JUNIPER NETWORKS

Junos® OS

CoS on Enhanced Queuing DPCs

Release
13.1

Published: 2013-02-13

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

Junos® OS CoS on Enhanced Queuing DPCs
13.1

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

YEAR 2000 NOTICE

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

END USER LICENSE AGREEMENT

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

Table of Contents

List of Tables

About the Documentation

Documentation and Release Notes

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

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

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

Supported Platforms

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

- MX Series

Using the Examples in This Manual

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

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

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

Merging a Full Example

To merge a full example, follow these steps:

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

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

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

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

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

Merging a Snippet

To merge a snippet, follow these steps:

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

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

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

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

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

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

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

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

Documentation Conventions

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

Table 1: Notice Icons

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

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

Table 2: Text and Syntax Conventions

Convention	Description	Examples
Bold text like this	Represents text that you type.	To enter configuration mode, type the**configure** command: user@host> **configure**
`Fixed-width text like this`	Represents output that appears on the terminal screen.	`user@host>` **show chassis alarms** `No alarms currently active`

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

Convention	Description	Examples
Italic text like this	• Introduces or emphasizes important new terms. • Identifies book names. • Identifies RFC and Internet draft titles.	• A policy *term* is a named structure that defines match conditions and actions. • *Junos OS System Basics Configuration 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)	Enclose 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)	Enclose a variable for which you can substitute one or more values.	**community name members [** ***community-ids*** **]**

Indention and braces ({ })	Identify 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.	
J-Web GUI Conventions		
Bold text like this	Represents J-Web graphical user interface (GUI) items you click or select.	• In the Logical Interfaces box, select **All Interfaces**. • To cancel the configuration, click **Cancel**.
> (bold right angle bracket)	Separates levels in a hierarchy of J-Web selections.	In the configuration editor hierarchy, select **Protocols>Ospf**.

Documentation Feedback

We encourage you to provide feedback, comments, and suggestions so that we can improve the documentation. You can send your comments to techpubs-comments@juniper.net, or fill out the documentation feedback form at https://www.juniper.net/cgi-bin/docbugreport/ . If you are using e-mail, be sure to include the following information with your comments:

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

Requesting Technical Support

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

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

Self-Help Online Tools and Resources

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

- Find CSC offerings: http://www.juniper.net/customers/support/
- Search for known bugs: http://www2.juniper.net/kb/
- Find product documentation: http://www.juniper.net/techpubs/
- Find solutions and answer questions using our Knowledge Base: http://kb.juniper.net/
- Download the latest versions of software and review release notes: http://www.juniper.net/customers/csc/software/
- Search technical bulletins for relevant hardware and software notifications: https://www.juniper.net/alerts/

- Join and participate in the Juniper Networks Community Forum: http://www.juniper.net/company/communities/
- Open a case online in the CSC Case Management tool: http://www.juniper.net/cm/

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

Opening a Case with JTAC

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

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

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

PART 1

Overview

CHAPTER 1

CoS on Enhanced Queuing DPCs

Enhanced Queuing DPC Hardware Properties

On a Juniper Networks MX Series 3D Universal Edge Router with Enhanced Queuing Dense Port Concentrators (DPCs), you can configure schedulers and queues. You can configure 15 VLAN sets per Gigabit Ethernet (1G) port and 255 VLAN sets per 10-Gigabit Ethernet (10G) port. The Enhanced Queuing DPC performs priority propagation from one hierarchy level to another and drop statistics are available on the Enhanced Queuing DPC per color per queue instead or just per queue.

Juniper Networks MX Series 3D Universal Edge Routers with Enhanced Queuing DPCs have Packet Forwarding Engines that can support up to 515 MB of frame memory, and packets are stored in 512-byte frames. Table 3 on page 3 compares the major properties of the Intelligent Queuing 2 (IQ2) PIC and the Packet Forwarding Engine within the Enhanced Queuing DPC.

Table 3: IQ2 PIC and Enhanced Queuing DPC Compared

Feature	IQ2 PIC	Packet Forwarding Engine Within Enhanced Queuing DPC
Number of usable queues	8,000	16,000
Number of shaped logical interfaces	1,000 with 8 queues each.	2,000 with 8 queues each, or 4,000 with 4 queues each.
Number of hardware priorities	2	4
Priority propagation	No	Yes
Dynamic mapping	No: schedulers/port are fixed.	Yes: schedulers/port are not fixed.
Drop statistics	Per queues	Per queue per color (PLP high, low)

In addition, the Enhanced Queuing DPC features support for hierarchical weighted random early detection (WRED) and enhanced queuing on aggregated Ethernet interfaces with link protection as well.

The Enhanced Queuing DPC supports the following hierarchical scheduler characteristics:

- Shaping at the physical interface level
- Shaping and scheduling at the service VLAN interface set level
- Shaping and scheduling at the customer VLAN logical interface level
- Scheduling at the queue level

VLAN (Level 3) shaping on a 10-Gigabit Ethernet MX Series Enhanced Queuing DPC differs from the VLAN (Level3) shaping on a 1-Gigabit Ethernet Enhanced Queuing DPC. To use the VLAN (Level 3) shaping on a 10-Gigabit Ethernet MX Series Enhanced Queuing DPC, configure an interface set at the **[edit interfaces interface-set]** hierarchy level. The interface set configuration is not required for configuring a 1-Gigabit Ethernet VLANs on the same Enhanced Queuing DPC.

The Enhanced Queuing DPC supports the following features for scalability:

- 16,000 queues per Packet Forwarding Engine
- 4 Packet Forwarding Engines per DPC
  - 4000 schedulers at logical interface level (Level 3) with 4 queues each
  - 2000 schedulers at logical interface level (Level 3) with 8 queues each
- 255 schedulers at the interface set level (Level 2) per 1-port Packet Forwarding Engine on a 10-Gigabit Ethernet DPC
- 15 schedulers at the interface set level (Level 2) per 10-port Packet Forwarding Engine on a 1-Gigabit Ethernet DPC
- About 400 milliseconds of buffer delay (this varies by packet size and if large buffers are enabled)
- 4 levels of priority (strict-high, high, medium, and low)

NOTE: Including the **transmit-rate** *rate* **exact** statement at the **[edit class-of-service schedulers** *scheduler-name***]** hierarchy level is not supported on Enhanced Queuing DPCs on MX Series routers.

The way that the Enhanced Queuing DPC maps a queue to a scheduler depends on whether 8 queues or 4 queues are configured. By default, a scheduler at level 3 has 4 queues. Level 3 scheduler X controls queue X*4 to X*4+3, so that scheduler 100 (for example) controls queues 400 to 403. However, when 8 queues per scheduler are enabled, the odd numbered schedulers are disabled, allowing twice the number of queues per subscriber as before. With 8 queues, level 3 scheduler X controls queue X*4 to X*4+7, so that scheduler 100 (for example) now controls queues 400 to 407.

You configure the **max-queues-per-interface** statement to set the number of queues at **4** or **8** at the FPC level of the hierarchy. Changing this statement results in a restart of the DPC. For more information about the **max-queues-per-interface** statement, see the Junos® OS Network Interfaces.

The Enhanced Queuing DPC maps level 3 (customer VLAN) schedulers in groups to level 2 (service VLAN) schedulers. Sixteen contiguous level 3 schedulers are mapped to level 2 when 4 queues are enabled, and 8 contiguous level 3 schedulers are mapped to level 2 when 8 queues are enabled. All of the schedulers in the group should use the same queue priority mapping. For example, if the queue priorities of one scheduler are high, medium, low, and low, then all members of this group should have the same queue priority.

Mapping of a group at level 3 to level 2 can be done at any time. However, a group at level 3 can only be unmapped from a level 2 scheduler only if all the schedulers in the group are free. Once unmapped, a level 3 group can be remapped to any level 2 scheduler. There is no restriction on the number of level 3 groups that can be mapped to a particular level 2 scheduler. There can be 256 level 3 groups, but fragmentation of the scheduler space can reduce the number of schedulers available. In other words, there are scheduler allocation patterns that might fail even though there are free schedulers.

In contrast to level-3-to-level-2 mapping, the Enhanced Queuing DPC maps level 2 (service VLAN) schedulers in a fixed mode to level 1 (physical interface) schedulers. On 40-port Gigabit Ethernet DPCs, there are 16 level 1 schedulers, and 10 of these are used for the physical interfaces. There are 256 level 2 schedulers, or 16 per level 1 scheduler. A level 1 scheduler uses level schedulers X*16 through X*16+15. So level 1 scheduler 0 uses level 2 schedulers 0 through 15, level 1 scheduler 1 uses level 2 schedulers 16 through 31, and so on. On 4-port 10-Gigabit Ethernet PICs, there is one level 1 scheduler for the physical interface, and 256 level 2 schedulers are mapped to the single level 1 scheduler.

The maximum number of level 3 (customer VLAN) schedulers that can be used is 4076 (4 queues) or 2028 (8 queues) for the 10-port Gigabit Ethernet Packet Forwarding Engine and 4094 (4 queues) or 2046 (8 queues) for the 10-Gigabit Ethernet Packet Forwarding Engine.

Enhanced Queuing is supported on aggregated Ethernet (AE) interfaces with two links in link protection mode. However, only one link in the AE bundle can be active at a time. Traffic is shaped independently on the two links, but the member's links do not need to reside in the same Packet Forwarding Engine or the same DPC. Finally, shared schedulers are not supported on the Enhanced Queuing DPC (use hierarchical schedulers to group logical interfaces).

Related Documentation

- Configuring Customer VLAN (Level 3) Shaping on Enhanced Queuing DPCs on page 21

PART 2

Configuration

CHAPTER 2

Configuration Tasks

Configuring Rate Limits on Enhanced Queuing DPCs

You can rate-limit the strict-high and high queues on the Enhanced Queuing DPC. Without rate limits, traffic in higher priority queues can block the transmission of lower priority packets. Unless limited, higher priority traffic is always sent before lower priority traffic, causing the lower priority queues to "starve" and cause timeouts and unnecessarily resent packets.

On the Enhanced Queuing DPC, you can rate-limit queues before the packets are queued for output. All packets exceeding the configured rate limit are dropped, so care is required when establishing this limit. This model is also supported on IQ2 PICs. For more information about configuring CoS on IQ2 PICs, see CoS on Enhanced IQ2 PICs Overview.

NOTE: Rate limiting is implemented differently on Enhanced Queuing DPCs and non-queuing Packet Forwarding Engines. On Enhanced Queuing DPCs, rate-limiting is implemented using a single rate two color policer. On non-queuing Packet Forwarding Engines, rate-limiting is achieved by shaping the queue to the transmit rate and keeping the queue delay buffers small to prevent too many packets from being queued once the shaping rate is reached.

To rate-limit queues, include the **transmit-rate** statement with the **rate-limit** option at the **[edit class-of-service schedulers *scheduler-name*]** hierarchy level:

```
[edit class-of-service schedulers scheduler-name]
transmit-rate rate rate-limit;
```

This example limits the transmit rate of a strict-high expedited-forwarding queue to 1 Mbps. The scheduler and scheduler map are defined, and then applied to the traffic at the **[edit interfaces]** and **[edit class-of-service]** hierarchy levels:

```
[edit class-of-service]
schedulers {
   scheduler-1 {
      transmit-rate 1m rate-limit; # This establishes the limit
      priority strict-high;
   }
}
scheduler-maps {
   scheduler-map-1 {
      forwarding-class expedited-forwarding scheduler scheduler-1;
   }
}

[edit interfaces]
s0-2/2/0 {
   per-unit-scheduler;
   encapsulation frame-relay;
   unit 0 {
      dlci 1;
   }
}

[edit class-of-service]
interfaces {
   so-2/2/0 {
      unit 0 {
         scheduler-map scheduler-map-1;
         shaping-rate 2m;
      }
   }
}
```

You can issue the following operational mode commands to verify your configuration (the first shows the rate limit in effect):

- **show class-of-service scheduler-map** *scheduler-map-name*
- **show class-of-service interface** *interface-name*

Configuring Simple Filters on Enhanced Queuing DPCs

You can configure and apply a simple filter to perform multifield classification on the ingress interfaces of an MX Series router with Enhanced Queuing DPCs. These simple filters can be used to override default CoS classification parameters such as forwarding class or loss priority. Simple filters, in contrast to other firewall filters, only support a subset of the full firewall filter syntax.

To configure a simple filter, include the **simple-filter** statement at the **[edit firewall family inet]** hierarchy level:

```
[edit firewall family inet]
simple-filter filter-name {
   term term-name {
      from {
         ... match-conditions...
      }
      then {
         forwarding-class class-name;
         loss-priority priority;
      }
   }
}
```

The following example configures a simple filter to detect ingress packets from various source addresses (**10.1.1.1/32**, **10.10.10.10/32**, and **10.4.0.0/8**), destination addresses (**10.6.6.6/32**), protocols (**tcp**), and source ports (**400-500**, **http**). The filter then assigns various forwarding classes and loss priorities to the filtered traffic. Finally, the filter is applied to the input side of an Enhanced Queuing DPC interface (**ge-2/3/3**).

```
[edit]
firewall {
   family inet {
      simple-filter sf-for-eq-dpc {
         term 1 {
            from {
               source-address 10.1.1.1/32;
               protocol tcp;
            }
            then loss-priority low;
         }
         term 2 {
            from {
               source-address 10.4.0.0/8;
               source-port http;
            }
            then loss-priority high;
         }
         term 3 {
            from {
               destination-address 10.6.6.6/32;
               source-port 400-500;
            }
            then {
               loss-priority low;
               forwarding-class best-effort;
            }
         }
         term 4 {
            from {
               forwarding-class expedited-forwarding;
               source-address 10.10.10.10/32;
            }
            then loss-priority low;
```

```
        }
        term 5 {
          from {
            source-address 10.10.10.10/32;
          }
          then loss-priority low;
        }
      }
    }
  }
interfaces { # Apply the simple filter above to the input side of the interface.
ge-2/3/3 {
  unit 0 {
    family inet {
      simple-filter {
        input sf-for-eq-dpc;
      }
    }
  }
}
```

Related Documentation

- Simple Filter Overview
- How Simple Filters Evaluate Packets
- Guidelines for Configuring Simple Filters
- Guidelines for Applying Simple Filters

Configuring WRED on Enhanced Queuing DPCs

Shaping to drop out-of-profile traffic is done on the Enhanced Queuing DPC at all levels but the queue level. However, weighed random early discard (WRED) is done at the queue level with much the same result. With WRED, the decision to drop or send the packet is made before the packet is placed in the queue.

WRED shaping on the Enhanced Queuing DPC is similar to the IQ2 PIC, but involves only two levels, not 64. The probabilistic drop region establishes a minimum and a maximum queue depth. Below the minimum queue depth, the drop probability is 0 (send). Above the maximum level, the drop probability is 100 (certainty).

There are four drop profiles associated with each queue. These correspond to each of four loss priorities (low, medium-low, medium-high, and high). Sixty-four sets of four drop profiles are available (32 for ingress and 32 for egress). In addition, there are eight WRED scaling profiles in each direction.

To configure WRED, include the **drop-profiles** statement at the **[edit class-of-service]** hierarchy level:

```
[edit class-of-service]
drop-profiles {
  profile-name {
    fill-level percentage drop-probability percentage;
  }
```

```
}
```

The following example is an Enhanced Queuing DPC drop profile for expedited forwarding traffic:

```
[edit class-of-service drop-profiles]
drop-ef {
   fill-level 20 drop-probability 0; # Minimum Q depth
   fill-level 100 drop-probability 100; # Maximum Q depth
}
```

Note that only two fill levels can be specified for the Enhanced Queuing DPC. You can configure the **interpolate** statement, but only two fill levels are used. The **delay-buffer-rate** statement in the traffic control profile determines the maximum queue size. This delay buffer rate is converted to a packet delay buffers, where one buffer is equal to 512 bytes. For example, at 10 Mbps, the Enhanced Queuing DPC allocates 610 delay buffers when the delay-buffer rate is set to 250 milliseconds. The WRED threshold values are specified in terms of absolute buffer values.

The WRED scaling factor multiples all WRED thresholds (both minimum and maximum) by the value specified. There are eight values in all: 1, 2, 4, 8, 16, 32, 64, and 128. The WRED scaling factor is chosen to best match the user-configured drop profiles. This is done because the hardware supports only certain values of thresholds (all values must be a multiple of 16). So if the configured value of a threshold is 500 (for example), the multiple of 16 is 256 and the scaling factor applied is 2, making the value 512, which allows the value of 500 to be used. If the configured value of a threshold is 1500, the multiple of 16 is 752 and the scaling factor applied is 2, making the value 1504, which allows the value of 1500 to be used

Hierarchical RED is used to support the oversubscription of the delay buffers (WRED is only configured at the queue, physical interface, and PIC level). Hierarchical RED works with WRED as follows:

- If any level accepts the packet (the queue depth is less than the minimum buffer level), then this level accepts the packet.
- If any level probabilistically drops the packet, then this level drops the packet.

However, these rules might lead to the accepting of packets under loaded conditions which might otherwise have been dropped. In other words, the logical interface accepts packets if the physical interface is not congested.

Related Documentation

- Shaping Granularity Values for Enhanced Queuing Hardware
- For more information about configuring RED drop profiles, see RED Drop Profiles Overview.

Configuring MDRR on Enhanced Queuing DPCs

The guaranteed rate (CIR) at the interface set level is implemented using modified deficit round-robin (MDRR). The Enhanced Queuing DPC hardware provides four levels of strict priority. There is no restriction on the number of queues for each priority. MDRR is used among queues of the same priority. Each queue has one priority when it is under the

guaranteed rate and another priority when it is over the guaranteed rate but under the shaping rate (PIR). The Enhanced Queuing DPC hardware implements the priorities with 256 service profiles. Each service profile assigns eight priorities for eight queues. One set is for logical interfaces under the guaranteed rate and another set is for logical interfaces over the guaranteed rate but under the shaping rate. Each service profile is associated with a group of 16 level 3 schedulers, so there is a unique service profile available for all 256 groups at level 3, giving 4096 logical interfaces.

The Junos OS provides three priorities for traffic under the guaranteed rate and one reserved priority for traffic over the guaranteed rate that is not configurable. The Junos OS provides three priorities when there is no guaranteed rate configured on any logical interface.

The relationship between Junos OS priorities and the Enhanced Queuing DPC hardware priorities below and above the guaranteed rate (CIR) is shown in Table 4 on page 14.

Table 4: Junos Priorities Mapped to Enhanced Queuing DPC Hardware Priorities

Junos OS Priority	Enhanced Queuing DPC Hardware Priority Below Guaranteed Rate	Enhanced Queuing DPC Hardware Priority Above Guaranteed Rate
Strict-high	High	High
High	High	Low
Medium-high	Medium-high	Low
Medium-low	Medium-high	Low
Low	Medium-low	Low

To configure MDRR, configure a scheduler at the **[edit class-of-service schedulers]** hierarchy level:

```
[edit class-of-service schedulers]
scheduler-name {
   buffer-size (seconds | percent percentage | remainder | temporal microseconds);
   priority priority-level;
   transmit-rate (percent percentage | rate | remainder) <exact | rate-limit>;
}
```

The following example creates two schedulers for MDRR:

```
[edit class-of-service schedulers]
best-effort-scheduler {
   transmit-rate percent 30; # if no shaping rate
   buffer-size percent 30;
   priority high;
}
expedited-forwarding-scheduler {
   transmit-rate percent 40; # if no shaping rate
   buffer-size percent 40;
   priority strict-high;
}
```

NOTE: The use of both shaping rate and a guaranteed rate at the interface set level (level 2) is not supported.

MDRR is provided at three levels of the scheduler hierarchy of the Enhanced Queuing DPC with a granularity of 1 through 255. There are 64 MDRR profiles at the queue level, 16 at the interface set level, and 32 at the physical interface level.

Queue transmit rates are used for queue level MDRR profile weight calculation. The queue MDRR weight is calculated differently based on the mode set for sharing excess bandwidth. If you configure the **equal** option for excess bandwidth, then the queue MDRR weight is calculated as:

Queue weight = (255 * Transmit-rate-percentage) / 100

If you configure the **proportional** option for excess bandwidth, which is the default, then the queue MDRR weight is calculated as:

Queue weight = Queue-transmit-rate / Queue-base-rate, where

Queue-transmit-rate = (Logical-interface-rate * Transmit-rate-percentage) / 100, and

Queue-base-rate = Excess-bandwidth-proportional-rate / 255

To configure the way that the Enhanced Queuing DPC should handle excess bandwidth, configure the **excess-bandwidth-share** statement at the **[edit interface-set *interface-set-name*]** hierarchy level. By default, the excess bandwidth is set to **proportional** with a default value of 32.64 Mbps. In this mode, the excess bandwidth is shared in the ratio of the logical interface shaping rates. If set to **equal**, the excess bandwidth is shared equally among the logical interfaces.

This example sets the excess bandwidth sharing to proportional at a rate of 100 Mbps with a shaping rate of 80 Mbps.

```
[edit interface-set example-interface-set]
excess-bandwidth-share proportional 100m;
output-traffic-control-profile PIR-80Mbps;
```

Shaping rates established at the logical interface level are used to calculate the MDRR weights used at the interface set level. The 16 MDRR profiles are set to initial values, and the closest profile with rounded values is chosen. By default, the physical port MDRR weights are preset to the full bandwidth on the interface.

Configuring Excess Bandwidth Sharing

When using the Enhanced Queuing DPC on an MX Series router, there are circumstances when you should configure excess bandwidth sharing and minimum logical interface shaping. This section details some of the guidelines for configuring excess bandwidth sharing.

- Excess Bandwidth Sharing and Minimum Logical Interface Shaping on page 16
- Selecting Excess Bandwidth Sharing Proportional Rates on page 16

Excess Bandwidth Sharing and Minimum Logical Interface Shaping

The default excess bandwidth sharing proportional rate is 32.65 Mbps (128 Kbps x 255). In order to have better weighed fair queuing (WFQ) accuracy among queues, the shaping rate configured should be larger than the excess bandwidth sharing proportional rate. Some examples are shown in Table 5 on page 16.

Table 5: Shaping Rates and WFQ Weights

Shaping Rate	Configured Queue Transmit Rate	WFQ Weight	Total Weights
10 Mbps	(30, 40, 25, 5)	(22, 30, 20, 4)	76
33 Mbps	(30, 40, 25, 5)	(76, 104, 64, 13)	257
40 Mbps	(30, 40, 25, 5)	(76, 104.64, 13)	257

With a 10-Mbps shaping rate, the total weights are 76. This is divided among the four queues according to the configured transmit rate. Note that when the shaping rate is larger than the excess bandwidth sharing proportional rate of 32.65 Mbps, the total weights on the logical interface are 257 and the WFQ accuracy is the same.

Selecting Excess Bandwidth Sharing Proportional Rates

A good excess bandwidth sharing proportional rate to configure is to choose the largest CIR (guaranteed rate) among all the logical interfaces (units). If the logical units have PIRs (shaping rates) only, then choose the largest PIR rate. However, this is not ideal if a single logical interface has a large weighed round-robin (WRR) rate. This can skew the distribution of traffic across the queues of the other logical interfaces. To avoid this issue, set the excess bandwidth sharing proportional rate to a lower value on the logical interfaces where the WRR rates are concentrated. This improves the bandwidth sharing accuracy among the queues on the same logical interface. However, the excess bandwidth sharing for the logical interface with the larger WRR rate is no longer proportional.

As an example, consider five logical interfaces on the same physical port, each with four queues, all with only PIRs configured and no CIRs. The WRR rate is the same as the PIR for the logical interface. The excess bandwidth is shared proportionally with a rate of 40 Mbps. The traffic control profiles for the logical interfaces are shown in Table 6 on page 16.

Table 6: Example Shaping Rates and WFQ Weights

Shaping Rate	Configured Queue Transmit Rate	WFQ Weight	Total Weights
(Unit 0) 10 Mbps	(95, 0, 0, 5)	(60, 0, 0, 3)	63
(Unit 1) 20 Mbps	(25, 25, 25, 25)	(32, 32, 32, 32)	128

Table 6: Example Shaping Rates and WFQ Weights *(continued)*

Shaping Rate	Configured Queue Transmit Rate	WFQ Weight	Total Weights
(Unit 2) 40 Mbps	(40, 30, 20, 10)	(102, 77, 51, 26)	255
(Unit 3) 200 Mbps	(70, 10, 10, 10)	(179, 26, 26, 26)	255
(Unit 4) 2 Mbps	(25, 25, 25, 25)	(5, 5, 5, 5)	20

Even though the maximum transmit rate for the queue on logical interface unit 3 is 200 Mbps, the excess bandwidth sharing proportional rate is kept at a much lower value. Within a logical interface, this method provides a more accurate distribution of weights across queues. However, the excess bandwidth is now shared equally between unit 2 and unit 3 (total weight of each = 255).

Mapping Calculated Weights to Hardware Weights

The calculated weight in a traffic control profile is mapped to hardware weight, but the hardware only supports a limited WFQ profile. The weights are rounded to the nearest hardware weight according to the values in Table 7 on page 17.

Table 7: Rounding Configured Weights to Hardware Weights

Traffic Control Profile Number	Number of Traffic Control Profiles	Weights	Maximum Error
1–16	16	1–16 (interval of 1)	50.00%
17–29	13	18–42 (interval of 2)	6.25%
30–35	6	45–60 (interval of 3)	1.35%
36–43	8	64–92 (interval of 4)	2.25%
44–49	6	98–128 (interval of 6)	3.06%
50–56	7	136–184 (interval of 8)	3.13%
57–62	6	194–244 (interval of 10)	2.71%
63–63	1	255–255 (interval of 11)	2.05%

From the table, as an example, the calculated weight of 18.9 is mapped to a hardware weight of 18, because 18 is closer to 18.9 than 20 (an interval of 2 applies in the range 18–42).

Allocating Weight with Only Shaping Rates or Unshaped Logical Interfaces

Logical interfaces with only shaping rates (PIRs) or unshaped logical interfaces (units) are given a weight of 10. A logical interface with a small guaranteed rate (CIR) might get

an overall weight less than 10. In order to allocate a higher share of the excess bandwidth to logical interfaces with a small guaranteed rate in comparison to the logical interfaces with only shaping rates configured, a minimum weight of 20 is given to the logical interfaces with guaranteed rates configured.

For example, consider a logical interface configuration with five units, as shown in Table 8 on page 18.

Table 8: Allocating Weights with PIR and CIR on Logical Interfaces

Logical Interface (Unit)	Traffic Control Profile	WRR Percentages	Weights
Unit 1	PIR 100 Mbps	95, 0, 0, 5	10, 1, 1, 1
Unit 2	CIR 20 Mbps	25, 25, 25, 25	64, 64, 64, 64
Unit 3	PIR 40 Mbps, CIR 20 Mbps	50, 30, 15, 5	128, 76, 38, 13
Unit 4	Unshaped	95, 0, 0, 5	10, 1, 1, 1
Unit 5	CIR 1 Mbps	95, 0, 0, 5	10, 1, 1, 1

The weights for these units are calculated as follows:

- Select the excess bandwidth sharing proportional rate to be the maximum CIR among all the logical interfaces: 20 Mbps (unit 2).
- Unit 1 has a PIR and unit 4 is unshaped. The weight for these units is 10.
- The weight for unit 1 queue 0 is 9.5 (10 x 95%), which translates to a hardware weight of 10.
- The weight for unit 1 queue 1 is 0 (0 x 0%), but although the weight is zero, a weight of 1 is assigned to give minimal bandwidth to queues with zero WRR.
- Unit 5 has a very small CIR (1 Mbps), and a weight of 20 is assigned to units with a small CIR.
- The weight for unit 5 queue 0 is 19 (20 x 95%), which translates to a hardware weight of 18.
- Unit 3 has a CIR of 20 Mbps, which is the same as the excess bandwidth sharing proportional rate, so it has a total weight of 255.
- The weight of unit 3 queue 0 is 127.5 (255 x 50%), which translates to a hardware weight of 128.

Sharing Bandwidth Among Logical Interfaces

As a simple example showing how bandwidth is shared among the logical interfaces, assume that all traffic is sent on queue 0. Assume also that there is a 40-Mbps load on all of the logical interfaces. Configuration details are shown in Table 9 on page 19.

NOTE: On the MX960 router, bandwidth sharing across high priority and strict-high priority schedulers configured on logical interfaces might not be as expected. This is a hardware limitation.

Table 9: Sharing Bandwidth Among Logical Interfaces

Logical Interface (Unit)	Traffic Control Profile	WRR Percentages	Weights
Unit 1	PIR 100 Mbps	95, 0, 0, 5	10, 1, 1, 1
Unit 2	CIR 20 Mbps	25, 25, 25, 25	64, 64, 64, 64
Unit 3	PIR 40 Mbps, CIR 20 Mbps	50, 30, 15, 5	128, 76, 38, 13
Unit 4	Unshaped	95, 0, 0, 5	10, 1, 1, 1

1. When the port is shaped at 40 Mbps, because units 2 and 3 have a guaranteed rate (CIR) configured, both units 2 and 3 get 20 Mbps of shared bandwidth.
2. When the port is shaped at 100 Mbps, because units 2 and 3 have a guaranteed rate (CIR) configured, each of them can transmit 20 Mbps. On units 1, 2, 3, and 4, the 60 Mbps of excess bandwidth is shaped according to the values shown in Table 10 on page 19.

Table 10: First Example of Bandwidth Sharing

Logical Interface (Unit)	Calculation	Bandwidth
Unit 1	10 / (10+64+128+10) x 60 Mbps	2.83 Mbps
Unit 2	64 / (10+64+128+10) x 60 Mbps	18.11 Mbps
Unit 3	128 / (10+64+128+10) x 60 Mbps	36.22 Mbps
Unit 4	10 (10+64+128+10) x 60 Mbps	2.83 Mbps

However, unit 3 only has 20 Mbps extra (PIR and CIR) configured. This means that the leftover bandwidth of 16.22 Mbps (36.22 Mbps – 20 Mbps) is shared among units 1, 2, and 4. This is shown in Table 11 on page 19.

Table 11: Second Example of Bandwidth Sharing

Logical Interface (Unit)	Calculation	Bandwidth
Unit 1	10 / (10+64+128+10) x 16.22 Mbps	1.93 Mbps
Unit 2	64 / (10+64+128+10) x 16.22 Mbps	12.36 Mbps
Unit 4	10 (10+64+128+10) x 16.22 Mbps	1.93 Mbps

Finally, Table 12 on page 20 shows the resulting allocation of bandwidth among the logical interfaces when the port is configured with a 100-Mbps shaping rate.

Table 12: Final Example of Bandwidth Sharing

Logical Interface (Unit)	Calculation	Bandwidth
Unit 1	2.83 Mbps + 1.93 Mbps	4.76 Mbps
Unit 2	20 Mbps + 18.11 Mbps + 12.36 Mbps	50.47 Mbps
Unit 3	20 Mbps + 20 Mbps	40 Mbps
Unit 4	2.83 Mbps + 1.93 Mbps	4.76 Mbps

Configuring Ingress Hierarchical CoS on Enhanced Queuing DPCs

You can configure ingress CoS parameters, including hierarchical schedulers, on MX Series routers with Enhanced Queuing DPCs. In general, the supported configuration statements apply to per-unit schedulers or to hierarchical schedulers.

To configure ingress CoS for per-unit schedulers, include the following statements at the **[edit class-of-service interfaces *interface-name*]** hierarchy level:

```
[edit class-of-service interfaces interface-name]
input-excess-bandwidth-share (proportional value | equal);
input-scheduler-map map-name;
input-shaping-rate rate;
input-traffic-control-profile profile-name shared-instance instance-name;
unit logical-unit-number;
  input-scheduler-map map-name;
  input-shaping-rate (percent percentage | rate);
  input-traffic-control-profile profile-name shared-instance instance-name;
}
```

To configure ingress CoS for hierarchical schedulers, include the **interface-set *interface-set-name*** statement at the **[edit class-of-service interfaces]** hierarchy level:

```
[edit class-of-service interfaces]
interface-set interface-set-name {
  input-excess-bandwidth-share (proportional value | equal);
  input-traffic-control-profile profile-name shared-instance instance-name;
  input-traffic-control-profile-remaining profile-name;
  interface interface-name {
    input-excess-bandwidth-share (proportional value | equal);
    input-traffic-control-profile profile-name shared-instance instance-name;
    input-traffic-control-profile-remaining profile-name;
    unit logical-unit-number;
      input-traffic-control-profile profile-name shared-instance instance-name;
    }
  }
}
```

By default, ingress CoS features are disabled on the Enhanced Queuing DPC.

For an Enhanced Queuing (EQ) DPC on an MX Series router, CoS queuing and scheduling are enabled on the egress side but disabled on the ingress side by default. To enable ingress CoS on the EQ DPC, you must configure the **traffic-manager** statement with **ingress-and-egress** mode:

```
[edit chassis fpc slot-number pic pic-number]
traffic-manager mode ingress-and-egress;
```

Configured CoS features on the ingress are independent of CoS features on the egress, with the following exceptions:

- If you configure a per-unit or hierarchical scheduler at the **[edit class-of-service interfaces]** hierarchy level, the schedulers apply in both the ingress and egress directions.
- You cannot configure the same logical interface on an ingress and an egress interface set. A logical interface can only belong to one interface set.
- The DPC's frame buffer of 512 MB is shared between ingress and egress configurations.

The following behavior aggregate (BA) classification tables are supported on the ingress side of the Enhanced Queuing DPC:

- inet-precedence
- DSCP
- exp (MPLS)
- DSCP for IPv6
- IEEE 802.1p

Configuring Customer VLAN (Level 3) Shaping on Enhanced Queuing DPCs

Customer VLAN (level 3) shaping on an MX Series 10-Gigabit Ethernet Enhanced Queuing DPC differs from the customer VLAN (level 3) shaping on an MX Series 1-Gigabit Ethernet Enhanced Queuing DPC. To use the customer VLAN (level 3) shaping on an MX Series 10-Gigabit Ethernet Enhanced Queuing DPC, configure an interface set at the **[edit interfaces interface-set]** hierarchy level. You do not need to configure the interface set while using customer VLAN (level 3) on an MX Series 1-Gigabit Ethernet Enhanced Queuing DPC.

To configure customer VLAN (level 3) shaping on an MX Series 10-Gigabit Ethernet Enhanced Queuing DPC:

1. Configure the interface set at the **[edit interfaces]** hierarchy level.

   ```
   [edit interfaces]
   user@host# set interface-set jnpr interface unit 100
   user@host# set interface-set jnpr interface xe-1/0/0 unit 101
   ```

2. Configure the hierarchical scheduler and enable VLAN tagging.

   ```
   [edit interfaces]
   user@host# set xe-1/0/0 hierarchical-scheduler
   user@host# set xe-1/0/0 vlan-tagging
   ```

3. Configure the logical interface properties.

```
[edit interfaces]
user@host# set xe-1/0/0 unit 100 vlan-id 100
user@host# set xe-1/0/0 unit 100 family inet address 10.1.0.1/24
user@host# set xe-1/0/0 unit 101 vlan-id 101
user@host# set xe-1/0/0 unit 101 family inet address 10.1.1.1/24
```

4. Configure the traffic control profiles at the **[edit class-of-service]** hierarchy level.

```
[edit class-of-service]
user@host# set traffic-control-profiles profile1 shaping-rate 10g burst-size 2k
user@host# set traffic-control-profiles profile1 guaranteed-rate 10g burst-size 2k
user@host# set traffic-control-profiles profile2 shaping-rate 50m burst-size 2k
user@host# set traffic-control-profiles profile2 guaranteed-rate 50m burst-size 2k
user@host# set traffic-control-profiles profile3 shaping-rate 80m burst-size 3k
user@host# set traffic-control-profiles profile3 guaranteed-rate 80m burst-size 3k
```

5. Configure the output traffic control profiles at the **[edit class-of-service interfaces]** hierarchy level.

```
[edit class-of-service interfaces]
user@host# set interface-set jnpr output-traffic-control-profiles profile1
user@host# set xe-1/0/0 unit 100 output-traffic-control-profiles profile2
user@host# set xe-1/0/0 unit 101 output-traffic-control-profiles profile3
```

To configure customer VLAN (level 3) shaping on an MX Series 1-Gigabit Ethernet Enhanced Queuing DPC:

1. Configure the interface set at the **[edit interfaces]** hierarchy level.

```
[edit interfaces]
user@host# set interface ge-1/0/0 unit 100
user@host# set interface ge-1/0/0 unit 101
```

2. Configure the hierarchical scheduler and enable the VLAN tagging.

```
[edit interfaces]
user@host# set ge-1/0/0 hierarchical-scheduler
user@host# set ge-1/0/0 vlan-tagging
```

3. Configure the logical interface properties.

```
[edit interfaces]
user@host# set ge-1/0/0 unit 100 vlan-id 100
user@host# set ge-1/0/0 unit 100 family inet address 10.1.0.1/24
user@host# set ge-1/0/0 unit 101 vlan-id 101
user@host# set ge-1/0/0 unit 101 family inet address 10.1.1.1/24
```

4. Configure the traffic control profiles at the **[edit class-of-service]** hierarchy level.

```
[edit class-of-service]
user@host# set traffic-control-profiles profile1 shaping-rate 10g burst-size 2k
user@host# set traffic-control-profiles profile1 guaranteed-rate 10g burst-size 2k
user@host# set traffic-control-profiles profile2 shaping-rate 50m burst-size 2k
user@host# set traffic-control-profiles profile2 guaranteed-rate 50m burst-size 2k
user@host# set traffic-control-profiles profile3 shaping-rate 80m burst-size 3k
user@host# set traffic-control-profiles profile3 guaranteed-rate 80m burst-size 3k
```


5. Configure the traffic control profiles at the **[edit class-of-service interfaces]** hierarchy level.

```
[edit class-of-service interfaces]
user@host# set ge-1/0/0 unit 100 output-traffic-control-profiles profile2
user@host# set ge-1/0/0 unit 101 output-traffic-control-profiles profile3
```

Related Documentation

- Enhanced Queuing DPC Hardware Properties on page 3

CHAPTER 3

Configuration Statements

buffer-size (Schedulers)

Syntax buffer-size (percent *percentage* | remainder | temporal *microseconds*);

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced in Junos OS Release 12.1X48 for PTX Series Packet Transport Switches.
Statement introduced in Junos OS Release 12.2 for ACX Series Routers.

Description Specify buffer size.

NOTE: On PTX Series Packet Transport Switches, buffer-size cannot be configured on rate-limited queues.

Default If you do not include this statement, the default scheduler transmission rate and buffer size percentages for queues 0 through 7 are 95, 0, 0, 5, 0, 0, 0, and 0 percent, respectively.

Options **percent *percentage***—Buffer size as a percentage of the total buffer.

Range: 0 through 100

remainder—Remaining buffer available.

temporal *microseconds*—Buffer size as a temporal value. The queuing algorithm starts dropping packets when it queues more than a computed number of bytes. This maximum is computed by multiplying the logical interface speed by the configured temporal value.

Range: The ranges vary by platform. See Buffer Size Temporal Value Ranges by Router Type.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring the Scheduler Buffer Size
- Example: Configuring CoS for a PBB Network on MX Series Routers

drop-probability (Interpolated Value)

Syntax drop-probability [*values*];

Hierarchy Level [edit class-of-service drop-profiles *profile-name* interpolate]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced before Junos OS 11.4 for EX Series switches.

Description Define up to 64 values for interpolating drop probabilities. On EX Series switches, this statement is supported only on EX8200 standalone switches and EX8200 Virtual Chassis.

Options ***percentage***—The probability (expressed in percentage) for a packet to be dropped from the queue.

Range: 0 through 100

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Default Drop Profile
- Configuring CoS Tail Drop Profiles (CLI Procedure)

drop-profiles (Class-of-Service)

Syntax

```
drop-profiles {
  profile-name {
    fill-level percentage drop-probability percentage;
    interpolate {
      drop-probability [values];
      fill-level [values]
    }
  }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced before Junos OS 11.4 for EX Series switches.

Description Define drop profiles for RED.

For a packet to be dropped, it must match the drop profile. When a packet arrives, RED checks the queue fill level. If the fill level corresponds to a nonzero drop probability, the RED algorithm determines whether to drop the 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.

Related Documentation

- Configuring RED Drop Profiles
- Understanding CoS Tail Drop Profiles
- Example: Configuring CoS on EX Series Switches
- Configuring CoS Tail Drop Profiles (CLI Procedure)

fill-level (Drop Profiles)

Syntax fill-level *percentage*;

Hierarchy Level [edit class-of-service drop-profiles *profile-name*]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced before Junos OS 11.4 for EX Series switches.

Description When configuring RED, map the fullness of a queue to a drop probability.

Options ***percentage***—How full the queue is, expressed as a percentage. You configure the **fill-level** and **drop-probability** statements in pairs. To specify multiple fill levels, include multiple **fill-level** and **drop-probability** statements. The values you assign to each statement pair must increase relative to the previous pair's values. This is shown in the segmented graph in RED Drop Profiles Overview.

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.

Related Documentation

- RED Drop Profiles Overview
- Configuring RED Drop Profiles
- Understanding CoS Tail Drop Profiles
- Configuring CoS Tail Drop Profiles (CLI Procedure)

forwarding-class (Multifield Classifiers)

Syntax forwarding-class *class-name*;

Hierarchy Level [edit firewall family *family-name* filter *filter-name* term *term-name* then]

Release Information Statement introduced before Junos OS Release 7.4.

Description Set the forwarding class of incoming packets.

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

- Configuring Multifield Classifiers

from

Syntax

```
from {
  applications [ application-name ];
  application-sets [ set-name ];
  destination-address (CoS) address;
  source-address address;
}
```

Hierarchy Level [edit services cos rule *rule-name* term *term-name*]

Release Information Statement introduced in Junos OS Release 8.1.

Description Specify input conditions for a CoS term.

Options 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.

Related Documentation

- Configuring CoS Rule Sets

input-excess-bandwidth-share

Syntax input-excess-bandwidth-share (proportional *value* | equal);

Hierarchy Level [edit class-of-service interfaces *interface-name*],
[edit class-of-service interfaces interface-set *interface-set-name*]

Release Information Statement introduced in Junos OS Release 9.0.

Description Determines the method of sharing excess bandwidth on the ingress interface in a hierarchical scheduler environment. If you do not include this statement, the node shares excess bandwidth proportionally at 32.64 Mbps.

Options **proportional *value***—(Default) Share ingress excess bandwidth proportionally (default value is 32.64 Mbps).

equal—Share ingress excess bandwidth equally.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring Ingress Hierarchical CoS on Enhanced Queuing DPCs on page 20

input-scheduler-map

Syntax input-scheduler-map *map-name*;

Hierarchy Level [edit class-of-service interfaces *interface-name*],
[edit class-of-service interfaces *interface-name* unit *logical-unit-number*]

Release Information Statement introduced in Junos OS Release 7.6.

Description Associate a scheduler map with a physical or logical interface. The **input-scheduler-map** and **input-traffic-control-profile** statements are mutually exclusive.

Options *map-name*—Name of scheduler map that you define at the **[edit interfaces** *interface-name* **atm-options scheduler-maps]** hierarchy level.

default—The default scheduler mapping.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring a Separate Input Scheduler for Each Interface
- Configuring Ingress Hierarchical CoS on Enhanced Queuing DPCs on page 20
- input-traffic-control-profile on page 33

input-shaping-rate (Physical Interface)

Syntax input-shaping-rate *rate*;

Hierarchy Level [edit class-of-service interfaces *interface-name*]

Release Information Statement introduced in Junos OS Release 7.6.

Description For Gigabit Ethernet IQ2, Enhanced Queuing DPC, MIC, and MPC interfaces, configure input traffic shaping by specifying the amount of bandwidth to be allocated to the physical interface. You can configure hierarchical shaping, meaning you can apply an input shaping rate to both the physical interface and the logical interface.

Options *rate*—Peak 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 160,000,000,000 bps

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring Hierarchical Input Shapers
- Configuring Ingress Hierarchical CoS on Enhanced Queuing DPCs on page 20
- input-traffic-control-profile on page 33

input-traffic-control-profile

Syntax input-traffic-control-profile *profile-name* shared-instance *instance-name*;

Hierarchy Level [edit class-of-service interfaces *interface-name* unit *logical-unit-number*]

Release Information Statement introduced in Junos OS Release 7.6.

Description For Gigabit Ethernet IQ2 and IQ2E PIC, Enhanced Queuing DPC , MIC, and MPC interfaces, apply an input traffic scheduling and shaping profile to the logical interface.

NOTE: The shared-instance statement applies only to Gigabit Ethernet IQ2 and IQ2E PICs.

Options ***profile-name***—Name of the traffic-control profile to be applied to this interface.

instance-name—Name of the shared scheduler and shaper to be applied to this interface.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring Traffic Control Profiles for Shared Scheduling and Shaping
- Configuring Ingress Hierarchical CoS on Enhanced Queuing DPCs on page 20
- input-shaping-rate (Logical Interface)
- traffic-control-profiles

input-traffic-control-profile-remaining

Syntax input-traffic-control-profile-remaining *profile-name*;

Hierarchy Level [edit class-of-service interfaces *interface-name*],
[edit class-of-service interfaces *interface-name* interface-set *interface-set-name*]

Release Information Statement introduced in Junos OS Release 9.0.

Description For Enhanced Queuing DPC, MICs, or MPC interfaces on MX Series routers, or for IQ2E PICs interfaces on M Series and T Series router, apply an input traffic scheduling and shaping profile for the remaining traffic to the logical interface or interface set.

Options ***profile-name***—Name of the traffic-control profile for the remaining traffic to be applied to this interface or interface set.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation
- Configuring Ingress Hierarchical CoS on Enhanced Queuing DPCs on page 20
- input-traffic-control-profile on page 33

interface-set (Hierarchical Schedulers)

Syntax
```
interface-set interface-set-name {
  excess-bandwidth-share (proportional value | equal);
  internal-node;
  output-traffic-control-profile profile-name;
  output-traffic-control-profile-remaining profile-name;
}
```

Hierarchy Level [edit class-of-service interfaces]

Release Information Statement introduced in Junos OS Release 8.5.

Description For Enhanced Queuing DPC, MIC, or MPC interfaces on MX Series routers, or for IQ2E PIC interfaces on M Series and T Series routers, configure hierarchical schedulers for an interface set.

Options ***interface-set-name***—Name of the interface set.

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.

Related Documentation
- Configuring Interface Sets
- Configuring Hierarchical Schedulers for CoS

interfaces

Syntax

```
interfaces {
  interface-name {
    classifiers{
      dscp(classifier-name | default) {
    }
      ieee-802.1 (classifier-name | default)  vlan-tag (inner | outer | classifier-name);
      inet-precedence (rewrite-name | default);
    }
    input-scheduler-map map-name;
    input-shaping-rate rate;
    irb {
      unit logical-unit-number {
        classifiers {
          type (classifier-name | default);
        }
        rewrite-rules {
          dscp (rewrite-name | default);
          dscp-ipv6 (rewrite-name | default);
          exp (rewrite-name | default)protocol protocol-types;
          ieee-802.1 (rewrite-name | default) vlan-tag (outer | outer-and-inner);
          inet-precedence (rewrite-name | default);
        }
      }
    }
    member-link-scheduler (replicate | scale);
    rewrite-rules {
      dscp (rewrite-name | default);
      ieee-802.1 (rewrite-name | default) vlan-tag (outer);
      inet-precedence (rewrite-name | default);
      }
    }
    scheduler-map map-name;
    scheduler-map-chassis map-name;
    shaping-rate rate;
    unit logical-unit-number {
      classifiers {
        type (classifier-name | default) family (mpls | inet);
      }
      forwarding-class class-name;
      fragmentation-map map-name;
      input-shaping-rate (percent percentage | rate);
      input-traffic-control-profile profile-name shared-instance instance-name;
      output-traffic-control-profile profile-name shared-instance instance-name;
      per-session-scheduler;
      rewrite-rules {
        dscp (rewrite-name | default);
        dscp-ipv6 (rewrite-name | default);
        exp (rewrite-name | default) protocol protocol-types;
        exp-push-push-push default;
        exp-swap-push-push default;
        ieee-802.1 (rewrite-name | default) vlan-tag (outer | outer-and-inner);
        inet-precedence (rewrite-name | default);
```

```
        }
        scheduler-map map-name;
        shaping-rate rate;
        translation-table (to-dscp-from-dscp | to-dscp-ipv6-from-dscp-ipv6 | to-exp-from-exp
          | to-inet-precedence-from-inet-precedence) table-name;
      }
    }
    interface-set interface-set-name {
      excess-bandwidth-share;
      internal-node;
      output-traffic-control-profile profile-name;
      output-traffic-control-profile-remaining profile-name;
    }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced before Junos OS Release 7.4.
Interface-set level added in Junos OS Release 8.5.

Description Configure interface-specific CoS properties for incoming packets.

NOTE: The dscp-ipv6 and ieee-802.1ad classifier types are not supported on ACX Series routers. For further information about support on ACX Series routers, see Understanding CoS CLI Configuration Statements on ACX Series Universal Access Routers.

Options 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.

Related Documentation

- Overview of BA Classifier Types
- Configuring Rewrite Rules
- Understanding CoS CLI Configuration Statements on ACX Series Universal Access Routers

loss-priority (Simple Filter)

Syntax loss-priority (high | low | medium);

Hierarchy Level [edit firewall family *family-name* simple-filter *filter-name* term *term-name* then]

Release Information Statement introduced in Junos OS Release 7.6.

Description Set the loss priority of incoming packets.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring Multifield Classifiers

priority (Schedulers)

Syntax priority *priority-level*;

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced in Junos OS Release 12.1X48 for PTX Series Packet Transport Switches.
Statement introduced in Junos OS Release 12.2 for ACX Series Routers.

Description Specify the packet-scheduling priority value.

Options *priority-level* can be one of the following:

- **low**—Scheduler has low priority.
- **medium-low**—Scheduler has medium-low priority.
- **medium-high**—Scheduler has medium-high priority.
- **high**—Scheduler has high priority. Assigning high priority to a queue prevents the queue from being underserved.
- **strict-high**—Scheduler has strictly high priority. Configure a **high** priority queue with unlimited transmission bandwidth available to it. As long as it has traffic to send, the **strict-high** priority queue receives precedence over **low**, **medium-low**, and **medium-high** priority queues, but not **high** priority queues. You can configure **strict-high** priority on only one queue per interface.

NOTE: The strict-high priority level is the only priority level supported on ACX Series Routers. However, multiple strict-high priority queues can be configured per interface on ACX Series Routers.

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation
- Configuring Schedulers for Priority Scheduling

schedulers (Class of Service)

Syntax

```
schedulers {
  scheduler-name {
    adjust-minimum rate;
    adjust-percent percentage;
    buffer-size (seconds | percent percentage | remainder | temporal microseconds);
    drop-profile-map loss-priority (any | low | medium-low | medium-high | high) protocol
      (any | non-tcp | tcp) drop-profile profile-name;
    excess-priority [ low | medium-low | medium-high | high | none];
    excess-rate (percent percentage | proportion value);
    priority priority-level;
    shaping-rate (percent percentage | rate);
    transmit-rate (percent percentage | rate | remainder) <exact | rate-limit>;
  }
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced before Junos OS Release 7.4.
Statement introduced in Junos OS Release 12.1X48 for PTX Series switches.

Description Specify the scheduler name and parameter values.

Options ***scheduler-name***—Name of the scheduler to be configured.

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.

Related Documentation

- Schedulers Overview
- Default Schedulers Overview
- Configuring Schedulers
- Configuring a Scheduler
- Example: Configuring CoS for a PBB Network on MX Series Routers

shared-instance

Syntax shared-instance *instance-name*;

Hierarchy Level [edit class-of-service interfaces *interface-name* unit *logical-unit-number* input-traffic-control-profile],
[edit class-of-service interfaces *interface-name* unit *logical-unit-number* output-traffic-control-profile]

Release Information Statement introduced in Junos OS Release 7.6.

Description For Gigabit Ethernet IQ2 and IQ2E PICs only, apply a shared traffic scheduling and shaping profile to the logical interface.

Options ***instance-name***—Name of the shared scheduler and shaper to be applied to this interface

Required Privilege Level interface—To view this statement in the configuration.
interface-control—To add this statement to the configuration.

Related Documentation

- Configuring Shaping on 10-Gigabit Ethernet IQ2 PICs
- traffic-control-profiles

simple-filter (Configuring)

Syntax

```
simple-filter filter-name {
  term term-name {
    from {
      match-conditions;
    }
    then {
      forwarding-class class-name;
      loss-priority (high | low | medium);
    }
  }
}
```

Hierarchy Level [edit firewall family inet filter *filter-name*]

Release Information Statement introduced in Junos OS Release 7.6.

Description Define a simple filter. Simple filters are recommended for metropolitan Ethernet applications.

Options

from—Match packet fields to values. If the **from** option is not included, all packets are considered to match and the actions and action modifiers in the **then** statement are taken.

match-conditions—One or more conditions to use to make a match.

term-name—Name that identifies the term. The name can contain letters, numbers, and hyphens (-), and can be up to 255 characters long. To include spaces in the name, enclose it in quotation marks (“ ”).

then—Actions to take on matching packets. If the **then** option is not included and a packet matches all the conditions in the **from** statement, the packet is accepted.

The remaining statements are explained separately. Only **forwarding-class** and **loss-priority** are valid in a simple filter configuration.

Required Privilege Level

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

Related Documentation

- Configuring Multifield Classifiers
- filter (Applying to an Interface)
- simple-filter (Applying to an Interface)

then

Syntax

```
then {
  application-profile profile-name;
  dscp (alias | bits);
  forwarding-class class-name;
  syslog;
  (reflexive | reverse) {
    application-profile profile-name;
    dscp (alias | bits);
    forwarding-class class-name;
    syslog;
  }
}
```

Hierarchy Level [edit services cos rule *rule-name* term *term-name*]

Release Information Statement introduced in Junos OS Release 8.1.

Description Define the CoS term actions.

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.

Related Documentation

- Configuring Actions in a CoS Rule
- Configuring Actions in CoS Rules

transmit-rate (Schedulers)

Syntax transmit-rate (*rate* | percent *percentage* | remainder) <exact | rate-limit>;

Hierarchy Level [edit class-of-service schedulers *scheduler-name*]

Release Information Statement introduced before Junos OS Release 7.4.
rate-limit option introduced in Junos OS Release 8.3. Applied to the Multiservices PICs in Junos OS Release 9.4.
Statement introduced in Junos OS Release 12.1X48 for PTX Series Packet Transport Switches.
Statement introduced in Junos OS Release 12.2 for ACX Series Routers.

Description Specify the transmit rate or percentage for a scheduler.

Default If you do not include this statement, the default scheduler transmission rate and buffer size percentages for queues 0 through 7 are 95, 0, 0, 5, 0, 0, 0, and 0 percent, respectively.

Options

exact—(Optional) Enforce the exact transmission rate. Under sustained congestion, a rate-controlled queue that goes into negative credit fills up and eventually drops packets. This value should never exceed the rate-controlled amount. For PTX Series Packet Transport Switches, this option is allowed only on the non-strict-high (high, medium-high, medium-low, or low) queues.

percent *percentage*—Percentage of transmission capacity. A percentage of zero drops all packets in the queue.

Range: 0 through 100 percent for M, MX and T Series routers; 1 through 100 percent for PTX Series Packet Transport Switches; 0 through 200 percent for the SONET/SDH OC48/STM16 IQE PIC

NOTE:

- On M Series Multiservice Edge Routers, for interfaces configured on 4-port E1 and 4-port T1 PICs only, you can configure a ***percentage*** value only from 11 through 100. These two PICs do not support transmission rates less than 11 percent.
- The configuration of the **transmit-rate percent 0 exact** statement at the **[edit class-of-service schedules *scheduler-name*]** hierarchy is ineffective on T4000 routers with Type 5 FPC.
- On MIC and MPC interfaces on MX Series routers, when the transmit rate is configured as a percentage and **exact** or **rate-limit** is enabled on a queue, the shaping rate of the parent node is used to compute the transmit rate. If **exact** or **rate-limit** is not configured, the guaranteed rate of the parent node is used to compute the transmit rate.

rate—Transmission rate, 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: 3200 through 160,000,000,000 bps

NOTE: For all MX Series interfaces, the rate can be from 65,535 through 160,000,000,000 bps.

rate-limit—(Optional) Limit the transmission rate to the rate-controlled amount. In contrast to the **exact** option, the scheduler with the **rate-limit** option shares unused bandwidth above the rate-controlled amount.

NOTE: For PTX Series Packet Transport Switches, this option is allowed only on the strict-high queue. We recommend that you configure rate limit on strict-high queues because the other queues may not meet their guaranteed bandwidths.

NOTE: The configuration of the **rate-limit** statement is supported on T4000 routers only with a Type 5 FPC.

remainder—Use the remaining rate available.

Required Privilege Level

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

Related Documentation

- Configuring Schedulers
- Configuring Scheduler Transmission Rate
- Example: Configuring CoS for a PBB Network on MX Series Routers

unit

Syntax

```
unit logical-unit-number {
  classifiers {
    type (classifier-name | default) family (mpls | all);
  }
  forwarding-class class-name;
  fragmentation-map map-name;
  input-traffic-control-profile profile-name shared-instance instance-name;
  output-traffic-control-profile profile-name shared-instance instance-name;
  per-session-scheduler;
  rewrite-rules {
    dscp (rewrite-name | default);
    dscp-ipv6 (rewrite-name | default);
    exp (rewrite-name | default)protocol protocol-types;
    exp-push-push-push default;
    exp-swap-push-push default;
    ieee-802.1 (rewrite-name | default) vlan-tag (outer | outer-and-inner);
    inet-precedence (rewrite-name | default);
  }
  scheduler-map map-name;
  shaping-rate rate;
}
```

Hierarchy Level [edit class-of-service interfaces *interface-name*]

Release Information Statement introduced before Junos OS Release 7.4.

Description Configure a logical interface on the physical device. You must configure a logical interface to be able to use the physical device.

Options ***logical-unit-number***—Number of the logical unit.

Range: 0 through 16,384

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.

Related Documentation

- Overview of BA Classifier Types
- Configuring Rewrite Rules

PART 3

Index

Index