CHAPTER 57

Configuring Fabric Congestion Control and QoS

Fibre Channel Congestion Control (FCC) is a Cisco proprietary flow control mechanism that alleviates congestion on Fibre Channel networks.

Quality of service (QoS) offers the following advantages:

- Provides relative bandwidth guarantee to application traffic.
- Controls latency experienced by application traffic.
- Prioritizes one application over another (for example, prioritizing transactional traffic over bulk traffic) through bandwidth and latency differentiation.

This chapter provides details on the QoS and FCC features provided in all switches. It includes the following sections:

- FCC, page 57-1
- QoS, page 57-3
- Example Configuration, page 57-10
- Ingress Port Rate Limiting, page 57-11
- Default Settings, page 57-12

FCC

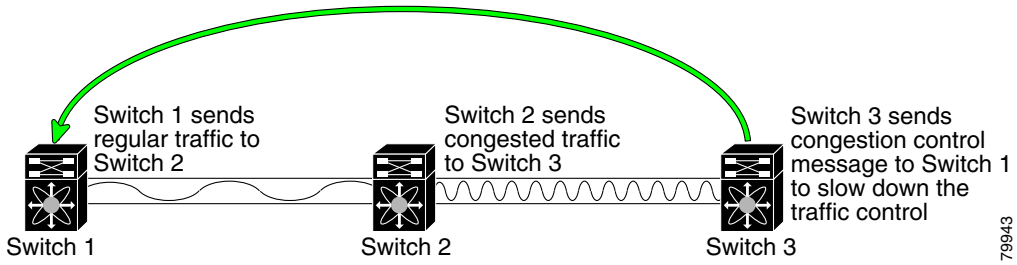
FCC reduces the congestion in the fabric without interfering with the standard Fibre Channel protocols. This section contains the following topics:

- About FCC, page 57-1
- FCC Process, page 57-2
- Enabling FCC, page 57-2
- Assigning FCC Priority, page 57-3

About FCC

The FCC protocol increases the granularity and the scale of congestion control applied to any class of traffic (see Figure 57-1).

Figure 57-1 FCC Mechanisms



Edge quench congestion control provides feedback to the source about the rate at which frames should be injected into the network (frame intervals).

Note FCC is not supported on the Cisco Fabric Switch for HP c-Class BladeSystem and Cisco Fabric Switch for IBM BladeCenter.

FCC Process

When a node in the network detects congestion for an output port, it generates an edge quench message. These frames are identified by the Fibre Channel destination ID (DID) and the source ID. A switch from other vendors simply forwards these frames.

Any receiving switch in the Cisco MDS 9000 Family handles frames in one of these ways:

- It forwards the frame.
- It limits the rate of the frame flow in the congested port.

The behavior of the flow control mechanism differs based on the Fibre Channel DID:

- If the Fibre Channel DID is directly connected to one of the switch ports, the input rate limit is applied to that port.
- If the destination of the edge quench frame is a Cisco domain or the next hop is a Cisco MDS 9000 Family switch, the frame is forwarded.
- If neither of these mechanisms is true, then the frame is processed in the port going towards the FC DID.

All switches (including the edge switch) along the congested path process path quench frames. However, only the edge switch processes edge quench frames.

Enabling FCC

By default, the FCC protocol is disabled. FCC can only be enabled for the entire switch.

Tip If you enable FCC, be sure to enable it in all switches in the fabric.

To enable or disable the FCC feature using Fabric Manager, follow these steps:

Step 1 Expand **Switches**, expand **FC Services** and then select **FCC** in the Physical Attributes pane.

The FCC information is displayed in the Information pane. The **General** tab is the default.

Step 2 Select the switch on which you want to enable FCC.

Step 3 Check the **Enable** check box.

Step 4 Click **Apply Changes to save your changes**.

Assigning FCC Priority

To assign FCC priority using Fabric Manager, follow these steps:

Step 1 Expand **Switches**, expand **FC Services** and then select **FCC** in the Physical Attributes pane.

The FCC information is displayed in the Information pane. The **General** tab is the default.

Step 2 Select the switch for which you want to assign the FCC priority.

Step 3 Enter the priority in the **Priority** column.

Step 4 Click **Apply Changes to save your changes**.

QoS

QoS implementation in the Cisco MDS 9000 Family follows the differentiated services (DiffServ) model. The DiffServ standard is defined in RFCs 2474 and 2475.

All switches support the following types of traffic:

- About Control Traffic, page 57-3
- Enabling or Disabling Control Traffic, page 57-4
- About Data Traffic, page 57-4
- VSAN Versus Zone-Based QoS, page 57-5
- Configuring Data Traffic, page 57-6
- About Class Map Creation, page 57-6
- Creating a Class Map, page 57-7
- About Service Policy Definition, page 57-8
- About Service Policy Enforcement, page 57-8
- About the DWRR Traffic Scheduler Queue, page 57-8
- Changing the Weight in a DWRR Queue, page 57-9

About Control Traffic

The Cisco MDS 9000 Family supports QoS for internally and externally generated control traffic. Within a switch, control traffic is sourced to the supervisor module and is treated as a high priority frame. A high priority status provides absolute priority over all other traffic and is assigned in the following cases:

- Internally generated time-critical control traffic (mostly Class F frames).
- Externally generated time-critical control traffic entering a switch in the Cisco MDS 9000 Family from a another vendor's switch. High priority frames originating from other vendor switches are marked as high priority as they enter a switch in the Cisco MDS 9000 Family.

Enabling or Disabling Control Traffic

By default, the QoS feature for certain critical control traffic is enabled. These critical control frames are assigned the highest (absolute) priority.

Tip We do not recommend disabling this feature as all critical control traffic is automatically assigned the lowest priority once you issue this command.

Toenable or disable the high priority assignment for control traffic using Fabric Manager, follow these steps:

Step 1 Expand **Switches**, expand **FC Services** and then select **QoS** in the Physical Attributes pane.

The QoS control traffic information is displayed in the Information pane. The **Control** tab is default.

Step 2 Select the switch on which you want to enable or disable control traffic.

Step 3 In the Command column, click the drop-down menu and select **enable** or **disable**.

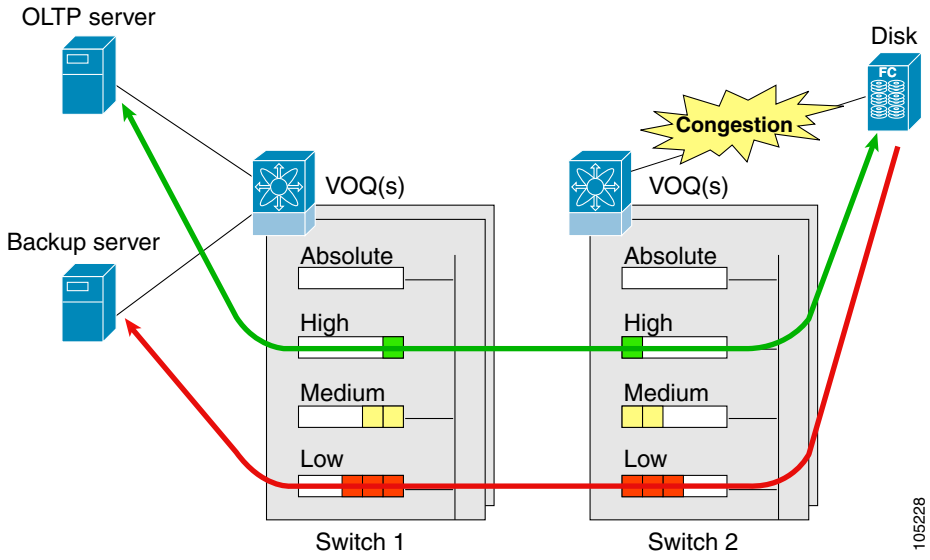
Step 4 Click **Apply Changes to save your changes**.

About Data Traffic

Online transaction processing (OLTP), which is a low volume, latency sensitive application, requires quick access to requested information. Backup processing application require high bandwidth but are not sensitive to latency. In a network that does not support service differentiation, all traffic is treated identically—they experience similar latency and are allocated similar bandwidths. The QoS feature in the Cisco MDS 9000 Family switches provides these guarantees.

Data traffic can be prioritized in distinct levels of service differentiation: low, medium, or high priority. You can apply QoS to ensure that Fibre Channel data traffic for your latency-sensitive applications receive higher priority over throughput-intensive applications such as data warehousing (see Figure 57-2).

Figure 57-2 Prioritizing Data Traffic



In Figure 57-2, the OLTP traffic arriving at Switch 1 is marked with a high priority level of throughput classification (class map) and marking (policy map). Similarly, the backup traffic is marked with a low priority level. The traffic is sent to the corresponding priority queue within a virtual output queue (VOQ).

A deficit weighted round robin (DWRR) scheduler configured in the first switch ensures that high priority traffic is treated better than low priority traffic. For example, DWRR weights of 70:20:10 implies that the high priority queue is serviced at 7 times the rate of the low priority queue. This guarantees lower delays and higher bandwidths to high priority traffic if congestion sets in. A similar configuration in the second switch ensures the same traffic treatment in the other direction.

If the ISL is congested when the OLTP server sends a request, the request is queued in the high priority queue and is serviced almost immediately since the high priority queue is not congested. The scheduler assigns its priority over the backup traffic in the low priority queue.

Note When the high priority queue does not have traffic flowing through, the low priority queue uses all the bandwidth and is not restricted to the configured value.

A similar occurrence in Switch 2 sends a response to the transaction request. The round trip delay experienced by the OLTP server is independent of the volume of low priority traffic or the ISL congestion. The backup traffic uses the available ISL bandwidth when it is not used by the OLTP traffic.

Tip To achieve this traffic differentiation, be sure to enable FCC (see the "Enabling FCC" section on page 57-2).

VSAN Versus Zone-Based QoS

While you can configure both zone-based QoS and VSAN-based QoS configurations in the same switch, both configurations have significant differences. Table 57-1 highlights the differences between configuring QoS priorities based on VSANs versus zones.

Table 57-1 QoS Configuration Differences

VSAN-Based QoS	Zone-Based QoS
If you configure the active zone set on a given VSAN and also configure QoS parameters in any of the member zones, you cannot associate the policy map with the VSAN.	You cannot activate a zone set on a VSAN that already has a policy map associated.
If the same flow is present in two class maps associated to a policy map, the QoS value of the class map attached first takes effect.	If the same flow is present in two zones in a given zone set with different QoS values, the higher QoS value is considered.
—	During a zone merge, if the Cisco SAN-OS software detects a mismatch for the QoS parameter, the link is isolated.
Takes effect only when QoS is enabled.	Takes effect only when QoS is enabled.

See the "About Zone-Based Traffic Priority" section on page 30-33 for details on configuring a zone-based QoS policy.

Configuring Data Traffic

To configure QoS using Fabric Manager, follow these steps:.

Step 1 Enable the QoS feature.

Step 2 Create and define class maps.

Step 3 Define service policies.

Step 4 Apply the configuration.

Tip QoS is supported in interoperability mode. For more information, refer to the *Cisco MDS 9000 Family Switch-to-Switch Interoperability Configuration Guide*.

About Class Map Creation

Use the class map feature to create and define a traffic class with match criteria to identify traffic belonging to that class. The class map name is restricted to 63 alphanumeric characters and defaults to the match-all option. Flow-based traffic uses one of the following values:

- WWN—The source WWN or the destination WWN.
- Fibre Channel ID (FC ID) —The source ID (SID) or the destination ID (DID). The possible values for mask are FFFFFF (the entire FC ID is used—this is the default), FFFF00 (only domain and area FC ID is used), or FF0000 (only domain FC ID is used).

Note An SID or DID of 0x000000 is not allowed.

- Source interface—The ingress interface.

Tip The order of entries to be matched within a class map is not significant.

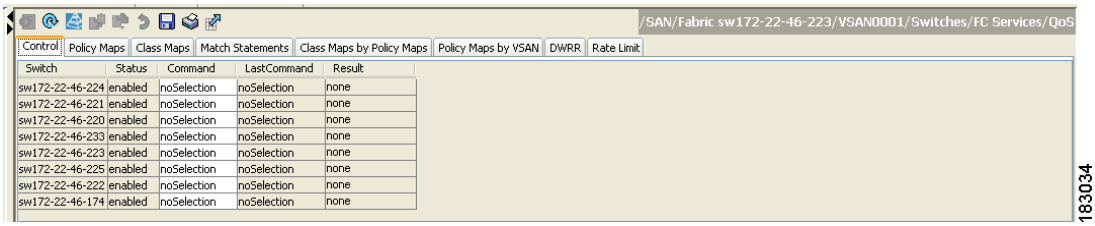
Creating a Class Map

To create a class map using Fabric Manager, follow these steps:

Step 1 Expand **Switches**, expand **FC Services** and then select **QoS** in the Physical Attributes pane.

The QoS information is displayed in the Information pane shown in Figure 57-3. The **Control** tab is the default.

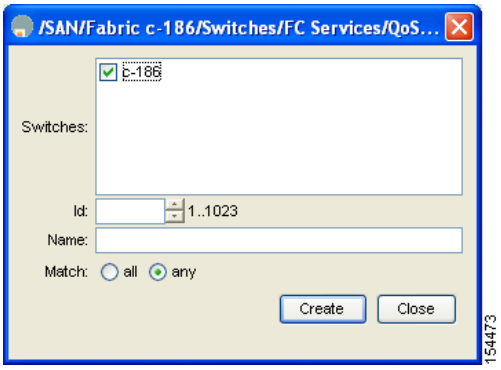
Figure 57-3 Quality of Service Control Tab



Step 2 In the **Class Maps** tab, click **Create Row** to create a new class map.

You see the Create Class Maps dialog box shown in Figure 57-4.

Figure 57-4 Create Class Maps Dialog Box



Step 3 Select the switches for the class map.

Step 4 Enter the source ID or the destination **ID** in the field.

Step 5 Enter a name for the class map.

Step 6 Select a Match mode.You can either match **any** or **all** criterion with one match statement from the class map configuration mode.

Step 7 Click **Create** to proceed with creating the class map.

About Service Policy Definition

Service policies are specified using policy maps. Policy maps provide an ordered mapping of class maps to service levels. You can specify multiple class maps within a policy map, and map a class map to a high, medium, or low service level. The default priority is low. The policy map name is restricted to 63 alphanumeric characters.

As an alternative, you can map a class map to a differentiated services code point (DSCP).The DSCP is an indicator of the service level for a specified frame. The DSCP value ranges from 0 to 63, and the default is 0. A DSCP value of 46 is disallowed.

The order of the class maps within a policy map is important to determine the order in which the frame is compared to class maps. The first matching class map has the corresponding priority marked in the frame.

Note Refer to http://www.cisco.com/warp/public/105/dscpvalues.html#dscpandassuredforwardingclasses for further information on implementing QoS DSCP values.

Note Class maps are processed in the order in which they are configured in each policy map.

About Service Policy Enforcement

When you have configured a QoS data traffic policy, you must enforce the data traffic configuration by applying that policy to the required VSAN(s). If you do not apply the policy to a VSAN, the data traffic configuration is not enforced. You can only apply one policy map to a VSAN.

Note You can apply the same policy to a range of VSANs.

About the DWRR Traffic Scheduler Queue

The Cisco SAN-OS software supports four scheduling queues:

- Strict priority queues are queues that are serviced in preference to other queues—it is always serviced if there is a frame queued in it regardless of the state of the other queues.
- QoS assigns all other traffic to the DWRR scheduling high, medium, and low priority traffic queues.

The DWRR scheduler services the queues in the ratio of the configured weights. Higher weights translate to proportionally higher bandwidth and lower latency. The default weights are 50 for the high queue, 30 for the medium queue, and 20 for the low queue. Decreasing order of queue weights is mandated to ensure the higher priority queues have a higher service level, though the ratio of the configured weights can vary (for example, one can configure 70:30:5 or 60:50:10 but not 50:70:10).

Table 57-2 describes the QoS behavior for Generation 1 and Generation 2 switching modules.

Table 57-2 QoS Behavior for Generation 1 and Generation 2 Switching Modules

Source Module Type	Destination Module Type	QoS Behavior Description
Generation 1	Generation 1	QoS behavior reflects the DWRR configuration for traffic coming in through a given port and queued to the same egress port. All the other traffic share equal bandwidth.
Generation 1	Generation 2	QoS behavior reflects the DWRR configuration for traffic coming in through a given port and queued to the same egress port. All the other streams share equal bandwidth.
Generation 2	Generation 1	Bandwidth partitioning is equal for all the traffic.
Generation 2	Generation 2	QoS behavior reflects the DWRR weights configuration for all possible streams.

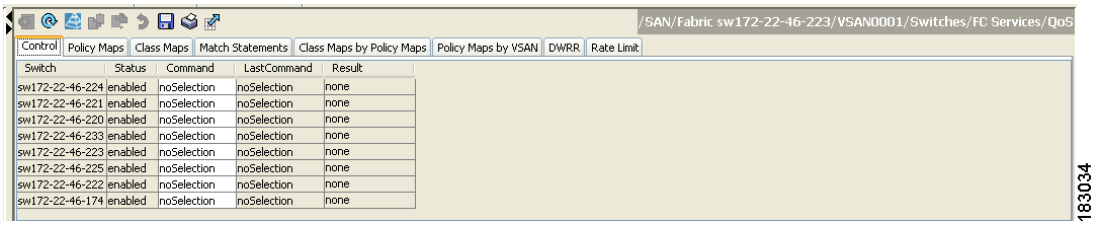
Changing the Weight in a DWRR Queue

To change the weight in a DWRR queue using Fabric Manager, follow these steps:

Step 1 Expand **Switches**, expand **FC Services** and then select **QoS** in the Physical Attributes pane.

The QoS control traffic information is displayed in the Information pane shown in Figure 57-5. The default is the **Control** tab.

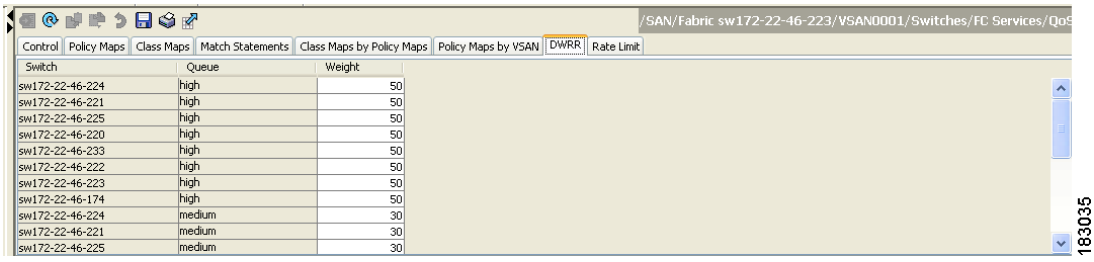
Figure 57-5 Quality of Service Control Tab



Step 2 Click the **DWRR** tab.

You see the queue status and weight (see Figure 57-6).

Figure 57-6 QoS Queue Status and Weight



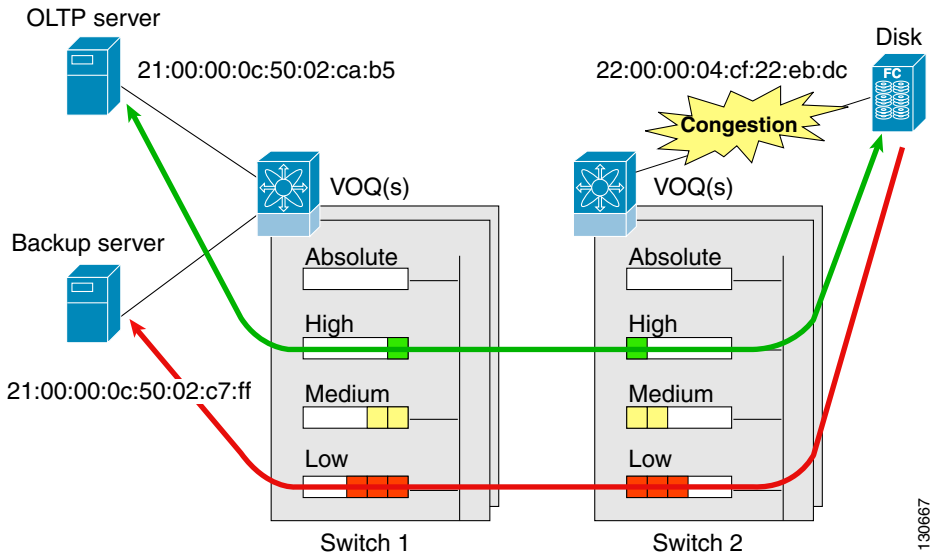
Step 3 Select a switch and change the weight.

Step 4 Click the **Apply Changes icon to save your changes**.

Example Configuration

This section describes a configuration example for the application illustrated in Figure 57-7.

Figure 57-7 Example Application for Traffic Prioritization



Both the OLTP server and the backup server are accessing the disk. The backup server is writing large amounts of data to the disk. This data does not require specific service guarantees. The volumes of data generated by the OLTP server to the disk are comparatively much lower but this traffic requires faster response because transaction processing is a low latency application.

The point of congestion is the link between Switch 2 and the disk, for traffic from the switch to the disk. The return path is largely uncongested as there is little backup traffic on this path.

Service differentiation is needed at Switch 2 to prioritize the OLTP-server-to-disk traffic higher than the backup-server-to-disk traffic.

To configure traffic prioritization for the example application, follow these steps:

Step 1 Create the class maps.

Step 2 Create the policy map.

Step 3 Assign the service policy.

Step 4 Assign the weights for the DWRR queues.

Step 5 Repeat Step 1 through Step 4 on Switch 1 to address forward path congestion at both switches.

Congestion could occur anywhere in the example configuration. To address congestion of the return path at both switches, you need to create two more class maps and include them in the policy map as follows:

Step 1 Create two more class maps.

Step 2 Assign the class maps to the policy map.

Step 3 Repeat Step 1 through Step 2 on Switch 1 to address return path congestion at both switches.

Ingress Port Rate Limiting

A port rate limiting feature helps control the bandwidth for individual Fibre Channel ports. Port rate limiting is also referred to as ingress rate limiting because it controls ingress traffic into a Fibre Channel port. The feature controls traffic flow by limiting the number of frames that are transmitted out of the exit point on the MAC. Port rate limiting works on all Fibre Channel ports. The rate limit ranges from 1 to 100% and the default is 100%.

Note Port rate limiting can only be configured on Cisco MDS 9100 Series switches, Cisco MDS 9216i switches, and MPS-14/2 modules.

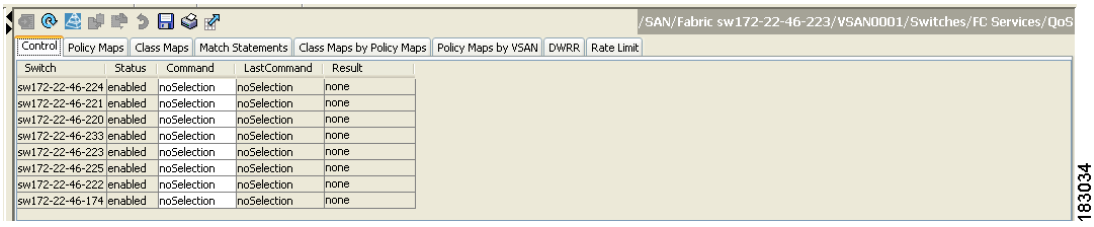
This feature can only be configured if the QoS feature is enabled and if this configuration is performed on a Cisco MDS 9100 series switch, Cisco MDS 9216i switch, or MPS-14/2 module.

To configure the port rate limiting value using Fabric Manager, follow these steps:

Step 1 Expand **Switches**, expand **FC Services** and then select **QoS** in the Physical Attributes pane.

The QoS control traffic information is displayed in the Information pane shown in Figure 57-8. The default is the **Control** tab.

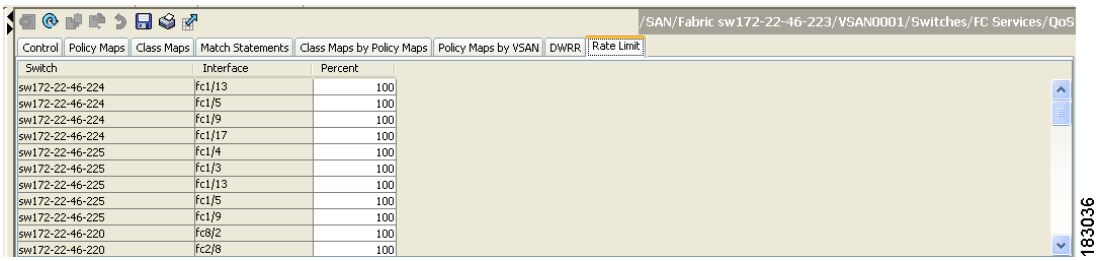
Figure 57-8 Quality of Service Control Tab



Step 2 Click the **Rate Limit** tab.

You see the information shown in Figure 57-9.

Figure 57-9 Rate Limits for Switch Interfaces



Step 3 Select the switch whose port rate limit you want to change.

Step 4 Enter the desired port rate limit in the Percent column.

Step 5 Click the **Apply Changes** icon to save your changes.

Default Settings

Table 57-3 lists the default settings for FCC, QoS, and rate limiting features.

Table 57-3 Default FCC, QoS, and Rate Limiting Settings

Parameters	Default
FCC protocol	Disabled.
QoS control traffic	Enabled.
QoS data traffic	Disabled.
Zone-based QoS priority	Low.
Rate limit	100%