Junos® OS

CoS on MIC and MPC Interfaces

Release
13.2

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 MIC and MPC Interfaces
13.2

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

YEAR 2000 NOTICE

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

END USER LICENSE AGREEMENT

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 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 xiii 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 xiii 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 MIC and MPC Interfaces

CoS Features on MIC and MPC Interfaces Overview

MIC and MPC interfaces on MX Series routers use the Trio chipset-based queuing model, which supports CoS characteristics that are optimized compared to CoS characteristics supported by the standard queuing model. This topic covers the following information:

Guaranteed Bandwidth and Weight of an Interface Node

The queuing model used by MIC and MPC interfaces separates the concepts of *guaranteed bandwidth* and *weight* of an interface node, although the two terms are often used interchangeably. The guaranteed bandwidth for an interface node is the bandwidth the node can use, independent of what is happening at the other nodes of the scheduling hierarchy. The weight of an interface node, on the other hand, is a value that determines how *excess bandwidth* is used. The weight of a node comes into play when other nodes at the same hierarchical scheduling level use less than the sum of their guaranteed bandwidths

For some application traffic types (such as constant bit rate voice, where there is little concern about excess bandwidth), the guaranteed bandwidth dominates the node. For other types of application traffic (such as bursty data, where a well-defined bandwidth is not always possible), the concept of weight dominates the node.

Hierarchical Scheduling Priority Levels

The queuing model used by MIC and MPC interfaces supports up to four levels of hierarchical scheduling. Scheduling node priority levels correspond to the physical interface, an interface set, a logical interface, and the output queue itself. The queuing model supports three scheduling priorities for guaranteed levels and two lower scheduling priorities for excess levels. You can configure one guaranteed priority and one excess priority. For example, you can configure a queue for guaranteed low (GL) as the guaranteed priority and configure excess high (EH) as the excess priority.

You can associate a guaranteed level with only one excess level. You can associate an excess level with any number of guaranteed priority levels, including none.

Interface nodes maintain their guaranteed priority level (for example, guaranteed high, GH) as long as they do not exceed their guaranteed bandwidth. If the queue bandwidth exceeds the guaranteed rate, then the priority drops to the excess priority (for example, excess high, EH). Because excess level priorities are lower than their guaranteed counterparts, the bandwidth guarantees for each of the other levels can be maintained.

CoS Capabilities and Limitations on MIC and MPC Interfaces

When configuring CoS features on MIC and MPC interfaces on MX Series router, keep the following limitations in mind:

- Input queuing is not supported on the MIC or MPC interfaces.
- On MPCs, you can configure up to 32 DCSP or Internet or EXP rewrite rules, and 32 IEEE rewrite rules. However, if you configure all 32 allowed rewrite rules, the class-of-service process intermittently fails and generates syslog entries.
- The MPC and MIC interfaces do not support the **q-pic-large-buffer** statement at the **[edit chassis fpc *mpc-number* pic *mic-number*]** hierarchy level. By default, 500 ms worth of buffer is supported when the delay buffer rate is less than 1 Gbps. By default, 100 ms worth of buffer is supported when the delay buffer rate is 1 Gbps or more. The maximum supported value for the delay buffer is 256 MB and the minimum value is 4 KB. However, due to the limited number of drop profiles supported and the large range of supported speeds, there can be differences between the user-configured value and the observed hardware value. The enhanced queuing (EQ) interfaces on MICs and MPCs support up to 255 drop profiles, and up to 128 tail-drop priorities for guaranteed low (GL) priorities and 64 each for guaranteed high and medium priorities.
- All tunnel interfaces have 100 ms buffers. The **huge-buffer-temporal** statement is not supported.
- The interfaces on MICs and MPCs take all Layer 1 and Layer 2 overhead bytes into account for all levels of the hierarchy, including preamble, interpacket gaps, frame check sequence, and cyclical redundancy check. Queue statistics also take these overheads into account when displaying byte statistics.
- The interfaces on MICs and MPCs do not support the **excess-bandwidth-sharing** statement. You can use the **excess-rate** statement in scheduler maps and traffic control profiles instead.

- To rewrite service VLAN tag COS bits on MIC and MPC interfaces for VPLS or bridge domains, configure the rewrite rules on the core-facing interface.
- When the interface's delay buffers are oversubscribed by configuration (that is, the user has configured more delay-buffer memory than the system can support), then the configured weighted random early detection (WRED) profiles are implicitly scaled down to drop packets more aggressively from the relatively full queues. This creates buffer space for packets in the relatively empty queues and provides a sense of fairness among the delay buffers. There is no configuration needed for this feature.
- When load balancing EQ MIC interfaces installed in Type 1 MPCs, you should configure odd- and even-numbered interfaces in the form ***interface-fpc/odd | even/*ports**. For example, if one link is **xe-1/0/0**, the other should be **xe-1/1/0**. If you do not configure odd and even load balancing, the system RED-drops packets when sending at line rate. This limitation does not apply to interfaces on EQ MICs installed in Type 2 MPCs.
- When you configure a behavior aggregate (BA) classifier that does not include a specific rewrite rule for MPLS packets, we highly recommend that you include the **rewrite-rules exp default** statement at the **[edit class-of-service interfaces *interface-name* unit *logical-unit-number*]** hierarchy level. Doing so ensures that MPLS exp value is rewritten according to the BA classifier rules configured for forwarding or packet loss priority.

Shaping Rate Granularity

The interfaces on MICs and MPCs have a certain granularity in the application of configured shaping and delay buffer parameters. In other words, the values used are not necessarily precisely the values configured. Nevertheless, the derived values are as close to the configured values as allowed. For the MPC, the shaping rate granularity is 250 kbps for coarse-grained queuing on the basic hardware and 24 kbps for fine-grained queuing on the enhanced queuing devices.

Table 3 on page 5 lists shaping granularities for each MPC type.

Table 3: MPC Interfaces Shaping Granularities

MPC	MPC Port Speed	Queue Level and Port Level	Logical Interface Level and Interface Set Level
Non-Queuing MPCs	1 Gbps / 10 Gbps	250 Kbps	n/a
Queuing MPCs	1 Gbps	2.4 Kbps	9.5 Kbps
	10 Gbps	9.6 Kbps	38 Kbps
Enhanced Queuing MPCs	1 Gbps	1.5 Kbps	6 Kbps
	10 Gbps	6 Kbps	24 Kbps

For delay buffers, the coarse-grained devices support 100 ms of transit rate by default, which can be changed by configuring an explicit buffer size. For fine-grained queuing on enhanced queuing devices, 500 ms of transmit rate is available by default, which can be changed by configuring an explicit buffer size. When this value is changed, there are 256 points available and the closest point is chosen. High-priority and medium-priority queues use 64 points, and the low-priority queues uses 128.

Overhead to Add to Egress Shaping Statistics

Another useful feature is the ability to control how much overhead to count with the **traffic-manager** statement and options. By default, overhead of 24 bytes (20 bytes for the header, plus 4 bytes of cyclical redundancy check [CRC]), is added to egress shaping statistics. You can configure the system to adjust the number of bytes to add to a packet to determine shaped session packet length by adding more bytes (up to 192) of overhead. You can also subtract bytes for egress shaping overhead (up to minus 63 bytes).

This example adds 12 more bytes of overhead to the egress shaping statistics:

```
[edit chassis fpc 0 pic 0]
traffic-manager egress-shaping-overhead 12;
```

Hierarchical Schedulers in Oversubscribed PIR Mode

In contrast to the Intelligent Queuing Enhanced (IQE) and Intelligent Queuing 2 Enhanced (IQ2E) PICs, the interfaces on MICs and MPCs set the guaranteed rate to zero in oversubscribed PIR mode for the per-unit scheduler. Also, the configured rate is scaled down to fit the oversubscribed value. For example, if there are two logical interface units with a shaping rate of 1 Gbps each on a 1-Gbps port (which is, therefore, oversubscribed 2 to 1), then the guaranteed rate on each unit is scaled down to 500 Mbps (scaled down by 2).

With hierarchical schedulers in oversubscribed PIR mode, the guaranteed rate for every logical interface unit is set to zero. This means that the queue transmit rates are always oversubscribed.

Because in oversubscribed PIR mode the queue transmit rates are always oversubscribed, the following are true:

- If the queue transmit rate is set as a percentage, then the guaranteed rate of the queue is set to zero; but the excess rate (weight) of the queue is set correctly.
- If the queue transmit rate is set as an absolute value and if the queue has guaranteed high or medium priority, then traffic up to the queue's transmit rate is sent at that priority level. However, for guaranteed low traffic, that traffic is demoted to the excess low region. This means that best-effort traffic well within the queue's transmit rate gets a lower priority than out-of-profile excess high traffic. This differs from the IQE and IQ2E PICs.

Related Documentation

- CoS Capabilities and Limitations on MIC and MPC Interfaces on page 7
- Configuring Hierarchical Schedulers for CoS

CoS Capabilities and Limitations on MIC and MPC Interfaces

This topic describes CoS scaling and performance parameters that apply to interfaces on MICs and MPCs.

Rewrite Operations

On MPCs, you can configure up to 32 DCSP, or Internet, or EXP rewrite rules, or 32 IEEE rewrite rules. However, if you configure all 32 allowed rewrite rules, the class-of-service process intermittently fails and generates system log entries.

Classification

MPCs support up to 32 classifiers of each type per module.

Related Documentation

- CoS Features on MIC and MPC Interfaces Overview on page 3

Junos Trio Chipset on MX Series FAQs

Junos Trio is a chipset with revolutionary 3D scaling technology that enables networks to dynamically scale for more bandwidth, subscribers, and services . This section answers questions related to this chipset used with Junos OS Release 10.1 and later.

Which product platforms use the Junos Trio chipset?

- MX Series MPC line cards

  The MX Series MPC provides the connection between the customer's Ethernet interfaces and the routing fabric of the MX Series chassis. The board features two Junos Trio chipsets.

- 16-port 10-Gigabit Ethernet MPC

  This MPC provides the connection between 10-Gigabit Ethernet LAN interfaces and the routing fabric of the MX Series chassis. It has four identical packet processing paths combined with a control plane.

- 100-Gigabit Ethernet MPC

  This MPC provides the connection between 100-Gigabit Ethernet LAN interfaces and the routing fabric of the MX Series chassis. It supports two separate slots for MICs.

- MX5, MX10, MX40, and MX80 routers.

What is the total number of Packet Forwarding Engines in each of the DPCs using the Junos Trio chipset?

- The 16-port 10-Gigabit Ethernet MPC on MX Series routers has a total of four Packet Forwarding Engines per MPC.
- Each MPC has two Packet Forwarding Engines.

What is the default power-on sequence when line cards with Junos Trio chipsets are in the same chassis with other types of line cards?

If the **set chassis network-services** attribute is not configured, the following is the line card power-up rule:

- If the first PIC concentrator powered up is a DPC, then only DPCs within the box are allowed to power up.
- If the first PIC concentrator powered up is an MPC, then only MPCs are allowed to power up.

As of Junos OS Release 10.2 and later, the power-up rule has been changed to the following:

- If the **set chassis network-services** attribute is configured as **ip** at start time, any MX Series device-supported boards (such as DPC, FPC, and MPC) will boot.
- If the **set chassis network-services** attribute is configured as **ethernet** at start time, any MX Series device-supported boards (such as DPC, FPC, and MPC) will boot.
- If the **set chassis network-services** attribute is configured as **enhanced-ip** at start time, only MPCs and MS-DPCs are powered on in the chassis. Non-service DPCs do not work with enhanced network services mode options.
- If the **set chassis network-services** attribute is configured as **enhanced-ethernet** at start time, only MPCs and MS-DPCs are powered on in the chassis.

What are the QoS differences between the 16-port 10-Gigabit Ethernet MPC and the I-chip-based DPC?

- Dynamic memory is not supported on the 16-port 10-Gigabit Ethernet MPC.
- A buffer configured on an MPC queue is treated as the maximum. However, it is treated as the minimum on the I-chip DPC.
- The MPC maintains packets in 128-byte chunks.
- Port shaping is supported on all MPCs.
- Queues can have unique shaping and guaranteed rate configuration.

What is the difference in buffer management on the MPCs compared with the DPCs?

On port-queuing DPCs, 64 byte-per-unit dynamic buffers are available per queue. If a queue is using more than its allocated bandwidth share due to excess bandwidth left over from other queues, its buffers are dynamically increased. This is feasible because the I-chip on the DPCs primarily perform weighted random early detection (WRED) drops at the head of the queue, as opposed to “tail-assisted” drops, which are performed only when a temporal buffer is configured or when the queue becomes full. When a temporal buffer is not configured, the allocated buffer is treated as the minimum for that queue, and can expand if other queues are not using their share.

With the Trio chipset on the MPCs, WRED drops are performed at the tail of the queue. The packet buffer is organized into 128-byte units. Before a packet is queued, buffer and

WRED checks are performed, and the decision to drop is made at this time. Once a packet is queued, it is not dropped. As a result, dynamic buffer allocation is not supported on the Packet Forwarding Engines containing the Trio chipset. The buffer allocation per queue on the Packet Forwarding Engines containing the Trio chipset is considered the maximum for that queue. Once the allocated buffer becomes full, subsequent packets are dropped until space is available, even if other queues are idle. Buffering is only required during oversubscription.

To provide larger buffers on Packet Forwarding Engines with the Trio chipset, the delay buffer can be increased from the default 100 ms to 200 ms of the port speed. the delay buffer can also be oversubscribed using the **delay-buffer-rate** configuration per port.

What are the supported QoS features for the 16-port 10-Gigabit Ethernet MPC and 100-Gigabit Ethernet MPC.

The Gigabit Ethernet MPCs supports the following QoS functionality:

- Port-based queuing
  - Per-port shaping
  - Eight queues per port
  - 100 ms of delay buffer by default per port
  - 200 ms of delay buffer configurable per port
  - Ability to oversubscribe the delay buffer beyond 200 ms per port
- Queue-level shaping and guaranteed rate
  - Separate guaranteed and shaping rates
  - Rate limit option to police a queue to act as a Low Latency Queue (LLQ)
- Four WRED profiles per queue
- Multiple queue priority levels
- Strict High, High, Medium, and Low guaranteed priority levels
  - Strict High and High are at the same hardware priority level
  - Round robin at each guaranteed priority level
- High and Low excess priority levels
  - Queues perform WRR at the excess priority levels
  - Strict priority scheduling at each excess priority level
- Classification per VLAN
  - MPLS EXP
  - IPv6, IPv4 ToS

    - Inner and outer tag 802.1p (and DEI7)
    - MF classifiers per VLAN
- Policers per VLAN
    - Single rate, single-rate tricolor marking, two-rate tricolor marking, hierarchical policers
    - Class-aware intelligent hierarchical policers
    - Physical interface policers
- Rewrites per VLAN
    - MPLS EXP, IP DSCP/PREC
    - Inner and outer tag 802.1p (and DEI7)
    - Ingress DSCP rewrite

Related Documentation

- MX Series QoS FAQ Overview
- Dense Port Concentrators FAQs

CHAPTER 2

Scheduling and Queuing

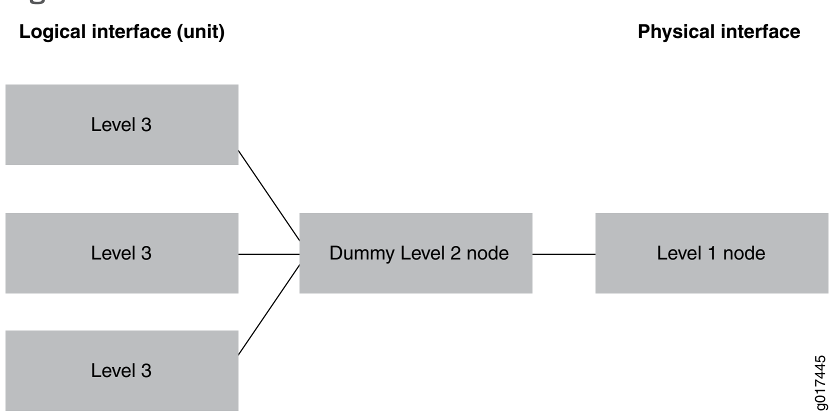
Scheduler Node Scaling on MIC and MPC Interfaces Overview

The MIC and MPC interface hardware (with the exception of the 10-Gigabit Ethernet MPC with SFP+) supports multiple levels of scheduler nodes. In per-unit-scheduling mode, each logical interface (unit) can have 4 or 8 queues and has a dedicated level 3 scheduler node. Scheduler nodes can be a one of four levels: the queue itself (level 4), the logical interface or unit (level 3), the interface set or virtual LAN (VLAN) collection (level 2), or the physical interface or port (level 1). For more information about hierarchical scheduling levels, see Configuring Hierarchical Schedulers for CoS.

The MIC and MPC interface hardware supports enhanced queuing in both per-unit scheduling and hierarchical scheduler modes. The way that the scheduler hierarchy is built depends on the scheduler mode configured.

In per-unit scheduling mode, each logical interface unit has its own dedicated level 3 node and all logical interface units share a common level 2 node (one per port). This scheduling mode is shown in Figure 1 on page 11.

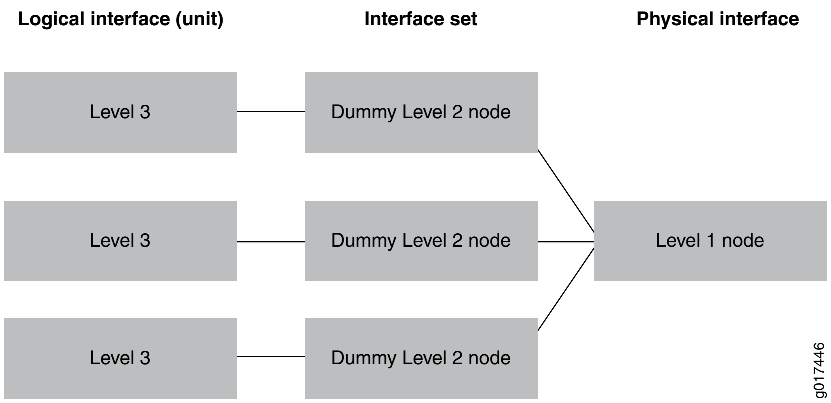
Figure 1: MIC and MPC Interface Per-Unit Scheduler Node Scaling



In this case, in per-unit scheduling mode, the level 2 node is a dummy node.

The case with hierarchical scheduling mode, for a similar configuration when there are no interface sets configured and only the logical interfaces have traffic control profiles is shown in Figure 2 on page 12.

Figure 2: MIC and MPC Interface Hierarchical Scheduling Node Scaling



When an interface set has a CoS scheduling policy but none of its child logical interfaces has a CoS scheduling policy, then the interface set is considered to be a leaf node and has one level 2 and one level 3 node.

In per-unit scheduling, the logical interfaces share a common level 2 node (one per port). In hierarchical-scheduling mode, each logical interface has its own level 2 node. So scaling is limited by the number of level 2 nodes. For hierarchical schedulers, in order to better control system resources in hierarchical-scheduling mode, you can limit the number of hierarchical levels in the scheduling hierarchy to two. In this case, all logical interfaces and interface sets with CoS scheduling policy share a single (dummy) level 2 node, so the maximum number of logical interfaces with CoS scheduling policies is increased (the interface sets must be at level 3). To configure scheduler node scaling, include the **hierarchical-scheduler** statement with the **maximum-hierarchy-levels** option at the **[edit interfaces xe-*fpc/pic/port*]** hierarchy level. The only supported value is 2.

```
[edit interfaces]
xe-2/0/0 {
  hierarchical-scheduler {
    maximum-hierarchy-levels 2;
  }
}
```

NOTE: Level 3 interface sets are supported by the **maximum-hierarchy-levels** option, but level 2 interface sets are not supported. If you configure level 2 interface sets with the **maximum-hierarchy-levels** option, you generate Packet Forwarding Engine errors.

Dedicated Queue Scaling for CoS Configurations on MIC and MPC Interfaces Overview

The 30-Gigabit Ethernet Queuing and 60-Gigabit Ethernet Queuing and Enhanced Queuing Ethernet Modular Port Concentrators (MPCs) provide a set of dedicated queues for subscriber interfaces configured with hierarchical scheduling or per-unit scheduling.

The dedicated queues offered on these MPCs enable service providers to reduce costs through different scaling configurations. For example, the 60-Gigabit Ethernet Enhanced Queuing MPC enables service providers to reduce the cost per subscriber by allowing many subscriber interfaces to be created with four or eight queues. Alternatively, the 30-Gigabit Ethernet and 60-Gigabit Ethernet Queuing MPCs enable service providers to reduce hardware costs, but allow fewer subscriber interfaces to be created with four or eight queues.

This topic describes the overall queue, scheduler node, and logical interface scaling for subscriber interfaces created on these MIC and MPC combinations.

Queue Scaling for MIC and MPC Combinations

Table 4 on page 13 lists the number of dedicated queues and number of subscribers supported per MPC.

Table 4: Dedicated Queues for MIC and MPC Interfaces

MPC	Dedicated Egress Queues	Supported Subscriber Interfaces	Logical Interfaces with 4 Queues	Logical Interfaces with 8 Queues
30-Gigabit Ethernet Queuing MPC	64,000	16,000	16,000 (8000 per PIC)	8000 (4000 per PIC)
60-Gigabit Ethernet Queuing MPC	128,000	32,000	32,000 (8000 per PIC)	16,000 (4000 per PIC)
60-Gigabit Ethernet Enhanced Queuing MPC	512,000	64,000	64,000 (16,000 per PIC)	64,000 (16,000 per PIC)

MPCs vary in the number of Packet Forwarding Engines on board. MPC1s, such as the 30-Gigabit Ethernet MPC, have one Packet Forwarding Engine. MPC2s, such as the 60-Gigabit Ethernet MPC, have two Packet Forwarding Engines. Each Packet Forwarding Engine has two schedulers that share the management of the queues.

A scheduler maps to one-half of a MIC; in CLI configuration statements, that one-half of a MIC corresponds to PIC 0, 1, 2, or 3. MIC ports are partitioned equally across the PICs. A two-port MIC has one port per PIC. A four-port MIC has two ports per PIC.

Each interface-set uses eight queues from total available egress queues.

Distribution of Queues on 30-Gigabit Ethernet Queuing MPCs

On 30-Gigabit Ethernet Queuing MPCs, each scheduler maps to different PICs. When only one MIC is installed, scheduler 0 maps to PIC 0 and scheduler 1 maps to PIC 1 on the MIC. When two MICs are installed, scheduler 0 can additionally distribute queues to PIC 2 on MIC 1, and scheduler 1 can additionally distribute queues to PIC 3 on MIC 1. However, the distribution of queues to the MICs is not hard-partitioned for 30-Gigabit Ethernet Queuing MPCs or other MPC1s. Distribution depends instead on how you allocate the queues to the PICs.

Figure 3 on page 14 shows the queue distribution on a 30-Gigabit Ethernet Queuing MPC with only one MIC installed. All 64,000 egress queues on the MPC are available to the single Packet Forwarding Engine. On the Packet Forwarding Engine, half of these queues (32,000) are managed by each scheduler. Scheduler 0 contributes all of its 32,000 queues to PIC 0. Scheduler 1 contributes all of its 32,000 queues to PIC 1.

Figure 3: Distribution of Queues on the 30-Gigabit Ethernet Queuing MPC with One MIC

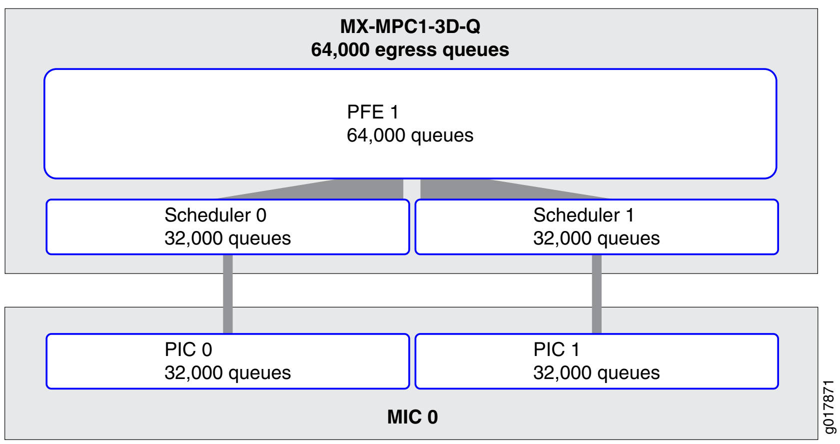
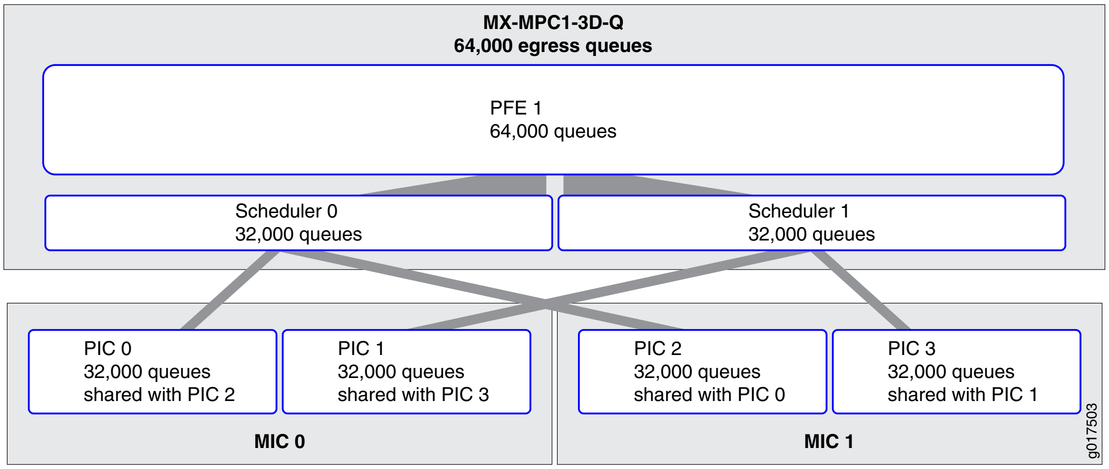


Figure 4 on page 15 shows the queue distribution on the same MPC with two MICs installed. In this case, each scheduler can supply two PICS, one on each MIC. Because the distribution of the queues across the MICs is not hard-partitioned, you can allocate from 0 to 32,000 queues from each scheduler's pool across the scheduler's associated PICs. For example, you can allocate 32,000 queues from Scheduler 0 to PIC 0, 4000 queues from Scheduler 1 to PIC 1, and 28,000 queues from Scheduler 1 to PIC 3. Alternatively, you can allocate the queues evenly across the PICs, or allocate them in other combinations with the limitation of 32,000 queues per PIC and 32,000 queues per port.

Figure 4: Distribution of Queues on the 30-Gigabit Ethernet Queuing MPC with Two MICs

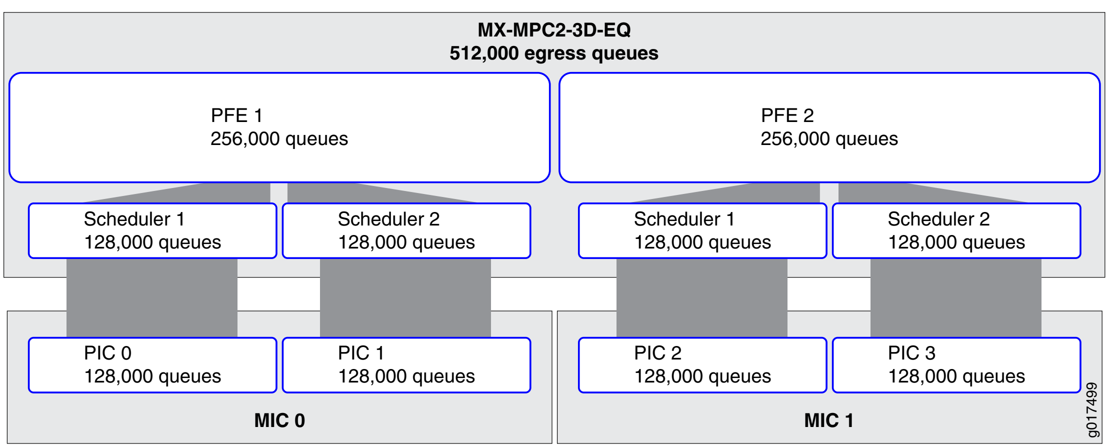


Distribution of Queues on 60-Gigabit Ethernet MPCs

On 60-Gigabit Ethernet Queuing and Enhanced Queuing Ethernet MPCs, each scheduler maps to a single PIC: PIC 0 or PIC 1 on MIC 0 and PIC 2 or PIC 3 on MIC 1. The distribution of the queues is hard-partitioned for these MPCs and other MPC2s; the only difference in distribution is in the total number of queues available.

For example, Figure 5 on page 15 shows how queues are distributed on a 60-Gigabit Ethernet Enhanced Queuing MPC. Of the 512,000 egress queues on the MPC, half (256,000) are available to each of the two Packet Forwarding Engines. On each Packet Forwarding Engine, half of these queues (128,000) are managed by each scheduler. The complete scheduler complement (128,000) is available to only one PIC in a MIC. Thus the total number of queues available depends on the number of MICs installed. The MPC must have 2 MICs to achieve the maximum of 512,000 queues. With a single MIC, the MPC can achieve only 256,000 queues.

Figure 5: Distribution of Queues on the 60-Gigabit Ethernet Enhanced Queuing MPC



Determining Maximum Egress Queues and Subscriber Interfaces per Port

The number of MICs installed in an MPC and the number of ports per MIC do not affect the maximum number of queues available on a given port. These factors affect only how you are able to allocate queues (and, therefore, subscribers) for your network.

For example, a 30-Gigabit Ethernet Queuing MPC supports a maximum of 16,000 subscriber interfaces and has a maximum of 32,000 queues available per PIC. On this card, you can allocate up to 32,000 queues to a single port in each PIC. If you dedicate 4 queues per subscriber interface, you can accommodate a maximum of 8000 subscriber interfaces on a single port, and therefore need at least two ports to reach the maximum 16,000 subscriber interfaces. If you dedicate 8 queues per subscriber interface, you can accommodate a maximum of 4000 subscriber interfaces on a single port, and you need 4 ports for the maximum of 16,000 subscriber interfaces.

The 60-Gigabit Ethernet Enhanced Queuing MPC supports a maximum of 64,000 subscriber interfaces and has a maximum of 128,000 queues per PIC. You can allocate up to 128,000 queues to a single port in each PIC. However, if you dedicate 4 queues per subscriber interface, you can accommodate a maximum of only 16,000 subscriber interfaces on a single MPC port—not 32,000—because the 60-Gigabit Ethernet Enhanced Queuing MPC is limited to 16,000 subscriber interfaces per PIC. If you dedicate 8 queues per subscriber interface, you can also accommodate a maximum of 16,000 subscriber interfaces on a single MPC port. In either case, you need at least 4 ports to reach the maximum of 64,000 subscriber interfaces.

Managing Remaining Queues

When the number of available dedicated queues on the MPC drops below 10 percent, an SNMP trap is generated to notify you .

When the maximum number of dedicated queues on the MPCs is reached, a system log message, **COSD_OUT_OF_DEDICATED_QUEUES**, is generated. The system does not provide subsequent subscriber interfaces with a dedicated set of queues. For per-unit scheduling configurations, there are no configurable queues remaining on the MPC.

For hierarchical scheduling configurations, remaining queues are available when the maximum number of dedicated queues is reached on the MPC. Traffic from these logical interfaces are considered unclassified and attached to a common set of queues that are shared by all subsequent logical interfaces. These common queues are the default port queues that are created for every port. You can configure a traffic control profile and attach that to the interface to provide CoS parameters for the remaining queues.

For example, when the 30-Gigabit Ethernet Queuing MPC is configured with 32,000 subscriber interfaces with four queues per subscriber, the MPC can support 16,000 subscribers with a dedicated set of queues. You can provide CoS shaping and scheduling parameters to the remaining queues for those subscriber interfaces by attaching a special traffic-control profile to the interface.

These subscriber interfaces remain with this traffic control profile, even if dedicated queues become available.

Related Documentation

- For information about managing dedicated queues in a static CoS configuration, see Managing Dedicated and Remaining Queues for Static CoS Configurations on MIC and MPC Interfaces on page 39
- For information about managing dedicated queues in a dynamic subscriber access configuration, see Managing Dedicated and Remaining Queues for Dynamic CoS Configurations on MIC and MPC Interfaces
- Scheduler Node Scaling on MIC and MPC Interfaces Overview on page 11
- COSD System Log Messages

Per-Priority Shaping on MIC and MPC Interfaces Overview

Per-priority shaping enables you to configure a separate shaping rate for each of the five priority levels supported by MIC and MPC interfaces. The main use of per-priority shaping rates is to ensure that higher priority services such as voice and video do not starve lower priority services such as data.

There are five scheduler priorities:

- Guaranteed high (GH)
- Guaranteed medium (GM)
- Guaranteed low (GL)
- Excess high (EH)
- Excess low (EL)

The five scheduler priorities support a shaping rate for each priority:

- Shaping rate priority high (GH)
- Shaping rate priority medium (GM)
- Shaping rate priority low (GL)
- Shaping rate excess high (EH)
- Shaping rate excess low (EL)

If each service is represented by a forwarding class queued at a separate priority, then assigning a per-priority shaping rate to higher priority services accomplishes the goal of preventing the starvation of lower priority services.

To configure per-priority shaping rates, include the **shaping-rate-excess-high** ***rate*** **<burst-size** ***burst*****>**, **shaping-rate-excess-low** ***rate*** **<burst-size** ***burst*****>**, **shaping-rate-priority-high** ***rate*** **<burst-size** ***burst*****>**, **shaping-rate-priority-low** ***rate*** **<burst-size** ***burst*****>**, or **shaping-rate-priority-medium** ***rate*** **<burst-size** ***burst*****>** at the **[edit class-of-service traffic-control-profiles** ***tcp-name*****]** hierarchy level and apply the traffic control profile at the **[edit interfaces]** hierarchy level. You can specify the rate in absolute values, or by using **k** (kilo-), **m** (mega-) or **g** (giga-) units.

You can include one or more of the per-priority shaping statements in a traffic control profile:

```
[edit class-of-service]
traffic-control-profiles {
  tcp-ge-port {
    shaping-rate-excess-high rate <burst-size bytes>;
    shaping-rate-excess-low rate <burst-size bytes>;
    shaping-rate-priority-high rate <burst-size bytes>;
    shaping-rate-priority-low rate <burst-size bytes>;
    shaping-rate-priority-medium rate <burst-size bytes>;
  }
}
```

BEST PRACTICE: When planning your implementation, consider the following behavior. You can configure independent burst-size values for each rate, but the system uses the maximum burst-size value configured in each rate family. For example, the system uses the highest configured value for the guaranteed rates (GH and GM) or the highest value of the excess rates (EH and EM).

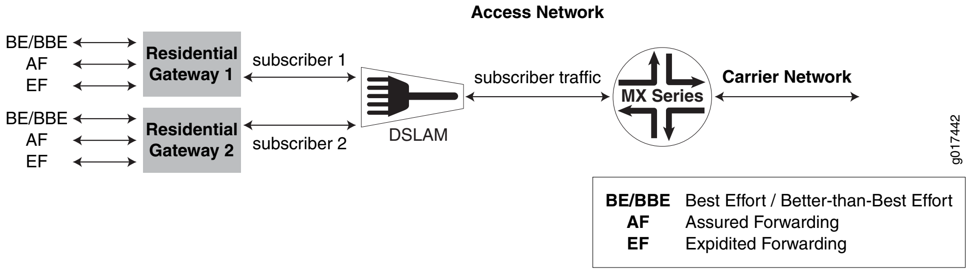
There are several important points about per-priority shaping rates:

- Per-priority shaping rates are only supported on MIC and MPC interfaces (with the exception of the 10-Gigabit Ethernet MPC with SFP+).
- Per-priority shaping is only available for level 1 and level 2 scheduler nodes. (For more information on hierarchical schedulers, see Configuring Hierarchical Schedulers for CoS.)
- Per-priority shaping rates are supported when level 1 or level 2 scheduler nodes have static or dynamic interfaces above them.
- Per-priority shaping rates are supported on aggregated Ethernet (AE) interfaces.
- Per-priority shaping rates are only supported in traffic control profiles.

Per-priority shaping rates can be helpful when the MX Series 3D Universal Edge Router is in a position between subscriber traffic on an access network and the carrier network, playing the role of a broadband services router. In that case, the MX Series router provides quality-of-service parameters on the subscriber access network so that each subscriber receives a minimum bandwidth (determined by the guaranteed rate) and a maximum bandwidth (determined by the shaping rate). This allows the devices closer to the carrier network to operate more efficiently and more simply and reduces operational network expenses because it allows more centralized network management.

One architecture for using per-priority shaping on the MX Series router is shown in Figure 6 on page 19. In the figure, subscribers use residential gateways with various traffic classes to support voice, video, and data services. The MX Series router sends this traffic from the carrier network to the digital subscriber line access multiplexer (DSLAM) and from the DSLAM on to the residential gateway devices.

Figure 6: Architecture for MIC and MPC Interface Per-Priority Shaping

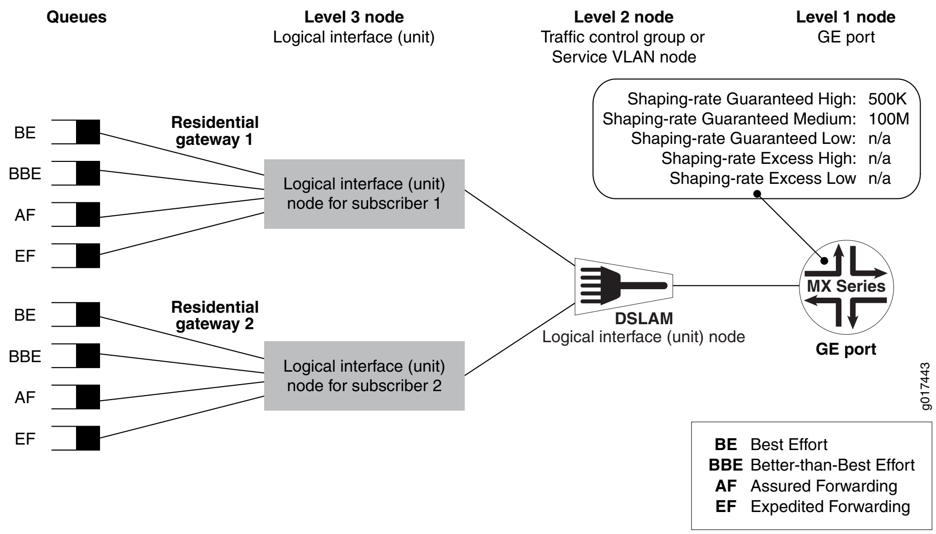


One way that the MX Series router can provide service classes for this physical network topology is shown in Figure 7 on page 20. In the figure, services such as voice and video are placed in separate forwarding classes and the services at different priority levels. For example:

- All expedited-forwarding queues are voice services at a priority level of guaranteed high.
- All assured-forwarding queues are video services at a priority level of guaranteed medium.
- All better-than-best-effort queues are services at a priority level of excess high.
- All best-effort queues are services at a priority level of excess low.

NOTE: This list covers only one possible configuration. Others are possible and reasonable, depending on the service provider's goals. For example, best-effort and better-than-best-effort traffic can have the same priority level, with the better-than-best-effort forwarding class having a higher scheduler weight than the best-effort forwarding class. For more information on forwarding classes, see Configuring Forwarding Classes.

Figure 7: Scheduling Hierarchy for Per-Priority Shaping



Aggregated voice traffic in this topology is shaped by applying a high-priority shaper to the port. Aggregated video traffic is shaped in the same way by applying a medium-priority shaper to the port. As long as the sum of the high- and medium-priority shapers is less than the port speed, some bandwidth is reserved for best-effort and better-than-best-effort traffic. So assured-forwarding and expedited-forwarding voice and video cannot starve best-effort and better-than-best-effort data services. One possible set of values for high-priority (guaranteed high) and medium-priority (guaranteed medium) traffic is shown in Figure 7 on page 20.

BEST PRACTICE: We recommend that you do not shape delay-sensitive traffic such as voice traffic because it adds delay (latency). Service providers often use connection admission control (CAC) techniques to limit aggregated voice traffic. However, establishing a shaping rate for other traffic guards against CAC failures and can be useful in pacing extreme traffic bursts.

Per-priority shaping statements:

```
[edit class-of-service]
traffic-control-profile {
  tcp-for-ge-port {
    shaping-rate-priority-high 500k;
    shaping-rate-priority-medium 100m;
  }
}
```

Apply (attach) the traffic control profile to the physical interface (port) at the **[edit class-of-services interfaces]** hierarchy level:

```
[edit class-of-service]
interfaces {
  ge-1/0/0 {
    output-traffic-control-profile tcp-for-ge-port;
```



```
  }
}
```

Traffic control profiles with per-priority shaping rates can only be attached to interfaces that support per-priority shaping.

You can apply per-priority shaping to levels other than the level 1 physical interface (port) of the scheduler hierarchy. Per-priority shaping can also be applied at level 2, the interface set level, which would typically represent the digital subscriber link access multiplexer (DSLAM). At this level you could use per-priority shaping to limit to total amount of video traffic reaching a DSLAM, for example.

You apply (attach) the traffic control profile to an interface set at the **[edit class-of-services interfaces]** hierarchy level:

```
[edit class-of-service]
interfaces {
  interface-set svlan-1 {
    output-traffic-control-profile tcp-for-ge-port;
  }
}
```

NOTE: **Although you can configure both input and output traffic control profiles, only output traffic control profiles are supported for per-priority shaping.**

You can configure per-priority shaping for the traffic remaining with the **output-traffic-control-profile-remaining** statement on a physical port (a level 2 node) but not for an interface set (a level 3 node).

Related Documentation

- Excess Bandwidth Distribution on MIC and MPC Interfaces Overview on page 26

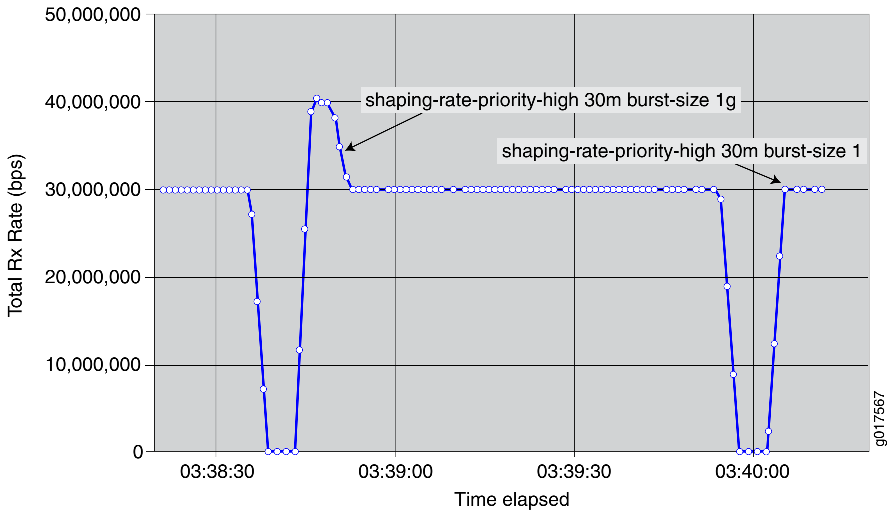
CHAPTER 3

Bandwidth Management

Traffic Burst Management on MIC and MPC Interfaces Overview

You can manage the impact of bursts of traffic on your network by configuring a burst-size value with the shaping rate or the guaranteed rate. The value is the maximum bytes of rate credit that can accrue for an idle queue or scheduler node. When a queue or node becomes active, the accrued rate credits enable the queue or node to catch up to the configured rate.

Figure 8: Sample Burst Shaping Rates



In Figure 8 on page 24, the network administrator configures a large burst-size value for the shaping rate, then configures a small burst-size value. The larger burst size is subject to a maximum value. The smaller burst size is subject to a minimum value that enables the system to achieve the configured rates.

In both configurations, the scheduler node can burst beyond its shaping rate for a brief interval. The burst of traffic beyond the shaping rate is more noticeable with the larger burst size than the smaller burst size.

- Guidelines for Configuring the Burst Size on page 24
- How the System Calculates the Burst Size on page 25

Guidelines for Configuring the Burst Size

Typically, the default burst-size (100 ms) for both scheduler nodes and queues on MIC and MPC interfaces is adequate for most networks. However, if you have intermediate equipment in your network that has very limited buffering and is intolerant of bursts of traffic, you might want to configure a lower value for the burst size.

Use caution when selecting a different burst size for your network. A burst size that is too high can overwhelm downstream networking equipment, causing dropped packets and inefficient network operation. Similarly, a burst size that is too low can prevent the network from achieving your configured rate.

When configuring a burst size, keep the following considerations in mind:

- The system uses an algorithm to determine the actual burst size that is implemented for a node or queue. For example, to reach a shaping rate of 8 Mbps, you must allocate 1Mb of rate credits every second. A shaping rate of 8 Mbps with a burst size of 500,000 bytes of rate-credit per seconds enables the system to transmit at most 500,000 bytes, or 4 Mbps. The system cannot implement a burst size that prevents the rate from being achieved.

  For more information, see “How the System Calculates the Burst Size” on page 25.

- There are minimum and maximum burst sizes for each platform, and different nodes and queue types have different scaling factors. For example, the system ensures the burst cannot be set lower than 1 Mbps for a shaping rate of 8 Mbps. To smoothly shape traffic, rate credits are sent much faster than once per second. The interval at which rate credits are sent varies depending on the platform, the type of rate, and the scheduler level.

- When you have configured adjustments for the shaping rate (either by percentage or through an application such as ANCP or Multicast OIF), the system bases the default and minimum burst-size calculations on the adjusted shaping rate.

- When you have configured cell shaping mode to account for ATM cell tax, the system bases the default and minimum burst-size calculations on the post-tax shaping rate.

- The guaranteed rate and shaping rate share the value specified for the burst size. If the guaranteed rate has a burst size specified, that burst size is used for the shaping rate; if the shaping rate has a burst size specified, that bursts size is used for the guaranteed rate. If you have specified a burst size for both rates, the system uses the lesser of the two values.

- The burst size configured for the guaranteed rate cannot exceed the burst-size configured for the shaping rate. The system generates a commit error.

- If you have not configured a guaranteed rate, logical interfaces and interface sets receive a default guaranteed rate from the port speed. Queues receive a default guaranteed rate from the parent logical interface or interface set.

How the System Calculates the Burst Size

When calculating the burst size, the system uses an exponent of a power of two. For example:

Shaping-rate in bps * 100 ms / (8 bits/byte * 1000 ms/s) = 1,875,000 bytes

The system then rounds this value up. For example, the system uses the following calculation to determine the burst size for a scheduler node with a shaping rate of 150 Mbps:

Max (Shaping rate, Guaranteed rate) bps * 100 ms / (8 bits/byte * 1000 ms/s) = 1,875,000 bytes

Rounded up to the next higher power of two = 2,097,150 (which is 221, or 0x200000)**

The system assigns a single burst size to each of the following rate pairs:

- Shaping rate and guaranteed rate
- Guaranteed high (GH) and guaranteed medium (GM)
- Excess high (EH) and excess low (EL)
- Guaranteed low (GL)

To calculate the burst size for each pair, the system:

- Uses the configured burst-size if only one of the pair is configured.
- Uses the lesser of the two burst sizes if both values are configured.
- Uses the next lower power of two.
- To calculate the minimum burst size, the system uses the greater of the two rates.

Related Documentation

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs on page 41

Excess Bandwidth Distribution on MIC and MPC Interfaces Overview

Service providers often used tiered services to provide bandwidth for excess traffic as traffic patterns vary. By default, excess bandwidth between a configured guaranteed rate and shaping rate is shared equally among all queues on MIC and MPC interfaces, which might not be optimal for all subscribers to a service.

You can adjust this distribution by configuring the rates and priorities for the excess bandwidth.

By default, when traffic exceeds the shaping or guaranteed rates, the system demotes traffic with guaranteed high (GH) priority and guaranteed medium (GM) priority. You can disable this priority demotion for the MIC and MPC interfaces in your router.

Related Documentation

- Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs on page 41

- Managing Excess Bandwidth Distribution for Dynamic CoS on MIC and MIC Interfaces

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Traffic Burst Management on MIC and MPC Interfaces Overview on page 24

Bandwidth Management for Downstream Traffic in Edge Networks Overview

In a subscriber access network, traffic with different encapsulations can be passed downstream to other customer premise equipment (CPE) through the MX Series router. Managing the bandwidth of downstream ATM traffic to Ethernet interfaces can be especially difficult because of the different Layer 2 encapsulations.

The *overhead accounting* feature enables you to shape traffic based on either frames or cells and assign a byte adjustment value to account for different encapsulations.

This feature is available on MIC and MPC interfaces.

Guidelines for Configuring the Shaping Mode

Frame shaping mode is useful for adjusting downstream traffic with different encapsulations. In frame shaping mode, shaping is based on the number of bytes in the frame, without regard to cell encapsulation or padding overhead. Frame is the default shaping mode on the router.

Cell shaping mode is useful for adjusting downstream cell-based traffic. In cell shaping mode, shaping is based on the number of bytes in cells, and accounts for the cell encapsulation and padding overhead.

When you specify cell mode, the resulting traffic stream conforms to the policing rates configured in downstream ATM switches, reducing the number of packet drops in the Ethernet network.

To account for ATM segmentation, the MX Series router adjusts all of the rates by 48/53 to account for ATM AAL5 encapsulation. In addition, the router accounts for cell padding, and internally adjusts each frame by 8 bytes to account for the ATM trailer.

Guidelines for Configuring Byte Adjustments

When the downstream traffic has different byte sizes per encapsulation, it is useful to configure a *byte adjustment* value to adjust the frame sizes. For example, you can configure the frame shaping mode and a byte adjustment value to account for differences in Layer 2 protocols for downstream Ethernet traffic.

We recommend that you specify a byte adjustment value that represents the difference between the CPE protocol overhead and B-RAS protocol overhead.

The system rounds up the byte adjustment value to the nearest multiple of 4. For example, a value of 6 is rounded to 8, and a value of –10 is rounded to –8.

You do not need to configure a byte adjustment value to account for the downstream ATM network. However, you can specify the byte value to account for additional encapsulations or decapsulations in the downstream network.

Relationship with Other CoS Features

Enabling the overhead accounting feature affects the resulting shaping rates, guaranteed rate, and excess rate parameters, if they are configured.

The overhead accounting feature also affects the egress shaping overhead feature that you can configure at the chassis level. We recommend that you use the egress shaping-overhead feature to account for the Layer 2 overhead of the outgoing interface, and use the overhead-accounting feature to account for downstream traffic with different encapsulations and cell-based networks.

When both features are configured together, the total byte adjustment value is equal to the adjusted value of the overhead-accounting feature plus the value of the egress-shaping-overhead feature. For example, if the configured byte adjustment value is 40, and the router internally adjusts the size of each frame by 8, the adjusted overhead accounting value is 48. That value is added to the egress shaping overhead of 24 for a total byte adjustment value of 72.

Setting the Shaping-Rate and Overhead-Accounting Class-of-Service Attributes Based on Access Line Information

You can also use access line parameters in PPPoE discovery packets to set the shaping-rate and overhead-accounting class-of-service attributes on dynamic subscriber interfaces in a broadband access network. This feature is supported on MIC and MPC 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).

You can mix ANCP and PPPoE vendor-specific tags for dynamically instantiated static and interface sets so that the shaping rate is first set using PPPoE vendor-specific tags and is later adjusted by ANCP. In this case, the shaping rate value overrides the PPPoE value.

The Access Node or DSLAM may forward access line information using several methods. This feature only uses access line information received in Vendor-Specific Point-to-Point Protocol over Ethernet (PPPoE) Tags [TR-101].

When you enable this feature, the values supplied by the PPPoE vendor-specific tags override the parameters that you have configured in the CLI for shaping-rate and overhead-accounting statements at the **[edit dynamic-profiles profile-name class-of-service traffic-control-profile]** hierarchy level. The shaping rate is based on the actual-data-rate-downstream attribute, and is only overridden if the vs-tag value is less than the configured value.

Related Documentation

- To configure overhead accounting for static Ethernet interfaces, see Configuring Static Shaping Parameters to Account for Overhead in Downstream Traffic Rates on page 43
- To configure overhead accounting for dynamic subscriber access, see Configuring Dynamic Shaping Parameters to Account for Overhead in Downstream Traffic Rates
- Setting Class-of-Service Parameters Using PPPoE Vendor-Specific Tags
- Configuring the Shaping Rate and Overhead Accounting Based on PPPoE Vendor-Specific Tags on Dynamic Subscriber Interfaces

Intelligent Oversubscription on MIC and MPC Interfaces Overview

On the MIC and MPC interfaces on MX Series routers, as on other types of interface hardware, arriving packets are assigned to one of two preconfigured traffic classes (network control and best effort) based on their header types and destination media access control (MAC) address. Oversubscription, the situation when the incoming packet rate is much higher than the Packet Forwarding Engine and system can handle, can cause key packets to be dropped and result in a flurry of resends, making the problem worse. However, MIC and MPC interfaces handle oversubscription more intelligently and drops lower priority packets when oversubscription occurs. Protocols such as routing protocols are classified as network control. Protocols such as telnet, FTP, and SSH are classified as best effort. No configuration is necessary.

The following frames and packets are assigned to the network control traffic class:

- ARPs: Ethertype **0x0806** for ARP and **0x8035** for dynamic RARP
- IEEE 802.3ad Link Aggregation Control Protocol (LACP): Ethertype **0x8809** and **0x01** or **0x02** (subtype) in first data byte
- IEEE 802.1ah: Ethertype **0x8809** and subtype **0x03**
- IEEE 802.1g: Destination MAC address **0x01–80–C2–00–00–02** with Logical Link Control (LLC) **0xAAAA03** and Ethertype **0x08902**
- PVST: Destination MAC address **0x01–00–0C–CC–CC–CD** with LLC **0xAAAA03** and Ethertype **0x010B**
- xSTP: Destination MAC address **0x01–80–C2–00–00–00** with LLC **0x424203**
- GVRP: Destination MAC address **0x01–80–C2–00–00–21** with LLC **0x424203**
- GMRP: Destination MAC address **0x01–80–C2–00–00–20** with LLC **0x424203**
- IEEE 802.1x: Destination MAC address **0x01–80–C2–00–00–03** with LLC **0x424203**
- Any per-port **my-mac** destination MAC address
- Any configured global Integrated Bridging and Routing (IRB) **my-mac** destination MAC address

In addition, the following Layer 3 control protocols are assigned to the network control traffic class:

- IGMP query and report: Ethertype **0x0800** and carrying an IPv4 protocol or IPv6 next header field set to 2 (IGMP)
- IGMP DVMRP: IGMP field version = **1** and type = **3**
- IPv4 ICMP: Ethertype **0x0800** and IPv4 protocols = **1** (ICMP)
- IPv6 ICMP: Ethertype **0x86DD** and IPv6 next header field = **0x3A** (ICMP)
- IPv4 or IPv6 OSPF: Ethertype **0x0800** and IPv4 protocol field or IPv6 next header field = **89** (OSPF)

- IPv4 or IPv6 VRRP: IPv4 Ethertype **0x0800** or IPv6 Ethertype **0x86DD** and IPv4 protocol field or IPv6 next header field = **112** (VRRP)
- IPv4 or IPv6 RSVP: IPv4 Ethertype **0x0800** or IPv6 Ethertype **0x86DD** and IPv4 protocol field or IPv6 next header field = **46** or **134**
- IPv4 or IPv6 PIM: IPv4 Ethertype **0x0800** or IPv6 Ethertype **0x86DD** and IPv4 protocol field or IPv6 next header field = **103**
- IPv4 or IPv6 IS-IS: IPv4 Ethertype **0x0800** or IPv6 Ethertype **0x86DD** and IPv4 protocol field or IPv6 next header field = **124**
- IPv4 router alert: IPv4 Ethertype **0x0800** and IPv4 option field = **0x94** (router alert)

Also, the following Layer 4 control protocols are assigned to the network control traffic class:

- IPv4 and IPv6 BGP: IPv4 Ethertype **0x0800** or IPv6 Ethertype **0x86DD**, TCP port = **179**, and carrying an IPv4 protocol or IPv6 next header field set to 6 (TCP)
- IPv4 and IPv6 LDP: IPv4 Ethertype **0x0800** or IPv6 Ethertype **0x86DD**, TCP or UDP port = **646**, and carrying an IPv4 protocol or IPv6 next header field set to 6 (TCP) or 17 (UDP)
- IPv4 UDP/L2TP control frames: IPv4 Ethertype **0x0800**, UDP port = **1701**, and carrying an IPv4 protocol field set to 17 (UDP)
- DHCP: Ethertype **0x0800**, IPv4 protocol field set to 17 (UDP), and UDP destination port = **0x43** (DHCP service) or **0x44** (DHCP host)
- IPv4 or IPv6 UDP/BFD: Ethertype **0x0800**, UDP port = **3784**, and IPv4 protocol field or IPv6 next header field set to 17 (UDP)

Finally, any PPP encapsulation (Ethertype **0x8863** (PPPoE Discovery) or **0x8864** (PPP0E Session Control)) is assigned to the network control traffic class (queue 3).

NOTE: These classifications are preconfigured.

CHAPTER 4

Specific Transports

CoS on Ethernet Pseudowires in Universal Edge Networks Overview

You can apply rewrite rules and classifiers to an Ethernet pseudowire on MIC and MPC interfaces on MX Series routers. In an edge network, the pseudowire can represent a single customer.

To create the pseudowires, you use logical tunnel (LT) interfaces that connect two virtual routing forwarding (VRF) instances. To provide CoS to the LT interface, you can apply classifiers and rewrite rules. Rewrite rules enable you to rewrite packet header information by specifying various CoS values, including DiffServ code point (DSCP) and IP precedence.

NOTE: Scheduling is not supported on LT interfaces in the current release.

For example, a VPLS instance is connected to a Layer 3 routing instance. The logical tunnel labeled **lt-9/0/0.0** is configured with **vpls** as the family, and **lt-9/0/0.1** is configured with **inet** as the family. You can apply a rewrite rule and classifier for DSCP to **lt-9/0/0.1**, which can represent a business subscriber.

Related Documentation

- Configuring CoS on an Ethernet Pseudowire for Multiservice Edge Networks on page 45

CoS Scheduling Policy on Logical Tunnel Interfaces Overview

You can configure a CoS scheduling policy on a logical tunnel interface (LT ifl). Logical tunnel interfaces can be used to terminate a pseudowire into a virtual routing and forwarding (VRF) instance. If an LT device is used to terminate a pseudowire, CoS scheduling policies can be applied on the LT interface to manage traffic entering the pseudowire. You accomplish this by configuring the hierarchical-scheduler attribute on the physical interface.

This feature is supported on MIC and MPC interfaces.

Related Documentation

- Configuring a CoS Scheduling Policy on Logical Tunnel Interfaces on page 46
- Configuring Hierarchical Schedulers for CoS
- CoS on Ethernet Pseudowires in Universal Edge Networks Overview on page 31
- Configuring CoS on an Ethernet Pseudowire for Multiservice Edge Networks on page 45

CoS for L2TP LNS Inline Services Overview

You can apply hierarchical scheduling and per-session shaping to Layer 2 Tunnel Protocol (L2TP) network server (LNS) inline services using a static or dynamic CoS configuration.

This feature is supported on MIC and MPC interfaces on MX240, MX480, and MX960 routers.

- Guidelines for Applying CoS to the LNS on page 32
- Hardware Requirements for Inline Services on the LNS on page 33

Guidelines for Applying CoS to the LNS

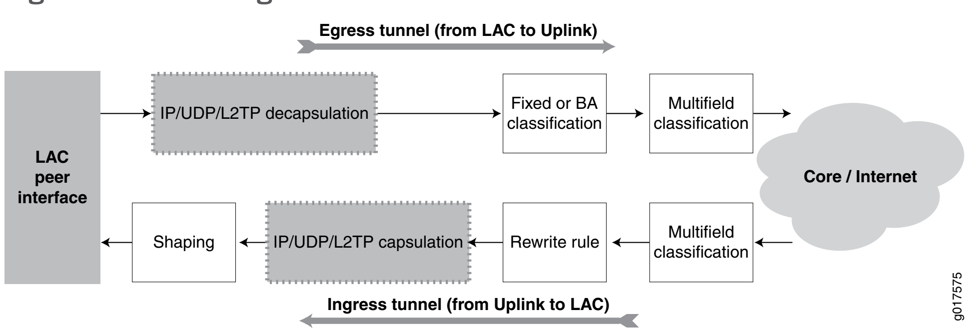
In L2TP configurations, IP, UDP, and L2TP headers are added to packets arriving at a PPP subscriber interface on the L2TP access concentrator (LAC) before being tunneled to the LNS.

When a service interface is configured for an L2TP LNS session, it has an *inner* IP header and an outer IP header. You can configure CoS for an LNS session that corresponds to the inner IP header only. The *outer* IP header is used for L2TP tunnel processing only.

However, we recommend that you configure classifiers and rewrite-rules to transfer the ToS (type of service) value from the inner IP header to the outer IP header of the L2TP packet.

Figure 9 on page 32 shows the classifier and rewrite rules that you can configure on an LNS inline service.

Figure 9: Processing of CoS Parameters in an L2TP LNS Inline Service



By default, the shaping calculation on the service interface includes the L2TP encapsulation. If necessary, you can configure additional adjustments for downstream ATM traffic from the LAC or differences in Layer 2 protocols.

Hardware Requirements for Inline Services on the LNS

Hierarchical scheduling for L2TP LNS inline services is supported on MIC and MPC interfaces only. The services that you can configure depend on the hardware combination. Table 5 on page 33 lists the supported inline services and peer interfaces for each MIC and MPC combination.

Table 5: Hardware Requirements for L2TP LNS Inline Services

MPC Module	Inline Service Support–With Per-Session Shaping	Inline Service Support–Without Per-Session Shaping
MX-MPC1-3D MX-MPC2-3D	No	Yes
MX-MPC1-3D-Q MX-MPC2-3D-Q MX-MPC2-3D-EQ	Yes	Yes
MPC-3D-16XGE-SFPP	No	No

Related Documentation

- Configuring Static CoS for an L2TP LNS Inline Service on page 48
- Configuring Dynamic CoS for an L2TP LNS Inline Service

CoS on Application Services Modular Line Card Overview

The Application Services Modular Line Card (AS MLC) is designed to run services for real-time traffic on MX240, MX480, and MX960 routers. It consists of three main components:

- Application Services Modular Carrier Card (AS MCC)
- Application Services Modular Processing Card (AS MXC)
- Application Services Modular Storage Card (AS MSC)

It supports class-of-service (CoS) features to ensure the quality of service (QoS) for real-time traffic that is sensitive to latency on a network. The AS MLC supports the following CoS features:

- Code-point Aliases—A code-point alias assigns a name to a pattern of code-point bits. On the AS MLC, you can use the code-point alias name for CoS components such as classifiers and drop-profile maps.
- Classification—Packet classification refers to the examination of ingress packets. On the AS MLC, the traffic flowing from the Modular Processing Card (AS MXC) towards the Modular Carrier Card (AS MCC) supports three types of classification:

- Behavior Aggregate (BA)—BA classifier can be configured on the aggregated logical interfaces to classify traffic flowing from the AS MXC towards the AS MCC. With BA classification you can set the forwarding class and loss priority of a packet based on its code points. The AS MLC only supports IP classification (classification based on Type of Service (ToS) and Differentiated Services Code Point (DSCP)) and classification is supported for the IPv4 family only. The Media Flow Controller application sets appropriate DSCP/ToS code-point in the packet that is evaluated by the BA classifier on the AS MCC to classify the packet.
- Multifield Classification—With multifield classifiers you can set the class and loss priority based on one or more of the following packet header fields: destination address, destination port, DSCP, IP protocol, and source address. Multifield classifiers are used when a simple BA classifier is insufficient to classify a packet.
- Fixed Classification—Fixed classification can be configured on logical interfaces by specifying a forwarding class to be applied to all packets received by the logical interface, regardless of the packet contents.

- Scheduling—Schedulers enable you to define the buffer sizes, delay buffer size, drop profile map, excess priority, excess rate percentage, output-traffic-control profile, priority, scheduler-map, shaping rate, transmit rate, and random early detection (RED) drop profiles to be applied to a particular queue for packet transmission.

The AS MLC provides CoS features in the following deployment scenarios:

- HTTP Reverse Proxy— In HTTP reverse proxy configurations, the service provider provides services to a set of domains (content providers) that buy content caching capability from the service provider. Clients connect to content providers through virtual IP (VIP) addresses. Service providers in the reverse proxy scenario generally deploy the routers with AS MLC hardware to honor service requests (such as caching) from the domain users.
- HTTP Transparent Proxy—In HTTP transparent proxy configurations, the service provider implements the AS MLC to improve its own caching capability and reduce the load on its own network. Implementing caching on an MX Series router with an AS MLC improves the retrieval speeds for data and optimizes the back-end network utilization.
- Mixed Mode—In mixed mode both reverse proxy and transparent proxy are configured on the same router.

CoS Implementation in HTTP Reverse Proxy Scenario

In the reverse proxy configuration, the AS MXC provides content to multiple domains. The Media Flow Controller application on an AS MXC implements the differentiated services by setting the DSCP or IP precedence value for the IP packets traversing from the AS MXC to an AS MCC on the AS MLC hardware. The Modular Carrier Card uses these values to classify the packet and provide a suitable level of service.

The Media Flow Controller application detects the domain it serves and marks the DSCP values or the IP precedence bit value based on how important the traffic corresponding to that particular domain is. The service provider operator also sets a behavior aggregate (BA) classifier on the aggregated interfaces on the AS MCC. Based on the DSCP/IP precedence bits, the classifier sets the forwarding class and packet loss priority for the packet. The forwarding class and the packet loss priority values govern the next-hop behavior of the packet traversing the Juniper Networks router.

Unlike a firewall, the Media Flow Controller application implemented on the AS MLC hardware marks the DSCP/IP precedence values based on the application layer protocols. This feature ensures that important traffic flowing from the AS MXC gets a higher priority and is processed accordingly. For example, if MPEG is implemented on the egress, the drop precedence for each frame can be different such that the P and B frames (which require more processing) are dropped before the I frames, resulting in a better quality video for the end user.

For traffic received on the ingress interfaces, end-to-end quality-of-service (QoS) policies ensure that the traffic arriving at the interface has the right DSCP values and the traffic is prioritized based on the forwarding class and packet loss priority values.

CoS Implementation in Transparent Proxy Scenario

In the HTTP transparent proxy configuration, the service provider deploys the AS MLC hardware to reduce its own traffic instead of serving a particular domain. The Media Flow Controller application marks the DSCP bits based on its own requirements rather than those of the domains. Besides this, the CoS implementation for the egress interface is similar to the reverse proxy configuration scenario. The incoming packets follow the QoS policies applied at the WAN interface.

CoS Implementation in Mixed-Mode Scenario

In mixed mode both reverse proxy and transparent proxy configuration coexist on the same AS MLC hardware. In such a scenario, reverse proxy is configured on an aggregated interface and transparent proxy is configured on a regular interface with the Media Flow Controller application marking the appropriate DSCP values for both the configurations. The individual CoS implementation in both the scenarios remains similar to the implementation discussed in “CoS Implementation in HTTP Reverse Proxy Scenario” on page 34 and “CoS Implementation in Transparent Proxy Scenario” on page 35

PART 2

Configuration

CHAPTER 5

Configuration Tasks for Bandwidth Management

Managing Dedicated and Remaining Queues for Static CoS Configurations on MIC and MPC Interfaces

This topic describes how to manage dedicated and remaining queues for static subscriber interfaces configured at the **[edit class-of-service]** hierarchy.

Configuring the Maximum Number of Queues for MIC and MPC Interfaces

30-Gigabit Ethernet Queuing MPCs and 60-Gigabit Ethernet Queuing and Enhanced Queuing MPCs support a dedicated number of queues when configured for hierarchical scheduling and per-unit scheduling configurations.

To scale the number of subscriber interfaces per queue, you can modify the number of queues supported on the MIC.

To configure the number of queues:

1. Specify that you want to configure the MIC.

   user@host# **edit chassis fpc** ***slot-number*** **pic** ***pic-number***

2. Configure the number of queues.

   [edit chassis fpc *slot-number* pic *pic-number*]
   user@host# **setmax-queues-per-interface** (8 | 4)

Configuring Remaining Common Queues on MIC and MPC Interfaces

30-Gigabit Ethernet Queuing MPCs and 60-Gigabit Ethernet Queuing and Enhanced Queuing MPCs support a dedicated set of queues when configured with hierarchical scheduling.

When the number of dedicated queues is reached on the module, there can be queues remaining. Traffic from these logical interfaces are considered unclassified and attached to a common set of queues that are shared by all subsequent logical interfaces.

You can configure traffic shaping and scheduling resources for the remaining queues by attaching a special traffic-control profile to the interface. This feature enables you to provide the same shaping and scheduling to remaining queues as the dedicated queues.

To configure the remaining queues on a MIC or MPC interface:

1. Configure CoS parameters in a traffic-control profile.

   [edit class-of-service]
   user@host# **edit traffic-control-profiles** ***profile-name***

2. Enable hierarchical scheduling for the interface.

   [edit interfaces *interface-name*]
   user@host# **set hierarchical-scheduler**

3. Attach the traffic control profiles for the dedicated and remaining queues to the port on which you enabled hierarchical scheduling.

   To provide the same shaping and scheduling parameters to dedicated and remaining queues, reference the same traffic-control profile.

   a. Attach the traffic-control profile for the dedicated queues on the interface.

      [edit class-of-service interfaces *interface-name*]
      user@host# **set output-traffic-control-profile** ***profile-name***

   b. Attach the traffic-control profile for the remaining queues on the interface.

      [edit class-of-service interfaces *interface-name*]
      user@host# **set output-traffic-control-profile-remaining** ***profile-name***

Related Documentation

- Dedicated Queue Scaling for CoS Configurations on MIC and MPC Interfaces Overview on page 13
- Verifying the Number of Dedicated Queues Configured on MIC and MPC Interfaces on page 75
- Configuring Hierarchical Schedulers for CoS
- Configuring Interface Sets

Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs

Service providers often used tiered services that must provide bandwidth for excess traffic as traffic patterns vary. By default, excess bandwidth between a configured guaranteed rate and shaping rate is shared equally among all queues, which might not be optimal for all subscribers to a service.

To manage excess bandwidth:

1. Configure the parameters for the interface.

   a. Configure the shaping rate.

      [edit class-of-service traffic-control-profiles *profile-name*]
      user@host# **set shaping-rate (percent** ***percentage*** **|** ***rate*****) <burst-size** ***bytes*****>**

      TIP: **On MIC and MPC interfaces, the guaranteed rate and the shaping rate share the value specified for the burst size. If the guaranteed rate has a burst size specified, it is used for the shaping rate; if the shaping rate has a burst size specified, it is used for the guaranteed rate. If you have specified a burst for both rates, the system uses the lesser of the two values.**

   b. Configure the excess rate.

      You can configure an excess rate for all priorities of traffic.

      [edit class-of-service traffic-control-profiles *profile-name*]
      user@host# **set** excess-rate **(percent** ***percentage*** **| proportion** ***value*****)**

      Optionally, you can configure an excess rate specifically for high- and low-priority traffic. When you configure the **excess-rate** statement for an interface, you cannot also configure the **excess-rate-low** and **excess-rate-high** statements.

      [edit class-of-service traffic-control-profiles *profile-name*]
      user@host# **set** excess-rate-high **(percent** ***percentage*** **| proportion** ***value*****)**
      user@host# **set**excess-rate-low **(percent** ***percentage*** **| proportion** ***value*****)**

      BEST PRACTICE: **We recommend that you configure either a percentage or a proportion of the excess bandwidth for all schedulers with the same parent in the hierarchy. For example, if you configure interface 1.1 with twenty percent of the excess bandwidth, configure interface 1.2 with eighty percent of the excess bandwidth.**

2. (Optional) Configure parameters for the queue.

   a. Configure the shaping rate.

      [edit class-of-service scheduler *scheduler-name*]
      user@host#**set shaping-rate (*****rate*** **| $junos-cos-scheduler-shaping-rate) <burst-size** ***bytes*****>**

b. Configure the excess rate.

[edit class-of-service scheduler *scheduler-name*]
user@host#**set** excess-rate **(percent** ***percentage*** **| proportion** ***value*****)**

c. (Optional) Configure the priority of excess bandwidth for the queue.

[edit class-of-service scheduler *scheduler-name*]
user@host#**set** excess-priority **(low | medium-low | medium-high | high | none)**

TIP:

For queues, you cannot configure the excess rate in these cases:

- When the **transmit-rate exact** statement is configured. In this case, the shaping rate is equal to the transmit rate and the queue does not operate in the excess region.
- When the scheduling priority is configured as **strict-high**. In this case, the queue gets all available bandwidth and never operates in the excess region.

By default, when traffic exceeds the shaping or guaranteed rates, the system demotes traffic configured with guaranteed high (GH) priority and guaranteed medium (GM) priority. To disable priority demotion, specify the **none** option. You cannot configure this option when the parent's guaranteed rate is set to zero.

For example, the following statements establish a traffic control profile with a shaping rate of 80 Mbps and an excess rate of 100 percent.

```
[edit class-of-service traffic-control-profiles]
tcp-example-excess {
  shaping-rate 80m;
  excess-rate percent 100;
}
```

The following statements establish a scheduler with an excess rate of 5 percent and a low priority for excess traffic.

```
[edit class-of-service scheduler]
example-scheduler {
  excess-priority low;
  excess-rate percent 5;
}
```

Related Documentation

- Excess Bandwidth Distribution on MIC and MPC Interfaces Overview on page 26
- For more information on hierarchical scheduling and operational modes, see Configuring Hierarchical Schedulers for CoS.

Configuring Static Shaping Parameters to Account for Overhead in Downstream Traffic Rates

The overhead accounting feature enables you to account for downstream traffic that has different encapsulations or downstream traffic from cell-based equipment, such as ATM switches.

You can configure the overhead accounting feature to shape downstream traffic based on frames or cell shaping mode.

You can also account for the different byte sizes per encapsulation by configuring a byte adjustment value for the shaping mode.

To configure the shaping mode and byte adjustment value for static CoS configurations:

1. Specify the shaping mode.

   Frame shaping mode is enabled by default.

   [edit class-of-service traffic-control-profiles *profile-name*]
   user@host# **set overhead-accounting (frame-mode | cell-mode)**

2. (Optional) Specify a byte adjustment value.

   [edit class-of-service traffic-control-profiles *profile-name*
   user@host# **set overhead-accounting bytes** ***byte-value***]

   BEST PRACTICE: We recommend that you specify a byte adjustment value that represents the difference between the customer premise equipment (CPE) protocol overhead and the B-RAS protocol overhead.

   The available range is –120 through 124 bytes. The system rounds up the byte adjustment value to the nearest multiple of 4. For example, a value of 6 is rounded to 8, and a value of –10 is rounded to –8.

Related Documentation

- Bandwidth Management for Downstream Traffic in Edge Networks Overview on page 27

CHAPTER 6

Configuration for Specific Transports

Configuring CoS on an Ethernet Pseudowire for Multiservice Edge Networks

You can configure rewrite rules and classifiers to logical tunnel (LT) interfaces that are configured to represent Ethernet pseudowires.

This feature is supported on MIC and MPC interfaces.

To configure CoS on an LT interface configured for an Ethernet pseudowire:

1. Configure a pair of LT interfaces to represent a pseudowire.

   To apply rewrite rules and classifiers to the pseudowire, you must assign one of the LT interfaces to the **inet** family.

   ```
   [edit]
   user@host#edit interfaces lt-fpc/pic/port
   user@host#edit unit logical-unit-number
   user@host#set encapsulation encapsulation
   user@host#set family (inet | inet6 | iso | mpls)
   user@host#set peer-unit unit-number
   ```

2. Configure the rewrite rule.

   The available rewrite rule types for an LT interface are **dscp** and **inet-precedence**.

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

3. Configure the classifier.

   The available classifier types for an LT interface are **dscp** and **inet-precedence**.

```
[edit class-of-service]
user@host#edit classifiers (dscp | inet-precedence) classifier-name
user@host#edit forwarding-class class-name
user@host#set loss-priority class-name code-points [aliases] [bit-patterns]
```

4. Apply the rewrite rule and classifier to the LT interface that you assigned to the **inet** family.

```
[edit class-of-service interfaces interface-name unit logical-unit-number]
user@host#set rewrite-rule (dscp | inet-precedence) (rewrite-name | default) protocol
   protocol-types
user@host# set classifiers (dscp | inet-precedence) (classifier-name | default)
```

Related Documentation

- CoS on Ethernet Pseudowires in Universal Edge Networks Overview on page 31
- For more information about rewrite rules and classifiers, see the Junos OS Class of Service Configuration Guide

Configuring a CoS Scheduling Policy on Logical Tunnel Interfaces

You can configure a CoS scheduling policy on a logical tunnel interface (LT ifl). Logical tunnel interfaces can be used to terminate a pseudowire into a virtual routing and forwarding (VRF) instance. If an lt device is used to terminate a pseudowire, CoS scheduling policies can be applied on the lt interface to manage traffic entering the pseudowire. You accomplish this by configuring the hierarchical-scheduler attribute on the physical interface.

NOTE: It is important to first commit the hierarchical-scheduler configuration under the logical tunnel physical interface (LT ifd), and subsequently add and commit the class-of-service configuration.

NOTE: The **output-traffic-control** statement applies only to the LT ifl that is part of an L3 VRF instance.

The following example shows two pseudowires (pw1 and pw2) over lt-1/0/10. pw1 carries data, voice, and video traffic, and pw2 carries only data and voice traffic. All pseudowire traffic is restricted to 800m bps. The shaping rate for traffic over pw1 is 400m bps and the shaping rate for traffic over pw2 is 400m bps.

```
[edit interfaces]
lt-1/0/10 {
   hierarchical-scheduler;
}
[edit class-of-service schedulers]
data_sch {
   buffer-size remainder;
   priority low;
}
voice_sch {
   transmit-rate 6k;
```



```
      priority strict-high;
    }
    video_sch {
      shaping-rate 1m;
      priority medium-low;
    }
    [edit class-of-service scheduler-maps]
    pw1-smap {
      forwarding-class be scheduler data_sch;
      forwarding-class ef scheduler voice_sch;
      forwarding-class af scheduler video_sch;
    }
    pw2-smap {
      forwarding-class be scheduler data_sch;
      forwarding-class ef scheduler voice_sch;
    }
    [edit class-of-service traffic-control-profiles]
    pw1-tcp {
      scheduler-map pw1-smap;
      shaping-rate 400m;
    }
    pw2-tcp {
      scheduler-map pw2-smap;
      shaping-rate 400m;
    }
    all-pw-tcp {
      shaping-rate 800m;
    }
    lt-ifd-remain {
      shaping-rate 10m;
    }
    [edit class-of-service interfaces]
    lt-1/0/10 {
      output-traffic-control-profile all-pw-tcp;
      output-traffic-control-profile-remaining lt-ifd-remain;
    unit 1 {
      output-traffic-control-profile pw1-tcp;
    }
    unit 3 {
      output-traffic-control-profile pw2-tcp;
    }
```

Related Documentation

- CoS Scheduling Policy on Logical Tunnel Interfaces Overview on page 31
- Configuring Hierarchical Schedulers for CoS
- Configuring Logical Tunnel Interfaces
- CoS on Ethernet Pseudowires in Universal Edge Networks Overview on page 31
- Configuring CoS on an Ethernet Pseudowire for Multiservice Edge Networks on page 45

Configuring Static CoS for an L2TP LNS Inline Service

You can configure hierarchical scheduling for an L2TP LNS inline service and manage the IP header values using rewrite rules and classifiers.

Before you begin, configure the L2TP LNS inline service interface. See Configuring an L2TP LNS with Inline Service Interfaces.

To configure static CoS for an L2TP LNS inline service:

1. Configure the hierarchical scheduler for the service interface (**si**) interface.

   [edit interfaces si-*fpc/port/pic*]
   user@host# **set hierarchical-scheduler maximum-hierarchy-levels 2**

   BEST PRACTICE: **To enable Level 3 nodes in the LNS scheduler hierarchy and to provide better scaling, we recommend that you also specify a maximum of two hierarchy levels.**

2. Configure the LNS to reflect the IP ToS value in the inner IP header to the outer IP header.

   [edit services l2tp tunnel-group *name*]
   user@host# **set tos-reflect**

3. Configure the classifier for egress traffic from the LAC:

   a. Define the fixed or behavior aggregate (BA) classifier.

      - To configure a fixed classifier:

        [edit class-of-service interfaces si-*fpc/port/pic* unit *logical-unit-number*]
        user@host# **set forwarding-class *class-name***

      - To configure a BA classifier:

        [edit class-of-service]
        user@host# **set classifiers (dscp | dscp-ipv6 | inet-precedence) *classifier-name* forwarding-class *class-name* loss-priority *level* code-points [*aliases*] [*bit-patterns*]**

   b. Apply the classifier to the service interface.

      - To apply the classifier for the DSCP or DSCP IPv6 value:

        [edit class-of-service interfaces si-*fpc/port/pic* unit *logical-unit-number* classifiers]
        user@host# **set dscp (*classifier-name* | default)**
        user@host# **set dscp-ipv6 (*classifier-name* | default)**

      - To apply the classifier for the ToS value:

        [edit class-of-service interfaces si-*fpc/port/pic* unit *logical-unit-number* classifiers]
        user@host# **set inet-precedence (*classifier-name* | default)**

4. Configure and apply a rewrite-rule to ingress traffic to the LAC:

   a. Configure the rewrite rule with the forwarding class and the loss priority value.

[edit class-of-service]
user@host# **set rewrite-rules (dscp | dscp-ipv6 | inet-precedence)** ***rewrite-name***
forwarding-class ***class-name*** **loss-priority** ***level*** **code-point (*****alias*** **|** ***bits*****)**

b. Apply the rewrite rule to the service interface.

- To apply the rewrite rule for the DSCP or DSCP IPv6 value:

  [edit class-of-service interfaces si-*fpc/port/pic* unit *logical-unit-number*
  rewrite-rules]
  user@host# **set dscp(*****rewrite-name*** **| <default>) protocol** ***protocol-types***
  user@host# **set dscp-ipv6 (*****rewrite-name*** **| <default>)**

- To apply the rewrite rule for the ToS value:

  [edit class-of-service interfaces si-*fpc/port/pic* unit *logical-unit-number*
  rewrite-rules]
  user@host# **set inet-precedence (*****rewrite-name*** **| <default>) protocol**
  protocol-types

5. (Optional) Configure additional adjustments for downstream ATM traffic.

   By default, the shaping calculation on the service interface includes the L2TP encapsulation. If necessary, you can configure additional adjustments for downstream ATM traffic from the LAC or differences in Layer 2 protocols.

   [edit class-of-service traffic-control-profiles *profile-name*]
   user@host# **set overhead-accounting (frame-mode | cell-mode) <bytes** ***byte-value***

6. Apply the traffic-control profile.

   [edit class-of-service interfaces si-*fpc/port/pic* unit *logical-unit-number*]
   user@host# **set output-traffic-control-profile** ***profile-name***

   BEST PRACTICE: To limit bandwidth for tunneled sessions with default CoS configurations, we recommend that you also configure CoS for the remaining traffic on the static service interface.

   [edit class-of-service interfaces si-*fpc/port/pic*]
   user@host# **set output-traffic-control-profile-remaining** ***profile-name***

Related Documentation

- CoS for L2TP LNS Inline Services Overview on page 32
- Configuring Static Shaping Parameters to Account for Overhead in Downstream Traffic Rates on page 43

Configuring Ingress Hierarchical CoS on MIC and MPC Interfaces

You can configure ingress CoS parameters, including hierarchical schedulers, on MIC and MPC interfaces on MX Series routers. 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;
unit logical-unit-number;
  input-scheduler-map map-name;
  input-shaping-rate (percent percentage | rate);
  input-traffic-control-profile profile-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-traffic-control-profile profile-name;
  input-traffic-control-profile-remaining profile-name;
  interface interface-name {
    input-excess-bandwidth-share (proportional value | equal);
    input-traffic-control-profile profile-name;
    input-traffic-control-profile-remaining profile-name;
    unit logical-unit-number;
  }
}
```

By default, ingress CoS features are disabled on MIC and MPC interfaces. To enable ingress CoS on a MIC or MPC interface, configure the **traffic-manager** statement with **ingress-and-egress** mode as shown in the following example:

```
chassis {
  fpc 7 {
    pic 0 {
      traffic-manager {
        mode ingress-and-egress;
      }
    }
  }
}
```

Configured CoS features on the ingress are independent of CoS features on the egress.

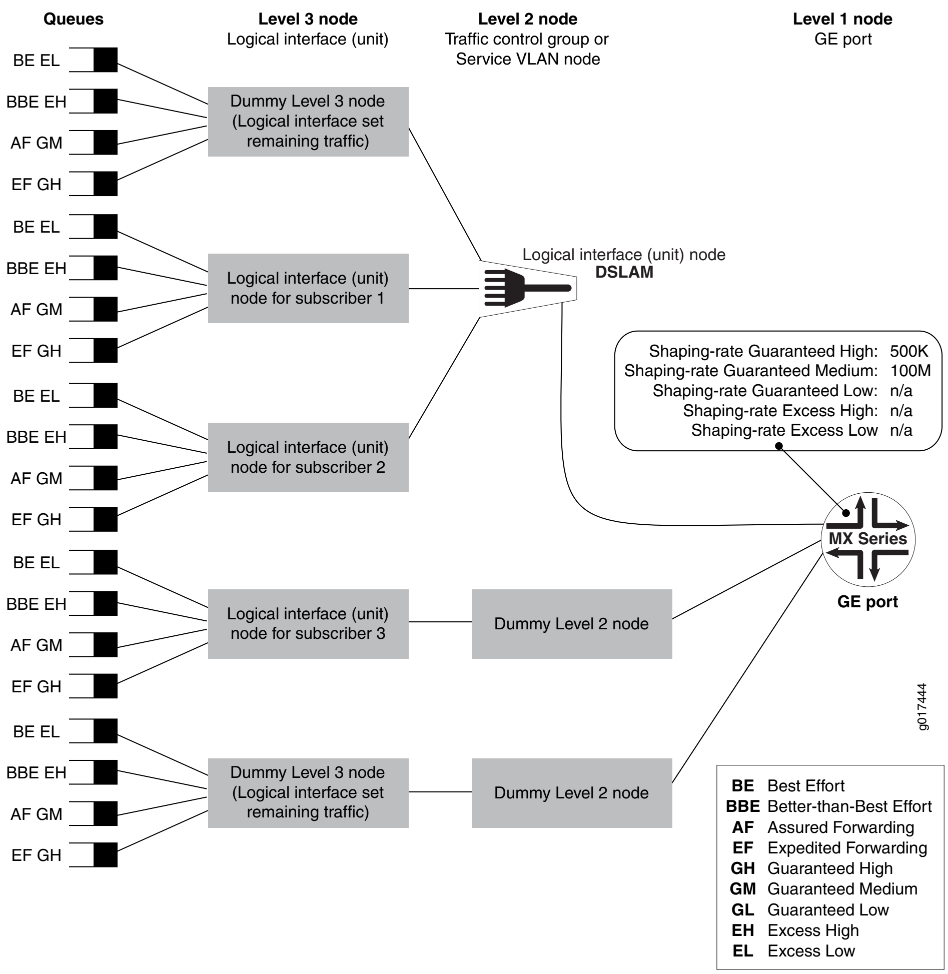
The following behavior aggregate (BA) classification tables are supported on the ingress side of MIC and MPC interfaces:

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

Example: Configuring Per-Priority Shaping on MIC and MPC Interfaces

In practice, per-priority shaping is used with other traffic control profiles to control traffic as a whole. Consider the traffic control profile applied to the physical interface (port), as shown in Figure 10 on page 51.

Figure 10: Example MIC and MPC Interface Scheduling Hierarchy



This example is more complex than those used before. In addition to a pair of subscribers in an interface set (DSLAM), the figure now adds the following:

- A dummy level 3 scheduler node (**interface-set-remaining-traffic**) that provides scheduling for interface set members that do not have explicit class-of-service parameters configured.
- A subscriber (Subscriber 3) that is not a member of an interface set. A dummy level 2 node connects Subscriber 3's level 3 node to level 1, making it appear to be at level 2.
- A dummy level 3 scheduler node (**port-remaining-traffic**) in order to provide queues for traffic that does not have explicit class-of-service parameters configured.
- A dummy level 2 scheduler node to connect level 1 and level 3 scheduler nodes. This dummy level 2 scheduler node is internal only.

This example uses a gigabit Ethernet interface with five logical interface units, each one representing one of the level 3 nodes in Figure 10 on page 51.

From the top of the figure to the bottom, the level 3 nodes are:

- Unit 3 is scheduled as a "dummy" level 3 node because unit 3 is a member of an interface set (**ifset-1**) but there is no explicit CoS configuration.
- Unit 1 is scheduled as a logical interface node for subscriber 1 because unit 1 is a member of an interface set (**ifset-1**) and has an explicit CoS configuration under the **[edit class-of-service interfaces]** hierarchy.
- Unit 2 is scheduled as a logical interface node for subscriber 2 because unit 2 is a member of an interface set (**ifset-1**) and has an explicit CoS configuration under the **[edit class-of-service interfaces]** hierarchy.
- Unit 4 is scheduled as a logical interface node for subscriber 3 because unit 4 is not a member of an interface set but has an explicit CoS configuration under the **[edit class-of-service interfaces]** hierarchy level.
- Unit 5 is scheduled by another "dummy" level 3 node, this one for remaining traffic at the port level, because unit 5 is not a member of an interface set and has no explicit CoS configuration.

In this example, per-priority shaping is applied at the physical port level. The example uses three priorities, but other parameters are possible. The example does not use shaping rates, transmit rates, excess priorities, or other options for reasons of simplicity. The example uses five forwarding classes and leaves out a network control forwarding class that would typically be included in real configurations.

The example configuration is presented in several parts:

- Interfaces configuration
- Class-of-service forwarding classes and traffic control profiles configuration
- Class-of-service interfaces configuration
- Class-of-service schedulers and scheduler map configuration

Interfaces configuration:

```
[edit]
interfaces {
   # A three member interface-set.
   interface-set ifset-1 {
      interface ge-1/1/0 {
         unit 1;
         unit 2;
         unit 3;
      }
   }
   # A ge port configured for "hierarchical-scheduling" and
   # vlans. 5 vlans are configured for the 5 level-3 scheduler
   # nodes
   #
   ge-1/1/0 {
      hierarchical-scheduler;
      vlan-tagging;
      unit 1 {
         vlan-id 1;
      }
      unit 2 {
         vlan-id 2;
      }
      unit 3 {
         vlan-id 3;
      }
      unit 4 {
         vlan-id 4;
      }
      unit 5 {
         vlan-id 5;
      }
   }
}
```

Class-of-service forwarding classes and traffic control profiles configuration:

```
[edit class-of-service]
forwarding-classes {
   queue 0 BE priority low;
   queue 1 BBE priority low;
   queue 2 AF priority low;
   queue 3 EF priority high;
}
traffic-control-profiles {
   tcp-if-portd {
      shaping-rate-priority-high 500k;
      shaping-rate-priority-medium 100m;
   }
   tcp-if-port-rem {
      scheduler-map smap-1;
   }
   tcp-ifset-rem {
      scheduler-map smap-1;
```

```
    }
    tcp-if-unit {
        scheduler-map smap-1;
        shaping-rate 10m;
    }
}
```

Class-of-service interfaces configuration:

```
[edit class-of-service]
interfaces {
    interface-set ifset-1 {
        output-traffic-control-profile-remaining tcp-ifset-rem;
    }
    ge-1/1/0 {
        output-traffic-control-profile tcp-if-port;
        output-traffic-control-profile-remaining tcp-if-port-rem;
        unit 1 {
            output-traffic-control-profile tcp-if-unit;
        }
        unit 2 {
            output-traffic-control-profile tcp-if-unit;
        }
        # Unit 3 present in the interface config and interface-set
        # config, but is absent in this CoS config so that we can
        # show traffic that uses the interface-set
        # remaining-traffic path.
        unit 4 {
            output-traffic-control-profile tcp-if-unit;
        }
        # Unit 5 is present in the interface config, but is absent
        # in this CoS config so that we can show traffic that
        # uses the if-port remaining-traffic path.
    }
}
```

Class-of-service schedulers and scheduler map configuration:

```
[edit class-of-service]
scheduler-maps {
    smap-1 {
        forwarding-class BE scheduler sched-be;
        forwarding-class BBE scheduler sched-bbe;
        forwarding-class AF scheduler sched-af;
    forwarding-class EF scheduler sched-ef;
}
schedulers {
    sched-be {
        priority low;
    }
    sched-bbe {
        priority low;
    }
    sched-af {
        priority medium-high;
    }
    sched-ef {
```



```
        priority high;
    }
}
```

You can configure both a shaping rate and a per-priority shaping rate. In this case, the legacy **shaping-rate** statement specifies the maximum rate for all traffic scheduled through the scheduler. Therefore, the per-priority shaping rates must be less than or equal to the overall shaping rate. So if there is a **shaping-rate 400m** statement configured in a traffic control profile, you cannot configure a higher value for a per-priority shaping rate (such as **shaping-rate-priority-high 500m**). However, the sum of the per-priority shaping rates can exceed the overall shaping rate: for **shaping-rate 400m** you can configure both **shaping-rate-priority-high 300m** and **shaping-rate-priority-low 200m** statements.

Generally, you cannot configure a shaping rate that is smaller than the guaranteed rate (which is why it is guaranteed). However, no such restriction is placed on per-priority shaping rates unless all shaping rates are for priority high or low or medium traffic.

This configuration is allowed (per-priority rates smaller than guaranteed rate):

```
[edit class-of-service]
traffic-control-profile {
    tcp-for-ge-port {
        guaranteed-rate 500m;
        shaping-rate-priority-high 400m;
        shaping-rate-priority-medium 300m;
        shaping-rate-excess-high 100m;
    }
}
```

However, this configuration generates an error (no excess per-priority rate, so the node can never achieve its guaranteed rate):

```
[edit class-of-service]
traffic-control-profile {
    tcp-for-ge-port {
        guaranteed-rate 301m;
        shaping-rate-priority-high 100m;
        shaping-rate-priority-medium 100m;
        shaping-rate-priority-low 100m;
    }
}
```

You verify configuration of per-priority shaping with the **show class-of-service traffic-control-profile** command. This example shows shaping rates established for the high and medium priorities for a traffic control profile named **tcp-ge-port**.

```
user@host# show class-of-service traffic-control-profile
Traffic control profile: tcp-ae, Index: 22093
   Shaping rate: 3000000000
   Scheduler map: <default>

Traffic control profile: tcp-ge-port, Index: 22093
   Shaping rate priority high: 1000000000
   Shaping rate priority medium: 9000000000
```



```
Scheduler map: <default>
```

There are no restrictions on or interactions between per-priority shaping rates and the excess rate. An excess rate (a weight) is specified as a percentage or proportion of excess bandwidth.

Table 6 on page 56 shows where traffic control profiles containing per-priority shaping rates can be attached for both per-unit schedulers and hierarchical schedulers.

Table 6: Applying Traffic Control Profiles

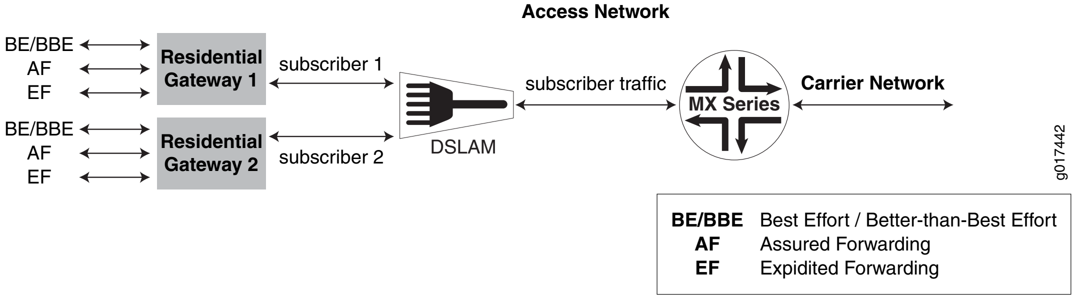
Type of Traffic Control Profile	Per-unit Allowed?	Hierarchical Allowed?
Port level **output-traffic-control-profile** with per-priority shaping	Yes	Yes
Port level **output-traffic-control-profile-remaining** with per-priority shaping	No	Yes
Port level **output-traffic-control-profile** and **output-traffic-control-profile-remaining** with per-priority shaping	No	Yes
Port level **input-traffic-control-profile** with per-priority shaping	No	No
Port level **input-traffic-control-profile-remaining** with per-priority shaping	No	No
Interface set **output-traffic-control-profile** with per-priority shaping	No	Yes
Interface set **output-traffic-control-profile-remaining** with per-priority shaping	No	No
Interface set **input-traffic-control-profile** with per-priority shaping	No	No
Interface set **input-traffic-control-profile-remaining** with per-priority shaping	No	No
Logical interface level **output-traffic-control-profile** with per-priority shaping	No	No
Logical interface level **input-traffic-control-profile** with per-priority shaping	No	No

Example: Configuring Static Shaping Parameters to Account for Overhead in Downstream Traffic Rates

This topic describes two scenarios for which you can configure static shaping parameters to account for packet overhead in a downstream network.

Figure 11 on page 57 shows the sample network that the examples reference.

Figure 11: Sample Network Topology for Downstream Traffic



Managing Traffic with Different Encapsulations

In this example, the MX Series router shown in Figure 11 on page 57 sends stacked VLAN frames to the DSLAM, and the DSLAM sends single-tagged VLAN frames to the residential gateway.

To accurately shape traffic at the residential gateway, the MX Series router must account for the different frame sizes. The difference between the stacked VLAN (S-VLAN) frames sent by the router and the single-tagged VLAN frames received at the residential gateway is a 4-byte VLAN tag. The residential gateway receives frames that are 4 bytes less.

To account for the different frame sizes, the network administrator configures the frame shaping mode with –4 byte adjustment:

1. The network administrator configure the traffic shaping parameters and attaches them to the interface.

   Enabling the overhead accounting feature affects the resulting shaping rate, guaranteed rate, and excess rate parameters, if they are configured.

```
[edit]
class-of-service {
  traffic-control-profiles {
    tcp-example-overhead-accounting-frame-mode {
      shaping-rate 10m;
      shaping-rate-priority-high 4m;
      guaranteed-rate 2m;
      excess-rate percent 50;
      overhead-accounting frame-mode bytes -4;
      }
    }
    interfaces {
      ge-1/0/0 {
        output-traffic-control-profile tcp-example-overhead-accounting-frame-mode;
      }
    }
  }
}
```

2. The network administrator verifies the adjusted rates.

```
user@host#show class-of-service traffic-control-profile
Traffic control profile: tcp-example-overhead-accounting-frame-mode, Index:
61785
```

```
Shaping rate: 10000000
Shaping rate priority high: 4000000
Excess rate 50
Guaranteed rate: 2000000
Overhead accounting mode: Frame Mode
Overhead bytes: -4
```

Managing Downstream Cell-Based Traffic

In this example, the DSLAM and residential gateway shown in Figure 11 on page 57 are connected through an ATM cell-based network. The MX Series router sends Ethernet frames to the DSLAM, and the DSLAM sends ATM cells to the residential gateway.

To accurately shape traffic at the residential gateway, the MX Series router must account for the different physical network characteristics.

To account for the different frame sizes, the network administrator configures the cell shaping mode with –4 byte adjustment:

1. Configure the traffic shaping parameters and attach them to the interface.

   Enabling the overhead accounting feature affects the resulting shaping rate, guaranteed rate, and excess rate parameters, if they are configured.

```
[edit]
class-of-service {
  traffic-control-profiles {
    tcp-example-overhead-accounting-cell-mode {
      shaping-rate 10m;
      shaping-rate-priority-high 4m;
      guaranteed-rate 2m;
      excess-rate percent 50;
      overhead-accounting cell-mode;
      }
    }
    interfaces {
      ge-1/0/0 {
        output-traffic-control-profile tcp-example-overhead-accounting-cell-mode;
      }
    }
  }
}
```

2. Verify the adjusted rates.

```
user@host#show class-of-service traffic-control-profile
Traffic control profile: tcp-example-overhead-accounting-cell-mode, Index:
61785
Shaping rate: 10000000
Shaping rate priority high: 4000000
Excess rate 50
Guaranteed rate: 2000000
Overhead accounting mode: Cell Mode
Overhead bytes: 0
```

   To account for ATM segmentation, the MX Series router adjusts all of the rates by 48/53 to account for ATM AAL5 encapsulation. In addition, the router accounts for cell padding, and internally adjusts each frame by 8 bytes to account for the ATM trailer.

Related Documentation

- Configuring Static Shaping Parameters to Account for Overhead in Downstream Traffic Rates on page 43

CHAPTER 7

Configuration Statements

adjust-minimum

Syntax adjust-minimum *rate*;

Hierarchy Level [edit class-of-service schedulers *scheduler-name*],
[edit class-of-service traffic-control-profiles *traffic-control-profile-name*]

Release Information Statement introduced in Junos OS Release 11.4.

Description For adjustments performed by the ANCP or multicast applications on EQ DPC, MIC, or MPC interfaces, specify the minimum shaping rate for an adjusted scheduler node. The node is associated with a traffic-control profile.

For adjustments performed by the multicast application on MIC or MPC interfaces, specify the minimum shaping rate for an adjusted queue. The queue is associated with a scheduler.

Options *rate*—Minimum shaping rate for a node or a queue, in Mbps

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

Related Documentation

- Configuring the Minimum Adjusted Shaping Rate on Scheduler Nodes for Subscribers

adjust-percent

Syntax	adjust-percent *percentage*;
Hierarchy Level	[edit class-of-service schedulers *scheduler-name*]
Release Information	Statement introduced in Junos OS Release 11.4.
Description	For a MIC or MPC interface, determine the percentage of adjustment for the shaping rate of a queue.
Options	***percentage***—Percentage of the shaping rate to adjust. **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	• Configuring Shaping-Rate Adjustments on Queues

excess-priority

Syntax excess-priority [low | medium-low | medium-high | high | none];

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

Release Information Statement introduced in Junos OS Release 9.3.
Option **none** introduced in Junos OS Release 11.4.

Description Determine the priority of excess bandwidth traffic on a scheduler.

NOTE: For Link Services IQ (LSQ) PICs or Multiservices PIC (MS-PICs), the **excess-priority** statement is allowed for consistency, but ignored. If an explicit priority is not configured for these interfaces, a default low priority is used. This default priority is also used in the excess region.

Options **low**—Excess traffic for this scheduler has low priority.

medium-low—Excess traffic for this scheduler has medium-low priority.

medium-high—Excess traffic for this scheduler has medium-high priority.

high—Excess traffic for this scheduler has high priority.

none—System does not demote the priority of guaranteed traffic when the bandwidth exceeds the shaping rate or the guaranteed rate.

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

Related Documentation

- Configuring Excess Bandwidth Sharing on IQE PICs
- Bandwidth Sharing on Nonqueuing Packet Forwarding Engines Overview
- Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs on page 41

excess-rate

Syntax excess-rate (percent *percentage* | proportion *value*);

Hierarchy Level [edit class-of-service schedulers *scheduler-name*],
[edit class-of-service traffic-control-profiles *traffic-control-profile-name*]

Release Information Statement introduced in Junos OS Release 9.3.
Application to the Multiservices PIC added in Junos OS Release 9.5.
Application to the MIC and MPC interfaces added in Junos OS Release 10.1.
Statement introduced in Junos OS Release 12.1X48R2 for PTX Series Packet Transport Switches.

Description For an Enhanced IQ PIC interfaces, Multiservices PIC interfaces, or MX Series router interfaces on MPCs or MICs, and T4000 router interfaces on Type 5 FPCs, determine the percentage or proportion of excess bandwidth traffic to share.

NOTE: **The proportion option provides a greater range of values over the percent option and hence influences the priorities assigned to the queues.**

Options *percentage*—Percentage of the excess bandwidth to share.

Range: 0 through 100 percent

Default: Excess bandwidth is shared in proportion to the configured transmit rate of each queue.

value—(M Series, MX Series, and T Series routers only) Proportion of the excess bandwidth to share. Option available at the **[edit class-of-service traffic-class-profiles *traffic-control-profile-name*]** hierarchy level only.

Range: 0 through 1000

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

Related Documentation

- Configuring Scheduler Transmission Rate
- Configuring Excess Bandwidth Sharing on IQE PICs
- Allocating Excess Bandwidth Among Frame Relay DLCIs on Multiservices PICs
- Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs on page 41

excess-rate-high

Syntax excess-rate-high (percent *percentage* | proportion *value*);

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

Release Information Statement introduced in Junos OS Release 11.4.

Description For a MIC or MPC interface, determine the percentage of excess bandwidth from high-priority traffic to share.

Options ***percentage***—Percentage of the excess bandwidth to share.

Range: 0 through 100 percent

proportion—Proportion of the excess bandwidth to share.

Range: 0 through 1000

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

Related Documentation

- Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs on page 41

excess-rate-low

Syntax excess-rate-low (percent *percentage* | proportion *value*);

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

Release Information Statement introduced in Junos OS Release 11.4.

Description For a MIC or MPC interface, determine the percentage of excess bandwidth from low-priority traffic to share.

Options ***percentage***—Percentage of the excess bandwidth to share.

Range: 0 through 100 percent

value—Proportion of the excess bandwidth to share.

Range: 0 through 1000

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

Related Documentation

- Managing Excess Bandwidth Distribution on Static Interfaces on MICs and MPCs on page 41

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

shaping-rate-excess-high

Syntax shaping-rate-excess-high *rate* [burst-size *bytes*];

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

Release Information Statement introduced in Junos OS Release 10.1.

Description For MIC and MPC interfaces on MX Series routers, configure a shaping rate and optional burst size for high-priority excess traffic. This can help to make sure higher priority services do not starve lower priority services.

Default If you do not include this statement, the default shaping rate for this priority is determined by the **shaping-rate** statement in the traffic control profile.

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

Default: None

burst-size ***bytes***—Maximum burst size, in bytes.

Range: 0 through 1,000,000,000

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

Related Documentation

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Oversubscribing Interface Bandwidth
- Configuring Traffic Control Profiles for Shared Scheduling and Shaping
- shaping-rate-excess-low on page 68
- shaping-rate-priority-high on page 69
- shaping-rate-priority-low on page 70
- shaping-rate-priority-medium on page 71

shaping-rate-excess-low

Syntax shaping-rate-excess-low *rate* [burst-size *bytes*];

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

Release Information Statement introduced in Junos OS Release 10.1.

Description For MIC and MPC interfaces on MX Series routers, configure a shaping rate and optional burst size for low-priority excess traffic. This can help to make sure higher priority services do not starve lower priority services.

Default If you do not include this statement, the default shaping rate for this priority is determined by the **shaping-rate** statement in the traffic control profile.

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

Default: None

burst-size *bytes*—Maximum burst size, in bytes.

Range: 0 through 1,000,000,000

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

Related Documentation

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Oversubscribing Interface Bandwidth
- Configuring Traffic Control Profiles for Shared Scheduling and Shaping
- shaping-rate-excess-high on page 67
- shaping-rate-priority-high on page 69
- shaping-rate-priority-low on page 70
- shaping-rate-priority-medium on page 71

shaping-rate-priority-high

Syntax shaping-rate-priority-high *rate* [burst-size *bytes*];

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

Release Information Statement introduced in Junos OS Release 10.1.

Description For MIC and MPC interfaces on MX Series routers, configure a shaping rate and optional burst size for high priority traffic. This can help to make sure higher priority services do not starve lower priority services.

Default If you do not include this statement, the default shaping rate for this priority is determined by the **shaping-rate** statement in the traffic control profile.

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

Default: None

burst-size ***bytes***—Maximum burst size, in bytes.

Range: 0 through 1,000,000,000

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

Related Documentation

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Oversubscribing Interface Bandwidth
- Configuring Traffic Control Profiles for Shared Scheduling and Shaping
- **shaping-rate-excess-high on page 67**
- **shaping-rate-excess-low on page 68**
- **shaping-rate-priority-low on page 70**
- **shaping-rate-priority-medium on page 71**

shaping-rate-priority-low

Syntax shaping-rate-priority-low *rate* [burst-size *bytes*];

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

Release Information Statement introduced in Junos OS Release 10.1.

Description For MIC and MPC interfaces on MX Series routers, configure a shaping rate and optional burst size for low priority traffic. This can help to make sure higher priority services do not starve lower priority services.

Default If you do not include this statement, the default shaping rate for this priority is determined by the **shaping-rate** statement in the traffic control profile.

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

Default: None

burst-size *bytes*—Maximum burst size, in bytes.

Range: 0 through 1,000,000,000

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

Related Documentation

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Oversubscribing Interface Bandwidth
- Configuring Traffic Control Profiles for Shared Scheduling and Shaping
- shaping-rate-excess-high on page 67
- shaping-rate-excess-low on page 68
- shaping-rate-priority-high on page 69
- shaping-rate-priority-medium on page 71

shaping-rate-priority-medium

Syntax shaping-rate-priority-medium *rate* [burst-size *bytes*];

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

Release Information Statement introduced in Junos OS Release 10.1.

Description For MIC and MPC interfaces on MX Series routers, configure a shaping rate and optional burst size for medium priority traffic. This can help to make sure higher priority services do not starve lower priority services.

Default If you do not include this statement, the default shaping rate for this priority is determined by the **shaping-rate** statement in the traffic control profile.

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

Default: None

burst-size ***bytes***—Maximum burst size, in bytes.

Range: 0 through 1,000,000,000

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

Related Documentation

- Per-Priority Shaping on MIC and MPC Interfaces Overview on page 17
- Oversubscribing Interface Bandwidth
- Configuring Traffic Control Profiles for Shared Scheduling and Shaping
- **shaping-rate-excess-high on page 67**
- **shaping-rate-excess-low on page 68**
- **shaping-rate-priority-high on page 69**
- **shaping-rate-priority-low on page 70**

traffic-control-profiles

Syntax

```
traffic-control-profiles profile-name {
  adjust-minimum rate;
  atm-service (cbr | rtvbr | nrtvbr);
  delay-buffer-rate (percent percentage | rate);
  peak-rate rate;
  sustained-rate rate;
  max-burst-size cells;
  excess-rate (percent percentage  | proportion value );
  excess-rate-high (percent percentage | proportion value);
  excess-rate-low (percent percentage | proportion value);
  guaranteed-rate (percent percentage | rate) <burst-size bytes>;
  overhead-accounting (frame-mode | cell-mode | frame-mode-bytes | cell-mode-bytes)
    <bytes (byte-value)>;
  scheduler-map map-name;
  shaping-rate (percent percentage | rate) <burst-size bytes>;
  shaping-rate-excess-high rate [ burst-size bytes ];
  shaping-rate-excess-low rate [ burst-size bytes ];
  shaping-rate-priority-high rate [ burst-size bytes ];
  shaping-rate-priority-low rate [ burst-size bytes ];
  shaping-rate-priority-medium rate [ burst-size bytes ];
}
```

Hierarchy Level [edit class-of-service]

Release Information Statement introduced in Junos OS Release 7.6.

Description For Gigabit Ethernet IQ, Channelized IQ PICs, FRF.15 and FRF.16 LSQ interfaces, and Enhanced Queuing (EQ) DPCs only, configure traffic shaping and scheduling profiles. For Enhanced EQ PICs and EQ DPCs only, you can include the **excess-rate** statement.

Options ***profile-name***—Name of the traffic-control 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

- Oversubscribing Interface Bandwidth
- Example: Configuring CoS for a PBB Network on MX Series Routers
- output-traffic-control-profile

PART 3

Administration

CHAPTER 8

Verifying and Managing CoS on MIC and MPC Interfaces

Verifying the Number of Dedicated Queues Configured on MIC and MPC Interfaces

Purpose Display the number of dedicated queue resources that are configured for the logical interfaces on a port.

Action

```
user@host#show class-of-service interface ge-1/1/0
Physical interface: ge-1/1/0, Index: 166
Queues supported: 4, Queues in use: 4
Total non-default queues created: 4
  Scheduler map: <default>, Index: 2
  Chassis scheduler map: <default-chassis>, Index: 4

  Logical interface: ge-1/1/0.100, Index: 72, Dedicated Queues: no
    Shaping rate: 32000
    Object            Name                   Type                 Index
    Scheduler-map     <remaining>                                 0
    Classifier        ipprec-compatibility   ip                   13

  Logical interface: ge-1/1/0.101, Index: 73, Dedicated Queues: no
    Shaping rate: 32000
    Object            Name                   Type                 Index
    Scheduler-map     <remaining>                                 0
    Classifier        ipprec-compatibility   ip                   13

  Logical interface: ge-1/1/0.102, Index: 74, Dedicated Queues: yes
    Shaping rate: 32000
    Object                 Name               Type                Index
    Traffic-control-profile  <control_tc_prof>  Output              45866
```

Related Documentation

- Managing Dedicated and Remaining Queues for Static CoS Configurations on MIC and MPC Interfaces on page 39
- Managing Dedicated and Remaining Queues for Dynamic CoS Configurations on MIC and MPC Interfaces

PART 4

Index

Index