Chapter 3

Configuring Queue Profiles for Buffer Management

This chapter provides information for configuring queue profiles for buffer management on the E-series router.

QoS topics are discussed in the following sections:

Queuing and Buffer Management Overview

A queue is a set of first-in, first-out (FIFO) buffers that buffer packets on the data path. QoS associates queues with a traffic class/interface pair. For example, if you create 4000 IP interfaces and configure each interface with four traffic classes, then 16,000 queues are created. For specific information about the maximum number of QoS queues supported, see *JUNOSe Release Notes, Appendix A, System Maximums*.

The E-series router dynamically manages the shared memory on egress line modules to provide a good balance between sharing the memory among queues and protecting an individual queue's claim on its fair share of the egress memory.

When egress packet memory is in high demand and aggregate utilization of the packet memory is high, queue lengths are set to lengths that strictly partition egress memory into per-queue memory sections. This conservative buffer-management strategy reserves a fair share of buffers for each queue, so that high bandwidth consumers cannot starve out moderate traffic consumers by allocating all the shared memory resource for themselves.

When egress packet memory is in low demand, a more liberal buffer management strategy is used to provide active queues with more access to the shared memory resource.

The router dynamically varies queue lengths for all queues as the real-time demand on the egress packet memory changes. You can configure limits to prevent the router from setting queue lengths too low or too high.

Static Oversubscription

The router uses static oversubscription to vary queue thresholds based on the number of queues currently configured, which is relatively static. Static oversubscription is based on the assumption that, when a few queues are configured, many of the queues are likely to be active at the same time. When a large number of queues are configured, fewer queues are likely to be active at the same time.

When few queues are configured, buffer memory is strictly partitioned between queues to ensure that buffers are available for all queues. As the number of configured queues increases, buffer memory is increasingly oversubscribed to allow more buffer sharing. Reserving buffer space for all queues when many are expected to be idle is unnecessary and wasteful.

Dynamic Oversubscription

The router uses dynamic oversubscription to vary queue thresholds based on the amount of egress buffer memory in use. The router divides egress buffer memory into eight regions of 4 MB each. When buffer memory is in low demand, queues are given large amounts of buffer memory. As the demand for buffer memory increases, queues are given progressively smaller amounts of buffer memory.

Color-Based Thresholding

Packets within the router are tagged with a drop precedence:

- Committed—Green
- Conformed—Yellow
- Exceeded—Red

When the queue fills above the exceeded threshold, the router drops red packets, but still queues yellow and green packets. When the queue fills above the conformed drop threshold, the router queues only green packets.

NOTE: All color-based thresholds vary in proportion to the dynamic queue length.

Related Topics

- Configuring Queue Profiles to Manage Buffers and Thresholds on page 23
- Guidelines for Managing Queue Thresholds on page 19
- Guidelines for Managing Buffers on page 20
- RED and WRED Overview on page 26

Memory Requirements for Queue and Buffers

JUNOSe software uses 128-byte buffers.

The egress memory available for queues available depends on the ASIC and the line module. Table 6 lists the egress memory.

Table 6: Egress Memory on ASIC Line Modules

ASIC	Line Module	Egress Memory (MB)
EFA	All EFA line modules	32
FFA	GE-2 and GE-HDE	64
	OC48	128
	ES2 4G LM	128
TFA	ES2 10G LM	96

Related Topics

- To identify the type of ASIC used by a line module, see the *ERX Module Guide* and the *E120 and E320 Module Guide*
- Guidelines for Managing Queue Thresholds on page 19
- Guidelines for Managing Buffers on page 20

Guidelines for Managing Queue Thresholds

To prevent the router from setting queue thresholds too low or too high, you can specify minimum and maximum queue thresholds. You can also specify the conformed length and exceeded length as percentages of the committed length.

Guidelines for Configuring a Maximum Threshold

We recommend that you constrain queue thresholds using committed or conformed threshold settings; any unused memory is redistributed to queues whose thresholds are not constrained. This use of thresholds is analogous to the way that shaping rates constrain bandwidth and cause bandwidth redistribution to unconstrained queues.

For example, voice queues are scheduled at strict priority; therefore, they require very little buffering. Configuring a maximum queue threshold enables the system to allocate more buffers to other queues in the system. Video queues are similar but because they are higher bandwidth, they might require higher maximum committed thresholds.

You might want to limit latency of your multicast traffic by bounding the queue length using a maximum committed threshold. The following example configures the multicast queues so that the committed threshold never exceeds 20 KB, even when the egress memory is lightly loaded. The forfeited buffers are allocated to other queues.

```
host1(config)#queue-profile multicast
host1(config-queue)#committed-length 0 20000
host1(config-queue)#exit
```

Be sure to include 0 in the syntax, or you will configure a minimum threshold.

Guidelines for Configuring a Minimum Threshold

Configuring a minimum threshold does not guarantee that a queue always obtains the minimum buffer allocation. You can configure 1000 queues with a minimum of 1 MB each, but the buffer memory is 32 MB or 128 MB, not 1 GB. In this case, the system moves into higher operating regions (global utilization) if all these queues buffer traffic, until it reaches 90 percent utilization. At that point, the thresholds must reduce to the reserved percentages, and the queue thresholds drop from a high threshold to a very low one. Queues are not guaranteed to obtain any buffering, and are buffered in the order in which they are received.

You can configure a minimum committed threshold by specifying a value such as 1000 with the **committed-length** command:

```
host1(config)#queue-profile multicast
host1(config-queue)#committed-length 1000 20000
host1(config-queue)#exit
```

Related Topics

- Memory Requirements for Queue and Buffers on page 19
- Configuring Queue Profiles to Manage Buffers and Thresholds on page 23

Guidelines for Managing Buffers

Queue profiles enable you to manage queue thresholds and buffers to manage the following common problems:

- Queues that back up and consume too many buffers
- Queues that cannot obtain buffers when they need them (called *buffer starvation*)

You can set the buffer weight to ensure that some sets of queues get higher thresholds than others. Buffer weight is analogous to weight in a scheduler profile. It directs the router to set the queue thresholds proportionately.

This feature provides graceful buffer allocation as the global utilization goes higher; queues with more buffer weight always obtain more buffers, but they do not undergo a dramatic drop in threshold when the system moves from region to region.

JUNOSe software uses 128-byte buffers. When setting very small queue thresholds, keep the following guidelines in mind:

- Specifying a maximum queue length of 0 bytes disables queuing of packets on the queue.
- Specifying a maximum queue length of 1–128 bytes creates a single 128-byte buffer for the queue.
- Specifying a maximum queue length of 129–256 bytes creates two 128-byte buffers for the queue.
- Packets and cells consume at least one buffer.

  For example, a 64-byte packet consumes a single 128-byte buffer. If you specify a maximum queue length of 256 bytes, then either two packets of 64–128 bytes in length or a single packet of 129–256 bytes can be queued.

For example, suppose a line module with 4000 IP interfaces is configured with four queues per IP interface, corresponding to four traffic classes. Suppose that queues in two of the traffic classes are configured with a buffer weight of 24 to increase burst tolerance. The following example configures the video queue:

```
host1(config)#queue-profile video
host1(config-queue)#buffer-weight 24
host1(config-queue)#exit
host1(config)#
```

When the egress memory is fully loaded, dynamic oversubscription is 0 percent, and the 8000 queues with the default buffer weight strictly partition 25 percent of the 32-MB memory, leaving 75 percent of the memory for the queues weighted 24 (corresponding to the ratio 75 percent:25 percent, or 24:8). Therefore, these queues have committed thresholds of 1 KB each, and queues with the buffer weight of 24 have committed thresholds of 3 KB each. As the egress memory becomes progressively less loaded, all the queue thresholds increase proportionally, based on dynamic oversubscription, but the queues with buffer weight 24 are always set with thresholds three times larger than the default thresholds.

Guidelines for Managing Buffer Starvation

Buffer starvation most commonly occurs when queues or nodes exist in a large round robin, usually in the default traffic-class group. When the round robin congests, the queues back up and require more buffers. The traffic in the round robin starts to burst based on a single node or queue. After a packet is dequeued, the node or queue can wait for thousands of other queues to dequeue a packet before it can dequeue again. During this time, the queue backs up.

If you configure different scheduler profile weights or assured rates for nodes in a large and congested round robin, the buffer starvation becomes apparent. The problem occurs when the heavy weighted nodes wait their turn in the round robin and thousands of other nodes dequeue. While the heavily weighted nodes wait, the system needs to buffer them. However, all queues receive the same buffer allocation by default. If the system goes to higher buffer regions, it starts dropping packets for all queues. When the heavy weight node finally transmits, it dequeues all buffers, but it cannot dequeue the packets that were dropped. You do not achieve the expected bandwidth based on scheduler profile weights.

To manage buffer starvation, configure buffer weights on queues so they are in the same ratio as the expected bandwidth for the queues. For example, if two queues have scheduler weight (or assured-rate) in the ratio of 2:1, then set the buffer weights to the same ratio.

To manage buffer starvation, set the **maximum-committed-threshold** on queues that do not need buffering, and increase the **buffer-weight** for the heavily weighted queues in the round robin.

The system calculates the correct ratio for you. Issue the **show egress queue rates** command to see the ratio:

```
host1#show egress-queue rates brief interface fastEthernet 9/0.2
                       traffic                forwarded aggregate minimum maximum
interface               class                   rate    drop rate  rate    rate
---------------------- ---------------------- --------- --------- ------- -------
ip FastEthernet9/0.2   best-effort                    0         0   25000 1000000
                       videoTrafficClass              0         0  375000 1000000
                       multicastTrafficClass          0         0  925000 1000000
                       internetTrafficClass           0         0   50000 1000000

   Total:                                             0         0

   Queues reported:                   4
   Queues filtered (under threshold):  0
   Queues disabled (no rate period):   0
   Queues disabled (no resources):     0
   Total queues:                       4
```

The minimum rate for each queue is the approximate rate the queue achieves if all configured queues in the line module run infinite traffic. Configure the buffer weights in proportion to the minimum rate displayed by the system.

Related Topics

- Memory Requirements for Queue and Buffers on page 19
- Configuring Queue Profiles to Manage Buffers and Thresholds on page 23
- Monitoring Forwarding and Drop Rates on the Egress Queue on page 330

Configuring Queue Profiles to Manage Buffers and Thresholds

A queue profile controls the buffering and dropping behavior of a set of egress queues by enabling you to set the buffer weight of the queue, the drop thresholds, and the constraints on queue lengths.

Set the queue lengths as follows:

- To oversubscribe buffer memory, set a minimum queue length.

NOTE: If the sum of the queue minimum lengths is greater than the amount of egress buffer memory, then the egress buffer memory is oversubscribed.

- To configure a minimal level of buffering or to limit the buffering in queues, set a maximum queue length. For example, if you want to control latency by configuring very small queues, set the maximum queue length to 256 bytes. The system queues no more than 256 bytes.

If you do not set the queue lengths, the router varies the queue length dynamically in the range 1 KB–7 MB.

1. Create a queue profile and enter Queue Configuration mode.

   ```
   host1(config)#queue-profile video
   host1(config-queue)#
   ```

   You can configure 16 queue profiles on an E-series router.

2. (Optional) Set the buffer weight of the queue.

   ```
   host1(config-queue)#buffer-weight 16
   ```

   Queues with a buffer weight of 16 are twice as long as queues with a buffer weight of 8. The range is 1–63; the default is 8.

3. (Optional) Set a minimum or maximum queue length for committed packets.

   ```
   host1(config-queue)#committed-length 11000 15000
   ```

   The range of minimum and maximum lengths is 0–1 GB. By default, there is no minimum or maximum length. The color for committed packets is green.

4. (Optional) Set a minimum or maximum queue length for conformed packets.

   ```
   host1(config-queue)#conformed-length 10000 14000
   ```

   The range of minimum and maximum lengths is 0–1 GB. By default, there is no minimum or maximum length. The color for conformed packets is yellow.

5. (Optional) Set a minimum or maximum queue length for exceeded packets.

   host1(config-queue)#**exceeded-length 9000 10000**

   The range of minimum and maximum lengths is 0–1 GB. By default, there is no minimum or maximum length. The color for exceeded packets is red.

6. (Optional) Set the conformed drop threshold as a percentage of the committed threshold.

   host1(config-queue)#**conformed-fraction 60**

   The range is 0–100 percent; the default is 50.

7. (Optional) Set the exceeded drop threshold as a percentage of the committed threshold.

   host1(config-queue)#**exceeded-fraction 40**

   The range is 0–100 percent; the default is 25.

Related Topics

- Queuing and Buffer Management Overview on page 17
- Guidelines for Managing Queue Thresholds on page 19
- Guidelines for Managing Buffers on page 20
- Memory Requirements for Queue and Buffers on page 19
- **buffer-weight** command
- **committed-length** command
- **conformed-fraction** command
- **conformed-length** command
- **exceeded-fraction** command
- **exceeded-length** command
- **queue-profile** command

Monitoring Queues and Buffers

To monitor queues and buffers, see:

- Monitoring Queue Thresholds on page 316
- Monitoring Queue Profiles on page 320