop probability at drop start (minimum drop rate): **0** percent Drop probability at drop end (maximum drop rate): **80** percent
Scheduler	Name: **be-sched** ECN: enabled Drop profile: **be-dp** Transmit rate: **25%** Buffer size: **25%** Priority: **low**
Scheduler map	Name: **be-map** Forwarding class: **best-effort** Scheduler: **be-sched** NOTE: By default, the **best-effort** forwarding class is mapped to output queue **0**.
Forwarding class set (ETS only)	Name: **be-pg** Forwarding class: **best-effort** (queue 0)
Traffic control profile (ETS only)	Name: **be-tcp** Scheduler map: **be-map**
Interface (ETS only)	Name: **xe-0/0/20** Forwarding class set: **be-pg** (Output) traffic control profile: **be-tcp**
Interface (port scheduling only)	Name: **xe-0/0/20**

NOTE: Only switches that support ETS hierarchical scheduling support forwarding class set and traffic control profile configuration. Direct port scheduling does not use the hierarchical scheduling structure.

NOTE: On QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, the WRED drop profile also controls packet drop behavior for traffic that is not ECN-capable (packets marked non-ECT, ECN bit code 00).

On QFX10000 switches, when ECN is enabled on a queue, the WRED drop profile only sets the ECN thresholds, it does not control packet drop on non-ECN packets. On ECN-enabled queues, QFX10000 switches use the tail-drop algorithm on non-ECN packets during periods of congestion. If you do not enable ECN, then the queue uses the WRED packet drop mechanism.

Configuration

CLI Quick Configuration

To quickly configure the drop profile, scheduler with ECN enabled, and to map the scheduler to an output queue on an interface, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the **[edit]** hierarchy level.

ETS Quick Configuration

```
[edit class-of-service]
set drop-profile be-dp interpolate fill-level 30 fill-level 75 drop-probability 0 drop-probability 80
set schedulers be-sched explicit-congestion-notification
set schedulers be-sched drop-profile-map loss-priority low protocol any drop-profile be-dp
set schedulers be-sched transmit-rate percent 25
set schedulers be-sched buffer-size percent 25
set schedulers be-sched priority low
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set forwarding-class-sets be-pg class best-effort
set traffic-control-profiles be-tcp scheduler-map be-map
set interfaces xe-0/0/20 forwarding-class-set be-pg output-traffic-control-profile be-tcp
```

Port Scheduling Quick Configuration (QFX10000 Switches)

```
[edit class-of-service]
set drop-profile be-dp interpolate fill-level 30 fill-level 75 drop-probability 0 drop-probability 80
set schedulers be-sched explicit-congestion-notification
set schedulers be-sched drop-profile-map loss-priority low protocol any drop-profile be-dp
set schedulers be-sched transmit-rate percent 25
set schedulers be-sched buffer-size percent 25
set schedulers be-sched priority low
set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
set interfaces xe-0/0/20 scheduler-map be-map
```


Configuring ECN

Step-by-Step Procedure

To configure ECN:

1. Configure the WRED packet drop profile **be-dp**. This example uses a drop start point of **30** percent, a drop end point of **75** percent, a minimum drop rate of **0** percent, and a maximum drop rate of **80** percent:

```
[edit class-of-service]
user@switch# set drop-profile be-dp interpolate fill-level 30 fill-level 75 drop-probability 0 drop-probability 80
```

2. Create the scheduler **be-sched** with ECN enabled and associate the drop profile **be-dp** with the scheduler:

```
[edit class-of-service]
user@switch# set schedulers be-sched explicit-congestion-notification
user@switch# set schedulers be-sched drop-profile-map loss-priority low protocol any drop-profile be-dp
user@switch# set be-sched transmit-rate percent 25
user@switch# set be-sched buffer-size percent 25
user@switch# set be-sched priority low
```

3. Map the scheduler **be-sched** to the **best-effort** forwarding class (output queue 0) using scheduler map **be-map**:

```
[edit class-of-service]
user@switch# set scheduler-maps be-map forwarding-class best-effort scheduler be-sched
```

4. If you are using ETS, add the forwarding class **best-effort** to the forwarding class set **be-pg**; if you are using direct port scheduling, skip this step:

```
[edit class-of-service]
user@switch# set forwarding-class-sets be-pg class best-effort
```

5. If you are using ETS, associate the scheduler map **be-map**with the traffic control profile **be-tcp**; if you are using direct port scheduling, skip this step:

```
[edit class-of-service]
user@switch# set traffic-control-profiles be-tcp scheduler-map be-map
```

6. If you are using ETS, associate the traffic control profile **be-tcp** and the forwarding class set **be-pg** with the interface on which you want to enable ECN on the best-effort queue:

```
[edit class-of-service]
user@switch# set interfaces xe-0/0/20 forwarding-class-set be-pg output-traffic-control-profile be-tcp
```

If you are using direct port scheduling, associate the scheduler map **be-map** with the interface on which you want to enable ECN on the best-effort queue:

```
[edit class-of-service]
user@switch# set interfaces xe-0/0/20 scheduler map be-map
```

Verification

Verifying That ECN Is Enabled

Purpose Verify that ECN is enabled in the scheduler **be-sched** by showing the configuration for the scheduler map **be-map**.

Action Display the scheduler map configuration using the operational mode command **show class-of-service scheduler-map be-map**:

```
user@switch> show class-of-service scheduler-map be-map
Scheduler map: be-map, Index: 12240

  Scheduler:be-sched, Forwarding class: best-effort, Index: 115
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified, Explicit Congestion Notification: enable
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any          3312    be-dp
      Medium-high     any          1       <default-drop-profile>
      High            any          1       <default-drop-profile>
```

Meaning The **show class-of-service scheduler-map** operational command shows the configuration of the scheduler associated with the scheduler map and the forwarding class mapped to that scheduler. The output shows that:

- The scheduler associated with the scheduler map is **be-sched**.
- The scheduler map applies to the forwarding class **best-effort** (output queue 0).
- The scheduler **be-sched** has a transmit rate of **25** percent, a queue buffer size of **25** percent, and a drop priority of **low**.
- Explicit congestion notification state is **enable**.
- The WRED drop profile used for low drop priority traffic is **be-dp**.

Release History Table

Release	Description
15.1	Starting in Junos OS 15.1, enhanced transmission selection (ETS) hierarchical scheduling is supported.

Related Documentation

- Understanding CoS Explicit Congestion Notification on page 380

Troubleshooting Egress Bandwidth That Exceeds the Configured Minimum Bandwidth

Problem

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

NOTE: On switches that support enhanced transmission selection (ETS) hierarchical scheduling, the switch allocates guaranteed minimum bandwidth first to a priority group using the guaranteed rate setting in the traffic control profile, and then allocates priority group minimum guaranteed bandwidth to forwarding classes in the priority group using the transmit rate setting in the queue scheduler.

On switches that support direct port scheduling, there is no scheduling hierarchy. The switch allocates port bandwidth to forwarding classes directly, using the transmit rate setting in the queue scheduler.

In this topic, if you are using direct port scheduling on your switch, ignore the references to priority groups and forwarding class sets (priority groups and forwarding class sets are only used for ETS hierarchical port scheduling). For direct port scheduling, only the transmit rate queue scheduler setting can cause the issue described in this topic.

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: For ETS, 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.)

For port scheduling, the sum of the queue transmit rates should not exceed the port bandwidth.

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

- **transmit-rate on page 426**
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350

Troubleshooting Egress Bandwidth That Exceeds the Configured Maximum Bandwidth

Problem **Description:** The maximum bandwidth of a queue when measured at the egress port exceeds the maximum bandwidth rate shaper (**shaping-rate** statement on QFX5200, QFX5100, EX4600, QFX3500, QFX3600, and OCX1100 switches, and on QFabric systems, and **transmit-rate (rate | percentage *percent* exact** statement on QFX10000 switches) configured for the queue.

Cause When you configure bandwidth for a queue (forwarding class) or a priority group (forwarding class set), 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.

Troubleshooting Egress Queue Bandwidth Impacted by Congestion

Problem

Description: 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 WRED 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 on page 406
- Example: Configuring WRED Drop Profiles on page 371
- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Understanding CoS WRED Drop Profiles on page 362

CHAPTER 8

Configuration Statements for Scheduling

buffer-size

Syntax buffer-size (percent *percent* | remainder);

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description On all switches, you configure the proportion of port buffers allocated to a particular output queue using the following process:

1. Configure a scheduler and set the **buffer-size** option.

2. Use a scheduler map to map the scheduler to the forwarding class that is mapped to the queue to which you want to apply the buffer size.

   For example, suppose that you want to change the dedicated buffer allocation for FCoE traffic. FCoE traffic is mapped to the fcoe forwarding class, and the fcoe forwarding class is mapped to queue 3 (this is the default configuration). To use default FCoE traffic mapping, in the scheduler map configuration, map the scheduler to the **fcoe** forwarding class.

3. If you are using enhanced transmission selection (ETS) hierarchical scheduling, associate the scheduler map with the traffic control profile you want to use on the egress ports that carry FCoE traffic. If you are using direct port scheduling, skip this step.

4. If you are using ETS, associate the traffic control profile that includes the scheduler map with the desired egress ports. For this example, you associate the traffic control profile with the ports that carry FCoE traffic. If you are using port scheduling, associate the scheduler map with the desired egress ports.

   Queue 3, which is mapped to the fcoe forwarding class and therefore to the FCoE traffic, receives the dedicated buffer allocation specified in the **buffer-size** statement.

NOTE: **The total of all of the explicitly configured buffer size percentages for all of the queues on a port cannot exceed 100 percent.**

QFX10000 Switches

On QFX10000 switches, the buffer size is the amount of time in milliseconds of port bandwidth that a queue can use to continue to transmit packets during periods of congestion, before the buffer runs out and packets begin to drop.

The switch can use up to 100 ms total (combined) buffer space for all queues on a port. A buffer-size configured as one percent is equal to 1 ms of buffer usage. A buffer-size of 15 percent (the default value for the best effort and network control queues) is equal to 15 ms of buffer usage.

The total buffer size of the switch is 4 GB. A 40-Gigabit port can use up to 500 MB of buffer space, which is equivalent to 100 ms of port bandwidth on a 40-Gigabit port. A 10-Gigabit port can use up to 125 MB of buffer space, which is equivalent to 100 ms of port bandwidth on a 10-Gigabit port. The total buffer sizes of the eight output queues on a port cannot exceed 100 percent, which is equal to the full 100 ms total buffer available to a port. The maximum amount of buffer space any queue can use is also 100 ms (which equates to a 100 percent buffer-size configuration), but if one queue uses all of the buffer, then no other queue receives buffer space.

There is no minimum buffer allocation, so you can set the buffer-size to zero (0) for a queue. However, we recommend that on queues on which you enable PFC to support lossless transport, you allocate a minimum of 5 ms (a minimum buffer-size of 5 percent). The two default lossless queues, fcoe and no-loss, have buffer-size default values of 35 ms (35 percent).

Queue buffer allocation is dynamic, shared among ports as needed. However, a queue cannot use more than its configured amount of buffer space. For example, if you are using the default CoS configuration, the best-effort queue receives a maximum of 15 ms of buffer space because the default transmit rate for the best-effort queue is 15 percent.

If a switch experiences congestion, queues continue to receives their full buffer allocation until 90 percent of the 4 GB buffer space is consumed. When 90 percent of the buffer space is in use, the amount of buffer space per port, per queue, is reduced in proportion to the configured buffer size for each queue. As the percentage of consumed buffer space rises above 90 percent, the amount of buffer space per port, per queue, continues to be reduced.

On 40-Gigabit ports, because the total buffer is 4 GB and the maximum buffer a port can use is 500 MB, up to seven 40-Gigabit ports can consume their full 100 ms allocation of buffer space. However, if an eighth 40-Gigabit port requires the full 500 MB of buffer space, then the buffer allocations are proportionally reduced because the buffer consumption is above 90 percent.

On 10-Gigabit ports, because the total buffer is 4 GB and the maximum buffer a port can use is 125 MB, up to 28 10-Gigabit ports can consume their full 100 ms allocation of buffer space. However, if a 29th 10-Gigabit port requires the full 125 MB of buffer space, then the buffer allocations are proportionally reduced because the buffer consumption is above 90 percent.

QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems

Set the dedicated buffer size of the egress queue that you bind the scheduler to in the scheduler map configuration. The switch allocates space from the global dedicated buffer pool to ports and queues in a hierarchical manner. The switch allocates an equal number of dedicated buffers to each egress port, so each egress port receives the same amount of dedicated buffer space. The amount of dedicated buffer space per port is not configurable.

However, the **buffer-size** statement allows you to control the way each port allocates its share of dedicated buffers to its queues. For example, if a port only uses two queues to forward traffic, you can configure the port to allocate all of its dedicated buffer space to those two ports and avoid wasting buffer space on queues that are not in use. We recommend that the buffer size should be the same size as the minimum guaranteed transmission rate (the **transmit-rate**).

Default The default behavior of the differs on different switches.

QFX10000 Switches

If you do not configure buffer-size and you do not explicitly configure a queue scheduler, the default buffer-size is the default transmit rate of the queue. If you explicitly configure a queue scheduler, the default buffer allocations are not used. If you explicitly configure a queue scheduler, configure the buffer-size for each queue in the scheduler, keeping in mind that the total buffer-size of the queues cannot exceed 100 percent (100 ms).

Table 78 on page 403 shows the default queue buffer sizes on QFX10000 switches. The default buffer size is the same as the default transmit rate for each default queue:

Table 78: Default Output Queue Buffer Sizes (QFX10000 Switches)

Queue Number	Forwarding Class	Transmit Rate	Buffer Size
0	best-effort	15%	15%
3	fcoe	35%	35%
4	no-loss	35%	35%
7	network-control	15%	15%

By default, only the queues mapped to the default forwarding classes receive buffer space from the port buffer pool. (Buffers are not wasted on queues that do not carry traffic.)

QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems

The port allocates dedicated buffers to queues that have an explicitly configured scheduler buffer size. If you do not explicitly configure a scheduler buffer size for a queue, the port serves the explicitly configured queues first. Then the port divides the remaining dedicated buffers equally among the queues that have an explicitly attached scheduler *without* an explicitly configured buffer size configuration. (If you configure a scheduler, but you do not configure the buffer size parameter, the default is equivalent to configuring the buffer size with the **remainder** option.)

If you use the default scheduler and scheduler map on a port (no explicit scheduler configuration), then the port allocates its dedicated buffer pool to queues based on the default scheduling. Table 79 on page 404 shows the default queue buffer sizes. The default buffer size is the same as the default transmit rate for each default queue:

Table 79: Default Output Queue Buffer Sizes (QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems)

Queue Number	Forwarding Class	Transmit Rate	Buffer Size
0	best-effort	5%	5%
3	fcoe	35%	35%
4	no-loss	35%	35%
7	network-control	5%	5%
8	mcast	20%	20%

By default, only the queues mapped to the default forwarding classes receive buffer space from the port buffer pool. (Buffers are not wasted on queues that do not carry traffic.)

NOTE: OCX Series switches do not support lossless transport. On OCX Series switches, do not map traffic to the lossless default fcoe and no-loss forwarding classes. OCX Series default DSCP classification does not map traffic to the fcoe and no-loss forwarding classes, so by default, the OCX system does not classify traffic into those forwarding classes. (On other switches, the fcoe and no-loss forwarding classes provide lossless transport for Layer 2 traffic. OCX Series switches do not support lossless Layer 2 transport.) The active forwarding classes (**best-effort**, **network-control**, and **mcast**) share the unused bandwidth assigned to the **fcoe** and **no-loss** forwarding classes.

Options

percent *percent*—Percentage of the port dedicated buffer pool allocated to the queue (or queues) mapped to the scheduler.

remainder—Remaining dedicated buffer pool after the port satisfies the needs of the explicitly configured buffers. The port divides the remaining buffers equally among the queues that are explicitly attached to a scheduler but that do not have an explicit buffer size configuration (or are configured with **remainder** as the buffer size).

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

drop-probability

List of Syntax

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems on page 405
QFX10000 Switches on page 405

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems

drop-probability 0 drop-probability *high-value*;

QFX10000 Switches

drop-probability *percentage1 percentage2 ... percentage32*;

Hierarchy Level

[edit class-of-service drop-profiles *profile-name* interpolate]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

When configuring WRED, map the packet **drop-probability** to the fullness of a queue (fill-level). You configure the **fill-level** and **drop-probability** statements in related pairs. The pairs of fill level and drop probability values set a probability of dropping packets at a specified queue fullness value.

On switches that support only two fill level/drop probability pairs, the first drop probability is always zero. The first fill level/drop probability pair sets the drop start point, and the second fill level/drop probability pair sets the drop end point.

On switches that support 32 fill level/drop probability pairs, the first fill level/drop probability pair sets the drop start point, and the last fill level/drop probability pair sets the drop end point.

As the queue fills from the drop start point to the drop end point, the rate of packet drop increases in a curve pattern. The higher the queue fill level, the higher the probability of dropping packets.

Options

0 (switches that support only two fill level/drop probability pairs)—Probability that packets will drop at the lowest **fill-level** value. This is always zero, because until the queue reaches the specified low **fill-level** value, no packets are scheduled to drop.

Range: 0

high-value (switches that support only two fill level/drop probability pairs)—The maximum probability that packets will drop before queue fullness exceeds the high value of the queue **fill-level**, expressed as a percentage. If the queue fills beyond the high **fill-level** value, all packets drop.

Range: 0 through 100 percent

percentage1 percentage2 ... percentage32 (switches that support 32 fill level/drop probability pairs)—The probability that packets will drop before the queue fullness exceeds the **fill-level** value, expressed as a percentage. Each drop probability pairs with a queue fill level to define the probability of a packet dropping at a specified queue fullness.

Range: 0 through 100 percent

Required Privilege Level
interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

drop-profile

Syntax drop-profile *profile-name*;

Hierarchy Level [edit class-of-service schedulers *scheduler-name* drop-profile-map loss-priority (low | medium-high | high) protocol *protocol*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Define drop profiles for weighted random early detection (WRED). When a packet arrives, WRED checks the queue fill level specified in the drop profile. If the fill level corresponds to a nonzero drop probability, the WRED algorithm determines whether to drop the arriving packet.

Options *profile-name*—Name of the drop profile.

Required Privilege Level
interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring Drop Profile Maps on page 377

drop-profile-map

Syntax drop-profile-map loss-priority (low | medium-high | high) protocol *protocol* drop-profile *drop-profile-name*;

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Map a drop profile to a loss priority and protocol for weighted random early detection (WRED). When a packet arrives, WRED checks the queue fill level. If the fill level corresponds to a nonzero drop probability, the WRED algorithm determines whether to drop the arriving packet.

Options The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring Drop Profile Maps on page 377

drop-profiles

List of Syntax

Switches Except QFX10000, QFabric Systems on page 408
QFX10000 Switches on page 408

Switches Except QFX10000, QFabric Systems

```
drop-profiles {
  profile-name {
    interpolate {
      fill-level low-value fill-level high-value drop-probability 0 drop-probability high-value;
    }
  }
}
```

QFX10000 Switches

```
drop-profiles {
  profile-name {
    interpolate {
      fill-level level1 level2 ... level32 drop-probability percent1 percent2 ... percent32;
    }
  }
}
```

Hierarchy Level

[edit class-of-service]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Define drop profiles for weighted random early detection (WRED).

For a packet to be dropped, it must match the drop profile. When a packet arrives, WRED checks the queue fill level. If the fill level corresponds to a nonzero drop probability, the WRED algorithm determines whether to drop the arriving packet.

Options

profile-name—Name of the drop profile.

The remaining statements are explained separately.

Required Privilege Level

interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

explicit-congestion-notification

Syntax explicit-congestion-notification;

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced in Junos OS Release 13.2X51 for EX Series switches.
Statement introduced in Junos OS Release 13.2X51-D20 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Enable explicit congestion notification (ECN) on the output queue (forwarding class) or output queues (forwarding classes) mapped to the scheduler. ECN enables end-to-end congestion notification between two endpoints on TCP/IP based networks. The two endpoints are an ECN-enabled sender and an ECN-enabled receiver. ECN must be enabled on both endpoints and on all of the intermediate devices between the endpoints for ECN to work properly. Any device in the transmission path that does not support ECN breaks the end-to-end ECN functionality.

A weighted random early detection (WRED) packet drop profile must be applied to the output queues on which ECN is enabled. ECN uses the WRED drop profile thresholds to mark packets when the output queue experiences congestion.

ECN reduces packet loss by forwarding ECN-capable packets during periods of network congestion instead of dropping those packets. (TCP notifies the network about congestion by dropping packets.) During periods of congestion, ECN marks ECN-capable packets that egress from congested queues. When the receiver receives an ECN packet that is marked as experiencing congestion, the receiver echoes the congestion state back to the sender. The sender then reduces its transmission rate to clear the congestion.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring ECN on page 389
- Understanding CoS Explicit Congestion Notification on page 380

fill-level

List of Syntax

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems on page 410
QFX10000 Switches on page 410

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems

fill-level *low-value* fill-level *high-value*;

QFX10000 Switches

fill-level *level1 level2* ... *level32*;

Hierarchy Level

[edit class-of-service drop-profiles *profile-name* interpolate]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

When configuring weighted random early detection (WRED), map the fullness of a queue to a packet drop-probability value. You configure the **fill-level** and **drop-probability** statements in related pairs. The pairs of fill level and drop probability values set a probability of dropping packets at a specified queue fullness value.

The first fill level is the packet drop start point. Packets do not drop until the queue fullness reaches the first fill level. The last fill level is the packet drop end point. After the queue exceeds the fullness set by the drop end point, all non-ECN packets are dropped. As the queue fills from the drop start point to the drop end point, the rate of packet drop increases in a curve pattern. The higher the queue fill level, the higher the probability of dropping packets.

On switches that support only two fill level/drop probability pairs, the two pairs are the drop start point and the drop end point. On switches that support up to 32 fill level/drop probability pairs, you can configure intermediate interpolations between the drop start point and the drop end point, which provides greater flexibility in controlling the packet drop curve.

NOTE: **Do not configure the last fill level as 100 percent.**

Options

low-value (switches that support only two fill level/drop probability pairs)—Fullness of the queue before packets begin to drop, expressed as a percentage. The low value must be less than the high value.

Range: 0 through 100

high-value (switches that support only two fill level/drop probability pairs)—Fullness of the queue before it reaches the maximum drop probability. If the queue fills beyond the fill level high value, all packets drop. The high value must be greater than the low value.

Range: 0 through 100

level1 level2 ... level32 (switches that support 32 fill level/drop probability pairs)—The queue fullness level, expressed as a percentage. Each fill level pairs with a drop probability to define the probability of a packet dropping at a specified queue fullness.

Range: 0 through 100

Required Privilege Level

interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

forwarding-class

List of Syntax

Classifier

```
forwarding-class class-name {
  loss-priority level {
    code-points [ aliases ] [ bit-patterns ];
  }
}
```

Rewrite Rule

```
forwarding-class class-name {
  loss-priority level {
    code-point [ aliases ] [ bit-patterns ];
  }
}
```

Scheduler Map

```
forwarding-class class-name {
  scheduler scheduler-name;
}
```

Interface

```
forwarding-class class-name;
```

Classifier Hierarchy Level

[edit class-of-service classifiers (dscp | dscp-ipv6 | ieee-802.1 | exp) *classifier-name*],

Rewrite Rule Hierarchy Level

[edit class-of-service rewrite-rules] (dscp | dscp-ipv6 | ieee-802.1) *rewrite-name* | exp],

Scheduler Map Hierarchy Level

[edit class-of-service scheduler-maps *map-name*],

Interface Hierarchy Level

[edit class-of-service interfaces *interface-name* unit *logical-unit-number*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

- Classifiers—Assign incoming traffic to the specified forwarding class based on the specified code point values and assign that traffic the specified loss priority
- Rewrite rules—At the egress interface, change (rewrite) the value of the code point bits and the loss priority to specified new values for traffic assigned to the specified forwarding class, before forwarding the traffic to the next hop.
- Scheduler maps—Apply the specified scheduler to the specified forwarding class.

- Interfaces—Assign the specified forwarding class to the interface to use as a fixed classifier (all incoming traffic on the interface is classified into that forwarding class).

NOTE: OCX Series switches do not support MPLS, so they do not support EXP classifiers or rewrite rules.

Options *class-name*—Name of the forwarding class.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

forwarding-class-set

Syntax

```
forwarding-class-set forwarding-class-set-name {
  output-traffic-control-profile profile-name;
}
```

Hierarchy Level [edit class-of-service interfaces *interface-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Apply a previously defined forwarding class set to an output traffic control profile.

Options ***fowarding-class-set-name***—Name of the forwarding class set.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Assigning CoS Components to Interfaces on page 21
- Understanding CoS Forwarding Class Sets (Priority Groups) on page 148

guaranteed-rate

Syntax guaranteed-rate (*rate*| percent *percentage*);

Hierarchy Level [edit class-of-service traffic-control-profiles *traffic-control-profile-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure a guaranteed minimum rate of transmission for a traffic control profile. The sum of the guaranteed rates of all of the forwarding class sets (priority groups) on a port should not exceed the total port bandwidth. The guaranteed rate also determines the amount of excess (extra) port bandwidth that the priority group (forwarding class set) can share. Extra port bandwidth is allocated among the priority groups on a port in proportion to the guaranteed rate of each priority group.

NOTE: **You cannot configure a guaranteed rate for a forwarding class set (priority group) that includes strict-high priority queues. If the traffic control profile is for a forwarding class set that contains strict-high priority queues, do not configure a guaranteed rate.**

Default If you do not specify a guaranteed rate, the guaranteed rate is zero (0) and there is no minimum guaranteed bandwidth.

NOTE: **If you do not configure a guaranteed rate for a traffic control profile, the queues that belong to any forwarding class set (priority group) that uses that traffic control profile cannot have a configured transmit rate. The result is that there is no minimum guaranteed bandwidth for those queues and that those queues can be starved during periods of congestion.**

Options percent *percentage*—Minimum percentage of transmission capacity allocated to the forwarding class set or logical interface.

Range: 1 through 100 percent

rate—Minimum transmission rate allocated to the forwarding class set or logical interface, in bits per second (bps). You can specify a value in bits per second either as a complete decimal number or as a decimal number followed by the abbreviation **k** (1000), **m** (1,000,000), or **g** (1,000,000,000).

Range: 1000 through 10,000,000,000 bps

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Minimum Guaranteed Output Bandwidth on page 350
- Understanding CoS Traffic Control Profiles on page 306
- **output-traffic-control-profile on page 417**

interpolate

List of Syntax QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems on page 415
QFX10000 Switches on page 415

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems

```
interpolate {
  fill-level low-value fill-level high-value;
  drop-probability 0 drop-probability high-value;
}
```

QFX10000 Switches

```
interpolate {
  fill-level level1 level2 ... level32 drop-probability percent1 percent2 ... percent32;
}
```

Hierarchy Level [edit class-of-service drop-profiles *profile-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Specify values for interpolating the relationship between queue fill level and drop probability for weighted random early detection (WRED) drop profiles.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

loss-priority (Drop Profiles)

Syntax loss-priority *level* protocol *protocol* drop-profile *profile-name*;

Hierarchy Level [edit class-of-service schedulers *scheduler-name* drop-profile-map]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure packet loss priority value for a weighted random early detection (WRED) drop profile mapped to a system drop profile.

Options *level*—Can be one of the following:

- **low**—Packet has low loss priority.
- **medium-high**—Packet has medium-high loss priority.
- **high**—Packet has high loss priority.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

output-traffic-control-profile

Syntax output-traffic-control-profile *profile-name*;

Hierarchy Level [edit class-of-service interfaces *interface-name* forwarding-class-set *forwarding-class-set-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Apply an output traffic scheduling and shaping profile to a forwarding class set (priority group).

Options ***profile-name***—Name of the traffic-control profile to apply to the specified forwarding class set.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Assigning CoS Components to Interfaces on page 21
- Understanding CoS Traffic Control Profiles on page 306

priority (Schedulers)

Syntax priority *priority*;

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Specify the packet bandwidth-scheduling priority value.

NOTE: On QFabric systems, the **priority** statement is valid only for Node device queue scheduling. The **priority** statement is notallowed for Interconnect device queue scheduling. If you map a schedulerthat includes a **priority** configuration to a fabric forwardingclass at the [edit class-of-service scheduler-map-fcset] hierarchylevel, the system generates a commit error. (On the Interconnect device,fabric fc-sets are not user-definable. Only the **fabric_fcset_strict_high** fabric fc-set is configured with high priority, and this configurationcannot be changed.)

Options *priority*—It can be one of the following:

- **low**—Scheduler has low priority.
- **high**—Scheduler has high priority (QFX10000 Series switches only)
- **strict-high**—Scheduler has strict high priority. On QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, you can configure only one queue as a strict-high priority queue. On QFX10000 switches, you can configure as many strict-high priority queues as you want. However, because strict-high priority traffic takes precedence over all other traffic, too much strict-high priority traffic can starve the other output queues.

  Strict-high priority allocates the scheduled bandwidth to the packets on the queue before any other queue receives bandwidth. Other queues receive the bandwidth that remains after the strict-high queue has been serviced.

  NOTE: On QFX10000 switches, we strongly recommend that you apply a transmit rate to strict-high priority queues to prevent them from starving other queues. A transmit rate configured on a strict-high priority queue limits the amount of traffic that receives strict-high priority treatment to the amount or percentage set by the transmit rate. The switch treats traffic in excess of the transmit rate as best-effort traffic that receives bandwidth from the leftover (excess) port bandwidth pool. On strict-high priority queues, all traffic that exceeds the transmit rate shares in the port excess bandwidth pool based on the strict-high priority excess bandwidth sharing weight of "1", which is not configurable. The actual amount of extra bandwidth that traffic exceeding the transmit rate receives depends on how many other queues consume excess bandwidth and the excess rates of those queues.

On QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, we recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. A shaping rate (shaper) sets the maximum amount of bandwidth a queue can consume. (Unlike using the transmit rate on a QFX10000 switch to limit traffic that receives strict-high priority treatment, traffic that exceeds the shaping rate is dropped, and is not treated as best-effort traffic that shares in excess bandwidth.) If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

Required Privilege Level
interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

protocol (Drop Profile Map)

Syntax
protocol *protocol* drop-profile *profile-name*;

Hierarchy Level
[edit class-of-service schedulers *scheduler-name* drop-profile-map loss-priority (low | medium-high | high)]

Release Information
Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description
Configure the protocol type for the specified weighted random early detection (WRED) drop profile.

Options
protocol—Type of protocol. The protocol can be:

- **any**—Accept any protocol type.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level
interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

scheduler

Syntax scheduler *scheduler-name*;

Hierarchy Level [edit class-of-service scheduler-maps *map-name* forwarding-class *class-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Map a scheduler to a forwarding class using a scheduler map.

NOTE: On QFX5200 only, absolute CoS rate limits for transmit rate and shaping rate do not reflect 50g and 100g interfaces. Therefore this statement does not affect those interfaces for QFX5200 in release 15.1X53-D30.

Options ***scheduler-name***—Name of the scheduler to map to the forwarding class.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

scheduler-map

Syntax scheduler-map *map-name*;

Enhanced Transmission Selection (ETS) Hierarchical Scheduling [edit class-of-service traffic-control-profiles *traffic-control-profile-name*]

Port Scheduling [edit class-of-service interfaces *interface-name*]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Associate a scheduler map with a traffic control profile.

Options ***map-name***—Name of the scheduler map.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

scheduler-maps

Syntax

```
scheduler-maps {
  map-name {
    forwarding-class class-name scheduler scheduler-name;
  }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Specify a scheduler map name to map a scheduler configuration to a forwarding class.

Options ***map-name***—Name of the scheduler map.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

schedulers

List of Syntax

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems on page 422
QFX10000 Switches on page 422

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems

```
schedulers {
  scheduler-name {
    buffer-size (percent percentage | remainder);
    drop-profile-map loss-priority (low | medium-high | high) protocol protocol drop-profile
      drop-profile-name;
    explicit-congestion-notification;
    priority priority;
    shaping-rate (rate | percent percentage);
    transmit-rate (percent percentage);
  }
}
```

QFX10000 Switches

```
schedulers {
  scheduler-name {
    buffer-size (percent percentage | remainder);
    drop-profile-map loss-priority (low | medium-high | high) protocol protocol drop-profile
      drop-profile-name;
    excess-rate;
    explicit-congestion-notification;
    priority priority;
    shaping-rate (rate | percent percentage);
    transmit-rate (percent percentage) <exact>;
  }
}
```

Hierarchy Level

[edit class-of-service]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Specify scheduler name and parameter values such as minimum bandwidth (**transmit-rate**), maximum bandwidth (**shaping-rate**), and priority (**priority**).

Options

scheduler-name —Name of the scheduler.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

shaping-rate

Syntax shaping-rate (*rate* | percent *percentage*);

Hierarchy Level [edit class-of-service schedulers *scheduler-name*],
[edit class-of-service traffic-control-profiles *profile-name*]

NOTE: Only switches that support enhanced transmission selection (ETS) hierarchical scheduling support the **traffic-control-profiles** hierarchy.

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure the shaping rate. The shaping rate throttles the rate of packet transmission by setting a maximum bandwidth (rate in bits per second) or a maximum percentage of bandwidth for a queue or a forwarding class set. You specify the maximum bandwidth for a queue by using a scheduler map to associate a forwarding class (queue) with a scheduler that has a configured shaping rate.

For ETS configuration, you specify the maximum bandwidth for a forwarding class set by setting the shaping rate for a traffic control profile, then you associate the scheduler map with the traffic control profile, and then you apply the traffic control profile and a forwarding class set to an interface.

For simple port scheduling configuration, you apply the scheduler map directly to an interface (instead of indirectly through the traffic control profile as in ETS).

We recommend that you configure the shaping rate as an absolute maximum usage and not as additional usage beyond the configured transmit rate (the minimum guaranteed bandwidth for a queue) or the configured guaranteed rate (the minimum guaranteed bandwidth for a forwarding class set).

NOTE: When you set the maximum bandwidth (**shaping-rate** value) for a queue or for a priority group at 100 Kbps or less, the traffic shaping behavior is accurate only within +/– 20 percent of the configured **shaping-rate** value.

NOTE: On QFX5200, QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, we recommend that you always apply a shaping rate to strict-high priority queues to prevent them from starving other queues. If you do not apply a shaping rate to limit the amount of bandwidth a strict-high priority queue can use, then the strict-high priority queue can use all of the available port bandwidth and starve other queues on the port.

NOTE: On QFX5200 Series switches, a granularity of 64kbps is supported for the shaping rate. Therefore, the shaping rate on queues for 100g interfaces might not be applied correctly.

NOTE: QFX10000 Series switches do not support the shaping-rate statement. However, you can configure the transmit-rate exact option to prevent a queue from consuming more bandwidth than you want the queue to consume.

On QFX10000 Series switches, we recommend that you use the transmit rate to set a limit on the amount of bandwidth that receives strict-high priority treatment on a strict-high priority queue. Traffic up to the transmit rate receives strict-high priority treatment. Traffic in excess of the transmit rate is treated as best-effort traffic that receives the strict-high priority queue excess rate weight of “1”. Do not use a shaping rate to set a maximum bandwidth limit on strict-high priority queues on QFX10000 Series switches.

Default If you do not configure a shaping rate, the default shaping rate is 100 percent (all of the available bandwidth), which is the equivalent of no rate shaping.

Options percent *percentage*—Shaping rate as a percentage of the available interface bandwidth.

Range: 1 through 100 percent

rate—Peak (maximum) rate, in bits per second (bps). You can specify a value in bits per second either as a complete decimal number or as a decimal number followed by the abbreviation k (1000), m (1,000,000), or g (1,000,000,000).

Range: 1000 through 10,000,000,000 bps

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- *Example: Configuring Queue Schedulers for Port Scheduling*
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Understanding CoS Output Queue Schedulers on page 283
- *Understanding CoS Port Schedulers on QFX Switches*
- Understanding CoS Traffic Control Profiles on page 306

traffic-control-profiles

Syntax

```
traffic-control-profiles profile-name {
  guaranteed-rate (rate| percent percentage);
  scheduler-map map-name;
  shaping-rate (rate| percent percentage);
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure traffic shaping and scheduling profiles for forwarding class sets (priority groups) to implement enhanced transmission selection (ETS) or for logical interfaces.

Options ***profile-name***—Name of the traffic-control profile. This name is also used to specify an output traffic control profile.

The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Traffic Control Profiles (Priority Group Scheduling) on page 311
- Example: Configuring Forwarding Class Sets on page 150
- Assigning CoS Components to Interfaces on page 21
- **output-traffic-control-profile on page 417**
- Understanding CoS Traffic Control Profiles on page 306

transmit-rate

List of Syntax

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems on page 426
QFX10000 Switches on page 426

QFX5100, EX4600, QFX3500, and QFX3600, Switches, QFabric Systems

transmit-rate (*rate* | percent *percentage*);

QFX10000 Switches

transmit-rate (*rate* | percent *percentage*) <exact>;

Hierarchy Level

[edit class-of-service schedulers *scheduler-name*]

Release Information

Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
Exact option introduced in Junos OS Release 15.1X53-D10 for the QFX Series.

Description

On QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, the transmit rate specifies the minimum guaranteed transmission rate or percentage for a queue (forwarding class) scheduler. The queue transmit rate also determines the amount of excess (extra) priority group bandwidth that the queue can share on switches that support enhanced transmission selection (ETS) hierarchical scheduling.

On QFX10000 switches, the transmit rate specifies the minimum guaranteed transmission rate or percentage for a queue (forwarding class) scheduler. The queue transmit rate also determines the amount of excess (extra) port bandwidth the queue can share if you do not explicitly configure an excess rate in the scheduler. The transmit rate also determines the amount of excess (extra) priority group bandwidth that the queue can share on switches that support enhanced transmission selection (ETS) hierarchical scheduling.

On QFX10000 switch strict-high priority queues, the transmit rate limits the amount of traffic the switch treats as strict-high priority traffic. Traffic up to the transmit rate receives strict-high priority treatment. The switch treats traffic that exceeds the transmit rate as best-effort traffic that receives an excess bandwidth sharing weight of “1”; you cannot configure an excess rate on a strict-high priority queue, and unlike queues with other scheduling priorities, the switch does not use the transmit rate to determine extra bandwidth sharing for strict-high priority queues.

CAUTION: We strongly recommend that you configure a transmit rate on strict-high priority queues to limit the amount of traffic the switch treats as strict-high priority traffic on those queues. This is especially important if you configure more than one strict-high priority queue on a port. To prevent a strict-high priority queue from starving the other queues on a port, we recommend that you always configure a transmit rate, even if you only configure one strict-high priority queue.

NOTE: For ETS, the **transmit-rate** setting works only if you also configure the guaranteed-rate in the traffic control profile that is attached to the forwarding class set to which the queue belongs. If you do not configure the guaranteed rate, the minimum guaranteed rate for individual queues that you set using the **transmit-rate** statement does not work. The sum of all queue transmit rates in a forwarding class set should not exceed the traffic control profile guaranteed rate.

NOTE: On QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, you cannot configure a transmit rate for a strict-high priority queue. Queues (forwarding classes) with a configured transmit rate cannot be included in a forwarding class set that has a strict-high priority queue. To prevent strict-high priority queues from consuming all of the available bandwidth on these switches, we recommend that you configure a shaping rate to set a maximum amount of bandwidth for strict-high priority queues.

NOTE: For transmit rates below 1 Gbps, we recommend that you configure the transmit rate as a percentage instead of as a fixed rate. This is because the system converts fixed rates into percentages and may round small fixed rates to a lower percentage. For example, a fixed rate of 350 Mbps is rounded down to 3 percent instead of 3.5 percent.

Default On QFX5100, EX4600, QFX3500, and QFX3600 switches, and on QFabric systems, if you do not configure the transmit rate, the default scheduler transmission rate and buffer size percentages for queues 0 through 11 are:

Table 80: Default Transmit Rates for QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems

Queue Number	Default Minimum Guaranteed Bandwidth (Transmit Rate)
0 (best-effort)	5 %
1	0
2	0
3 (fcoe)	35 %
4 (no-loss)	35 %
5	0
6	0
7 (network control)	5 %
8 (mcast)	20 %
9	0
10	0
11	0

NOTE: OCX Series switches do not support lossless transport. The OCX Series default DSCP classifier does not classify traffic into the default lossless fcoe and no-loss forwarding classes. The bandwidth that the default scheduler allocates to the default fcoe and no-loss forwarding classes on other switches is allocated to the default best-effort, network-control, and mcast forwarding classes on OCX Series switches.

On QFX10000 switches, if you do not configure the transmit rate, the default scheduler transmission rate and buffer size percentages for queues 0 through 7 are:

Table 81: Default Transmit Rates for QFX10000 Switches

Queue Number	Default Minimum Guaranteed Bandwidth (Transmit Rate)
0 (best-effort)	15 %
1	0
2	0
3 (fcoe)	35 %
4 (no-loss)	35 %
5	0
6	0
7 (network control)	15 %

Configure schedulers if you want to change the minimum guaranteed bandwidth and other queue characteristics.

Options

rate—Minimum transmission rate for the queue, in bps. You can specify a value in bits-per-second either as a complete decimal number or as a decimal number followed by the abbreviation **k** (1000), m (1,000,000), or **g** (1,000,000,000).

Range: 1000 through 10,000,000,000 bps on 10-Gigabit interfaces, 1000 through 40,000,000,000 bps on 40-Gigabit interfaces.

percent ***percentage***—Minimum percentage of transmission capacity allocated to the queue.

Range: 1 through 100 percent

exact—(QFX10000 switches only) Shape queues that are not strict-high priority queues to the transmit rate so that the transmit rate is the maximum bandwidth limit. Traffic that exceeds the exact transmit rate is dropped. You cannot set an excess rate on queues configured as **transmit-rate** ***(rate*** **|** ***percentage)*** **exact** because the purpose of setting an exact transmit rate is to set a maximum bandwidth (shaping rate) on the traffic.

NOTE: On QFX10000 switches, oversubscribing all 8 queues configured with the **transmit rate exact** (shaping) statement at the **[edit class-of-service schedulers** ***scheduler-name*****]** hierarchy level might result in less than 100 percent utilization of port bandwidth.

Required Privilege Level
interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Configuring CoS Hierarchical Port Scheduling (ETS) on page 321
- Example: Configuring Queue Schedulers on page 293
- *Example: Configuring Queue Schedulers for Port Scheduling*
- Understanding CoS Output Queue Schedulers on page 283
- *Understanding CoS Port Schedulers on QFX Switches*

CHAPTER 9

Monitoring Commands for Scheduling

Monitoring CoS Scheduler Maps

Purpose Use the monitoring functionality to display assignments of CoS forwarding classes to schedulers.

Action To monitor CoS scheduler maps in the CLI, enter the CLI command:

user@switch> **show class-of-service scheduler-map**

To monitor a specific scheduler map in the CLI, enter the CLI command:

user@switch> **show class-of-service scheduler-map *scheduler-map-name***

Meaning Table 82 on page 431 summarizes key output fields for CoS scheduler maps.

Table 82: Summary of Key CoS Scheduler Maps Output Fields

Field	Values
Scheduler map	Name of a scheduler map that maps forwarding classes to schedulers.
Index	Index of a specific object—scheduler maps, schedulers, or drop profiles.

Table 82: Summary of Key CoS Scheduler Maps Output Fields (continued)

Field	Values
Scheduler	Name of a scheduler that controls queue properties such as bandwidth and scheduling priority.
Forwarding class	Name(s) of the forwarding class(es) to which the scheduler is mapped.
Transmit rate	Guaranteed minimum bandwidth configured on the queue mapped to the scheduler. On strict-high priority queues on QFX10000 switches, defines the maximum amount of traffic on the queue that is treated as strict-high priority traffic.
Priority	Scheduling priority of traffic on a queue: • **strict-high** or **high**—Packets on a strict-high priority queue are transmitted first, before all other traffic, up to the configured maximum bandwidth (shaping rate). On QFX3500, QFX3600, EX4600, and OCX series switches, and on QFabric system, only one queue can be configured as **strict-high** or **high** priority. On QFX10000 switches, you can configure more than one strict-high priority queue. • **low**—Packets in this queue are transmitted after packets in the **strict-high** queue.
Drop Profiles	Name and index of a drop profile that is mapped to a specific loss priority and protocol pair. The drop profile determines the way best effort queues drop packets during periods of congestion.
Loss Priority	Packet loss priority mapped to the drop profile. You can configure different drop profiles for **low**, **medium-high**, and **high** loss priority traffic.
Protocol	Transport protocol of the drop profile for the particular priority.
Name	Name of the drop profile.

show class-of-service drop-profile

Syntax

show class-of-service drop-profile
<profile-name *profile-name*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display data points for each class-of-service (CoS) random early detection (RED) drop profile.

Options

none—Display all drop profiles.

profile-name ***profile-name***—(Optional) Display the specified profile only.

Required Privilege Level

view

List of Sample Output

show class-of-service drop-profile on page 434
show class-of-service drop-profile (EX4200 Switch) on page 434
show class-of-service drop-profile (EX8200 Switch) on page 434

Output Fields

Table 83 on page 433 describes the output fields for the **show class-of-service drop-profile** command. Output fields are listed in the approximate order in which they appear.

Table 83: show class-of-service drop-profile Output Fields

Field Name	Field Description
Drop profile	Name of a drop profile.
Type	Type of drop profile: • **discrete** (default) • **interpolated** (EX8200 switches, QFX Series switches, QFabric systems, EX4600 switches, OCX Series switches only)
Index	Internal index of this drop profile.
Fill Level	Percentage fullness of a queue.
Drop probability	Drop probability at this fill level.

Sample Output

show class-of-service drop-profile

```
user@host> show class-of-service drop-profile
Drop profile: <default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: user-drop-profile, Type: interpolated, Index: 2989
  Fill level    Drop probability
           0                   0
           1                   1
           2                   2
           4                   4
           5                   5
           6                   6
           8                   8
          10                  10
          12                  15
          14                  20
          15                  23
        ... 64 entries total
          90                  96
          92                  96
          94                  97
          95                  98
          96                  98
          98                  99
          99                  99
         100                 100
```

show class-of-service drop-profile (EX4200 Switch)

```
user@switch> show class-of-service drop-profile
Drop profile: <default-drop-profile>, Type: discrete, Index: 1
  Fill level
         100
Drop profile: dp1, Type: discrete, Index: 40496
  Fill level
          10
```

show class-of-service drop-profile (EX8200 Switch)

```
user@switch> show class-of-service drop-profile
Drop profile: <default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: dp1, Type: interpolated, Index: 40496
  Fill level    Drop probability
           0                   0
           1                  80
           2                  90
           4                  90
           5                  90
           6                  90
           8                  90
          10                  90
          12                  91
          14                  91
```



```
              15                 91
              16                 91
              18                 91
              20                 91
              22                 92
              24                 92
              25                 92
              26                 92
              28                 92
              30                 92
              32                 93
              34                 93
              35                 93
              36                 93
              38                 93
              40                 93
              42                 94
              44                 94
              45                 94
              46                 94
              48                 94
              49                 94
              51                 95
              52                 95
              54                 95
              55                 95
              56                 95
              58                 95
              60                 95
              62                 96
              64                 96
              65                 96
              66                 96
              68                 96
              70                 96
              72                 97
              74                 97
              75                 97
              76                 97
              78                 97
              80                 97
              82                 98
              84                 98
              85                 98
              86                 98
              88                 98
              90                 98
              92                 99
              94                 99
              95                 99
              96                 99
              98                 99
              99                 99
             100                100
Drop profile: dp2, Type: discrete, Index: 40499
  Fill level    Drop probability
              10                  5
              50                 50
```

show class-of-service forwarding-table

List of Syntax

Syntax on page 436
Syntax (TX Matrix and TX Matrix Plus Router) on page 436

Syntax

show class-of-service forwarding-table

Syntax (TX Matrix and TX Matrix Plus Router)

show class-of-service forwarding-table
<lcc *number*> | <sfc *number*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display the entire class-of-service (CoS) configuration as it exists in the forwarding table. Executing this command is equivalent to executing all **show class-of-service forwarding-table** commands in succession.

Options

lcc ***number***—(TX Matrix and TX Matrix Plus router only) (Optional) On a TX Matrix router, display the forwarding table configuration for a specific T640 router (or line-card chassis) configured in a routing matrix. On a TX Matrix Plus router, display the forwarding table configuration for a specific router (or line-card chassis) configured in the routing matrix.

Replace *number* with the following values depending on the LCC configuration:

- 0 through 3, when T640 routers are connected to a TX Matrix router in a routing matrix.
- 0 through 3, when T1600 routers are connected to a TX Matrix Plus router in a routing matrix.
- 0 through 7, when T1600 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.
- 0, 2, 4, or 6, when T4000 routers are connected to a TX Matrix Plus router with 3D SIBs in a routing matrix.

sfc ***number***—(TX Matrix Plus routers only) (Optional) Display the forwarding table configuration for the TX Matrix Plus router. Replace ***number*** with **0**.

Required Privilege Level

view

List of Sample Output

show class-of-service forwarding-table on page 437
show class-of-service forwarding-table lcc (TX Matrix Plus Router) on page 438

Output Fields

See the output field descriptions for **show class-of-service forwarding-table** commands:

- show class-of-service forwarding-table classifier
- show class-of-service forwarding-table classifier mapping
- show class-of-service forwarding-table drop-profile
- *show class-of-service forwarding-table fabric scheduler-map*
- show class-of-service forwarding-table rewrite-rule
- show class-of-service forwarding-table rewrite-rule mapping
- show class-of-service forwarding-table scheduler-map

Sample Output

show class-of-service forwarding-table

```
user@host> show class-of-service forwarding-table
Classifier table index: 9, # entries: 8, Table type: EXP
Entry #    Code point    Forwarding-class #    PLP
   0           000         0                    0
   1           001         0                    1
   2           010         1                    0
   3           011         1                    1
   4           100         2                    0
   5           101         2                    1
   6           110         3                    0
   7           111         3                    1

                       Table Index/
Interface      Index      Q num      Table type
sp-0/0/0.1001     66          11      IPv4 precedence
sp-0/0/0.2001     67          11      IPv4 precedence
sp-0/0/0.16383    68          11      IPv4 precedence
fe-0/0/0.0        69          11      IPv4 precedence

Interface: sp-0/0/0 (Index: 129, Map index: 2, Map type: FINAL,
 Num of queues: 2):
  Entry 0 (Scheduler index: 16, Forwarding-class #: 0):
    Tx rate: 0 Kb (95%), Buffer size: 95 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
  Entry 1 (Scheduler index: 18, Forwarding-class #: 3):
    Tx rate: 0 Kb (5%), Buffer size: 5 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1

Interface: fe-0/0/0 (Index: 137, Map index: 2, Map type: FINAL,
 Num of queues: 2):
  Entry 0 (Scheduler index: 16, Forwarding-class #: 0):
    Tx rate: 0 Kb (95%), Buffer size: 95 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
  Entry 1 (Scheduler index: 18, Forwarding-class #: 3):
    Tx rate: 0 Kb (5%), Buffer size: 5 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
Interface: fe-0/0/1 (Index: 138, Map index: 2, Map type: FINAL,
 Num of queues: 2):
  Entry 0 (Scheduler index: 16, Forwarding-class #: 0):
```



```
    Tx rate: 0 Kb (95%), Buffer size: 95 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1
  Entry 1 (Scheduler index: 18, Forwarding-class #: 3):
    Tx rate: 0 Kb (5%), Buffer size: 5 percent
Priority low
    PLP high: 1, PLP low: 1, PLP medium-high: 1, PLP medium-low: 1

...

RED drop profile index: 1, # entries: 1
                             Drop
Entry      Fullness(%)   Probability(%)
    0              100              100
```

show class-of-service forwarding-table lcc (TX Matrix Plus Router)

```
user@host> show class-of-service forwarding-table lcc 0
lcc0-re0:
--------------------------------------------------------------------------

Classifier table index: 9, # entries: 64, Table type: IPv6 DSCP
Entry #   Code point   Forwarding-class #   PLP
   0        000000        0        0
   1        000001        0        0
   2        000010        0        0
   3        000011        0        0
   4        000100        0        0
   5        000101        0        0
   6        000110        0        0
   7        000111        0        0
   8        001000        0        0
   9        001001        0        0
  10        001010        0        0
  11        001011        0        0
  12        001100        0        0
  13        001101        0        0
  14        001110        0        0
  15        001111        0        0
  16        010000        0        0
  17        010001        0        0
  18        010010        0        0
  19        010011        0        0
  20        010100        0        0
  21        010101        0        0
  22        010110        0        0
  23        010111        0        0
  24        011000        0        0
  25        011001        0        0
  26        011010        0        0
  27        011011        0        0
  28        011100        0        0
  29        011101        0        0
  30        011110        0        0
  31        011111        0        0
  32        100000        0        0
  33        100001        0        0
  34        100010        0        0
  35        100011        0        0
  36        100100        0        0
  37        100101        0        0
```



```
  38          100110          0          0
  39          100111          0          0
  40          101000          0          0
  41          101001          0          0
  42          101010          0          0
  43          101011          0          0
  44          101100          0          0
  45          101101          0          0
  46          101110          0          0
...
```

show class-of-service forwarding-table drop-profile

Syntax show class-of-service forwarding-table drop-profile

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Display the data points of all random early detection (RED) drop profiles as they exist in the forwarding table.

Options This command has no options.

Required Privilege Level view

List of Sample Output show class-of-service forwarding-table drop-profile on page 440

Output Fields Table 84 on page 440 describes the output fields for the **show class-of-service forwarding-table drop-profile** command. Output fields are listed in the approximate order in which they appear.

Table 84: show class-of-service forwarding-table drop-profile Output Fields

Field Name	Field Description
RED drop profile index	Index of this drop profile.
# entries	Number of entries in a particular RED drop profile index.
Entry	Drop profile entry number.
Fullness(%)	Percentage fullness of a queue.
Drop probability(%)	Drop probability at this fill level.

Sample Output

show class-of-service forwarding-table drop-profile

```
user@host> show class-of-service forwarding-table drop-profile
RED drop profile index: 4, # entries: 1
                                 Drop
Entry       Fullness(%)    Probability(%)
    0               100               100

RED drop profile index: 8742, # entries: 3
                                 Drop
Entry       Fullness(%)    Probability(%)
```



```
    0              10              10
    1              20              20
    2              30              30

RED drop profile index: 24627, # entries: 64
                             Drop
Entry      Fullness(%)   Probability(%)
    0               0               0
    1               1               1
    2               2               2
    3               4               4
 ...
   61              98              99
   62              99              99
   63             100             100

RED drop profile index: 25393, # entries: 64
                             Drop
Entry      Fullness(%)   Probability(%)
    0               0               0
    1               1               1
    2               2               2
    3               4               4
 ...
   61              98              98
   62              99              99
   63             100             100
```

show class-of-service forwarding-table scheduler-map

Syntax show class-of-service forwarding-table scheduler-map

Release Information Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description For each physical interface, display the scheduler map information as it exists in the forwarding table.

Options This command has no options.

Required Privilege Level view

List of Sample Output show class-of-service forwarding-table scheduler-map on page 443

Output Fields Table 85 on page 442 describes the output fields for the **show class-of-service forwarding-table scheduler-map** command. Output fields are listed in the approximate order in which they appear.

Table 85: show class-of-service forwarding-table scheduler-map Output Fields

Field Name	Field Description
Interface	Name of the physical interface.
Index	Physical interface index.
Map index	Scheduler map index.
Num of queues	Number of queues defined in this scheduler map.
Entry	Number of this entry in the scheduler map.
Scheduler index	Scheduler policy index.
Forwarding-class #	Forwarding class number to which this entry is applied.
Tx rate	Configured transmit rate of the scheduler (in bps). The rate is a percentage of the total interface bandwidth, or the keyword **remainder**, which indicates that the scheduler receives the remaining bandwidth of the interface.
Max buffer delay	Amount of transmit delay (in milliseconds) or buffer size of the queue. This amount is a percentage of the total interface buffer allocation or the keyword **remainder**, which indicates that the buffer is sized according to what remains after other scheduler buffer allocations.

Table 85: show class-of-service forwarding-table scheduler-map Output Fields (continued)

Field Name	Field Description
Priority	• **high**—Queue priority is high. • **low**—Queue priority is low.
PLP high	Drop profile index for a high packet loss priority profile.
PLP low	Drop profile index for a low packet loss priority profile.
PLP medium-high	Drop profile index for a medium-high packet loss priority profile.
PLP medium-low	Drop profile index for a medium-low packet loss priority profile.
TCP PLP high	Drop profile index for a high TCP packet loss priority profile.
TCP PLP low	Drop profile index for a low TCP packet loss priority profile.
Policy is exact	If this line appears in the output, exact rate limiting is enabled. Otherwise, no rate limiting is enabled.

Sample Output

show class-of-service forwarding-table scheduler-map

```
user@host> show class-of-service forwarding-table scheduler-map
Interface: so-5/0/0 (Index: 9, Map index: 17638, Num of queues: 2):
  Entry 0 (Scheduler index: 6090, Forwarding-class #: 0):
    Tx rate: 0 Kb (30%), Max buffer delay: 39 bytes (0%)
    Priority low
    PLP high: 25393, PLP low: 24627, TCP PLP high: 25393, TCP PLP low:8742
    Policy is exact
  Entry 1 (Scheduler index: 38372, Forwarding-class #: 1):
    Traffic chunk: Max = 0 bytes, Min = 0 bytes
    Tx rate: 0 Kb (40%), Max buffer delay: 68 bytes (0%)
    Priority high
    PLP high: 25393, PLP low: 24627, TCP PLP high: 25393, TCP PLP low: 8742

Interface: at-6/1/0 (Index: 10, Map index: 17638, Num of queues: 2):
  Entry 0 (Scheduler index: 6090, Forwarding-class #: 0):
    Traffic chunk: Max = 0 bytes, Min = 0 bytes
    Tx rate: 0 Kb (30%), Max buffer delay: 39 bytes (0%)
    Priority high
    PLP high: 25393, PLP low: 24627, TCP PLP high: 25393, TCP PLP low: 8742
  Entry 1 (Scheduler index: 38372, Forwarding-class #: 1):
    Traffic chunk: Max = 0 bytes, Min = 0 bytes
    Tx rate: 0 Kb (40%), Max buffer delay: 68 bytes (0%)
    Priority low
    PLP high: 25393, PLP low: 24627, TCP PLP high: 25393, TCP PLP low: 8742
```


show class-of-service interface

Syntax

show class-of-service interface
<comprehensive | detail> <*interface-name*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 9.0 for EX Series switches.
Forwarding class map information added in Junos OS Release 9.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 12.1 for the PTX Series Packet Transport routers.
Command introduced in Junos OS Release 12.2 for the ACX Series Universal Metro routers.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.
Options **detail** and **comprehensive** introduced in Junos OS Release 11.4.
Command introduced in Junos OS Release 15.1R3 on MX Series routers for enhanced subscriber management.

Description

Display the logical and physical interface associations for the classifier, rewrite rules, and scheduler map objects.

NOTE: On routing platforms with dual Routing Engines, running this command on the backup Routing Engine, with or without any of the available options, is not supported and produces the following error message:

error: the class-of-service subsystem is not running

Options

none—Display CoS associations for all physical and logical interfaces.

comprehensive—(M Series, MX Series, and T Series routers) (Optional) Display comprehensive quality-of-service (QoS) information about all physical and logical interfaces.

detail—(M Series, MX Series, and T Series routers) (Optional) Display QoS and CoS information based on the interface.

If the **interface** *interface-name* is a physical interface, the output includes:

- Brief QoS information about the physical interface
- Brief QoS information about the logical interface
- CoS information about the physical interface
- Brief information about filters or policers of the logical interface
- Brief CoS information about the logical interface

If the **interface** *interface-name* is a logical interface, the output includes:

- Brief QoS information about the logical interface
- Information about filters or policers for the logical interface
- CoS information about the logical interface

interface-name—(Optional) Display class-of-service (CoS) associations for the specified interface.

none—Display CoS associations for all physical and logical interfaces.

NOTE: ACX5000 routers do not support classification on logical interfaces and therefore do not show CoS associations for logical interfaces with this command.

Required Privilege Level view

Related Documentation

- *Verifying and Managing Junos OS Enhanced Subscriber Management*

List of Sample Output

Output Fields Table 22 on page 63 describes the output fields for the **show class-of-service interface** command. Output fields are listed in the approximate order in which they appear.

Table 86: show class-of-service interface Output Fields

Field Name	Field Description
Physical interface	Name of a physical interface.
Index	Index of this interface or the internal index of this object. (Enhanced subscriber management for MX Series routers) Index values for dynamic CoS traffic control profiles and dynamic scheduler maps are larger for enhanced subscriber management than they are for legacy subscriber management.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Dedicated Queues	Status of dedicated queues configured on an interface. Supported only on Trio MPC/MIC interfaces on MX Series routers. (Enhanced subscriber management for MX-Series routers) This field is not displayed for enhanced subscriber management.
Maximum usable queues	Number of queues you can configure on the interface.
Maximum usable queues	Maximum number of queues you can use.
Total non-default queues created	Number of queues created in addition to the default queues. Supported only on Trio MPC/MIC interfaces on MX Series routers. (Enhanced subscriber management for MX Series routers) This field is not displayed for enhanced subscriber management.
Rewrite Input IEEE Code-point	(QFX3500 switches only) IEEE 802.1p code point (priority) rewrite value. Incoming traffic from the Fibre Channel (FC) SAN is classified into the forwarding class specified in the native FC interface (NP_Port) fixed classifier and uses the priority specified as the IEEE 802.1p rewrite value.
Shaping rate	Maximum transmission rate on the physical interface. You can configure the shaping rate on the physical interface, or on the logical interface, but not on both. Therefore, the **Shaping rate** field is displayed for either the physical interface or the logical interface.
Scheduler map	Name of the output scheduler map associated with this interface. (Enhanced subscriber management for MX Series routers) The name of the dynamic scheduler map object is associated with a generated UID (for example, **SMAP-1_UID1002**) instead of with a subscriber interface.
Scheduler map forwarding class sets	(QFX Series only) Name of the output fabric scheduler map associated with a QFabric system Interconnect device interface.
Input shaping rate	For Gigabit Ethernet IQ2 PICs, maximum transmission rate on the input interface.
Input scheduler map	For Gigabit Ethernet IQ2 PICs, name of the input scheduler map associated with this interface.
Chassis scheduler map	Name of the scheduler map associated with the packet forwarding component queues.
Rewrite	Name and type of the rewrite rules associated with this interface.
Traffic-control-profile	Name of the associated traffic control profile. (Enhanced subscriber management for MX Series routers) The name of the dynamic traffic control profile object is associated with a generated UID (for example, **TC_PROF_100_199_SERIES_UID1006**) instead of with a subscriber interface.
Classifier	Name and type of classifiers associated with this interface.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Forwarding-class-map	Name of the forwarding map associated with this interface.
Congestion-notification	(QFX Series and EX4600 switches only) Congestion notification state, **enabled** or **disabled**.
Logical interface	Name of a logical interface.
Object	Category of an object: **Classifier**, **Fragmentation-map** (for LSQ interfaces only), **Scheduler-map**, **Rewrite**, **Translation Table** (for IQE PICs only), or **traffic-class-map** (for T4000 routers with Type 5 FPCs).
Name	Name of an object.
Type	Type of an object: **dscp**, **dscp-ipv6**, **exp**, **ieee-802.1**, **ip**, **inet-precedence**, or **ieee-802.1ad** (for traffic class map on T4000 routers with Type 5 FPCs)..
Link-level type	Encapsulation on the physical interface.
MTU	MTU size on the physical interface.
Speed	Speed at which the interface is running.
Loopback	Whether loopback is enabled and the type of loopback.
Source filtering	Whether source filtering is enabled or disabled.
Flow control	Whether flow control is enabled or disabled.
Auto-negotiation	(Gigabit Ethernet interfaces) Whether autonegotiation is enabled or disabled.
Remote-fault	(Gigabit Ethernet interfaces) Remote fault status. • **Online**—Autonegotiation is manually configured as online. • **Offline**—Autonegotiation is manually configured as offline.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Device flags	The **Device flags** field provides information about the physical device and displays one or more of the following values: • **Down**—Device has been administratively disabled. • **Hear-Own-Xmit**—Device receives its own transmissions. • **Link-Layer-Down**—The link-layer protocol has failed to connect with the remote endpoint. • **Loopback**—Device is in physical loopback. • **Loop-Detected**—The link layer has received frames that it sent, thereby detecting a physical loopback. • **No-Carrier**—On media that support carrier recognition, no carrier is currently detected. • **No-Multicast**—Device does not support multicast traffic. • **Present**—Device is physically present and recognized. • **Promiscuous**—Device is in promiscuous mode and recognizes frames addressed to all physical addresses on the media. • **Quench**—Transmission on the device is quenched because the output buffer is overflowing. • **Recv-All-Multicasts**—Device is in multicast promiscuous mode and therefore provides no multicast filtering. • **Running**—Device is active and enabled.

Interface flags	The **Interface flags** field provides information about the physical interface and displays one or more of the following values:
	• **Admin-Test**—Interface is in test mode and some sanity checking, such as loop detection, is disabled.
	• **Disabled**—Interface is administratively disabled.
	• **Down**—A hardware failure has occurred.
	• **Hardware-Down**—Interface is nonfunctional or incorrectly connected.
	• **Link-Layer-Down**—Interface keepalives have indicated that the link is incomplete.
	• **No-Multicast**—Interface does not support multicast traffic.
	• **No-receive No-transmit**—Passive monitor mode is configured on the interface.
	• **Point-To-Point**—Interface is point-to-point.
	• **Pop all MPLS labels from packets of depth**—MPLS labels are removed as packets arrive on an interface that has the **pop-all-labels** statement configured. The depth value can be one of the following:
	  ◦ **1**—Takes effect for incoming packets with one label only.
	  ◦ **2**—Takes effect for incoming packets with two labels only.
	  ◦ **[1 2]**—Takes effect for incoming packets with either one or two labels.
	• **Promiscuous**—Interface is in promiscuous mode and recognizes frames addressed to all physical addresses.
	• **Recv-All-Multicasts**—Interface is in multicast promiscuous mode and provides no multicast filtering.
	• **SNMP-Traps**—SNMP trap notifications are enabled.
	• **Up**—Interface is enabled and operational.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Flags	The **Logical interface flags** field provides information about the logical interface and displays one or more of the following values: • **ACFC Encapsulation**—Address control field Compression (ACFC) encapsulation is enabled (negotiated successfully with a peer). • **Device-down**—Device has been administratively disabled. • **Disabled**—Interface is administratively disabled. • **Down**—A hardware failure has occurred. • **Clear-DF-Bit**—GRE tunnel or IPsec tunnel is configured to clear the Don't Fragment (DF) bit. • **Hardware-Down**—Interface protocol initialization failed to complete successfully. • **PFC**—Protocol field compression is enabled for the PPP session. • **Point-To-Point**—Interface is point-to-point. • **SNMP-Traps**—SNMP trap notifications are enabled. • **Up**—Interface is enabled and operational.
Encapsulation	Encapsulation on the logical interface.
Admin	Administrative state of the interface (**Up** or **Down**)
Link	Status of physical link (**Up** or **Down**).
Proto	Protocol configured on the interface.
Input Filter	Names of any firewall filters to be evaluated when packets are received on the interface, including any filters attached through activation of dynamic service.
Output Filter	Names of any firewall filters to be evaluated when packets are transmitted on the interface, including any filters attached through activation of dynamic service.
Link flags	Provides information about the physical link and displays one or more of the following values: • **ACFC**—Address control field compression is configured. The Point-to-Point Protocol (PPP) session negotiates the ACFC option. • **Give-Up**—Link protocol does not continue connection attempts after repeated failures. • **Loose-LCP**—PPP does not use the Link Control Protocol (LCP) to indicate whether the link protocol is operational. • **Loose-LMI**—Frame Relay does not use the Local Management Interface (LMI) to indicate whether the link protocol is operational. • **Loose-NCP**—PPP does not use the Network Control Protocol (NCP) to indicate whether the device is operational. • **Keepalives**—Link protocol keepalives are enabled. • **No-Keepalives**—Link protocol keepalives are disabled. • **PFC**—Protocol field compression is configured. The PPP session negotiates the PFC option.
Hold-times	Current interface hold-time up and hold-time down, in milliseconds.
CoS queues	Number of CoS queues configured.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Last flapped	Date, time, and how long ago the interface went from down to up. The format is **Last flapped:** ***year-month-day hour*:*minute*:*second*:*timezone*** **(*hour*:*minute*:*second* ago)**. For example, **Last flapped: 2002-04-26 10:52:40 PDT (04:33:20 ago).**
Statistics last cleared	Number and rate of bytes and packets received and transmitted on the physical interface. • **Input bytes**—Number of bytes received on the interface. • **Output bytes**—Number of bytes transmitted on the interface. • **Input packets**—Number of packets received on the interface. • **Output packets**—Number of packets transmitted on the interface.
Exclude Overhead Bytes	Exclude the counting of overhead bytes from aggregate queue statistics. • **Disabled**—Default configuration. Includes the counting of overhead bytes in aggregate queue statistics. • **Enabled**—Excludes the counting of overhead bytes from aggregate queue statistics for just the physical interface. • **Enabled for hierarchy**—Excludes the counting of overhead bytes from aggregate queue statistics for the physical interface as well as all child interfaces, including logical interfaces and interface sets.
IPv6 transit statistics	Number of IPv6 transit bytes and packets received and transmitted on the logical interface if IPv6 statistics tracking is enabled.
Input errors	Input errors on the interface. The labels are explained in the following list: • **Errors**—Sum of the incoming frame aborts and FCS errors. • **Drops**—Number of packets dropped by the input queue of the I/O Manager ASIC. If the interface is saturated, this number increments once for every packet that is dropped by the ASIC's RED mechanism. • **Framing errors**—Number of packets received with an invalid frame checksum (FCS). • **Runts**—Number of frames received that are smaller than the runt threshold. • **Giants**—Number of frames received that are larger than the giant threshold. • **Bucket Drops**—Drops resulting from the traffic load exceeding the interface transmit or receive leaky bucket configuration. • **Policed discards**—Number of frames that the incoming packet match code discarded because they were not recognized or not of interest. Usually, this field reports protocols that Junos OS does not handle. • **L3 incompletes**—Number of incoming packets discarded because they failed Layer 3 (usually IPv4) sanity checks of the header. For example, a frame with less than 20 bytes of available IP header is discarded. Layer 3 incomplete errors can be ignored by configuring the **ignore-l3-incompletes** statement. • **L2 channel errors**—Number of times the software did not find a valid logical interface for an incoming frame. • **L2 mismatch timeouts**—Number of malformed or short packets that caused the incoming packet handler to discard the frame as unreadable. • **HS link CRC errors**—Number of errors on the high-speed links between the ASICs responsible for handling the router interfaces. • **HS link FIFO overflows**—Number of FIFO overflows on the high-speed links between the ASICs responsible for handling the router interfaces.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Output errors	Output errors on the interface. The labels are explained in the following list: • **Carrier transitions**—Number of times the interface has gone from **down** to **up**. This number does not normally increment quickly, increasing only when the cable is unplugged, the far-end system is powered down and up, or another problem occurs. If the number of carrier transitions increments quickly (perhaps once every 10 seconds), the cable, the far-end system, or the PIC is malfunctioning. • **Errors**—Sum of the outgoing frame aborts and FCS errors. • **Drops**—Number of packets dropped by the output queue of the I/O Manager ASIC. If the interface is saturated, this number increments once for every packet that is dropped by the ASIC's RED mechanism. NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), the **Drops** field does not always use the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs. • **Aged packets**—Number of packets that remained in shared packet SDRAM so long that the system automatically purged them. The value in this field should never increment. If it does, it is most likely a software bug or possibly malfunctioning hardware. • **HS link FIFO underflows**—Number of FIFO underflows on the high-speed links between the ASICs responsible for handling the router interfaces. • **MTU errors**—Number of packets whose size exceeds the MTU of the interface.
Egress queues	Total number of egress Maximum usable queues on the specified interface.
Queue counters	CoS queue number and its associated user-configured forwarding class name. • **Queued packets**—Number of queued packets. • **Transmitted packets**—Number of transmitted packets. • **Dropped packets**—Number of packets dropped by the ASIC's RED mechanism. NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), the **Dropped packets** field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.
SONET alarms **SONET defects**	(SONET) SONET media-specific alarms and defects that prevent the interface from passing packets. When a defect persists for a certain period, it is promoted to an alarm. Based on the router configuration, an alarm can ring the red or yellow alarm bell on the router or light the red or yellow alarm LED on the craft interface. See these fields for possible alarms and defects: **SONET PHY**, **SONET section**, **SONET line**, and **SONET path**.
SONET PHY	Counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET PHY** field has the following subfields: • **PLL Lock**—Phase-locked loop • **PHY Light**—Loss of optical signal

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
SONET section	Counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET section** field has the following subfields: • **BIP-B1**—Bit interleaved parity for SONET section overhead • **SEF**—Severely errored framing • **LOS**—Loss of signal • **LOF**—Loss of frame • **ES-S**—Errored seconds (section) • **SES-S**—Severely errored seconds (section) • **SEFS-S**—Severely errored framing seconds (section)
SONET line	Active alarms and defects, plus counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET line** field has the following subfields: • **BIP-B2**—Bit interleaved parity for SONET line overhead • **REI-L**—Remote error indication (near-end line) • **RDI-L**—Remote defect indication (near-end line) • **AIS-L**—Alarm indication signal (near-end line) • **BERR-SF**—Bit error rate fault (signal failure) • **BERR-SD**—Bit error rate defect (signal degradation) • **ES-L**—Errored seconds (near-end line) • **SES-L**—Severely errored seconds (near-end line) • **UAS-L**—Unavailable seconds (near-end line) • **ES-LFE**—Errored seconds (far-end line) • **SES-LFE**—Severely errored seconds (far-end line) • **UAS-LFE**—Unavailable seconds (far-end line)

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
SONET path	Active alarms and defects, plus counts of specific SONET errors with detailed information. • **Seconds**—Number of seconds the defect has been active. • **Count**—Number of times that the defect has gone from inactive to active. • **State**—State of the error. A state other than **OK** indicates a problem. The **SONET path** field has the following subfields: • **BIP-B3**—Bit interleaved parity for SONET section overhead • **REI-P**—Remote error indication • **LOP-P**—Loss of pointer (path) • **AIS-P**—Path alarm indication signal • **RDI-P**—Path remote defect indication • **UNEQ-P**—Path unequipped • **PLM-P**—Path payload (signal) label mismatch • **ES-P**—Errored seconds (near-end STS path) • **SES-P**—Severely errored seconds (near-end STS path) • **UAS-P**—Unavailable seconds (near-end STS path) • **ES-PFE**—Errored seconds (far-end STS path) • **SES-PFE**—Severely errored seconds (far-end STS path) • **UAS-PFE**—Unavailable seconds (far-end STS path)
Received SONET overhead **Transmitted SONET overhead**	Values of the received and transmitted SONET overhead: • **C2**—Signal label. Allocated to identify the construction and content of the STS-level SPE and for PDI-P. • **F1**—Section user channel byte. This byte is set aside for the purposes of users. • **K1** and **K2**—These bytes are allocated for APS signaling for the protection of the multiplex section. • **J0**—Section trace. This byte is defined for STS-1 number 1 of an STS-*N* signal. Used to transmit a 1-byte fixed-length string or a 16-byte message so that a receiving terminal in a section can verify its continued connection to the intended transmitter. • **S1**—Synchronization status. The S1 byte is located in the first STS-1 number of an STS-*N* signal. • **Z3** and **Z4**—Allocated for future use.
Received path trace **Transmitted path trace**	SONET/SDH interfaces allow path trace bytes to be sent inband across the SONET/SDH link. Juniper Networks and other router manufacturers use these bytes to help diagnose misconfigurations and network errors by setting the transmitted path trace message so that it contains the system hostname and name of the physical interface. The received path trace value is the message received from the router at the other end of the fiber. The transmitted path trace value is the message that this router transmits.
HDLC configuration	Information about the HDLC configuration. • **Policing bucket**—Configured state of the receiving policer. • **Shaping bucket**—Configured state of the transmitting shaper. • **Giant threshold**—Giant threshold programmed into the hardware. • **Runt threshold**—Runt threshold programmed into the hardware.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Packet Forwarding Engine configuration	Information about the configuration of the Packet Forwarding Engine: • **Destination slot**—FPC slot number. • **PLP byte**—Packet Level Protocol byte.
CoS information	Information about the CoS queue for the physical interface. • **CoS transmit queue**—Queue number and its associated user-configured forwarding class name. • **Bandwidth %**—Percentage of bandwidth allocated to the queue. • **Bandwidth bps**—Bandwidth allocated to the queue (in bps). • **Buffer %**—Percentage of buffer space allocated to the queue. • **Buffer usec**—Amount of buffer space allocated to the queue, in microseconds. This value is nonzero only if the buffer size is configured in terms of time. • **Priority**—Queue priority: **low** or **high**. • **Limit**—Displayed if rate limiting is configured for the queue. Possible values are **none** and **exact**. If **exact** is configured, the queue transmits only up to the configured bandwidth, even if excess bandwidth is available. If **none** is configured, the queue transmits beyond the configured bandwidth if bandwidth is available.
Forwarding classes	Total number of forwarding classes supported on the specified interface.
Egress queues	Total number of egress Maximum usable queues on the specified interface.
Queue	Queue number.
Forwarding classes	Forwarding class name.
Queued Packets	Number of packets queued to this queue.
Queued Bytes	Number of bytes queued to this queue. The byte counts vary by PIC type.
Transmitted Packets	Number of packets transmitted by this queue. When fragmentation occurs on the egress interface, the first set of packet counters shows the postfragmentation values. The second set of packet counters (displayed under the **Packet Forwarding Engine Chassis Queues** field) shows the prefragmentation values.
Transmitted Bytes	Number of bytes transmitted by this queue. The byte counts vary by PIC type.
Tail-dropped packets	Number of packets dropped because of tail drop.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description

RED-dropped packets	Number of packets dropped because of random early detection (RED).
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, the total number of dropped packets is displayed. On all other M Series routers, the output classifies dropped packets into the following categories:
	  • **Low, non-TCP**—Number of low-loss priority non-TCP packets dropped because of RED.
	  • **Low, TCP**—Number of low-loss priority TCP packets dropped because of RED.
	  • **High, non-TCP**—Number of high-loss priority non-TCP packets dropped because of RED.
	  • **High, TCP**—Number of high-loss priority TCP packets dropped because of RED.
	• (MX Series routers with enhanced DPCs, and T Series routers with enhanced FPCs only) The output classifies dropped packets into the following categories:
	  • **Low**—Number of low-loss priority packets dropped because of RED.
	  • **Medium-low**—Number of medium-low loss priority packets dropped because of RED.
	  • **Medium-high**—Number of medium-high loss priority packets dropped because of RED.
	  • **High**—Number of high-loss priority packets dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

RED-dropped bytes	Number of bytes dropped because of RED. The byte counts vary by PIC type.
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, only the total number of dropped bytes is displayed. On all other M Series routers, the output classifies dropped bytes into the following categories:
	  • **Low, non-TCP**—Number of low-loss priority non-TCP bytes dropped because of RED.
	  • **Low, TCP**—Number of low-loss priority TCP bytes dropped because of RED.
	  • **High, non-TCP**—Number of high-loss priority non-TCP bytes dropped because of RED.
	  • **High, TCP**—Number of high-loss priority TCP bytes dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

Transmit rate	Configured transmit rate of the scheduler. The rate is a percentage of the total interface bandwidth.
Rate Limit	Rate limiting configuration of the queue. Possible values are : • **None**—No rate limit. • **exact**—Queue transmits at the configured rate.
Buffer size	Delay buffer size in the queue.
Priority	Scheduling priority configured as **low** or **high**.
Excess Priority	Priority of the excess bandwidth traffic on a scheduler: **low**, **medium-low**, **medium-high**, **high**, or **none**.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description
Drop profiles	Display the assignment of drop profiles. • **Loss priority**—Packet loss priority for drop profile assignment. • **Protocol**—Transport protocol for drop profile assignment. • **Index**—Index of the indicated object. Objects that have indexes in this output include schedulers and drop profiles. • **Name**—Name of the drop profile. • **Type**—Type of the drop profile: **discrete** or **interpolated**. • **Fill Level**—Percentage fullness of a queue. • **Drop probability**—Drop probability at this fill level.
Excess Priority	Priority of the excess bandwidth traffic on a scheduler.
Drop profiles	Display the assignment of drop profiles. • **Loss priority**—Packet loss priority for drop profile assignment. • **Protocol**—Transport protocol for drop profile assignment. • **Index**—Index of the indicated object. Objects that have indexes in this output include schedulers and drop profiles. • **Name**—Name of the drop profile. • **Type**—Type of the drop profile: **discrete** or **interpolated**. • **Fill Level**—Percentage fullness of a queue. • **Drop probability**—Drop probability at this fill level.

Table 86: show class-of-service interface Output Fields (continued)

Field Name	Field Description

Adjustment information	Display the assignment of shaping-rate adjustments on a scheduler node or queue.
	• **Adjusting application**—Application that is performing the shaping-rate adjustment.
	  ◦ The adjusting application can appear as **ancp LS-0**, which is the Junos OS Access Node Control Profile process (**ancpd**) that performs shaping-rate adjustments on schedule nodes.
	  ◦ The adjusting application can appear as **DHCP**, which adjusts the shaping-rate and overhead-accounting class-of-service attributes based on DHCP option 82, suboption 9 (Vendor Specific Information). The shaping rate is based on the actual-data-rate-downstream attribute. The overhead accounting value is based on the access-loop-encapsulation attribute and specifies whether the access loop uses Ethernet (frame mode) or ATM (cell mode).
	  ◦ The adjusting application can also appear as **pppoe**, which adjusts the shaping-rate and overhead-accounting class-of-service attributes on dynamic subscriber interfaces in a broadband access network based on access line parameters in Point-to-Point Protocol over Ethernet (PPPoE) Tags [TR-101]. This feature is supported on MPC/MIC interfaces on MX Series routers. The shaping rate is based on the actual-data-rate-downstream attribute. The overhead accounting value is based on the access-loop-encapsulation attribute and specifies whether the access loop uses Ethernet (frame mode) or ATM (cell mode).
	• **Adjustment type**—Type of adjustment: **absolute** or **delta**.
	• **Configured shaping rate**—Shaping rate configured for the scheduler node or queue.
	• **Adjustment value**—Value of adjusted shaping rate.
	• **Adjustment target**—Level of shaping-rate adjustment performed: **node** or **queue**.
	• **Adjustment overhead-accounting mode**—Configured shaping mode: **frame** or **cell**.
	• **Adjustment overhead bytes**—Number of bytes that the ANCP agent adds to or subtracts from the actual downstream frame overhead before reporting the adjusted values to CoS.
	• **Adjustment target**—Level of shaping-rate adjustment performed: **node** or **queue**.
	• **Adjustment multicast index**—

Sample Output

show class-of-service interface (Physical)

```
user@host> show class-of-service interface so-0/2/3
Physical interface: so-0/2/3, Index: 135
Maximum usable queues: 8, Queues in use: 4
Total non-default queues created: 4
  Scheduler map: <default>, Index: 2032638653

  Logical interface: fe-0/0/1.0, Index: 68, Dedicated Queues: no
    Shaping rate: 32000
    Object                  Name                  Type           Index
    Scheduler-map           <default>                             27
    Rewrite                 exp-default           exp             21
    Classifier              exp-default           exp             5
    Classifier              ipprec-compatibility  ip              8
    Forwarding-class-map    exp-default           exp             5
```

show class-of-service interface (Logical)

```
user@host> show class-of-service interface so-0/2/3.0
Logical interface: so-0/2/3.0, Index: 68, Dedicated Queues: no
    Shaping rate: 32000
```



```
Object                   Name                  Type             Index
Scheduler-map            <default>                               27
Rewrite                  exp-default           exp               21
Classifier               exp-default           exp               5
Classifier               ipprec-compatibility  ip                8
Forwarding-class-map     exp-default           exp               5
```

show class-of-service interface (Gigabit Ethernet)

```
user@host> show class-of-service interface ge-6/2/0
Physical interface: ge-6/2/0, Index: 175
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Input scheduler map: <default>, Index: 3
  Chassis scheduler map: <default-chassis>, Index: 4
```

show class-of-service interface (ANCP)

```
user@host> show class-of-service interface pp0.1073741842
  Logical interface: pp0.1073741842, Index: 341
Object                 Name                   Type                   Index
Traffic-control-profile TCP-CVLAN             Output                 12408
Classifier             dscp-ipv6-compatibility dscp-ipv6                 9
Classifier             ipprec-compatibility   ip                        13

    Adjusting application: ancp LS-0
      Adjustment type: absolute
      Configured shaping rate: 4000000
      Adjustment value: 11228000
      Adjustment overhead-accounting mode: Frame Mode
      Adjustment overhead bytes: 50
      Adjustment target: node
```

show class-of-service interface (PPPoE Interface)

```
user@host> show class-of-service interface pp0.1
Logical interface: pp0.1, Index: 85
    Object                 Name                   Type           Index
    Traffic-control-profile tcp-pppoe.o.pp0.1     Output         2726446535
    Classifier             ipprec-compatibility   ip             13

    Adjusting application: PPPoE
      Adjustment type: absolute
      Adjustment value: 5000000
      Adjustment overhead-accounting mode: cell
      Adjustment target: node
```

show class-of-service interface (DHCP Interface)

```
user@host> show class-of-service interface demux0.1
Logical interface: pp0.1, Index: 85
    Object                 Name                   Type           Index
    Traffic-control-profile tcp-dhcp.o.demux0.1   Output         2726446535
    Classifier             ipprec-compatibility   ip             13

    Adjusting application: DHCP
      Adjustment type: absolute
      Adjustment value: 5000000
      Adjustment overhead-accounting mode: cell
      Adjustment target: node
```


show class-of-service interface (T4000 Routers with Type 5 FPCs)

```
user@host> show class-of-service interface xe-4/0/0
Physical interface: xe-4/0/0, Index: 153
    Maximum usable queues: 8, Queues in use: 4
      Shaping rate: 5000000000 bps
      Scheduler map: <default>, Index: 2
      Congestion-notification: Disabled

      Logical interface: xe-4/0/0.0, Index: 77
        Object                 Name                   Type
Index
        Classifier             ipprec-compatibility   ip
  13
```

show class-of-service interface detail

```
user@host> show class-of-service interface ge-0/3/0 detail

Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Link-level type: Ethernet, MTU: 1518, Speed: 1000mbps, Loopback: Disabled,
Source filtering: Disabled, Flow control: Enabled, Auto-negotiation: Enabled,
Remote fault: Online
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000

  Physical interface: ge-0/3/0, Index: 138
  Maximum usable queues: 4, Queues in use: 5
  Shaping rate: 50000 bps
  Scheduler map: interface-schedular-map, Index: 58414
  Input shaping rate: 10000 bps
  Input scheduler map: schedular-map, Index: 15103
  Chassis scheduler map: <default-chassis>, Index: 4
  Congestion-notification: Disabled

  Logical interface ge-0/3/0.0
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    inet
    mpls
Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.0      up    up   inet
                           mpls
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.0      up    up   inet
                           mpls

  Logical interface: ge-0/3/0.0, Index: 68
    Object                 Name                   Type                 Index
    Rewrite                exp-default            exp (mpls-any)         33
    Classifier             exp-default            exp                    10
    Classifier             ipprec-compatibility   ip                     13


  Logical interface ge-0/3/0.1
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    inet
Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.1      up    up   inet
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.1      up    up   inet
```



```
  Logical interface: ge-0/3/0.1, Index: 69
    Object                  Name                   Type            Index
    Classifier              ipprec-compatibility   ip               13
```

show class-of-service interface comprehensive

```
user@host> show class-of-service interface ge-0/3/0 comprehensive
Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Interface index: 138, SNMP ifIndex: 601, Generation: 141
  Link-level type: Ethernet, MTU: 1518, Speed: 1000mbps, BPDU Error: None,
MAC-REWRITE Error: None, Loopback: Disabled, Source filtering: Disabled, Flow
control: Enabled,
  Auto-negotiation: Enabled, Remote fault: Online
  Device flags   : Present Running
  Interface flags: SNMP-Traps Internal: 0x4000
  CoS queues     : 4 supported, 4 maximum usable queues
  Schedulers     : 256
  Hold-times     : Up 0 ms, Down 0 ms
  Current address: 00:14:f6:f4:b4:5d, Hardware address: 00:14:f6:f4:b4:5d
  Last flapped   : 2010-09-07 06:35:22 PDT (15:14:42 ago)
  Statistics last cleared: Never  Exclude Overhead Bytes: Disabled
  Traffic statistics:
   Input  bytes  :                   0                    0 bps
   Output bytes  :                   0                    0 bps
   Input  packets:                   0                    0 pps
   Output packets:                   0                    0 pps
   IPv6 total statistics:
    Input  bytes  :                  0
    Output bytes  :                  0
    Input  packets:                  0
    Output packets:                  0
  Ingress traffic statistics at Packet Forwarding Engine:
   Input  bytes  :                   0                    0 bps
   Input  packets:                   0                    0 pps
   Drop   bytes  :                   0                    0 bps
   Drop   packets:                   0                    0 pps
  Label-switched interface (LSI) traffic statistics:
   Input  bytes  :                   0                    0 bps
   Input  packets:                   0                    0 pps
  Input errors:
    Errors: 0, Drops: 0, Framing errors: 0, Runts: 0, Policed discards: 0, L3
incompletes: 0, L2 channel errors: 0, L2 mismatch timeouts: 0, FIFO errors: 0,
Resource errors: 0
  Output errors:
    Carrier transitions: 5, Errors: 0, Drops: 0, Collisions: 0, Aged packets: 0,
 FIFO errors: 0, HS link CRC errors: 0, MTU errors: 0, Resource errors: 0
  Ingress queues: 4 supported, 5 in use
  Queue counters:       Queued packets  Transmitted packets      Dropped packets

    0 af3                            0                    0                    0

    1 af2                            0                    0                    0

    2 ef2                            0                    0                    0

    3 ef1                            0                    0                    0

  Egress queues: 4 supported, 5 in use
  Queue counters:       Queued packets  Transmitted packets      Dropped packets
```



```
    0 af3                          0                    0                    0

    1 af2                          0                    0                    0

    2 ef2                          0                    0                    0

    3 ef1                          0                    0                    0

  Active alarms  : None
  Active defects : None
  MAC statistics:                      Receive         Transmit
    Total octets                             0                0
    Total packets                            0                0
    Unicast packets                          0                0
    Broadcast packets                        0                0
    Multicast packets                        0                0
    CRC/Align errors                         0                0
    FIFO errors                              0                0
    MAC control frames                       0                0
    MAC pause frames                         0                0
    Oversized frames                         0
    Jabber frames                            0
    Fragment frames                          0
    VLAN tagged frames                       0
    Code violations                          0
  Filter statistics:
    Input packet count                       0
    Input packet rejects                     0
    Input DA rejects                         0
    Input SA rejects                         0
    Output packet count                                       0
    Output packet pad count                                   0
    Output packet error count                                 0
    CAM destination filters: 0, CAM source filters: 0
  Autonegotiation information:
    Negotiation status: Complete
    Link partner:
       Link mode: Full-duplex, Flow control: Symmetric/Asymmetric, Remote fault:
 OK
    Local resolution:
      Flow control: Symmetric, Remote fault: Link OK
  Packet Forwarding Engine configuration:
    Destination slot: 0
  CoS information:
    Direction : Output
    CoS transmit queue               Bandwidth               Buffer Priority
Limit
                              %            bps     %            usec
    2 ef2                    39          19500     0             120     high
none
    Direction : Input
    CoS transmit queue               Bandwidth               Buffer Priority
Limit
                              %            bps     %            usec
    0 af3                    30           3000    45               0      low
none

Physical interface: ge-0/3/0, Enabled, Physical link is Up
  Interface index: 138, SNMP ifIndex: 601
Forwarding classes: 16 supported, 5 in use
Ingress queues: 4 supported, 5 in use
```

```
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Forwarding classes: 16 supported, 5 in use
Egress queues: 4 supported, 5 in use
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                     0                     0 pps
    RL-dropped bytes     :                     0                     0 bps
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                      0 pps
    Bytes                :                     0                     0 bps
```



```
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps

Packet Forwarding Engine Chassis Queues:
Queues: 4 supported, 5 in use
Queue: 0, Forwarding classes: af3
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available
Queue: 1, Forwarding classes: af2
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available
Queue: 2, Forwarding classes: ef2
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                      0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
```

```
    RED-dropped bytes    : Not Available
Queue: 3, Forwarding classes: ef1
  Queued:
    Packets              :               108546                     0 pps
    Bytes                :             12754752                   376 bps
  Transmitted:
    Packets              :               108546                      0 pps
    Bytes                :             12754752                   376 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  : Not Available
    RED-dropped bytes    : Not Available

Physical interface: ge-0/3/0, Index: 138
Maximum usable queues: 4, Queues in use: 5
  Shaping rate: 50000 bps

Scheduler map: interface-schedular-map, Index: 58414

  Scheduler: ef2, Forwarding class: ef2, Index: 39155
    Transmit rate: 39 percent, Rate Limit: none, Buffer size: 120 us, Buffer
Limit: none, Priority: high
    Excess Priority: unspecified
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
  Input shaping rate: 10000 bps
  Input scheduler map: schedular-map

Scheduler map: schedular-map, Index: 15103

  Scheduler: af3, Forwarding class: af3, Index: 35058
    Transmit rate: 30 percent, Rate Limit: none, Buffer size: 45 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: unspecified
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any         40582    green
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any         18928    yellow
Drop profile: green, Type: discrete, Index: 40582
  Fill level    Drop probability
          50                   0
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
```



```
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: yellow, Type: discrete, Index: 18928
  Fill level    Drop probability
         50                  0
        100                100
  Chassis scheduler map: < default-drop-profile>
Scheduler map: < default-drop-profile>, Index: 4

  Scheduler: < default-drop-profile>, Forwarding class: af3, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100

  Scheduler: < default-drop-profile>, Forwarding class: af2, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any             1    < default-drop-profile>
      Medium low      any             1    < default-drop-profile>
      Medium high     any             1    < default-drop-profile>
      High            any             1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
        100                100

  Scheduler: < default-drop-profile>, Forwarding class: ef2, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
```

```
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any            1    < default-drop-profile>
      Medium low      any            1    < default-drop-profile>
      Medium high     any            1    < default-drop-profile>
      High            any            1    < default-drop-profile>
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100

  Scheduler: < default-drop-profile>, Forwarding class: ef1, Index: 25
    Transmit rate: 25 percent, Rate Limit: none, Buffer size: 25 percent, Buffer
 Limit: none, Priority: low
    Excess Priority: low
    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any            1    < default-drop-profile>
      Medium low      any            1    < default-drop-profile>
      Medium high     any            1    < default-drop-profile>
      High            any            1    < default-drop-profile>
Drop profile: , Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
Drop profile: < default-drop-profile>, Type: discrete, Index: 1
  Fill level    Drop probability
         100                 100
  Congestion-notification: Disabled
Forwarding class                       ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                   0       0          0             low
            normal
  af2                                   1       1          1             low
            normal
  ef2                                   2       2          2             high
            normal
  ef1                                   3       3          3             high
            normal
  af1                                   4       4          0             low
            normal

  Logical interface ge-0/3/0.0 (Index 68) (SNMP ifIndex 152) (Generation 159)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    Traffic statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
```



```
     Output packets:                   0
    Local statistics:
     Input  bytes  :                   0
     Output bytes  :                   0
     Input  packets:                   0
     Output packets:                   0
    Transit statistics:
     Input  bytes  :                   0                    0 bps
     Output bytes  :                   0                    0 bps
     Input  packets:                   0                    0 pps
     Output packets:                   0                    0 pps
    Protocol inet, MTU: 1500, Generation: 172, Route table: 0
      Flags: Sendbcast-pkt-to-re
      Input Filters: filter-in-ge-0/3/0.0-i,
      Policer: Input: p1-ge-0/3/0.0-inet-i
    Protocol mpls, MTU: 1488, Maximum labels: 3, Generation: 173, Route table: 0

      Flags: Is-Primary
      Output Filters: exp-filter,,,,,

Logical interface ge-1/2/0.0 (Index 347) (SNMP ifIndex 638) (Generation 156)

Forwarding class ID  Queue  Restricted queue  Fabric priority  Policing priority
   SPU priority
best-effort      0   0      0                 low              normal
    low

Aggregate Forwarding-class statistics per forwarding-class
  Aggregate Forwarding-class statistics:
   Forwarding-class statistics:

Forwarding-class best-effort statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

IPv4 protocol forwarding-class statistics:
  Forwarding-class statistics:
  Forwarding-class best-effort statistics:

    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0
```



```
    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

    IPv6  protocol forwarding-class statistics:
    Forwarding-class statistics:
     Forwarding-class best-effort statistics:

    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Forwarding-class expedited-forwarding statistics:
    Input unicast bytes:      0
    Output unicast bytes:     0
    Input unicast packets:    0
    Output unicast packets:   0

    Input multicast bytes:    0
    Output multicast bytes:   0
    Input multicast packets:  0
    Output multicast packets: 0

Logical interface ge-0/3/0.0 (Index 68) (SNMP ifIndex 152)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.1 ]  Encapsulation: ENET2
    Input packets : 0
    Output packets: 0

Interface       Admin Link Proto Input Filter         Output Filter
ge-0/3/0.0      up    up   inet  filter-in-ge-0/3/0.0-i
                           mpls                       exp-filter
Interface       Admin Link Proto Input Policer         Output Policer
ge-0/3/0.0      up    up
                           inet  p1-ge-0/3/0.0-inet-i
                           mpls

Filter: filter-in-ge-0/3/0.0-i
Counters:
Name                                                Bytes              Packets
count-filter-in-ge-0/3/0.0-i                            0                    0

Filter: exp-filter
```

```
Counters:
Name                                                Bytes              Packets
count-exp-seven-match                                   0                    0
count-exp-zero-match                                    0                    0
Policers:
Name                                              Packets
p1-ge-0/3/0.0-inet-i                                    0


  Logical interface: ge-0/3/0.0, Index: 68
    Object                 Name                   Type                    Index

    Rewrite                exp-default            exp (mpls-any)             33

Rewrite rule: exp-default, Code point type: exp, Index: 33
  Forwarding class                   Loss priority       Code point
  af3                                low                 000
  af3                                high                001
  af2                                low                 010
  af2                                high                011
  ef2                                low                 100
  ef2                                high                101
  ef1                                low                 110
  ef1                                high                111
    Object                 Name                   Type                    Index

    Classifier             exp-default            exp                        10

Classifier: exp-default, Code point type: exp, Index: 10
  Code point         Forwarding class                   Loss priority
  000                af3                                low
  001                af3                                high
  010                af2                                low
  011                af2                                high
  100                ef2                                low
  101                ef2                                high
  110                ef1                                low
  111                ef1                                high
    Object                 Name                   Type                    Index

    Classifier             ipprec-compatibility   ip                         13

Classifier: ipprec-compatibility, Code point type: inet-precedence, Index: 13
  Code point         Forwarding class                   Loss priority
  000                af3                                low
  001                af3                                high
  010                af3                                low
  011                af3                                high
  100                af3                                low
  101                af3                                high
  110                ef1                                low
  111                ef1                                high
Forwarding class                   ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                               0       0          0              low
          normal
  af2                               1       1          1              low
          normal
  ef2                               2       2          2              high
          normal
  ef1                               3       3          3              high
```

```
          normal
  af1                                 4       4          0             low
          normal

  Logical interface ge-0/3/0.1 (Index 69) (SNMP ifIndex 154) (Generation 160)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    Traffic statistics:
     Input  bytes  :                  0
     Output bytes  :                  0
     Input  packets:                  0
     Output packets:                  0
    Local statistics:
     Input  bytes  :                  0
     Output bytes  :                  0
     Input  packets:                  0
     Output packets:                  0
    Transit statistics:
     Input  bytes  :                  0                    0 bps
     Output bytes  :                  0                    0 bps
     Input  packets:                  0                    0 pps
     Output packets:                  0                    0 pps
    Protocol inet, MTU: 1500, Generation: 174, Route table: 0
      Flags: Sendbcast-pkt-to-re

  Logical interface ge-0/3/0.1 (Index 69) (SNMP ifIndex 154)
    Flags: SNMP-Traps 0x4000 VLAN-Tag [ 0x8100.2 ]  Encapsulation: ENET2
    Input packets : 0
    Output packets: 0


Interface       Admin Link Proto Input Filter        Output Filter
ge-0/3/0.1      up    up   mpls
Interface       Admin Link Proto Input Policer        Output Policer
ge-0/3/0.1      up    up
                           mpls

  Logical interface: ge-0/3/0.1, Index: 69
    Object                 Name                  Type                    Index

    Classifier             ipprec-compatibility   ip                        13

Classifier: ipprec-compatibility, Code point type: inet-precedence, Index: 13
  Code point         Forwarding class                    Loss priority
  000                af3                                 low
  001                af3                                 high
  010                af3                                 low
  011                af3                                 high
  100                af3                                 low
  101                af3                                 high
  110                ef1                                 low
  111                ef1                                 high
Forwarding class                       ID      Queue  Restricted queue  Fabric
priority  Policing priority
  af3                                 0       0          0             low
          normal
  af2                                 1       1          1             low
          normal
  ef2                                 2       2          2             high
          normal
  ef1                                 3       3          3             high
          normal
```



```
  af1                                  4       4          0              low
            normal
```

show class-of-service interface (ACX Series Routers)

```
user@host-g11# show class-of-service interface
Physical interface: at-0/0/0, Index: 130
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

  Logical interface: at-0/0/0.0, Index: 69

  Logical interface: at-0/0/0.32767, Index: 70

Physical interface: at-0/0/1, Index: 133
Maximum usable queues: 4, Queues in use: 4
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled

  Logical interface: at-0/0/1.0, Index: 71

  Logical interface: at-0/0/1.32767, Index: 72

Physical interface: ge-0/1/0, Index: 146
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Rewrite                 dscp-default          dscp                       31
Classifier              d1                    dscp                    11331
Classifier              ci                    ieee8021p                 583

  Logical interface: ge-0/1/0.0, Index: 73
Object                  Name                  Type                    Index
Rewrite                 custom-exp            exp (mpls-any)          46413

  Logical interface: ge-0/1/0.1, Index: 74

  Logical interface: ge-0/1/0.32767, Index: 75

Physical interface: ge-0/1/1, Index: 147
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Classifier              ipprec-compatibility  ip                         13

  Logical interface: ge-0/1/1.0, Index: 76

Physical interface: ge-0/1/2, Index: 148
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                  Name                  Type                    Index
Rewrite                 ri                    ieee8021p (outer)       35392
Classifier              ci                    ieee8021p                 583

Physical interface: ge-0/1/3, Index: 149
```



```
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

  Logical interface: ge-0/1/3.0, Index: 77
Object                 Name                   Type                    Index
Rewrite                custom-exp2            exp (mpls-any)          53581

Physical interface: ge-0/1/4, Index: 150
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/5, Index: 151
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/6, Index: 152
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/1/7, Index: 153
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             d1                     dscp                    11331

Physical interface: ge-0/2/0, Index: 154
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

Physical interface: ge-0/2/1, Index: 155
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
Classifier             ipprec-compatibility   ip                         13

  Logical interface: ge-0/2/1.0, Index: 78

  Logical interface: ge-0/2/1.32767, Index: 79

Physical interface: xe-0/3/0, Index: 156
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                 Name                   Type                    Index
```

```
Classifier            ipprec-compatibility   ip                       13

  Logical interface: xe-0/3/0.0, Index: 80

Physical interface: xe-0/3/1, Index: 157
Maximum usable queues: 8, Queues in use: 5
  Scheduler map: <default>, Index: 2
  Congestion-notification: Disabled
Object                Name                   Type                   Index
Classifier            ipprec-compatibility   ip                       13

  Logical interface: xe-0/3/1.0, Index: 81

[edit]
user@host-g11#
```

show class-of-service interface (PPPoE Subscriber Interface for Enhanced Subscriber Management)

```
user@host> show class-of-service interface pp0.3221225474
  Logical interface: pp0.3221225475, Index: 3221225475
Object                Name                   Type                   Index
Traffic-control-profile TC_PROF_100_199_SERIES_UID1006 Output      4294967312
Scheduler-map         SMAP-1_UID1002         Output                4294967327
Rewrite-Output        ieee-rewrite           ieee8021p              60432
Rewrite-Output        rule1                  ip                     50463

    Adjusting application: PPPoE IA tags
      Adjustment type: absolute
      Configured shaping rate: 11000000
      Adjustment value: 5000000
      Adjustment target: node

    Adjusting application: ucac
      Adjustment type: delta
      Configured shaping rate: 5000000
      Adjustment value: 100000
      Adjustment target: node
```

show class-of-service scheduler-map

Syntax

show class-of-service scheduler-map
<*name*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 15.1R3 on MX Series routers for enhanced subscriber management.

Description

Display the mapping of schedulers to forwarding classes and a summary of scheduler parameters for each entry.

Options

none—Display all scheduler maps.

name—(Optional) Display a summary of scheduler parameters for each forwarding class to which the named scheduler is assigned.

Required Privilege Level

view

Related Documentation

- *Verifying and Managing Junos OS Enhanced Subscriber Management*

List of Sample Output

show class-of-service scheduler-map on page 475
show class-of-service scheduler-map (QFX Series) on page 476

Output Fields

Table 87 on page 474 describes the output fields for the **show class-of-service scheduler-map** command. Output fields are listed in the approximate order in which they appear.

Table 87: show class-of-service scheduler-map Output Fields

Field Name	Field Description
Scheduler map	Name of the scheduler map. (Enhanced subscriber management for MX Series routers) The name of the dynamic scheduler map object is associated with a generated UID (for example, **SMAP-1_UID1002**) instead of with a subscriber interface.
Index	Index of the indicated object. Objects having indexes in this output include scheduler maps, schedulers, and drop profiles. (Enhanced subscriber management for MX Series routers) Index values for dynamic CoS traffic control profiles are larger for enhanced subscriber management than they are for legacy subscriber management.
Scheduler	Name of the scheduler.

Table 87: show class-of-service scheduler-map Output Fields (continued)

Field Name	Field Description
Forwarding class	Classification of a packet affecting the forwarding, scheduling, and marking policies applied as the packet transits the router.
Transmit rate	Configured transmit rate of the scheduler (in bps). The rate is a percentage of the total interface bandwidth, or the keyword **remainder**, which indicates that the scheduler receives the remaining bandwidth of the interface.
Rate Limit	Rate limiting configuration of the queue. Possible values are **none**, meaning no rate limiting, and **exact**, meaning the queue only transmits at the configured rate.
Maximum buffer delay	Amount of transmit delay (in milliseconds) or the buffer size of the queue. The buffer size is shown as a percentage of the total interface buffer allocation, or by the keyword **remainder** to indicate that the buffer is sized according to what remains after other scheduler buffer allocations.
Priority	Scheduling priority: **low** or **high**.
Excess priority	Priority of excess bandwidth: **low**, **medium-low**, **medium-high**, **high**, or **none**.
Explicit Congestion Notification	(QFX Series, OCX Series, and EX4600 switches only) Explicit congestion notification (ECN) state: • Disable—ECN is disabled on the specified scheduler • Enable—ECN is enabled on the specified scheduler ECN is disabled by default.
Adjust minimum	Minimum shaping rate for an adjusted queue, in bps.
Adjust percent	Bandwidth adjustment applied to a queue, in percent.
Drop profiles	Table displaying the assignment of drop profiles by name and index to a given loss priority and protocol pair.
Loss priority	Packet loss priority for drop profile assignment.
Protocol	Transport protocol for drop profile assignment.
Name	Name of the drop profile.

Sample Output

show class-of-service scheduler-map

```
user@host> show class-of-service scheduler-map
Scheduler map: dd-scheduler-map, Index: 84

  Scheduler: aa-scheduler, Index: 8721, Forwarding class: aa-forwarding-class
   Transmit rate: 30 percent, Rate Limit: none, Maximum buffer delay: 39 ms,
   Priority: high
   Drop profiles:
    Loss priority   Protocol    Index    Name
```

```
    Low           non-TCP     8724     aa-drop-profile
    Low           TCP         9874     bb-drop-profile
    High          non-TCP     8833     cc-drop-profile
    High          TCP         8484     dd-drop-profile

Scheduler: bb-scheduler, Forwarding class: aa-forwarding-class
 Transmit rate: 40 percent, Rate limit: none, Maximum buffer delay: 68 ms,
 Priority: high
 Drop profiles:
  Loss priority   Protocol    Index    Name
    Low           non-TCP     8724     aa-drop-profile
    Low           TCP         9874     bb-drop-profile
    High          non-TCP     8833     cc-drop-profile
    High          TCP         8484     dd-drop-profile
```

show class-of-service scheduler-map (QFX Series)

```
user@switch# show class-of-service scheduler-map
Scheduler map: be-map, Index: 12240

  Scheduler:be-sched, Forwarding class: best-effort, Index: 115
    Transmit rate: 30 percent, Rate Limit: none, Buffer size: remainder,
    Buffer Limit: none, Priority: low
    Excess Priority: unspecified, Explicit Congestion Notification: disable

    Drop profiles:
      Loss priority   Protocol    Index    Name
      Low             any          3312    lan-dp
      Medium-high     any          2714    be-dp1
      High            any          3178    be-dp2
```


show class-of-service traffic-control-profile

Syntax

show class-of-service traffic-control-profile
<*profile-name*>

Release Information

Command introduced before Junos OS Release 7.4.
Command introduced in Junos OS Release 11.1 for the QFX Series.
Command introduced in Junos OS Release 12.2 for ACX Series Routers.
Command introduced in Junos OS Release 15.1R3 on MX Series routers for enhanced subscriber management.

Description

For Gigabit Ethernet IQ PICs, Channelized IQ PICs, EQ DPCs, and MPC/MIC interfaces only, display traffic shaping and scheduling profiles.

(ACX Series routers) For ATM IMA pseudowire interfaces, display traffic shaping and scheduling profiles.

Options

none—Display all profiles.

profile-name—(Optional) Display information about a single profile.

Required Privilege Level

view

Related Documentation

- *Verifying and Managing Junos OS Enhanced Subscriber Management*

List of Sample Output

show class-of-service traffic-control-profile on page 479
show class-of-service traffic-control-profile (MX Series routers with Clear Channel Multi-Rate CE MIC) on page 480
show class-of-service traffic-control-profile (ACX Series routers with ATM IMA pseudowire interfaces) on page 480
show class-of-service traffic-control-profile (Enhanced Subscriber Management) on page 480

Output Fields

Table 88 on page 477 describes the output fields for the **show class-of-service traffic-control-profile** command. Output fields are listed in the approximate order in which they appear.

Table 88: show class-of-service traffic-control-profile Output Fields

Field Name	Field Description
Traffic control profile	Name of the traffic control profile. (Enhanced subscriber management for MX Series routers) The name of the dynamic traffic control profile object is associated with a generated UID (for example, **TC_PROF_100_199_SERIES_UID1000**) instead of with a subscriber interface.

Table 88: show class-of-service traffic-control-profile Output Fields (continued)

Field Name	Field Description
Index	Index number of the traffic control profile. (Enhanced subscriber management for MX Series routers) Index values for dynamic CoS traffic control profiles are larger for enhanced subscriber management than they are for legacy subscriber management.
ATM Service	(MX Series routers with ATM Multi-Rate CE MIC) Configured category of ATM service. Possible values: • cbr—Constant bit rate. • rtvbr—Real time variable bit rate. • nrtvbr—Non real time variable bit rate. • ubr—Unspecified bit rate.
Maximum Burst Size	Configured maximum burst size, in cells.
Peak rate	Configured peak rate, in cps.
Sustained rate	Configured sustained rate, in cps.
Shaping rate	Configured shaping rate, in bps. NOTE: (MX Series routers with ATM Multi-Rate CE MIC) Configured peak rate, in cps.
Shaping rate burst	Configured burst size for the shaping rate, in bytes. NOTE: (MX Series routers with ATM Multi-Rate CE MIC) Configured maximum burst rate, in cells.
Shaping rate priority high	Configured shaping rate for high-priority traffic, in bps.
Shaping rate priority medium	Configured shaping rate for medium-priority traffic, in bps.
Shaping rate priority low	Configured shaping rate for low-priority traffic, in bps.
Shaping rate excess high	Configured shaping rate for high-priority excess traffic, in bps.
Shaping rate excess low	Configured shaping rate for low-priority excess traffic, in bps.
Scheduler map	Name of the associated scheduler map. (Enhanced subscriber management for MX Series routers) The name of the dynamic scheduler map object is associated with a generated UID (for example, **SMAP-1_UID1002**) instead of with a subscriber interface.
Delay Buffer rate	Configured delay buffer rate, in bps.

Table 88: show class-of-service traffic-control-profile Output Fields (continued)

Field Name	Field Description
Excess rate	Configured excess rate, in percent or proportion.
Excess rate high	Configured excess rate for high priority traffic, in percent or proportion.
Excess rate low	Configured excess rate for low priority traffic, in percent or proportion.
Guaranteed rate	Configured guaranteed rate, in bps or cps. NOTE: (MX Series routers with ATM Multi-Rate CE MIC) This value depends on the ATM service category chosen. Possible values: • **cbr**—Guaranteed rate is equal to the configured peak rate in cps. • **rtvbr**—Guaranteed rate is equal to the configured sustained rate in cps. • **nrtvbr**—Guaranteed rate is equal to the configured sustained rate in cps.
Guaranteed rate burst	Configured burst size for the guaranteed rate, in bytes.
adjust-minimum	Configured minimum shaping rate for an adjusted queue, in bps.
overhead accounting mode	Configured shaping mode: **Frame Mode** or **Cell Mode**.
Overhead bytes	Configured byte adjustment value.
Adjust parent	Configured shaping-rate adjustment for parent scheduler nodes. If enabled, this field appears. **flow-aware** indicates that the parent scheduler node is adjusted only once per multicast channel.

Sample Output

show class-of-service traffic-control-profile

```
user@host> show class-of-service traffic-control-profile
Traffic control profile: Profile1, Index: 57625
  Scheduler map: m1
  Delay Buffer rate: 500000
  Guaranteed rate: 1000000

Traffic control profile: Profile2, Index: 57624
  Scheduler map: m2
  Delay Buffer rate: 600000
  Guaranteed rate: 2000000

Traffic control profile: Profile3, Index: 57627
  Scheduler map: m3
  Delay Buffer rate: 800000
```

```
  Guaranteed rate: 3000000
 .Excess rate high: proportion 4

Traffic control profile: Profile4, Index: 57626
  Scheduler map: m4
  Delay Buffer rate: 750000
  Guaranteed rate: 4000000
..adjust-minimum 20000000

 Traffic control profile: foo, Index: 57626
   Shaping rate: 100000000
   Scheduler map: <default>
   Overhead accounting mode: Frame Mode
   Frame mode overhead accounting bytes: -12
   Adjust parent: flow-aware
```

show class-of-service traffic-control-profile (MX Series routers with Clear Channel Multi-Rate CE MIC)

```
user@host> show class-of-service traffic-control-profile
Traffic control profile: at-vbr1, Index: 11395
  ATM Service: RTVBR
  Scheduler map: m3
  overhead accounting mode: Frame Mode
  Shaping rate: 1000 cps
  Shaping rate burst: 500 cells
  Delay Buffer rate: 2000 cps
  Guaranteed rate: 1000 cps

Traffic control profile: foo, Index: 38286
  ATM Service: UBR
  Scheduler map: m3
  overhead accounting mode: Frame Mode
```

show class-of-service traffic-control-profile (ACX Series routers with ATM IMA pseudowire interfaces)

```
user@host> show class-of-service traffic-control-profile
Traffic control profile: foo, Index: 38286
  ATM Service: RTVBR
  Shaping rate: 2000 cps
  Shaping rate burst: 200 cells
  Scheduler map: <default>
  Delay Buffer rate: 1000 cps
  Guaranteed rate: 1700 cps
```

show class-of-service traffic-control-profile (Enhanced Subscriber Management)

```
user@host> show class-of-service traffic-control-profile
Traffic control profile: TC_PROF_100_199_SERIES_UID1000, Index: 4294967313
  Shaping rate: 11000000
  Shaping rate burst: 1 bytes
  Scheduler map: SMAP-1_UID1002
  Delay Buffer rate: 5000000
  Overhead accounting mode: Cell Mode
  Frame mode overhead accounting bytes: -4
  Cell mode overhead accounting bytes: 20
```

show interfaces queue

Syntax

show interfaces queue
<aggregate | remaining-traffic>
<both-ingress-egress>
<egress>
<forwarding-class *forwarding-class*>
<ingress>
<interface-name *interface-name*>
<l2-statistics>

Release Information

Command introduced before Junos OS Release 7.4.
both-ingress-egress, **egress**, and **ingress** options introduced in Junos OS Release 7.6.
Command introduced in Junos OS Release 11.1 for the QFX Series.
l2-statistics option introduced in Junos OS Release 12.1.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display class-of-service (CoS) queue information for physical interfaces.

Options

none—Show detailed CoS queue statistics for all physical interfaces.

aggregate—(Optional) Display the aggregated queuing statistics of all logical interfaces that have traffic-control profiles configured. (Not on the QFX Series.)

both-ingress-egress—(Optional) On Gigabit Ethernet Intelligent Queuing 2 (IQ2) PICs, display both ingress and egress queue statistics. (Not on the QFX Series.)

egress—(Optional) Display egress queue statistics.

forwarding-class *forwarding-class*—(Optional) Forwarding class name for this queue. Shows detailed CoS statistics for the queue associated with the specified forwarding class.

ingress—(Optional) On Gigabit Ethernet IQ2 PICs, display ingress queue statistics. (Not on the QFX Series.)

interface-name *interface-name*—(Optional) Show detailed CoS queue statistics for the specified interface.

l2-statistics—(Optional) Display Layer 2 statistics for MLPPP, FRF.15, and FRF.16 bundles

remaining-traffic—(Optional) Display the remaining-traffic queue statistics of all logical interfaces that have traffic-control profiles configured.

Overhead for Layer 2 Statistics

Transmitted packets and transmitted byte counts are displayed for the Layer 2 level with the addition of encapsulation overheads applied for fragmentation, as shown in Table 89 on page 482. Others counters, such as packets and bytes queued (input) and drop counters, are displayed at the Layer 3 level. In the case of link fragmentation and interleaving (LFI) for which fragmentation is not applied, corresponding Layer 2 overheads are added, as shown in Table 89 on page 482.

Table 89: Layer 2 Overhead and Transmitted Packets or Byte Counts

Protocol	Fragmentation		LFI
	First fragmentation	Second to *n* fragmentations	
	Bytes	Bytes	
MLPPP (Long)	13	12	8
MLPPP (short)	11	10	8
MLFR (FRF15)	12	10	8
MFR (FRF16)	10	8	-
MCMLPPP(Long)	13	12	-
MCMLPPP(Short)	11	10	-

Layer 2 Statistics—Fragmentation Overhead Calculation

```
MLPPP/MC-MLPPP Overhead details:
   ===============================
   Fragment 1:

    Outer PPP header                        : 4 bytes
    Long or short sequence MLPPP header     : 4 bytes or 2 bytes
    Inner PPP header                        : 1 byte
    HDLC flag and FCS bytes                 : 4 bytes

   Fragments 2 .. n :

    Outer PPP header                        : 4 bytes
    Long or short sequence MLPPP header     : 4 bytes or 2 bytes
    HDLC flag and FCS bytes                 : 4 bytes


   MLFR (FRF15) Overhead details:
   =============================
   Fragment 1:
       Framerelay header      :  2 bytes
       Control,NLPID          :  2 bytes
       Fragmentaion header    :  2 bytes
       Inner proto            :  2 bytes
       HDLC flag and  FCS     :  4 bytes

   Fragments 2 ...n :

       Framerelay header      :  2 bytes
       Control,NLPID          :  2 bytes
       Fragmentaion header    :  2 bytes
       HDLC flag and  FCS     :  4 bytes
```



```
MFR (FRF16) Overhead details:
==============================
Fragment 1:
   Fragmentaion header   :  2 bytes
   Framerelay header     :  2 bytes
   Inner proto           :  2 bytes
   HDLC flag and  FCS    :  4 bytes

Fragments 2 ...n :
   Fragmentaion header   :  2 bytes
   Framerelay header     :  2 bytes
   HDLC flag and  FCS    :  4 bytes
```

Overhead with LFI

```
MLPPP(Long & short sequence):
   ============================
       Outer PPP header      :  4 bytes
       HDLC flag and  FCS    :  4 bytes

  MLFR (FRF15):
  =============
       Framerelay header     :  2 bytes
       Control,NLPID         :  2 bytes
       HDLC flag and  FCS    :  4 bytes
```

The following examples show overhead for different cases:

- A 1000-byte packet is sent to a mlppp bundle without any fragmentation. At the Layer 2 level, bytes transmitted is 1013 in 1 packet. This overhead is for MLPPP long sequence encap.
- A 1000-byte packet is sent to a mlppp bundle with a fragment threshold of 250byte. At the Layer 2 level, bytes transmitted is 1061 bytes in 5 packets.
- A 1000-byte LFI packet is sent to an mlppp bundle. At the Layer 2 level, bytes transmitted is 1008 in 1 packet.

remaining-traffic—(Optional) Display the queuing statistics of all logical interfaces that do not have traffic-control profiles configured. (Not on the QFX Series.)

Additional Information

For rate-limited interfaces hosted on Modular Interface Cards (MICs), Modular Port Concentrators (MPCs), or Enhanced Queuing DPCs, rate-limit packet-drop operations occur *before* packets are queued for transmission scheduling. For such interfaces, the statistics for queued traffic do not include the packets that have already been dropped due to rate limiting, and consequently the displayed statistics for queued traffic are the same as the displayed statistics for transmitted traffic.

NOTE: For rate-limited interfaces hosted on other types of hardware, rate-limit packet-drop operations occur *after* packets are queued for transmission scheduling. For these other interface types, the statistics for queued traffic include the packets that are later dropped due to rate limiting, and consequently the displayed statistics for queued traffic equals the sum of the statistics for transmitted and rate-limited traffic.

On M Series routers (except for the M320 and M120 routers), this command is valid only for a PIC installed on an enhanced Flexible PIC Concentrator (FPC).

Queue statistics for aggregated interfaces are supported on the M Series and T Series routers only. Statistics for an aggregated interface are the summation of the queue statistics of the child links of that aggregated interface. You can view the statistics for a child interface by using the **show interfaces statistics** command for that child interface.

When you configure tricolor marking on a 10-port 1-Gigabit Ethernet PIC, for queues 6 and 7 only, the output does not display the number of queued bytes and packets, or the number of bytes and packets dropped because of RED. If you do not configure tricolor marking on the interface, these statistics are available for all queues.

For the 4-port Channelized OC12 IQE PIC and 1-port Channelized OC48 IQE PIC, the **Packet Forwarding Engine Chassis Queues** field represents traffic bound for a particular physical interface on the PIC. For all other PICs, the **Packet Forwarding Engine Chassis Queues** field represents the total traffic bound for the PIC.

For Gigabit Ethernet IQ2 PICs, the **show interfaces queue** command output does not display the number of tail-dropped packets. This limitation does not apply to Packet Forwarding Engine chassis queues.

When fragmentation occurs on the egress interface, the first set of packet counters shows the postfragmentation values. The second set of packet counters (under the **Packet Forwarding Engine Chassis Queues** field) shows the prefragmentation values.

The behavior of the **egress** queues for the **Routing Engine-Generated Traffic** is not same as the configured queue for MLPPP and MFR configurations.

For related CoS operational mode commands, see the CLI Explorer.

Required Privilege Level

view

List of Sample Output

show interfaces queue (Rate-Limited Interface on a Gigabit Ethernet MIC in an MPC) on page 490
show interfaces queue (Aggregated Ethernet on a T320 Router) on page 491
show interfaces queue (Gigabit Ethernet on a T640 Router) on page 493
show interfaces queue aggregate (Gigabit Ethernet Enhanced DPC) on page 493
show interfaces queue (Gigabit Ethernet IQ2 PIC) on page 497
show interfaces queue both-ingress-egress (Gigabit Ethernet IQ2 PIC) on page 500

show interfaces queue ingress (Gigabit Ethernet IQ2 PIC) on page 502
show interfaces queue egress (Gigabit Ethernet IQ2 PIC) on page 503
show interfaces queue remaining-traffic (Gigabit Ethernet Enhanced DPC) on page 505
show interfaces queue (Channelized OC12 IQE Type 3 PIC in SONET Mode) on page 507
show interfaces queue (QFX Series) on page 517
show interfaces queue l2-statistics (lsq interface) on page 518
show interfaces queue lsq (lsq-ifd) on page 519
show interfaces queue (Aggregated Ethernet on a MX series Router) on page 520

Output Fields Table 90 on page 485 lists the output fields for the **show interfaces queue** command. Output fields are listed in the approximate order in which they appear.

Table 90: show interfaces queue Output Fields

Field Name	Field Description
Physical interface	Name of the physical interface.
Enabled	State of the interface. Possible values are described in the “Enabled Field” section under *Common Output Fields Description*.
Interface index	Physical interface's index number, which reflects its initialization sequence.
SNMP ifIndex	SNMP index number for the interface.
Forwarding classes supported	Total number of forwarding classes supported on the specified interface.
Forwarding classes in use	Total number of forwarding classes in use on the specified interface.
Ingress queues supported	On Gigabit Ethernet IQ2 PICs only, total number of ingress queues supported on the specified interface.
Ingress queues in use	On Gigabit Ethernet IQ2 PICs only, total number of ingress queues in use on the specified interface.
Output queues supported	Total number of output queues supported on the specified interface.
Output queues in use	Total number of output queues in use on the specified interface.
Egress queues supported	Total number of egress queues supported on the specified interface.
Egress queues in use	Total number of egress queues in use on the specified interface.

Table 90: show interfaces queue Output Fields (continued)

Field Name	Field Description
Queue counters (Ingress)	CoS queue number and its associated user-configured forwarding class name. Displayed on IQ2 interfaces. • **Queued packets**—Number of queued packets. NOTE: This field is not supported on QFX5100, QFX5110, QFX5200, and QFX5210 switches due to hardware limitations. • **Transmitted packets**—Number of transmitted packets. • **Dropped packets**—Number of packets dropped by the ASIC's RED mechanism.
Burst size	(Logical interfaces on IQ PICs only) Maximum number of bytes up to which the logical interface can burst. The burst size is based on the shaping rate applied to the interface.
The following output fields are applicable to both interface component and Packet Forwarding component in the **show interfaces queue** command:	
Queue	Queue number.
Forwarding classes	Forwarding class name.
Queued Packets	Number of packets queued to this queue. NOTE: For Gigabit Ethernet IQ2 interfaces, the Queued Packets count is calculated by the Junos OS interpreting one frame buffer as one packet. If the queued packets are very large or very small, the calculation might not be completely accurate for transit traffic. The count is completely accurate for traffic terminated on the router. For rate-limited interfaces hosted on MICs or MPCs only, this statistic does not include traffic dropped due to rate limiting. For more information, see “Additional Information” on page 483. NOTE: This field is not supported on QFX5100, QFX5110, QFX5200, and QFX5210 switches due to hardware limitations.
Queued Bytes	Number of bytes queued to this queue. The byte counts vary by interface hardware. For more information, see Table 91 on page 489. For rate-limited interfaces hosted on MICs or MPCs only, this statistic does not include traffic dropped due to rate limiting. For more information, see “Additional Information” on page 483. NOTE: This field is not supported on QFX5100, QFX5110, QFX5200, and QFX5210 switches due to hardware limitations.
Transmitted Packets	Number of packets transmitted by this queue. When fragmentation occurs on the egress interface, the first set of packet counters shows the postfragmentation values. The second set of packet counters (displayed under the **Packet Forwarding Engine Chassis Queues** field) shows the prefragmentation values. NOTE: For Layer 2 statistics, see “Overhead for Layer 2 Statistics” on page 481

Table 90: show interfaces queue Output Fields (continued)

Field Name	Field Description
Transmitted Bytes	Number of bytes transmitted by this queue. The byte counts vary by interface hardware. For more information, see Table 91 on page 489. NOTE: On MX Series routers, this number can be inaccurate when you issue the command for a physical interface repeatedly and in quick succession, because the statistics for the child nodes are collected infrequently. Wait ten seconds between successive iterations to avoid this situation. NOTE: For Layer 2 statistics, see "Overhead for Layer 2 Statistics" on page 481
Tail-dropped packets	Number of packets dropped because of tail drop. NOTE: Starting with Junos OS 18.3R1, the **Tail-dropped packets** counter is supported on PTX Series Packet Transport Routers.
RL-dropped packets	Number of packets dropped due to rate limiting. For rate-limited interfaces hosted on MICs, MPCs, and Enhanced Queuing DPCs only, this statistic is not included in the queued traffic statistics. For more information, see "Additional Information" on page 483. NOTE: The **RL-dropped packets** counter is not supported on the PTX Series Packet Transport Routers, and is omitted from the output.
RL-dropped bytes	Number of bytes dropped due to rate limiting. For rate-limited interfaces hosted on MICs, MPCs, and Enhanced Queuing DPCs only, this statistic is not included in the queued traffic statistics. For more information, see "Additional Information" on page 483.

RED-dropped packets	Number of packets dropped because of random early detection (RED).
	
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, the total number of dropped packets is displayed. On all other M Series routers, the output classifies dropped packets into the following categories:
	  • **Low, non-TCP**—Number of low-loss priority non-TCP packets dropped because of RED.
	  • **Low, TCP**—Number of low-loss priority TCP packets dropped because of RED.
	  • **High, non-TCP**—Number of high-loss priority non-TCP packets dropped because of RED.
	  • **High, TCP**—Number of high-loss priority TCP packets dropped because of RED.
	
	• (MX Series routers with enhanced DPCs, and T Series routers with enhanced FPCs only) The output classifies dropped packets into the following categories:
	  • **Low**—Number of low-loss priority packets dropped because of RED.
	  • **Medium-low**—Number of medium-low loss priority packets dropped because of RED.
	  • **Medium-high**—Number of medium-high loss priority packets dropped because of RED.
	  • **High**—Number of high-loss priority packets dropped because of RED.
	
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

Table 90: show interfaces queue Output Fields (continued)

Field Name	Field Description

RED-dropped bytes	Number of bytes dropped because of RED. The byte counts vary by interface hardware. For more information, see Table 91 on page 489.
	• (M Series and T Series routers only) On M320 and M120 routers and the T Series routers, only the total number of dropped bytes is displayed. On all other M Series routers, the output classifies dropped bytes into the following categories:
	  • **Low, non-TCP**—Number of low-loss priority non-TCP bytes dropped because of RED.
	  • **Low, TCP**—Number of low-loss priority TCP bytes dropped because of RED.
	  • **High, non-TCP**—Number of high-loss priority non-TCP bytes dropped because of RED.
	  • **High, TCP**—Number of high-loss priority TCP bytes dropped because of RED.
	NOTE: Due to accounting space limitations on certain Type 3 FPCs (which are supported in M320 and T640 routers), this field does not always display the correct value for queue 6 or queue 7 for interfaces on 10-port 1-Gigabit Ethernet PICs.

Queue-depth bytes	Displays queue-depth average, current, peak, and maximum values for RTP queues. Because queue-depth values cannot be aggregated, displays the values for RTP queues regardless of whether **aggregate**, **remaining-traffic**, or neither option is selected.
Queue-depth bytes	Displays queue-depth average, current, peak, and maximum values for RTP queues. Because queue-depth values cannot be aggregated, displays the values for RTP queues regardless of whether **aggregate**, **remaining-traffic**, or neither option is selected.
Last-packet enqueued	Starting with Junos OS Release 16.1, **Last-packet enqueued** output field is introduced. If **packet-timestamp** is enabled for an FPC, shows the day, date, time, and year in the format *day-of-the-week month day-date hh:mm:ss yyyy* when a packet was enqueued in the CoS queue. When the timestamp is aggregated across all active Packet Forwarding Engines, the latest timestamp for each CoS queue is reported.

Byte counts vary by interface hardware. Table 91 on page 489 shows how the byte counts on the outbound interfaces vary depending on the interface hardware. Table 91 on page 489 is based on the assumption that outbound interfaces are sending IP traffic with 478 bytes per packet.

Table 91: Byte Count by Interface Hardware

Interface Hardware	Output Level	Byte Count Includes	Comments
Gigabit Ethernet IQ and IQE PICs	Interface	Queued: 490 bytes per packet, representing 478 bytes of Layer 3 packet + 12 bytes Transmitted: 490 bytes per packet, representing 478 bytes of Layer 3 packet + 12 bytes RED dropped: 496 bytes per packet representing 478 bytes of Layer 3 packet + 18 bytes	The 12 additional bytes include 6 bytes for the destination MAC address + 4 bytes for the VLAN + 2 bytes for the Ethernet type. For RED dropped, 6 bytes are added for the source MAC address.
	Packet forwarding component	Queued: 478 bytes per packet, representing 478 bytes of Layer 3 packet Transmitted: 478 bytes per packet, representing 478 bytes of Layer 3 packet	–
Non-IQ PIC	Interface	T Series, TX Series, T1600, and MX Series routers: • Queued: 478 bytes of Layer 3 packet. • Transmitted: 478 bytes of Layer 3 packet. T4000 routers with Type 5 FPCs : • Queued: 478 bytes of Layer 3 packet + the full Layer 2 overhead including 4 bytes CRC + the full Layer 1 overhead 8 bytes preamble + 12 bytes Inter frame Gap. • Transmitted: 478 bytes of Layer 3 packet + the full Layer 2 overhead including 4 bytes CRC + the full Layer 1 overhead 8 bytes preamble + 12 bytes Interframe Gap. M Series routers: • Queued: 478 bytes of Layer 3 packet. • Transmitted: 478 bytes of Layer 3 packet + the full Layer 2 overhead. PTX Series Packet Transport Routers: • Queued: The sum of the transmitted bytes and the RED dropped bytes. • Transmitted: Full Layer 2 overhead (including all L2 encapsulation and CRC) + 12 inter-packet gap + 8 for the preamble. • RED dropped: Full Layer 2 overhead (including all L2 encapsulation and CRC) + 12 inter-packet gap + 8 for the preamble (does not include the VLAN header or MPLS pushed bytes).	The Layer 2 overhead is 14 bytes for non-VLAN traffic and 18 bytes for VLAN traffic.

Table 91: Byte Count by Interface Hardware (continued)

Interface Hardware	Output Level	Byte Count Includes	Comments
IQ and IQE PICs with a SONET/SDH interface	Interface	Queued: 482 bytes per packet, representing 478 bytes of Layer 3 packet + 4 bytes Transmitted: 482 bytes per packet, representing 478 bytes of Layer 3 packet + 4 bytes RED dropped: 482 bytes per packet, representing 478 bytes of Layer 3 packet + 4 bytes	The additional 4 bytes are for the Layer 2 Point-to-Point Protocol (PPP) header.
	Packet forwarding component	Queued: 478 bytes per packet, representing 478 bytes of Layer 3 packet Transmitted: 486 bytes per packet, representing 478 bytes of Layer 3 packet + 8 bytes	For transmitted packets, the additional 8 bytes includes 4 bytes for the PPP header and 4 bytes for a cookie.
Non-IQ PIC with a SONET/SDH interface	Interface	T Series, TX Series, T1600, and MX Series routers: • Queued: 478 bytes of Layer 3 packet. • Transmitted: 478 bytes of Layer 3 packet. M Series routers: • Queued: 478 bytes of Layer 3 packet. • Transmitted: 483 bytes per packet, representing 478 bytes of Layer 3 packet + 5 bytes • RED dropped: 478 bytes per packet, representing 478 bytes of Layer 3 packet	For transmitted packets, the additional 5 bytes includes 4 bytes for the PPP header and 1 byte for the packet loss priority (PLP).
Interfaces configured with Frame Relay Encapsulation	Interface	The default Frame Relay overhead is 7 bytes. If you configure the Frame Check Sequence (FCS) to 4 bytes, then the overhead increases to 10 bytes.	
1-port 10-Gigabit Ethernet IQ2 and IQ2–E PICs 4-port 1G IQ2 and IQ2-E PICs 8-port 1G IQ2 and IQ2-E PICs	Interface	Queued: 478 bytes of Layer 3 packet + the full Layer 2 overhead including CRC. Transmitted: 478 bytes of Layer 3 packet + the full Layer 2 overhead including CRC.	The Layer 2 overhead is 18 bytes for non-VLAN traffic and 22 bytes for VLAN traffic.
	Packet forwarding component	Queued: 478 bytes of Layer 3 packet. Transmitted: 478 bytes of Layer 3 packet.	–

Sample Output

show interfaces queue (Rate-Limited Interface on a Gigabit Ethernet MIC in an MPC)

The following example shows queue information for the rate-limited interface ge-4/2/0 on a Gigabit Ethernet MIC in an MPC. For rate-limited queues for interfaces hosted on MICs or MPCs, rate-limit packet drops occur prior to packet output queuing. In the command output, the nonzero statistics displayed in the **RL-dropped packets** and **RL-dropped bytes** fields quantify the traffic dropped to rate-limit queue 0 output to 10 percent of 1 gigabyte (100 megabits) per second. Because the RL-dropped traffic is not included in the **Queued** statistics, the statistics displayed for queued traffic are the same as the statistics for transmitted traffic.

```
user@host> show interfaces queue ge-4/2/0
Physical interface: ge-4/2/0, Enabled, Physical link is Up
  Interface index: 203, SNMP ifIndex: 1054
Forwarding classes: 16 supported, 4 in use
Egress queues: 8 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :             131300649                141751 pps
    Bytes                :           11287964840              99793248 bps
  Transmitted:
    Packets              :             131300649                141751 pps
    Bytes                :           11287964840              99793248 bps
    Tail-dropped packets :                     0                     0 pps
    RL-dropped packets   :             205050862                602295 pps
    RL-dropped bytes     :           13595326612             327648832 bps
    RED-dropped packets  :                     0                     0 pps
     Low                 :                     0                     0 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
     Low                 :                     0                     0 bps
     Medium-low          :                     0                     0 bps
     Medium-high         :                     0                     0 bps
     High                :                     0                     0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
```

show interfaces queue (Aggregated Ethernet on a T320 Router)

The following example shows that the aggregated Ethernet interface, **ae1**, has traffic on queues **af1** and **af12**:

```
user@host> show interfaces queue ae1
Physical interface: ae1, Enabled, Physical link is Up
Interface index: 158, SNMP ifIndex: 33 Forwarding classes: 8 supported, 8 in use
Output queues: 8 supported, 8 in use
Queue: 0, Forwarding classes: be
  Queued:
    Packets              :                     5                     0 pps
    Bytes                :                   242                     0 bps
  Transmitted:
    Packets              :                     5                     0 pps
    Bytes                :                   242                     0 bps
    Tail-dropped packets :                     0                     0 pps
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: af1
  Queued:
    Packets              :              42603765                595484 pps
```

```
    Bytes              :          5453281920           609776496 bps
  Transmitted:
    Packets            :            42603765              595484 pps
    Bytes              :          5453281920           609776496 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 2, Forwarding classes: ef1
  Queued:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
  Transmitted:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 3, Forwarding classes: nc
  Queued:
    Packets            :                  45                   0 pps
    Bytes              :                3930                   0 bps
  Transmitted:
    Packets            :                  45                   0 pps
    Bytes              :                3930                   0 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 4, Forwarding classes: af11
  Queued:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
  Transmitted:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 5, Forwarding classes: ef11
  Queued:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
  Transmitted:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 6, Forwarding classes: af12
  Queued:
    Packets            :            31296413              437436 pps
    Bytes              :          4005940864           447935200 bps
  Transmitted:
    Packets            :            31296413              437436 pps
    Bytes              :          4005940864           447935200 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 7, Forwarding classes: nc2
  Queued:
    Packets            :                   0                   0 pps
    Bytes              :                   0                   0 bps
```



```
 Transmitted:
   Packets              :                   0                    0 pps
   Bytes                :                   0                    0 bps
   Tail-dropped packets :                   0                    0 pps
   RED-dropped packets  :                   0                    0 pps
   RED-dropped bytes    :                   0                    0 bps
```

show interfaces queue (Gigabit Ethernet on a T640 Router)

```
user@host> show interfaces queue
Physical interface: ge-7/0/1, Enabled, Physical link is Up
   Interface index: 150, SNMP ifIndex: 42
 Forwarding classes: 8 supported, 8 in use
 Output queues: 8 supported, 8 in use
 Queue: 0, Forwarding classes: be
   Queued:
     Packets              :                  13                    0 pps
     Bytes                :                 622                    0 bps
   Transmitted:
     Packets              :                  13                    0 pps
     Bytes                :                 622                    0 bps
     Tail-dropped packets :                   0                    0 pps
     RED-dropped packets  :                   0                    0 pps
     RED-dropped bytes    :                   0                    0 bps
 Queue: 1, Forwarding classes: af1
   Queued:
     Packets              :          1725947945               372178 pps
     Bytes                :        220921336960            381110432 bps
   Transmitted:
     Packets              :          1725947945               372178 pps
     Bytes                :        220921336960            381110432 bps
     Tail-dropped packets :                   0                    0 pps
     RED-dropped packets  :                   0                    0 pps
     RED-dropped bytes    :                   0                    0 bps
 Queue: 2, Forwarding classes: ef1
   Queued:
     Packets              :                   0                    0 pps
     Bytes                :                   0                    0 bps
   Transmitted:
     Packets              :                   0                    0 pps
     Bytes                :                   0                    0 bps
     Tail-dropped packets :                   0                    0 pps
     RED-dropped packets  :                   0                    0 pps
     RED-dropped bytes    :                   0                    0 bps
 Queue: 3, Forwarding classes: nc
   Queued:
     Packets              :                 571                    0 pps
     Bytes                :               49318                  336 bps
   Transmitted:
     Packets              :                 571                    0 pps
     Bytes                :               49318                  336 bps
     Tail-dropped packets :                   0                    0 pps
     RED-dropped packets  :                   0                    0 pps
     RED-dropped bytes    :                   0                    0 bps
```

show interfaces queue aggregate (Gigabit Ethernet Enhanced DPC)

```
user@host> show interfaces queue ge-2/2/9 aggregate
```

```
Physical interface: ge-2/2/9, Enabled, Physical link is Up
  Interface index: 238, SNMP ifIndex: 71
Forwarding classes: 16 supported, 4 in use
Ingress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :             148450735                947295 pps
    Bytes                :            8016344944             409228848 bps
  Transmitted:
    Packets              :              76397439                487512 pps
    Bytes                :            4125461868             210602376 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :              72053285                459783 pps
     Low                 :              72053285                459783 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :            3890877444             198626472 bps
     Low                 :            3890877444             198626472 bps
     Medium-low          :                     0                     0 bps
     Medium-high         :                     0                     0 bps
     High                :                     0                     0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
     Low                 :                     0                     0 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
     Low                 :                     0                     0 bps
     Medium-low          :                     0                     0 bps
     Medium-high         :                     0                     0 bps
     High                :                     0                     0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets              :             410278257                473940 pps
    Bytes                :           22156199518             204742296 bps
  Transmitted:
    Packets              :               4850003                  4033 pps
    Bytes                :             261900162               1742256 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :             405425693                469907 pps
     Low                 :             405425693                469907 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :           21892988124             203000040 bps
     Low                 :           21892988124             203000040 bps
     Medium-low          :                     0                     0 bps
     Medium-high         :                     0                     0 bps
     High                :                     0                     0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets              :                     0                     0 pps
```



```
      Bytes              :                0                    0 bps
    Transmitted:
      Packets            :                0                    0 pps
      Bytes              :                0                    0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :              0                    0 pps
       Low               :                0                    0 pps
       Medium-low        :                0                    0 pps
       Medium-high       :                0                    0 pps
       High              :                0                    0 pps
      RED-dropped bytes  :                0                    0 bps
       Low               :                0                    0 bps
       Medium-low        :                0                    0 bps
       Medium-high       :                0                    0 bps
       High              :                0                    0 bps
  Forwarding classes: 16 supported, 4 in use
  Egress queues: 4 supported, 4 in use
  Queue: 0, Forwarding classes: best-effort
    Queued:
      Packets            :         76605230               485376 pps
      Bytes              :       5209211400            264044560 bps
    Transmitted:
      Packets            :         76444631               484336 pps
      Bytes              :       5198235612            263478800 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :         160475                 1040 pps
       Low               :           160475                 1040 pps
       Medium-low        :                0                    0 pps
       Medium-high       :                0                    0 pps
       High              :                0                    0 pps
      RED-dropped bytes  :         10912300               565760 bps
       Low               :         10912300               565760 bps
       Medium-low        :                0                    0 bps
       Medium-high       :                0                    0 bps
       High              :                0                    0 bps
  Queue: 1, Forwarding classes: expedited-forwarding
    Queued:
      Packets            :                0                    0 pps
      Bytes              :                0                    0 bps
    Transmitted:
      Packets            :                0                    0 pps
      Bytes              :                0                    0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :              0                    0 pps
       Low               :                0                    0 pps
       Medium-low        :                0                    0 pps
       Medium-high       :                0                    0 pps
       High              :                0                    0 pps
      RED-dropped bytes  :                0                    0 bps
       Low               :                0                    0 bps
       Medium-low        :                0                    0 bps
       Medium-high       :                0                    0 bps
       High              :                0                    0 bps
  Queue: 2, Forwarding classes: assured-forwarding
    Queued:
      Packets            :          4836136                 3912 pps
      Bytes              :        333402032              2139056 bps
    Transmitted:
      Packets            :          3600866                 1459 pps
      Bytes              :        244858888               793696 bps
      Tail-dropped packets : Not Available
```

```
      RED-dropped packets  :             1225034                 2450 pps
       Low                 :             1225034                 2450 pps
       Medium-low          :                   0                    0 pps
       Medium-high         :                   0                    0 pps
       High                :                   0                    0 pps
      RED-dropped bytes    :            83302312              1333072 bps
       Low                 :            83302312              1333072 bps
       Medium-low          :                   0                    0 bps
       Medium-high         :                   0                    0 bps
       High                :                   0                    0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets              :                   0                    0 pps
      Bytes                :                   0                    0 bps
    Transmitted:
      Packets              :                   0                    0 pps
      Bytes                :                   0                    0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                   0                    0 pps
       Low                 :                   0                    0 pps
       Medium-low          :                   0                    0 pps
       Medium-high         :                   0                    0 pps
       High                :                   0                    0 pps
      RED-dropped bytes    :                   0                    0 bps
       Low                 :                   0                    0 bps
       Medium-low          :                   0                    0 bps
       Medium-high         :                   0                    0 bps
       High                :                   0                    0 bps

  Packet Forwarding Engine Chassis Queues:
  Queues: 4 supported, 4 in use
  Queue: 0, Forwarding classes: best-effort
    Queued:
      Packets              :            77059796               486384 pps
      Bytes                :          3544750624            178989576 bps
    Transmitted:
      Packets              :            77059797               486381 pps
      Bytes                :          3544750670            178988248 bps
      Tail-dropped packets :                   0                    0 pps
      RED-dropped packets  :                   0                    0 pps
       Low                 :                   0                    0 pps
       Medium-low          :                   0                    0 pps
       Medium-high         :                   0                    0 pps
       High                :                   0                    0 pps
      RED-dropped bytes    :                   0                    0 bps
       Low                 :                   0                    0 bps
       Medium-low          :                   0                    0 bps
       Medium-high         :                   0                    0 bps
       High                :                   0                    0 bps
  Queue: 1, Forwarding classes: expedited-forwarding
    Queued:
      Packets              :                   0                    0 pps
      Bytes                :                   0                    0 bps
    Transmitted:
      Packets              :                   0                    0 pps
      Bytes                :                   0                    0 bps
      Tail-dropped packets :                   0                    0 pps
      RED-dropped packets  :                   0                    0 pps
       Low                 :                   0                    0 pps
       Medium-low          :                   0                    0 pps
       Medium-high         :                   0                    0 pps
```



```
       High               :                   0                    0 pps
      RED-dropped bytes   :                   0                    0 bps
       Low                :                   0                    0 bps
       Medium-low         :                   0                    0 bps
       Medium-high        :                   0                    0 bps
       High               :                   0                    0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets               :             4846580                 3934 pps
    Bytes                 :           222942680              1447768 bps
  Transmitted:
    Packets               :             4846580                 3934 pps
    Bytes                 :           222942680              1447768 bps
    Tail-dropped packets  :                   0                    0 pps
    RED-dropped packets   :                   0                    0 pps
     Low                  :                   0                    0 pps
     Medium-low           :                   0                    0 pps
     Medium-high          :                   0                    0 pps
     High                 :                   0                    0 pps
    RED-dropped bytes     :                   0                    0 bps
     Low                  :                   0                    0 bps
     Medium-low           :                   0                    0 bps
     Medium-high          :                   0                    0 bps
     High                 :                   0                    0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets               :                   0                    0 pps
    Bytes                 :                   0                    0 bps
  Transmitted:
    Packets               :                   0                    0 pps
    Bytes                 :                   0                    0 bps
    Tail-dropped packets  :                   0                    0 pps
    RED-dropped packets   :                   0                    0 pps
     Low                  :                   0                    0 pps
     Medium-low           :                   0                    0 pps
     Medium-high          :                   0                    0 pps
     High                 :                   0                    0 pps
    RED-dropped bytes     :                   0                    0 bps
     Low                  :                   0                    0 bps
     Medium-low           :                   0                    0 bps
     Medium-high          :                   0                    0 bps
     High                 :                   0                    0 bps
```

show interfaces queue (Gigabit Ethernet IQ2 PIC)

```
user@host> show interfaces queue ge-7/1/3
Physical interface: ge-7/1/3, Enabled, Physical link is Up
  Interface index: 170, SNMP ifIndex: 70 Forwarding classes: 16 supported, 4 in
use Ingress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets               :           418390039                   10 pps
    Bytes                 :         38910269752                 7440 bps
  Transmitted:
    Packets               :           418390039                   10 pps
    Bytes                 :         38910269752                 7440 bps
    Tail-dropped packets  : Not Available
    RED-dropped packets   :                   0                    0 pps
    RED-dropped bytes     :                   0                    0 bps
Queue: 1, Forwarding classes: expedited-forwarding
```

```
    Queued:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
    Transmitted:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                  0                    0 pps
      RED-dropped bytes    :                  0                    0 bps
  Queue: 2, Forwarding classes: assured-forwarding
    Queued:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
    Transmitted:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                  0                    0 pps
      RED-dropped bytes    :                  0                    0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets           :                  7055                    1 pps
      Bytes             :                451552                  512 bps
    Transmitted:
      Packets           :                  7055                    1 pps
      Bytes             :                451552                  512 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                  0                    0 pps
      RED-dropped bytes    :                  0                    0 bps
  Forwarding classes: 16 supported, 4 in use Egress queues: 4 supported, 4 in use
  Queue: 0, Forwarding classes: best-effort
    Queued:
      Packets           :                  1031                    0 pps
      Bytes             :                143292                    0 bps
    Transmitted:
      Packets           :                  1031                    0 pps
      Bytes             :                143292                    0 bps
      Tail-dropped packets : Not Available
      RL-dropped packets   :                  0                    0 pps
      RL-dropped bytes     :                  0                    0 bps
      RED-dropped packets  :                  0                    0 pps
      RED-dropped bytes    :                  0                    0 bps
  Queue: 1, Forwarding classes: expedited-forwarding
    Queued:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
    Transmitted:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
      Tail-dropped packets : Not Available
      RL-dropped packets   :                  0                    0 pps
      RL-dropped bytes     :                  0                    0 bps
      RED-dropped packets  :                  0                    0 pps
      RED-dropped bytes    :                  0                    0 bps
  Queue: 2, Forwarding classes: assured-forwarding
    Queued:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
    Transmitted:
      Packets           :                     0                    0 pps
      Bytes             :                     0                    0 bps
```



```
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets              :                77009                    11 pps
    Bytes                :              6894286                  7888 bps
  Transmitted:
    Packets              :                77009                    11 pps
    Bytes                :              6894286                  7888 bps
    Tail-dropped packets : Not Available
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps

Packet Forwarding Engine Chassis Queues:
Queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :                 1031                     0 pps
    Bytes                :               147328                     0 bps
  Transmitted:
    Packets              :                 1031                     0 pps
    Bytes                :               147328                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  :                    0                     0 pps
     Low, non-TCP        :                    0                     0 pps
     Low, TCP            :                    0                     0 pps
     High, non-TCP       :                    0                     0 pps
     High, TCP           :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
     Low, non-TCP        :                    0                     0 bps
     Low, TCP            :                    0                     0 bps
     High, non-TCP       :                    0                     0 bps
     High, TCP           :                    0                     0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RED-dropped packets  :                    0                     0 pps
     Low, non-TCP        :                    0                     0 pps
     Low, TCP            :                    0                     0 pps
     High, non-TCP       :                    0                     0 pps
     High, TCP           :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
     Low, non-TCP        :                    0                     0 bps
     Low, TCP            :                    0                     0 bps
     High, non-TCP       :                    0                     0 bps
     High, TCP           :                    0                     0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
```

```
    Packets             :                  0                   0 pps
    Bytes               :                  0                   0 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
     Low, non-TCP        :                 0                   0 pps
     Low, TCP            :                 0                   0 pps
     High, non-TCP       :                 0                   0 pps
     High, TCP           :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
     Low, non-TCP        :                 0                   0 bps
     Low, TCP            :                 0                   0 bps
     High, non-TCP       :                 0                   0 bps
     High, TCP           :                 0                   0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets             :              94386                  12 pps
    Bytes               :           13756799                9568 bps
  Transmitted:
    Packets             :              94386                  12 pps
    Bytes               :           13756799                9568 bps
    Tail-dropped packets :                 0                   0 pps
    RED-dropped packets  :                 0                   0 pps
     Low, non-TCP        :                 0                   0 pps
     Low, TCP            :                 0                   0 pps
     High, non-TCP       :                 0                   0 pps
     High, TCP           :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
     Low, non-TCP        :                 0                   0 bps
     Low, TCP            :                 0                   0 bps
     High, non-TCP       :                 0                   0 bps
     High, TCP           :                 0                   0 bps
```

show interfaces queue both-ingress-egress (Gigabit Ethernet IQ2 PIC)

```
user@host> show interfaces queue ge-6/2/0 both-ingress-egress
Physical interface: ge-6/2/0, Enabled, Physical link is Up
  Interface index: 175, SNMP ifIndex: 121
Forwarding classes: 8 supported, 4 in use
Ingress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets             : Not Available
    Bytes               :                  0                   0 bps
  Transmitted:
    Packets             :                254                   0 pps
    Bytes               :              16274                   0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets             : Not Available
    Bytes               :                  0                   0 bps
  Transmitted:
    Packets             :                  0                   0 pps
    Bytes               :                  0                   0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                 0                   0 pps
    RED-dropped bytes    :                 0                   0 bps
Queue: 2, Forwarding classes: assured-forwarding
```

```
    Queued:
      Packets              : Not Available
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets              : Not Available
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
  Forwarding classes: 8 supported, 4 in use
  Egress queues: 4 supported, 4 in use
  Queue: 0, Forwarding classes: best-effort
    Queued:
      Packets              : Not Available
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     3                     0 pps
      Bytes                :                   126                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
  Queue: 1, Forwarding classes: expedited-forwarding
    Queued:
      Packets              : Not Available
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
  Queue: 2, Forwarding classes: assured-forwarding
    Queued:
      Packets              : Not Available
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets              : Not Available
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
```

```
Packet Forwarding Engine Chassis Queues:
Queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :           80564692                    0 pps
    Bytes                :         3383717100                    0 bps
  Transmitted:
    Packets              :           80564692                    0 pps
    Bytes                :         3383717100                    0 bps
    Tail-dropped packets :                  0                    0 pps
    RED-dropped packets  :                  0                    0 pps
    RED-dropped bytes    :                  0                    0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              :           80564685                    0 pps
    Bytes                :         3383716770                    0 bps
  Transmitted:
    Packets              :           80564685                    0 pps
    Bytes                :         3383716770                    0 bps
    Tail-dropped packets :                  0                    0 pps
    RED-dropped packets  :                  0                    0 pps
    RED-dropped bytes    :                  0                    0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets              :                  0                    0 pps
    Bytes                :                  0                    0 bps
  Transmitted:
    Packets              :                  0                    0 pps
    Bytes                :                  0                    0 bps
    Tail-dropped packets :                  0                    0 pps
    RED-dropped packets  :                  0                    0 pps
    RED-dropped bytes    :                  0                    0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets              :               9397                    0 pps
    Bytes                :            3809052                  232 bps
  Transmitted:
    Packets              :               9397                    0 pps
    Bytes                :            3809052                  232 bps
    Tail-dropped packets :                  0                    0 pps
    RED-dropped packets  :                  0                    0 pps
    RED-dropped bytes    :                  0                    0 bps
```

show interfaces queue ingress (Gigabit Ethernet IQ2 PIC)

```
user@host> show interfaces queue ge-6/2/0 ingress
Physical interface: ge-6/2/0, Enabled, Physical link is Up
  Interface index: 175, SNMP ifIndex: 121
Forwarding classes: 8 supported, 4 in use
Ingress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              : Not Available
    Bytes                :                  0                    0 bps
  Transmitted:
    Packets              :                288                    0 pps
    Bytes                :              18450                    0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                  0                    0 pps
    RED-dropped bytes    :                  0                    0 bps
```



```
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              : Not Available
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets              : Not Available
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets              : Not Available
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
```

show interfaces queue egress (Gigabit Ethernet IQ2 PIC)

```
user@host> show interfaces queue ge-6/2/0 egress
Physical interface: ge-6/2/0, Enabled, Physical link is Up
  Interface index: 175, SNMP ifIndex: 121
Forwarding classes: 8 supported, 4 in use
Egress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              : Not Available
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     3                     0 pps
    Bytes                :                   126                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              : Not Available
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
```

```
      Packets            : Not Available
      Bytes              :                    0                     0 bps
    Transmitted:
      Packets            :                    0                     0 pps
      Bytes              :                    0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                  0                     0 pps
      RED-dropped bytes    :                  0                     0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets            : Not Available
      Bytes              :                    0                     0 bps
    Transmitted:
      Packets            :                    0                     0 pps
      Bytes              :                    0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                  0                     0 pps
      RED-dropped bytes    :                  0                     0 bps
  Packet Forwarding Engine Chassis Queues:
  Queues: 4 supported, 4 in use
  Queue: 0, Forwarding classes: best-effort
    Queued:
      Packets            :             80564692                     0 pps
      Bytes              :           3383717100                     0 bps
    Transmitted:
      Packets            :             80564692                     0 pps
      Bytes              :           3383717100                     0 bps
      Tail-dropped packets :                  0                     0 pps
      RED-dropped packets  :                  0                     0 pps
      RED-dropped bytes    :                  0                     0 bps
  Queue: 1, Forwarding classes: expedited-forwarding
    Queued:
      Packets            :             80564685                     0 pps
      Bytes              :           3383716770                     0 bps
    Transmitted:
      Packets            :             80564685                     0 pps
      Bytes              :           3383716770                     0 bps
      Tail-dropped packets :                  0                     0 pps
      RED-dropped packets  :                  0                     0 pps
      RED-dropped bytes    :                  0                     0 bps
  Queue: 2, Forwarding classes: assured-forwarding
    Queued:
      Packets            :                    0                     0 pps
      Bytes              :                    0                     0 bps
    Transmitted:
      Packets            :                    0                     0 pps
      Bytes              :                    0                     0 bps
      Tail-dropped packets :                  0                     0 pps
      RED-dropped packets  :                  0                     0 pps
      RED-dropped bytes    :                  0                     0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets            :                 9538                     0 pps
      Bytes              :              3819840                     0 bps
    Transmitted:
      Packets            :                 9538                     0 pps
      Bytes              :              3819840                     0 bps
      Tail-dropped packets :                  0                     0 pps
      RED-dropped packets  :                  0                     0 pps
      RED-dropped bytes    :                  0                     0 bps
```


show interfaces queue remaining-traffic (Gigabit Ethernet Enhanced DPC)

```
user@host> show interfaces queue ge-2/2/9 remaining-traffic
Physical interface: ge-2/2/9, Enabled, Physical link is Up
  Interface index: 238, SNMP ifIndex: 71
Forwarding classes: 16 supported, 4 in use
Ingress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :             110208969                472875 pps
    Bytes                :            5951284434             204282000 bps
  Transmitted:
    Packets              :             110208969                472875 pps
    Bytes                :            5951284434             204282000 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
     Low                 :                     0                     0 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
     Low                 :                     0                     0 bps
     Medium-low          :                     0                     0 bps
     Medium-high         :                     0                     0 bps
     High                :                     0                     0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
     Low                 :                     0                     0 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
     Low                 :                     0                     0 bps
     Medium-low          :                     0                     0 bps
     Medium-high         :                     0                     0 bps
     High                :                     0                     0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
  Transmitted:
    Packets              :                     0                     0 pps
    Bytes                :                     0                     0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                     0                     0 pps
     Low                 :                     0                     0 pps
     Medium-low          :                     0                     0 pps
     Medium-high         :                     0                     0 pps
     High                :                     0                     0 pps
    RED-dropped bytes    :                     0                     0 bps
     Low                 :                     0                     0 bps
     Medium-low          :                     0                     0 bps
```

```
      Medium-high        :                    0                    0 bps
      High               :                    0                    0 bps
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
  Transmitted:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                    0                    0 pps
     Low                 :                    0                    0 pps
     Medium-low          :                    0                    0 pps
     Medium-high         :                    0                    0 pps
     High                :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
     Low                 :                    0                    0 bps
     Medium-low          :                    0                    0 bps
     Medium-high         :                    0                    0 bps
     High                :                    0                    0 bps
Forwarding classes: 16 supported, 4 in use
Egress queues: 4 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :            109355853               471736 pps
    Bytes                :           7436199152            256627968 bps
  Transmitted:
    Packets              :            109355852               471736 pps
    Bytes                :           7436198640            256627968 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                    0                    0 pps
     Low                 :                    0                    0 pps
     Medium-low          :                    0                    0 pps
     Medium-high         :                    0                    0 pps
     High                :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
     Low                 :                    0                    0 bps
     Medium-low          :                    0                    0 bps
     Medium-high         :                    0                    0 bps
     High                :                    0                    0 bps
Queue: 1, Forwarding classes: expedited-forwarding
  Queued:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
  Transmitted:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
    Tail-dropped packets : Not Available
    RED-dropped packets  :                    0                    0 pps
     Low                 :                    0                    0 pps
     Medium-low          :                    0                    0 pps
     Medium-high         :                    0                    0 pps
     High                :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
     Low                 :                    0                    0 bps
     Medium-low          :                    0                    0 bps
     Medium-high         :                    0                    0 bps
     High                :                    0                    0 bps
Queue: 2, Forwarding classes: assured-forwarding
  Queued:
    Packets              :                    0                    0 pps
```



```
      Bytes                :                    0                     0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                    0                     0 pps
       Low                 :                    0                     0 pps
       Medium-low          :                    0                     0 pps
       Medium-high         :                    0                     0 pps
       High                :                    0                     0 pps
      RED-dropped bytes    :                    0                     0 bps
       Low                 :                    0                     0 bps
       Medium-low          :                    0                     0 bps
       Medium-high         :                    0                     0 bps
       High                :                    0                     0 bps
  Queue: 3, Forwarding classes: network-control
    Queued:
      Packets              :                    0                     0 pps
      Bytes                :                    0                     0 bps
    Transmitted:
      Packets              :                    0                     0 pps
      Bytes                :                    0                     0 bps
      Tail-dropped packets : Not Available
      RED-dropped packets  :                    0                     0 pps
       Low                 :                    0                     0 pps
       Medium-low          :                    0                     0 pps
       Medium-high         :                    0                     0 pps
       High                :                    0                     0 pps
      RED-dropped bytes    :                    0                     0 bps
       Low                 :                    0                     0 bps
       Medium-low          :                    0                     0 bps
       Medium-high         :                    0                     0 bps
       High                :                    0                     0 bps
```

show interfaces queue (Channelized OC12 IQE Type 3 PIC in SONET Mode)

```
user@host> show interfaces queue t3-1/1/0:7
 Physical interface: t3-1/1/0:7, Enabled, Physical link is Up

   Interface index: 192, SNMP ifIndex: 1948

   Description: full T3 interface connect to 6ce13 t3-3/1/0:7 for FR testing -
Lam

 Forwarding classes: 16 supported, 9 in use

 Egress queues: 8 supported, 8 in use

 Queue: 0, Forwarding classes: DEFAULT

   Queued:

     Packets              :               214886                 13449 pps

     Bytes                :              9884756               5164536 bps

   Transmitted:

     Packets              :               214886                 13449 pps
```

```
      Bytes                :             9884756               5164536 bps
      Tail-dropped packets :                   0                     0 pps
      RED-dropped packets  :                   0                     0 pps
       Low                 :                   0                     0 pps
       Medium-low          :                   0                     0 pps
       Medium-high         :                   0                     0 pps
       High                :                   0                     0 pps
      RED-dropped bytes    :                   0                     0 bps
       Low                 :                   0                     0 bps
       Medium-low          :                   0                     0 bps
       Medium-high         :                   0                     0 bps
       High                :                   0                     0 bps
  Queue: 1, Forwarding classes: REALTIME
    Queued:
      Packets              :                   0                     0 pps
      Bytes                :                   0                     0 bps
    Transmitted:
      Packets              :                   0                     0 pps
      Bytes                :                   0                     0 bps
      Tail-dropped packets :                   0                     0 pps
      RED-dropped packets  :                   0                     0 pps
       Low                 :                   0                     0 pps
       Medium-low          :                   0                     0 pps
       Medium-high         :                   0                     0 pps
       High                :                   0                     0 pps
      RED-dropped bytes    :                   0                     0 bps
       Low                 :                   0                     0 bps
       Medium-low          :                   0                     0 bps
       Medium-high         :                   0                     0 bps
       High                :                   0                     0 bps
  Queue: 2, Forwarding classes: PRIVATE
```

```
    Queued:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
    Transmitted:
      Packets              :                     0                     0 pps
      Bytes                :                     0                     0 bps
      Tail-dropped packets :                     0                     0 pps
      RED-dropped packets  :                     0                     0 pps
       Low                 :                     0                     0 pps
       Medium-low          :                     0                     0 pps
       Medium-high         :                     0                     0 pps
       High                :                     0                     0 pps
      RED-dropped bytes    :                     0                     0 bps
       Low                 :                     0                     0 bps
       Medium-low          :                     0                     0 bps
       Medium-high         :                     0                     0 bps
       High                :                     0                     0 bps
  Queue: 3, Forwarding classes: CONTROL
    Queued:
      Packets              :                    60                     0 pps
      Bytes                :                  4560                     0 bps
    Transmitted:
      Packets              :                    60                     0 pps
      Bytes                :                  4560                     0 bps
      Tail-dropped packets :                     0                     0 pps
      RED-dropped packets  :                     0                     0 pps
       Low                 :                     0                     0 pps
       Medium-low          :                     0                     0 pps
       Medium-high         :                     0                     0 pps
       High                :                     0                     0 pps
```



```
    RED-dropped bytes    :                    0                    0 bps
     Low                 :                    0                    0 bps
     Medium-low          :                    0                    0 bps
     Medium-high         :                    0                    0 bps
     High                :                    0                    0 bps
Queue: 4, Forwarding classes: CLASS_B_OUTPUT
  Queued:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
  Transmitted:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
    Tail-dropped packets :                    0                    0 pps
    RED-dropped packets  :                    0                    0 pps
     Low                 :                    0                    0 pps
     Medium-low          :                    0                    0 pps
     Medium-high         :                    0                    0 pps
     High                :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
     Low                 :                    0                    0 bps
     Medium-low          :                    0                    0 bps
     Medium-high         :                    0                    0 bps
     High                :                    0                    0 bps
Queue: 5, Forwarding classes: CLASS_C_OUTPUT
  Queued:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
  Transmitted:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
    Tail-dropped packets :                    0                    0 pps
```



```
    RED-dropped packets  :                     0                    0 pps
     Low                 :                     0                    0 pps
     Medium-low          :                     0                    0 pps
     Medium-high         :                     0                    0 pps
     High                :                     0                    0 pps
    RED-dropped bytes    :                     0                    0 bps
     Low                 :                     0                    0 bps
     Medium-low          :                     0                    0 bps
     Medium-high         :                     0                    0 bps
     High                :                     0                    0 bps
Queue: 6, Forwarding classes: CLASS_V_OUTPUT
  Queued:
    Packets              :                     0                    0 pps
    Bytes                :                     0                    0 bps
  Transmitted:
    Packets              :                     0                    0 pps
    Bytes                :                     0                    0 bps
    Tail-dropped packets :                     0                    0 pps
    RED-dropped packets  :                     0                    0 pps
     Low                 :                     0                    0 pps
     Medium-low          :                     0                    0 pps
     Medium-high         :                     0                    0 pps
     High                :                     0                    0 pps
    RED-dropped bytes    :                     0                    0 bps
     Low                 :                     0                    0 bps
     Medium-low          :                     0                    0 bps
     Medium-high         :                     0                    0 bps
     High                :                     0                    0 bps
Queue: 7, Forwarding classes: CLASS_S_OUTPUT, GETS
  Queued:
```



```
    Packets            :                   0                    0 pps
    Bytes              :                   0                    0 bps
  Transmitted:
    Packets            :                   0                    0 pps
    Bytes              :                   0                    0 bps
    Tail-dropped packets :                 0                    0 pps
    RED-dropped packets  :                 0                    0 pps
     Low               :                   0                    0 pps
     Medium-low        :                   0                    0 pps
     Medium-high       :                   0                    0 pps
     High              :                   0                    0 pps
    RED-dropped bytes  :                   0                    0 bps
     Low               :                   0                    0 bps
     Medium-low        :                   0                    0 bps
     Medium-high       :                   0                    0 bps
     High              :                   0                    0 bps

Packet Forwarding Engine Chassis Queues:
Queues: 8 supported, 8 in use
Queue: 0, Forwarding classes: DEFAULT
  Queued:
    Packets            :              371365                23620 pps
    Bytes              :            15597330              7936368 bps
  Transmitted:
    Packets            :              371365                23620 pps
    Bytes              :            15597330              7936368 bps
    Tail-dropped packets :                 0                    0 pps
    RED-dropped packets  :                 0                    0 pps
     Low               :                   0                    0 pps
     Medium-low        :                   0                    0 pps
     Medium-high       :                   0                    0 pps
```



```
      High              :              0                   0 pps
     RED-dropped bytes   :              0                   0 bps
      Low               :              0                   0 bps
      Medium-low        :              0                   0 bps
      Medium-high       :              0                   0 bps
      High              :              0                   0 bps
Queue: 1, Forwarding classes: REALTIME
  Queued:
    Packets             :              0                   0 pps
    Bytes               :              0                   0 bps
  Transmitted:
    Packets             :              0                   0 pps
    Bytes               :              0                   0 bps
    Tail-dropped packets :             0                   0 pps
    RED-dropped packets  :             0                   0 pps
     Low                :              0                   0 pps
     Medium-low         :              0                   0 pps
     Medium-high        :              0                   0 pps
     High               :              0                   0 pps
    RED-dropped bytes    :             0                   0 bps
     Low                :              0                   0 bps
     Medium-low         :              0                   0 bps
     Medium-high        :              0                   0 bps
     High               :              0                   0 bps
Queue: 2, Forwarding classes: PRIVATE
  Queued:
    Packets             :              0                   0 pps
    Bytes               :              0                   0 bps
  Transmitted:
    Packets             :              0                   0 pps
```

```
      Bytes                :                    0                    0 bps
      Tail-dropped packets :                    0                    0 pps
      RED-dropped packets  :                    0                    0 pps
       Low                 :                    0                    0 pps
       Medium-low          :                    0                    0 pps
       Medium-high         :                    0                    0 pps
       High                :                    0                    0 pps
      RED-dropped bytes    :                    0                    0 bps
       Low                 :                    0                    0 bps
       Medium-low          :                    0                    0 bps
       Medium-high         :                    0                    0 bps
       High                :                    0                    0 bps
  Queue: 3, Forwarding classes: CONTROL
    Queued:
      Packets              :                32843                    0 pps
      Bytes                :              2641754                   56 bps
    Transmitted:
      Packets              :                32843                    0 pps
      Bytes                :              2641754                   56 bps
      Tail-dropped packets :                    0                    0 pps
      RED-dropped packets  :                    0                    0 pps
       Low                 :                    0                    0 pps
       Medium-low          :                    0                    0 pps
       Medium-high         :                    0                    0 pps
       High                :                    0                    0 pps
      RED-dropped bytes    :                    0                    0 bps
       Low                 :                    0                    0 bps
       Medium-low          :                    0                    0 bps
       Medium-high         :                    0                    0 bps
       High                :                    0                    0 bps
  Queue: 4, Forwarding classes: CLASS_B_OUTPUT
```



```
  Queued:
    Packets              :                 0                 0 pps
    Bytes                :                 0                 0 bps
  Transmitted:
    Packets              :                 0                 0 pps
    Bytes                :                 0                 0 bps
    Tail-dropped packets :                 0                 0 pps
    RED-dropped packets  :                 0                 0 pps
     Low                 :                 0                 0 pps
     Medium-low          :                 0                 0 pps
     Medium-high         :                 0                 0 pps
     High                :                 0                 0 pps
    RED-dropped bytes    :                 0                 0 bps
     Low                 :                 0                 0 bps
     Medium-low          :                 0                 0 bps
     Medium-high         :                 0                 0 bps
     High                :                 0                 0 bps
Queue: 5, Forwarding classes: CLASS_C_OUTPUT
  Queued:
    Packets              :                 0                 0 pps
    Bytes                :                 0                 0 bps
  Transmitted:
    Packets              :                 0                 0 pps
    Bytes                :                 0                 0 bps
    Tail-dropped packets :                 0                 0 pps
    RED-dropped packets  :                 0                 0 pps
     Low                 :                 0                 0 pps
     Medium-low          :                 0                 0 pps
     Medium-high         :                 0                 0 pps
     High                :                 0                 0 pps
```

```
    RED-dropped bytes    :                      0                     0 bps
     Low                 :                      0                     0 bps
     Medium-low          :                      0                     0 bps
     Medium-high         :                      0                     0 bps
     High                :                      0                     0 bps
Queue: 6, Forwarding classes: CLASS_V_OUTPUT
  Queued:
    Packets              :                      0                     0 pps
    Bytes                :                      0                     0 bps
  Transmitted:
    Packets              :                      0                     0 pps
    Bytes                :                      0                     0 bps
    Tail-dropped packets :                      0                     0 pps
    RED-dropped packets  :                      0                     0 pps
     Low                 :                      0                     0 pps
     Medium-low          :                      0                     0 pps
     Medium-high         :                      0                     0 pps
     High                :                      0                     0 pps
    RED-dropped bytes    :                      0                     0 bps
     Low                 :                      0                     0 bps
     Medium-low          :                      0                     0 bps
     Medium-high         :                      0                     0 bps
     High                :                      0                     0 bps
Queue: 7, Forwarding classes: CLASS_S_OUTPUT, GETS
  Queued:
    Packets              :                      0                     0 pps
    Bytes                :                      0                     0 bps
  Transmitted:
    Packets              :                      0                     0 pps
    Bytes                :                      0                     0 bps
    Tail-dropped packets :                      0                     0 pps
```



```
RED-dropped packets  :              0                 0 pps
 Low                 :              0                 0 pps
 Medium-low          :              0                 0 pps
 Medium-high         :              0                 0 pps
 High                :              0                 0 pps
RED-dropped bytes    :              0                 0 bps
 Low                 :              0                 0 bps
 Medium-low          :              0                 0 bps
 Medium-high         :              0                 0 bps
 High                :              0                 0 bps
```

show interfaces queue (QFX Series)

```
user@switch> show interfaces queue xe-0/0/15
   Physical interface: xe-0/0/15, Enabled, Physical link is Up
     Interface index: 49165, SNMP ifIndex: 539
   Forwarding classes: 12 supported, 8 in use
   Egress queues: 12 supported, 8 in use
   Queue: 0, Forwarding classes: best-effort
     Queued:
       Packets              :                     0                     0 pps
       Bytes                :                     0                     0 bps
     Transmitted:
       Packets              :                     0                     0 pps
       Bytes                :                     0                     0 bps
       Tail-dropped packets : Not Available
       Total-dropped packets:                     0                     0 pps
       Total-dropped bytes  :                     0                     0 bps
   Queue: 3, Forwarding classes: fcoe
     Queued:
       Packets              :                     0                     0 pps
       Bytes                :                     0                     0 bps
     Transmitted:
       Packets              :                     0                     0 pps
       Bytes                :                     0                     0 bps
       Tail-dropped packets : Not Available
       Total-dropped packets:                     0                     0 pps
       Total-dropped bytes  :                     0                     0 bps
0 bps
   Queue: 4, Forwarding classes: no-loss
     Queued:
       Packets              :                     0                     0 pps
       Bytes                :                     0                     0 bps
     Transmitted:
       Packets              :                     0                     0 pps
       Bytes                :                     0                     0 bps
       Tail-dropped packets : Not Available
       Total-dropped packets:                     0                     0 pps
       Total-dropped bytes  :                     0                     0 bps
```

```
Queue: 7, Forwarding classes: network-control
  Queued:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
  Transmitted:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
    Tail-dropped packets : Not Available
    Total-dropped packets:                    0                    0 pps
    Total-dropped bytes  :                    0                    0 bps
Queue: 8, Forwarding classes: mcast
  Queued:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
  Transmitted:
    Packets              :                    0                    0 pps
    Bytes                :                    0                    0 bps
    Tail-dropped packets : Not Available
    Total-dropped packets:                    0                    0 pps
    Total-dropped bytes  :                    0                    0 bps
```

show interfaces queue l2-statistics (lsq interface)

```
user@switch> show interfaces queue lsq-2/2/0.2 l2-statistics
Logical interface lsq-2/2/0.2 (Index 69) (SNMP ifIndex 1598)
Forwarding classes: 16 supported, 4 in use
Egress queues: 8 supported, 4 in use
Burst size: 0
Queue: 0, Forwarding classes: be
  Queued:
    Packets              :                    1                    0 pps
    Bytes                :                 1001                    0 bps
  Transmitted:
    Packets              :                    5                     0 pps
    Bytes                :                 1062                    0 bps
    Tail-dropped packets :                    0                    0 pps
    RED-dropped packets  :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
Queue: 1, Forwarding classes: ef
  Queued:
    Packets              :                    1                    0 pps
    Bytes                :                 1500                    0 bps
  Transmitted:
    Packets              :                    6                     0 pps
    Bytes                :                 1573                    0 bps
    Tail-dropped packets :                    0                    0 pps
    RED-dropped packets  :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
Queue: 2, Forwarding classes: af
  Queued:
    Packets              :                    1                    0 pps
    Bytes                :                  512                    0 bps
  Transmitted:
    Packets              :                    3                     0 pps
    Bytes                :                  549                    0 bps
    Tail-dropped packets :                    0                    0 pps
    RED-dropped packets  :                    0                    0 pps
    RED-dropped bytes    :                    0                    0 bps
Queue: 3, Forwarding classes: nc
  Queued:
```

```
    Packets            :                 0                    0 pps
    Bytes              :                 0                    0 bps
  Transmitted:
    Packets            :                 0                     0 pps
    Bytes              :                 0                    0 bps
    Tail-dropped packets :               0                    0 pps
    RED-dropped packets  :               0                    0 pps
    RED-dropped bytes    :               0                    0 bps
=========
```

show interfaces queue lsq (lsq-ifd)

```
user@switch> show interfaces queue lsq-1/0/0
Logical interface lsq-1/0/0 (Index 348) (SNMP ifIndex 660)
Forwarding classes: 16 supported, 4 in use
Egress queues: 8 supported, 4 in use
Burst size: 0
Queue: 0, Forwarding classes: be
  Queued:
    Packets            :             55576                 1206 pps
    Bytes              :          29622008              5145472 bps
  Transmitted:
    Packets            :             55576                 1206 pps
    Bytes              :          29622008              5145472 bps
    Tail-dropped packets :               0                    0 pps
    RL-dropped packets   :               0                    0 pps
    RL-dropped bytes     :               0                    0 bps
    RED-dropped packets  :               0                    0 pps
     Low               :                 0                    0 pps
     Medium-low        :                 0                    0 pps
     Medium-high       :                 0                    0 pps
     High              :                 0                    0 pps
    RED-dropped bytes    :               0                    0 bps
     Low               :                 0                    0 bps
     Medium-low        :                 0                    0 bps
     Medium-high       :                 0                    0 bps
     High              :                 0                    0 bps
Queue: 1, Forwarding classes: ef
  Queued:
    Packets            :                 0                    0 pps
    Bytes              :                 0                    0 bps
  Transmitted:
    Packets            :                 0                    0 pps
    Bytes              :                 0                    0 bps
    Tail-dropped packets :               0                    0 pps
    RL-dropped packets   :               0                    0 pps
    RL-dropped bytes     :               0                    0 bps
    RED-dropped packets  :               0                    0 pps
     Low               :                 0                    0 pps
     Medium-low        :                 0                    0 pps
     Medium-high       :                 0                    0 pps
     High              :                 0                    0 pps
    RED-dropped bytes    :               0                    0 bps
     Low               :                 0                    0 bps
     Medium-low        :                 0                    0 bps
     Medium-high       :                 0                    0 bps
     High              :                 0                    0 bps
Queue: 2, Forwarding classes: af
  Queued:
    Packets            :                 0                    0 pps
```



```
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
     Low                 :                    0                     0 pps
     Medium-low          :                    0                     0 pps
     Medium-high         :                    0                     0 pps
     High                :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
     Low                 :                    0                     0 bps
     Medium-low          :                    0                     0 bps
     Medium-high         :                    0                     0 bps
     High                :                    0                     0 bps
Queue: 3, Forwarding classes: nc
  Queued:
    Packets              :                22231                   482 pps
    Bytes                :             11849123               2057600 bps
  Transmitted:
    Packets              :                22231                   482 pps
    Bytes                :             11849123               2057600 bps
    Tail-dropped packets :                    0                     0 pps
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
     Low                 :                    0                     0 pps
     Medium-low          :                    0                     0 pps
     Medium-high         :                    0                     0 pps
     High                :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
     Low                 :                    0                     0 bps
     Medium-low          :                    0                     0 bps
     Medium-high         :                    0                     0 bps
     High                :                    0                     0 bps
```

Sample Output

show interfaces queue (Aggregated Ethernet on a MX series Router)

user@host> **show interfaces queue ae0 remaining-traffic**

```
Physical interface: ae0    , Enabled, Physical link is Up
  Interface index: 128, SNMP ifIndex: 543
Forwarding classes: 16 supported, 4 in use
Egress queues: 8 supported, 4 in use
Queue: 0, Forwarding classes: best-effort
  Queued:
    Packets              :                   16                     0 pps
    Bytes                :                 1896                     0 bps
  Transmitted:
    Packets              :                   16                     0 pps
    Bytes                :                 1896                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
     Low                 :                    0                     0 pps
     Medium-low          :                    0                     0 pps
```



```
      Medium-high        :                     0                     0 pps
      High               :                     0                     0 pps
     RED-dropped bytes   :                     0                     0 bps
      Low                :                     0                     0 bps
      Medium-low         :                     0                     0 bps
      Medium-high        :                     0                     0 bps
      High               :                     0                     0 bps
   Queue-depth bytes     :
     Average             :                     0
     Current             :                     0
     Peak                :                     0
     Maximum             :             119013376
 Queue: 1, Forwarding classes: expedited-forwarding
   Queued:
     Packets             :                     0                     0 pps
     Bytes               :                     0                     0 bps
   Transmitted:
     Packets             :                     0                     0 pps
     Bytes               :                     0                     0 bps
     Tail-dropped packets :                    0                     0 pps
     RL-dropped packets  :                     0                     0 pps
     RL-dropped bytes    :                     0                     0 bps
     RED-dropped packets :                     0                     0 pps
      Low                :                     0                     0 pps
      Medium-low         :                     0                     0 pps
      Medium-high        :                     0                     0 pps
      High               :                     0                     0 pps
     RED-dropped bytes   :                     0                     0 bps
      Low                :                     0                     0 bps
      Medium-low         :                     0                     0 bps
      Medium-high        :                     0                     0 bps
      High               :                     0                     0 bps
   Queue-depth bytes     :
     Average             :                     0
     Current             :                     0
     Peak                :                     0
     Maximum             :                 32768
 Queue: 2, Forwarding classes: assured-forwarding
   Queued:
     Packets             :                     0                     0 pps
     Bytes               :                     0                     0 bps
   Transmitted:
     Packets             :                     0                     0 pps
     Bytes               :                     0                     0 bps
     Tail-dropped packets :                    0                     0 pps
     RL-dropped packets  :                     0                     0 pps
     RL-dropped bytes    :                     0                     0 bps
     RED-dropped packets :                     0                     0 pps
      Low                :                     0                     0 pps
      Medium-low         :                     0                     0 pps
      Medium-high        :                     0                     0 pps
      High               :                     0                     0 pps
     RED-dropped bytes   :                     0                     0 bps
      Low                :                     0                     0 bps
      Medium-low         :                     0                     0 bps
      Medium-high        :                     0                     0 bps
      High               :                     0                     0 bps
   Queue-depth bytes     :
     Average             :                     0
     Current             :                     0
     Peak                :                     0
```



```
    Maximum              :                32768
Queue: 3, Forwarding classes: network-control
  Queued:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
  Transmitted:
    Packets              :                    0                     0 pps
    Bytes                :                    0                     0 bps
    Tail-dropped packets :                    0                     0 pps
    RL-dropped packets   :                    0                     0 pps
    RL-dropped bytes     :                    0                     0 bps
    RED-dropped packets  :                    0                     0 pps
     Low                 :                    0                     0 pps
     Medium-low          :                    0                     0 pps
     Medium-high         :                    0                     0 pps
     High                :                    0                     0 pps
    RED-dropped bytes    :                    0                     0 bps
     Low                 :                    0                     0 bps
     Medium-low          :                    0                     0 bps
     Medium-high         :                    0                     0 bps
     High                :                    0                     0 bps
  Queue-depth bytes      :
    Average              :                    0
    Current              :                    0
    Peak                 :                    0
    Maximum              :              6258688
```


PART 4

Ethernet PAUSE

CHAPTER 10

Using Ethernet PAUSE

Understanding CoS Flow Control (Ethernet PAUSE and PFC)

Flow control supports lossless transmission by regulating traffic flows to avoid dropping frames during periods of congestion. Flow control stops and resumes the transmission of network traffic between two connected peer nodes on a full-duplex Ethernet physical link. Controlling the flow by pausing and restarting it prevents buffers on the nodes from overflowing and dropping frames. You configure flow control on a per-interface basis.

Two methods of peer-to-peer flow control are supported:

- IEEE 802.3X Ethernet PAUSE

  NOTE: QFX10000 switches do not support Ethernet PAUSE. Information about Ethernet PAUSE does not apply to QFX10000 switches.

  OCX Series switches support symmetric Ethernet PAUSE flow control on Layer 3 tagged interfaces. OCX Series switches do not support asymmetric Ethernet PAUSE flow control. Information about asymmetric flow control does not apply to OCX Series switches.

- IEEE 802.1Qbb priority-based flow control (PFC)

  NOTE: OCX Series switches do not support PFC or lossless Layer 2 transport. Information about PFC, lossless transport, and congestion notification profiles do not apply to OCX Series switches.

Video: Why Use PFC in a Data Center Network?

- General Information about Ethernet PAUSE and PFC and When to Use Them on page 526
- Ethernet PAUSE on page 526

General Information about Ethernet PAUSE and PFC and When to Use Them

Ethernet PAUSE and PFC are link-level flow control mechanisms.

NOTE: For end-to-end congestion control for best-effort traffic, see "Understanding CoS Explicit Congestion Notification" on page 380.

Ethernet PAUSE pauses transmission of all traffic on a physical Ethernet link.

PFC decouples the pause function from the physical Ethernet link and enables you to divide traffic on one link into eight priorities. You can think of the eight priorities as eight "lanes" of traffic that are mapped to forwarding classes and output queues. Each priority maps to a 3-bit IEEE 802.1p CoS code point value in the VLAN header. You can enable PFC on one or more priorities (IEEE 802.1p code points) on a link. When PFC-enabled traffic is paused on a link, traffic that is not PFC-enabled continues to flow (or is dropped if congestion is severe enough).

Use Ethernet PAUSE when you want to prevent packet loss on all of the traffic on a link. Use PFC to prevent traffic loss only on a specified type of traffic that require lossless treatment, for example, Fibre Channel over Ethernet (FCoE) traffic.

NOTE: Depending on the amount of traffic on a link or assigned to a priority, pausing traffic can cause ingress port congestion and spread congestion through the network.

Ethernet PAUSE and PFC are mutually exclusive configurations on an interface. Attempting to configure both Ethernet PAUSE and PFC on a link causes a commit error.

By default, all forms of flow control are disabled. You must explicitly enable flow control on interfaces to pause traffic.

Ethernet PAUSE

Ethernet PAUSE is a congestion relief feature that works by providing link-level flow control for all traffic on a full-duplex Ethernet link. Ethernet PAUSE works in both directions on the link. In one direction, an interface generates and sends Ethernet PAUSE messages to stop the connected peer from sending more traffic. In the other direction, the interface responds to Ethernet PAUSE messages it receives from the connected peer to stop sending traffic.

NOTE: QFX10000 switches do not support Ethernet PAUSE. Information about Ethernet PAUSE does not apply to QFX10000 switches.

OCX Series switches support symmetric Ethernet PAUSE flow control on Layer 3 tagged interfaces. OCX Series switches do not support asymmetric Ethernet PAUSE flow control. Information about asymmetric flow control does not apply to OCX Series switches.

Ethernet PAUSE also works on aggregated Ethernet interfaces. For example, if the connected peer interfaces are called Node A and Node B:

- When the receive buffers on interface Node A reach a certain level of fullness, the interface generates and sends an Ethernet PAUSE message to the connected peer (interface Node B) to tell the peer to stop sending frames. The Node B buffers store frames until the time period specified in the Ethernet PAUSE frame elapses; then Node B resumes sending frames to Node A.
- When interface Node A receives an Ethernet PAUSE message from interface Node B, interface Node A stops transmitting frames until the time period specified in the Ethernet PAUSE frame elapses; then Node A resumes transmission. (The Node A transmit buffers store frames until Node A resumes sending frames to Node B.)

  In this scenario, if Node B sends an Ethernet PAUSE frame with a time value of 0 to Node A, the 0 time value indicates to Node A that it can resume transmission. This happens when the Node B buffer empties to below a certain threshold and the buffer can once again accept traffic.

Symmetric flow control means an interface has the same Ethernet PAUSE configuration in both directions. The Ethernet PAUSE generation and Ethernet PAUSE response functions are both configured as enabled, or they are both disabled. You configure symmetric flow control by including the **flow-control** statement at the **[edit interfaces *interface-name* ether-options]** hierarchy level.

Asymmetric flow control allows you to configure the Ethernet PAUSE functionality in each direction independently on an interface. The configuration for generating Ethernet PAUSE messages and for responding to Ethernet PAUSE messages does not have to be the same. It can be enabled in both directions, disabled in both directions, or enabled in one direction and disabled in the other direction. You configure asymmetric flow control by including the **configured-flow-control** statement at the **[edit interfaces *interface-name* ether-options]** hierarchy level.

On any particular interface, symmetric and asymmetric flow control are mutually exclusive. Asymmetric flow control overrides and disables symmetric flow control. (If PFC is configured on an interface, you cannot commit an Ethernet PAUSE configuration on the interface. Attempting to commit an Ethernet PAUSE configuration on an interface with PFC enabled on one or more queues results in a commit error. To commit the PAUSE configuration, you must first delete the PFC configuration.) Both symmetric and asymmetric flow control are supported.

- Symmetric Flow Control on page 528
- Asymmetric Flow Control on page 528

Symmetric Flow Control

Symmetric flow control configures both the receive and transmit buffers in the same state. The interface can both send Ethernet PAUSE messages and respond to them (flow control is enabled), or the interface cannot send Ethernet PAUSE messages or respond to them (flow control is disabled).

When you enable symmetric flow control on an interface, the Ethernet PAUSE behavior depends on the configuration of the connected peer. With symmetric flow control enabled, the interface can perform any Ethernet PAUSE functions that the connected peer can perform. (When symmetric flow control is disabled, the interface does not send or respond to Ethernet PAUSE messages.)

Asymmetric Flow Control

Asymmetric flow control enables you to specify independently whether or not the interface receive buffer generates and sends Ethernet PAUSE messages to stop the connected peer from transmitting traffic, and whether or not the interface transmit buffer responds to Ethernet PAUSE messages it receives from the connected peer and stops transmitting traffic. The receive buffer configuration determines if the interface transmits Ethernet PAUSE messages, and the transmit buffer configuration determines if the interface receives and responds to Ethernet PAUSE messages:

- Receive buffers on—Enable Ethernet PAUSE transmission (generate and send Ethernet PAUSE frames)
- Transmit buffers on—Enable Ethernet PAUSE reception (respond to received Ethernet PAUSE frames)

You must explicitly set the flow control for both the receive buffer and the transmit buffer (**on** or **off**) to configure asymmetric Ethernet PAUSE. Table 92 on page 528 describes the configured flow control state when you set the receive (Rx) and transmit (Tx) buffers on an interface:

Table 92: Asymmetric Ethernet PAUSE Flow Control Configuration

Receive (Rx) Buffer	Transmit (Tx) Buffer	Configured Flow Control State
On	**Off**	Interface generates and sends Ethernet PAUSE messages. Interface does not respond to Ethernet PAUSE messages (interface continues to transmit even if peer requests that the interface stop sending traffic).
Off	**On**	Interface responds to Ethernet PAUSE messages received from the connected peer, but does not generate or send Ethernet PAUSE messages. (The interface does not request that the connected peer stop sending traffic.)

Table 92: Asymmetric Ethernet PAUSE Flow Control Configuration (continued)

Receive (Rx) Buffer	Transmit (Tx) Buffer	Configured Flow Control State
On	On	Same functionality as symmetric Ethernet PAUSE. Interface generates and sends Ethernet PAUSE messages and responds to received Ethernet PAUSE messages.
Off	Off	Ethernet PAUSE flow control is disabled.

The configured flow control is the Ethernet PAUSE state configured on the interface.

On 1-Gigabit Ethernet interfaces, autonegotiation of Ethernet PAUSE with the connected peer is supported. (Autonegotiation on 10-Gigabit Ethernet interfaces is not supported.) Autonegotiation enables the interface to exchange state advertisements with the connected peer so that the two devices can agree on the Ethernet PAUSE configuration. Each interface advertises its flow control state to the connected peer using a combination of the Ethernet PAUSE and ASM_DIR bits, as described in Table 93 on page 529:

Table 93: Flow Control State Advertised to the Connected Peer (Autonegotiation)

Rx Buffer State	Tx Buffer State	PAUSE Bit	ASM_DIR Bit	Description
Off	Off	0	0	The interface advertises no Ethernet PAUSE capability. This is equivalent to disabling flow control on an interface.
On	On	1	0	The interface advertises symmetric flow control (both the transmission of Ethernet PAUSE messages and the ability to receive and respond to Ethernet PAUSE messages).
On	Off	0	1	The interface advertises asymmetric flow control (the transmission of Ethernet PAUSE messages, but not the ability to receive and respond to Ethernet PAUSE messages).

Table 93: Flow Control State Advertised to the Connected Peer (Autonegotiation) (continued)

Rx Buffer State	Tx Buffer State	PAUSE Bit	ASM_DIR Bit	Description
Off	On	1	1	The interface advertises both symmetric and asymmetric flow control. Although the interface does not generate and send Ethernet PAUSE requests to the peer, the interface supports both symmetric and asymmetric Ethernet PAUSE configuration on the peer because the peer is not affected if the peer does not receive Ethernet PAUSE requests. (If the interface responds to the peer's Ethernet PAUSE requests, that is sufficient to support either symmetric or asymmetric flow control on the peer.)

The flow control configuration on each switch interface interacts with the flow control configuration of the connected peer. Each peer advertises its state to the other peer. The interaction of the flow control configuration of the peers determines the flow control behavior (resolution) between them, as shown in Table 94 on page 531. The first four columns show the Ethernet PAUSE configuration on the local QFX Series or EX4600 switch and on the connected peer (also known as the *link partner*). The last two columns show the Ethernet PAUSE resolution that results from the local and peer configurations on each interface. This illustrates how the Ethernet PAUSE configuration of each interface affects the Ethernet PAUSE behavior on the other interface.

NOTE: In the Resolution columns of the table, disabling Ethernet PAUSE transmit means that the interface receive buffers do not generate and send Ethernet PAUSE messages to the peer. Disabling Ethernet PAUSE receive means that the interface transmit buffers do not respond to Ethernet PAUSE messages received from the peer.

Table 94: Asymmetric Ethernet PAUSE Behavior on Local and Peer Interfaces

Local Interface (QFX Series or EX4600 Switch)		Peer Interface			
PAUSE Bit	ASM_DIR Bit	PAUSE Bit	ASM_DIR Bit	Local Resolution	Peer Resolution
0	0	Don't care	Don't care	Disable Ethernet PAUSE transmit and receive	Disable Ethernet PAUSE transmit and receive
0	1	0	Don't care	Disable Ethernet PAUSE transmit and receive	Disable Ethernet PAUSE transmit and receive
0	1	1	0	Disable Ethernet PAUSE transmit and receive	Disable Ethernet PAUSE transmit and receive
0	1	1	1	Enable Ethernet PAUSE transmit and disable Ethernet PAUSE receive	Disable Ethernet PAUSE transmit and enable Ethernet PAUSE receive
1	0	0	Don't care	Disable Ethernet PAUSE transmit and receive	Disable Ethernet PAUSE transmit and receive
1	0	1	Don't care	Enable Ethernet PAUSE transmit and receive	Enable Ethernet PAUSE transmit and receive
1	1	0	0	Disable Ethernet PAUSE transmit and receive	Disable Ethernet PAUSE transmit and receive
1	1	0	1	Enable Ethernet PAUSE receive and disable Ethernet PAUSE transmit	Enable Ethernet PAUSE transmit and disable Ethernet PAUSE receive
1	1	Don't care	Don't care	Enable Ethernet PAUSE transmit and receive	Enable Ethernet PAUSE transmit and receive

NOTE: For your convenience, Table 94 on page 531 replicates Table 28B-3 of Section 2 of the IEEE 802.X specification.

PFC

PFC is a lossless transport and congestion relief feature that works by providing granular link-level flow control for each IEEE 802.1p code point (priority) on a full-duplex Ethernet link. When the receive buffer on a switch interface fills to a threshold, the switch transmits a pause frame to the sender (the connected peer) to temporarily stop the sender from transmitting more frames. The buffer threshold must be low enough so that the sender has time to stop transmitting frames and the receiver can accept the frames already on the wire before the buffer overflows. The switch automatically sets queue buffer thresholds to prevent frame loss.

When congestion forces one priority on a link to pause, all of the other priorities on the link continue to send frames. Only frames of the paused priority are not transmitted. When the receive buffer empties below another threshold, the switch sends a message that starts the flow again.

You configure PFC using a congestion notification profile (CNP). A CNP has two parts:

- Input—Specify the code point (or code points) on which to enable PFC, and optionally specify the maximum receive unit (MRU) and the cable length between the interface and the connected peer interface.
- Output—Specify the output queue or output queues that respond to pause messages from the connected peer.

You apply a PFC configuration by configuring a CNP on one or more interfaces. Each interface that uses a particular CNP is enabled to pause traffic identified by the priorities (code points) specified in that CNP. You can configure one CNP on an interface, and you can configure different CNPs on different interfaces. When you configure a CNP on an interface, ingress traffic that is mapped to a priority that the CNP enables for PFC is paused whenever the queue buffer fills to the pause threshold. (The pause threshold is not user-configurable.)

Configure PFC for a priority end to end along the entire data path to create a lossless lane of traffic on the network. You can selectively pause the traffic in any queue without pausing the traffic for other queues on the same link. You can create lossless lanes for traffic such as FCoE, LAN backup, or management, while using standard frame-drop congestion management for IP traffic on the same link.

Potential consequences of flow control are:

- Ingress port congestion (configuring too many lossless flows can cause ingress port congestion)
- A paused priority that causes upstream devices to pause the same priority, thus spreading congestion back through the network

By definition, PFC supports symmetric pause only (as opposed to Ethernet PAUSE, which supports symmetric and asymmetric pause). With symmetric pause, a device can:

- Transmit pause frames to pause incoming traffic. (You configure this using the input stanza of a congestion notification profile.)
- Receive pause frames and stop sending traffic to a device whose buffer is too full to accept more frames. (You configure this using the output stanza of a congestion notification profile.)

Receiving a PFC frame from a connected peer pauses traffic on egress queues based on the IEEE 802.1p priorities that the PFC pause frame identifies. The priorities are 0 through 7. By default, the priorities map to queue numbers 0 through 7, respectively, and to specific forwarding classes, as shown in Table 95 on page 533:

Table 95: Default PFC Priority to Queue and Forwarding Class Mapping

IEEE 802.1p Priority (Code Point)	Queue	Forwarding Class
0 (000)	0	best-effort
1 (001)	1	best-effort
2 (010)	2	best-effort
3 (011)	3	fcoe
4 (100)	4	no-loss
5 (101)	5	best-effort
6 (110)	6	network-control
7 (111)	7	network-control

For example, a received PFC pause frame that pauses priority 3 pauses output queue 3. If you do not want to use the default configuration, you can configure customized mapping of priorities to queues and forwarding classes.

NOTE: By convention, deployments with converged server access typically use IEEE 802.1p priority 3 for FCoE traffic. The default configuration sets the **fcoe** forwarding class as a lossless forwarding class that is mapped to queue 3. The default classifier maps incoming priority 3 traffic to the **fcoe** forwarding class. *However, you must apply PFC to the entire FCoE data path to configure the end-to-end lossless behavior that FCoE traffic requires.*

If your network uses priority 3 for FCoE traffic, we recommend that you use the default configuration. If your network uses a priority other than 3 for FCoE traffic, you can configure lossless FCoE transport on any IEEE 80.21p priority as described in *Understanding CoS IEEE 802.1p Priorities for Lossless Traffic Flows* and *Understanding CoS IEEE 802.1p Priority Remapping on an FCoE-FC Gateway*.

To enable PFC on a priority:

1. Specify the IEEE 802.1p code point to pause in the input stanza of a CNP.
2. If you are not using the default lossless forwarding classes, specify the IEEE 802.1p code point to pause and the corresponding output queue in the output stanza of the CNP.
3. Apply the CNP to the ingress interfaces on which you want to pause the traffic.
4. If you are not using the default lossless forwarding classes, apply the CNP to the ingress interfaces on which you want to pause the traffic.

CAUTION: Any change to the PFC configuration on a port temporarily blocks the entire port (not just the priorities affected by the PFC change) so that the port can implement the change, then unblocks the port. Blocking the port stops ingress and egress traffic, and causes packet loss on all queues on the port until the port is unblocked.

A change to the PFC configuration means any change to a CNP, including changing the input portion of the CNP (enabling or disabling PFC on a priority, or changing the MRU or cable-length values) or changing the output portion of the CNP that enables or disables output flow control on a queue. A PFC configuration change only affects ports that use the changed CNP.

The following actions change the PFC configuration:

- Deleting or disabling a PFC configuration (input or output) in a CNP that is in use on one or more interfaces. For example:
  1. An existing CNP with an input stanza that enables PFC on priorities 3, 5, and 6 is configured on interfaces xe-0/0/20 and xe-0/0/21.
  2. We disable the PFC configuration for priority 6 in the input CNP, and then commit the configuration.
  3. The PFC configuration change causes all traffic on interfaces xe-0/0/20 and xe-0/0/21 to stop until the PFC change has been implemented. When the PFC change has been implemented, traffic resumes.
- Configuring a CNP on an interface. (This changes the PFC state by enabling PFC on one or more priorities.)
- Deleting a CNP from an interface. (This changes the PFC state by disabling PFC on one or more priorities.)

When you associate the CNP with an interface, the interface uses PFC to send pause requests when the output queue buffer for the lossless traffic fills to the pause threshold.

On switches that use different classifiers for unicast and multidestination traffic, you can map a unicast queue (queue 0 through 7) and a multidestination queue (queue 8, 9, 10, or 11) to the same IEEE 802.1p code point (priority) so that both unicast and multicast traffic use that priority. However, do not map multidestination traffic to lossless output queues. Starting with Junos OS Release 12.3, you can map one priority to multiple output queues.

NOTE: You can attach a maximum of one CNP to an interface, but you can create an unlimited number of CNPs that explicitly configure only the input stanza and use the default output stanza.

The output stanza of the CNP maps to a profile that interfaces use to respond to pause messages received from the connected peer. On standalone switches, you can create two CNPs with an explicitly configured output stanza.

When a switch is a Node device in a QFabric system, you can create one CNP with an explicitly configured output stanza. (One fewer profile is available on QFabric systems because the system needs a default profile for fabric interfaces, which are not used as fabric interfaces when the switches are not part of a QFabric system. *Understanding CoS IEEE 802.1p Priorities for Lossless Traffic Flows* describes configuring output flow control.

Lossless Transport Support Summary

The switch supports up to six lossless forwarding classes. For lossless transport, you must enable PFC on the IEEE 802.1p priorities (code points) mapped to lossless forwarding classes.

CAUTION: Any change to the PFC configuration on a port temporarily blocks the entire port (not just the priorities affected by the PFC change) so that the port can implement the change, then unblocks the port. Blocking the port stops ingress and egress traffic, and causes packet loss on all queues on the port until the port is unblocked.

The following limitation applies to support lossless transport on QFabric systems only:

- The internal fiber cable length from the QFabric system Node device to the QFabric system Interconnect device cannot exceed 150 meters.

The default CoS configuration provides two lossless forwarding classes, *fcoe* and *no-loss*. If you explicitly configure lossless forwarding classes, you must include the **no-loss** packet drop attribute to enable lossless behavior, or the traffic is not lossless. For both default and explicit lossless forwarding class configuration, you must configure CNP input stanzas to enable PFC on the priority of the lossless traffic and apply the CNPs to ingress interfaces.

NOTE: The information in this note applies only to systems that do not run the ELS CLI.

Junos OS Release 12.2 introduced changes to the way the switch handles lossless forwarding classes (including the default **fcoe** and **no-loss** forwarding classes).

In Junos OS Release 12.1, either explicitly configuring the **fcoe** and **no-loss** forwarding classes or using the default configuration for these forwarding classes resulted in the same lossless behavior for traffic mapped to those forwarding classes.

However, in Junos OS Release 12.2, if you explicitly configure the **fcoe** or the **no-loss** forwarding class, that forwarding class is no longer treated as a lossless forwarding class. Traffic mapped to these forwarding classes is treated as lossy (best-effort) traffic. This is true even if the explicit configuration is exactly the same as the default configuration.

If your CoS configuration from Junos OS Release 12.1 or earlier includes the explicit configuration of the **fcoe** or the **no-loss** forwarding class, then when you upgrade to Junos OS Release 12.2, those forwarding classes are not lossless. To preserve the lossless treatment of these forwarding classes, delete the the explicit **fcoe** and **no-loss** forwarding class configuration before you upgrade to Junos OS Release 12.2.

See *Overview of CoS Changes Introduced in Junos OS Release 12.2* for detailed information about this change and how to delete an existing lossless configuration.

In Junos OS Release 12.3, the default behavior of the **fcoe** and **no-loss** forwarding classes is the same as in Junos OS Release 12.2. However, in Junos OS Release 12.3, you can configure up to six lossless forwarding classes. All explicitly configured lossless forwarding classes must include the new **no-loss** packet drop attribute or the forwarding class is lossy.

Understanding CoS IEEE 802.1p Priorities for Lossless Traffic Flows provides detailed information about the explicit configuration of lossless priorities and about the default configuration of lossless priorities, including the input and output stanzas of the CNP.

NOTE: PFC and Ethernet PAUSE are used only on Ethernet interfaces. Fabric (fte) ports on QFabric systems (Node device fabric ports and Interconnect device fabric ports) use link-layer flow control (LLFC) to ensure the appropriate treatment of lossless traffic.

Release History Table

Release	Description
12.3	Starting with Junos OS Release 12.3, you can map one priority to multiple output queues.

Related Documentation

- *Understanding DCB Features and Requirements*
- Understanding CoS Explicit Congestion Notification on page 380
- *Configuring CoS PFC (Congestion Notification Profiles)*
- *Example: Configuring CoS PFC for FCoE Traffic*

Enabling and Disabling CoS Symmetric Ethernet PAUSE Flow Control

Ethernet PAUSE flow control is a congestion relief feature that works by providing link-level flow control for all traffic on a full-duplex Ethernet link, including Ethernet links that belong to Ethernet link aggregated (LAG) interfaces. Ethernet PAUSE works in both directions on the link. In one direction, an interface generates and sends PAUSE messages to stop the connected peer from sending more traffic. In the other direction, the interface responds to PAUSE messages it receives from the connected peer to stop sending traffic.

Symmetric flow control means that an interface has the same PAUSE configuration in both directions. The PAUSE generation and PAUSE response functions are both configured as enabled, or they are both disabled.

Asymmetric flow control allows you to configure the PAUSE functionality in each direction independently on an interface. The configuration for generating PAUSE messages and for responding to PAUSE messages does not have to be the same. It can be enabled in both directions, disabled in both directions, or enabled in one direction and disabled in the other direction. If you do not want to PAUSE all of the traffic on a link, you can use priority-based flow control (PFC) to selectively pause traffic based on its IEEE 802.1p code point.

NOTE: OCX Series switches do not support PFC.

On any particular interface, symmetric and asymmetric flow control are mutually exclusive. If you attempt to configure both features, the switch returns a commit error. Ethernet PAUSE and PFC are also mutually exclusive features, so you cannot configure both of them on the same interface. If you attempt to configure both Ethernet PAUSE and PFC on an interface, the switch returns a commit error.

By default, all flow control features are disabled. You enable symmetric flow control on the interfaces on which you want to PAUSE all of the traffic on a link.

- To enable symmetric flow control on an interface:

```
[edit interfaces interface-name ether-options]
```

```
user@switch# set flow-control
```

- To disable symmetric flow control on an interface:

```
[edit interfaces interface-name ether-options]
user@switch# set no-flow-control
```

Related Documentation

- *Configuring CoS Asymmetric Ethernet PAUSE Flow Control*
- *Configuring CoS PFC (Congestion Notification Profiles)*
- Understanding CoS Flow Control (Ethernet PAUSE and PFC) on page 525

CHAPTER 11

Configuration Statements for Ethernet PAUSE

flow-control

Syntax (flow-control | no-flow-control);

Hierarchy Level [edit interfaces *interface-name* ether-options]

Release Information Statement introduced in Junos OS Release 11.1 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Explicitly enable or disable symmetric Ethernet PAUSE flow control, which regulates the flow of packets from the switch to the remote side of the connection by pausing all traffic flows on a link during periods of network congestion. Symmetric flow control means that Ethernet PAUSE is enabled in both directions. The interface generates and sends Ethernet PAUSE messages when the receive buffers fill to a certain threshold and the interface responds to PAUSE messages received from the connected peer. By default, flow control is disabled.

You can configure asymmetric flow control by including the **configured-flow-control** statement at the **[edit interfaces *interface-name* ether-options** hierarchy level. Symmetric flow control and asymmetric flow control are mutually exclusive features. If you attempt to configure both, the switch returns a commit error.

NOTE: Ethernet PAUSE temporarily stops transmitting all traffic on a link when the buffers fill to a certain threshold. To temporarily pause traffic on individual “lanes” of traffic (each lane contains the traffic associated with a particular IEEE 802.1p code point, so there can be eight lanes of traffic on a link), use priority-based flow control (PFC).

Ethernet PAUSE and PFC are mutually exclusive features, so you cannot configure both of them on the same interface. If you attempt to configure both Ethernet PAUSE and PFC on an interface, the switch returns a commit error.

OCX Series switches do not support PFC.

- **flow-control**—Enable flow control; flow control is useful when the remote device is a Gigabit Ethernet switch.
- **no-flow-control**—Disable flow control.

Default Flow control is disabled.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- *configured-flow-control*
- *Configuring Gigabit and 10-Gigabit Ethernet Interfaces for EX4600 and QFX Series Switches*
- Understanding CoS Flow Control (Ethernet PAUSE and PFC) on page 525
- *Junos OS Network Interfaces Library for Routing Devices*

PART 5

Buffers

CHAPTER 12

Using Buffers

Understanding CoS Buffer Configuration

Packet Forwarding Engine (PFE) wide common packet buffer memory is used to store packets on interface queues. The buffer memory has separate ingress and egress accounting to make accept, drop, or pause decisions. Because the switch has a single pool of memory with separate ingress and egress accounting, the full amount of buffer memory is available from both the ingress and the egress perspective. Packets are accounted for as they enter and leave the switch, but there is no concept of a packet arriving at an ingress buffer and then being moved to an egress buffer. Specific common buffer memory amounts for individual switches is listed in Table 96 on page 547.

Table 96: Common Packet Buffer Memory on Switches

Switch	Common Packet Buffer Memory
QFX3500, QFX3600	9 MB
QFX5100, EX4600, and OCX Series	12 MB
QFX5110, QFX5200-32C	16MB
QFX5200-48Y	22MB
QFX5210	42MB

NOTE: QFX10000 does not have a shared buffer.

The buffers are divided into two pools from both an ingress and an egress perspective:

1. *Shared buffers* are a global memory pool that the switch allocates dynamically to ports as needed, so the buffers are shared among the switch ports.
2. *Dedicated buffers* are a memory pool divided equally among the switch ports. Each port receives a minimum guaranteed amount of buffer space, dedicated to each port, not shared among ports.

NOTE: Lossless traffic is traffic on which you enable priority-based flow control (PFC) to ensure lossless transport. Lossless traffic does not refer to best-effort traffic on a link enabled for Ethernet PAUSE (IEEE 802.3x).

OCX Series switches do not support lossless transport or PFC. In this topic, references to lossless transport do not apply to OCX Series switches.

The switch reserves nonconfigurable buffer space to ensure that ports and queues receive a minimum memory allocation. You can configure how the system uses the rest of the buffer space to optimize the allocation for your mix of network traffic. You can configure the percentage of available buffer space used as shared buffer space versus dedicated buffer space. You can also configure how shared buffer space is allocated to different types of traffic. You can optimize the buffer settings for the traffic on your network.

The default buffer configuration is designed for networks that have a balance of best-effort and lossless traffic. Because OCX Series switches do not support lossless traffic, instead of using the default buffer configuration on OCX Series switches, consider configuring the buffers as recommended for networks with mostly best-effort traffic as shown in Table 115 on page 563 and Table 116 on page 563.

The default class-of-service configuration provides two lossless forwarding classes (**fcoe** and **no-loss)**, a best-effort unicast forwarding class, a network control traffic forwarding class, and one multidestination (multicast, broadcast, and destination lookup fail) forwarding class.

NOTE: On OCX Series switches, do not map traffic to the default lossless forwarding classes.

Each default forwarding class maps to a different default output queue. The default configuration allocates the buffers in a manner that supports a moderate amount of lossless traffic while still providing the ability to absorb bursts in best-effort traffic transmission.

Changing the buffer settings changes the abilities of the buffers to absorb traffic bursts and handle lossless traffic. For example, networks with mostly best-effort traffic require allocating most of the shared buffer space to best-effort buffers. This provides deep, flexible buffers that can absorb traffic bursts with minimal packet loss, at the expense of buffer availability for lossless traffic.

Conversely, networks with mostly lossless traffic require allocating most of the shared buffer space to lossless headroom buffers. This prevents packet loss on lossless flows at the expense of absorbing bursty best-effort traffic efficiently.

CAUTION: Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.

This topic describes the buffer architecture and settings:

Buffer Pools

From both an ingress and an egress perspective, the PFE buffer is split into two main pools, a shared buffer pool and a dedicated buffer pool that ensures a minimum allocation to each port. You can configure the amount of buffer space allocated to each of the two pools. A portion of the buffer space is reserved so that there is always a minimum amount of shared and dedicated buffer space available to each port.

- Shared buffer pool—A global memory space that all of the ports on the switch share dynamically as they need buffers. The shared buffer pool is further partitioned into buffers for best-effort unicast, best-effort multidestination (broadcast, multicast, and destination lookup fail), and PFC (lossless) traffic types. You can allocate global shared memory space to buffer partitions to better support different mixes of network traffic. The larger the shared buffer pool, the better the switch can absorb traffic bursts because more shared memory is available for the traffic.
- Dedicated buffer pool—A reserved global memory space allocated equally to each port. The switch reserves a minimum dedicated buffer pool that is not user-configurable. You can divide the dedicated buffer allocation for a port among the port queues on a per-port, per-queue basis. (For example, this enables you to dedicate more buffer space to queues that transport lossless traffic.)

  A larger dedicated buffer pool means a larger amount of dedicated buffer space for each port, so congestion on one port is less likely to affect traffic on another port because the traffic does not need to use as much shared buffer space. However, the larger the dedicated buffer pool, the less bursty traffic the switch can handle because there is less dynamic shared buffer memory.

You can configure the way the available unreserved portion of the buffer space is allocated to the global shared buffer pool and to the dedicated shared buffer pool by configuring the ingress and egress shared buffer percentages.

By default, 100 percent of the available unreserved buffer space is allocated to the shared buffer pool. If you change the percentage of space allocated to the shared buffer, the available buffer space that is not allocated to the shared buffer is allocated to the dedicated buffer. For example, if you configure the ingress shared buffer pool as 80 percent, the remaining 20 percent of the available buffer space is allocated to the dedicated buffer pool and divided equally across the ports.

NOTE: When 100 percent of the available (user-configurable) buffers are allocated to the shared buffer pool, the switch still reserves a minimum dedicated buffer pool.

You can separately configure ingress and egress shared buffer pool allocations. You can also partition the ingress and egress shared buffer pool to allocate percentages of the shared buffer pool to specific types of traffic. If you do not use the default configuration or one of the recommended configurations, pay particular attention to the ingress configuration of the lossless headroom buffers (these buffers handle PFC pause during periods of congestion) and to the egress configuration of the best-effort buffers to handle incast congestion (multiple synchronized sources sending data to the same receiver in parallel).

In addition to the shared buffer pool and the dedicated buffer pool, there is also a small ingress global headroom buffer pool that is reserved and is not configurable.

When contention for buffer space occurs, the switch uses an internal algorithm to ensure that the buffer pools are distributed fairly among competing flows. When traffic for a given flow exceeds the amount of dedicated port buffer reserved for that flow, the flow begins to consume memory from the dynamic shared buffer pool. Competing flows compete for shared buffer memory with other flows that also have exhausted their dedicated buffers. When there is no congestion, there are no competing flows.

Buffer Handling of Lossless Flows (PFC) Versus Ethernet PAUSE

When we discuss lossless buffers in the following sections, we mean buffers that handle traffic on which you enable PFC to ensure lossless transport. The lossless buffers are not used for best-effort traffic on a link on which you enable Ethernet PAUSE (IEEE 802.3x). The lossless ingress and egress shared buffers, and the ingress lossless headroom shared buffer, are used only for traffic on which you enable PFC.

NOTE: To support lossless flows, you must configure the appropriate data center bridging capabilities (PFC, DCBX, and ETS) and scheduling properties.

NOTE: OCX Series switches do not support PFC or lossless transport. OCX Series switches support symmetric Ethernet PAUSE.

Shared Buffer Pool and Partitions

The shared buffer pool is a global memory space that all of the ports on the switch share dynamically as they need buffers. The switch uses the shared buffer pool to absorb traffic bursts after the dedicated buffer pool for a port is exhausted.

You can divide both the ingress shared buffer pool and the egress shared buffer pool into three partitions to allocate percentages of each buffer pool to different types of traffic. When you partition the ingress or egress shared buffer pool:

- If you explicitly configure one ingress shared buffer partition, you must explicitly configure all three ingress shared buffer partitions. (You either explicitly configure all three ingress partitions or you use the default setting for all three ingress partitions.)

  If you explicitly configure one egress shared buffer partition, you must explicitly configure all three egress shared buffer partitions. (You either explicitly configure all three egress partitions or you use the default setting for all three egress partitions.)

  The switch returns a commit error if you do not explicitly configure all three partitions when configuring the ingress or egress shared buffer partitions.

- The combined percentages of the three ingress shared buffer partitions must total exactly 100 percent.

  The combined percentages of the three egress shared buffer partitions must total exactly 100 percent.

  When you explicitly configure ingress or egress shared buffer partitions, the switch returns a commit error if the total percentage of the three partitions does not equal 100 percent.
- If you explicitly partition one set of shared buffers, you do not have to explicitly partition the other set of shared buffers. For example, you can explicitly configure the ingress shared buffer partitions and use the default egress shared buffer partitions. However, if you change the buffer partitions for the ingress buffer pool to match the expected types of traffic flows, you would probably also want to change the buffer partitions for the egress buffer pool to match those traffic flows.

You can configure the percentage of available unreserved buffer space allocated to the shared buffer pool. Space that you do not allocate to the shared buffer pool is added to the dedicated buffer pool and divided equally among the ports. The default configuration allocates 100 percent of the unreserved ingress and egress buffer space to the shared buffers.

Configuring the ingress and egress shared buffer pool partitions enables you to allocate more buffers to the types of traffic your network predominantly carries, and fewer buffers to other traffic.

Ingress Shared Buffer Pool Partitions

You can configure three ingress buffer pool partitions:

- Lossless buffers—Shared buffer pool for all lossless ingress traffic. We recommend 5 percent as the minimum value for lossless buffers.
- Lossless headroom buffers—Shared buffer pool for packets received while a pause is asserted. If PFC is enabled on priorities on a port, when the port sends a pause message to the connected peer, the port uses the headroom buffers to store the packets that arrive between the time the port sends the pause message and the time the last packet arrives after the peer pauses traffic. The minimum value for lossless headroom buffers is 0 (zero) percent. (Lossless headroom buffers are the only buffers for which the recommended value can be less than 5 percent.)
- Lossy buffers—Shared buffer pool for all best-effort ingress traffic (best-effort unicast, multidestination, and strict-high priority traffic). We recommend 5 percent as the minimum value for best-effort buffers.

The combined percentage values of the ingress lossless, lossless headroom, and best-effort buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. If you explicitly configure an ingress shared buffer partition, you must explicitly configure all three ingress buffer partitions, even if the lossless headroom buffer partition has a value of 0 (zero) percent.

Egress Shared Buffer Pool Partitions

You can configure three egress buffer pool partitions:

- Lossless buffers—Shared buffer pool for all lossless egress queues. We recommend 5 percent as the minimum value for lossless buffers.
- Lossy buffers—Shared buffer pool for all best-effort egress queues (best-effort unicast, and strict-high priority queues). We recommend 5 percent as the minimum value for best-effort buffers.
- Multicast buffers—Shared buffer pool for all multidestination (multicast, broadcast, and destination lookup fail) egress queues. We recommend 5 percent as the minimum value for multicast buffers.

The combined percentage values of the egress lossless, lossy, and multicast buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All egress buffer partitions must be explicitly configured and should have a value of at least 5 percent. If you explicitly configure an egress shared buffer partition, you must explicitly configure all three egress buffer partitions, and each partition should have a value of at least 5 percent.

NOTE: QFX5200-32C does not replicate all multicast streams when two or more downstream interface packet sizes are higher than ~6k and have an 1000pps packet ingress rate. This is because the number of working flows on QFX5200-32C is indirectly proportional to the packet size and directly proportional to available multicast shared buffers.

Dedicated Port Buffer Pool and Buffer Allocation to Queues

The global dedicated buffer pool is memory that is allocated equally to each port, so each port receives a guaranteed minimum amount of buffer space. Dedicated buffers are not shared among ports. Each port receives an equal proportion of the dedicated buffer pool.

The amount of dedicated buffer space is not user-configurable and depends on the percentage of available nonreserved buffers allocated to the shared buffers. (The dedicated buffer space is equal to the minimum reserved port buffers plus the remainder of the available nonreserved buffers that are not allocated to the shared buffer pool.)

When traffic enters and exits the switch, the switch ports use their dedicated buffers to store packets. If the dedicated buffers are not sufficient to handle the traffic, the switch uses shared buffers. The only way to increase the dedicated buffer pool is to decrease the shared buffer pool from its default value of 100 percent of available unreserved buffers.

NOTE: If 100 percent of the available unreserved buffers are allocated to the shared buffer pool, the switch still reserves a minimum dedicated buffer pool.

The larger the shared buffer pool, the better the burst absorption across the ports. The larger the dedicated buffer pool, the larger the amount of dedicated buffer space for each port. The greater the dedicated buffer space, the less likely that congestion on one port can affect traffic on another port, because the traffic does not need to use as much shared buffer space.

Allocating Dedicated Port Buffers to Queues

You can divide the dedicated buffer allocation for an egress port among the port queues by including the **buffer-size** statement in the scheduler configuration. This enables you to control the egress port dedicated buffer allocation on a per-port, per-queue basis. (For example, this enables you to dedicate more buffer space to queues that transport lossless traffic, or to stop the port from reserving buffers for queues that do not carry traffic.) Egress dedicated port buffer allocation is a hierarchical structure that allocates a global dedicated buffer pool evenly among ports, and then divides the allocation for each port among the port queues.

By default, ports divide their allocation of dedicated buffers among their egress queues in the same proportion as the default scheduler sets the minimum guaranteed transmission rates (the **transmit-rate** option) for traffic. Only the queues included in the default scheduler receive bandwidth and dedicated buffers, in the proportions shown in Table 97 on page 553:

Table 97: Default Dedicated Buffer Allocation to Egress Queues (Based on Default Scheduler)

Forwarding Class	Queue	Minimum Guaranteed Bandwidth (transmit-rate)	Proportion of Reserved Dedicated Port Buffers
best-effort	0	5%	5%
fcoe	3	35%	35%
no-loss	4	35%	35%
network-control	7	5%	5%
mcast	8	20%	20%

In the default configuration, no egress queues other than the ones shown in Table 97 on page 553 receive an allocation of dedicated port buffers.

NOTE: The switch uses hierarchical scheduling to control port and queue bandwidth allocation, as described in "Understanding CoS Hierarchical Port Scheduling (ETS)" on page 315 and shown in "Example: Configuring CoS Hierarchical Port Scheduling (ETS)" on page 321. For egress queue buffer size configuration, when you attach a traffic control profile (includes the queue scheduler information) to a port, the dedicated egress buffers on the port are divided among the queues as configured in the scheduler.

If you do not want to use the default allocation of dedicated port buffers to queues, use the **buffer-size** option in the scheduler that is attached to the port to configure the queue allocation. You can configure the dedicated buffer allocation to queues in two ways:

- As a percentage—The queue receives the specified percentage of dedicated port buffers when the queue is mapped to the scheduler and the scheduler is attached to a port.
- As a remainder—After the port services the queues that have an explicit percentage buffer size configuration, the remaining dedicated port buffer space is divided equally among the other queues to which a scheduler is attached. (No default or explicit scheduler for a queue means no dedicated buffer allocation for that queue.) If you configure a scheduler and you do not specify a buffer size as a percentage, *remainder* is the default setting.

NOTE: **The total of all of the explicitly configured buffer size percentages for all of the queues on a port cannot exceed 100 percent.**

Configuring Dedicated Port Buffer Allocation to Queues

In a port configuration that includes multiple forwarding class sets, with multiple forwarding classes mapped to multiple schedulers, the allocation of port dedicated buffers to queues depends on the mix of queues with buffer sizes configured as explicit percentages and queues configured with (or defaulted to) the **remainder** option.

The best way to demonstrate how using the percentage and remainder options affects dedicated port buffer allocation to queues is by showing an example of queue buffer allocation, and then showing how the queue buffer allocation changes when you add another forwarding class (queue) to the port.

Table 98 on page 554 shows an initial configuration that includes four forwarding class sets, the five default forwarding classes (mapped to the five default queues for those forwarding classes), the **buffer-size** option configuration, and the resulting buffer allocation for each queue. Table 99 on page 555 shows the same configuration after we add another forwarding class (best-effort-2, mapped to queue 1) to the best-effort forwarding class set. Comparing the buffer allocations in each table shows you how adding another queue affects buffer allocation when you use remainders and explicit percentages to configure the buffer allocation for different queues.

Table 98: Egress Queue Dedicated Buffer Allocation (Example 1)

Forwarding Class Set (Priority Group)	Forwarding Class	Queue	Scheduler Buffer Size Configuration	Buffer Allocation per Queue (Percentage)
fc-set-be	best-effort	0	10%	10%
fc-set-lossless	fcoe	3	20%	20%
	no-loss	4	40%	40%

Table 98: Egress Queue Dedicated Buffer Allocation (Example 1) (continued)

Forwarding Class Set (Priority Group)	Forwarding Class	Queue	Scheduler Buffer Size Configuration	Buffer Allocation per Queue (Percentage)
fc-set-strict-high	network-control	7	remainder	15%
fc-set-mcast	mcast	8	remainder	15%

In this first example, 70 percent of the egress port dedicated buffer pool is explicitly allocated to the best-effort, fcoe, and no-loss queues. The remaining 30 percent of the port dedicated buffer pool is split between the two queues that use the **remainder** option (network-control and mcast), so each queue receives 15 percent of the dedicated buffer pool.

Now we add another forwarding class (queue) to the best-effort priority group (fc-set-be) and configure it with a buffer size of *remainder* instead of configuring a specific percentage. Because a third queue now shares the remaining dedicated buffers, the queues that share the remainder receive fewer dedicated buffers, as shown in Table 99 on page 555. The queues with explicitly configured percentages receive the configured percentage of dedicated buffers.

Table 99: Egress Queue Dedicated Buffer Allocation with Another Remainder Queue (Example 2)

Priority Group (fc-set)	Forwarding Class	Queue	Scheduler Buffer Size Configuration	Buffer Allocation per Queue (Percentage)
fc-set-be	best-effort	0	10%	10%
	best-effort-2	1	remainder	10%
fc-set-lossless	fcoe	3	20%	20%
	no-loss	4	40%	40%
fc-set-strict-high	network-control	7	remainder	10%
fc-set-mcast	mcast	8	remainder	10%

The two tables show how the port divides the dedicated buffer space that remains after servicing the queues that have an explicitly configured percentage of dedicated buffer space.

Trade-off Between Shared Buffer Space and Dedicated Buffer Space

The trade-off between shared buffer space and dedicated buffer space is:

- Shared buffers provide better absorption of traffic bursts because there is a larger pool of dynamic buffers that ports can use as needed to handle the bursts. However, all flows that exhaust their dedicated buffer space compete for the shared buffer pool. A larger shared buffer pool means a smaller dedicated buffer pool, and therefore more competition for the shared buffer pool because more flows exhaust their dedicated buffer allocation. Too much shared buffer space results in no single flow receiving very much shared buffer space, to maintain fairness when many flows contend for that space.

- Dedicated buffers provide guaranteed buffer space to each port. The larger the dedicated buffer pool, the less likely that congestion on one port affects traffic on another port, because the traffic does not need to use as much shared buffer space. However, less shared buffer space means less ability to dynamically absorb traffic bursts.

For optimal burst absorption, the switch needs enough dedicated buffer space to avoid persistent competition for the shared buffer space. When fewer flows compete for the shared buffers, the flows that need shared buffer space to absorb bursts receive more of the shared buffer because fewer flows exhaust their dedicated buffer space.

The default configuration and the configurations recommended for different traffic scenarios allocate 100 percent of the user-configurable memory space to the global shared buffer pool because the amount of space reserved for dedicated buffers provides enough space to avoid persistent competition for dynamic shared buffers. This results in fewer flows competing for the shared buffers, so the competing flows receive more of the buffer space.

Order of Buffer Consumption

The total buffer pool is divided into ingress and egress shared buffer pools and dedicated buffer pools. When traffic flows through the switch, the buffer space is used in a particular order that depends on the type of traffic.

On ingress, the order of buffer consumption is:

- Best-effort unicast traffic:
  1. Dedicated buffers
  2. Shared buffers
  3. Global headroom buffers (very small)
- Lossless unicast traffic:
  1. Dedicated buffers
  2. Shared buffers
  3. Lossless headroom buffers
  4. Global headroom buffers (very small)
- Multidestination traffic:
  1. Dedicated buffers
  2. Shared buffers

3. Global headroom buffers (very small)

On egress, the order of buffer consumption is the same for unicast best-effort, lossless unicast, and multidestination traffic:

- Dedicated buffers
- Shared buffers

In all cases on all ports, the switch uses the dedicated buffer pool first and the shared buffer pool only after the dedicated buffer pool for the port or queue is exhausted. This reserves the maximum amount of dynamic shared buffer space to absorb traffic bursts.

Default Buffer Pool Values

You can view the default or configured ingress and egress buffer pool values in KB units using the **show class-of-service shared-buffer** operational command. You can view the configured shared buffer pool values in percent units using the **show configuration class-of-service shared-buffer** operational command.

This section provides the default total buffer, shared buffer, and dedicated buffer values.

Total Buffer Pool Size

The total buffer pool is common memory that has separate ingress and egress accounting, so the full buffer pool is available from both the ingress and egress perspective. The total buffer pool consists of the dedicated buffer space and the shared buffer space. The size of the total buffer pool is not user-configurable, but the allocation of buffer space to the dedicated and shared buffer pools is user-configurable.

On QFX3500 and QFX3600 switches, the combined total size of the ingress and egress buffer pools is approximately 9 MB (exactly 9360 KB).

On QFX5100, EX4600, and OCX Series switches, the combined total size of the ingress and egress buffer pools is approximately 12 MB (exactly 12480 KB).

On QFX5110 and QFX5200-32C switches, the combined total size of the ingress and egress buffer pools is approximately 16 MB.

On QFX5200-48Y switches, the combined total size of the ingress and egress buffer pools is approximately 22 MB.

On QFX5210 switches, the combined total size of the ingress and egress buffer pools is approximately 42 MB.

Shared Buffer Pool Default Values

Some switches have a larger shared buffer pool than other switches. However, the allocation of shared buffer space to the individual ingress and egress buffer pools is the same on a percentage basis, even though the absolute values are different. For example, the default ingress lossless buffer is 9 percent of the total shared ingress buffer space on all of the switches, even though the default absolute value of the ingress lossless buffer differs from switch to switch.

This section describes the default values in percent and in KB for the shared ingress and shared egress buffers.

- Shared Ingress Buffer Default Values on page 558
- Shared Egress Buffer Default Values on page 559

Shared Ingress Buffer Default Values

Table 100 on page 558 shows the default ingress shared buffer allocation values in KB units for QFX5210 switches.

Table 100: QFX5210 Switch Default Shared Ingress Buffer Values (KB)

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
29224	2630.16	13150.80	13443.04

Table 101 on page 558 shows the default ingress shared buffer allocation values in KB units for QFX5200-48Y switches.

Table 101: QFX5200-48Y Switch Default Shared Ingress Buffer Values (KB)

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
19154.69	1723.92	8619.61	8811.16

Table 102 on page 558 shows the default ingress shared buffer allocation values in KB units for QFX5110 and QFX5200-32C switches.

Table 102: QFX5110 and QFX5200-32C Switch Default Shared Ingress Buffer Values (KB)

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
11779.62	1060.17	5300.83	5418.63

Table 103 on page 559 shows the default ingress shared buffer allocation values in KB units for QFX5100, EX4600, and OCX Series switches.

Table 103: QFX5100, EX4600, and OCX Series Switch Default Shared Ingress Buffer Values (KB)

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
9567.19 KB	861.05 KB	4305.23 KB	4400.91 KB

Table 104 on page 559 shows the default ingress shared buffer allocation values in KB units for QFX3500 and QFX3600 switches.

Table 104: QFX3500 and QFX3600 Switch Default Shared Ingress Buffer Values (KB)

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
7202 KB	648.18 KB	3240.9 KB	3312.92 KB

Table 105 on page 559 shows the default ingress shared buffer allocation values as percentages for all switches. (If you change the default shared buffer allocation, you configure the change as a percentage.)

Table 105: Default Shared Ingress Buffer Values (Percentage)

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
100%	9%	45%	46%

Shared Egress Buffer Default Values

Table 106 on page 559 shows the default egress shared buffer allocation values in KB units for QFX5210 switches.

Table 106: QFX5210 Switch Default Shared Egress Buffer Values (KB)

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
28080	14040	8704.80	5335.20

Table 107 on page 559 shows the default egress shared buffer allocation values in KB units for QFX5200-48Y switches.

Table 107: QFX5200-48Y Switch Default Shared Egress Buffer Values (KB)

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
19115.69	9557.84	5925.86	3631.98

Table 108 on page 560 shows the default egress shared buffer allocation values in KB units for QFX5110 and QFX5200-32C switches.

Table 108: QFX5110 and QFX5200-32C Switch Default Shared Egress Buffer Values (KB)

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
11232	5616	3481.92	2134

NOTE: QFX5200-32C does not replicate all multicast streams when two or more downstream interface packet sizes are higher than ~6k and have an 1000pps packet ingress rate. This is because the number of working flows on QFX5200-32C is indirectly proportional to the packet size and directly proportional to available multicast shared buffers.

Table 109 on page 560 shows the default egress shared buffer allocation values in KB units for QFX5100, EX4600, and OCX Series switches.

Table 109: QFX5100, EX4600, and OCX Series Switch Default Shared Egress Buffer Values (KB)

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
8736 KB	4368 KB	2708.16 KB	1659.84 KB

Table 110 on page 560 shows the default egress shared buffer allocation values in KB units.

Table 110: QFX3500 and QFX3600 Switch Default Shared Egress Buffer Values (KB)

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
6656 KB	3328 KB	2063.36 KB	1264.64 KB

Table 111 on page 560 shows the default egress shared buffer allocation values for all switches as percentages.

Table 111: Default Shared Egress Buffer Values (Percentage)

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	50%	31%	19%

Dedicated Buffer Pool Default Values

The system reserves ingress and egress dedicated buffer pools that are divided equally among the switch ports. By default, the system allocates 100 percent of the available unreserved buffer space to the shared buffer pool. If you reduce the percentage of available unreserved buffer space allocated to the shared buffer pool, the remaining unreserved buffer space is added to the dedicated buffer pool allocation. You configure the amount of dedicated buffer pool space by reducing (or increasing) the percentage of buffer space allocated to the shared buffer pool. You do not directly configure the dedicated buffer pool allocation.

Table 112 on page 561 shows the default ingress and egress dedicated buffer pool values in KB units for QFX5210, QFX5200, QFX5110, QFX5100, QFX3500, QFX3600, EX4600, and OCX Series switches.

Table 112: Default Ingress and Egress Dedicated Buffer Pool Values KB) per Switch (

Dedicated Buffer Type	QFX5210	QFX5200-48Y	QFX5110, QFX5200-32C	QFX5100, EX4600, OCX Series	QFX3500, QFX3600
Ingress	14040	3373.50	4860.38	2912.81	2158
Egress	15184	3412.50	5408	3744	2704

Shared Buffer Configuration Recommendations for Different Network Traffic Scenarios

The way you configure the shared buffer pool depends on the mix of traffic on your network. This section provides shared buffer configuration recommendations for five basic network traffic scenarios:

- Balanced traffic—The network carries a balanced mix of unicast best-effort, lossless, and multicast traffic. (This is the default configuration.)
- Best-effort unicast traffic—The network carries mostly unicast best-effort traffic.
- Best-effort traffic with Ethernet PAUSE (IEEE 802.3X) enabled—The network carries mostly best-effort traffic with Ethernet PAUSE enabled on the links.
- Best-effort multicast traffic—The network carries mostly multicast best-effort traffic.
- Lossless traffic—The network carries mostly lossless traffic (traffic on which PFC is enabled).

NOTE: Lossless traffic is defined as traffic on which you enable PFC to ensure lossless transport. Lossless traffic does not refer to best-effort traffic on a link on which you enable Ethernet PAUSE. Start with the recommended profiles for each network traffic scenario, and adjust them if necessary for your network traffic conditions.

OCX Series switches do not support lossless transport or PFC. In this topic, references to lossless transport do not apply to OCX Series switches. OCX Series switches support symmetric Ethernet PAUSE.

CAUTION: Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete. This includes changing the default configuration to one of the recommended configurations.

Because you configure buffer allocations in percentages, the recommended allocations for each network traffic scenario are valid for all QFX Series switches, EX4600 switches, and OCX Series switches. Use one of the following recommended shared buffer configurations for your network traffic conditions. Start with a recommended configuration, then make small adjustments to the buffer allocations to fine-tune the buffers if necessary as described in “Optimizing Buffer Configuration” on page 565.

Balanced Traffic (Default Configuration)

The default shared buffer configuration is optimized for networks that carry a balanced mix of best-effort unicast, lossless, and multidestination (multicast, broadcast, and destination lookup fail) traffic. The default class-of-service (CoS) configuration is also optimized for networks that carry a balanced mix of traffic.

NOTE: On OCX Series switches, the default CoS configuration optimization does not include lossless traffic because OCX Series switches do not support lossless transport.

Except on OCX Series switches, we recommend that you use the default shared buffer configuration for networks that carry a balanced mix of traffic, especially if you are using the default CoS settings. Table 113 on page 563 shows the default ingress shared buffer allocations:

Table 113: Default Ingress Shared Buffer Configuration

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
100%	9%	45%	46%

Table 114 on page 563 shows the default egress shared buffer allocations:

Table 114: Default Egress Shared Buffer Configuration

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	50%	31%	19%

Best-Effort Unicast Traffic

If your network carries mostly best-effort (lossy) unicast traffic, then the default shared buffer configuration allocates too much buffer space to support lossless transport. Instead of wasting those buffers, we recommend that you use the following ingress shared buffer settings (see Table 115 on page 563) and egress shared buffer settings (see Table 116 on page 563):

Table 115: Recommended Ingress Shared Buffer Configuration for Networks with Mostly Best-Effort Unicast Traffic

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
100%	5%	0%	95%

Table 116: Recommended Egress Shared Buffer Configuration for Networks with Mostly Best-Effort Unicast Traffic

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	5%	75%	20%

See "Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic" on page 570 for an example that shows you how to configure the recommended buffer settings shown in Table 115 on page 563 and Table 116 on page 563.

Ethernet PAUSE Traffic

If your network carries mostly best-effort (lossy) traffic *and* enables Ethernet PAUSE on links, then the default shared buffer configuration allocates too much buffer space to the shared ingress buffer (Ethernet PAUSE traffic uses the dedicated buffers instead of shared buffers) and not enough space to the lossless-headroom buffers. We recommend that you use the following ingress shared buffer settings (see Table 117 on page 564) and egress shared buffer settings (see Table 118 on page 564):

Table 117: Recommended Ingress Shared Buffer Configuration for Networks with Mostly Best-Effort Traffic and Ethernet PAUSE Enabled

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
70%	5%	80%	15%

Table 118: Recommended Egress Shared Buffer Configuration for Networks with Mostly Best-Effort Traffic and Ethernet PAUSE Enabled

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	5%	75%	20%

See "Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links with Ethernet PAUSE Enabled" on page 576 for an example that shows you how to configure the recommended buffer settings shown in Table 115 on page 563 and Table 116 on page 563.

Best-Effort Multicast (Multidestination) Traffic

If your network carries mostly best-effort (lossy) multicast traffic, then the default shared buffer configuration allocates too much buffer space to support lossless transport. Instead of wasting those buffers, we recommend that you use the following ingress shared buffer settings (see Table 119 on page 564) and egress shared buffer settings (see Table 120 on page 564):

Table 119: Recommended Ingress Shared Buffer Configuration for Networks with Mostly Best -Effort Multicast Traffic

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
100%	5%	0%	95%

Table 120: Recommended Egress Shared Buffer Configuration for Networks with Mostly Best-Effort Multicast Traffic

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	5%	20%	75%

See "Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic" on page 583 for an example that shows you how to configure the recommended buffer settings shown in Table 119 on page 564 and Table 120 on page 564.

Lossless Traffic

If your network carries mostly lossless traffic, then the default shared buffer configuration allocates too much buffer space to support best-effort traffic. Instead of wasting those buffers, we recommend that you use the following ingress shared buffer settings (see Table 121 on page 565) and egress shared buffer settings (see Table 122 on page 565):

Table 121: Recommended Ingress Shared Buffer Configuration for Networks with Mostly Lossless Traffic

Total Shared Ingress Buffer	Lossless Buffer	Lossless-Headroom Buffer	Lossy Buffer
100%	15%	80%	5%

Table 122: Recommended Egress Shared Buffer Configuration for Networks with Mostly Lossless Traffic

Total Shared Egress Buffer	Lossless Buffer	Lossy Buffer	Multicast Buffer
100%	90%	5%	5%

See *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic* for an example that shows you how to configure the recommended buffer settings shown in Table 121 on page 565 and Table 122 on page 565.

Optimizing Buffer Configuration

Starting from the default configuration or from a recommended buffer configuration, you can further optimize the buffer allocation to best support the mix of traffic on your network. Adjust the settings gradually to fine-tune the shared buffer allocation. Use caution when adjusting the shared buffer configuration, not just when you fine-tune the ingress and egress buffer partitions, but also when you fine-tune the total ingress and egress shared buffer percentage. (Remember that if you allocate less than 100 percent of the available buffers to the shared buffers, the remaining buffers are added to the dedicated buffers). Tuning the buffers incorrectly can cause problems such as ingress port congestion.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

The relationship between the sizes of the ingress buffer pool and the egress buffer pool affects when and where packets are dropped. The buffer pool sizes include the shared buffers and the dedicated buffers. In general, if there are more ingress buffers than egress buffers, the switch can experience ingress port congestion because egress queues fill before ingress queues can empty.

Use the show class-of-service shared-buffer operational command to see the sizes in kilobytes (KB) of the dedicated and shared buffers and of the shared buffer partitions.

For best-effort traffic (unicast and multidestination), the combined ingress lossy shared buffer partition and ingress dedicated buffers must be *less than* the combined egress lossy and multicast shared buffer partitions plus the egress dedicated buffers. This prevents ingress port congestion by ensuring that egress best-effort buffers are deeper than ingress best-effort buffers, and ensures that if packets are dropped, they are dropped at the egress queues. (Packets dropping at the ingress prevents the egress schedulers from working properly.)

For lossless traffic (traffic on which you enable PFC), the combined ingress lossless shared buffer partition and a reasonable portion of the ingress headroom buffer partition, plus the dedicated buffers, must be *less than* the total egress lossless shared buffer partition and dedicated buffers. (A reasonable portion of the ingress headroom buffer is approximately 20 to 25 percent of the buffer space, but this varies depending on how much buffer headroom is required to support the lossless traffic.) When these conditions are met, if there is ingress port congestion, the ingress port congestion triggers PFC on the ingress port to prevent packet loss. If the total lossless ingress buffers exceed the total lossless egress buffers, packets could be dropped at the egress instead of PFC being applied at the ingress to prevent packet loss.

NOTE: If you commit a buffer configuration for which the switch does not have sufficient resources, the switch might log an error instead of returning a commit error. In that case, a syslog message is displayed on the console. For example:

```
user@host# commit
configuration check succeeds

Message from syslogd@host at Jun 13 11:11:10  ...
host dc-pfe: Not enough Ingress Lossless headroom.(Already allocated
more). Dedicated : 14340 Lossy : 47100 Lossless 4239 Headroom 21195
Avail : 20781
commit complete
```

If the buffer configuration commits but you receive a syslog message that indicates the configuration cannot be implemented, you can:

- Reconfigure the buffers or reconfigure other parameters (for example, the PFC configuration, which affects the need for lossless headroom buffers and lossless buffers—the more priorities you pause, the more lossless and lossless headroom buffer space you need), then attempt the commit operation again.
- Roll back the switch to the last successful configuration.

If you receive a syslog message that says the buffer configuration cannot be implemented, you must take corrective action. If you do not fix the configuration or roll back to a previous successful configuration, the system behavior is unpredictable.

General Buffer Configuration Rules and Considerations

Keep the following rules and considerations in mind when you configure the buffers:

- Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.
- If you configure the ingress or egress shared buffer percentages as less than 100 percent, the remaining percentage of buffer space is added to the dedicated buffer pool.
- The sum of all of the ingress shared buffer partitions must equal 100 percent. Each partition must be configured with a value of at least 5 percent except the lossless headroom buffer, which can have a value of 0 percent.
- The sum of all of the egress shared buffer partitions must equal 100 percent. Each partition must be configured with a value of at least 5 percent.
- Lossless and lossless headroom shared buffers serve traffic on which you enable PFC, and do not serve traffic subject to Ethernet PAUSE.
- The switch uses the dedicated buffer pool first and the shared buffer pool only after the dedicated buffer pool for a port or queue is exhausted.
- Too little dedicated buffer space results in too much competition for shared buffer space.
- Too much dedicated buffer space results in poorer burst absorption because there is less available shared buffer space.
- Always check the syslog messages after you commit a new buffer configuration.
- The optimal buffer configuration for your network depends on the types of traffic on the network. If your network carries less traffic of a certain type (for example, lossless traffic), then you can reduce the size of the buffers allocated to that type of traffic (for example, you can reduce the sizes of the lossless and lossless headroom buffers).

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links with Ethernet PAUSE Enabled on page 576
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Example: Configuring Queue Schedulers on page 293
- Configuring Global Ingress and Egress Shared Buffers on page 568

Configuring Global Ingress and Egress Shared Buffers

Although the switch reserves some buffer space to ensure a minimum memory allocation for ports and queues, you can configure how the system uses the rest of the buffer space to optimize the buffer allocation for your particular mix of network traffic. The global shared buffer pool is memory space that all of the ports on the switch share dynamically as they need buffers. You can allocate global shared memory space to different types of ingress and egress buffers to better support different mixes of network traffic.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

Use the default shared buffer settings (for a network with a balanced mix of lossless, best-effort, and multicast traffic) or one of the recommended shared buffer configurations for your mix of network traffic (mostly best-effort unicast traffic, mostly best-effort traffic on links enabled for Ethernet PAUSE, mostly multicast traffic, or mostly lossless traffic). Either the default configuration or one of the recommended configurations provides a buffer allocation that satisfies the needs of most networks.

After starting from one of the recommended configurations, you can fine-tune the shared buffer settings, but do so with caution to prevent traffic loss due to buffer misconfiguration.

You can configure the percentage of available (user-configurable) buffer space allocated to the global shared buffers. Any space that you do not allocate to the global shared buffer pool is added to the dedicated buffer pool. The default configuration allocates 100 percent of the available buffer space to the global shared buffers.

You can partition the ingress and egress shared buffer pools to allocate more buffers to the types of traffic your network predominantly carries, and fewer buffers to other traffic. From the buffer space allocated to the ingress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless ingress traffic. The minimum value for the lossless buffers is 5 percent.
- Lossless headroom buffers—Percentage of shared buffer pool for packets received while a pause is asserted. If Ethernet PAUSE is configured on a port or if priority-based flow control (PFC) is configured on priorities on a port, when the port sends a pause message to the connected peer, the port uses the headroom buffers to store the packets that arrive between the time the port sends the pause message and the time the last packet arrives after the peer pauses traffic. The minimum value for the lossless headroom buffers is 0 (zero) percent. (Lossless headroom buffers are the only buffers that can have a minimum value of less than 5 percent.)
- Lossy buffers—Percentage of shared buffer pool for all best-effort ingress traffic (best-effort unicast, multidestination, and strict-high priority traffic). The minimum value for the lossy buffers is 5 percent.

The combined percentage values of the ingress lossless, lossless headroom, and lossy buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All ingress buffer partitions must be explicitly configured, even when the lossless headroom buffer partition has a value of 0 (zero) percent.

From the buffer space allocated to the egress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless egress queues. The minimum value for the lossless buffers is 5 percent.
- Lossy buffers—Percentage of shared buffer pool for all best-effort egress queues (best-effort unicast and strict-high priority queues). The minimum value for the lossy buffers is 5 percent.
- Multicast buffers—Percentage of shared buffer pool for all multidestination (multicast, broadcast, and destination lookup fail) egress queues. The minimum value for the multicast buffers is 5 percent.

The combined percentage values of the egress lossless, lossy, and multicast buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All egress buffer partitions must be explicitly configured and must have a value of at least 5 percent.

To configure the shared buffer allocation and partitioning using the CLI:

1. Configure the percentage of available (nonreserved) buffers used for the ingress global shared buffer pool:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set ingress percent percent
   ```

2. Configure the global ingress buffer partitions for lossless, lossless-headroom, and lossy traffic:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set ingress buffer-partition lossless percent percent
   user@switch# set ingress buffer-partition lossless-headroom percent percent
   user@switch# set ingress buffer-partition lossy percent percent
   ```

3. Configure the percentage of available (nonreserved) buffers used for the egress global shared buffer pool:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set egress percent percent
   ```

4. Configure the global egress buffer partitions for lossless, lossy, and multicast queues:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set egress buffer-partition lossless percent percent
   ```



```
user@switch# set egress buffer-partition lossy percent percent
user@switch# set egress buffer-partition multicast percent percent
```

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links with Ethernet PAUSE Enabled on page 576
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Understanding CoS Buffer Configuration on page 546

Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic

Although the switch reserves some buffer space to ensure a minimum memory allocation for ports and queues, you can configure how the system uses the rest of the buffer space to optimize the buffer allocation for your particular mix of network traffic.

This example shows you the recommended configuration of the global shared buffer pool to support a network that carries mostly best-effort (lossy) unicast traffic. The global shared buffer pool is memory space that all of the ports on the switch share dynamically as they need buffers. You can allocate global shared memory space to different types of buffers to better support different mixes of network traffic.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

Use the default shared buffer settings (for a network with a balanced mix of lossless, best effort, and multicast traffic) or one of the recommended shared buffer configurations for your mix of network traffic (mostly best-effort unicast traffic, mostly best-effort traffic on links enabled for Ethernet PAUSE, mostly multicast traffic, or mostly lossless traffic). Either the default configuration or one of the recommended configurations provides a buffer allocation that satisfies the needs of most networks.

NOTE: **OCX Series switches do not support lossless transport.**

After starting from the recommended configuration, you can fine-tune the shared buffer settings, but do so with caution to prevent traffic loss due to buffer misconfiguration.

- Requirements on page 571
- Overview on page 571

- Configuration on page 573
- Verification on page 574

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 12.3 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

You can configure the percentage of available (user-configurable) buffer space allocated to the global shared buffers. Any space that you do not allocate to the global shared buffer pool is added to the dedicated buffer pool. The default configuration allocates 100 percent of the available buffer space to the global shared buffers.

You can partition the ingress and egress shared buffer pools to allocate more buffers to the types of traffic your network predominantly carries, and fewer buffers to other traffic. From the buffer space allocated to the ingress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless ingress traffic. The minimum value for the lossless buffers is 5 percent.
- Lossless headroom buffers—Percentage of shared buffer pool for packets received while a pause is asserted. If Ethernet PAUSE is configured on a port or if priority-based flow control (PFC) is configured on priorities on a port, when the port sends a pause message to the connected peer, the port uses the headroom buffers to store the packets that arrive between the time the port sends the pause message and the time the last packet arrives after the peer pauses traffic. The minimum value for the lossless headroom buffers is 0 (zero) percent. (Lossless headroom buffers are the only buffers that can have a minimum value of less than 5 percent.)
- Lossy buffers—Percentage of shared buffer pool for all best-effort ingress traffic (best-effort unicast, multidestination, and strict-high priority traffic). The minimum value for the lossy buffers is 5 percent.

The combined percentage values of the ingress lossless, lossless headroom, and lossy buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All ingress buffer partitions must be explicitly configured, even when the lossless headroom buffer partition has a value of 0 (zero) percent.

From the buffer space allocated to the egress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless egress queues. The minimum value for the lossless buffers is 5 percent.

- Lossy buffers—Percentage of shared buffer pool for all best-effort egress queues (best-effort unicast, and strict-high priority queues). The minimum value for the lossy buffers is 5 percent.
- Multicast buffers—Percentage of shared buffer pool for all multidestination (multicast, broadcast, and destination lookup fail) egress queues. The minimum value for the multicast buffers is 5 percent.

The combined percentage values of the egress lossless, lossy, and multicast buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All egress buffer partitions must be explicitly configured and must have a value of at least 5 percent.

To configure the shared buffers to support a network that carries mostly best-effort unicast traffic, more buffer space needs to be allocated to lossy buffers, and less buffer space should be allocated to lossless buffers. This example shows you how to configure the global shared buffer pool allocation that we recommend to support a network that carries mostly unicast traffic.

Topology

Table 123 on page 572 shows the configuration components for this example.

Table 123: Components of the Recommended Shared Buffer Configuration for Best-Effort Unicast Network Topologies

Component	Settings
Hardware	QFX3500 switch
Ingress shared buffer	Percentage of available ingress buffer space allocated to the ingress shared buffer: 100% Percentage of ingress buffer space allocated to lossless traffic (lossless buffer partition): 5% Percentage of ingress buffer space allocated to lossless headroom traffic (lossless-headroom buffer partition): 0% Percentage of ingress buffer space allocated to best-effort traffic (lossy buffer partition): 95%
Egress shared buffer	Percentage of available egress buffer space allocated to the egress shared buffer: 100% Percentage of egress buffer space allocated to lossless queues (lossless buffer partition): 5% Percentage of egress buffer space allocated to best-effort queues (lossy buffer partition): 75% Percentage of egress buffer space allocated to multicast traffic (multicast buffer partition): 20%

Configuration

CLI Quick Configuration

To quickly configure the recommended shared buffer settings for networks that carry mostly best-effort unicast traffic, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the **[edit class-of-service shared-buffer]** hierarchy level:

```
[edit class-of-service shared-buffer]
set ingress percent 100
set ingress buffer-partition lossless percent 5
set ingress buffer-partition lossless-headroom percent 0
set ingress buffer-partition lossy percent 95
set egress percent 100
set egress buffer-partition lossless percent 5
set egress buffer-partition lossy percent 75
set egress buffer-partition multicast percent 20
```

Configuring the Global Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic

Step-by-Step Procedure

To configure the global ingress and egress shared buffer allocations and partitions for a network that carries mostly best-effort unicast traffic:

1. Configure the percentage of available (nonreserved) buffers used for the ingress global shared buffer pool:

```
[edit class-of-service shared-buffer]
user@switch# set ingress percent 100
```

2. Configure the global ingress buffer partitions for lossless, lossless-headroom, and lossy traffic:

```
[edit class-of-service shared-buffer]
user@switch# set ingress buffer-partition lossless percent 5
user@switch# set ingress buffer-partition lossless-headroom percent 0
user@switch# set ingress buffer-partition lossy percent 95
```

3. Configure the percentage of available (nonreserved) buffers used for the egress global shared buffer pool:

```
[edit class-of-service shared-buffer]
user@switch# set egress percent 100
```

4. Configure the global egress buffer partitions for lossless, lossy, and multicast queues:

```
[edit class-of-service shared-buffer]
user@switch# set egress buffer-partition lossless percent 5
user@switch# set egress buffer-partition lossy percent 75
user@switch# set egress buffer-partition multicast percent 20
```

Results

Display the results of the configuration:

```
root@dcbg-tp-pa-02> show configuration class-of-service shared-buffer
ingress {
    percent 100;
    buffer-partition lossless {
        percent 5;
    }
    buffer-partition lossy {
        percent 95;
    }
    buffer-partition lossless-headroom {
        percent 0;
    }
}
egress {
    percent 100;
    buffer-partition lossless {
        percent 5;
    }
    buffer-partition lossy {
        percent 75;
    }
    buffer-partition multicast {
        percent 20;
    }
}
```

Verification

Verify that you correctly configured the shared buffer.

Verifying the Shared Buffer Configuration

Purpose Verify that the ingress and egress global shared buffer pools are correctly configured and partitioned among the shared buffer types.

Action List the global shared buffer configuration using the operational mode command **show class-of-service shared-buffer**:

```
user@switch> show class-of-service shared-buffer
root@dcbg-tp-pa-02> show class-of-service shared-buffer
Ingress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2158.00 KB
  Shared Buffer    :  7202.00 KB
    Lossless          :  360.10 KB
    Lossless Headroom :  0.00 KB
    Lossy             :  6841.90 KB

  Lossless Headroom Utilization:
  Node Device         Total          Used                 Free
  0                   0.00 KB        0.00 KB              0.00 KB
```

```
Egress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2704.00 KB
  Shared Buffer    :  6656.00 KB
    Lossless          :  332.80 KB
    Multicast         :  1331.20 KB
    Lossy             :  4992.00 KB
```

Meaning The **show class-of-service shared-buffer** operational command shows all of the ingress and egress global shared buffer settings, including the buffer partitioning.

For the ingress shared buffers, the command output shows:

- The total switch buffer pool is 9360 KB (9 MB).
- The dedicated buffer pool is 2158 KB. This is the size of the global ingress dedicated buffer pool when you configure the ingress shared buffer pool as 100 percent of the available (user-configurable) buffer space. This is the minimum size of the reserved, ingress dedicated ingress buffer pool (not user-configurable). If you configure the shared buffer as less than 100 percent of the available buffer pool, the remaining buffer space is added to the dedicated buffer pool.
- With the ingress shared buffer pool configured as 100 percent of the available buffers, the total size of the ingress shared buffer pool is 7202 KB.
- The ingress shared buffer pool is partitioned to allocate:
  - 360.10 KB to lossless traffic
  - No space to lossless headroom traffic
  - 6841.90 KB to lossy unicast traffic
- The Lossless Headroom Utilization field shows how much of the buffer space reserved for paused traffic is used. Because the lossless headroom buffer partition is set to 0 (zero) percent, the total amount of lossless headroom buffer space is 0 KB; therefore the amount of used and free lossless headroom buffer space is also 0 KB.

For the egress shared buffers, the command output shows:

- The total switch buffer pool is 9360 KB (9 MB).
- The dedicated buffer pool is 2704 KB. This is the size of the global egress dedicated buffer pool when you configure the egress shared buffer pool as 100 percent of the available (user-configurable) buffer space. This is the minimum size of the reserved, egress dedicated buffer pool (not user-configurable). If you configure the shared buffer as less than 100 percent of the available buffer pool, the remaining buffer space is added to the dedicated buffer pool.
- With the egress shared buffer pool configured as 100 percent of the available buffers, the total size of the egress shared buffer pool is 6656 KB. This is less than the ingress shared buffer pool because the switch reserves more egress dedicated buffer space

than ingress dedicated buffer space. (More dedicated buffer space means less shared buffer space, and more shared buffer space means less dedicated buffer space.)

- The egress shared buffer pool is partitioned to allocate:
  - 332.80 KB to lossless traffic
  - 1331.20 KB to multicast traffic
  - 4992 KB to lossy unicast traffic

NOTE: The output values are valid for QFX3500 and QFX3600 switches. QFX5100, EX4600, and OCX Series switches have larger buffers (12 MB instead of 9 MB), so the total buffer size and the sizes of each buffer partition are larger on those switches.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links with Ethernet PAUSE Enabled on page 576
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links with Ethernet PAUSE Enabled

Although the switch reserves some buffer space to ensure a minimum memory allocation for ports and queues, you can configure how the system uses the rest of the buffer space to optimize the buffer allocation for your particular mix of network traffic.

This example shows you the recommended configuration of the global shared buffer pool to support a network that carries mostly best-effort (lossy) traffic on links with Ethernet PAUSE (IEEE 802.3X) enabled.

NOTE: OCX Series switches support symmetric Ethernet PAUSE flow control, but do not support asymmetric Ethernet PAUSE flow control.

The global shared buffer pool is memory space that all of the ports on the switch share dynamically as they need buffers. You can allocate global shared memory space to different types of buffers to better support different mixes of network traffic.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

Use the default shared buffer settings (for a network with a balanced mix of lossless, best effort, and multicast traffic) or one of the recommended shared buffer configurations for your mix of network traffic (mostly best-effort unicast traffic, mostly best-effort traffic on links enabled for Ethernet PAUSE, mostly multicast traffic, or mostly lossless traffic). Either the default configuration or one of the recommended configurations provides a buffer allocation that satisfies the needs of most networks.

After starting from the recommended configuration, you can fine-tune the shared buffer settings, but do so with caution to prevent traffic loss due to buffer misconfiguration.

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 12.3 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

You can configure the percentage of available (user-configurable) buffer space allocated to the global shared buffers. Any space that you do not allocate to the global shared buffer pool is added to the dedicated buffer pool. The default configuration allocates 100 percent of the available buffer space to the global shared buffers.

You can partition the ingress and egress shared buffer pools to allocate more buffers to the types of traffic your network predominantly carries, and fewer buffers to other traffic. From the buffer space allocated to the ingress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless ingress traffic. The minimum value for the lossless buffers is 5 percent.
- Lossless headroom buffers—Percentage of shared buffer pool for packets received while a pause is asserted. If Ethernet PAUSE is configured on a port or if priority-based flow control (PFC) is configured on priorities on a port, when the port sends a pause message to the connected peer, the port uses the headroom buffers to store the packets that arrive between the time the port sends the pause message and the time the last packet arrives after the peer pauses traffic. The minimum value for the lossless

headroom buffers is 0 (zero) percent. (Lossless headroom buffers are the only buffers that can have a minimum value of less than 5 percent.)

NOTE: OCX Series switches do not support PFC.

- Lossy buffers—Percentage of shared buffer pool for all best-effort ingress traffic (best-effort unicast, multidestination, and strict-high priority traffic). The minimum value for the lossy buffers is 5 percent.

The combined percentage values of the ingress lossless, lossless headroom, and lossy buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All ingress buffer partitions must be explicitly configured, even when the lossless headroom buffer partition has a value of 0 (zero) percent.

From the buffer space allocated to the egress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless egress queues. The minimum value for the lossless buffers is 5 percent.
- Lossy buffers—Percentage of shared buffer pool for all best-effort egress queues (best-effort unicast and strict-high priority queues). The minimum value for the lossy buffers is 5 percent.
- Multicast buffers—Percentage of shared buffer pool for all multidestination (multicast, broadcast, and destination lookup fail) egress queues. The minimum value for the multicast buffers is 5 percent.

The combined percentage values of the egress lossless, lossy, and multicast buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All egress buffer partitions must be explicitly configured and must have a value of at least 5 percent.

To configure the shared buffers to support a network that carries mostly best-effort traffic on links enabled for Ethernet PAUSE, more buffer space needs to be allocated to ingress dedicated port buffers, and less buffer space should be allocated to ingress shared buffers. Also, more buffer space needs to be allocated to lossless-headroom buffers, and less space to ingress lossy buffers. This example shows you how to configure the global shared buffer pool allocation that we recommend to support a network that carries mostly best-effort traffic on links enabled for Ethernet PAUSE.

Topology

Table 124 on page 578 shows the configuration components for this example.

Table 124: Components of the Recommended Shared Buffer Configuration for Best-Effort Network Topologies with Links Enabled for Ethernet PAUSE

Component	Settings
Hardware	QFX3500 switch

Table 124: Components of the Recommended Shared Buffer Configuration for Best-Effort Network Topologies with Links Enabled for Ethernet PAUSE (continued)

Component	Settings
Ingress shared buffer	Percentage of available ingress buffer space allocated to the ingress shared buffer: 70% Percentage of ingress buffer space allocated to lossless traffic (lossless buffer partition): 5% Percentage of ingress buffer space allocated to lossless headroom traffic (lossless-headroom buffer partition): 80% Percentage of ingress buffer space allocated to best-effort traffic (lossy buffer partition): 15%
Egress shared buffer	Percentage of available egress buffer space allocated to the egress shared buffer: 100% Percentage of egress buffer space allocated to lossless queues (lossless buffer partition): 5% Percentage of egress buffer space allocated to best-effort queues (lossy buffer partition): 75% Percentage of egress buffer space allocated to multicast traffic (multicast buffer partition): 20%

Configuration

CLI Quick Configuration

To quickly configure the recommended shared buffer settings for networks that carry mostly best-effort unicast traffic, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the **[edit class-of-service shared-buffer]** hierarchy level:

```
[edit class-of-service shared-buffer]
set ingress percent 70
set ingress buffer-partition lossless percent 5
set ingress buffer-partition lossless-headroom percent 80
set ingress buffer-partition lossy percent 15
set egress percent 100
set egress buffer-partition lossless percent 5
set egress buffer-partition lossy percent 75
set egress buffer-partition multicast percent 20
```

Configuring the Global Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links Enabled for Ethernet PAUSE

Step-by-Step Procedure

To configure the global ingress and egress shared buffer allocations and partitions:

1. Configure the percentage of available (nonreserved) buffers used for the ingress global shared buffer pool:

```
[edit class-of-service shared-buffer]
user@switch# set ingress percent 70
```

2. Configure the global ingress buffer partitions for lossless, lossless-headroom, and lossy traffic:

```
[edit class-of-service shared-buffer]
user@switch# set ingress buffer-partition lossless percent 5
user@switch# set ingress buffer-partition lossless-headroom percent 80
user@switch# set ingress buffer-partition lossy percent 15
```

3. Configure the percentage of available (nonreserved) buffers used for the egress global shared buffer pool:

```
[edit class-of-service shared-buffer]
user@switch# set egress percent 100
```

4. Configure the global egress buffer partitions for lossless, lossy, and multicast queues:

```
[edit class-of-service shared-buffer]
user@switch# set egress buffer-partition lossless percent 5
user@switch# set egress buffer-partition lossy percent 75
user@switch# set egress buffer-partition multicast percent 20
```

Results

Display the results of the configuration:

```
root@dcbg-tp-pa-02> show configuration class-of-service shared-buffer
ingress {
    percent 70;
    buffer-partition lossless {
        percent 5;
    }
    buffer-partition lossy {
        percent 15;
    }
    buffer-partition lossless-headroom {
        percent 80;
    }
}
egress {
    percent 100;
```

```
    buffer-partition lossless {
        percent 5;
    }
    buffer-partition lossy {
        percent 75;
    }
    buffer-partition multicast {
        percent 20;
    }
}
```

Verification

Verify that you correctly configured the shared buffer.

Verifying the Shared Buffer Configuration

Purpose Verify that the ingress and egress global shared buffer pools are correctly configured and partitioned among the shared buffer types.

Action List the global shared buffer configuration using the operational mode command **show class-of-service shared-buffer**:

```
user@switch> show class-of-service shared-buffer
root@dcbg-tp-pa-02> show class-of-service shared-buffer
Ingress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  4318.60 KB
  Shared Buffer    :  5041.40 KB
    Lossless          :  252.07 KB
    Lossless Headroom :  4033.12 KB
    Lossy             :  756.21 KB

Egress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2704.00 KB
  Shared Buffer    :  6656.00 KB
    Lossless          :  332.80 KB
    Multicast         :  1331.20 KB
    Lossy             :  4992.00 KB
```

Meaning The **show class-of-service shared-buffer** operational command shows all of the ingress and egress global shared buffer settings, including the buffer partitioning.

For the ingress shared buffers, the command output shows:

- The total switch buffer pool is 9360 KB (9 MB).
- The dedicated buffer pool is 4318.6 KB. This is the size of the global ingress dedicated buffer pool when you configure the ingress shared buffer pool as 70 percent of the available (user-configurable) buffer space.

- With the ingress shared buffer pool configured as 70 percent of the available buffers, the total size of the ingress shared buffer pool is 5041.4 KB.
- The ingress shared buffer pool is partitioned to allocate:
  - 252.07 KB to lossless traffic
  - 4033.12 KB to lossless headroom traffic
  - 756.21 KB to lossy unicast traffic

For the egress shared buffers, the command output shows:

- The total switch buffer pool is 9360 KB (9 MB).
- The dedicated buffer pool is 2704 KB. This is the size of the global egress dedicated buffer pool when you configure the egress shared buffer pool as 100 percent of the available (user-configurable) buffer space. This is the minimum size of the reserved, egress dedicated buffer pool (not user-configurable). If you configure the shared buffer as less than 100 percent of the available buffer pool, the remaining buffer space is added to the dedicated buffer pool.
- With the egress shared buffer pool configured as 100 percent of the available buffers, the total size of the egress shared buffer pool is 6656 KB. This is less than the ingress shared buffer pool because the switch reserves more egress dedicated buffer space than ingress dedicated buffer space. (More dedicated buffer space means less shared buffer space, and more shared buffer space means less dedicated buffer space.)
- The egress shared buffer pool is partitioned to allocate:
  - 332.80 KB to lossless traffic
  - 1331.20 KB to multicast traffic
  - 4992 KB to lossy unicast traffic

NOTE: **The output values are valid for QFX3500 and QFX3600 switches. QFX5100, EX4600, and OCX Series switches have larger buffers (12 MB instead of 9 MB), so the total buffer size and the sizes of each buffer partition are larger on those switches.**

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic

Although the switch reserves some buffer space to ensure a minimum memory allocation for ports and queues, you can configure how the system uses the rest of the buffer space to optimize the buffer allocation for your particular mix of network traffic.

This example shows you the recommended configuration of the global shared buffer pool to support a network that carries mostly multicast traffic. The global shared buffer pool is memory space that all of the ports on the switch share dynamically as they need buffers. You can allocate global shared memory space to different types of buffers to better support different mixes of network traffic.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

Use the default shared buffer settings (for a network with a balanced mix of lossless, best effort, and multicast traffic) or one of the recommended shared buffer configurations for your mix of network traffic (mostly best-effort unicast traffic, mostly best-effort traffic on links enabled for Ethernet PAUSE, mostly multicast traffic, or mostly lossless traffic). Either the default configuration or one of the recommended configurations provides a buffer allocation that satisfies the needs of most networks.

After starting from the recommended configuration, you can fine-tune the shared buffer settings, but do so with caution to prevent traffic loss due to buffer misconfiguration.

Requirements

This example uses the following hardware and software components:

- One switch (this example was tested on a Juniper Networks QFX3500 Switch)
- Junos OS Release 12.3 or later for the QFX Series or Junos OS Release 14.1X53-D20 or later for the OCX Series

Overview

You can configure the percentage of available (user-configurable) buffer space allocated to the global shared buffers. Any space that you do not allocate to the global shared buffer pool is added to the dedicated buffer pool. The default configuration allocates 100 percent of the available buffer space to the global shared buffers.

You can partition the ingress and egress shared buffer pools to allocate more buffers to the types of traffic your network predominantly carries, and fewer buffers to other traffic. From the buffer space allocated to the ingress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless ingress traffic. The minimum value for the lossless buffers is 5 percent.
- Lossless headroom buffers—Percentage of shared buffer pool for packets received while a pause is asserted. If Ethernet PAUSE is configured on a port or if priority-based flow control (PFC) is configured on priorities on a port, when the port sends a pause message to the connected peer, the port uses the headroom buffers to store the packets that arrive between the time the port sends the pause message and the time the last packet arrives after the peer pauses traffic. The minimum value for the lossless headroom buffers is 0 (zero) percent. (Lossless headroom buffers are the only buffers that can have a minimum value of less than 5 percent.)
- Lossy buffers—Percentage of shared buffer pool for all best-effort ingress traffic (best-effort unicast, multidestination, and strict-high priority traffic). The minimum value for the lossy buffers is 5 percent.

The combined percentage values of the ingress lossless, lossless headroom, and lossy buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All ingress buffer partitions must be explicitly configured, even when the lossless headroom buffer partition has a value of 0 (zero) percent.

From the buffer space allocated to the egress shared buffer pool, you can allocate space to:

- Lossless buffers—Percentage of shared buffer pool for all lossless egress queues. The minimum value for the lossless buffers is 5 percent.
- Lossy buffers—Percentage of shared buffer pool for all best-effort egress queues (best-effort unicast, and strict-high priority queues). The minimum value for the lossy buffers is 5 percent.
- Multicast buffers—Percentage of shared buffer pool for all multidestination (multicast, broadcast, and destination lookup fail) egress queues. The minimum value for the multicast buffers is 5 percent.

The combined percentage values of the egress lossless, lossy, and multicast buffer partitions must total exactly 100 percent. If the buffer percentages total more than 100 percent or less than 100 percent, the switch returns a commit error. All egress buffer partitions must be explicitly configured and must have a value of at least 5 percent.

To configure the shared buffers to support a network that carries mostly multicast traffic, more buffer space needs to be allocated to lossy buffers, less buffer space should be allocated to lossless buffers, and more space needs to be allocated to egress multicast buffers. This example shows you how to configure the global shared buffer pool allocation that we recommend to support a network that carries mostly multicast traffic.

Topology

Table 125 on page 585 shows the configuration components for this example.

Table 125: Components of the Recommended Shared Buffer Configuration for Multicast Network Topologies

Component	Settings
Hardware	QFX3500 switch
Ingress shared buffer	Percentage of available ingress buffer space allocated to the ingress shared buffer: 100% Percentage of ingress buffer space allocated to lossless traffic (lossless buffer partition): 5% Percentage of ingress buffer space allocated to lossless headroom traffic (lossless-headroom buffer partition): 0% Percentage of ingress buffer space allocated to best-effort traffic (lossy buffer partition): 95%
Egress shared buffer	Percentage of available egress buffer space allocated to the egress shared buffer: 100% Percentage of egress buffer space allocated to lossless queues (lossless buffer partition): 5% Percentage of egress buffer space allocated to best-effort queues (lossy buffer partition): 20% Percentage of egress buffer space allocated to multicast traffic (multicast buffer partition): 75%

Configuration

CLI Quick Configuration

To quickly configure the recommended shared buffer settings for networks that carry mostly multicast traffic, copy the following commands, paste them in a text file, remove line breaks, change variables and details to match your network configuration, and then copy and paste the commands into the CLI at the **[edit class-of-service shared-buffer]** hierarchy level:

```
[edit class-of-service shared-buffer]
set ingress percent 100
set ingress buffer-partition lossless percent 5
set ingress buffer-partition lossless-headroom percent 0
set ingress buffer-partition lossy percent 95
set egress percent 100
set egress buffer-partition lossless percent 5
set egress buffer-partition lossy percent 20
set egress buffer-partition multicast percent 75
```

Configuring the Global Shared Buffer Pool for Networks with Mostly Multicast Traffic

Step-by-Step Procedure

To configure the global ingress and egress shared buffer allocations and partitions for a network that carries mostly multicast traffic:

1. Configure the percentage of available (nonreserved) buffers used for the ingress global shared buffer pool:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set ingress percent 100
   ```

2. Configure the global ingress buffer partitions for lossless, lossless-headroom, and lossy traffic:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set ingress buffer-partition lossless percent 5
   user@switch# set ingress buffer-partition lossless-headroom percent 0
   user@switch# set ingress buffer-partition lossy percent 95
   ```

3. Configure the percentage of available (nonreserved) buffers used for the egress global shared buffer pool:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set egress percent 100
   ```

4. Configure the global egress buffer partitions for lossless, lossy, and multicast queues:

   ```
   [edit class-of-service shared-buffer]
   user@switch# set egress buffer-partition lossless percent 5
   user@switch# set egress buffer-partition lossy percent 20
   user@switch# set egress buffer-partition multicast percent 75
   ```

Results

Display the results of the configuration:

```
root@dcbg-tp-pa-02> show configuration class-of-service shared-buffer
ingress {
    percent 100;
    buffer-partition lossless {
        percent 5;
    }
    buffer-partition lossy {
        percent 95;
    }
    buffer-partition lossless-headroom {
        percent 0;
    }
}
egress {
```



```
        percent 100;
        buffer-partition lossless {
            percent 5;
        }
        buffer-partition lossy {
            percent 20;
        }
        buffer-partition multicast {
            percent 75;
        }
    }
```

Verification

Verify that you correctly configured the shared buffer.

Verifying the Shared Buffer Configuration

Purpose Verify that you correctly configured the ingress and egress global shared buffer pools and that you correctly partitioned the buffer among the shared buffer types.

Action List the global shared buffer configuration using the operational mode command **show class-of-service shared-buffer**:

```
user@switch> show class-of-service shared-buffer
root@dcbg-tp-pa-02> show class-of-service shared-buffer
Ingress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2158.00 KB
  Shared Buffer    :  7202.00 KB
    Lossless          :  360.10 KB
    Lossless Headroom :  0.00 KB
    Lossy             :  6841.90 KB

  Lossless Headroom Utilization:
  Node Device         Total          Used                  Free
  0                   0.00 KB        0.00 KB               0.00 KB

Egress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2704.00 KB
  Shared Buffer    :  6656.00 KB
    Lossless          :  332.80 KB
    Multicast         :  4992.00 KB
    Lossy             :  1331.20 KB
```

Meaning The **show class-of-service shared-buffer** operational command shows all of the ingress and egress global shared buffer settings, including the buffer partitioning.

For the ingress shared buffers, the command output shows:

- The total switch buffer pool is 9360 KB (9 MB).
- The dedicated buffer pool is 2158 KB. This is the size of the global ingress dedicated buffer pool when you configure the ingress shared buffer pool as 100 percent of the available (user-configurable) buffer space. This is the minimum size of the reserved, ingress dedicated ingress buffer pool (not user-configurable). If you configure the shared buffer as less than 100 percent of the available buffer pool, the remaining buffer space is added to the dedicated buffer pool.
- With the ingress shared buffer pool configured as 100 percent of the available buffers, the total size of the ingress shared buffer pool is 7202 KB.
- The ingress shared buffer pool is partitioned to allocate:
  - 360.10 KB to lossless traffic
  - No space to lossless headroom traffic
  - 6841.90 KB to lossy unicast traffic
- The Lossless Headroom Utilization field shows how much of the buffer space reserved for paused traffic is used. Because the lossless headroom buffer partition is set to 0 (zero) percent, the total amount of lossless headroom buffer space is 0 KB; therefore the amount of used and free lossless headroom buffer space is also 0 KB.

For the egress shared buffers, the command output shows:

- The total switch buffer pool is 9360 KB (9 MB).
- The dedicated buffer pool is 2704 KB. This is the size of the global egress dedicated buffer pool when you configure the egress shared buffer pool as 100 percent of the available (user-configurable) buffer space. This is the minimum size of the reserved, egress dedicated buffer pool (not user-configurable). If you configure the shared buffer as less than 100 percent of the available buffer pool, the remaining buffer space is added to the dedicated buffer pool.
- With the egress shared buffer pool configured as 100 percent of the available buffers, the total size of the egress shared buffer pool is 6656 KB. This is less than the ingress shared buffer pool because the switch reserves more egress dedicated buffer space than ingress dedicated buffer space. (More dedicated buffer space means less shared buffer space, and more shared buffer space means less dedicated buffer space.)
- The egress shared buffer pool is partitioned to allocate:
  - 332.80 KB to lossless traffic
  - 4992 KB to multicast traffic
  - 1331.20 KB to lossy unicast traffic

NOTE: The output values are valid for QFX3500 and QFX3600 switches. QFX5100, EX4600, and OCX Series switches have larger buffers (12 MB instead of 9 MB), so the total buffer size and the sizes of each buffer partition are larger on those switches.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Traffic on Links with Ethernet PAUSE Enabled on page 576
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

CHAPTER 13

Configuration Statements for Buffers

buffer-partition (Egress)

Syntax

```
buffer-partition (lossless | lossy | multicast) {
  percent percent;
}
```

Hierarchy Level [edit class-of-service shared-buffer egress]

Release Information Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description The egress shared buffer pool is divided into three partitions. Each partition reserves a percentage of the available shared buffer pool for a type of traffic, so that the switch provides enough resources to support a mix of best-effort, lossless, and multicast traffic (multicast also includes broadcast and destination lookup fail traffic). To better support the mix of traffic on your network, you can optimize the allocation of egress shared buffers to different types of traffic by fine-tuning the shared buffer partitions.

NOTE: **OCX Series switches do not support lossless transport.**

The percentages you configure for the three egress shared buffer partitions must total exactly 100 percent. If the total of the three shared buffer percentages is not 100 percent, the system returns a commit error and does not commit the configuration. You can configure any partition to 0 (zero) percent as long as the allocation to other partitions totals 100 percent.

This is a global allocation that applies to all ports. All ports on the switch receive the same allocation of egress shared buffers.

If you do not configure buffer partitions, the switch uses the default partitioning.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

Default The default egress buffer partition shown in Table 126 on page 592 supports networks with a balanced mix of best-effort, multicast, and lossless traffic. It is the recommended configuration if you are using the default configuration with two lossless forwarding classes.

Table 126: Default Egress Shared Buffer Partitioning

Lossless Partition	Lossy Partition	Multicast Partition
50%	31%	19%

The sum of the default percentages configured for each partition is 100 percent. The sum of the partition percentages must always total 100 percent.

Options

lossless—Shared buffer space reserved for all lossless egress traffic.

lossy—Shared buffer space for best-effort unicast egress traffic.

multicast—Shared buffer space reserved for all multicast (including broadcast and destination lookup fail) egress traffic.

percent *percent*—The percentage of buffer space to allocate to the specified buffer partition (lossless, lossy, or multicast buffers). The sum of the percentages for the three buffer partitions must total 100 percent.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

buffer-partition (Ingress)

Syntax

```
buffer-partition (lossless | lossless-headroom | lossy) {
  percent percentage;
}
```

Hierarchy Level [edit class-of-service shared-buffer ingress]

Release Information Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description The ingress shared buffer pool is divided into three partitions. Each partition reserves a percentage of the available shared buffer pool for a type of traffic, so that the switch provides enough resources to support a mix of best effort (best-effort unicast and multicast) and lossless traffic. To better support the mix of traffic on your network, you can optimize the allocation of ingress shared buffers to different types of traffic by fine-tuning the shared buffer partitions.

NOTE: **OCX Series switches do not support lossless transport.**

The percentages you configure for the three ingress shared buffer partitions must total exactly 100 percent. If the total of the three shared buffer percentages is not 100 percent, the system returns a commit error and does not commit the configuration. You can configure any partition to 0 (zero) percent as long as the allocation to other partitions totals 100 percent.

This is a global allocation that applies to all ingress traffic. All ports on the switch receive the same allocation of ingress shared buffers.

If you do not configure buffer partitions, the switch uses the default partitioning.

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

Default The default ingress buffer partition shown in Table 127 on page 594 supports networks with a balanced mix of best-effort, multicast, and lossless traffic. It is the recommended configuration if you are using the default configuration with two lossless forwarding classes.

Table 127: Default Ingress Shared Buffer Partitioning

Lossless Partition	Lossless-Headroom Partition	Lossy Partition
9%	45%	46%

The sum of the default percentages configured for each partition is 100 percent. The sum of the partition percentages always must total 100 percent.

Options

lossless—Shared buffer space reserved for all lossless ingress traffic.

lossless-headroom—Shared buffer space reserved to store packets received while either an 802.3x Ethernet PAUSE or a priority-based flow control (PFC) pause is asserted. (When an ingress interface pauses traffic, it must have the buffer space to store all of the packets currently in the buffer, and also all of the packets received before the connected peer stops sending traffic and the wire is cleared of packets.)

lossy—Shared buffer space for best-effort ingress traffic.

percent ***percent***—The percentage of buffer space to allocate to the specified buffer partition (lossless, lossless-headroom, or lossy buffers). The sum of the percentages for the three buffer partitions must total 100 percent.

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

buffer-size

Syntax buffer-size (percent *percent* | remainder);

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description On all switches, you configure the proportion of port buffers allocated to a particular output queue using the following process:

1. Configure a scheduler and set the **buffer-size** option.

2. Use a scheduler map to map the scheduler to the forwarding class that is mapped to the queue to which you want to apply the buffer size.

   For example, suppose that you want to change the dedicated buffer allocation for FCoE traffic. FCoE traffic is mapped to the fcoe forwarding class, and the fcoe forwarding class is mapped to queue 3 (this is the default configuration). To use default FCoE traffic mapping, in the scheduler map configuration, map the scheduler to the **fcoe** forwarding class.

3. If you are using enhanced transmission selection (ETS) hierarchical scheduling, associate the scheduler map with the traffic control profile you want to use on the egress ports that carry FCoE traffic. If you are using direct port scheduling, skip this step.

4. If you are using ETS, associate the traffic control profile that includes the scheduler map with the desired egress ports. For this example, you associate the traffic control profile with the ports that carry FCoE traffic. If you are using port scheduling, associate the scheduler map with the desired egress ports.

   Queue 3, which is mapped to the fcoe forwarding class and therefore to the FCoE traffic, receives the dedicated buffer allocation specified in the **buffer-size** statement.

NOTE: **The total of all of the explicitly configured buffer size percentages for all of the queues on a port cannot exceed 100 percent.**

QFX10000 Switches

On QFX10000 switches, the buffer size is the amount of time in milliseconds of port bandwidth that a queue can use to continue to transmit packets during periods of congestion, before the buffer runs out and packets begin to drop.

The switch can use up to 100 ms total (combined) buffer space for all queues on a port. A buffer-size configured as one percent is equal to 1 ms of buffer usage. A buffer-size of 15 percent (the default value for the best effort and network control queues) is equal to 15 ms of buffer usage.

The total buffer size of the switch is 4 GB. A 40-Gigabit port can use up to 500 MB of buffer space, which is equivalent to 100 ms of port bandwidth on a 40-Gigabit port. A 10-Gigabit port can use up to 125 MB of buffer space, which is equivalent to 100 ms of port bandwidth on a 10-Gigabit port. The total buffer sizes of the eight output queues on a port cannot exceed 100 percent, which is equal to the full 100 ms total buffer available to a port. The maximum amount of buffer space any queue can use is also 100 ms (which equates to a 100 percent buffer-size configuration), but if one queue uses all of the buffer, then no other queue receives buffer space.

There is no minimum buffer allocation, so you can set the buffer-size to zero (0) for a queue. However, we recommend that on queues on which you enable PFC to support lossless transport, you allocate a minimum of 5 ms (a minimum buffer-size of 5 percent). The two default lossless queues, fcoe and no-loss, have buffer-size default values of 35 ms (35 percent).

Queue buffer allocation is dynamic, shared among ports as needed. However, a queue cannot use more than its configured amount of buffer space. For example, if you are using the default CoS configuration, the best-effort queue receives a maximum of 15 ms of buffer space because the default transmit rate for the best-effort queue is 15 percent.

If a switch experiences congestion, queues continue to receives their full buffer allocation until 90 percent of the 4 GB buffer space is consumed. When 90 percent of the buffer space is in use, the amount of buffer space per port, per queue, is reduced in proportion to the configured buffer size for each queue. As the percentage of consumed buffer space rises above 90 percent, the amount of buffer space per port, per queue, continues to be reduced.

On 40-Gigabit ports, because the total buffer is 4 GB and the maximum buffer a port can use is 500 MB, up to seven 40-Gigabit ports can consume their full 100 ms allocation of buffer space. However, if an eighth 40-Gigabit port requires the full 500 MB of buffer space, then the buffer allocations are proportionally reduced because the buffer consumption is above 90 percent.

On 10-Gigabit ports, because the total buffer is 4 GB and the maximum buffer a port can use is 125 MB, up to 28 10-Gigabit ports can consume their full 100 ms allocation of buffer space. However, if a 29th 10-Gigabit port requires the full 125 MB of buffer space, then the buffer allocations are proportionally reduced because the buffer consumption is above 90 percent.

QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems

Set the dedicated buffer size of the egress queue that you bind the scheduler to in the scheduler map configuration. The switch allocates space from the global dedicated buffer pool to ports and queues in a hierarchical manner. The switch allocates an equal number of dedicated buffers to each egress port, so each egress port receives the same amount of dedicated buffer space. The amount of dedicated buffer space per port is not configurable.

However, the **buffer-size** statement allows you to control the way each port allocates its share of dedicated buffers to its queues. For example, if a port only uses two queues to forward traffic, you can configure the port to allocate all of its dedicated buffer space to those two ports and avoid wasting buffer space on queues that are not in use. We recommend that the buffer size should be the same size as the minimum guaranteed transmission rate (the **transmit-rate**).

Default The default behavior of the differs on different switches.

QFX10000 Switches

If you do not configure buffer-size and you do not explicitly configure a queue scheduler, the default buffer-size is the default transmit rate of the queue. If you explicitly configure a queue scheduler, the default buffer allocations are not used. If you explicitly configure a queue scheduler, configure the buffer-size for each queue in the scheduler, keeping in mind that the total buffer-size of the queues cannot exceed 100 percent (100 ms).

Table 78 on page 403 shows the default queue buffer sizes on QFX10000 switches. The default buffer size is the same as the default transmit rate for each default queue:

Table 128: Default Output Queue Buffer Sizes (QFX10000 Switches)

Queue Number	Forwarding Class	Transmit Rate	Buffer Size
0	best-effort	15%	15%
3	fcoe	35%	35%
4	no-loss	35%	35%
7	network-control	15%	15%

By default, only the queues mapped to the default forwarding classes receive buffer space from the port buffer pool. (Buffers are not wasted on queues that do not carry traffic.)

QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems

The port allocates dedicated buffers to queues that have an explicitly configured scheduler buffer size. If you do not explicitly configure a scheduler buffer size for a queue, the port serves the explicitly configured queues first. Then the port divides the remaining dedicated buffers equally among the queues that have an explicitly attached scheduler *without* an explicitly configured buffer size configuration. (If you configure a scheduler, but you do not configure the buffer size parameter, the default is equivalent to configuring the buffer size with the **remainder** option.)

If you use the default scheduler and scheduler map on a port (no explicit scheduler configuration), then the port allocates its dedicated buffer pool to queues based on the default scheduling. Table 79 on page 404 shows the default queue buffer sizes. The default buffer size is the same as the default transmit rate for each default queue:

Table 129: Default Output Queue Buffer Sizes (QFX5100, EX4600, QFX3500, and QFX3600 Switches, and QFabric Systems)

Queue Number	Forwarding Class	Transmit Rate	Buffer Size
0	best-effort	5%	5%
3	fcoe	35%	35%
4	no-loss	35%	35%
7	network-control	5%	5%
8	mcast	20%	20%

By default, only the queues mapped to the default forwarding classes receive buffer space from the port buffer pool. (Buffers are not wasted on queues that do not carry traffic.)

NOTE: OCX Series switches do not support lossless transport. On OCX Series switches, do not map traffic to the lossless default fcoe and no-loss forwarding classes. OCX Series default DSCP classification does not map traffic to the fcoe and no-loss forwarding classes, so by default, the OCX system does not classify traffic into those forwarding classes. (On other switches, the fcoe and no-loss forwarding classes provide lossless transport for Layer 2 traffic. OCX Series switches do not support lossless Layer 2 transport.) The active forwarding classes (**best-effort**, **network-control**, and **mcast**) share the unused bandwidth assigned to the **fcoe** and **no-loss** forwarding classes.

Options

percent *percent*—Percentage of the port dedicated buffer pool allocated to the queue (or queues) mapped to the scheduler.

remainder—Remaining dedicated buffer pool after the port satisfies the needs of the explicitly configured buffers. The port divides the remaining buffers equally among the queues that are explicitly attached to a scheduler but that do not have an explicit buffer size configuration (or are configured with **remainder** as the buffer size).

Required Privilege Level

interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

egress (Buffer Configuration)

Syntax

```
egress {
  percent percent;
  buffer-partition (lossless | lossy | multicast) {
    percent percent;
  }
}
```

Hierarchy Level [edit class-of-service shared-buffer]

Release Information Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure the global shared buffer pool allocation for egress traffic. The system allocates the shared buffer pool dynamically across its ports as the ports require memory space. Some buffer space is reserved for other buffers such as dedicated buffers (buffers allocated permanently to ports).

The percentage you specify is the percentage of available (user-configurable) buffer space allocated to the global shared egress buffer pool. If you allocate less than 100 percent of the available buffer space to the shared buffer pool, the remaining buffer space is added to the dedicated buffer pool. (You cannot directly configure the dedicated buffer pool for each port; dedicated buffers are allocated evenly across all the ports. However, on a port, you can configure the portion of dedicated port buffer space allocated to each queue in the scheduler configuration using the **buffer-size** option.)

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

You can also partition the shared buffer pool to adjust the egress buffer allocations for different mixes of network traffic using the **buffer-partition** statement.

Default The default shared buffer percentage is 100 percent. (All available buffer space is allocated to the shared buffer pool.)

Options **percent *percent***—Percentage of available egress buffer space allocated to the shared buffer pool. If the percentage is less than 100 percent, the remaining buffer space is allocated to the dedicated buffer pool.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

ingress (Buffer Configuration)

Syntax

```
ingress {
  buffer-partition (lossless | lossless-headroom | lossy) {
    percent percent;
  }
  percent percent;
}
```

Hierarchy Level [edit class-of-service shared-buffer]

Release Information Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description Configure the global shared buffer pool allocation for ingress traffic. The system allocates the shared buffer pool dynamically across its ports as the ports require memory space. Some buffer space is reserved for buffers such as dedicated buffers (buffers allocated permanently to ports) and headroom buffers (buffers that help prevent packet loss on lossless flows).

The percentage you specify is the percentage of available (user-configurable) buffer space allocated to the global shared ingress buffer pool. If you allocate less than 100 percent of the available buffer space to the shared buffer pool, the remaining buffer space is added to the dedicated buffer pool. (You cannot directly configure the dedicated buffer pool for each port; dedicated buffers are allocated evenly across all the ports.)

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until buffer reprogramming is complete.**

You can also partition the shared buffer pool to adjust the ingress buffer allocations for different mixes of network traffic using the **buffer-partition** statement.

Default The default shared buffer percentage is 100 percent. (All available buffer space is allocated to the shared buffer pool.)

Options **percent *percent***—Percentage of available ingress buffer space allocated to the shared buffer pool. If the percentage is less than 100 percent, the remaining buffer space is allocated to the dedicated buffer pool.

The remaining statement is explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

shared-buffer

Syntax

```
shared-buffer {
  egress {
    buffer-partition (lossless | lossy | multicast) {
      percent percent
    }
    percent percent;
  }
  ingress {
    percent percent;
    buffer-partition (lossless | lossless-headroom | lossy) {
      percent percent
    }
  }
}
```

Hierarchy Level

[edit class-of-service]

Release Information

Statement introduced in Junos OS Release 12.3 for the QFX Series.
Statement introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Configure the global shared buffer pool allocation to ports. Shared buffers are a pool of buffer space that the system can allocate dynamically across all of its ports as memory space is needed. Some buffer space is reserved for dedicated buffers (buffers allocated permanently to ports), headroom buffers (buffers that help prevent packet loss on lossless flows), and other buffers.

Configure the way the system uses the available (user-configurable) buffer space by setting the **shared-buffer** percentage for the ingress buffer pool and for the egress buffer pool.

The percentage you specify is the percentage of available buffer space allocated to the global shared ingress buffer pool or to the global shared egress buffer pool. If you allocate less than 100 percent of the available buffer space to the shared buffer pool, the remaining buffer space is added to the dedicated buffer pool. (You cannot directly configure the dedicated buffer pool for each port; dedicated buffers are allocated evenly across all the ports.)

CAUTION: **Changing the buffer configuration is a disruptive event. Traffic stops on *all* ports until the buffer reprogramming is complete.**

You can also partition the ingress shared buffer pool and the egress shared buffer pool to adjust the buffer allocations for different mixes of network traffic (best-effort, lossless, multicast) using the **buffer-partition** statement.

NOTE: If you commit a buffer configuration for which the switch does not have sufficient resources, the switch might log an error instead of returning a commit error. In that case, a syslog message is displayed on the console. For example:

```
user@host# commit
configuration check succeeds

Message from syslogd@host at Jun 13 11:11:10  ...
host dc-pfe: Not enough Ingress Lossless headroom.(Already allocated
more). Dedicated : 14340 Lossy : 47100 Lossless 4239 Headroom 21195
Avail : 20781
commit complete
```

If the buffer configuration commits but you receive a syslog message that indicates the configuration cannot be implemented, you can:

- Reconfigure the buffers or reconfigure other parameters (for example, the PFC configuration, which affects the need for lossless headroom buffers and lossless buffers—the more priorities you pause, the more lossless and lossless headroom buffer space you need), then attempt the commit operation again.
- Roll back the switch to the last successful configuration.

If you receive a syslog message that says the buffer configuration cannot be implemented, you must take corrective action. If you do not fix the configuration or roll back to a previous successful configuration, the system behavior is unpredictable.

Options The remaining statements are explained separately. See CLI Explorer.

Required Privilege Level interfaces—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

CHAPTER 14

Monitoring Commands for Buffers

show class-of-service shared-buffer

Syntax

show class-of-service shared-buffer
<egress | ingress>

Release Information

Command introduced in Junos OS Release 12.3 for the QFX Series.
Command introduced in Junos OS Release 14.1X53-D20 for the OCX Series.

Description

Display the shared buffer allocation and partitioning configuration.

NOTE: Due to QFX5200 cross-point architecture, all buffer usage counters are maintained separately. When usage counters are displayed with the command show class-of-service shared-buffer on QFX5200, various pipe counters are displayed separately.

Options

none—Display ingress and egress shared buffer settings.

egress—(Optional) Display the egress shared buffer settings.

ingress—(Optional) Display the ingress shared buffer settings.

Required Privilege Level

view

Related Documentation

- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Best-Effort Unicast Traffic on page 570
- Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Multicast Traffic on page 583
- *Example: Recommended Configuration of the Shared Buffer Pool for Networks with Mostly Lossless Traffic*
- Configuring Global Ingress and Egress Shared Buffers on page 568
- Understanding CoS Buffer Configuration on page 546

List of Sample Output

show class-of-service shared-buffer on page 609

Output Fields

Table 130 on page 609 describes the output fields for the **show class-of-service shared-buffer** command. Output fields are listed in the approximate order in which they appear.

Table 130: show class-of-service shared-buffer Output Fields

Field Name	Field Description
Ingress	Ingress shared buffer configuration.
Total Buffer	Total buffer space available to the ports in KB. This is the combined dedicated buffer pool and shared buffer pool.
Dedicated Buffer	Buffer space allocated to the dedicated buffer pool in KB.
Shared Buffer	Buffer space allocated to the shared buffer pool in KB.
Lossless	Buffer space allocated to the lossless traffic buffer pool in KB.
Lossless Headroom	Buffer space allocated to the lossless headroom traffic buffer pool to support priority-based flow control (PFC) and Ethernet PAUSE in KB. (Ingress ports only.) NOTE: OCX Series switches do not support PFC.
Lossy	Buffer space allocated to the lossy (best-effort) traffic buffer pool in KB.
Lossless Headroom Utilization	Utilization of the ingress lossless headroom buffer pool. (These fields can help you to determine how much headroom buffer space you need to reserve to support PFC and Ethernet PAUSE for lossless flows.) NOTE: OCX Series switches do not support PFC.
Node Device	Index number that identifies the switch. On a QFX3500 switch, this field always has a value of zero (0).
Total	Size of the lossless headroom ingress buffer pool in KB.
Used	Amount in KB of lossless headroom ingress buffer used.
Free	Amount in KB of lossless headroom ingress buffer free (unused).
Egress	Egress shared buffer configuration.
Multicast	Buffer space allocated to the multicast traffic buffer pool in KB. (Egress ports only.)

Sample Output

show class-of-service shared-buffer

```
user@switch> show class-of-service shared-buffer
Ingress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2158.00 KB
  Shared Buffer    :  7202.00 KB
    Lossless          :  648.18 KB
```

```
    Lossless Headroom :  3240.90 KB
    Lossy             :  3312.92 KB

  Lossless Headroom Utilization:
  Node Device         Total          Used                 Free
  0                   3240.90 KB     0.00 KB              3240.90 KB

Egress:
  Total Buffer     :  9360.00 KB
  Dedicated Buffer :  2704.00 KB
  Shared Buffer    :  6656.00 KB
    Lossless          :  3328.00 KB
    Multicast         :  1264.64 KB
    Lossy             :  2063.36 KB
```