



MQC Hierarchical Queuing with 3 Level Scheduler

The MQC Hierarchical Queuing with 3 Level Scheduler feature provides a flexible packet scheduling and queuing system in which you can specify how excess bandwidth is to be allocated among the subscriber (logical) queues.

History for the MQC Hierarchical Queuing with 3 Level Scheduler Feature

Release	Modification
12.2(31)SB2	This feature was introduced and implemented on the Cisco 10000 series router for the PRE3.

Finding Support Information for Platforms and Cisco IOS Software Images

Use Cisco Feature Navigator to find information about platform support and Cisco IOS software image support. Access Cisco Feature Navigator at <http://www.cisco.com/go/fn>. You must have an account on Cisco.com. If you do not have an account or have forgotten your username or password, click **Cancel** at the login dialog box and follow the instructions that appear.

- [Prerequisites for the Three-Level Scheduler, page 1](#)
- [Restrictions for the Three-Level Scheduler, page 2](#)
- [Information About the MQC Hierarchical Queuing with 3 Level Scheduler, page 2](#)
- [How to Configure Bandwidth-Remaining Ratios, page 6](#)
- [Configuration Examples for the Three-Level Scheduler, page 6](#)
- [Additional References, page 10](#)

Prerequisites for the Three-Level Scheduler

Traffic classes must be configured on the router using the class-map command.

Restrictions for the Three-Level Scheduler

- The priority queue in a child policy must be policed to 90 percent of the parent's shaped bandwidth.
- The three-level scheduler does not support bandwidth propagation. Therefore, you cannot configure a bandwidth guarantee for any queue other than a priority queue.
- To allow oversubscription provisioning, the admission control check is not performed.
- The three-level scheduler does not allocate an implicit bandwidth guarantee for the parent class-default class. Instead, the scheduler uses the ratio of the classes to allocate bandwidth.
- When hierarchical policies are enabled on multiple VLANs and each VLAN hierarchical policy has priority services configured in a child policy, the three-level scheduler first services the priority traffic from all VLANs and then proportionally shares the remaining bandwidth of the interface among all of the VLANs.



Note

The two-level scheduler allocates an equal share of interface bandwidth to each VLAN. After the two-level scheduler serves priority services, best-effort traffic from a VLAN uses the remaining bandwidth. If priority traffic is not configured, instead of proportionally allocating the remaining bandwidth available to each VLAN, the two-level scheduler allocates the entire interface bandwidth to the VLAN's best-effort traffic.

- The sum of all priority traffic running on a given port must be less than or equal to 90 percent of the port bandwidth.

Information About the MQC Hierarchical Queuing with 3 Level Scheduler

The MQC Hierarchical Queuing with 3 Level Scheduler feature provides a flexible packet scheduling and queuing system in which you can specify how excess bandwidth is to be allocated among the subscriber queues and logical interfaces. Rather than allocating an implicit minimum bandwidth guarantee to each queue, the three-level scheduler uses the bandwidth-remaining ratio parameter to allocate unused bandwidth to each logical queue. The three-level scheduler services queues based on the following user-configurable parameters:

- Maximum rate—The specified shape rate of the parent queue.
- Bandwidth-remaining ratio—The value used to determine the portion of unused, non-guaranteed bandwidth allocated to a logical queue relative to other queues competing for the unused bandwidth.



Note

At the class level, the router converts the values specified in the bandwidth bps and bandwidth remaining percent commands to a bandwidth-remaining ratio value. The router does not allow you to configure the bandwidth bps and bandwidth remaining percent commands on the physical and logical layers.

The three-level scheduler on the PRE3 supports priority propagation by propagating the priority guarantees you configure for subscriber services down to the logical interface level. Therefore, the priority traffic is

serviced first at the logical and class level. After servicing the priority traffic bandwidth, the three-level scheduler allocates unused bandwidth to the logical queues based on the configured bandwidth-remaining ratio. In this default case, the three-level scheduler allocates an equal share of the unused bandwidth to each logical queue.

The three-level scheduler supports shaping and scheduling only on the egress interface. The bandwidth command must be configured as a percentage of the available bandwidth or as an absolute bandwidth. You cannot concurrently configure the bandwidth and bandwidth remaining commands on the same class queue or the same policy map.

For more information about the bandwidth-remaining ratio, see the Distribution of Remaining Bandwidth Using Ratio feature module.

Modular QoS Command-Line Interface

The Modular Quality of Service Command-Line Interface (MQC) is designed to simplify the configuration of Quality of Service (QoS) on Cisco routers and switches by defining a common command syntax and resulting set of QoS behaviors across platforms. This model replaces the previous model of defining unique syntaxes for each QoS feature and for each platform.

The MQC contains the following three steps:

- Define a traffic class using the class-map command.
- Create a traffic policy by associating the traffic class with one or more QoS features using the policy-map command.
- Attach the traffic policy to the interface, subinterface, or virtual circuit (VC) using the service-policy command.

For more information about MQC, see the Modular Quality of Service Command-Line Interface document.

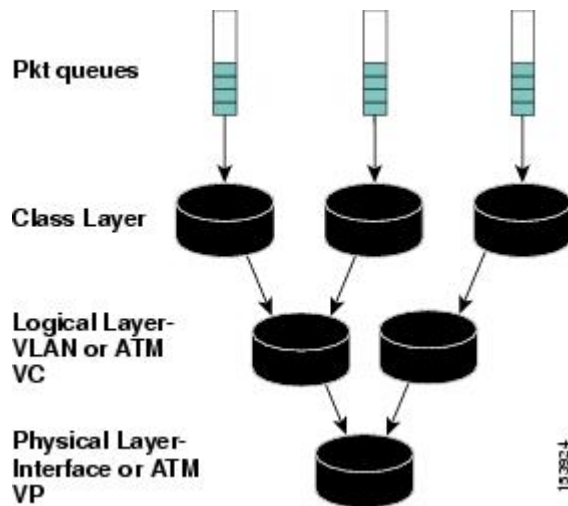
Scheduling Hierarchy

As shown in the figure below, the three-level scheduler uses the following scheduling hierarchy to allocate bandwidth for subscriber traffic:

- Class layer—The three-level scheduler uses virtual-time calendars to schedule class queues and logical interfaces.
- Logical layer (VLAN or ATM VC)—Virtual-time calendars perform weighted round robin based on the weight of the logical interface and the number of bytes dequeued.

- Physical layer (interface or ATM virtual path)—Token buckets ensure that the maximum rate for the class and the logical interface are not exceeded.

Figure 1: Scheduling Hierarchy



The figure above provides an example of how the scheduling hierarchy can apply to Ethernet and ATM topologies. For Ethernet, you cannot oversubscribe the Queue-in-Queue (qinq) into the interface. For ATM, you cannot oversubscribe the virtual path (VP) into the interface.

Table 1: Applying the Scheduling Hierarchy to Ethernet and ATM

Scheduling Hierarchy	Ethernet	ATM
Class layer (virtual time)	MQC-defined queues	MQC-defined queues
Logical layer (virtual time)	VLAN (inner tag)	Virtual channel (VC)
	Session	
Physical (real time)	Queue-in-Queue (outer tag)	Virtual path (VP)
	VLAN (inner tag), if session is the logical layer identifier	

By using VP and VC scheduling with existing Cisco 10000 ATM line cards, the scheduler supports priority propagation: cell-based VP shaping in the segmentation and reassembly (SAR) mechanism with frame-based VC scheduling in the performance routing engine 3 (PRE3).

Priority Service and Latency

The three-level scheduler supports multiple levels of priority service that you can use for such purposes as control traffic, delay-sensitive traffic (for example, voice), minimum guarantees, and excess bandwidth

allocation. Each level of priority supports multiple queues, which allows for multiple types of delay-sensitive traffic (for example, voice and video).

The three-level scheduler can service the same queue from multiple levels of priority service. For example, the three-level scheduler uses priority level 1 for voice, priority level 2 for video, and the excess bandwidth for data.

For a priority class with policing configured, the three-level scheduler always polices the priority traffic to the rate specified in the police command (1000 kbps as shown in the following example configuration), regardless of whether or not the underlying interface is congested.

```
Router(config-pmap-c)# priority
Router(config-pmap-c)# police 1000
```

**Note**

The three-level scheduler does not support the priority kbps command.

Latency Requirements

Delay-sensitive traffic incurs a maximum of 10 milliseconds (ms) of latency on edge router interfaces and a maximum of 1 ms of latency on core router interfaces. For interface speeds at T1/E1 and below, the three-level scheduler services 2 maximum transmission units (MTUs) of nonpriority traffic before servicing a priority packet. Requirements for high-speed interfaces are not as strict as 2 MTUs, but are always bound by 10 ms on edge interfaces and 1 ms on core interfaces.

The three-level scheduler also supports the minimal latency requirement (2 MTUs of nonpriority traffic in front of priority traffic) at the physical link rate. However, in some cases, it is impossible for the three-level scheduler to service all competing packets with a latency of 2 MTUs. For example, if many priority packets compete at the same time for bandwidth, the last one serviced may incur latency that is greater than 2 MTUs.

The table below lists the maximum latency requirements for various interface speeds.

Table 2: Maximum Latency Requirements

Interface Speed	Maximum Latency
Greater than 2 Mbps	2 MTU + 6 ms
2 Mbps to 1 Gbps	2 MTU
1 Gbps or greater	1 ms

Priority Propagation with Imposed Burstiness

A single physical interface can have large numbers of logical interfaces and each of these logical interfaces can have both priority and nonpriority traffic competing for the physical link. To minimize latency, the priority traffic of one logical interface has priority over the nonpriority traffic of other logical interfaces, thereby imposing burstiness on the minimum rate traffic of other logical interfaces. The latency that the priority traffic incurs results from the rate constraining the delivered rate of the priority traffic. In many cases, this constraining rate is not the rate of the priority class's parent policy.

For example, suppose a 10 Gigabit Ethernet (GE) interface has 100 VLANs that are shaped to various rates. Each VLAN has a priority class and additional classes configured. Through priority propagation, the scheduler delivers latency to the priority traffic based on the 10 GE rate and not the VLAN rate.

**Note**

The VLAN rate is at most 1 to 2 MTUs of nonpriority traffic in front of priority traffic, which would bound the latency incurred by priority traffic (due to non-priority traffic) at 1 to 2 MTUs served at the 10 GE rate.

The priority traffic of one logical interface cannot only impose burstiness on other traffic, but also starve other traffic. The only way to prevent the starvation of other traffic is by configuring a policer on the priority queue by limiting the percent of priority traffic to less than 90 percent of the parent bandwidth and the port bandwidth.

Configuration Granularity

The table below describes the configuration granularity for the three-level scheduler.

Table 3: Three-Level Scheduler Configuration Granularity

Interface Bandwidth	Granularity
Less than or equal to 2 Mbps	.4%
Greater than 2 Mbps and less than 1 Gbps	.2%
Greater than or equal to 1 Gbps	.1%

How to Configure Bandwidth-Remaining Ratios

To configure bandwidth-remaining ratios on subinterface-level and class-level queues, see the Distribution of Remaining Bandwidth Using Ratio, Release 12.2(31)SB2 feature module.

Configuration Examples for the Three-Level Scheduler

This section provides the following configuration examples:

Bandwidth Allocation—Policy Attached to an Interface Example

The following example configuration consists of one policy map named Child with the following traffic classes defined: prec0, prec2, and class-default. The policy is attached to the ATM interface 1/0/0, which has a configured rate of 1000 kbps.

```
policy-map Child
class prec0
bandwidth 300
class prec2
```

```

bandwidth 100
class class-default
bandwidth 50
!
interface atm 1/0/0
bandwidth 1000
service-policy output Child

```

Assuming that the traffic flow through each class is enough to require maximum possible bandwidth, the three-level scheduler allocates bandwidth as described in the table below.

Table 4: Queuing Presentation—Policy Attached to an Interface

Traffic Class	Bandwidth Ratio	Total Bandwidth Allocated
prec0	6	666 kbps
prec2	2	222 kbps
class-default	1	111 kbps

Bandwidth Allocation—Parent Policy Attached to Two Subinterfaces Example

The following example configuration contains a hierarchical policy consisting of two policy maps: Child and Parent. The Child policy has two traffic classes (voice and video) with each configured as a priority class with policing enabled. The Parent policy has its class-default class shaped to 1000 kbps. The Parent policy is attached to the ATM subinterface 1/0/1.1 and to subinterface 1/0/1.2. ATM interface 1/0/1 has a configured rate of 2100 kbps.

```

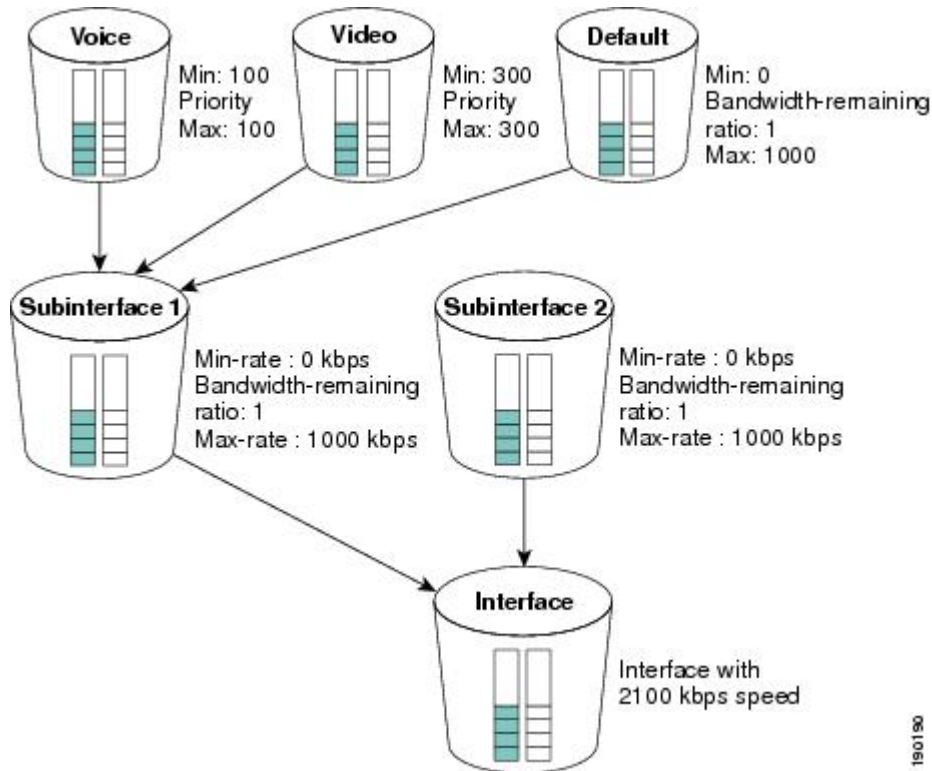
policy-map Child
class voice
priority level 1
police 100
!
class video
priority level 2
police 300
!
policy-map Parent
class class-default
shape average 1000
service-policy Child
!
interface atm 1/0/1
atm pvp 1 1400
!
interface atm 1/0/1.1
bandwidth remaining ratio 1
service-policy output Parent
!
interface atm 1/0/1.2
bandwidth remaining ratio 1
service-policy output Parent
!

```

The figure below shows an example of the queuing presentation based on the above configuration. The service rates for all Child classes under each subinterface might differ from the rates shown in the figure below,

depending on the presence or absence of priority propagation and how the class's bandwidth usage is accounted against the Parent queue.

Figure 2: Queuing Presentation—Parent Enabled on Two Subinterfaces



Each subinterface receives an equal share of bandwidth. Based on the bandwidth-remaining ratio of 1, each subinterface-level queue receives a rate of 700 kbps (subinterfaces 1 and 2 queues, and default queue at subinterface-level).

- For subinterface 1, assume that only the voice traffic is active. From the 700-kbps bandwidth allocated to subinterface 1, the voice traffic receives a bandwidth rate of 100 kbps and the default traffic receives a rate of 600 kbps.
- For subinterface 2, assume that only the video traffic is active. From the 700-kbps bandwidth allocated to subinterface 2, the video traffic receives a bandwidth rate of 300 kbps and the default traffic receives a rate of 400 kbps.

Tuning the Bandwidth-Remaining Ratio Example

The following example configuration shows how to tune the bandwidth-remaining ratio using the bandwidth remaining ratio command. In the example, the class-default class of Parent1 has a bandwidth-remaining ratio of 9 and the class-default class of Parent2 has a bandwidth-remaining ratio of 7.

```
policy-map Child
class prec0
priority level 1
police 100
```



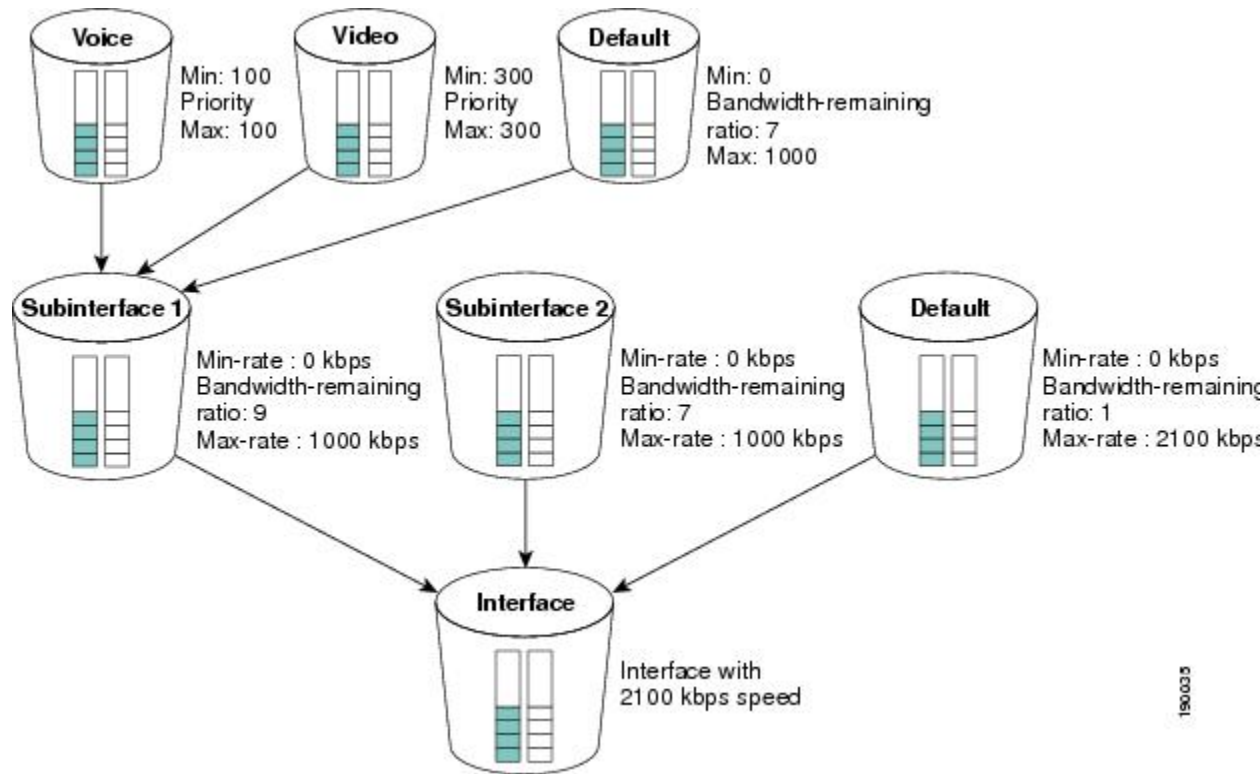
```

!
class prec2
priority level 2
police 300
!
policy-map Parent1
class class-default
shape average 10000
bandwidth remaining ratio 9
!
policy-map Parent2
class class-default
shape average 1000
bandwidth remaining ratio 7

```

The figure below shows an example of the queuing presentation based on the above configuration and assuming that the Parent1 policy is enabled on subinterface 1 and the Parent2 policy is enabled on subinterface 2, and that the interface speed is 2100 kbps.

Figure 3: Queuing Presentation—Tuning the Bandwidth-Remaining Ratio



Based on the preceding configuration, the three-level scheduler distributes bandwidth in the following way (assuming that the voice traffic is active on subinterface 1 only and the video traffic is active on subinterface 2 only):

- A total of 400 kbps of bandwidth is used from the interface: 100 kbps-bandwidth guarantee for voice traffic on subinterface 1 and 300-kbps bandwidth guarantee for video traffic on subinterface 2.
- The remaining 1700-kbps bandwidth is distributed across the subinterface-level queues based on their bandwidth-remaining ratios:
 - Subinterface 1 with bandwidth-remaining ratio 9 receives 956 kbps

- Subinterface 2 with bandwidth-remaining ratio 7 receives 743 kbps

Additional References

The following sections provide references related to the MQC Hierarchical Queuing with 3 Level Scheduler feature.

Related Documents

Related Topic	Document Title
Bandwidth	<i>Cisco 10000 Series Router Quality of Service Configuration Guide</i> “Distributing Bandwidth Between Queues”
Bandwidth-remaining ratio	“Distribution of Remaining Bandwidth Using Ratio” feature module
Hierarchical policies	<i>Cisco 10000 Series Router Quality of Service Configuration Guide</i> “Defining QoS for Multiple Policy Levels”
Policy maps	<i>Cisco 10000 Series Router Quality of Service Configuration Guide</i> “Configuring QoS Policy Actions and Rules”
Shaping traffic	<i>Cisco IOS Quality of Service Solutions Configuration Guide</i> , Release 12.2 Part 4: Policing and Shaping > Configuring Class-Based Shaping Part 4: Policing and Shaping > Policing and Shaping Overview > Traffic Shaping > Class-Based Shaping
Traffic policing and shaping	Comparing Traffic Policing and Traffic Shaping for Bandwidth Limiting

Standards

Standard	Title
No new or modified standards are supported by this feature, and support for existing standards has not been modified by this feature.	—

MIBs

MIB	MIBs Link
No new or modified MIBs are supported by this feature, and support for existing MIBs has not been modified by this feature.	To locate and download MIBs for selected platforms, Cisco IOS releases, and feature sets, use Cisco MIB Locator found at the following URL: http://www.cisco.com/go/mibs

RFCs

RFC	Title
No new or modified RFCs are supported by this feature, and support for existing RFCs has not been modified by this feature.	—

Technical Assistance

Description	Link
The Cisco Technical Support & Documentation website contains thousands of pages of searchable technical content, including links to products, technologies, solutions, technical tips, and tools. Registered Cisco.com users can log in from this page to access even more content.	http://www.cisco.com/techsupport

