



CHAPTER 28

Configuring QoS

QoS refers to the capability of a network to provide better service to selected network traffic over various technologies for the best overall services with limited bandwidth of the underlying technologies.

The primary goal of QoS in the security appliance is to provide rate limiting on selected network traffic for both individual flow and VPN tunnel flow to ensure that all traffic gets its fair share of limited bandwidth. A flow can be defined in a number of ways. In the security appliance, QoS can apply to a combination of source and destination IP addresses, source and destination port number, and the TOS byte of the IP header.

This section contains the following topics:

- [Configuring a QoS Service Policy, page 28-1](#)
- [Priority Queue, page 28-3](#)

Configuring a QoS Service Policy

A QoS service policy is created in the same manner as a regular service policy. This procedure provides an overview of the regular service policy creation process and focuses on the configuration of the QoS features for that service policy. For more information about creating service policy rules, see [Adding a Service Policy Rule for Through Traffic, page 23-4](#).

To configure a QoS service policy, perform the following steps:

-
- Step 1** Open the **Configuration > Firewall > Service Policy Rules** pane.
 - Step 2** Click **Add** to create a new service policy rule.
The Service Policy Wizard opens.
 - Step 3** Define the scope of the service policy rule. The service policy rule can apply globally (to all interfaces) or to a specific interface. Click **Next**.
 - Step 4** Define the traffic matched by the service policy. Depending upon the match criteria that you select, you may have several wizard screens to walk through. See [Adding a Service Policy Rule for Through Traffic, page 23-4](#) for more information about configuring the match criteria. Click **Next**.
The Rule Actions screen appears.
 - Step 5** Click the **QoS** tab.
 - Step 6** To configure QoS, perform one of the following actions:

- To define the specified traffic as high priority traffic, click **Enable priority for this flow**. This defined the traffic as high priority and allows you to configure priority queues for it. You cannot enable traffic policing if you select this option.
 - To configuring rate limiting on the traffic, click **Enable policing**. This allows you to limit the input or the output (or both) traffic rates, define the burst rate, and specify the action to take on conforming and nonconforming traffic. See [QoS Tab Field Information, page 28-2](#), for more information about these settings.
- Step 7** Click **Finish**. The service policy rule is added to the rule table. Click **Apply** to send the configuration to the device.
- Step 8** If you enabled priority for the traffic, you must configure the priority queue for the specific interfaces. See [Priority Queue, page 28-3](#), for information about configuring the priority queue.

QoS Tab Field Information

The **QoS** tab lets you apply strict scheduling priority and rate-limit traffic.

Restrictions

If a service policy is applied or removed from an interface that has existing VPN client/LAN-to-LAN or non-tunneled traffic already established, the QoS policy is not applied or removed from the traffic stream. To apply or remove the QoS policy for such connections, you must clear (that is, drop) the connections and reestablish them.

Fields

- **Enable Priority for this flow**—Enables or disables strict scheduling priority for this flow. The priority does not take effect until the priority queues are set. To configure priority queues, see [Priority Queue, page 28-3](#).
- **Enable policing**—Check this check box to enable input and output traffic policing. Then, check the **Input policing** or **Output policing** (or both) check boxes to enable the specified type of traffic policing. For each type of traffic policing, configure the following fields:
 - **Committed Rate**—The rate limit for this traffic flow; this is a value in the range 8000-2000000000, specifying the maximum speed (bits per second) allowed.
 - **Conform Action**—The action to take when the rate is less than the conform-burst value. Values are transmit or drop.
 - **Exceed Action**—Take this action when the rate is between the conform-rate value and the conform-burst value. Values are transmit or drop.
 - **Burst Rate**—A value in the range 1000-512000000, specifying the maximum number of instantaneous bytes allowed in a sustained burst before throttling to the conforming rate value.



Note The **Enable Policing** check box merely enforces the maximum speed and burst rate, forcing them to the conforming rate value. It does not enforce the **conform-action** or the **exceed-action** specification if these are present.

Modes

The following table shows the modes in which this feature is available:

Firewall Mode		Security Context		
Routed	Transparent	Single	Multiple	
			Context	System
•	—	•	—	—

Priority Queue

If you define a QoS service policy rule that defines traffic as high priority, you must enable a priority queue on one or more interfaces for that service rule to be enabled for traffic flowing through those interfaces. Priority queueing is disabled by default. Go to **Configuration > Device Management > Advanced > Priority Queue** to configure the priority queues.

The Priority Queue pane shows the priority queue table. The Priority Queue table displays the following information for each interface on which the priority queue is configured:

- **Interface**—The interface for which the queue is configured.
- **Queue Limit**—The maximum number of packets that can be enqueued to a normal or priority queue before it drops the connection. Both queues have the same limit. Packets in the priority queue are totally drained before packets in the normal priority queue are transmitted.
- **Transmission Ring Limit**—Specifies the depth of the priority queues. If priority queueing is not enabled, this column shows the message: “Ring Disabled.”

To add or change a priority queue configuration, perform one of the following:

- To add a new priority queue, click **Add**. The Add Priority Queue dialog box appears.
- To edit an existing priority queue, click the queue entry in the table and click **Edit**. Alternately, you can double click the queue entry in the table. The Edit Priority Queue dialog box appears.

Fields

The Add/Edit Priority Queue dialog box contains the following fields:

- **Interface**—Select the interface on which to enable the priority queue. You can only configure one priority queue per interface. This field displays all interfaces for which a priority queue has not been set.
- **Queue Limit**—Specifies the maximum number of packets that can be enqueued to a normal or priority queue before it drops the connection. The minimum is 0 packets, and the maximum is determined dynamically at run time based on available memory. The theoretical maximum number of packets is 2147483647.



Note Both queues have the same limit. Packets in the priority queue are totally drained before packets in the normal priority queue are transmitted.

- **Transmission Ring Limit**—Specifies the maximum number of low-latency or normal priority packets allowed into the Ethernet transmit driver before the driver pushes back to the queues on the interface to let them buffer packets until the congestion clears. The minimum value is 3. The upper limit of the range of values is determined dynamically at run time. The key determinant is the memory needed to support the queues and the memory available on the device. The queues must not

exceed the available memory. The theoretical maximum number of packets is 2147483647 (that is, up to line speed at full duplex). If priority queuing is not enabled, this column shows the message: “Ring Disabled.”

The transmission ring limit is the number of either type of packets allowed into the driver before the driver pushes back to the queues sitting in front of the interface to let them buffer packets until the congestion clears. In general, you can adjust the queue-limit and transmission ring limit parameters to optimize the flow of low-latency traffic.

Because queues are not of infinite size, they can fill and overflow. When a queue is full, any additional packets cannot get into the queue and are dropped. This is *tail drop*. To avoid having the queue fill up, you can adjust the queue-limit parameter to increase the queue buffer size.

Modes

The following table shows the modes in which this feature is available:

Firewall Mode		Security Context		
Routed	Transparent	Single	Multiple	
			Context	System
•	•	•	•	—

