NOKIA

7450 Ethernet Service Switch
7750 Service Router
7950 Extensible Routing System
Virtualized Service Router

Releases up to 25.10.R3

Quality of Service Advanced Configuration Guide for Classic CLI

Table of contents

List of tables

List of figures

Preface

About This Guide

Each Advanced Configuration Guide is organized alphabetically and provides feature and configuration explanations, CLI descriptions, and overall solutions. The Advanced Configuration Guide chapters are written for and based on several Releases, up to 25.10.R3. The Applicability section in each chapter specifies on which release the configuration is based.

The Advanced Configuration Guides supplement the user configuration guides listed in the *7450 ESS, 7750 SR, and 7950 XRS Guide to Documentation*.

Audience

This manual is intended for network administrators who are responsible for configuring the routers. It is assumed that the network administrators have a detailed understanding of networking principles and configurations.

Class Fair Hierarchical Policing for SAPs

This chapter provides information to configure Class Fair Hierarchical Policing for SAPs.

Topics in this chapter include:

- Applicability

- Overview

- Configuration

- Conclusion

Applicability

This chapter was written based on SR OS Release 9.0.R1. There are no specific pre-requisites for this configuration.

Overview

The Quality of Service (QoS) features of the SR-series platforms provide traffic control with both shaping and policing.

Shaping is achieved using a queue; packets are placed on the queue and a scheduler removes packets from the queue at a given rate. This provides an upper bound to the traffic rate sent, thereby protecting down stream devices from bursts. However, shaping can introduce latency and jitter as packets are delayed in the queue. Packets can be dropped when the queue is full or statistically when weighted random early discard is applied. Configuration of shaping on the SR OS is described in QoS Architecture and Basic Operation.

Policing is another mechanism for controlling traffic rates but it does not introduce latency/jitter. This is achieved using a token bucket mechanism which drops certain packets from the traffic. A common disadvantage of policing implementations is that they are usually applicable to a single level of traffic priority and have no way to fairly share capacity between multiple streams at the same priority level. Nokia's Class Fair Hierarchical Policing (CFHP) addresses these problems by implementing a four level prioritized policing hierarchy which also provides weighted fairness for traffic at a given priority.

Regardless of whether shaping or policing is being used, the preceding QoS classification and subsequent packet marking functionality is similar for both and is covered in more detail in QoS Architecture and Basic Operation.

This chapter describes the configuration and operation of CFHP when applied to Service Access Points (SAPs). It is also possible to use CFHP for subscribers in a Triple Play Service Delivery Architecture (TPSDA) environment but that is beyond the scope of this chapter.

Policers

CFHP can be used both for ingress and egress QoS. The basic element is a policer which can apply both a committed information rate (CIR) and peak information rate (PIR) to a traffic flow (determined by the ingress classification). Traffic is directed to a policer by assigning a forwarding class (FC) to the policer.

To describe the operation of a policer we will use a token bucket model, this is shown in Figure 1: Policer Token Bucket Model.

Figure 1: Policer Token Bucket Model
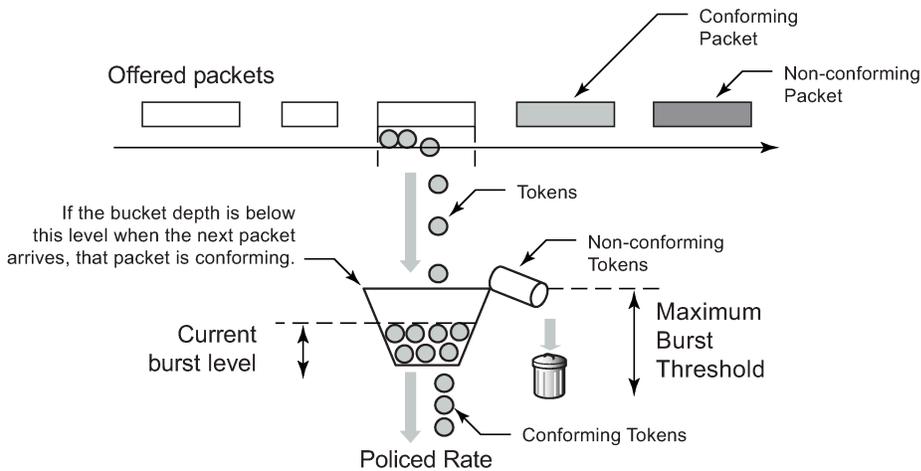
OSSG513

The policer is modeled by a bucket being filling with tokens which represent the bytes in the packets passing through the policer. The bucket drains at a given rate (the policed rate) and if the token (byte) arrival rate exceeds the drain rate then the bucket will fill. The bucket has a maximum depth, defined by a maximum burst threshold. If tokens for a packet arrive in the bucket when the current burst level of tokens is below the maximum burst threshold then the packet is considered to be conforming and all of its tokens are accepted into the bucket. If a packet's tokens arrive when the current burst level has exceeded the maximum burst threshold then none its tokens are accepted into the bucket and the packet is considered to be non-conforming (in the representation, these tokens over-flow into a waste bin).

Table 1: Burst Levels shows an example of the two possibilities.

Table 1: Burst Levels

Maximum burst threshold = 2000 tokens (bytes)				
Policed rate 2 Mbps = 250000 bytes/sec (250 tokens/ms)				
Arrival Time	**Packet Size**	**Current Burst Level**	**Conforming Packet**	**New Burst Level**
T0	1024	1500	Yes	1500 + 1024 = 2524
T0 + 1ms	128	2524 - 250 = 2274	No	2274

When the first packet arrives the current burst level is below the maximum burst threshold so it is conforming, however, when the second packet arrives the current burst level is above the maximum burst threshold so it is non-conforming.

An important aspect of the implementation of hierarchical policing is the ability of a policer bucket to have multiple burst thresholds. The tokens for each arriving packet are only compared against a single threshold relating to the characteristics of packet. These burst thresholds allow specific granular QoS control.

Policer Buckets

A policer uses up to 3 buckets depending on its configuration. A PIR bucket to control the traffic rate which is always used though its rate could be max, there can be an optional CIR bucket if a CIR rate is defined for dynamically profiling (in-profile/out-of-profile) packets, finally there may be a fair information rate (FIR) bucket used to maintain traffic fairness in a hierarchical policing scenario when multiple child policers are configured at the same parent priority level.

The PIR bucket is drained at the PIR rate and has two burst thresholds, one for high burst priority traffic (defined by the maximum burst size (MBS)) and a second for low burst priority traffic (defined by the MBS minus high-prio-only), see Figure 2: Peak Information Rate (PIR) Bucket. The traffic burst priority is determined at ingress by the configured priority of either high or low, and at the egress by the profile state of the packets (in-profile=high, out-of-profile=low). Note that by default all FCs are low burst priority. If a packet conforms at the PIR bucket (its tokens enter the bucket) then the packet is forwarded, otherwise the packet is discarded. Discarding logically results in the packet's tokens not being placed into the CIR, FIR or parent policer buckets.

Figure 2: Peak Information Rate (PIR) Bucket
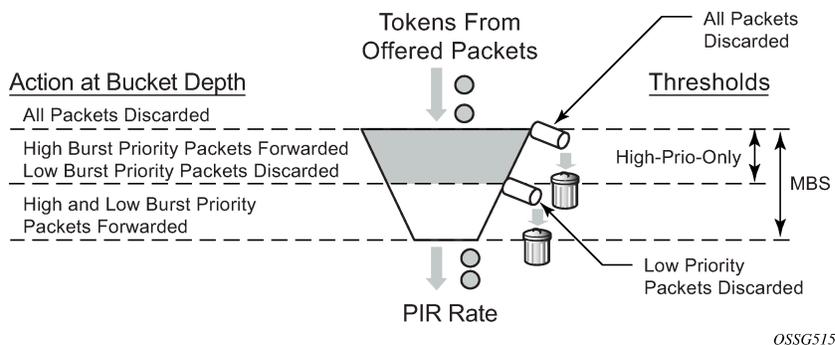
The CIR bucket is drained at the CIR rate and has one configurable burst threshold (defined by the committed burst size (CBS)). At the ingress, if the bucket level is below this threshold traffic is determined to be in-profile so the only action of the CIR bucket is to set the state of dynamically profiled packets to be either in-profile or out-of-profile. At the egress, re-profiling only affects Dot1P and DEI (Layer 2) egress marking (if the frame is double tagged, only the outer VLAN tag is remarked).

The CBS threshold is used when operating in color-blind mode, the profile of incoming packets is undefined and dynamically set based on the current burst level in the CIR bucket compared to the CBS threshold. It is also possible to operate (simultaneously) in color-aware mode, where the classification of incoming packets is used to explicitly determine whether a packet is in-profile or out-of-profile. For color-aware mode, the CIR bucket does not change the packet profile state.

In order to ensure that the overall amount of in-profile traffic takes into account both the explicit and dynamic in-profile packets, tokens from the explicit in-profile packets are allowed to fill the bucket above

the CBS threshold. By doing this, dynamically profiled packets are only marked as in-profile after the token level representing dynamically in-profile and explicit in-profile packets have fallen below the CBS threshold (as the bucket drains). Note that explicitly marked out-of-profile packets remain out-of-profile, so the bottom of the bucket can be considered to be an implicit burst threshold for these packets. This is shown in Figure 3: Committed Information Rate (CIR) Bucket.

Figure 3: Committed Information Rate (CIR) Bucket
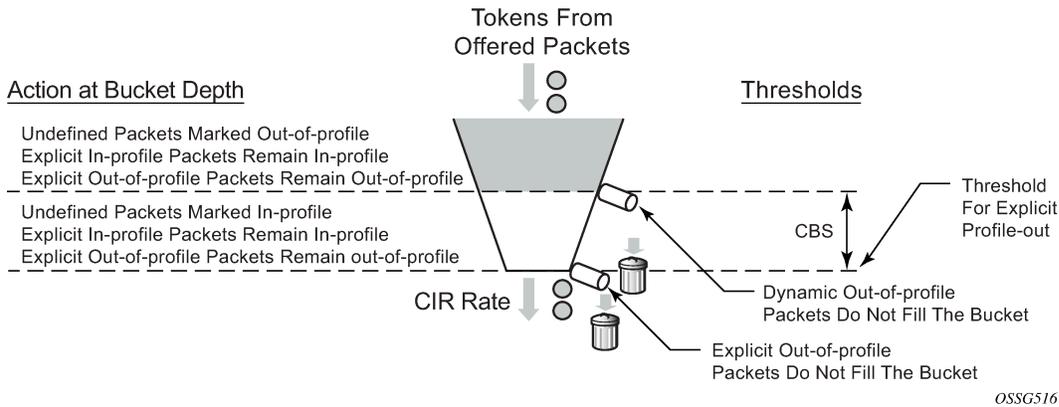
As the depths of the PIR and CIR buckets (MBS and CBS, respectively) are configured independently it is possible to have, for example, the CBS to be larger than the MBS (which is not possible for a queue). This could result in traffic being discarded because it is non-conforming at the PIR bucket but would have been conforming at the CIR bucket. Conversely, if the CBS is smaller than the MBS and the PIR=CIR traffic can be forwarded as out-of-profile, which would not be the case with a queue.

The FIR bucket is controlled by the system and is only used in hierarchical policing scenarios to determine a child's fair access to the available capacity at a parent priority level relative to other children at the same level. This bucket is only used when there is more than one child policer assigned to a given parent policer priority level. The drain rate of the FIR bucket is dynamically set proportionally to the weight configured for the child. This is shown in Figure 4: Fair Information Rate (FIR) Bucket.

Figure 4: Fair Information Rate (FIR) Bucket
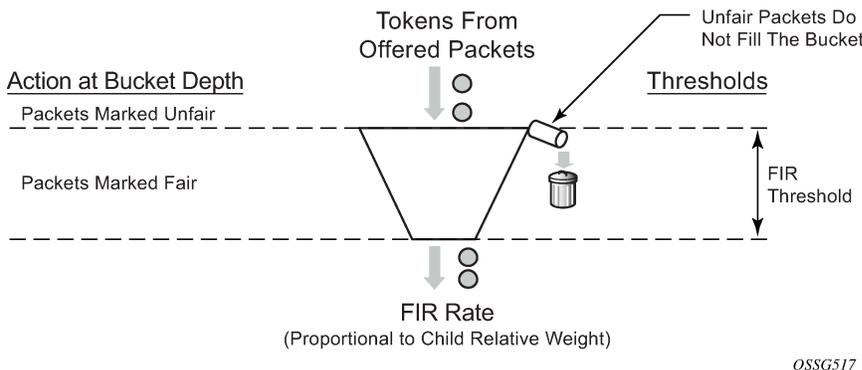
Hierarchical Policing

Policers can be used standalone or with a parent policer to provide hierarchical policing. Up to four stages can be configured in the hierarchy: the child policer, tier 1 and 2 intermediate arbiters, and a root arbiter (which is associated with the parent policer). The arbiters are logical entities that distribute bandwidth at a particular tier to their children in a priority level order, see Figure 5: Policer and Arbiter Hierarchy .

This may result in the drain rates for the child policer buckets being modified, so each child policer PIR and CIR bucket has an administrative rate value (what it is configured to) and an operational rate value (the current operating rate) based on the bandwidth distribution by the parent arbiters.

Each stage in the hierarchy connects to its parent at a priority level and a weight. There are eight available priorities which are serviced in a strict order (8 to 1, highest to lowest, respectively). The weight is used to define relative fairness when multiple children are configured in the same priority level. Note that the child access to parent policer burst capacity is governed by the level at which the child ultimately connects into the root arbiter, not by its connection level at any intermediate arbiters.

Figure 5: Policer and Arbiter Hierarchy
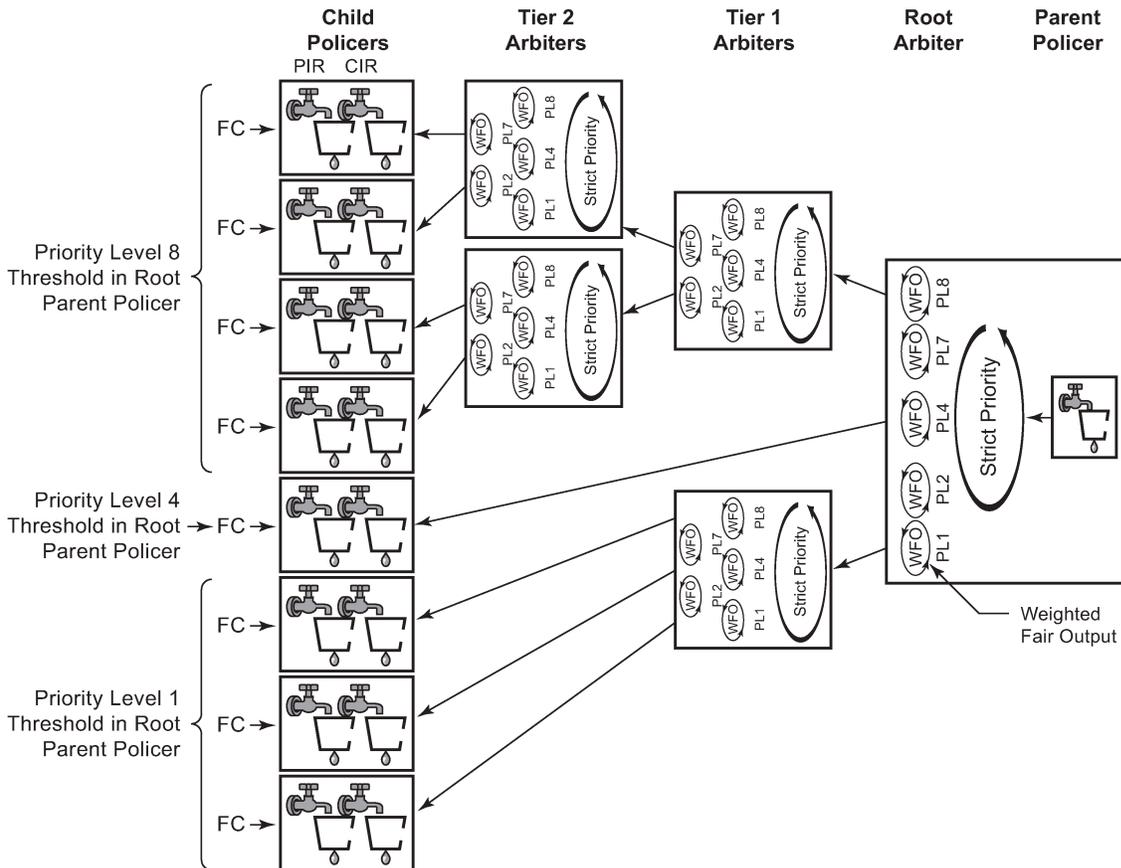
OSSG518

The final configuration aspect to consider is the parent policer, specifically its multiple thresholds and how they relate to the child policers. See Figure 6: Parent Policer and Root Arbiter.

There are 8 priority levels at the parent policer, each having an associated discard-fair and discard-unfair threshold.

The discard-fair threshold is the upper burst limit for all tokens (consequently, all packets) at the given priority, all traffic at a given priority level is discarded when its tokens arrive with this threshold being exceeded. The discard-fair thresholds enable prioritization at the parent policer by having the burst capacity for each priority threshold be larger (or equal) to those of lower priorities. For example, referring to Figure 6: Parent Policer and Root Arbiter, the priority 6 (P6) discard-fair threshold is larger than the priority 5 (P5) discard-fair threshold with the result that even if the priority 5 and below traffic is overloading the parent policer, the priority 6 traffic has burst capacity available in order to allow some of its packets to conform and get forwarded through the parent policer.

Note that if a packet is discarded at the parent policer, the discard needs to be reflected in the associated child policer, this is achieved by also logically removing the related tokens from the child policer buckets.

Figure 6: Parent Policer and Root Arbiter
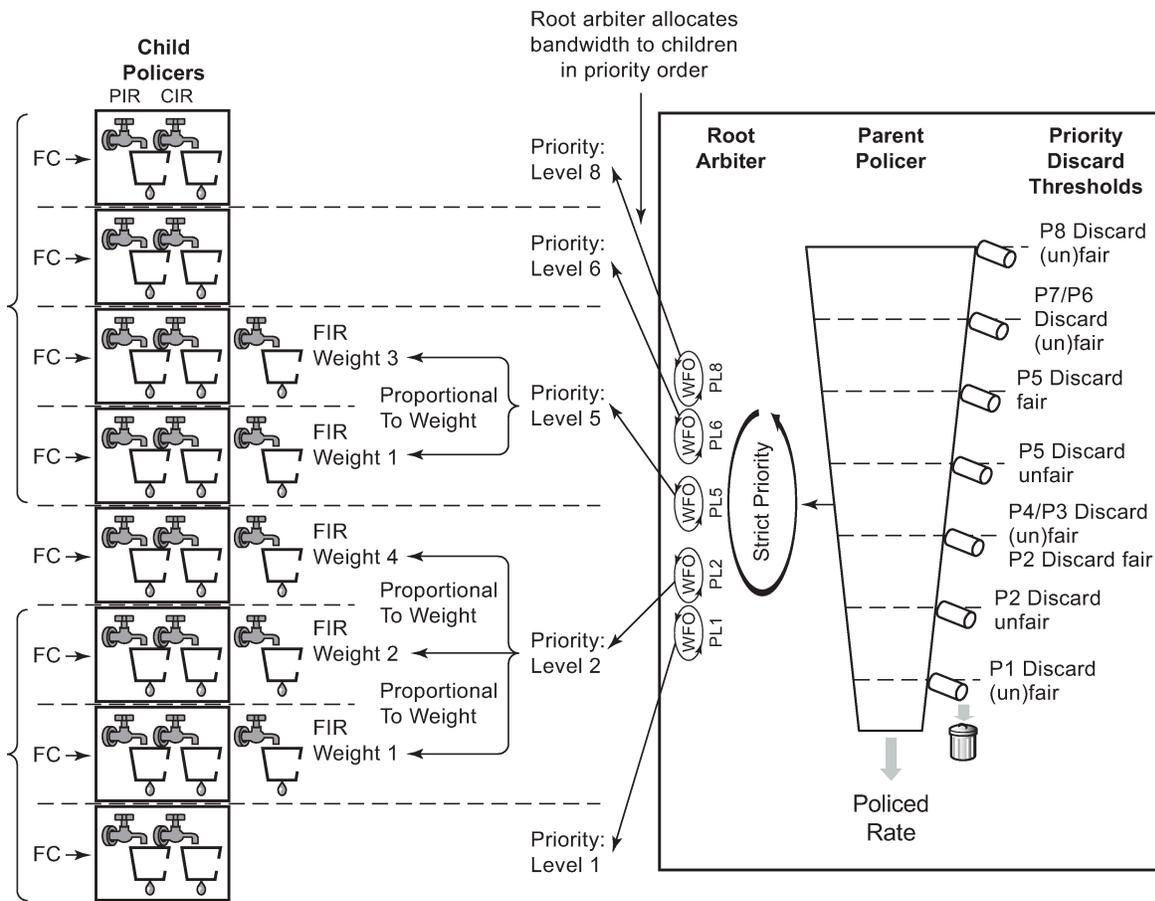
OSSG519

Each priority also has a discard-unfair threshold which discards only unfair traffic of that priority, remembering that fair and unfair are determined by the FIR bucket based on the relative weights of the children.

By default, if there are no children configured at a given priority level then both its discard-fair and discard-unfair thresholds are set to zero bytes above the previous priority's discard-fair threshold.

If there is only a single child at a priority level, the discard-fair will be greater than the previous priority's discard-fair value (by an amount equal to the maximum of the min-thresh-separation and the mbs-contribution, see below) but the discard-unfair will be the same as the previous priority's discard-fair threshold (there is no need for a fairness function when there is only a single child at that priority).

If there is more than one child at a priority level, the discard-unfair threshold will be greater than the previous priority's discard-fair threshold by min-thresh-separation (see below) and the discard-fair threshold will be adjusted upwards by an amount equal to mbs-contribution minus min-thresh-separation.

The result can be summarized as follows:

- With no children at a priority level, the discard fair and unfair thresholds match the values of the previous priority.

- If there are at least two children at a priority level, the discard-unfair burst capacity equals min-thresh-separation.

- The burst capacity for a given priority level with at least one child equals the mbs-contribution, unless this is less than min-thresh-separation in which case the min-thresh-separation is used.

The burst tolerance for each threshold is its own burst capacity plus the sum of the burst capacities of all lower thresholds. Referring to Figure 6: Parent Policer and Root Arbiter, the total burst capacity for priority 6 is the sum of the burst capacities for priorities 1 to 6. Note that the burst for a given FC is normally controlled by the burst allowed at the child PIR threshold, not by the parent policer.

As the burst capacity at the parent policer for a given priority level can change when adding or removing children at lower priority levels, a parameter (fixed) is available per priority threshold which causes the discard-fair and discard-unfair thresholds to be non-zero and so greater than the previous priority's thresholds, calculated as above, even when there are no children at that priority level. An exception to this is when the mbs-contribution is set to zero with the fixed parameter configured, in which case both the discard unfair and fair for that priority level are set to zero bytes above the previous level's thresholds (which results in the corresponding traffic being dropped).

A specific configuration and associated show output is included below to highlight the different threshold options described above.

The QoS example shown in Figure 7: Configuration Example is used to describe the configuration of CFHP.

Figure 7: Configuration Example
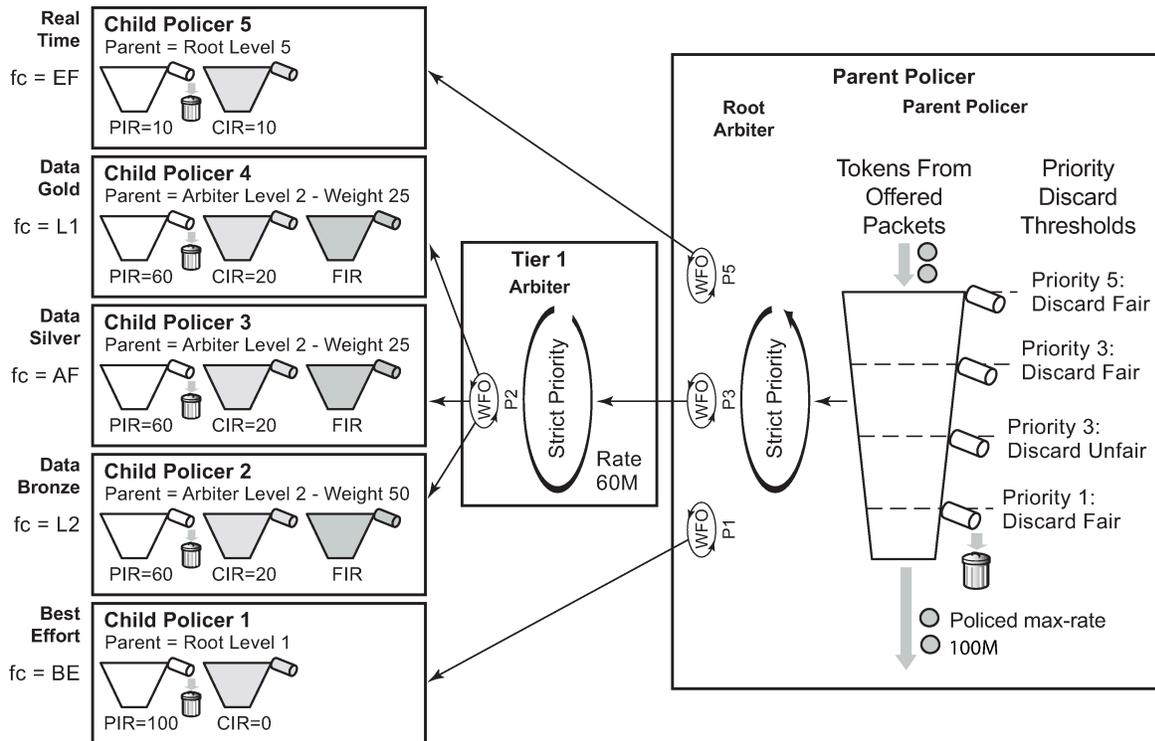
Five classes of services are accepted, each with a specific CIR and PIR. The data classes, bronze, silver and gold (L2/AF/L1), have a relative weighting of 50/25/25 at priority Level 2 of an intermediate arbiter which is constrained to 60Mbps. At the parent policer, the real time traffic (EF) is defined at level 5, with the data classes at Level 3 and a best effort class (BE) at Level 1. The overall traffic is constrained to 100Mbps at the parent policer. Only unicast traffic is policed in this example.

This example focuses on ingress policing, however, the configuration of policers, arbiters and the parent policer at the egress is almost identical to that at the ingress, the only difference being the particular statistics that can be collected.

There is a difference between ingress and egress policing in terms of how the ingress traffic accesses the switch fabric and the egress traffic access the port after it has been policed. In both cases, unicast access is enabled through a set of policer-output-queues, which are shared-queues at the ingress and queue-groups at the egress (at the egress, user defined queue-groups can be used). It is also possible to use a single service queue to access the egress port. Ingress multipoint traffic accesses the switch fabric using the Ingress Multicast Path Management (IMPM) queues.

This is shown in Figure 8: Post Policing Queues on a line card.

Figure 8: Post Policing Queues
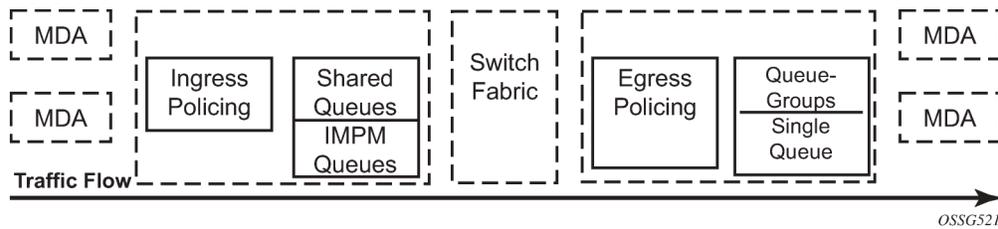
The differences between the ingress and egress policing configuration will be high-lighted in the associated sections.

Configuration

To achieve the QoS shown in Figure 7: Configuration Example, configure a SAP-ingress QoS policy to define the child policers and a policer-control-policy to define the intermediate arbiter and the root arbiter/parent policer. As this example is for ingress, the unicast traffic will pass through a set of shared queues called policer-output-queues, which could be modified if required.

Policers

Policers control the CIR and PIR rates for each of the traffic classes and are defined in a SAP-ingress QoS policy. The focus here are parameters related to policing.

The configuration of a child (or standalone) policer is similar to that of a queue.

```
config>qos>sap-ingress# policer policer-id [create]
    description "description-string"
    adaptation-rule [pir {max | min | closest}] [cir {max | min | closest}]
    stat-mode {no-stats|minimal|offered-profile-no-cir|
            offered-priority-no-cir|offered-profile-cir|offered-priority-cir|
            offered-total-cir|offered-limited-profile-cir}
    rate {max | kilobits-per-second} [cir {max | kilobits-per-second}]
    percent-rate pir-percent [cir cir-percent]
    mbs size [bytes | kilobytes]
    cbs size [bytes | kilobytes]
    high-prio-only [default | percent-of-mbs]
    parent {root | arbiter-name} [level level] [weight weight-within-level]
    packet-byte-offset {add bytes | subtract bytes}
```

Parameters:

- description — This configures a text string, up to 80 characters, which can be used to describe the use of the policy.

- adaptation-rule — The hardware supports distinct values for the rates. This parameter tells the system how the rate configured should be mapped onto the possible hardware values. min results in the next higher hardware value being used, max results in the next lower hardware value being used and closest results in the closest available hardware value being used. As can be seen, it is possible to set the adaptation-rule independently for the CIR and PIR. Default: closest

- stat-mode — This defines the traffic statistics collected by the policer, summarized in Table 2: Policer stat-mode.

Table 2: Policer stat-mode

stat-mode	Ingress		Egress	
no-stats	0	Neither policer nor parent arbiter account are required.	0	Neither policer nor parent arbiter accounting are required.
minimal (default)	1	Basic policer accounting (default).	1	Basic policer accounting (default).
offered-profile-no-cir	2	All ingress packets are either in-profile or out-of-profile.	2	Accounting for egress offered profile is required. No visibility for CIR profile state output.
offered-priority-no-cir	2	Ingress packet burst priority accounting is the primary requirement.		N/A
offered-limited-profile-cir	3	Ingress color-aware profiling is in use but packets are not being classified as in-profile.		N/A
offered-profile-cir	4	Ingress color-aware profiling is in use and packets are undefined or classified as out-of-profile or in-profile.	4	Egress profile reclassification is performed.
offered-priority-cir	4	Ingress policer is used in color-blind mode and ingress packet priority and CIR state output accounting is needed.		N/A
offered-total-cir	2	Ingress priority and ingress profile accounting is not needed. CIR profiling is in use.	2	Offered profile visibility is not required (such as, all offered packets have the same profile) and CIR profiling is in use.
	Counter resources needed for this stat mode			

- rate and cir — The rate defines the PIR and the cir defines the CIR, both are in Kbps. The parameters rate and percent-rate are mutually exclusive and will overwrite each other when configured in the same policy. Range: PIR=1 to 20,000,000 Kbps or max ; CIR=0 to 20,000,000 Kbps or max Default: rate(PIR)=max ; cir=0

- percent-rate and cir — The percent-rate defines the PIR and the cir defines the CIR with their values being a percentage of the maximum policer rate of 20Gbps. The parameters rate and percent-rate are mutually exclusive and will overwrite each other when configured in the same policy. Range: pir-percent = [0.01..100.00]; cir-percent = [0.00..100.00] Default: pir-percent = 100; cir-percent = 0.00

- mbs and cbs — The mbs defines the MBS for the PIR bucket and the cbs defines the CBS for the CIR bucket, both can be configured in bytes or kilobytes. Note that the PIR MBS applies to high burst priority packets (these are packets whose classification match criteria is configured with priority high at the ingress and are in-profile packets at the egress). Range: mbs=0 to 4194304 bytes; cbs=0 to 4194304 bytes Note: mbs=0 prevents any traffic from being forwarded. Default: mbs=10ms of traffic or 64KB if PIR=max; cbs=10ms of traffic or 64KB if CIR=max

- high-prio-only — This defines a second burst threshold within the PIR bucket to give a maximum burst size for low burst priority packets (these are packets whose classification match criteria is configured with priority low at the ingress and are out-of-profile packets at the egress). It is configured as a percentage of the MBS. Default: 10%

- parent — This parameter is used when hierarchical policing is being performed and points to the parent arbiter (which could be the root arbiter or an intermediate arbiter), giving the level to which this policer connects to its parent arbiter and its relative weight compared to other children at the same level. Note that for a child policer to be associated with a parent, its stat-mode cannot be no-stats. Range: level=1 to 8; weight=1 to 100 Default: level=1; weight=1

- packet-byte-offset — This changes the packet size used for accounting purposes, both in terms of the CIR and PIR rates and what is reported in the statistics. The change can either add or subtract a number of bytes. For example:

  - To have the policer work on Layer 2 frame size including inter-frame gap and preamble, add 20 bytes.

  - To have the policer work on IP packet size instead of the default layer 2 frame size, subtract the encapsulation overhead:

    14 bytes L2 + 4bytes VLAN ID + 4 bytes FCS = 22 bytes

    Range: add-bytes=0 to 31; sub-bytes=1 to 32 Default: add-bytes=0; sub-bytes=0

A FC must be assigned to the policer in order for the policer to be instantiated (allocating a hardware policer).

By default, any unicast traffic assigned to the FC at the ingress will be processed by the policer, non-unicast traffic would continue to use the multipoint queue. At the egress all traffic assigned to the FC is processed by the policer (as there is no distinction between unicast and non-unicast traffic at the egress).

If required, non-unicast traffic can be policed in IES/VPRN and VPLS services at the ingress (note: all Epipe traffic is treated as unicast). Within an IES/VPRN service, multicast traffic can be assigned to a specific ingress policer on a PIM enabled IP interface. When the service is VPLS, broadcast, unknown unicast and multicast traffic can be individually assigned to ingress policers. In each of these cases, the policers used could be separate from the unicast policer, resulting in the instantiation of additional hardware policers, or a single policer could be used for multiple traffic types (this differs from the queuing implementation where separate queue types are used for unicast and non-unicast traffic).

```
config>qos>sap-ingress>fc#
    broadcast-policer <policer-id>
    unknown-policer <policer-id>
    multicast-policer <policer-id>
```

As mentioned above, the ingress policed unicast traffic passes through a set of shared-queues (policer-output-queues) to access the switch fabric with the multipoint traffic using the IMPM queues.

When policers are required at the egress, a SAP-egress policy is used. The configuration of the policers is almost identical to that used in the SAP-ingress policy, the only difference being the available stat-modes (as shown above).

At the egress, the policed traffic can also be directed to a specific queue-group (instead of the default policer-output-queues) and to a specific queue within that queue-group, as follows:

```
config>qos>sap-egress>fc# policer <policer-id> [group <queue-group-name> [queue <queueid>]]
```

It is also possible to direct the egress policed traffic to a single service queue if specific egress queuing is required, as follows:

```
config>qos>sap-egress>fc# policer <policer-id> queue <queue-id>
```

Multiple egress policers in a SAP-egress policy can use the same local queue and other forwarding classes can directly use the same local queue that is being used by policers.

Parent Policer and Arbiters

The parent policer and its associated root arbiter, together with the tier 1 and 2 arbiters, are configured within a policer-control-policy.

```
config>qos# policer-control-policy policy-name [create]
    description description-string
    root
        max-rate {kilobits-per-second | max}
        priority-mbs-thresholds
            min-thresh-separation size [bytes|kilobytes]
            priority level
                mbs-contribution size [bytes|kilobytes] [fixed]
    tier 1
        arbiter arbiter-name [create]
            description escription-string
            rate {kilobits-per-second|max}
            parent root [level priority-level] [weight weight-within-level]
    tier 2
        arbiter arbiter-name [create]
            description description-string
            rate {kilobits-per-second | max}
            parent {root|arbiter-name} [level priority-level] [weight weight-within-level]
```

Parameters:

- description — This configures a text string, up to 80 characters, which can be used to describe the use of the policy.

- root — This section defines the configuration of the parent policer and the root arbiter.

  - max-rate — This defines the policed rate of the parent policer, the rate at which the bucket is drained. It is defined in Kbps with an option to use max, in which case the maximum possible rate is used. Range: 1 to 20,000,000Kbps or max Default: max

  - priority-mbs-thresholds This section defines the thresholds used for the 8 priorities available in the parent policer.

    - min-thresh-separation — This defines the minimum separation between any two active thresholds in the parent policer in units of bytes or kilobytes.

      It should be set to a value greater than the maximum packet size used for traffic passing through the policer. This ensures that a single packet arriving in the parent policer will not cause the depth of tokens to cross two burst thresholds, if this did happen it would result in the prioritization failing as a given priority level could be starved of burst capacity by a lower priority traffic.

      This parameter is also used as the burst capacity for each priority level's unfair packets. Range: 0 to 4194304 bytes Default: 1536 bytes

- mbs-contribution — This is normally used to define the amount of packet burst capacity required at the parent policer for a particular priority level with at least one child, keeping in mind that the total capacity is the sum of this plus that of all lower thresholds. The actual burst capacity used depends also on the setting of min-thresh-separation, as described earlier. This permits the tuning of the burst capacity at the parent for any children at a given priority level. A conservative setting would ensure that the burst at the parent policer for a given priority is the sum of the bursts of all children at that priority. Less conservative settings could use a lower value and assume some level of oversubscription.

  The use of the fixed parameter causes both the fair and unfair discard thresholds to be non-zero even when there are no children assigned to this priority level (unless the mbs-contribution is set to zero). Range: 0 to 4194304 Default: 8192 bytes

  The relationship between these two parameters is shown in Figure 9: Parent Policer Thresholds.

Figure 9: Parent Policer Thresholds
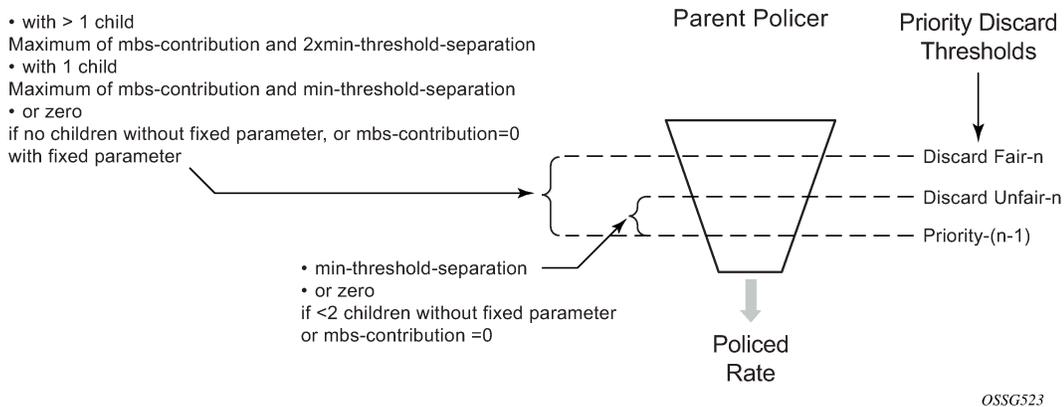
- tier

  This section defines the configuration of any intermediate tier 1 or 2 arbiters.

  – arbiter

    This specifies the name of the arbiter.

    - description — This configures a text string, up to 80 characters, which can be used to describe the use of the policy.

    - rate — This defines the rate of the arbiter, it is the maximum rate at which the arbiter will distribute burst capacity to its children. It is defined in Kbps with an option to use max, in which case the maximum possible rate is used. Range: 1 to 20,000,000Kbps or max Default: max

    - parent — This parameter is used when hierarchical policing is being performed and points to the parent arbiter (which could be the root arbiter or a tier 1 arbiter), giving the level to which this arbiter connects to its parent arbiter and its relative weight compared to other children at the same level. Range: level=1 to 8; weight=1 to 100 Default: level=1; weight=1

Access to Switch Fabric and Egress Port

After the traffic has been processed by the policers it must pass through a set of queues in order to access the switch fabric at the ingress or the port at the egress.

For the ingress unicast traffic, there is a set of shared-queues (one queue per FC for each possible switch fabric destination) called policer output queues. Note that only their queue characteristics can be configured, the FC to queue mapping is fixed. Also, the PIR/CIR rates only affect the packet scheduling, they do not alter the packet profile state. The details of shared-queues are beyond the scope of this chapter.

```
config>qos# shared-queue "policer-output-queues"
    description description-string
    fc <fc-name> [create]
        broadcast-queue <queue-id>
        multicast-queue <queue-id>
        queue <queue-id>
        unknown-queue <queue-id>
    queue queue-id [queue-type] [multipoint] [create]
        cbs percent
        mbs percent
        high-prio-only percent
        pool pool-name
        rate percent [cir percent]
```

Multipoint traffic uses the IMPM queues to access the switch fabric. For the egress to the port, either a queue-group or a single service queue is used. There is a default queue-group called policer-output-queues or a user configured queue-group can also be used.

As mentioned above, when a policer is assigned to a specific queue-group (default or user defined) it is optionally possible to configure explicitly the queue to be used. Within the queue-group it is also possible to redirect a FC for policed traffic to a specific queue, using the FC parameter. The preference of the FC to queue mapping is (in order, highest to lowest):

1. Explicitly configured in SAP-egress FC definition

2. Mapped using FC parameter within queue-group definition

3. Default is to use queue 1

```
config>qos>qgrps>egr# queue-group queue-group-name [create]
    description description-string
    queue queue-id [queue-type] [create]
        adaptation-rule [pir adaptation-rule] [cir adaptation-rule]
        burst-limit size [bytes|kilobytes]
        cbs size-in-kbytes
        high-prio-only percent
        mbs size [bytes|kilobytes]
        parent scheduler-name [weight weight] [level level] [cir-weight cir-weight]
                              [cir-level cir-level]
        percent-rate pir-percent [cir cir-percent]
        pool pool-name
        port-parent [weight weight] [level level] [cir-weight cir-weight]
                    [cir-level cir-level]
        rate pir-rate [cir cir-rate]
        xp-specific
            wred-queue [policy slope-policy-name]
    fc fc-name [create]
        queue queue-id
```

The default policer-output-queues queue-group consists of two queues; queue 1 being best-effort and queue 2 being expedite. The lowest four FCs (BE, L2, AF, L1) are assigned to queue 1 and the highest four queues (H2, EF, H1, NC) are assigned to queue 2. It may be important to change the queue 2 definition in the queue-group to have CIR=PIR when there are other best-effort queues using a non-zero CIR on the

same egress port. This ensures that the policed traffic using queue 2 will be scheduled before any other best-effort within CIR traffic. It also results in the queue CBS being non-zero, allowing the queue 2 traffic access to reserved buffer space.

```
A:PE-1>config>qos# queue-group-templates egress queue-group "policer-output-queues"
A:PE-1>cfg>qos>qgrps>egr>qgrp# info detail
----------------------------------------------
                    description "Default egress policer output queues."
                    queue 1 best-effort create
                        no parent
                        no port-parent
                        adaptation-rule pir closest cir closest
                        rate max cir 0
                        cbs default
                        mbs default
                        high-prio-only default
                        no pool
                        xp-specific
                            no wred-queue
                        exit
                        no burst-limit
                    exit
                    queue 2 expedite create
                        no parent
                        no port-parent
                        adaptation-rule pir closest cir closest
                        rate max cir 0
                        cbs default
                        mbs default
                        high-prio-only default
                        no pool
                        xp-specific
                            no wred-queue
                        exit
                        no burst-limit
                    exit
                    fc af create
                        queue 1
                    exit
                    fc be create
                        queue 1
                    exit
                    fc ef create
                        queue 2
                    exit
                    fc h1 create
                        queue 2
                    exit
                    fc h2 create
                        queue 2
                    exit
                    fc l1 create
                        queue 1
                    exit
                    fc l2 create
                        queue 1
                    exit
                    fc nc create
                        queue 2
                    exit
```

The remaining details of queue-groups are beyond the scope of this section.

Applying the SAP Ingress and Policer Control Policy

The SAP ingress policy and policer control policy are both applied under the associated SAP. After applying these, it is possible to override the configuration of specific policers and/or the policer control policy. This is shown below. The parameter values are the same as detailed for the policies, as above.

```
config>service><service>#
    sap sap-id [create]
        [ingress|egress]
            qos policy-id
            policer-control-policy policy-name
            policer-override
                policer policer-id [create]
cbs size [bytes|kilobytes]
     mbs size [bytes|kilobytes]
packet-byte-offset {add add-bytes | subtract sub-bytes}
rate {rate | max} [cir {max | rate}
percent-rate <pir-percent> [cir <cir-percent>]
stat-mode stat-mode
            policer-control-override [create]
                max-rate {rate | max}
                priority-mbs-thresholds
min-thresh-separation size [bytes | kilobytes]
priority level
                        mbs-contribution size [bytes | kilobytes]
```

The SAP ingress policy and policer control policy required for the configuration example in Figure 7: Configuration Example is shown below.

```
#------------------------------------------------
echo "QoS Policy Configuration"
#------------------------------------------------
    qos
        policer-control-policy "cfhp-1" create
            root
                max-rate 100000
            exit
            tier 1
                arbiter "a3" create
                    parent "root" level 3
                    rate 60000
                exit
            exit
        exit
        sap-ingress 10 create
            queue 1 create
            exit
            queue 11 multipoint create
            exit
            policer 1 create
                stat-mode offered-total-cir
                parent "root"
                rate 100000
                high-prio-only 0
            exit
            policer 2 create
                stat-mode offered-total-cir
                parent "a3" level 2 weight 50
                rate 60000 cir 20000
                high-prio-only 0
```

```
              exit
              policer 3 create
                  stat-mode offered-total-cir
                  parent "a3" level 2 weight 25
                  rate 60000 cir 20000
                  high-prio-only 0
              exit
              policer 4 create
                  stat-mode offered-total-cir
                  parent "a3" level 2 weight 25
                  rate 60000 cir 20000
                  high-prio-only 0
              exit
              policer 5 create
                  stat-mode offered-total-cir
                  parent "root" level 5
                  rate 10000 cir 10000
                  high-prio-only 0
              exit
              fc "af" create
                  policer 3
              exit
              fc "be" create
                  policer 1
              exit
              fc "ef" create
                  policer 5
              exit
              fc "l1" create
                  policer 4
              exit
              fc "l2" create
                  policer 2
              exit
              dot1p 1 fc "be"
              dot1p 2 fc "l2"
              dot1p 3 fc "af"
              dot1p 4 fc "l1"
              dot1p 5 fc "ef"
          exit
```

Traffic is classified based on dot1p values, each of which is assigned to an individual FC which in turn is assigned to a policer. The policer rates are configured as required for the example with an appropriate stat-mode. Default values are used for the policer burst thresholds. As all FCs are low burst priority by default, the high-prio-only has been set to zero in order to allow the traffic to use all of the MBS available at the PIR bucket.

Policers 2, 3 and 4 are parented to the arbiter "a3" with the required weights and at a single level (Level 2). In this example it does not matter which level of "a3" is used to parent these policers, the important aspect is the level at which "a3" is parented to the root. Consequently, these policers use the Level 3 parent policer thresholds (not the level they are parented on a"a3" not Level 2). Arbiter "a3" has a rate of 60Mbps so that its children cannot exceed this rate (except up to the burst tolerances).

Policers 1 and 5 are directly parented to the root arbiter, together with tier 1 arbiter "a3".

The total capacity for the 5 traffic streams is constrained to 100Mbps by the parent policer, again with the default burst tolerances at the root arbiter.

The SAP-ingress and policer-control-policies are applied to a SAP within an Epipe.

```
#--------------------------------------------------
echo "Service Configuration"
```

```
#-------------------------------------------------
    service
        epipe 1 customer 1 create
            sap 1/1/3:1 create
                ingress
                    policer-control-policy "cfhp-1"
                    qos 10
                exit
            exit
            sap 1/1/4:1 create
            exit
            no shutdown
        exit
    exit
```

The following configuration is used to highlight the relative thresholds in the parent policer when a priority
level has 0, 1 or 2 associated children, both with and without using the fixed parameter.

```
-------------------------------------------------
echo "QoS Policy Configuration"
#-------------------------------------------------
    qos
        policer-control-policy "cfhp-2" create
            root
                max-rate 100000
                priority-mbs-thresholds
                    min-thresh-separation 256 bytes
                    priority 1
                        mbs-contribution 1 kilobytes
                    exit
                    priority 2
                        mbs-contribution 1 kilobytes
                    exit
                    priority 3
                        mbs-contribution 1 kilobytes
                    exit
                    priority 4
                        mbs-contribution 1 kilobytes fixed
                    exit
                    priority 5
                        mbs-contribution 1 kilobytes fixed
                    exit
                    priority 6
                        mbs-contribution 1 kilobytes fixed
                    exit
                exit
            exit
        exit
        sap-ingress 20 create
            queue 1 create
            exit
            queue 11 multipoint create
            exit
            policer 1 create
                parent "root" level 2
            exit
            policer 2 create
                parent "root" level 3
            exit
            policer 3 create
                parent "root" level 3
            exit
            policer 4 create
```

```
                parent "root" level 5
            exit
            policer 5 create
                parent "root" level 6
            exit
            policer 6 create
                parent "root" level 6
            exit
            fc "af" create
                policer 3
            exit
            fc "be" create
                policer 1
            exit
            fc "ef" create
                policer 6
            exit
            fc "h2" create
                policer 5
            exit
            fc "l1" create
                policer 4
            exit
            fc "l2" create
                policer 2
            exit
        exit
#-------------------------------------------------
echo "Service Configuration"
#-------------------------------------------------
    service
        epipe 2 customer 1 create
            sap 1/1/3:2 create
                ingress
                    policer-control-policy "cfhp-2"
                    qos 20
                exit
            exit
            sap 1/1/4:2 create
            exit
            no shutdown
        exit
```

A policer-control-policy can also be applied under a multi-service site (MSS) so that the hierarchical policing applies to traffic on multiple SAPs, potentially from different services. The MSS can only be assigned to a port, which could be a LAG, but it is not possible to assign an MSS to a card. When MSS are used, policer overrides are not supported.

```
config>service><service>#
    service
        customer customer-id [create]
            multi-service-site customer-site-name [create]
                assignment port port-id
                egress
                    policer-control-policy name
                ingress
                    policer-control-policy name
        service-type
            sap sap-id
                multi-service-site customer-site-name
                ingress
                    qos policy-id
                egress
```

```
      qos policy-id
```

Show Output

After configuring the example as described in the previous section, steady state traffic was sent through the Epipe to overload each of the policers and the show output below was collected. This output focuses on the policer and arbiter details.

The following shows the policers on the SAP and their current state.

```
A:PE-1# show qos policer sap 1/1/3:1
===============================================================================
Policer Information (Summary), Slot 1
===============================================================================
-------------------------------------------------------------------------------
Name            FC-Maps       MBS       HP-Only A.PIR     A.CIR
Direction                     CBS       Depth   O.PIR     O.CIR     O.FIR
-------------------------------------------------------------------------------
1->1/1/3:1->1
Ingress         be            124 KB    0 KB    100000    0
                              0 KB      82      30000     0         30000
1->1/1/3:1->2
Ingress         l2            76 KB     0 KB    60000     20000
                              25 KB     77846   30000     20000     30000
1->1/1/3:1->3
Ingress         af            76 KB     0 KB    60000     20000
                              25 KB     77824   15000     15000     15000
1->1/1/3:1->4
Ingress         l1            76 KB     0 KB    60000     20000
                              25 KB     77868   15000     15000     15000
1->1/1/3:1->5
Ingress         ef            12800 B   0 KB    10000     10000
                              12800 B   12834   10000     10000     10000
===============================================================================
A:PE-1#
```

The output above shows the configured values for the policers, e.g. PIR and CIR, together with their operational (current) state, such as PIR, CIR and FIR. The depth of each of the PIR buckets is also shown.

The detailed state of each policer can be seen by adding the parameter detail. The following is the output for policer 3.

```
A:PE-1# show qos policer sap 1/1/3:1 ingress detail
...
===============================================================================
Policer Info (1->1/1/3:1->3), Slot 1
===============================================================================
Policer Name       : 1->1/1/3:1->3
Direction          : Ingress         Fwding Plane      : 1
FC-Map             : af
Depth PIR          : 77842 Bytes     Depth CIR         : 25618 Bytes
Depth FIR          : 77842 Bytes
MBS                : 76 KB           CBS               : 25 KB
Hi Prio Only       : 0 KB            Pkt Byte Offset   : 0
Admin PIR          : 60000 Kbps      Admin CIR         : 20000 Kbps
Oper PIR           : 15000 Kbps      Oper CIR          : 15000 Kbps
Oper FIR           : 15000 Kbps
Stat Mode          : offered-total-cir
PIR Adaption       : closest         CIR Adaption      : closest
Parent Arbiter Name: a3
```

```
-------------------------------------------------------------------------------
Arbiter Member Information
-------------------------------------------------------------------------------
Offered Rate      : 45800 Kbps
Level             : 2                 Weight            : 25
Parent PIR        : 15000 Kbps        Parent FIR        : 15000 Kbps
Consumed          : 15000 Kbps
-------------------------------------------------------------------------------
===============================================================================...
A:PE-1#
```

Notice that the above output shows the depth of the PIR, CIR and FIR buckets together with their operational rates. This can be used to explain the operation of the policers in this example and is discussed later in this section.

The stat-mode of offered-total-cir configured on policer 3 results in these statistics being collected.

```
A:PE-1# show service id 1 sap 1/1/3:1 stats
===============================================================================
...
-------------------------------------------------------------------------------
Sap per Policer stats
-------------------------------------------------------------------------------
                       Packets                 Octets

Ingress Policer 1 (Stats mode: offered-total-cir)
Off. All           : 2690893                 172217152
Dro. InProf        : 0                       0
Dro. OutProf       : 967465                  61917760
For. InProf        : 0                       0
For. OutProf       : 1723428                 110299392

Ingress Policer 2 (Stats mode: offered-total-cir)
Off. All           : 2690988                 172223232
Dro. InProf        : 0                       0
Dro. OutProf       : 909492                  58207488
For. InProf        : 1178507                 75424448
For. OutProf       : 602989                  38591296
...
```

The following output is included for reference and shows the statistics which are collected for each of the ingress and egress stat-modes.

```
PE-1# show service id 2 sap 1/1/1:2 stats
...
-------------------------------------------------------------------------------
Sap per Policer stats
-------------------------------------------------------------------------------
                       Packets                 Octets

Ingress Policer 1 (Stats mode: no-stats)

Ingress Policer 2 (Stats mode: minimal)
Off. All           : 0                       0
For. All           : 0                       0
Dro. All           : 0                       0

Ingress Policer 3 (Stats mode: offered-profile-no-cir)
Off. InProf        : 0                       0
Off. OutProf       : 0                       0
For. InProf        : 0                       0
```

```
For. OutProf       : 0                    0
Dro. InProf        : 0                    0
Dro. OutProf       : 0                    0

Ingress Policer 4 (Stats mode: offered-priority-no-cir)
Off. HiPrio        : 0                    0
Off. LowPrio       : 0                    0
For. HiPrio        : 0                    0
For. LoPrio        : 0                    0
Dro. HiPrio        : 0                    0
Dro. LowPrio       : 0                    0

Ingress Policer 5 (Stats mode: offered-profile-cir)
Off. InProf        : 0                    0
Off. OutProf       : 0                    0
Off. Uncolor       : 0                    0
For. InProf        : 0                    0
For. OutProf       : 0                    0
Dro. InProf        : 0                    0
Dro. OutProf       : 0                    0

Ingress Policer 6 (Stats mode: offered-priority-cir)
Off. HiPrio        : 0                    0
Off. LowPrio       : 0                    0
For. InProf        : 0                    0
For. OutProf       : 0                    0
Dro. InProf        : 0                    0
Dro. OutProf       : 0                    0

Ingress Policer 7 (Stats mode: offered-total-cir)
Off. All           : 0                    0
For. InProf        : 0                    0
For. OutProf       : 0                    0
Dro. InProf        : 0                    0
Dro. OutProf       : 0                    0

Ingress Policer 8 (Stats mode: offered-limited-profile-cir)
Off. OutProf       : 0                    0
Off. Uncolor       : 0                    0
For. InProf        : 0                    0
For. OutProf       : 0                    0
Dro. InProf        : 0                    0
Dro. OutProf       : 0                    0

Egress Policer 1 (Stats mode: no-stats)

Egress Policer 2 (Stats mode: minimal)
Off. All           : 0                    0
For. All           : 0                    0
Dro. All           : 0                    0

Egress Policer 3 (Stats mode: offered-profile-no-cir)
Off. InProf        : 0                    0
Off. OutProf       : 0                    0
For. InProf        : 0                    0
For. OutProf       : 0                    0
Dro. InProf        : 0                    0
Dro. OutProf       : 0                    0

Egress Policer 4 (Stats mode: offered-profile-cir)
Off. InProf        : 0                    0
Off. OutProf       : 0                    0
Off. Uncolor       : 0                    0
For. InProf        : 0                    0
```

```
For. OutProf         : 0                    0
Dro. InProf          : 0                    0
Dro. OutProf         : 0                    0

Egress Policer 5 (Stats mode: offered-total-cir)
Off. All             : 0                    0
For. InProf          : 0                    0
For. OutProf         : 0                    0
Dro. InProf          : 0                    0
Dro. OutProf         : 0                    0
===============================================================================
```

It is possible to show the policer-control-policy details and the SAPs with which it is associated, as shown here.

```
A:PE-1# show qos policer-control-policy cfhp-1
===============================================================================
QoS Policer Control Policy
===============================================================================
Policy-Name       : cfhp-1
Description       : (Not Specified)
Min Threshold Sep : Def

-------------------------------------------------------------------------------
Priority MBS Thresholds
-------------------------------------------------------------------------------
Priority          MBS Contribution
-------------------------------------------------------------------------------
1                 none
2                 none
3                 none
4                 none
5                 none
6                 none
7                 none
8                 none


-------------------------------------------------------------------------------
Tier/Arbiter                  Lvl/Wt    Rate      Parent
-------------------------------------------------------------------------------
  root                        N/A       100000    None
1 a3                          3/1       60000     root

===============================================================================
A:PE-1# show qos policer-control-policy "cfhp-1" association

===============================================================================
QoS Policer Control Policy
===============================================================================
Policy-Name       : cfhp-1
Description       : (Not Specified)

-------------------------------------------------------------------------------
Associations
-------------------------------------------------------------------------------
Service-Id        : 1 (Epipe)         Customer-Id       : 1
 - SAP : 1/1/3:1 (Ing)

===============================================================================
A:PE-1
```

The following command shows the policer hierarchy, including the child policers and their relationship to the intermediate arbiter (a3) and the root arbiter. It can be used to monitor the status of the child policers in the hierarchy. The output shows the assigned, offered and consumed capacity for each policer.

```
A:PE-1# show qos policer-hierarchy sap 1/1/3:1
===============================================================================
Policer Hierarchy - Sap 1/1/3:1
===============================================================================
Ingress Policer Control Policy : cfhp-1
Egress Policer Control Policy  :
-------------------------------------------------------------------------------
root (Ing)
|
| slot(1)
|
|--(A) : a3 (Sap 1/1/3:1)
|  |
|  |--(P) : Policer 1->1/1/3:1->4
|  |  |
|  |  |    [Level 2 Weight 25]
|  |  |    Assigned PIR:15000      Offered:45800
|  |  |    Consumed:15000
|  |  |
|  |  |    Assigned FIR:15000
|  |
|  |--(P) : Policer 1->1/1/3:1->3
|  |  |
|  |  |    [Level 2 Weight 25]
|  |  |    Assigned PIR:15000      Offered:45800
|  |  |    Consumed:15000
|  |  |
|  |  |    Assigned FIR:15000
|  |
|  |--(P) : Policer 1->1/1/3:1->2
|  |  |
|  |  |    [Level 2 Weight 50]
|  |  |    Assigned PIR:30000      Offered:45800
|  |  |    Consumed:30000
|  |  |
|  |  |    Assigned FIR:30000
|
|--(P) : Policer 1->1/1/3:1->5
|  |
|  |    [Level 5 Weight 1]
|  |    Assigned PIR:10000      Offered:10000
|  |    Consumed:10000
|  |
|  |    Assigned FIR:10000
|
|--(P) : Policer 1->1/1/3:1->1
|  |
|  |    [Level 1 Weight 1]
|  |    Assigned PIR:30000      Offered:45800
|  |    Consumed:30000
|  |
|  |    Assigned FIR:30000
root (Egr)
|
No Active Members Found on slot 1
===============================================================================
A:PE-1#
```

The complete information about the policer hierarchy can be seen by adding the detail parameter, as shown below, with alternative parameters to select more specific information.

- root-detail — Rates, depth and thresholds for the root arbiter.

- thresholds — CBS, MBS and high-prio-only thresholds with associated rates of child policers.

- priority-info — Discard-fair and discard-unfair thresholds, with number of associated children, for each of the root priority levels.

- depth — Parent policer and child PIR buckets depth, with PIR and FIR rate information.

- arbiter — Specific information of a given arbiter.

- port — For use with LAGs in different line cards or using adapt-qos link.

The output adds a good representation of the root arbiter thresholds, indicating the priority levels, discard-unfair and discard-fair thresholds, and how many child policers are associated with each level. It also includes the current depth of the child policer PIR buckets and the parent policer bucket.

```
A:PE-1# show qos policer-hierarchy sap 1/1/3:1 detail
===============================================================================
Policer Hierarchy - Sap 1/1/3:1
===============================================================================
Ingress Policer Control Policy : cfhp-1
Egress Policer Control Policy  :
-------------------------------------------------------------------------------
Legend :
(*) real-time dynamic value
(w) Wire rates
-------------------------------------------------------------------------------
root (Ing)
|
| slot(1)
|    MaxPIR:100000
|    ConsumedByChildren:100000
|    OperPIR:100000      OperFIR:100000
|
|    DepthPIR:8111 bytes
|  Priority 8
|    Oper Thresh Unfair:17408      Oper Thresh Fair:25600
|    Association count:0
|  Priority 7
|    Oper Thresh Unfair:17408      Oper Thresh Fair:25600
|    Association count:0
|  Priority 6
|    Oper Thresh Unfair:17408      Oper Thresh Fair:25600
|    Association count:0
|  Priority 5
|    Oper Thresh Unfair:17408      Oper Thresh Fair:25600
|    Association count:1
|  Priority 4
|    Oper Thresh Unfair:9728       Oper Thresh Fair:17408
|    Association count:0
|  Priority 3
|    Oper Thresh Unfair:9728       Oper Thresh Fair:17408
|    Association count:3
|  Priority 2
|    Oper Thresh Unfair:0          Oper Thresh Fair:8192
|    Association count:0
|  Priority 1
|    Oper Thresh Unfair:0          Oper Thresh Fair:8192
|    Association count:1
|
```

```
|--(A) : a3 (Sap 1/1/3:1)
|   |     MaxPIR:60000
|   |     ConsumedByChildren:60000
|   |     OperPIR:60000        OperFIR:60000
|   |
|   |     [Level 3 Weight 1]
|   |     Assigned PIR:60000      Offered:60000
|   |     Consumed:60000
|   |
|   |     Assigned FIR:60000
|   |
|   |--(P) : Policer 1->1/1/3:1->4
|   |   |     MaxPIR:60000        MaxCIR:20000
|   |   |     CBS:25600           MBS:77824
|   |   |     HiPrio:0
|   |   |     Depth:77876
|   |   |
|   |   |     OperPIR:15000        OperCIR:15000
|   |   |     OperFIR:15000
|   |   |     PacketByteOffset:0
|   |   |     StatMode: offered-total-cir
|   |   |
|   |   |     [Level 2 Weight 25]
|   |   |     Assigned PIR:15000      Offered:45800
|   |   |     Consumed:15000
|   |   |
|   |   |     Assigned FIR:15000
|   |
|   |--(P) : Policer 1->1/1/3:1->3
|   |   |     MaxPIR:60000        MaxCIR:20000
|   |   |     CBS:25600           MBS:77824
|   |   |     HiPrio:0
|   |   |     Depth:77834
|   |   |
|   |   |     OperPIR:15000        OperCIR:15000
|   |   |     OperFIR:15000
|   |   |     PacketByteOffset:0
|   |   |     StatMode: offered-total-cir
|   |   |
|   |   |     [Level 2 Weight 25]
|   |   |     Assigned PIR:15000      Offered:45800
|   |   |     Consumed:15000
|   |   |
|   |   |     Assigned FIR:15000
|   |
|   |--(P) : Policer 1->1/1/3:1->2
|   |   |     MaxPIR:60000        MaxCIR:20000
|   |   |     CBS:25600           MBS:77824
|   |   |     HiPrio:0
|   |   |     Depth:77848
|   |   |
|   |   |     OperPIR:30000        OperCIR:20000
|   |   |     OperFIR:30000
|   |   |     PacketByteOffset:0
|   |   |     StatMode: offered-total-cir
|   |   |
|   |   |     [Level 2 Weight 50]
|   |   |     Assigned PIR:30000      Offered:45800
|   |   |     Consumed:30000
|   |   |
|   |   |     Assigned FIR:30000
|   |
|--(P) : Policer 1->1/1/3:1->5
|   |     MaxPIR:10000        MaxCIR:10000
```

```
|   |     CBS:12800              MBS:12800
|   |     HiPrio:0
|   |     Depth:12854
|   |
|   |     OperPIR:10000         OperCIR:10000
|   |     OperFIR:10000
|   |     PacketByteOffset:0
|   |     StatMode: offered-total-cir
|   |
|   |     [Level 5 Weight 1]
|   |     Assigned PIR:10000      Offered:10000
|   |     Consumed:10000
|   |
|   |     Assigned FIR:10000
|
|--(P) : Policer 1->1/1/3:1->1
|   |     MaxPIR:100000       MaxCIR:0
|   |     CBS:0               MBS:126976
|   |     HiPrio:0
|   |     Depth:135
|   |
|   |     OperPIR:30000        OperCIR:0
|   |     OperFIR:30000
|   |     PacketByteOffset:0
|   |     StatMode: offered-total-cir
|   |
|   |     [Level 1 Weight 1]
|   |     Assigned PIR:30000      Offered:45800
|   |     Consumed:30000
|   |
|   |     Assigned FIR:30000

root (Egr)
|
No Active Members Found on slot 1


===========================================================================
A:PE-1#
```

The output above gives the depth of the parent policer, which can be used with the output below to explain the operation of the policing in this example.

```
A:PE-1# show qos policer sap 1/1/3:1 detail | match expression "Slot | Bytes | Kbps"
Policer Info (1->1/1/3:1->1), Slot 1
Depth PIR          : 153 Bytes        Depth CIR          : 0 Bytes
Depth FIR          : 153 Bytes
Admin PIR          : 100000 Kbps      Admin CIR          : 0 Kbps
Oper PIR           : 30000 Kbps       Oper CIR           : 0 Kbps
Oper FIR           : 30000 Kbps
Offered Rate       : 45800 Kbps
Parent PIR         : 30000 Kbps       Parent FIR         : 30000 Kbps
Consumed           : 30000 Kbps
Policer Info (1->1/1/3:1->2), Slot 1
Depth PIR          : 77828 Bytes      Depth CIR          : 25624 Bytes
Depth FIR          : 77828 Bytes
Admin PIR          : 60000 Kbps       Admin CIR          : 20000 Kbps
Oper PIR           : 30000 Kbps       Oper CIR           : 20000 Kbps
Oper FIR           : 30000 Kbps
Offered Rate       : 45800 Kbps
Parent PIR         : 30000 Kbps       Parent FIR         : 30000 Kbps
Consumed           : 30000 Kbps
Policer Info (1->1/1/3:1->3), Slot 1
Depth PIR          : 77858 Bytes      Depth CIR          : 25634 Bytes
```

```
Depth FIR          : 77858 Bytes
Admin PIR          : 60000 Kbps        Admin CIR         : 20000 Kbps
Oper PIR           : 15000 Kbps        Oper CIR          : 15000 Kbps
Oper FIR           : 15000 Kbps
Offered Rate       : 45800 Kbps
Parent PIR         : 15000 Kbps        Parent FIR        : 15000 Kbps
Consumed           : 15000 Kbps
Policer Info (1->1/1/3:1->4), Slot 1
Depth PIR          : 77838 Bytes       Depth CIR         : 25614 Bytes
Depth FIR          : 77838 Bytes
Admin PIR          : 60000 Kbps        Admin CIR         : 20000 Kbps
Oper PIR           : 15000 Kbps        Oper CIR          : 15000 Kbps
Oper FIR           : 15000 Kbps
Offered Rate       : 45800 Kbps
Parent PIR         : 15000 Kbps        Parent FIR        : 15000 Kbps
Consumed           : 15000 Kbps
Policer Info (1->1/1/3:1->5), Slot 1
Depth PIR          : 12814 Bytes       Depth CIR         : 12814 Bytes
Depth FIR          : 12814 Bytes
Admin PIR          : 10000 Kbps        Admin CIR         : 10000 Kbps
Oper PIR           : 10000 Kbps        Oper CIR          : 10000 Kbps
Oper FIR           : 10000 Kbps
Offered Rate       : 10000 Kbps
Parent PIR         : 10000 Kbps        Parent FIR        : 10000 Kbps
Consumed           : 10000 Kbps
A:PE-1#
```

From the output above, it can be seen that the offered rate for policers 1-4 is 45800Kbps, in fact it is the same for policer 5 but this is capped at the admin PIR rate, 10000Kbps.

The depth of the parent policer is only 8111 bytes, so this is not causing any discarding of priority 2-5 traffic at the parent policer as their discard thresholds are all above this value. Therefore the drops in policers 2-5 are all occurring in the child policers.

Policer 5 is consuming all of its operational capacity (PIR, CIR and FIR), and it can be seen that the level of the PIR bucket is 12814 bytes, which is slightly above its MBS of 12800 bytes. The level of the PIR bucket will oscillate around the MBS value as tokens are added to exceed the threshold (causing discards) then the draining reduces the level to just below the threshold (allowing forwarding).

Policers 2-4 are functioning in the same way as policer 5, as can be seen from their PIR bucket levels (levels are 77828 bytes with MBS of 77824), resulting in the PIR buckets constraining the rates of the traffic through these policers. This is happening because the arbiter "a3" is distributing its 60000Kbps in the configured ratio to these policers, which changes the operational PIR to 30000Kbps for policer 2 and 15000Kbps for policers 3 and 4, all being below the offered traffic rate. A similar effect can be seen with the CIR rates and bucket depths, as the operational CIR rate of policer 2 has reached its administrative value with those of policer 3 and 4 being constrained by the operational PIR. The CIR bucket depths are just above the CBS, again this will oscillate causing traffic to both in-profile and out-of-profile. As this is steady state traffic, the operational FIR rates for these policers have settled to match their operational PIR rates.

Policer 1 is also discarding traffic at the PIR bucket but it is also discarding traffic at the parent policer. This can be seen by the fact that policer 1 PIR depth is nowhere near its MBS whereas the parent policer level is just below the priority 1 discard-fair threshold. The level of the parent policer bucket will oscillate around this threshold causing policer 1 traffic to be discarded, which in turn is reflected back into the level of tokens in the policer 1 PIR bucket.

As this example is based on ingress unicast policing, the traffic exits the policers and then accesses the switch fabric using a set of shared-queue (policer-output-queues). The parameters for these queues can be seen using the following **show** command.

```
A:PE-1# show qos shared-queue "policer-output-queues" detail
```

```
===============================================================================
QoS Shared Queue Policy
===============================================================================
-------------------------------------------------------------------------------
Shared Queue Policy (policer-output-queues)
-------------------------------------------------------------------------------
Policy        : policer-output-queues
Description   : Default Policer Output Shared Queue Policy

-------------------------------------------------------------------------------
Queue CIR       PIR       CBS       MBS       HiPrio  Multipoint Pool-Name
-------------------------------------------------------------------------------
1     0         100       1         50        10      FALSE
2     25        100       3         50        10      FALSE
3     25        100       10        50        10      FALSE
4     25        100       3         25        10      FALSE
5     100       100       10        50        10      FALSE
6     100       100       10        50        10      FALSE
7     10        100       3         25        10      FALSE
8     10        100       3         25        10      FALSE
9     0         100       1         50        10      TRUE
10    25        100       3         50        10      TRUE
11    25        100       10        50        10      TRUE
12    25        100       3         25        10      TRUE
13    100       100       10        50        10      TRUE
14    100       100       10        50        10      TRUE
15    10        100       3         25        10      TRUE
16    10        100       3         25        10      TRUE

-------------------------------------------------------------------------------
FC    UCastQ  MCastQ  BCastQ  UnknownQ
-------------------------------------------------------------------------------
be    1       9       9       9
l2    2       10      10      10
af    3       11      11      11
l1    4       12      12      12
h2    5       13      13      13
ef    6       14      14      14
h1    7       15      15      15
nc    8       16      16      16

-------------------------------------------------------------------------------
Associations
-------------------------------------------------------------------------------
Service : 1              SAP : 1/1/3:1
===============================================================================
A:PE-1#
```

For egress policing, policed traffic can access the exit port by a queue-group, the default being called policer-output-queues. The following shows the parameters for these queues.

```
A:PE-1# show qos queue-group "policer-output-queues" detail
===============================================================================
QoS Queue-Group Ingress
===============================================================================
===============================================================================
QoS Queue-Group Egress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name     : policer-output-queues
```

```
Description   : Default egress policer output queues.

----------------------------------------------------------------------
Q  CIR Admin PIR Admin CBS        HiPrio PIR Lvl/Wt    Parent     BurstLimit(B)
   CIR Rule  PIR Rule  MBS               CIR Lvl/Wt    Wred-Queue   Slope
   Named-Buffer Pool
----------------------------------------------------------------------
1  0         max       def        def    1/1           None         default
   closest   closest   def               0/1           disabled     default
   (not-assigned)
2  0         max       def        def    1/1           None         default
   closest   closest   def               0/1           disabled     default
   (not-assigned)


======================================================================
Queue Group Ports (access)
======================================================================
Port              Sched Pol         Acctg Pol Stats     Description
----------------------------------------------------------------------
1/1/3                               0          No
1/1/4                               0          No
----------------------------------------------------------------------


======================================================================
Queue Group Ports (network)
======================================================================
Port              Sched Pol         Acctg Pol Stats     Description
----------------------------------------------------------------------
No Matching Entries


======================================================================
Queue Group Sap FC Maps
======================================================================
Sap Policy     FC Name           Queue Id
----------------------------------------------------------------------
No Matching Entries
======================================================================
A:PE-1#
```

The following output shows the relative thresholds in the parent policer when a priority level has 0, 1 or 2 associated children, both with and without using the fixed parameter.

```
A:PE-1# show qos policer-hierarchy sap 1/1/3:2 ingress priority-info
===============================================================================
Policer Hierarchy - Sap 1/1/3:2
===============================================================================
Ingress Policer Control Policy : cfhp-2
-------------------------------------------------------------------------------
root (Ing)
|
| slot(1)
|  Priority 8
|    Oper Thresh Unfair:4352       Oper Thresh Fair:5120
|    Association count:0
|  Priority 7
|    Oper Thresh Unfair:4352       Oper Thresh Fair:5120
|    Association count:0
|  Priority 6
|    Oper Thresh Unfair:4352       Oper Thresh Fair:5120
|    Association count:2 fixed
|  Priority 5
|    Oper Thresh Unfair:3328       Oper Thresh Fair:4096
```

```
|     Association count:1 fixed
|   Priority 4
|     Oper Thresh Unfair:2304         Oper Thresh Fair:3072
|     Association count:0 fixed
|   Priority 3
|     Oper Thresh Unfair:1280         Oper Thresh Fair:2048
|     Association count:2
|   Priority 2
|     Oper Thresh Unfair:0            Oper Thresh Fair:1024
|     Association count:1
|   Priority 1
|     Oper Thresh Unfair:0            Oper Thresh Fair:0
|     Association count:0


===============================================================================
A:PE-1#
```

Where

- Priority Level 1 has no children so both its fair and unfair thresholds are 0.

- Priority Level 2 has one child so its unfair threshold is 0 and its fair threshold is at the configured mbs-contribution [1024 bytes] (given that this is larger than the min-thresh-separation).

- Priority Level 3 has two children so its unfair threshold is equal to the min-thresh-separation plus the fair threshold of priority 2 [256+1024=1280 bytes]. Its fair threshold is effectively the mbs-contribution plus the fair threshold of priority 2 [1024+1024=2048 bytes] (given that the mbs-contribution is larger than 2x min-thresh-separation).

- Priorities 4, 5 and 6 have the fixed parameter configured. Even though priority 4 has no children, priority 5 has only one child and priority 6 has two children, all three priorities have the same incremental values for their unfair and fair discard threshold. This result in

  - Priority 4's unfair threshold being equal to the min-thresh-separation plus the fair threshold of priority 3 [256+2048=2304 bytes]. Its fair threshold is effectively the mbs-contribution plus the fair threshold of priority 3 [1024+2048=3072 bytes] (given that the mbs-contribution is larger than 2x min-thresh-separation).

  - Priority 5's unfair threshold being equal to the min-thresh-separation plus the fair threshold of priority 4 [256+3072=3328 bytes]. Its fair threshold is effectively the mbs-contribution plus the fair threshold of priority 4 [1024+3072=4096 bytes] (given that the mbs-contribution is larger than 2x min-thresh-separation).

  - Priority 6's unfair threshold being equal to the min-thresh-separation plus the fair threshold of priority 5 [256+4096=4352 bytes]. Its fair threshold is effectively the mbs-contribution plus the fair threshold of priority 5 [1024+4096=5120 bytes] (given that the mbs-contribution is larger than 2x min-thresh-separation).

Note that the above parameter values were chosen to exactly match available hardware values to simplify the output.

Conclusion

This chapter has described the configuration of Class Fair Hierarchical Policing for SAPs. This hardware policing provides low latency ingress and egress prioritized traffic control with the ability to provide fairness between child policers at the same parent policer priority level.

FP and Port Queue Groups

This chapter provides information about FP and port queue groups.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This chapter is applicable to the 7950 XRS-16c/20, 7750 SR-7/12, 7750 SR-a4/8, 7750 SR-c4/12, and 7450 ESS-6/6v/7/12 platforms and assumes only FP2- and higher-based line cards are used.

Port queue groups can be configured on FP1 line cards but not all functions described in this chapter are supported on FP1 line cards. FP queue groups are not supported on FP1 line cards because FP queue groups rely on hardware policing.

The configuration was tested on Release 13.0.R7. There are no other specific prerequisites for this configuration.

Overview

Queue groups provide flexible QoS control beyond that available by default for SAPs and network interfaces.

Many applications require detailed QoS control for SAPs, with aggregated QoS control across the core on network interfaces. Queue groups allow the reverse, specifically aggregated QoS control for multiple SAPs and per-network interface QoS control. This is summarized in Table 3: Default QoS and Queue Group Comparison .

Table 3: Default QoS and Queue Group Comparison

	Default	**Queue group**
SAPs	• Per-SAP ingress queues and policers • Per-SAP egress queues and policers	• Policers for a set of SAPs at ingress • Queues for a set of SAPs at egress
Network interfaces	• Per-MDA ingress network queues	• Policers per network interface at ingress

	Default	Queue group
	• Per-port egress network queues	• Policers and queues per network interface at egress

Queue groups were introduced in SR OS Release 7 as a mechanism of grouping a set of queues and enhanced in subsequent Releases. Release 10 added the ability to configure policers within queue groups and introduced a more flexible configuration where the queue group template could be applied multiple times with different instances. The term queue groups was retained even though they can now contain queues, policers, or both. This chapter describes this new configuration.

Queue groups can be used for other applications than those listed in Table 3: Default QoS and Queue Group Comparison ; for example:

- Pseudowire (PW) QoS

  Providing QoS control for spoke SDPs in the various pipe services, different types of VPLS services, and associated IES/VPRN interfaces (see chapter Pseudowire QoS).

- Carrier Supporting Carrier (CSC) services

  Providing QoS control for CSC network interfaces in VPRN services.

- Ingress QoS control on VPRN network interfaces

  Providing for ingress QoS control of unicast traffic into a VPRN over automatically created or manually created bindings in a VPRN service.

- Network egress QoS

  Providing queue rates in kbps instead of the percentage of the port rate, and queue CBS/MBS in kbytes instead of a fractional percentage of the pool size.

When egress queue groups are used for SAPs, the groups provide a similar functionality to multi-service sites in that both can provide an overall rate for a set of SAPs. However, multi-service sites also provide per-SAP QoS control, whereas the queue groups do not.

Queue groups are not applicable to subscriber management except when egress policing is used.

> **Note:**
> The queue groups described in this chapter are different from those that are configured when using a high-scale Ethernet MDA (HS-MDAv2).

Configuration

Following the description of the overall configuration of queue groups, their configuration is illustrated using examples on a SAP, a network interface, and with egress subscriber policing.

The steps required to configure a queue group are summarized as follows:

- Create a queue group template:

  - Ingress or egress

- Apply an instance of the queue group template to:

  - FP ingress

- • Access or network
    - – Ethernet egress port
        - • Access or network
- • Redirect traffic to a policer or queue in the FP or port queue group instance per forwarding class (FC) within a:
    - – SAP ingress or egress QoS policy
    - – Network QoS policy for both ingress and egress
- • Consider post-egress policer queuing

Creating Queue Group Templates

Queue group templates are configured separately for ingress and egress; an ingress queue group template with the same name as an egress queue group template is a different object.

Ingress queue group templates can contain policers or queues, but not both.

Egress queue group templates can contain policers or queues, or both; queue 1 is created by default and cannot be deleted. A forwarding class (FC) can also be configured to determine the queue mapping of the FC for forwarded SAP egress post-policer traffic.

To summarize the use of policers with respect to queue groups:

- • Ingress FP queue groups can only use policers.
- • Port egress access queue groups cannot use policers.
- • Network ingress and egress applications can only use policers in a queue group.

The configuration of an ingress and egress queue group template is as follows:

```
configure
    qos
        queue-group-templates
            ingress
                queue-group <queue-group-name> [create]
                    description <description-string>
                    policer <policer-id> [create]
                    queue <queue-id> [multipoint] [<queue-type>]
                                        [<queue-mode>] [create]
            egress
                queue-group <queue-group-name> [create]
                    description <description-string>
                    fc <fc-name> [create]
                        queue <queue-id>
                    policer <policer-id> [create]
                    queue <queue-id> [<queue-type>] [create]
```

The configuration of the policer and queue is similar to the configuration in a SAP ingress and egress QoS policy. However, the SAP ingress QoS policy allows the configuration of a percent-rate and the SAP egress QoS policy allows the configuration of an avg-frame-overhead, neither of which is available in a queue group template. An ingress queue group template supports the creation of up to 32 queues and 32 policers, while an egress queue group template supports the creation of up to 8 queues and 8 policers.

Because an egress queue group template uses kbps for the CIR/PIR rates and bytes/kbytes for the CBS/MBS, unlike a network-queue QoS policy, a network egress interface can use them when an egress queue group template is applied.

The system instantiates the following queue group templates by default. Table 5: Queue Group Templates - Egress shows the ingress queue group templates:

Table 4: Queue Group Templates - Ingress

Group name	Description
_tmnx_nat_ing_q_grp	NAT/LNS Ingress Queue Group Template
_tmnx_nat_ing_q_grp_v2	NAT/LNS Ingress Queue Group Template for ISAv2
_tmnx_lns_esm_ing_q_grp	LNS ESM Ingress Queue Group Template

Table 5: Queue Group Templates - Egress shows the egress queue group templates:

Table 5: Queue Group Templates - Egress

Group name	Description
_tmnx_nat_egr_q_grp	NAT/LNS Egress Queue Group Template
policer-output-queues	Default egress policer output queues
_tmnx_nat_egr_q_grp_v2	NAT/LNS Egress Queue Group Template for ISAv2
_tmnx_lns_esm_egr_q_grp	LNS ESM Egress Queue Group Template

This chapter will only discuss the policer-output-queues queue group from the preceding table, an instance of which is created on all egress access and hybrid ports, to be used for egress policed traffic.

Applying Queue Group Templates

Ingress

An ingress queue group template containing only policers can be applied to an FP ingress. When applied, an instance identifier must be specified, which represents the instantiated instance of the related queue group template.

If an attempt is made to configure a queue group template containing queues for an FP ingress, the following error message appears:

```
*A:PE-1>config#  card 5 fp 1 ingress access queue-group "qg2" instance 1 create
MINOR: CHMGR #1164 Cannot attach a Queue Group containing queues
```

Because 7750 SR-a4/8 platforms do not support SAP or network hardware policers, FP ingress queue groups are not supported on those platforms.

Queue group templates containing only queues can be applied to port ingress. However, only one such queue group can be applied per ingress port.

An ingress template can be applied to an FP ingress as either an ingress access queue group or an ingress network queue group, or both. In each case:

- An accounting policy can be applied.

- Statistics collection can be enabled.

- A description can be configured.

- A policer control policy can be applied, containing parent arbiters for the queue group policers, and its parameters can be overridden.

- Policer parameters configured within the queue group template can also be overridden.

> **Note:**
> When the queue group applies to a tunnel object that can move between different network interfaces, and consequently different network ingress FPs (for example, PW QoS), a network ingress queue group instance must be applied to each FP ingress that could be used.

The configuration of ingress queue group templates for ingress FPs is as follows:

```
configure
    card <slot-number>
        fp [<fp-number>]
            ingress
                access
                    queue-group <queue-group-name>
                            instance <[1..65535]> [create]
                        accounting-policy <acct-policy-id>
                        collect-stats
                        description <description-string>
                        policer-control-override [create]
                            max-rate {<rate> | max}
                            priority-mbs-thresholds
                                min-thresh-separation <size> [bytes | kilobytes]
                                priority <level>
                                    mbs-contribution <size> [bytes | kilobytes]
                        policer-control-policy
                                <policer-control-policy-name>
                        policer-override
                            policer <policer-id> [create]
                network
                    queue-group <queue-group-name>
                            instance <[1..65535]> [create]
                        accounting-policy <acct-policy-id>
                        collect-stats
                        description <description-string>
                        policer-control-override [create]
                            max-rate {<rate> | max}
                            priority-mbs-thresholds
                                min-thresh-separation <size> [bytes | kilobytes]
                                priority <level>
                                    mbs-contribution <size> [bytes | kilobytes]
                        policer-control-policy
                                <policer-control-policy-name>
                        policer-override
                            policer <policer-id> [create]
```

> **Note:**
> Configure FP ingress access and network queue group instances consistently across FPs relating to a LAG.

A single ingress template containing only queues can also be applied to an Ethernet port ingress, which can be used only for access ingress:

- An accounting policy can be applied.

- Statistics collection can be enabled.

- A description can be configured.

- Queue parameters configured within the queue group template can be overridden.

- A scheduler policy can be applied containing parent schedulers for the queue group queues.

The configuration of an ingress queue group template for an ingress port is as follows:

```
configure
    port <port-id>
        ethernet
            access
                ingress
                    queue-group <queue-group-name> [create]
                        accounting-policy <acct-policy-id>
                        collect-stats
                        description <description-string>
                        queue-overrides
                            queue <queue-id> [create]
                        scheduler-policy <scheduler-policy-name>
```

Egress

An egress queue group template containing policers or queues, or both, can be applied to an Ethernet port network egress. Only queue group templates containing queues (not policers) can be applied to an Ethernet port access egress. When applied, an instance identifier can be specified, which represents the instantiated instance of the related queue group template; the default being 1.

If an attempt is made to configure a queue group template containing policers for an Ethernet port access egress, the following error message appears:

```
*A:PE-1# configure port 5/1/5 ethernet access egress queue-group "qg1" instance 1 create
MINOR: CLI Could not create/change "qg1".
MINOR: PMGR #1324 Cannot attach a Queue Group containing policers
```

An egress template can be applied as either an egress access queue group or an egress network queue group, or both. In each case:

- An accounting policy can be applied.

- Statistics collection can be enabled.

- A description can be configured.

- Queue parameters configured within the queue group template can be overridden.

- A scheduler policy can be applied containing parent schedulers for the queue group queues.

When applied to the Ethernet access egress, a host-match can be configured, which is used to select a queue group for subscriber egress policed traffic; see later in this chapter for details.

When applied to the Ethernet network egress, a policer control policy can be applied, which contains parent arbiters for the queue group policers.

Note:
When the queue group applies to a tunnel object that can move between different network interfaces, and consequently different network egress ports (for example, PW QoS), a network egress queue group instance must be applied to each network egress port that could be used.

The configuration of egress queue group templates for egress ports is as follows:

```
configure
    port <port-id>
        ethernet
            access
                egress
                    queue-group <queue-group-name> [create]
                            [instance <instance-id>]
                        accounting-policy <acct-policy-id>
                        agg-rate
                            limit-unused-bandwidth
                            queue-frame-based-accounting
                            rate <kilobits-per-second>
                        collect-stats
                        description <description-string>
                        host-match dest <destination-string> [create]
                        queue-overrides
                            queue <queue-id> [create]
                        scheduler-policy <scheduler-policy-name>
            network
                    queue-group <queue-group-name> [create]
                            [instance <instance-id>]
                        accounting-policy <acct-policy-id>
                        agg-rate
                            limit-unused-bandwidth
                            queue-frame-based-accounting
                            rate <kilobits-per-second>
                        collect-stats
                        description <description-string>
                        policer-control-policy
                                <policer-control-policy-name>
                        queue-overrides
                            queue <queue-id> [create]
                        scheduler-policy <scheduler-policy-name>
```

Note:
When port egress queue groups are used with a LAG, the system enforces a consistent configuration between the ports (based on only the configuration of the primary LAG port).

Redirecting Traffic to a Queue Group Queue or Policer

There are multiple ways to redirect traffic to a queue group queue or policer. The redirection is always on a per-FC basis so that different FCs within the same policy can use a mix of a local SAP policer, a local SAP or network queue, or a queue group policer or queue (where applicable).

Redirection to a queue group queue or policer is not supported for subscribers; attempting to do so will display the following errors:

```
*A:PE-1>config>qos>sap-ingress# fc "af" policer 1 fp-redirect-group
MINOR: QOS #1489 Cannot assign a queue-group because SLA profile references to this policy
 exist
```

```
*A:PE-1>config>qos>sap-egress# fc af policer 1 port-redirect-group-queue
MINOR: QOS #1628 Cannot assign a queue-group because SLA profile references to this policy
 exist

*A:PE-1>config>qos>sap-egress# fc af queue 1 port-redirect-group-queue
MINOR: QOS #1628 Cannot assign a queue-group because SLA profile references to this policy
 exist

*A:PE-1>config>subscr-mgmt>sla-prof>ingress# qos 40
MINOR: SUBMGR #1110 Cannot assign Qos Policy, QOS Plcy contains queue-group references

*A:PE-1>config>subscr-mgmt>sla-prof>egress# qos 40
MINOR: SUBMGR #1110 Cannot assign Qos Policy, QOS Plcy contains queue-group references
```

SAP Ingress Redirection

Traffic can be redirected per FC to a policer in an FP ingress access queue group, using the fp-redirect-group parameter in the SAP ingress QoS policy, as follows:

```
configure
    qos
        sap-ingress <policy-id> [create]
            fc <fc-name> [create]
                policer <policer-id> fp-redirect-group
```

The policer-id is referencing a policer in the FP ingress queue group.

The queue group name and instance to be used is not yet configured, allowing this QoS policy to be applied to objects using different FP ingress queue group instances, which results in greater flexibility.

Compare this preferred configuration to the original configuration used for ingress port queue groups, as follows, where the queue group name is specified within the SAP ingress QoS policy and the instance is not available:

```
configure
    qos
        sap-ingress <policy-id> [create]
            fc <fc-name> [create]
                queue <queue-id> group <queue-group-name>
```

In the preferred configuration, redirection is completed by applying the SAP ingress QoS policy to the SAP with the queue group name and instance to be used, as follows:

```
configure
    service
        {apipe|cpipe|epipe|fpipe|ipipe} <service-id>
            sap <sap-id>
                ingress
                    qos <policy-id> fp-redirect-group <queue-group-name>
                                                    instance <instance-id>
        {ies|vprn} <service-id>
            interface <ip-int-name>
                sap <sap-id>
                    ingress
                        qos <policy-id>
                            fp-redirect-group <queue-group-name>
                                    instance <instance-id>
```

SAP Egress Redirection

Traffic can be redirected per FC to a queue in a port egress access queue group, using the port-redirect-group parameter, as follows:

```
configure
    qos
        sap-egress <policy-id> [create]
            fc <fc-name> [create]
                queue <queue-id> port-redirect-group-queue
```

The queue-id is referencing a queue in the port egress queue group.

The queue group name and instance to be used is not yet configured, allowing this QoS policy to be applied to objects using different port egress queue groups.

Compare this preferred configuration to the original configuration, which allows the queue group name with an instance to be specified:

```
queue <id> {group <grp-name> [instance instance-id]}
```

In the preferred configuration, redirection is completed by applying the SAP egress QoS policy to the SAP with the queue group name and instance to be used, as follows:

```
configure
    service
        {apipe|cpipe|epipe|fpipe|ipipe} <service-id>
            sap <sap-id>
                egress
                    qos <policy-id>
                        port-redirect-group <queue-group-name>
                                            instance <instance-id>
        {ies|vprn} <service-id>
            interface <ip-int-name>
                sap <sap-id>
                    egress
                        qos <policy-id>
                            port-redirect-group <queue-group-name>
                                                instance <instance-id>
```

Network Ingress Redirection

Traffic can be redirected per FC to a policer in an FP ingress network queue group, using the fp-redirect-group parameter, as follows:

```
configure
    qos
        network <network-policy-id> [create]
            ingress
                fc <fc-name>
                    fp-redirect-group <policer-type> <policer-id>
```

The policer-id is referencing a policer in the FP ingress queue group.

The traffic usage for each policer type for the supported services is shown in Table 4.

Table 6: Network Ingress FP Queue Group Policer Usage

Policer Type	Usage
broadcast-policer	Broadcast traffic for PW QoS in a VPLS service
mcast-policer	Multipoint traffic (except for ingress QoS control on VPRN network interfaces and CSC services where the IP multicast traffic uses the ingress network queues or queue group related to the network interface) Multicast traffic for PW QoS in a VPLS service
unknown-policer	Unknown traffic for PW QoS in a VPLS service
policer	Unicast traffic

The queue group name and instance to be used is not yet configured, allowing this QoS policy to be applied to objects using different FP ingress queue group instances.

The redirection is completed by applying the network QoS policy to the required object with the queue group name and instance to be used, as follows (only the network interface redirection is shown):

```
configure
    router
        interface <interface-name>
            qos <network-policy-id>
                        ingress-fp-redirect-group <queue-group-name>
                        ingress-instance <instance-id>
```

Network Egress Redirection

Traffic can be redirected per FC to a policer or a queue in a port egress network queue group, using the port-redirect-group parameter, as follows:

```
configure
    qos
        network <network-policy-id> [create]
            egress
                fc <fc-name>
                    port-redirect-group {queue <queue-id> |
                                    policer <plcr-id> [queue <queue-id>]}
```

The queue-id and policer-id are referencing a queue and a policer in the port egress queue group.

The queue group name and instance to be used are not yet configured, allowing this QoS policy to be applied to objects using different port egress queue groups.

The redirection is completed by applying the network QoS policy to the required object with the queue group name and instance to be used, as follows (only the network interface redirection is shown):

```
configure
    router
        interface <interface-name>
            qos <network-policy-id>
```

```
                    egress-fp-redirect-group <queue-group-name>
                    egress-instance <instance-id>
```

> **Note:**
> Non-IPv4/non-IPv6/non-MPLS packets are not subject to the redirection to an egress queue group instance and will remain on the regular port network egress queues. When using an egress port scheduler, parent the related regular network port queues to appropriate port scheduler priority levels, to ensure the required operation under port congestion. This is important for protocol traffic, such as LACP, EFMOAM, ETH-CFM, ARP, and IS-IS, which by default use the FC nc regular network port queue.

Post-egress Policer Queuing

The queuing of traffic exiting an egress policer is described as follows for SAP egress and network egress.

SAP egress policed traffic exits a port, using either a local queue or an egress queue group instance queue. If an egress queue group instance queue is to be used, the policed traffic can be mapped to the queue in the following ways, listed in reverse preference order (first is lowest preference):

1. By default, SAP egress policed traffic exits using a queue in the policer-output-queues queue group instance when only a policer is configured:

```
configure
    qos
        sap-egress <policy-id> [create]
            fc <fc-name> [create]
                policer <policer-id>
```

The system always creates this egress queue group template and applies it to all access and hybrid ports (for use by the access part of the hybrid port) as instance 1. The queue group template contains the mapping for all FCs to one of the two created queues to determine which queue is to be used.

The default configuration of the policer-output-queues group template is as follows (without the queue details):

```
configure
    qos
        queue-group-templates
            egress
                queue-group "policer-output-queues" create
                    description "Default egress policer
                                 output queues."
                    queue 1 best-effort create
                    exit
                    queue 2 expedite create
                    exit
                    fc af create
                        queue 1
                    exit
                    fc be create
                        queue 1
                        exit
                    fc ef create
                        queue 2
                    exit
                    fc h1 create
                        queue 2
                    exit
```

```
                    fc h2 create
                        queue 2
                    exit
                    fc l1 create
                        queue 1
                    exit
                    fc l2 create
                        queue 1
                    exit
                    fc nc create
                        queue 2
                    exit
                exit
```

This queue group template cannot be deleted, but its configuration can be modified. The configuration commands under a port access egress queue group instance are also available under the policer-output-queues queue group instance.

2. The FC in the SAP egress QoS policy can be mapped to a policer with its traffic redirected to a port access egress queue group, as follows:

```
configure
    qos
        sap-egress <policy-id> [create]
            fc <fc-name> [create]
                policer <policer-id> port-redirect-group-queue
```

The QoS policy must be applied to the SAP with a redirection to a queue group instance. By default, queue 1 in the queue group will be used, but other queues can be created and used by creating an FC-to-queue mapping to them.

3. Similar to step 2, the FC in the SAP egress QoS policy can be mapped to a policer with its traffic redirected to a port access egress queue group instance and the queue within that queue group instance to be used, as follows:

```
configure
    qos
        sap-egress <policy-id> [create]
            fc <fc-name> [create]
                policer <policer-id> port-redirect-group-queue queue <id>
```

The QoS policy must again be applied to the SAP with a redirection to a queue group instance. In this case, the queue specified with the port-redirect-group-queue will be used.

Network egress can only use policers that are created within an egress queue group instance, and this requires the egress FC to have the port-redirect-group parameter configured.

With the following configuration, traffic exits the port on the network egress queue to which its FC is mapped (not on a queue group queue):

```
configure
    qos
        network <network-policy-id> [create]
            egress
                fc <fc-name>
                    port-redirect-group policer <plcr-id>
```

If a queue is specified on the port-redirect-group statement, as follows, the traffic exits on the referenced queue within the port network egress queue group instance:

```
configure
    qos
        network <network-policy-id> [create]
            egress
                fc <fc-name>
                    port-redirect-group policer <plcr-id> queue <queue-id>
```

Configuration Examples

The following configuration examples are for the use of queue groups with a SAP, network interface, and egress policed traffic subscriber traffic. Different queue group templates and instances have been used to highlight the flexibility of the configuration.

SAP Configuration Example

In this example, a SAP is created in an IES service using ingress queue group qg1 instance 1 and egress queue group qg1 instance 1.

First, the queue group templates are configured, as follows:

```
configure
    qos
        queue-group-templates
            ingress
                queue-group "qg1" create
                    policer 1 create
                    exit
                    policer 2 create
                    exit
                exit
            exit
            egress
                queue-group "qg1" create
                    queue 1 best-effort create
                    exit
                    queue 2 expedite create
                    exit
                exit
                queue-group "policer-output-queues" create
                    queue 3 best-effort create
                    exit
                    fc l2 create
                        queue 3
                    exit
                exit
```

The policer-output-queues queue group template has been modified to add an extra queue and map FC l2 to it. This is used by the traffic mapped to policer 1 in the SAP egress QoS policy.

The ingress template is applied to card 5 fp 1 to create an FP access queue group instance, as follows:

```
configure
    card 5
        fp 1
```

```
                ingress
                    access
                        queue-group "qg1" instance 1 create
                        exit
                    exit
```

The egress template is applied to port 5/1/5 to create a port access queue group instance, as follows:

```
configure
    port 5/1/5
        ethernet
            mode hybrid
            encap-type dot1q
            access
                egress
                    queue-group "qg1" instance 1 create
                    exit
                exit
```

The SAP ingress QoS policy is created to redirect FC af to policer 1 and FC ef to policer 2 in the FP ingress queue group. To emphasize that the redirection is per FC and can coexist with locally mapped queues or policers, FC be is mapped to the local queue 1 and FC l2 to local queue 2, as follows:

```
configure
    qos
        sap-ingress 10 create
            queue 1 create
            exit
            queue 2 create
            exit
            queue 11 multipoint create
            exit
            fc "af" create
                policer 1 fp-redirect-group
            exit
            fc "be" create
                queue 1
            exit
            fc "ef" create
                policer 2 fp-redirect-group
            exit
            fc "l2" create
                queue 2
            exit
            dot1p 0 fc "be"
            dot1p 1 fc "l2"
            dot1p 2 fc "af"
            dot1p 3 fc "ef"
        exit
```

The SAP egress QoS policy is created to redirect FC af to queue 1 in the port egress queue group and FC ef to local policer 2, with its traffic exiting through queue 2 in the port egress queue group. FC be is mapped to the local queue 1 and FC l2 to local policer 1. The policer 1 traffic will exit using queue 3 of the policer-output-queues queue group, because the mapping for FC l2 has been modified in its template, as follows:

```
configure
    qos
        sap-egress 10 create
            queue 1 create
```

```
            exit
            policer 1 create
            exit
            policer 2 create
            exit
            fc af create
                queue 1 port-redirect-group-queue
            exit
            fc be create
                queue 1
            exit
            fc ef create
                policer 2 port-redirect-group-queue queue 2
            exit
            fc l2 create
                policer 1
            exit
        exit
```

The IES service is created with an interface on SAP 5/1/5:1, using the preceding ingress and egress QoS policies and queue group instances. A second interface (not shown) is used to forward the traffic to and from this interface, as follows:

```
configure
    service
        ies 1 customer 1 create
            interface "PE-1-int1-2" create
                address 10.1.2.1/24
                sap 5/1/5:1 create
                    ingress
                        qos 10 fp-redirect-group "qg1" instance 1
                    exit
                    egress
                        qos 10 port-redirect-group "qg1" instance 1
                    exit
                exit
            exit
            no shutdown
        exit
```

The following output shows the configuration and the QoS state after sending traffic through the service.

The traffic statistics are either counted in the SAP queues and policers, or in the queue group instance queues and policers, but not in both. However, summary statistics per SAP are available when using FP ingress queue groups, as shown.

The queue group templates can be shown. The output shows the details related to the ingress queue group template:

```
*A:PE-1# show qos queue-group "qg1" ingress detail
===============================================================================
QoS Queue-Group Ingress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name    : qg1
Description    : (Not Specified)
-------------------------------------------------------------------------------
Q   Mode     CIR Admin  PIR Admin  CBS          HiPrio  PIR Lvl/Wt Parent
             CIR Rule   PIR Rule   MBS                  CIR Lvl/Wt BurstLimit(B)
             Named-Buffer Pool     Pkt Bt Ofst          Adv Config Policy Name
```

```
-------------------------------------------------------------------
No Matching Entries
===================================================================
Queue Group Ports
===================================================================
Port                Sched Pol          Acctg Pol Stats    Description
-------------------------------------------------------------------
No Matching Entries
===================================================================
Queue Group Sap FC Maps
===================================================================
Sap Policy     FC Name           Queue (id type)
-------------------------------------------------------------------
No Matching Entries
===================================================================
Queue Group FP Maps
===================================================================
Card Num       Fp Num             Instance          Type
-------------------------------------------------------------------
5              1                  1                 Access
-------------------------------------------------------------------
Entries found: 1
-------------------------------------------------------------------
===================================================================
Queue Group Policer
===================================================================
Policer Id     : 1
Description    : (Not Specified)
PIR Adptn      : closest            CIR Adptn    : closest
Parent         : none               Level        : 1
Weight         : 1                  Adv. Cfg Plcy: none
Admin PIR      : max                Admin CIR    : 0
CBS            : def                MBS          : def
Hi Prio Only   : def                Pkt Offset   : 0
Profile Capped : Disabled
StatMode       : minimal
===================================================================
Policer Id     : 2
Description    : (Not Specified)
PIR Adptn      : closest            CIR Adptn    : closest
Parent         : none               Level        : 1
Weight         : 1                  Adv. Cfg Plcy: none
Admin PIR      : max                Admin CIR    : 0
CBS            : def                MBS          : def
Hi Prio Only   : def                Pkt Offset   : 0
Profile Capped : Disabled
StatMode       : minimal
===================================================================
*A:PE-1#
```

The following output shows the details related to the egress queue group template:

```
*A:PE-1# show qos queue-group "qg1" egress detail
===================================================================
QoS Queue-Group Egress
===================================================================
-------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------
Group-Name     : qg1
Description    : (Not Specified)
-------------------------------------------------------------------
Q CIR Admin  PIR Admin  CBS        HiPrio PIR Lvl/Wt Parent     BurstLimit(B)
```

```
   CIR Rule   PIR Rule   MBS                CIR Lvl/Wt Wred-Queue Slope
   Named-Buffer Pool     Pkt Bt Ofst        Adv Config Policy Name
  -------------------------------------------------------------------------
1 0          max        def        def      1/1        None       default
  closest    closest    def                 0/1        disabled   default
  (not-assigned)        add 0               (not-assigned)
2 0          max        def        def      1/1        None       default
  closest    closest    def                 0/1        disabled   default
  (not-assigned)        add 0               (not-assigned)
===========================================================================
Queue Group FC Mapping
===========================================================================
FC Name                             Queue-Id
---------------------------------------------------------------------------
No Matching Entries
===========================================================================
===========================================================================
Queue Group Ports (access)
===========================================================================
Port       Sched Pol         Acctg Pol Stats Description         QGrp-Instance
---------------------------------------------------------------------------
5/1/5                         0         No                        1
---------------------------------------------------------------------------
===========================================================================
Queue Group Ports (network)
===========================================================================
Port    Sched Pol    Policer-Ctrl-Pol  Acctg Pol Stats Description QGrp-Instance
---------------------------------------------------------------------------
No Matching Entries
===========================================================================
Qos Sap-Egress FC Group-Queue References
===========================================================================
Sap Policy     FC Name           Queue Id
---------------------------------------------------------------------------
No Matching Entries
===========================================================================
Qos Sap-Egress FC Port-Redirect-Group-Queue References
===========================================================================
Sap Policy     FC Name           Queue Id
---------------------------------------------------------------------------
10             af                1
10             ef                2
---------------------------------------------------------------------------
Entries found: 2
---------------------------------------------------------------------------
===========================================================================
Queue Group Policer
===========================================================================
No Matching Entries
---------------------------------------------------------------------------
HSMDA PIR Admin  Packet  WRR    MBS        Slope Plcy       WRR Plcy
Queue PIR Rule   Offset  Weight            Max Class        Burst Lmt
---------------------------------------------------------------------------
1     max        add 0   1      default    default          n/a
      closest                              8                default
2     max        add 0   1      default    default          n/a
      closest                              8                default
3     max        add 0   1      default    default          n/a
      closest                              8                default
4     max        add 0   1      default    default          n/a
      closest                              8                default
5     max        add 0   1      default    default          n/a
      closest                              8                default
6     max        add 0   1      default    default          n/a
```

```
         closest                              8              default
7        max        add 0   1     default   default         n/a
         closest                              8              default
8        max        add 0   1     default   default         n/a
         closest                              8              default
---------------------------------------------------------------------------
===========================================================================
*A:PE-1#
```

The preceding output shows that the ingress template has an instance 1, which is applied to card 5 fp 1, and the egress template has an instance 1, which is applied to port 5/1/5.

The remapping of FC l2 to queue 3 in the policer-output-queues queue group template is as follows:

```
*A:PE-1# show qos queue-group "policer-output-queues" egress detail | match post-lines 12
 "Queue Group FC Mapping"
Queue Group FC Mapping
===========================================================================
FC Name                             Queue-Id
---------------------------------------------------------------------------
af                                  1
be                                  1
ef                                  2
h1                                  2
h2                                  2
l1                                  1
l2                                  3
nc                                  2
===========================================================================
*A:PE-1#
```

For SAP ingress QoS policy 10, the redirection is true for FC af using policer 1 and FC ef using policer 2 to an FP ingress queue group, as follows:

```
*A:PE-1# show qos sap-ingress 10 detail | match post-lines 5 expression "af Unicast|ef Unicast"
FC                  : af Unicast
---------------------------------------------------------------------------
Policer             : 1              Queue            : def
Queue-Group         : None
FP Queue-Group      : True
FC                  : ef Unicast
---------------------------------------------------------------------------
Policer             : 2              Queue            : def
Queue-Group         : None
FP Queue-Group      : True
```

In the following output for SAP egress QoS policy 10:

- The redirection is true for FC af using queue 1 in the queue group instance, to be specified when the policy is applied to a SAP.

- The redirection is true for FC ef using local policer 2 with its traffic exiting queue 2 in the queue group instance, to be specified when the policy is applied to a SAP.

- FC l2 is using local policer 1 with its traffic exiting using queue 3 in the policer-output-queues port access egress queue group.

```
*A:PE-1# show qos sap-egress 10 detail | match post-lines 6 expression "Queue-Group"
FC  Queue  Queue-Group                      InstanceId   SapBReDir  Plcr
---------------------------------------------------------------------------
be  1      n/a                              n/a          False      None
```

```
l2   3        policer-output-queues              1             False      1
af   1        n/a                                n/a           True       None
ef   2        n/a                                n/a           True       2
```

The QoS information, including the queue group instances being used for the IES service SAP, is shown as
follows:

```
*A:PE-1# show service id 1 sap 5/1/5:1 detail | match post-lines 4 QOS
QOS
-------------------------------------------------------------------------------
Ingress qos-policy : 10                         Egress qos-policy : 10
Ingress FP QGrp    : qg1                         Egress Port QGrp  : qg1
Ing FP QGrp Inst   : 1                           Egr Port QGrp Inst: 1
```

In the following output for the egress SAP, FC ef is using policer 2 and its traffic exits using queue 2 in
queue group qg1 instance 1, while FC l2 is using policer 1 and its traffic exits using queue 3 of the policer-
output-queues queue group instance 1:

```
*A:PE-1# show qos policer sap 5/1/5:1 egress detail | match post-lines 4 "Policer Info"
Policer Info (1->5/1/5:1->1), Slot 5
===============================================================================
Policer Name      : 1->5/1/5:1->1
Direction         : Egress             Fwding Plane      : 1
FC->[QGrp:Inst->]Q : l2->policer-output-queues:1->3
Policer Info (1->5/1/5:1->2), Slot 5
===============================================================================
Policer Name      : 1->5/1/5:1->2
Direction         : Egress             Fwding Plane      : 1
FC->[QGrp:Inst->]Q : ef->qg1:1->2
```

The associations of the port access egress queue group are shown as follows:

```
*A:PE-1# show port 5/1/5 queue-group "qg1" instance 1 egress access associations
===============================================================================
Ethernet port 5/1/5 Access Egress queue-group
===============================================================================
Queue-Group Name   : qg1
Instance-Id        : 1
-------------------------------------------------------------------------------
Subscriber-Host Queue-Group Associations
-------------------------------------------------------------------------------
No associations
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
SAP Based Queue-Group Association
-------------------------------------------------------------------------------
Service-Id              SAP                     Qos Policy-Id
-------------------------------------------------------------------------------
1 (IES)                 5/1/5:1                 10
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
Qos Policy Based Queue-Group Association
-------------------------------------------------------------------------------
FC Name   Qos Policy-Id     Local Policer-Id    Local Queue-Id     QGrp-Q
-------------------------------------------------------------------------------
No associations
-------------------------------------------------------------------------------
===============================================================================
===============================================================================
*A:PE-1#
```

After traffic is sent through the service, the FP ingress access queue group policer statistics are as follows:

```
*A:PE-1# show card 5 fp 1 ingress queue-group "qg1" instance 1 mode access statistics
===============================================================================
Card:5  Acc.QGrp: qg1  Instance: 1
===============================================================================
Group Name   : qg1
Description  : (Not Specified)
Pol Ctl Pol  : None                        Acct Pol      : None
Collect Stats : disabled
-------------------------------------------------------------------------------
Statistics
-------------------------------------------------------------------------------
                      Packets               Octets
Ing. Policer:  1  Grp: qg1 (Stats mode: minimal)
Off. All            : 1000                  128000
Dro. All            : 0                     0
For. All            : 1000                  128000
Ing. Policer:  2  Grp: qg1 (Stats mode: minimal)
Off. All            : 1000                  128000
Dro. All            : 0                     0
For. All            : 1000                  128000
===============================================================================
*A:PE-1#
```

The traffic sent through the port egress access queue group queues is as follows:

```
*A:PE-1# show port 5/1/5 queue-group "qg1" instance 1 egress access statistics
-------------------------------------------------------------------------------
Ethernet port 5/1/5 Access Egress queue-group
-------------------------------------------------------------------------------
                      Packets               Octets
Egress Queue: 1 Group: qg1 Instance: 1
In Profile forwarded  : 0                   0
In Profile dropped    : 0                   0
Out Profile forwarded : 1000                128000
Out Profile dropped   : 0                   0
Egress Queue: 2 Group: qg1 Instance: 1
In Profile forwarded  : 0                   0
In Profile dropped    : 0                   0
Out Profile forwarded : 1000                128000
Out Profile dropped   : 0                   0
-------------------------------------------------------------------------------
*A:PE-1#
```

Finally, the FC l2 traffic using the egress policer 1 is in queue 3, in the policer-output-queues queue group instance statistics, as follows:

```
*A:PE-1# show port 5/1/5 queue-group "policer-output-queues" instance 1 egress access
 statistics
-------------------------------------------------------------------------------
Ethernet port 5/1/5 Access Egress queue-group
-------------------------------------------------------------------------------
                      Packets               Octets
Egress Queue: 1 Group: policer-output-queues Instance: 1
In Profile forwarded  : 0                   0
In Profile dropped    : 0                   0
Out Profile forwarded : 0                   0
Out Profile dropped   : 0                   0
Egress Queue: 2 Group: policer-output-queues Instance: 1
In Profile forwarded  : 0                   0
In Profile dropped    : 0                   0
```

```
Out Profile forwarded : 0                      0
Out Profile dropped   : 0                      0
Egress Queue: 3 Group: policer-output-queues Instance: 1
In Profile forwarded  : 0                      0
In Profile dropped    : 0                      0
Out Profile forwarded : 1000                   128000
Out Profile dropped   : 0                      0
-------------------------------------------------------------------------
*A:PE-1#
```

The number of valid ingress packets received on a SAP, or subscribers on that SAP, in the sap-stats
output, are shown as follows. The received valid counter includes both the local SAP counters and the
counters from the related FP ingress queue group instance. This is useful to display SAP-level traffic
statistics when forwarding classes in a SAP ingress policy have been redirected to an ingress queue
group.

```
*A:PE-1# show service id 1 sap 5/1/5:1 sap-stats | match post-lines 6 "Forwarding Engine Stats"
Forwarding Engine Stats
Dropped               : 0                      0
Received Valid        : 4000                   512000
Off. HiPrio           : 0                      0
Off. LowPrio          : 2000                   256000
Off. Uncolor          : 0                      0
Off. Managed          : 0                      0
```

Traffic forwarded through FP ingress access, port ingress access, and port egress access queue groups
can be monitored, as follows:

```
monitor card <slot-number> fp <fp-number>  ingress {access|network} queue-group <queue-
group-name> instance <instance-id> [interval <seconds>][repeat<repeat>] policer <policer-id>
 [absolute | percent-rate [<reference-rate>]]

monitor port queue-group <queue-group-name> ingress <access> ingress-queue <ingress-queue-id>
 [interval <seconds>] [repeat <repeat>] [absolute|rate]

monitor port queue-group <queue-group-name> egress <access> [instance <instance-id>] [egress-
queue <egress-queue-id>] [interval <seconds>] [repeat <repeat>] [absolute|rate]
```

The summary of the queue groups applied to a port is shown as follows:

```
*A:PE-1# show port 5/1/5 queue-group summary
===============================================================================
Port queue-group summary
===============================================================================
Access-egress queue groups:
---------------------------
qg1
policer-output-queues
Total number of access-egress queue groups  : 2
Network-egress queue groups:
---------------------------
Total number of network-egress queue groups : 0
Access-ingress queue groups:
---------------------------
Total number of access-ingress queue groups : 0
===============================================================================
*A:PE-1#
```

The total usage of queue groups is shown as follows:

```
*A:PE-1# show qos queue-group summary
===============================================================================
Queue-group instances per card
===============================================================================
card       port-acc-ing  port-acc-egr  port-nw-egr   fp-acc-ing    fp-nw-ing
-------------------------------------------------------------------------------
1          0             0             0             0             0
2          0             0             0             0             0
3          0             0             0             0             0
4          0             0             0             0             0
5          0             3             0             1             0
-------------------------------------------------------------------------------
Total ingress QG templates per system :  4
Total egress  QG templates per system :  5
===============================================================================
*A:PE-1#
```

The preceding output includes the created ingress template plus the three system-created ingress templates (making four in total), and the created egress template plus the four system-created egress templates (making five in total). There is one applied FP access ingress queue group instance on card 5. There are three port access egress queue group instances (the applied queue group instance and two instances of the policer-output-queues queue group), one on each access port used for IES service interfaces.

Network Interface Configuration Example

A network interface is created using ingress queue group qg2 instance 2 and egress queue group qg2 instance 2. If the goal is to provide per-network interface QoS on a single port, each network interface would be configured on a separate VLAN.

First, the queue group templates are configured, as follows:

```
configure
    qos
        queue-group-templates
            ingress
                queue-group "qg2" create
                    policer 1 create
                    exit
                    policer 2 create
                    exit
                exit
            exit
            egress
                queue-group "qg2" create
                    queue 1 best-effort create
                    exit
                    queue 2 expedite create
                    exit
                    policer 1 create
                    exit
                    policer 2 create
                    exit
                exit
            exit
```

The ingress template is applied to card 5 fp 1 to create an FP network queue group instance, as follows:

```
configure
    card 5
        fp 1
            ingress
                network
                    queue-group "qg2" instance 2 create
                    exit
                exit
```

The egress template is applied to port 5/1/5 to create a port network queue group instance, as follows:

```
configure
    port 5/1/5
        ethernet
            mode hybrid
            encap-type dot1q
            network
                egress
                    queue-group "qg2" instance 2 create
                    exit
                exit
```

The network QoS policy is created to:

- Ingress

  Redirect FC af to policer 1 and FC ef to policer 2 in the FP ingress queue group. FC be and FC l2 continue to use their default mapping to the local network ingress queues 1 and 2, respectively.

- Egress

  Redirect FC af to queue 1 in the port egress queue group and FC ef to policer 2 in the port egress queue group, with its traffic exiting through queue 2 of the port egress queue group. FC l2 is redirected to policer 1 in the port egress queue group, with its traffic exiting through the default network egress queue mapped by FC l2; that is, queue 2. FC be continues to use the default network egress queue 1.

```
configure
    qos
        network 10 create
            ingress
                dot1p 0 fc be profile out
                dot1p 1 fc l2 profile out
                dot1p 2 fc af profile out
                dot1p 3 fc ef profile in
                fc af
                    fp-redirect-group policer 1
                exit
                fc ef
                    fp-redirect-group policer 2
                exit
            exit
            egress
                fc af
                    port-redirect-group queue 1
                exit
                fc ef
                    port-redirect-group policer 2 queue 2
                exit
                fc l2
                    port-redirect-group policer 1
```

```
            exit
        exit
    exit
```

The network interface is created on port 5/1/5:2 using the preceding network QoS policy with the ingress and egress being redirected to created queue group instances. A second interface (not shown) is used to forward the traffic to and from this network interface.

```
configure
    router Base
        interface "PE-1-int2-2"
            address 10.2.2.1/24
            port 5/1/5:2
            qos 10 egress-port-redirect-group "qg2" egress-instance 2
                   ingress-fp-redirect-group "qg2" ingress-instance 2
            no shutdown
        exit
```

The following output shows the details of the configuration and the QoS state after sending traffic through the network interface.

The traffic statistics are either counted in the network interface queues or in the queue group instance queues and policers, but not in both.

The queue group templates can be shown. The output shows the details related to the ingress queue group template:

```
*A:PE-1# show qos queue-group "qg2" ingress detail
===============================================================================
QoS Queue-Group Ingress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name    : qg2
Description    : (Not Specified)
-------------------------------------------------------------------------------
Q   Mode    CIR Admin  PIR Admin  CBS          HiPrio  PIR Lvl/Wt Parent
            CIR Rule   PIR Rule   MBS                  CIR Lvl/Wt BurstLimit(B)
            Named-Buffer Pool     Pkt Bt Ofst          Adv Config Policy Name
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
Queue Group Ports
===============================================================================
Port              Sched Pol          Acctg Pol Stats     Description
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
Queue Group Sap FC Maps
===============================================================================
Sap Policy    FC Name          Queue (id type)
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
Queue Group FP Maps
===============================================================================
Card Num      Fp Num            Instance          Type
-------------------------------------------------------------------------------
5             1                 2                 Network
-------------------------------------------------------------------------------
Entries found: 1
```

```
-------------------------------------------------------------------------
=========================================================================
Queue Group Policer
=========================================================================
Policer Id    : 1
Description   : (Not Specified)
PIR Adptn     : closest                      CIR Adptn    : closest
Parent        : none                         Level        : 1
Weight        : 1                            Adv. Cfg Plcy: none
Admin PIR     : max                          Admin CIR    : 0
CBS           : def                          MBS          : def
Hi Prio Only  : def                          Pkt Offset   : 0
Profile Capped : Disabled
StatMode      : minimal
=========================================================================
Policer Id    : 2
Description   : (Not Specified)
PIR Adptn     : closest                      CIR Adptn    : closest
Parent        : none                         Level        : 1
Weight        : 1                            Adv. Cfg Plcy: none
Admin PIR     : max                          Admin CIR    : 0
CBS           : def                          MBS          : def
Hi Prio Only  : def                          Pkt Offset   : 0
Profile Capped : Disabled
StatMode      : minimal
=========================================================================
*A:PE-1#
```

The following output shows the details related to the egress queue group template:

```
*A:PE-1# show qos queue-group "qg2" egress detail
===============================================================================
QoS Queue-Group Egress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name    : qg2
Description   : (Not Specified)
-------------------------------------------------------------------------------
Q CIR Admin  PIR Admin  CBS          HiPrio PIR Lvl/Wt Parent      BurstLimit(B)
  CIR Rule   PIR Rule   MBS                 CIR Lvl/Wt Wred-Queue Slope
  Named-Buffer Pool     Pkt Bt Ofst         Adv Config Policy Name
-------------------------------------------------------------------------------
1 0          max        def          def    1/1        None        default
  closest    closest    def                 0/1        disabled    default
  (not-assigned)        add 0               (not-assigned)
2 0          max        def          def    1/1        None        default
  closest    closest    def                 0/1        disabled    default
  (not-assigned)        add 0               (not-assigned)
===============================================================================
Queue Group FC Mapping
===============================================================================
FC Name                                 Queue-Id
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
===============================================================================
Queue Group Ports (access)
===============================================================================
Port      Sched Pol        Acctg Pol Stats Description        QGrp-Instance
-------------------------------------------------------------------------------
No Matching Entries
```

```
===============================================================================
Queue Group Ports (network)
===============================================================================
Port    Sched Pol    Policer-Ctrl-Pol  Acctg Pol Stats Description QGrp-Instance
-------------------------------------------------------------------------------
5/1/5                                             No               2
-------------------------------------------------------------------------------
===============================================================================
Qos Sap-Egress FC Group-Queue References
===============================================================================
Sap Policy     FC Name           Queue Id
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
Qos Sap-Egress FC Port-Redirect-Group-Queue References
===============================================================================
Sap Policy     FC Name           Queue Id
-------------------------------------------------------------------------------
No Matching Entries
===============================================================================
Queue Group Policer
===============================================================================
Policer Id     : 1
Description    : (Not Specified)
PIR Adptn      : closest               CIR Adptn    : closest
Parent         : none                  Level        : 1
Weight         : 1                     Adv. Cfg Plcy: none
Admin PIR      : max                   Admin CIR    : 0
CBS            : def                   MBS          : def
Hi Prio Only   : def                   Pkt Offset   : 0
Profile Capped : Disabled
StatMode       : minimal
===============================================================================
Policer Id     : 2
Description    : (Not Specified)
PIR Adptn      : closest               CIR Adptn    : closest
Parent         : none                  Level        : 1
Weight         : 1                     Adv. Cfg Plcy: none
Admin PIR      : max                   Admin CIR    : 0
CBS            : def                   MBS          : def
Hi Prio Only   : def                   Pkt Offset   : 0
Profile Capped : Disabled
StatMode       : minimal
-------------------------------------------------------------------------------
HSMDA PIR Admin Packet  WRR     MBS        Slope Plcy      WRR Plcy
Queue PIR Rule  Offset  Weight             Max Class       Burst Lmt
-------------------------------------------------------------------------------
1     max       add 0   1       default    default         n/a
      closest                              8               default
2     max       add 0   1       default    default         n/a
      closest                              8               default
3     max       add 0   1       default    default         n/a
      closest                              8               default
4     max       add 0   1       default    default         n/a
      closest                              8               default
5     max       add 0   1       default    default         n/a
      closest                              8               default
6     max       add 0   1       default    default         n/a
      closest                              8               default
7     max       add 0   1       default    default         n/a
      closest                              8               default
8     max       add 0   1       default    default         n/a
      closest                              8               default
-------------------------------------------------------------------------------
```

```
===============================================================================
*A:PE-1#
```

The preceding output shows that the ingress template has an instance 2, which is applied to card 5 fp 1, and the egress template has an instance 2, which is applied to port 5/1/5.

The details of network QoS policy 10 shows the redirection to the ingress and egress queue groups, as follows:

```
*A:PE-1# show qos network 10 detail | match expression "Egress Forwarding Class Mapping|Ingress
 Forwarding Class Mapping|FC Value|Redirect"
Egress Forwarding Class Mapping
FC Value        : 0                      FC Name          : be
Redirect Grp Q  :   None                 Redirect Grp Plcr:  None
FC Value        : 1                      FC Name          : l2
Redirect Grp Q  :   None                 Redirect Grp Plcr:  1
FC Value        : 2                      FC Name          : af
Redirect Grp Q  :   1                    Redirect Grp Plcr:  None
FC Value        : 3                      FC Name          : l1
Redirect Grp Q  :   None                 Redirect Grp Plcr:  None
FC Value        : 4                      FC Name          : h2
Redirect Grp Q  :   None                 Redirect Grp Plcr:  None
FC Value        : 5                      FC Name          : ef
Redirect Grp Q  :   2                    Redirect Grp Plcr:  2
FC Value        : 6                      FC Name          : h1
Redirect Grp Q  :   None                 Redirect Grp Plcr:  None
FC Value        : 7                      FC Name          : nc
Redirect Grp Q  :   None                 Redirect Grp Plcr:  None
Ingress Forwarding Class Mapping
FC Value              : 0                FC Name              : be
Redirect UniCast Plcr  : None            Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 1                FC Name              : l2
Redirect UniCast Plcr  : None            Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 2                FC Name              : af
Redirect UniCast Plcr  : 1               Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 3                FC Name              : l1
Redirect UniCast Plcr  : None            Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 4                FC Name              : h2
Redirect UniCast Plcr  : None            Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 5                FC Name              : ef
Redirect UniCast Plcr  : 2               Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 6                FC Name              : h1
Redirect UniCast Plcr  : None            Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
FC Value              : 7                FC Name              : nc
Redirect UniCast Plcr  : None            Redirect MultiCast Plcr : None
Redirect BroadCast Plcr : None           Redirect Unknown Plcr   : None
*A:PE-1#
```

The preceding output shows that:

- Ingress

    - FC af is redirected to unicast policer 1 in the queue group instance, to be specified when the policy is applied to a network interface.

- FC ef is redirected to unicast policer 2 in the queue group instance, to be specified when the policy is applied to a network interface.

- Egress

    - FC af is redirected to queue 1 in the queue group instance, to be specified when the policy is applied to a network interface.

    - FC ef is redirected to policer 2 and queue 2 in the queue group instance, to be specified when the policy is applied to a network interface, so that the policer 2 traffic exits using queue 2 in the port network egress queue group instance.

The queue group instances used by the network interface are shown as follows:

```
*A:PE-1# show router interface "PE-1-int2-2" detail | match post-lines 3 "QoS Queue-Group
 Redirection Details"
QoS Queue-Group Redirection Details
-------------------------------------------------------------------------------
Ingress FP QGrp  : qg2                   Egress Port QGrp  : qg2
Ing FP QGrp Inst : 2                     Egr Port QGrp Inst: 2
```

After traffic is sent through the network, it can be shown in the FP ingress network queue group policers, as follows:

```
*A:PE-1# show card 5 fp 1 ingress queue-group "qg2" instance 2 mode network statistics
===============================================================================
Card:5  Net.QGrp: qg2  Instance: 2
===============================================================================
Group Name     : qg2
Description    : (Not Specified)
Pol Ctl Pol    : None                    Acct Pol      : None
Collect Stats : disabled
-------------------------------------------------------------------------------
Statistics
-------------------------------------------------------------------------------
                    Packets               Octets
Ing. Policer:  1  Grp: qg2 (Stats mode: minimal)
Off. All           : 1000                 128000
Dro. All           : 0                    0
For. All           : 1000                 128000
Ing. Policer:  2  Grp: qg2 (Stats mode: minimal)
Off. All           : 1000                 128000
Dro. All           : 0                    0
For. All           : 1000                 128000
===============================================================================
*A:PE-1#
```

The traffic sent through the port egress network queue group queues can also be shown, as follows:

```
*A:PE-1# show port 5/1/5 queue-group "qg2" instance 2 egress network statistics
-------------------------------------------------------------------------------
Ethernet port 5/1/5 Network Egress queue-group
-------------------------------------------------------------------------------
                    Packets               Octets
Egress Queue:  1    Group: qg2    Instance-Id:  2
In Profile forwarded  : 0                  0
In Profile dropped    : 0                  0
Out Profile forwarded : 1000               128000
Out Profile dropped   : 0                  0
Egress Queue:  2    Group: qg2    Instance-Id:  2
In Profile forwarded  : 0                  0
```

```
In Profile dropped   : 0                       0
Out Profile forwarded : 1000                    128000
Out Profile dropped   : 0                       0
Egress Policer:  1  Group: qg2  Instance-Id: 2
Stats mode: minimal
Off. All             : 1000                     128000
Dro. All             : 0                        0
For. All             : 1000                     128000
Egress Policer:  2  Group: qg2  Instance-Id: 2
Stats mode: minimal
Off. All             : 1000                     128000
Dro. All             : 0                        0
For. All             : 1000                     128000
------------------------------------------------------------------------
*A:PE-1#
```

In the preceding output, the traffic on queue 1 is FC af, on policer 1 is FC l2, and on policer 2 is FC ef, with the post-policed traffic on queue 2.

Finally, the FC l2 traffic using the network egress queue group policer 1 can be shown in queue 2 of the default network egress queues, as follows:

```
*A:PE-1# show port 5/1/5 detail | match post-lines 4 "Egress Queue  2"
Egress Queue  2                 Packets                 Octets
    In Profile forwarded  :    0                        0
    In Profile dropped    :    0                        0
    Out Profile forwarded :    1000                     128000
    Out Profile dropped   :    0                        0
```

Traffic forwarded through both FP ingress network and port egress network queue groups can be monitored, as follows:

```
monitor card <slot-number> fp <fp-number>  ingress {access|network} queue-group <queue-
group-name> instance <instance-id> [interval <seconds>][repeat<repeat>] policer <policer-id>
 [absolute | percent-rate [<reference-rate>]]

monitor port queue-group <queue-group-name> egress <access> [instance <instance-id>] [egress-
queue <egress-queue-id>] [interval <seconds>] [repeat <repeat>] [absolute|rate]
```

The summary of the queue groups applied to a port is shown as follows:

```
*A:PE-1# show port 5/1/5 queue-group summary

===============================================================================
Port queue-group summary
===============================================================================
Access-egress queue groups:
---------------------------
policer-output-queues
Total number of access-egress queue groups  : 1

Network-egress queue groups:
---------------------------
qg2
Total number of network-egress queue groups : 1

Access-ingress queue groups:
---------------------------
Total number of access-ingress queue groups : 0
===============================================================================
*A:PE-1#
```

The total usage of queue groups is shown as follows:

```
*A:PE-1# show qos queue-group summary


===============================================================================
Queue-group instances per card
===============================================================================
card       port-acc-ing  port-acc-egr  port-nw-egr   fp-acc-ing   fp-nw-ing
-------------------------------------------------------------------------------
1          0             0             0             0            0
2          0             0             0             0            0
3          0             0             0             0            0
4          0             0             0             0            0
5          0             2             1             0            1
-------------------------------------------------------------------------------
Total ingress QG templates per system :  4
Total egress  QG templates per system :  5
===============================================================================
*A:PE-1#
```

The preceding output includes the created ingress template plus the three system-created ingress templates (making four in total), and the created egress template plus the four system-created egress templates (making five in total). There is the one applied FP network ingress queue group instance on card 5. There is the one created port network egress queue group instance. There are also two port access egress queue group instances, which are the policer-output-queues queue group instances associated with the access port used for IES service interface (not discussed) and the access side of the hybrid port 5/1/5.

Egress Policed Subscriber Configuration Example

Queue groups are only applicable to subscribers for egress policed traffic. By default, subscriber egress policed traffic exits the port using a queue in the egress access policer-output-queues queue group instance. The queue used is determined by the FC mapping in the policer-output-queues queue group template. This is the same default operation as in the SAP example.

The subscriber policed traffic can be sent to a different queue group instance using the inter-dest-id and a host-match (described when applying an egress access queue group template to a port), which represent an intermediate destination, such as a downstream DSLAM or GPON OLT. The inter-dest-id can be associated with a subscriber host when it is created; this would usually be received from the DHCP, or RADIUS server, or Diameter Gx (Policy and Rule Charging Function), or the local user database, or configured under a static host.

An alternative to host matching on an inter-dest-id is to match on the top VLAN tag when a QinQ SAP is configured using a default inter-dest-id.

A default inter-dest-id can be configured in IES, VPRN, and VPLS services, and under an MSAP policy, as follows:

```
configure
    service
        {ies|vprn} <service-id>
            subscriber-interface <ip-int-name>
                group-interface <ip-int-name> [create]
                    sap <sap-id>
                        sub-sla-mgmt
                            def-inter-dest-id {<inter-dest-string>|use-top-q}
```

```
configure
```

```
    service
        vpls <service-id>
            sap <sap-id>
                sub-sla-mgmt
                    def-inter-dest-id {<inter-dest-string>|use-top-q}
```

```
configure
    subscriber-mgmt
        msap-policy <msap-policy-name> [create]
                sub-sla-mgmt
                    def-inter-dest-id {<inter-dest-string>|use-top-q}
```

The egress queue group template is configured as follows:

```
configure
    qos
        queue-group-templates
            egress
                queue-group "qg3" create
                    queue 1 best-effort create
                    exit
                    queue 2 expedite create
                    exit
                    fc af create
                        queue 1
                    exit
                    fc ef create
                        queue 2
                    exit
                exit
            exit
```

The egress template is applied to port 5/1/5 to create a port access queue group instance. A host match is configured under the created queue group instance on an egress access port. In the following, the host match dslam-1 is used:

```
configure
    port 5/1/5
        ethernet
            mode hybrid
            encap-type dot1q
            access
                egress
                    queue-group "qg3" instance 1 create
                        host-match dest "dslam-1" create
                    exit
                exit
```

A host-match can only be configured under instance 1 of a port access egress queue group; if its configuration is attempted on a different instance, the following error is displayed:

```
*A:PE-1# configure port 5/1/5 ethernet access egress queue-group "qg3" instance 3 create host-
match dest another-dslam create
MINOR: PMGR #1337 Host match entries only supported on port access egress queue groups with
 system default instance 1
*A:PE-1#
```

The subscriber host uses a SAP egress QoS policy in an egress SLA profile to map FCs to egress queue and policers. The SAP egress QoS policy is created with FC af using policer 1 and FC ef using policer 2, as follows:

```
configure
    qos
        sap-egress 30 create
            queue 1 create
            exit
            queue 2 create
            exit
            policer 1 create
            exit
            policer 2 create
            exit
            fc af create
                policer 1
            exit
            fc be create
                queue 1
            exit
            fc ef create
                policer 2
            exit
            fc l2 create
                queue 2
            exit
        exit
```

To redirect the subscriber egress policed traffic to the access egress queue group qg3 instance 1 on port 5/1/5, which is configured with the host-match, an inter-dest-id is configured for the created subscriber static host, as follows:

```
configure
    service
        vprn 3 customer 1 create
            route-distinguisher 65536:200
            subscriber-interface "sub-int-1" create
                address 10.3.2.1/24
                group-interface "group-int-1" create
                    arp-populate
                    sap 5/1/5:3 create
                        sub-sla-mgmt
                            def-sub-profile "basic-sub"
                            def-sla-profile "basic-sla"
                            multi-sub-sap 200
                            single-sub-parameters
                                profiled-traffic-only
                            exit
                            no shutdown
                        exit
                        static-host ip 10.3.2.2 mac 00:00:10:03:02:02 create
                            inter-dest-id "dslam-1"
                            sla-profile "basic-sla"
                            sub-profile "basic-sub"
                            subscriber "sub1"
                            no shutdown
                        exit
                    exit
```

The host match configured under the egress queue group instance is shown as follows:

```
A:PE-1# show port 5/1/5 queue-group "qg3" instance 1 access | match post-lines 3 Host-Matches
Host-Matches
-------------------------------------------------------------------------------
Dest: dslam-1
-------------------------------------------------------------------------------
```

When the subscriber host is created, the inter-dest-id for the subscriber host is shown as follows:

```
A:PE-1# show service active-subscribers subscriber "sub1" detail | match expression "Subscriber
 sub1 |Sub. Int Dest Id"
Subscriber sub1 (basic-sub)
Sub. Int Dest Id : "dslam-1"
```

The inter-dest-id dslam-1 is matched against the host-match destination configured on the access egress group instances on the port on which the host is being created. If a match is found, the subscriber egress policed traffic will use that egress queue group instance, with the actual queue used being selected by the FC-to-queue mapping in the related queue group template. Otherwise, the default policer-output-queues queue group instance will be used.

The egress queue group instance subscriber host associations are shown as follows:

```
A:PE-1# show port 5/1/5 queue-group egress "qg3" associations | match post-lines 6 Subscriber-
Host
Subscriber-Host Queue-Group Associations
-------------------------------------------------------------------------------
svc-id : 3 (VPRN)
 sap   : 5/1/5:3
 subscr: sub1
 ip    : 10.3.2.2
 mac   : 00:00:10:03:02:02  pppoe-sid: N/A
```

The following output shows that FC ef is using policer 2 and its traffic exits using queue 2 in queue group qg3, while FC af is using policer 1 and its traffic exits using queue 1 in queue group qg3.

```
A:PE-1# show qos policer subscriber "sub1" egress detail | match post-lines 4 "Policer Info"
Policer Info (Sub=1:1 3->5/1/5:3->2), Slot 5
===============================================================================
Policer Name       : Sub=1:1 3->5/1/5:3->2
Direction          : Egress             Fwding Plane      : 1
FC->[QGrp:Inst->]Q : ef->qg3->2
Policer Info (Sub=1:1 3->5/1/5:3->1), Slot 5
===============================================================================
Policer Name       : Sub=1:1 3->5/1/5:3->1
Direction          : Egress             Fwding Plane      : 1
FC->[QGrp:Inst->]Q : af->qg3->1
```

After traffic is sent through the subscriber, the egress policed traffic can be shown in the port egress access queue group instance, as follows:

```
A:PE-1# show port 5/1/5 queue-group "qg3" instance 1 access egress statistics

-------------------------------------------------------------------------------
Ethernet port 5/1/5 Access Egress queue-group
-------------------------------------------------------------------------------
                     Packets               Octets

Egress Queue: 1 Group: qg3 Instance: 1
```

```
In Profile forwarded  : 0                       0
In Profile dropped    : 0                       0
Out Profile forwarded : 1000                    128000
Out Profile dropped   : 0                       0
Egress Queue: 2 Group: qg3 Instance: 1
In Profile forwarded  : 0                       0
In Profile dropped    : 0                       0
Out Profile forwarded : 1000                    128000
Out Profile dropped   : 0                       0
-------------------------------------------------------------------------------
A:PE-1#
```

The traffic statistics are either counted in the subscriber queues or policers, or in the queue group instance queues, but not in both. However, summary statistics per SAP are available when using FP ingress queue groups.

The number of valid ingress packets received on a SAP, or subscribers on that SAP, can be shown in the sap-stats output, as follows. The received valid counter includes both the local SAP counters and the counters from the related FP ingress queue group instance. This is useful to display SAP-level traffic statistics when forwarding classes in a SAP ingress policy have been redirected to an ingress queue group.

```
*A:PE-1# show service id 3 sap 5/1/5:3 sap-stats | match post-lines 6 "Forwarding Engine Stats"
Forwarding Engine Stats
Dropped             : 0                       0
Received Valid      : 4000                    512000
Off. HiPrio         : 0                       0
Off. LowPrio        : 0                       0
Off. Uncolor        : 0                       0
Off. Managed        : 0                       0
```

Traffic forwarded through port egress access queue groups can be monitored, as follows:

```
monitor port queue-group <queue-group-name> egress <access> [instance <instance-id>] [egress-
queue <egress-queue-id>] [interval <seconds>] [repeat <repeat>] [absolute|rate]
```

The summary of the queue groups applied to a port is shown as follows:

```
A:PE-1# show port 5/1/5 queue-group summary

===============================================================================
Port queue-group summary
===============================================================================
Access-egress queue groups:
---------------------------
qg3
policer-output-queues
Total number of access-egress queue groups  : 2

Network-egress queue groups:
---------------------------
Total number of network-egress queue groups : 0

Access-ingress queue groups:
---------------------------
Total number of access-ingress queue groups : 0
===============================================================================
A:PE-1#
```

The total usage of queue groups is shown as follows:

```
A:PE-1# show qos queue-group summary


===============================================================================
Queue-group instances per card
===============================================================================
card      port-acc-ing  port-acc-egr  port-nw-egr   fp-acc-ing    fp-nw-ing
-------------------------------------------------------------------------------
1         0             0             0             0             0
2         0             0             0             0             0
3         0             0             0             0             0
4         0             0             0             0             0
5         0             3             0             0             0
-------------------------------------------------------------------------------
Total ingress QG templates per system :  3
Total egress  QG templates per system :  5
===============================================================================
A:PE-1#
```

The preceding output includes the three system-created ingress templates, and the created egress
template plus the four system-created egress templates (making five in total). There are three port access
egress queue group instances (the applied queue group instance and two instances of the policer-output-
queues queue group), one on each access port used for VPRN service interfaces.

Conclusion

This chapter described the use of queue groups as a mechanism to provide an aggregate QoS control
for multiple SAPs and per-network interface QoS control. The configuration steps and commands are
described, followed by example configurations on a SAP, network interface, and for egress policed traffic
subscriber traffic.

High Scale QoS IOM: QoS, Service, and Network Configuration

This chapter provides information about High Scale QoS IOM: QoS, Service, and Network Configuration.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This chapter is applicable to the 7750 SR-7/12/12e platforms and describes the High Scale QoS (HSQ) IOM. The configuration was tested on Release 15.0.R5.

Overview

This chapter describes the QoS operation and configuration of the HSQ IOM, with a focus on services and network interfaces. For the subscriber management configuration, see chapters *High Scale QoS IOM in ESM Context: Single SLA Mode* and *High Scale QoS IOM in ESM Context: Expanded SLA Mode* in the *7450 ESS, 7750 SR, and VSR Triple Play Service Delivery Architecture Advanced Configuration Guide for Classic CLI*.

The HSQ IOM is an FP3-based IOM that has a multicore CPU and accepts up to two MDA-e cards. The HSQ IOM supports an enhanced egress QoS architecture to provide scalable network, service, and subscriber QoS. At ingress, the HSQ IOM supports regular FP3 QoS with a high ingress policer scaling. This chapter focuses on the HSQ IOM egress QoS.

The HSQ IOM supports six scheduling classes across multiple hierarchical levels of hardware egress shaping with very stringent egress burst control. The scheduling allows a mix of strict priority and weighted round-robin (WRR). A flexible buffer pool structure permits both buffer isolation and buffer oversubscription for the queue buffer allocation.

The HSQ IOM supports 768k queues, which are grouped into 96k queue groups; each comprises eight queues (referred to as HSQ queue groups). HSQ queue groups are used for SAP egress queues, network egress queues, and both access and network egress queue group instance queues.

The SAP egress, network egress, and access and network egress queue group related commands that are not supported with an HSQ IOM are provided in the associated configuration chapters following. In addition, the following are not applicable to the HSQ:

- QoS related

    - Egress access and network MDA and port pools

      – All HSMDA commands

      – All VPORT related commands

      – PBB egress B-SAP per ISID shaping

      – Port **hybrid-buffer-allocation egr-weight**

- MPLS related

      – Generalized Multiprotocol Label Switching (GMPLS) UNI

- Service related

      – G.8031 protected Ethernet tunnels

- System related

      – Port cross-connects (PXC)

      – Ethernet satellite host ports

      – Soft reset

The operation of the HSQ IOM is described in the following sections:

- Shaping

- Scheduling

- Buffer Management

- LAGs

Shaping

The HSQ egress shaping uses the following objects:

- HSQ queue groups

  An HSQ queue group comprises eight egress queues with two WRR groups. One HS queue group is allocated to each of the following:

      – An egress SAP

      – An egress network port

      – An egress access queue group instance

      – An egress network queue group instance

      – A subscriber egress (single SLA profile instance in single HS SLA mode). Enhanced Subscriber Management (ESM) is beyond the scope of this chapter.

- Primary shapers

  In the context of this chapter, a primary shaper is allocated for each secondary shaper because it is required in the hierarchy, but it does not perform any QoS control. Primary shapers are also allocated for each subscriber egress configured with multiple SLA profile instances in extended HS SLA mode, however, ESM is beyond the scope of this chapter.

- Secondary shapers

  Secondary shapers provide an abstraction to be used for QoS control of traffic to a downstream device such as an access node. Shaping can be performed on the entire traffic or on each scheduling class within the secondary shaper.

- Ports

  The traffic forwarded to each port can be shaped. In addition, traffic in each scheduling class within a port can be shaped individually or within a single WRR group.

Six scheduling classes are supported across all the preceding objects.

The egress QoS scheduling hierarchy is shown in Figure 10: Egress HSQ IOM Scheduling Hierarchy.

Figure 10: Egress HSQ IOM Scheduling Hierarchy
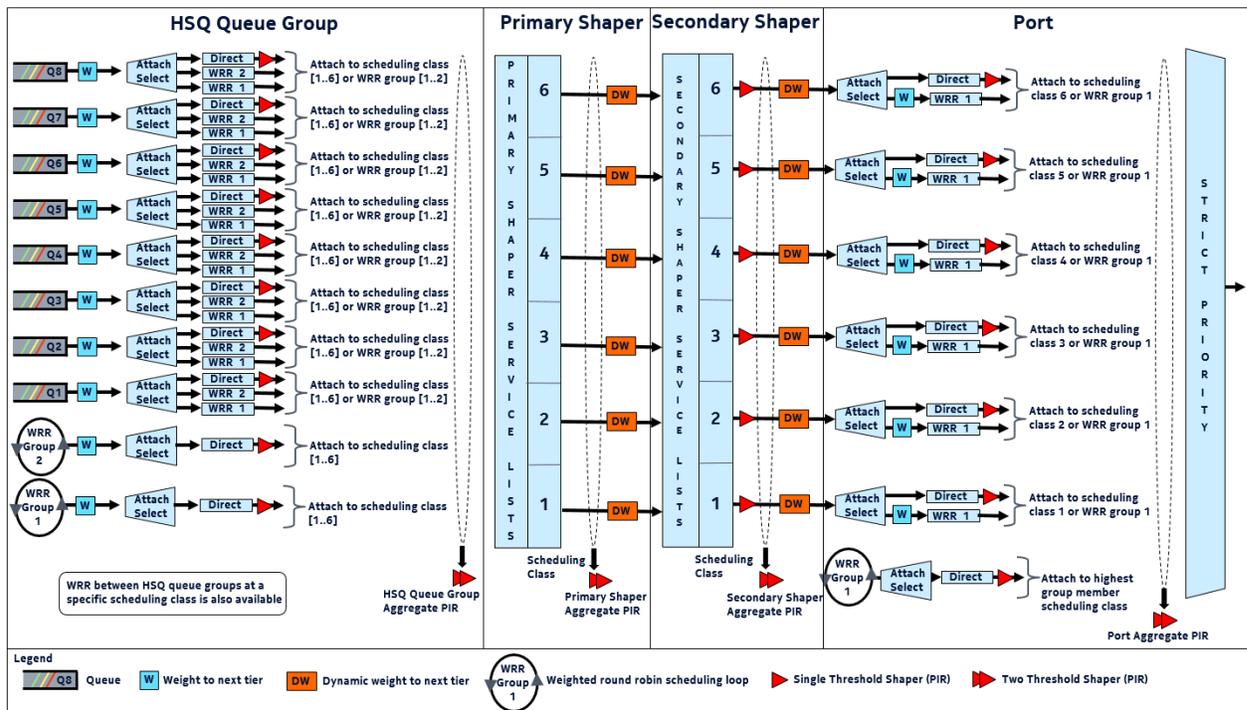
The available egress shaping is described in detail, as follows:

- Per-queue or per-WRR group of queues
- Per-HSQ queue group aggregate
- Per-primary shaper aggregate
- Per-secondary shaper aggregate
- Per-secondary shaper per scheduling class
- Per-port aggregate
- Per-port per scheduling class

Per-Queue or Per-WRR Group of Queues

Each queue can be independently shaped by configuring its PIR and attaching it to a primary shaper scheduling class. Alternatively, it can be shaped together with other queues in the same HSQ queue group as part of a WRR group. The WRR group can have a configured rate, and also needs to be attached to a primary shaper scheduling class.

There are eight queues and two WRR groups available within an HSQ queue group, which attach to the six primary shaper scheduling classes. Only one object (queue or WRR group) per HSQ queue group can attach to a scheduling class at any time, so to make use of all queues in an HSQ queue group, at least three queues must be attached to a WRR group. Queues and WRR groups can remain unattached from a scheduling class, in which case the related queues discard all received packets.

The queue PIR is configured under the queue within a SAP egress QoS policy for services, in a network queue policy for network interfaces, or in an egress queue group template for both access and network egress queue group instances. The queue CIR is ignored when the policy is applied to an HSQ IOM. The per-WRR group PIR is configured within the same policies under the **hs-wrr-group** context. Queue and WRR group PIR use packet-based accounting (L2 rate), which can be adjusted using the queue **packet-byte-offset** parameter for SAP egress and egress queue group instances.

The attachment of a queue or WRR group to a scheduling class is configured within an **hs-attachment-policy**. A default **hs-attachment-policy** (which is not configurable) is created by the system and is applied to all SAP egress QoS policies, network queue policies, and egress queue group templates. The default policy has queues 1 to 3 attached to WRR group 1, which is attached to scheduling class 1, and queues 4 to 8 attached directly to scheduling classes 2 to 6.

When creating a new **hs-attachment-policy**, the following rules apply to the queue and WRR attachment:

- A queue must be attached to a scheduling class, or a WRR group, which is also attached to a scheduling class, so as to forward packets.

- Only one queue or WRR group can be attached to a scheduling class per HSQ queue group.

- Queues can only be attached to scheduling classes in an ascending order; for example, if queue 2 is attached to scheduling class 2, then queue 1 cannot attach to scheduling classes 3 to 6.

- The queue identifiers must be contiguous when attaching queues to a WRR group.

- Queues attached to WRR group 1 must have lower queue identifiers than those attached to WRR group 2.

- The maximum number of queues attached to a WRR group is six: six to group 1 or six to group 2, or six to a combination of groups 1 and 2.

- WRR group 2 can only be attached to a scheduling class after WRR group 1 has at least one attached queue and has been attached to a scheduling class.

- WRR group 2 must be attached to a higher scheduling class than WRR group 1.

Per-HSQ Queue Group Aggregate

A per-HSQ queue group aggregate shapes traffic forwarded by all the queues in its associated HSQ queue group to an aggregate rate. This is applicable to SAP egress queues, and to both access and network egress queue group instances. It is not applicable to network egress queues.

The per-HSQ queue group aggregate PIR is configurable as an egress aggregate rate limit applied under a SAP or a port access or network egress queue group instance. The HSQ queue group PIR uses packet-based accounting (L2 rate), which can be adjusted using the queue **packet-byte-offset** parameter for SAP egress and egress queue group instances.

When using HSQ queue groups with access or network egress queue group instances on 100G ports, the **hs-turbo** parameter can be configured under the port queue group instance to allow the corresponding HSQ queue group queues to achieve a higher throughput. The **hs-turbo** parameter is not applicable to 10G ports and so is ignored when configured under a queue group instance on a 10G port.

Per-Primary Shaper Aggregate

A primary shaper aggregate shapes the traffic forwarded by all of the HSQ queue groups connected to the primary shaper to an aggregate rate.

User-configured primary shapers are not applicable to SAP egress HSQ queue groups, network egress HSQ queue groups, or both access and network egress queue group instance HSQ queue groups. However, the hierarchy shown in Figure 10: Egress HSQ IOM Scheduling Hierarchy is always conformed to, so by default these HSQ queue groups always connect to a system-created per-port default primary shaper that has its aggregate PIR rate set to the maximum rate, so as not to constrain the traffic rate at this level.

The system also instantiates a primary shaper, again with its aggregate PIR set to the maximum rate, when the first egress SAP or pseudowire SAP (PW-SAP) is associated with a secondary shaper. This primary shaper is then used by all HSQ queue groups associated with that secondary shaper. User-configured primary shaper aggregates are applicable to ESM, which is beyond the scope of this chapter.

Per-Secondary Shaper Aggregate

Secondary shapers are aimed at providing QoS control for traffic forwarded to a specific downstream device, such as an access node.

A secondary shaper aggregate shapes the traffic forwarded by all of its connected primary shapers (and HSQ queue groups). Secondary shapers are applicable to SAP egress queues, but not to network egress or to both access and network egress queue group instance HSQ queue groups. The hierarchy shown in Figure 10: Egress HSQ IOM Scheduling Hierarchy is always conformed to, so by default all primary shapers (and their HSQ queue groups) always connect to a system-created per-port default secondary shaper that has its aggregate PIR rate set to the maximum rate, so as not to constrain the traffic rate at this level. Secondary shaper aggregates are also applicable to ESM, which is beyond the scope of this chapter.

Multiple HS secondary shapers can be created under the **config>port>ethernet>egress** context using the **hs-secondary-shaper** statement. The HS secondary shaper aggregate PIR is configured under the associated secondary shaper. A default hs-secondary-shaper is applied under each HSQ egress port with an aggregate PIR rate **max**, which can be configured if required. The secondary shaper PIR uses frame-based accounting (L1 rate) and is not affected by a queue **packet-byte-offset** parameter.

SAP egress HSQ queue groups are connected to an HS secondary shaper using the **hs-secondary-shaper** parameter under a queue override, which is configured under the SAP egress context. When the first egress SAP or PW-SAP is associated with a user-configured HS secondary shaper, the system instantiates a default primary shaper for that secondary shaper.

Per-Secondary Shaper per Scheduling Class

Each of the six scheduling classes can be individually shaped within an HS secondary shaper. The HS secondary shaper scheduling class PIR is configured under the associated secondary shaper. The default HS secondary shaper scheduling class PIRs are set to **max** and can also be modified. The secondary shaper scheduling class PIR uses frame-based accounting (L1 rate) and is not affected by a queue **packet-byte-offset** parameter.

Per-Port Aggregate

A per-port aggregate shapes the traffic forwarded by its connected secondary shapers, that is, all the traffic egressing out of the physical port. It is applicable to SAP egress, network egress, and both access and network egress queue group instance traffic. A default HS scheduler policy (which is not configurable) is applied to all ports.

A user-defined HS scheduler policy can be created in which the port aggregate PIR (**max-rate**) can be configured and the policy then applied under the **config>port>ethernet>egress** context. Only a single HS scheduler policy is supported on each port. The port aggregate PIR uses frame-based accounting (L1 rate) and is not affected by a queue **packet-byte-offset** parameter.

An alternative to configuring a per-port aggregate is to configure an **egress-rate** on the port. This provides more granular control as it is configured in kb/s (whereas the per-port aggregate is in Mb/s). The HSQ **egress-rate** is based on the Ethernet size of the packet including the IFG (Inter-Frame-Gap) and preamble.

Per-Port per Scheduling Class

Each of the six scheduling classes can also be individually shaped per port by configuring a scheduling class PIR within an HS scheduler policy. The scheduling classes can also be grouped in a single WRR group at each egress port with each class being assigned a weight within the group.

The scheduling class identifiers must be contiguous within the WRR group and the group is scheduled at the scheduling class of its highest member scheduling class. Both the scheduling class PIR and the WRR group PIR are set to **max** in the default HS scheduler policy, with the WRR group being unused. The port scheduling class PIR uses frame-based accounting (L1 rate) and is not affected by a queue **packet-byte-offset** parameter.

Scheduling

The scheduling allows a mix of strict priority and WRR. There are six scheduling classes, which are implemented from the HSQ queue group queues through the primary shaper, secondary shaper, and port. The scheduling classes are serviced in a strict priority order (scheduling class 6 having the highest priority and scheduling class 1 having the lowest priority), with WRR groups at the HSQ queue group and port levels, and a dynamic weight at the primary and secondary shaper levels.

Packet forwarding is achieved using service lists; the objects at each level are on a service list at that level if they are in a state ready to send packets, or are off the service list if they have exceeded their configured PIR together with its related burst. When a port has a scheduling opportunity, it selects the secondary shaper to be serviced next, which selects the primary shaper to be serviced next, which selects the HSQ queue group to be serviced next, which selects a queue to be serviced next, resulting in a packet from that queue being forwarded.

At the HSQ queue group level, queues can be attached to one of two WRR groups, each of which is scheduled at a single scheduling class with packets being taken from the constituent queues based on a configured queue weighting. The weight is configured using the **hs-wrr-weight** under the **queue** statement within a SAP egress QoS policy, a network queue policy, or in an egress queue group template.

Weighting is also supported between queues and WRR groups in different HSQ queue groups per-primary shaper scheduling class. This allows the capacity available at the primary shaper scheduling class to be shared in a WRR manner between the HSQ queue group queues and WRR groups attached to that

scheduling class. This is configured within a SAP egress QoS policy, network queue policy, and egress queue group template, using the **hs-class-weight** parameter under the respective **queue** or **hs-wrr-group** statement.

This weighting should not be confused with the **hs-wrr-weight** parameter, which specifies the relative weights of different queues within the same HSQ queue group WRR group. This **hs-class-weight** parameter could be used to give unequal shares of the available capacity to different types of service offerings.

There is a single WRR group at the port level that allows multiple scheduling classes to be collapsed to a single class per port with each class in the group being assigned a weight. The weight of each scheduling class in the group is configured within the applied HS scheduler policy.

The dynamic weights at the primary and secondary shapers are managed by the system, based on the number of pending packets for each of the shapers, not on the number of attached objects in each. The more pending packets a shaper has, the higher the weight it gets. The goal is to ensure a balanced distribution of capacity between each of the primary shapers and each of the secondary shapers. For example, this allows a secondary shaper with 10 000 active HSQ queue groups to receive proportionately more scheduling opportunities than another secondary shaper with only 100 active HSQ queue groups.

The HSQ queue group and secondary shaper aggregate rates are implemented as a set of token buckets to control the aggregate rates. As packets are transmitted from each, the scheduler updates its bucket states based on the number of bytes forwarded. Two thresholds are used within each bucket to provide more granular control over this scheduling behavior: a low burst limit threshold and a high burst limit threshold.

These thresholds control when their respective queues are removed from the scheduler list, thereby allowing the queues using the high threshold to continue to forward packets even after the queues using the low threshold are no longer being serviced. This is shown in Figure 11: HSQ Queue Group and Secondary Shaper Aggregate Scheduler Bucket. The **low-burst-max-class** parameter defines which queues use each of the thresholds, and is described following.

Figure 11: HSQ Queue Group and Secondary Shaper Aggregate Scheduler Bucket
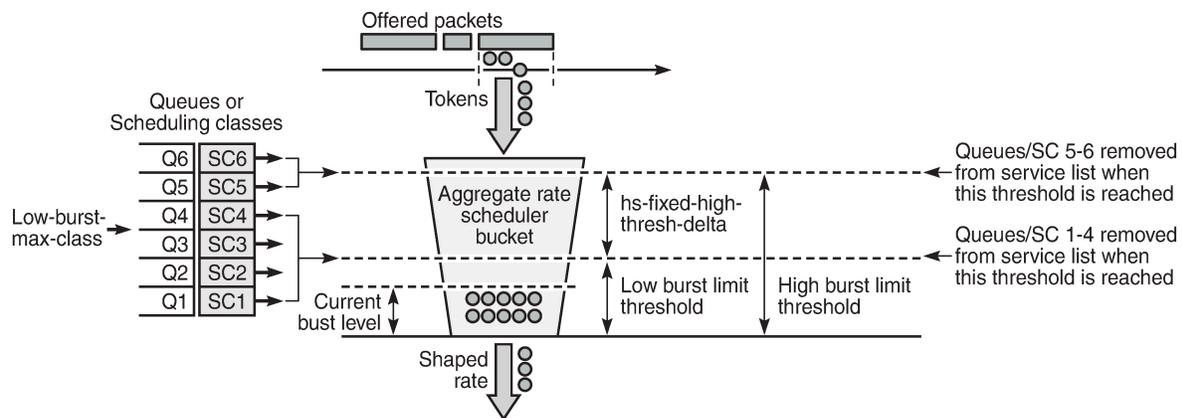
Tokens representing the bytes in the packets are added to the bucket as packets are forwarded. Tokens are drained from the scheduler bucket at the configured aggregate PIR rate. If the rate at which packets are forwarded (tokens are added) exceeds the shaping rate (tokens drained), a depth of tokens builds up in the bucket. If the depth reaches the low burst limit threshold, the queues using the low threshold are

removed from the scheduling list. If the depth continues to increase and reaches the high threshold, the remaining queues are removed from the scheduler list.

The low burst threshold depth is determined by the system. It is equivalent to the burst control group visitation time used by the FP egress queue scheduler. The shaping rate tokens are periodically removed from the bucket by the system by decrementing the current burst size. This period must be small enough to ensure that the resulting decrement does not cause the bucket depth to be negative, which is not permitted. Because the bucket depth cannot be negative, any potential negative decrement is lost, which equates to a loss of scheduling opportunities and the queue would underrun.

The high burst limit threshold uses a fixed increment on top of the low burst limit threshold. This fixed increment is configured under **card>fp>egress** using the **hs-fixed-high-thresh-delta** parameter and has a default value of 4000 bytes. It is recommended to set this parameter to a value at least two times the maximum packet size to prevent the classes using the low burst threshold from affecting those using the high burst threshold when forwarding larger packets. An insufficient burst threshold delta defeats the intended purpose of mapping classes to the high burst threshold.

The **low-burst-max-class** parameter in the HS attachment policy (for the HSQ queue group aggregate rate) or under the secondary shaper configuration (for the secondary shaper aggregate rate) configures which queues use the low burst limit threshold and which use the high burst limit threshold. This parameter has a default max class of 6 in both contexts. As the name of the parameter implies, the specified class is the highest class that uses the low burst threshold; classes above the specified class use the high burst threshold.

Buffer Management

The HSQ supports a flexible buffer management configuration that allows both buffer isolation and buffer oversubscription for the queue buffer allocation. There are four levels to the buffer hierarchy, which are shown in Figure 12: HSQ Buffer Pool Hierarchy:

- Root pools
- Mid pools
- Port class pools
- Queue group queues

Figure 12: HSQ Buffer Pool Hierarchy
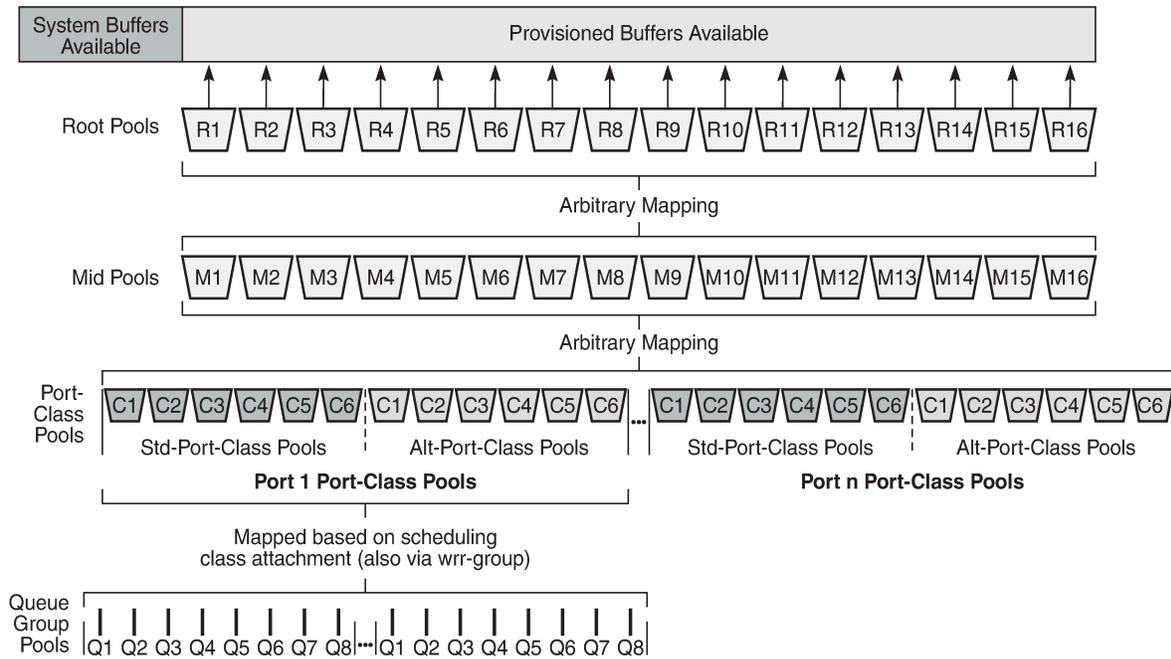
The total buffer allocation is divided into a system-reserved portion and a user-provisioned portion. The system buffers are allocated 5% of the total buffers in the default **hs-pool-policy**, which is applied to all HSQ IOMs under card>fp>egress. This value can be modified by creating a new **hs-pool-policy**, setting its **system-reserve** parameter, and applying the policy on an HSQ IOM. The user-provisioned portion is allocated the remainder of the available buffers, which can be configured as follows.

Root Pools

The root pools represent the total number of available buffers that can be provisioned. Up to 16 root pools can be configured, each having an allocation weight to determine its allocation of the available buffers. A root pool with an allocation weight of zero is not allocated any buffers. Root pools cannot oversubscribe the real buffers on the IOM. The use of multiple root pools provides buffer isolation between the queues using each root pool. At least one root pool (root pool 1) must be assigned buffers by having a non-zero allocation weight.

A high watermark is maintained for the buffer usage in each root pool. A slope policy is applied to each root pool to handle congestion control, the default being the _tmnx_hs_default slope policy. Root pools are configured per FP in an **hs-pool-policy** applied under **card>fp>egress**. Root pools 1 and 2 have an allocation weight of 75 and 25, respectively, in the default **hs-pool-policy**, with the remaining pools having a weight of 0.

Mid Pools

The mid pools are an abstract pool mapping mechanism. Each mid pool can be parented to a single parent root pool using its **parent-root-pool** parameter. Mid pools cannot be parented to a root pool without buffers

and mid pools are unused if not parented to a root pool. Up to 16 mid pools are available and at least one mid pool must be parented to a root pool for its queues to buffer packets. The number of buffers in a mid pool is configured as a percentage of its parent root pool size using the **allocation-percent** parameter.

Mid pools can facilitate buffer isolation by being mapped to different root pools. Mapping multiple mid pools to the same root pool allows the buffers of that root pool to be shared by those child mid pools, and if the sum of the child mid pool allocation percent is greater than 100, then the root pool will be oversubscribed accordingly.

An oversubscription factor can also be applied to each mid pool (using the **port-bw-oversub-factor** parameter) to permit its child class pools to oversubscribe it. This does not change the size of the mid pool, but allows the mid pool size to be increased in the calculation of each of its child port class pools.

A high watermark is maintained for the buffer usage in each mid pool. A slope policy is applied to each mid pool to handle congestion control, the default being the _tmnx_hs_default slope policy. Mid pools are configured per FP egress in an **hs-pool-policy** applied under **card>fp>egress**. In the default **hs-pool-policy**, mid pools 1 to 4 are parented to root pool 1 with allocation percentages of 40, 35, 30, and 25; mid pools 5 and 6 are parented to root pool 2 with allocation percentages of 80 and 20; and mid pools 7 to 16 are not parented to any root pool. All mid pools have a **port-bw-oversub-factor** of 1.

Port Class Pools

Port class pools, as the name implies, are per-class pools that exist at the port level. There are two sets of port class pools per port: six standard port class pools and six alternative port class pools. The alternative set of port class pools enables additional flexibility for both buffer isolation and oversubscription by providing a simple mechanism to parent queues to different port class pools and, therefore, to different mid and root pools.

HSQ queue group queues are statically assigned to the port class pool associated with the scheduling class that they have been attached to (via a WRR group, if used): scheduling class 1 to port class pool 1, up to scheduling class 6 to port class pool 6. Port class pools are configured in an **hs-port-pool-policy**, which is applied under **config>port>ethernet>egres**s. A default **hs-port-pool-policy** in which only the standard port class pools are used is applied to all HSQ ports.

Queues can be assigned to an alternative class pool (again based on the associated scheduling class) using the **hs-alt-port-class-pool** parameter under the queue in the SAP egress QoS policy, network queue policy, or egress queue group template.

Each port class pool parents to a single mid pool using its **parent-mid-pool** parameter. Port class pools are unused if not parented to a mid pool. Each port class pool must be parented to a mid pool that is parented to a root pool for queues to buffer packets. Port class pools can facilitate buffer isolation by being parented to different mid pools that are parented to different root pools. The standard port class pools are parented to their respective mid pool (port class pool 1 to mid pool 1, up to port class pool 6 to mid pool 6) in the default **hs-port-pool-policy**, with the alternative port class pools not parented to any mid pool.

The oversubscription of port class pools in a mid pool can be achieved by configuring the **port-bw-oversub-factor** under the parent mid pool (in the **hs-pool-policy**), which is multiplied by the size of the mid pool when calculating the size of each child class pool.

A weight is configurable per port to handle the allocation of buffers to different class pools parented to the same mid pool. This is configured using the **allocation port-bw-weight** under the class pool statement, where the weight configured for a port class pool is divided by the sum of the weights of the port class pools parented to the same mid pool, to determine the proportion of the allocated buffers for that port class pool. It is also possible to configure an **explicit-percent** for a port class pool, in which case that port class

pool will be allocated the configured explicit percentage of the mid pool (without any mid pool **port-bw-oversub-factor** being applied).

If there are multiple port class pools parented to the same mid pool, their buffer allocation is determined using the weighting mechanism based on the port class pool **allocation port-bw-weight** parameter. Port class pools configured with an **explicit-percent** have a weight of zero (that is, they do not participate in the weighting buffer allocation). The port class pools in the default **hs-port-pool-policy** are configured with an **allocation port-bw-weight** of 1.

The port class pools are sized dynamically to provide a fair share of a mid pool size to each of its child port class pools, based on the potential bandwidth represented by each port on which the port class pools exist. The first step is to determine the usable bandwidth of each port. The mid pool buffers are then shared between its child port class pools, based on their related port usable bandwidth. An oversubscription factor is then applied to allow the port class pools to oversubscribe their mid pool. Finally, each port mid pool buffer allocation is shared between the child port class pools on that port.

No buffers are allocated to port class pools if there are no SAPs or network interfaces configured on that port and the port is shutdown. The details of the port class pool sizing calculation are as follows (examples of each are shown in the Buffer Pools configuration section):

1. Determine each port bandwidth value.

   a. This is the minimum of the port current line rate, the port **egress-rate** limit, and the **hs-scheduler-policy max-rate** configured on the port.

   b. The port bandwidth may be further modified by the port **modify-buffer-allocation-rate egr-percentage-of-rate** command, which can increase or decrease the port bandwidth by the specified percent. This allows the port to have a higher or lower bandwidth derived weight, based on how the port is being used, instead of bandwidth alone.

2. Determine each port portion of each mid pool.

   The port class pools are configured to map to the mid pools, so it is possible that not every port will have a port class pool associated with a mid pool. This requires that the system perform the relative bandwidth calculations separately per mid pool. A port without any port class pools associated with a mid pool will have a port portion of zero for that mid pool.

   Per mid pool, each port portion of the mid pool size is calculated based on:

   Port_Portion = (Port_Adj_Bw / Sigma_Mid_Pool_Ports_Adj_Bw) * Mid_Pool_Size

   Where:

   - *Port_Adj_Bw* is calculated in (1).
   - *Sigma_Mid_Pool_Ports_Adj_Bw* is the sum of the adjusted bandwidths for all ports, with port class pools mapped to the mid pool that are not sized configured with **explicit-percent** (see (4)).
   - *Mid_Pool_Size* is the mid pool parent root pool size multiplied by the mid pool allocation weight.

3. Modify the mid pool sizes by their **port-bw-oversub-factor**.

   The port bandwidth weighting mechanism allocates 100% of the mid pool size to the associated port class pools. To allow the port class pools to oversubscribe their parent mid pool, the mid pool **port-bw-oversub-factor** parameter can be used to increase the apparent size of the mid pool (this does not change the mid pool size) in the calculation in (2). This potentially provides a more efficient use of the mid pool available buffers since it is not expected that all port class pools will be using their allotted size simultaneously.

4. Determine each port class pool share of the mid pool port share.

Multiple port class pools on the same port may be mapped to the same mid pool. This requires a mechanism to distribute the portion of the mid pool allocated to each port class pool on that port.

Each port class pool **allocation port-bw-weight** parameter is used to determine how much of the port mid pool is given to each port class pool associated with the mid pool. A port class pool is allocated the portion of its mid pool size multiplied by its port class pool **port-bw-weight** divided by the sum of the **port-bw-weight** for all port class pools associated with that mid pool on that port.

Alternatively, port class pools can be sized using an **explicit-percent** of the actual mid pool size (without applying the **port-bw-oversub-factor**). These class pools are assigned a **port-bw-weight** equal to zero, causing them to be excluded from the port portion distribution. It is expected (but not required) that either port bandwidth-based sizing or explicit percent-based sizing will be used, with concurrent use of both mechanisms being transitory in nature.

Whenever one of the inputs to the preceding calculations changes, the bandwidth weighted sizes for the corresponding pool class pools are recalculated.

A high watermark is maintained for the buffer usage in each port class pool. A slope policy is applied to each port class pool to handle congestion control, the default being the *_tmnx_hs_default* slope policy.

Queue Group Queues

HSQ queue group queues always operate in WRED per queue mode, supporting three WRED slopes. The total number of buffers usable by the queue is limited by the queue MBS configuration, and each packet profile type (exceed, out, in) is limited by the respective slope configuration (exceed, low, high) in the applied slope policy, if the slope is not shutdown. For a buffer to be allocated, the applicable WRED slope processing (if enabled) must accept the packet, the MBS must not be exceeded, and there must be available buffers in its parent port class pool, mid pool, and root pool.

The regular **mbs** queue parameter configuration is used within SAP egress QoS policies and egress queue group templates, using the regular defaults. In the network queue policy, the MBS is configured using the **hs-mbs** parameter, which allows a different default to be used, with its value calculated based on a percentage of one second of the queue PIR converted to bytes (the regular **mbs** parameter is ignored in the network queue policy). The queue CBS and drop tail configuration is ignored on an HSQ queue group queue.

A default slope (named *_tmnx_hs_default*) is applied to each HSQ queue, using the **policy** parameter on the **hs-wred-queue** statement within a SAP egress QoS policy, network queue policy, and egress queue group template. A user-configured regular slope policy can be applied using the same parameter and statement. The **highplus-slope** and **time-average-factor** in the applied slope policy are ignored on HSQ queue group queues.

LAGs

LAGs are supported on HSQ ports. The LAG **port-type** must be set to **hs** to add an HSQ port to a LAG, at which point only HSQ ports can be added to that LAG. When an HSQ queue group is created on a LAG, an HSQ queue group is allocated on each LAG port.

LAG **access adapt-qos** modes **link** and **port-fair** are supported; **distribute** mode is not supported.

LAG **access per-fp-egr-queuing** is supported and, when configured, either **per-link-hash** or **per-service-hashing** (supported service types only) must be enabled under the **LAG access**. LAG **access per-fp-sap-instance** is supported (this requires **per-fp-egr-queuing** to be enabled).

The full configured queue MBS is applied to all the related HSQ queue group queues on the individual LAG ports.

Configuration

This section describes simple configurations using an HSQ IOM for SAP egress, network egress, and access and network egress queue group instances.

Each configuration uses the same four queues:

- Queue 7 at scheduling class 5

- Queue 6 at scheduling class 4

- Queues 1 and 2 in WRR group 1 using scheduling class 1, with queue 1 having a weight of 2 and queue 2 having a weight of 1

Scheduling class 6 has been reserved for queue 8 to be used for network protocol traffic. Queue 3 is unattached, but could at some point be added to WRR group 1. The configuration provides the future flexibility to either add queues 4 and 5 to WRR group 1, to WRR group 2, or attach them to scheduling classes 2 and 3. Although eight queues are always allocated in an HSQ queue group, only the queues to be used need to be configured.

The QoS path for the configured SAP, network interface, and access and network queue group instances with respect to their HSQ queue group, primary shaper, secondary shaper, and port scheduler, is shown in Figure 13: Configured QoS Paths.

Figure 13: Configured QoS Paths
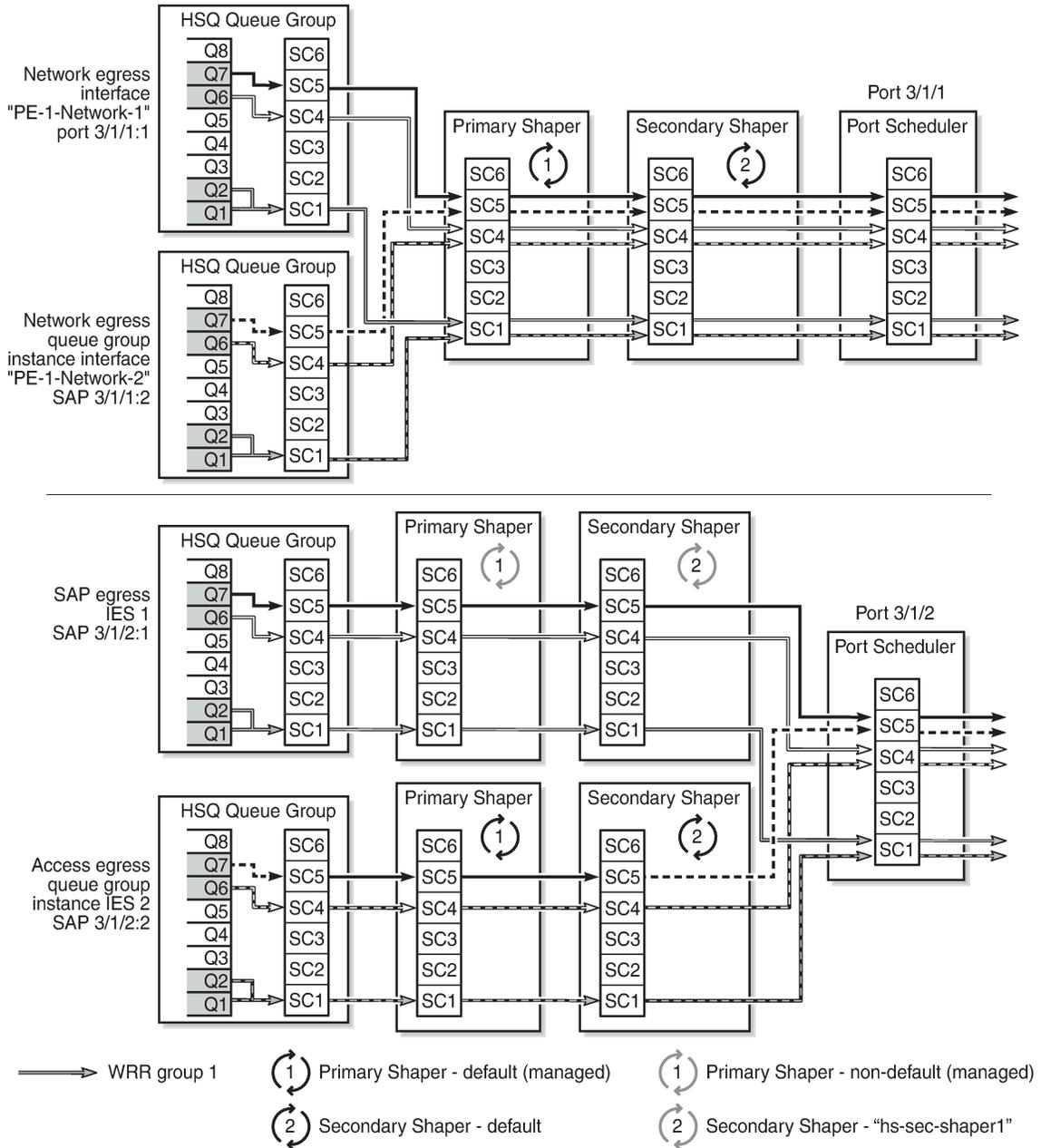
The configurations start with the generic aspects:

- Card configuration
- Buffer pools
- Shaping and scheduling
  - HSQ queue groups

- HS secondary shapers

  - These are specific to SAP egress in the context of this chapter; however, as secondary shapers can also be used by subscribers, they are included with the generic aspects.

- Ports

This is followed by the specific configuration related to:

- SAP egress

- Network egress

- Access and network egress queue groups

Card Configuration

An HSQ IOM is configured with the card type *iom4-e-hs* and the associated supported MDAs:

```
A:PE-1# configure card 3
A:PE-1>config>card# info
----------------------------------------------
        card-type iom4-e-hs
        mda 1
            mda-type me10-10gb-sfp+
            no shutdown
        exit
        no shutdown
----------------------------------------------
A:PE-1>config>card# exit all
A:PE-1# show card


===============================================================================
Card Summary
===============================================================================
Slot      Provisioned Type                        Admin Operational   Comments
             Equipped Type (if different)         State State
-------------------------------------------------------------------------------
2         imm-2pac-fp3                             up    up
3         iom4-e-hs                                up    up
A         cpm5                                     up    up/active
B         cpm5                                     up    up/standby
===============================================================================
A:PE-1#
```

The supported MDA types are displayed by entering a "?" after the **mda-type** parameter:

```
A:PE-1# configure card 3 mda 1 mda-type
  - mda-type <mda-type>
  - no mda-type

 <mda-type>            : me1-100gb-cfp2|me10-10gb-sfp+|me12-10/1gb-sfp+|me2-100gb-cfp4|me2-
100gb-qsfp28|me40-1gb-csfp|me6-10gb-sfp+
```

The **hs-fixed-high-thresh-delta** on card 3 fp 1 is *default*, resulting in the high burst limit threshold (which is used by queues and WRR groups attached to scheduling classes above the **low-burst-max-class**) being 4000 bytes larger than the low burst limit threshold:

```
*A:PE-1# show card 3 detail | match "HS Fixed High Threshold Delta"
    HS Fixed High Threshold Delta : default
```

```
*A:PE-1#
```

The HSQ-specific resource usage is displayed as follows:

```
*A:PE-1# tools dump resource-usage card 3 fp 1

===============================================================================
Resource Usage Information for Card Slot #3 FP #1
===============================================================================
                                        Total    Allocated      Free
-------------------------------------------------------------------------------

                        Egress Queues |   786432        123      786309
                      Ingress Policers |   511999          1      511998
                  Ingress Policer Stats |   511967          0      511967

           Egress HS Turbo Queue Group |       64         10          54
                Egress HS Queue Group |    98240         36       98204
                   HS Primary Shapers +    16384         22       16362
           HS Explicit Primary Shapers -                  0
            HS Managed Primary Shapers -                 22
                 HS Secondary Shapers |     4096         22        4074
===============================================================================
*A:PE-1#
```

The preceding output displays the usage information of the egress queues, the ingress policers and their statistics, the HS turbo queue groups, HS queue groups, HS primary shapers (the managed primary shapers are system-created, whereas the explicit primary shapers are used for ESM), and the HS secondary shapers.

The following card commands are ignored on HSQ IOMs:

- All regular **pool** commands
- **ingress-buffer-allocation**
- **reset-on-recoverable-error**
- **virtual-scheduler-adjustment** (egress only)

The following card commands are not configurable on HSQ IOMs:

- **named-pool-mode**
- **stable-pool-sizing**
- **egress wred-queue-control**

Buffer Pools

The buffer pool configuration used in this example provides buffer isolation between traffic in queue 7, queue 6, and WRR group 1, by assigning each to a different root and mid pool. The combined queue 1 and 2 traffic share a root and mid pool.

Root and Mid Pools

The HSQ root and mid pools for an IOM are configured in an HS pool policy, which is applied under **card>fp>egress**.

An HS pool policy is configured as follows:

```
configure
    qos
        hs-pool-policy <policy-name> [create]
            description <description-string>
            mid-tier
                mid-pool <mid-pool-id>
                    allocation-percent <percent-of-parent-pool>
                    parent-root-pool <root-pool-id>
                    port-bw-oversub-factor <oversubscription-factor>
                    slope-policy <policy-name>
            root-tier
                root-pool <root-pool-id>
                    allocation-weight <pool-weight>
                    slope-policy <policy-name>
            system-reserve <percent-of-buffers>
```

Where (in order):

```
<policy-name>        : [32 chars max]
<description-string> : [80 chars max]
<mid-pool-id>        : [1..16]
<percent-of-parent*> : [0.01..100.00]
<root-pool-id>       : [1..16]
<oversubscription-*> : [1..10]
<policy-name>        : [32 chars max]
<root-pool-id>       : [1..16 | none]
<pool-weight>        : [0..100]
<policy-name>        : [32 chars max]
<percent-of-buffers> : [1.00..30.00]
```

A default HS pool policy is created by the system with the following configuration:

hs-pool-policy default						
System reserve: 5%						
Root pools			**Mid pools**			
Root Pool ID	**Allocation weight**	**Slope policy**	**Mid Pool ID**	**Parent mid pool**	**Allocation %**	**Port BW oversub factor**	**Slope policy**
1	75	_tmnx_hs_default	1	1	40%	1	_tmnx_hs_default
2	25	_tmnx_hs_default	2	1	35%	1	_tmnx_hs_default
3-16	0		3	1	30%	1	_tmnx_hs_default
			4	1	25%	1	_tmnx_hs_default
			5	2	80%	1	_tmnx_hs_default
			6	2	20%	1	_tmnx_hs_default
			7-16	None			

If a new HS pool policy is created, its initial configuration is as follows:

hs-pool-policy <new>								
System reserve: 5%								
Root pools			**Mid pools**					
Root Pool ID	**Allocation weight**	**Slope policy**	**Mid Pool ID**	**Parent mid pool**	**Allocation %**	**Port BW oversub factor**		**Slope policy**
1	100	_tmnx_hs_default	1-16	1	1%	1		_tmnx_hs_default
2-16								

The HS pool policy (*hs-pool-pol-1*) used for this example is shown following. Root pool 1 and mid pool 1 have been reserved for network protocol traffic and are not used. Root pool 2 and mid pool 2 are used for queue 7 traffic, root pool 3 and mid pool 3 are used for queue 6 traffic, and root pool 4 and mid pool 4 are used for queue 1 and queue 2 (WRR group 1) traffic. The buffer allocation for each pool is based on the expected traffic volumes.

Root pool 4 and mid pool 4 have a more aggressive slope policy (*hs-slope-1*) than the default HSQ slope policy (*_tmnx_hs_default*). The default HSQ slope policy is as follows (the HSQ slopes use the instantaneous queue depth so the **time-average-factor** is ignored and the **highplus-slope** is also ignored):

```
*A:PE-1>config>qos# slope-policy "_tmnx_hs_default"
*A:PE-1>config>qos>slope-policy# info detail
----------------------------------------------
            description "Default HS slope policy."
            highplus-slope
                shutdown
                start-avg 100
                max-avg 100
                max-prob 100
            exit
            high-slope
                start-avg 100
                max-avg 100
                max-prob 100
                no shutdown
            exit
            low-slope
                start-avg 90
                max-avg 90
                max-prob 100
                no shutdown
            exit
            exceed-slope
                start-avg 80
                max-avg 80
                max-prob 100
                no shutdown
            exit
            time-average-factor 7
```

The slope policy *hs-slope-1* is configured as follows:

```
A:PE-1>config>qos# slope-policy "hs-slope-1"
```

```
A:PE-1>config>qos>slope-policy# info
----------------------------------------------
            highplus-slope
                shutdown
            exit
            high-slope
                start-avg 85
                max-avg 100
                no shutdown
            exit
            low-slope
                no shutdown
            exit
            exceed-slope
                shutdown
            exit
```

Mid pool 4 has been configured to allow a 4 times oversubscription by its child class pools.

Root pools 5 to 16 and mid pools 5 to 16 are unused.

HS pool policy *hs-pool-pol-1* is summarized as follows:

hs-pool-policy hs-pool-pol-1							
System reserve: 5%							
Root pools			**Mid pools**				
Root Pool ID	**Allocation weight**	**Slope policy**	**Mid Pool ID**	**Parent mid pool**	**Allocation %**	**Port BW oversub factor**	**Slope policy**
1	5	_tmnx_hs_default	1	1	100%	1	_tmnx_hs_default
2	10	_tmnx_hs_default	2	2	100%	1	_tmnx_hs_default
3	20	_tmnx_hs_default	3	3	100%	1	_tmnx_hs_default
4	65	hs-slope-1	4	4	100%	4	hs-slope-1
5-16	0	_	5-16	None			

The HS pool policy hs-pool-pol-1 is configured as follows:

```
A:PE-1>config>qos>hs-pool-policy# info
----------------------------------------------
            root-tier
                root-pool 1
                    allocation-weight 5
                exit
                root-pool 2
                    allocation-weight 10
                exit
                root-pool 3
                    allocation-weight 20
                exit
                root-pool 4
                    allocation-weight 65
                    slope-policy "hs-slope-1"
                exit
            exit
```

```
            mid-tier
                mid-pool 1
                    allocation-percent 100.00
                exit
                mid-pool 2
                    parent-root-pool 2
                    allocation-percent 100.00
                exit
                mid-pool 3
                    parent-root-pool 3
                    allocation-percent 100.00
                exit
                mid-pool 4
                    parent-root-pool 4
                    allocation-percent 100.00
                    port-bw-oversub-factor 4
                    slope-policy "hs-slope-1"
                exit
                mid-pool 5
                    parent-root-pool none
                exit
                mid-pool 6
                    parent-root-pool none
                exit
                mid-pool 7
                    parent-root-pool none
                exit
                mid-pool 8
                    parent-root-pool none
                exit
                mid-pool 9
                    parent-root-pool none
                exit
                mid-pool 10
                    parent-root-pool none
                exit
                mid-pool 11
                    parent-root-pool none
                exit
                mid-pool 12
                    parent-root-pool none
                exit
                mid-pool 13
                    parent-root-pool none
                exit
                mid-pool 14
                    parent-root-pool none
                exit
                mid-pool 15
                    parent-root-pool none
                exit
                mid-pool 16
                    parent-root-pool none
                exit
            exit
```

The HS pool policy is shown as follows:

```
*A:PE-1# show qos hs-pool-policy "hs-pool-pol-1"


===============================================================================
HS Pool Policy Information
===============================================================================
```

```
Policy Name            : hs-pool-pol-1
Description            : (Not Specified)
System Reserve         : 5.00


-------------------------------------------------------------------------------
Root Pool Information
-------------------------------------------------------------------------------
Pool Id                : 1             Allocation Weight   : 5
Slope Policy           : _tmnx_hs_default

Pool Id                : 2             Allocation Weight   : 10
Slope Policy           : _tmnx_hs_default

Pool Id                : 3             Allocation Weight   : 20
Slope Policy           : _tmnx_hs_default

Pool Id                : 4             Allocation Weight   : 65
Slope Policy           : hs-slope-1

Pool Id                : 5             Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 6             Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 7             Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 8             Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 9             Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 10            Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 11            Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 12            Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 13            Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default
Pool Id                : 14           Alloc
ation Weight   : 0
Slope Policy           : _tmnx_hs_default
Pool Id                : 15            Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default

Pool Id                : 16            Allocation Weight   : 0
Slope Policy           : _tmnx_hs_default


-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
Mid Pool Information
-------------------------------------------------------------------------------
Pool Id                : 1             Allocation Percent  : 100.00
Port BW Oversub Factor : 1             Parent Root Pool    : 1
Slope Policy           : _tmnx_hs_default

Pool Id                : 2             Allocation Percent  : 100.00
```

```
Port BW Oversub Factor  : 1              Parent Root Pool    : 2
Slope Policy            : _tmnx_hs_default

Pool Id                 : 3              Allocation Percent  : 100.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 3
Slope Policy            : _tmnx_hs_default

Pool Id                 : 4              Allocation Percent  : 100.00
Port BW Oversub Factor  : 4              Parent Root Pool    : 4
Slope Policy            : hs-slope-1

Pool Id                 : 5              Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 6              Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 7              Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 8              Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 9              Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 10             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 11             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 12             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 13             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 14             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 15             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

Pool Id                 : 16             Allocation Percent  : 1.00
Port BW Oversub Factor  : 1              Parent Root Pool    : 0
Slope Policy            : _tmnx_hs_default

-------------------------------------------------------------------------------
===============================================================================
*A:PE-1#
```

This HS pool policy is configured for the HSQ IOM as follows:

```
A:PE-1# configure card 3
A:PE-1>config>card# info
----------------------------------------------
        card-type iom4-e-hs
        fp 1
            egress
                hs-pool-policy "hs-pool-pol-1"
            exit
        exit
        no shutdown
----------------------------------------------
A:PE-1>config>card#
```

The association of this HS pool policy is shown as follows:

```
*A:PE-1# show qos hs-pool-policy "hs-pool-pol-1" association


===============================================================================
HS Pool Policy Information
===============================================================================
Policy Name            : hs-pool-pol-1
Description            : (Not Specified)
System Reserve         : 5.00

-------------------------------------------------------------------------------
Card Forwarding Plane (FP) Associations
-------------------------------------------------------------------------------
Card                   FP
-------------------------------------------------------------------------------
3                      1
-------------------------------------------------------------------------------
===============================================================================
*A:PE-1#
```

The resulting system and user-provisioned pool information is shown following. This output shows the total buffer allocation, number of allocated buffers, available buffer allocation, and buffer high watermarks for the system pools and user-provisioned pools. The output shows the hierarchy of the root and mid pools, with their applied slope policy and the related instantaneous slope drop probabilities (as a percentage):

```
*A:PE-1# show hs-pools 3 fp 1 egress


===============================================================================
HS Pools Card Forwarding Plane Information
===============================================================================
Card                 : 3                    FP                   : 1

-------------------------------------------------------------------------------
System Pool Information
-------------------------------------------------------------------------------

Total Buffers      : 209412 KB      Allocated         : 0 KB
Available          : 209412 KB      High Water Mark   : 0 KB

-------------------------------------------------------------------------------
Buffer Pool Hierarchy Information
-------------------------------------------------------------------------------

Root Pool : 1
|   Total               : 198942 KB    Allocated          : 0 KB
```

```
| Available         : 198942 KB   High Water Mark   : 0 KB
| Hi-Slope Drop Prob : 0          Lo-Slope Drop Prob: 0
| Excd-Slope Drop Prob: 0
| Hs Slope Policy    : _tmnx_hs_default
|
|--- Mid Pool : 1
|    | Total              : 198942 KB   Allocated       : 0 KB
|    | Available          : 198942 KB   High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy     : _tmnx_hs_default
|    |
Root Pool : 2
| Total              : 397886 KB   Allocated       : 0 KB
| Available          : 397886 KB   High Water Mark   : 0 KB
| Hi-Slope Drop Prob : 0          Lo-Slope Drop Prob: 0
| Excd-Slope Drop Prob: 0
| Hs Slope Policy    : _tmnx_hs_default
|
|--- Mid Pool : 2
|    | Total              : 397886 KB   Allocated       : 0 KB
|    | Available          : 397886 KB   High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy     : _tmnx_hs_default
|    |
Root Pool : 3
| Total              : 795772 KB   Allocated       : 0 KB
| Available          : 795772 KB   High Water Mark   : 0 KB
| Hi-Slope Drop Prob : 0          Lo-Slope Drop Prob: 0
| Excd-Slope Drop Prob: 0
| Hs Slope Policy    : _tmnx_hs_default
|
|--- Mid Pool : 3
|    | Total              : 795772 KB   Allocated       : 0 KB
|    | Available          : 795772 KB   High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy     : _tmnx_hs_default
|    |
Root Pool : 4
| Total              : 2586262 KB  Allocated       : 0 KB
| Available          : 2586262 KB  High Water Mark   : 0 KB
| Hi-Slope Drop Prob : 0          Lo-Slope Drop Prob: 0
| Excd-Slope Drop Prob: 0
| Hs Slope Policy    : hs-slope-1
|
|--- Mid Pool : 4
|    | Total              : 2586262 KB  Allocated       : 0 KB
|    | Available          : 2586262 KB  High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy     : hs-slope-1
|    |
```

Port Class Pools

The HSQ port class pools for a port are configured in an HS port pool policy, which is applied under **config>port>ethernet>egress**.

An HS port pool policy is configured as follows:

```
configure
    qos
        hs-port-pool-policy <policy-name> [create]
            description <description-string>
            std-port-class-pools
                class-pool <std-class-pool-id>
                    allocation explicit-percent <percent-of-parent-pool>
                    allocation port-bw-weight <pool-weight>
                    parent-mid-pool <mid-pool-id>
                    slope-policy <policy-name>
            alt-port-class-pools
                class-pool <alt-class-pool-id>
                    allocation explicit-percent <percent-of-parent-pool>
                    allocation port-bw-weight <pool-weight>
                    parent-mid-pool <mid-pool-id>
                    slope-policy <policy-name>
```

Where (in order):

```
<policy-name>        : [32 chars max]
<description-string> : [80 chars max]
<std-class-pool-id>  : [1..6]
<percent-of-parent*> : [0.01..100.00]
<pool-weight>        : [1..100]
<mid-pool-id>        : [1..16 | none]
<policy-name>        : [32 chars max]
<alt-class-pool-id>  : [1..6]
<percent-of-parent*> : [0.01..100.00]
<pool-weight>        : [1..100]
<mid-pool-id>        : [1..16 | none]
<policy-name>        : [32 chars max]
```

A default HS pool policy is created by the system with the following configuration:

hs-port-pool-policy default							
Standard port class pools				**Alternative port class pools**			
Class Pool ID	**Parent mid pool**	**Allocation port bw weight**	**Slope policy**	**Class Pool ID**	**Parent mid pool**	**Allocation port bw weight**	**Slope policy**
1	1	1	_tmnx_hs_default	1-6	None		
2	2	1	_tmnx_hs_default				
3	3	1	_tmnx_hs_default				
4	4	1	_tmnx_hs_default				
5	5	1	_tmnx_hs_default				
6	6	1	_tmnx_hs_default				

Newly created HS port pool policies have the following parameters:

hs-port-pool-policy <new>

Standard port class pools				Alternative port class pools			
Class Pool ID	Parent mid pool	Allocation port bw weight	Slope policy	Class Pool ID	Parent mid pool	Allocation port bw weight	Slope policy
1-6	1	1	_tmnx_hs_default	1-6	None		

The HS port pool policy used for traffic in this example is shown following. Only port class pools 1, 4, 5, and 6 are used, which are parented to mid pools 4, 3, 2, and 1, respectively. As scheduling classes 2 and 3 are unused, their associated standard port class pools are not parented to a mid pool. The alternative port class pools are also unused, so are not parented to a mid pool. Standard port class pools 4 to 6 use the default HSQ slope policy with standard port class pool 1 using slope policy *hs-slope-1*.

hs-port-pool-policy hs-port-pool-pol-1

Standard port class pools				Alternative port class pools			
Class Pool ID	Parent mid pool	Allocation port bw weight	Slope policy	Class Pool ID	Parent mid pool	Allocation port bw weight	Slope policy
1	4	1	hs-slope-1	1-6	None		
2-3	None						
4	3	1	_tmnx_hs_default				
5	2	1	_tmnx_hs_default				
6	1	1	_tmnx_hs_default				

The HS port pool policy is configured as follows:

```
*A:PE-1>config>qos# hs-port-pool-policy "hs-port-pool-pol-1"
*A:PE-1>config>qos>hs-port-pool-policy# info
--------------------------------------------
            std-port-class-pools
                class-pool 1
                    parent-mid-pool 4
                    slope-policy "hs-slope-1"
                exit
                class-pool 2
                    parent-mid-pool none
                exit
                class-pool 3
                    parent-mid-pool none
                exit
                class-pool 4
                    parent-mid-pool 3
                exit
                class-pool 5
                    parent-mid-pool 2
                exit
            exit
```

The HS port pool policy is shown as follows:

```
*A:PE-1# show qos hs-port-pool-policy "hs-port-pool-pol-1"

===============================================================================
HS Port Pool Policy Information
===============================================================================
Policy Name             : hs-port-pool-pol-1
Description             : (Not Specified)


-------------------------------------------------------------------------------
Standard Port Class Pool Information
-------------------------------------------------------------------------------
Class Id                : 1            Parent Mid Pool    : 4
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : hs-slope-1

Class Id                : 2            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 3            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 4            Parent Mid Pool    : 3
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 5            Parent Mid Pool    : 2
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 6            Parent Mid Pool    : 1
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default


-------------------------------------------------------------------------------


-------------------------------------------------------------------------------
Alternate Port Class Pool Information
-------------------------------------------------------------------------------
Class Id                : 1            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 2            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 3            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 4            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 5            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
Slope Policy            : _tmnx_hs_default

Class Id                : 6            Parent Mid Pool    : 0
Alloc Port BW Weight    : 1            Alloc Explicit Prcnt: 0.00
```

```
Slope Policy             : _tmnx_hs_default

        ----------------------------------------------------------------------
        ======================================================================
*A:PE-1#
```

The preceding HS port pool policy applied to ports 3/1/1 and 3/1/2 is shown as follows:

```
*A:PE-1# show qos hs-port-pool-policy "hs-port-pool-pol-1" association

===============================================================================
HS Port Pool Policy Information
===============================================================================
Policy Name             : hs-port-pool-pol-1
Description             : (Not Specified)

        -----------------------------------------------------------------------
Port Ethernet Egress Associations
        -----------------------------------------------------------------------
3/1/1
3/1/2
        -----------------------------------------------------------------------
===============================================================================
*A:PE-1#
```

The remaining ports (3/1/[3..10]) are unused in this example, so their port class pools are not parented, by applying the following HS port pool policy to each:

hs-port-pool-policy no-class-pools							
Standard port class pools				Alternative port class pools			
Class Pool ID	Parent mid pool	Allocation port bw weight	Slope policy	Class Pool ID	Parent mid pool	Allocatior port bw weight	Slope policy
1-6	None			1-6	None		

```
*A:PE-1>config>qos# hs-port-pool-policy "no-class-pools"
*A:PE-1>config>qos>hs-port-pool-policy# info
----------------------------------------------
            std-port-class-pools
                class-pool 1
                    parent-mid-pool none
                exit
                class-pool 2
                    parent-mid-pool none
                exit
                class-pool 3
                    parent-mid-pool none
                exit
                class-pool 4
                    parent-mid-pool none
                exit
                class-pool 5
                    parent-mid-pool none
                exit
                class-pool 6
                    parent-mid-pool none
                exit
```

```
            exit
```

The HS port pool policies applied to the ports on card 3 are configured as follows:

```
configure
    port 3/1/1
        ethernet
            egress
                hs-port-pool-policy "hs-port-pool-pol-1"
    port 3/1/2
        ethernet
            egress
                hs-port-pool-policy "hs-port-pool-pol-1"
    port 3/1/[3..10]
        ethernet
            egress
                hs-port-pool-policy "no-class-pools"
```

The pools created on port 3/1/1, after applying the preceding HS pool policy and HS port pool policies, are shown following. The root and mid pools in this output are the same as in the **show hs-pools 3 fp 1 egress** output; these pools are configured per FP so are the same for all ports. The additional information shows the details of the port class pools on this port and to which mid pools they are parented:

```
*A:PE-1# show hs-pools port 3/1/1 egress

===============================================================================
HS Pools Port Information
===============================================================================
Port               : 3/1/1


-------------------------------------------------------------------------------
System Pool Information
-------------------------------------------------------------------------------


Total Buffers      : 209412 KB        Allocated          : 0 KB
Available          : 209412 KB        High Water Mark    : 0 KB


-------------------------------------------------------------------------------
Buffer Pool Hierarchy Information
-------------------------------------------------------------------------------


Root Pool : 1
|  Total              : 198942 KB   Allocated         : 0 KB
|  Available          : 198942 KB   High Water Mark   : 0 KB
|  Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|  Excd-Slope Drop Prob: 0
|  Hs Slope Policy     : _tmnx_hs_default
|
|--- Mid Pool : 1
|    | Total              : 198942 KB   Allocated         : 0 KB
|    | Available          : 198942 KB   High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy     : _tmnx_hs_default
|    |
|    |--- Std Port Class Pool : 6
|    |     Total              : 99470 KB    Allocated         : 0 KB
|    |     Available          : 99470 KB    High Water Mark   : 0 KB
|    |     Hi-Slope Drop Prob  : 0          Lo-Slope Drop Prob: 0
|    |     Excd-Slope Drop Prob: 0
|    |     Hs Slope Policy     : _tmnx_hs_default
|    |
```

```
Root Pool : 2
|  Total              : 397886 KB    Allocated         : 0 KB
|  Available          : 397886 KB    High Water Mark   : 0 KB
|  Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|  Excd-Slope Drop Prob: 0
|  Hs Slope Policy    : _tmnx_hs_default
|
|--- Mid Pool : 2
|    | Total              : 397886 KB    Allocated         : 0 KB
|    | Available          : 397886 KB    High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy    : _tmnx_hs_default
|    |
|    |--- Std Port Class Pool : 5
|    |       Total              : 198942 KB    Allocated         : 0 KB
|    |       Available          : 198942 KB    High Water Mark   : 0 KB
|    |       Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|    |       Excd-Slope Drop Prob: 0
|    |       Hs Slope Policy    : _tmnx_hs_default
|    |
Root Pool : 3
|  Total              : 795772 KB    Allocated         : 0 KB
|  Available          : 795772 KB    High Water Mark   : 0 KB
|  Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|  Excd-Slope Drop Prob: 0
|  Hs Slope Policy    : _tmnx_hs_default
|
|--- Mid Pool : 3
|    | Total              : 795772 KB    Allocated         : 0 KB
|    | Available          : 795772 KB    High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy    : _tmnx_hs_default
|    |
|    |--- Std Port Class Pool : 4
|    |       Total              : 397886 KB    Allocated         : 0 KB
|    |       Available          : 397886 KB    High Water Mark   : 0 KB
|    |       Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|    |       Excd-Slope Drop Prob: 0
|    |       Hs Slope Policy    : _tmnx_hs_default
|    |
Root Pool : 4
|  Total              : 2586262 KB  Allocated         : 0 KB
|  Available          : 2586262 KB  High Water Mark   : 0 KB
|  Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|  Excd-Slope Drop Prob: 0
|  Hs Slope Policy    : hs-slope-1
|
|--- Mid Pool : 4
|    | Total              : 2586262 KB  Allocated         : 0 KB
|    | Available          : 2586262 KB  High Water Mark   : 0 KB
|    | Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|    | Excd-Slope Drop Prob: 0
|    | Hs Slope Policy    : hs-slope-1
|    |
|    |--- Std Port Class Pool : 1
|    |       Total              : 4194302 KB  Allocated         : 0 KB
|    |       Available          : 4194302 KB  High Water Mark   : 0 KB
|    |       Hi-Slope Drop Prob  : 0            Lo-Slope Drop Prob: 0
|    |       Excd-Slope Drop Prob: 0
|    |       Hs Slope Policy    : hs-slope-1
|    |
```

> **Note:**
> The maximum size of a port class pool is capped here at 4194302 kbytes, as shown for the port
> class pool parented to mid/root pool 4.

The following show commands are available to display the pool configuration and queue details:

```
show hs-pools port <port-id> egress
show hs-pools port <port-id> egress network-queues
show hs-pools port <port-id> egress queue-group <queue-group-name>
                            [instance <instance-id>]
show hs-pools port <port-id> egress sap <sap-id>
show hs-pools port <port-id> egress subscriber <sub-ident-string>
```

The root and mid pool information in each command is the same as in the command **show hs-pools
<card-slot-number> fp <forwarding-plane> egress** but the commands also show the port class pool
information on the specified port, together with the queue information for that port.

Pool Sizing and Oversubscription

The logic to size the root, mid, and port class pools is described in the Buffer Management section.

Root pools are sized using their **allocation-weight** parameter, which is divided by the sum of all root pool
allocation-weight to give the portion of the total user-provisioned buffers allocated to the root pool.

The mid pools are sized using their **allocation-percent** parameter, which is a percentage of their parent
root pool size.

The port class pools size calculation has more factors. The following output shows the effect of the different
sizing parameters on the port class pools size. The steps are:

- An HS pool policy and HS port pool policies are applied to an HSQ IOM and its ports to configure one
  root pool with one child mid pool that has one child standard port class pool on port 3/1/1. Each pool
  has the same size matching the total available buffers:

```
*A:PE-1# show hs-pools port 3/1/1 egress |
          match "Root Pool : 1" post-lines 26 | match "Pool" post-lines 1
Root Pool : 1
|  Total             : 3978866 KB  Allocated        : 0 KB
|--- Mid Pool : 1
|     | Total              : 3978866 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |      Total            : 3978866 KB  Allocated        : 0 KB
```

- A standard port class pool is added to port 3/1/2, causing the mid pool size to be shared between port
  class pool 1 on port 3/1/1 and 3/1/2:

```
*A:PE-1# configure port 3/1/2 ethernet egress
                           hs-port-pool-policy "hs-port-pool-policy-test"
*A:PE-1# show hs-pools port 3/1/1 egress |
          match "Root Pool : 1" post-lines 26 | match "Pool" post-lines 1
Root Pool : 1
|  Total             : 3978866 KB  Allocated        : 0 KB
|--- Mid Pool : 1
|     | Total              : 3978866 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |      Total            : 1989432 KB  Allocated        : 0 KB
*A:PE-1# show hs-pools port 3/1/2 egress | match "Root Pool : 1" post-lines 26 | match
 "Pool" post-lines 1
```

```
Root Pool : 1
|  Total               : 3978866 KB  Allocated          : 0 KB
|--- Mid Pool : 1
|      | Total             : 3978866 KB  Allocated        : 0 KB
|      |--- Std Port Class Pool : 1
|      |      Total             : 1989432 KB  Allocated        : 0 KB
```

- The egress rate is set on port 3/1/1 (which is a 10 Gb/s port) to 5 Gb/s. This reduces the size of the port class pools on port 3/1/1 to one-third of the mid pool size and increases the size of the port class pools on port 3/1/2 to two-thirds of the mid pool size, after which the egress rate is removed:

```
*A:PE-1# configure port 3/1/1 ethernet egress-rate 5000000
*A:PE-1# show hs-pools port 3/1/1 egress |
           match "Root Pool : 1" post-lines 26 | match "Pool" post-lines 1
Root Pool : 1
|  Total               : 3978866 KB  Allocated          : 0 KB
|--- Mid Pool : 1
|      | Total             : 3978866 KB  Allocated        : 0 KB
|      |--- Std Port Class Pool : 1
|      |      Total             : 1326288 KB  Allocated        : 0 KB
*A:PE-1# show hs-pools port 3/1/2 egress |
           match "Root Pool : 1" post-lines 26 | match "Pool" post-lines 1
Root Pool : 1
|  Total               : 3978866 KB  Allocated          : 0 KB
|--- Mid Pool : 1
|      | Total             : 3978866 KB  Allocated        : 0 KB
|      |--- Std Port Class Pool : 1
|      |      Total             : 2652576 KB  Allocated        : 0 KB
*A:PE-1# configure port 3/1/1 ethernet no egress-rate
```

- The egr-percentage-of-rate is set to 200% on port 3/1/1 to increase its port class pool to two-thirds of the mid pools size and reduce the port class pool on port 3/1/2 to one-third of the mid pools size, after which the egr-percentage-of-rate is removed:

```
*A:PE-1# configure port 3/1/1 modify-buffer-allocation-rate egr-percentage-of-rate 200
*A:PE-1# show hs-pools port 3/1/1 egress |
           match "Root Pool : 1" post-lines 26 | match "Pool" post-lines 1
Root Pool : 1
|  Total               : 3978866 KB  Allocated          : 0 KB
|--- Mid Pool : 1
|      | Total             : 3978866 KB  Allocated        : 0 KB
|      |--- Std Port Class Pool : 1
|      |      Total             : 2652576 KB  Allocated        : 0 KB
*A:PE-1# show hs-pools port 3/1/2 egress |
           match "Root Pool : 1" post-lines 26 | match "Pool" post-lines 1
Root Pool : 1
|  Total               : 3978866 KB  Allocated          : 0 KB
|--- Mid Pool : 1
|      | Total             : 3978866 KB  Allocated        : 0 KB
|      |--- Std Port Class Pool : 1
|      |      Total             : 1326288 KB  Allocated        : 0 KB
*A:PE-1# configure port 3/1/1 modify-buffer-allocation-rate
                          no egr-percentage-of-rate
```

- Standard port class pool 2 is parented to the mid pool with a **port-bw-weight** set to 2. The **port-bw-weight** of port class pool 1 is the default of 1. This causes this port mid pool size to be shared in a 1:2 ratio between port class pool 1 and 2 on both ports 3/1/1 and 3/1/2 (only port 3/1/1 is shown). The total port class pool size is shown in the second step preceding, that is, 1989432 kbytes.

```
*A:PE-1# configure qos hs-port-pool-policy "hs-port-pool-policy-test"
```

```
*A:PE-1>config>qos>hs-port-pool-policy# std-port-class-pools
                                      class-pool 2 parent-mid-pool 1
*A:PE-1>config>qos>hs-port-pool-policy# std-port-class-pools
                                      class-pool 2 allocation port-bw-weight 2
*A:PE-1>config>qos>hs-port-pool-policy# exit all
*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 1" post-lines 47 | match "Pool" post-lines 1
Root Pool : 1
|  Total              : 3978866 KB  Allocated        : 0 KB
|--- Mid Pool : 1
|     | Total              : 3978866 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |     Total              : 663144 KB   Allocated        : 0 KB
|     |--- Std Port Class Pool : 2
|     |     Total              : 1326288 KB  Allocated        : 0 KB
```

- "The two port class pools 1 and 2 on port 3/1/1 are modified to use an **explicit-percent** of 40% and 60%, respectively:

```
*A:PE-1# configure qos hs-port-pool-policy "hs-port-pool-policy-test"
*A:PE-1>config>qos>hs-port-pool-policy# std-port-class-pools
                                      class-pool 1 allocation explicit-percent 40
*A:PE-1>config>qos>hs-port-pool-policy# std-port-class-pools
                                      class-pool 2 allocation explicit-percent 60
*A:PE-1>config>qos>hs-port-pool-policy# exit all
*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 1" post-lines 47 | match "Pool" post-lines 1
Root Pool : 1
|  Total              : 3978866 KB  Allocated        : 0 KB
|--- Mid Pool : 1
|     | Total              : 3978866 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |     Total              : 1591546 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 2
|     |     Total              : 2387318 KB  Allocated        : 0 KB
```

To assist with sizing the buffer pools, each pool has a high watermark, which can be displayed using the **show hs-pools** command and cleared using the following commands:

```
clear card <slot-number> fp <[1..2]> hs-pool high-water-mark
clear card <slot-number> fp <[1..2]> hs-pool high-water-mark mid-pool <[1..16]>
clear card <slot-number> fp <[1..2]> hs-pool high-water-mark root-pool <[1..16]>
clear card <slot-number> fp <[1..2]> hs-pool high-water-mark system
clear port <port-id> hs-pool high-water-mark
                              { [standard <1..6>] | [alternate <1..6>] }
```

The HSQ ingress and non-HSQ egress line cards support a **stable-pool-sizing** command under **card>fp**, which avoids pool sizes changing when MDAs and ports are configured.

An equivalent effect can be achieved at the egress of an HSQ by creating two sets of root pools and two sets of mid pools in the **hs-pool-policy** applied to the IOM under **card>fp>egress**. The first set of mid pools parent to the first set of root pools and the second set of mid pools parent to the second set of root pools. Then create two hs-port-pool-policies: one applied to the ports on the first MDA with its port class pools parented to the first set of mid pools and the other applied to the ports on the second MDA with its port class pools parented to the second set of mid pools. This provides deterministic pool sizing independent of MDAs being inserted or removed.

Further control at the port class level can be obtained by using port class pool **explicit-percent** based sizing to eliminate the effect of changing port states, including bandwidth changes.

The root pools cannot oversubscribe the user-provisioned buffers, but the mid pools can oversubscribe their root pool and the port class pools can oversubscribe their mid pool.

The following configuration and output shows the oversubscription possibilities.

The default HS pool policy is applied to the HSQ IOM, so root pools 1 and 2 are allocated 75% and 25% of the user-provisioned buffers:

```
*A:PE-1>config>qos>hs-pool-policy# info detail |
                match expression " root-pool 1$| root-pool 2$" post-lines 1
            root-pool 1
                allocation-weight 75
            root-pool 2
                allocation-weight 25

*A:PE-1# show hs-pools 3 fp 1 egress | match "Root Pool" post-lines 1
Root Pool : 1
|  Total             : 2984148 KB  Allocated        : 0 KB
Root Pool : 2
|  Total             : 994716 KB   Allocated        : 0 KB
```

If a new HS pool policy is applied to this IOM, only one root pool is allocated buffers with an allocation weight of 100:

```
*A:PE-1>config>qos>hs-pool-policy$ info detail |
                match expression " root-pool 1$" post-lines 1
            root-pool 1
                allocation-weight 100

*A:PE-1# show hs-pools port 3/1/1 egress | match "Root Pool : 1" post-lines 1
Root Pool : 1
|  Total             : 3978866 KB  Allocated        : 0 KB
```

Root pool 16 is configured with the same allocation weight of 100, causing the two root pools to share the available user-provisioned buffers:

```
*A:PE-1>config>qos>hs-pool-policy# info detail |
                match expression " root-pool 1$| root-pool 16$" post-lines 1
            root-pool 1
                allocation-weight 100
            root-pool 16
                allocation-weight 100

*A:PE-1# show hs-pools port 3/1/1 egress | match "Root Pool : 1" post-lines 1
Root Pool : 1
|  Total             : 1989432 KB  Allocated        : 0 KB
Root Pool : 16
|  Total             : 1989432 KB  Allocated        : 0 KB
```

Mid pool 16 is parented to root pool 16 with an allocation percent of 100, so it has the same number of allocated buffers as root pool 16:

```
*A:PE-1>config>qos>hs-pool-policy# info detail |
      match expression " root-pool 16$| mid-pool 16$" post-lines 2 |
      match invert-match slope
            root-pool 16
                allocation-weight 100
            mid-pool 16
                parent-root-pool 16
                allocation-percent 100.00
```

```
*A:PE-1# show hs-pools port 3/1/1 egress |
              match "Root Pool : 16" post-lines 19 | match "Pool" post-lines 1
Root Pool : 16
|  Total               : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|     | Total             : 1989432 KB  Allocated        : 0 KB
```

Mid pool 15 is also parented to root pool 16 with an allocation percent of 100, causing both mid pools
15 and 16 to have the same number of allocated buffers as root pool 16; therefore, the root pool is
oversubscribed two times:

```
*A:PE-1>config>qos>hs-pool-policy# info detail |
    match expression " root-pool 16$| mid-pool 15$| mid-pool 16$" post lines 2 |
    match invert-match slope
              root-pool 16
                  allocation-weight 100
              mid-pool 15
                  parent-root-pool 16
                  allocation-percent 100.00
              mid-pool 16
                  parent-root-pool 16
                  allocation-percent 100.00

*A:PE-1# show hs-pools port 3/1/1 egress |
              match "Root Pool : 16" post-lines 19 |
              match "Pool" post-lines 1
Root Pool : 16
|  Total               : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|     | Total             : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|     | Total             : 1989432 KB  Allocated        : 0 KB
```

An HS port pool policy is applied to port 3/1/1 with standard port class pool 1 parented to mid pool 16 with
an **allocation port-bw-weight** of 1. This port class pool has the same number of allocated buffers as mid
pool 16:

```
*A:PE-1>config>qos>hs-port-pool-policy>std-port-class-pools# info
----------------------------------------------
              class-pool 1
                  parent-mid-pool 16
              exit

*A:PE-1# show hs-pools port 3/1/1 egress |
              match "Root Pool : 16" post-lines 26 | match "Pool" post-lines 1
Root Pool : 16
|  Total               : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|     | Total             : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|     | Total             : 1989432 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |     Total             : 1989432 KB  Allocated        : 0 KB
```

The **port-bw-oversub-factor** is set to 2 for mid pool 16 in the HS pool policy, after which the size of mid
pool 16 does not change. However, its apparent size for the calculation of its port class pool doubles, which
causes the size of port class pool 1 to be twice that of mid pool 16, thereby oversubscribing it:

```
*A:PE-1>config>qos>hs-pool-policy# info
```

```
                -----------------------------------------------
                    root-tier
                        root-pool 16
                            allocation-weight 100
                        exit
                    exit
                    mid-tier
                        mid-pool 15
                            parent-root-pool 16
                            allocation-percent 100.00
                        exit
                        mid-pool 16
                            parent-root-pool 16
                            allocation-percent 100.00
                            port-bw-oversub-factor 2
                        exit
                    exit

*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 16" post-lines 26 |
            match "Pool" post-lines 1
Root Pool : 16
|  Total            : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|    | Total            : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|    | Total            : 1989432 KB  Allocated        : 0 KB
|    |--- Std Port Class Pool : 1
|    |     Total            : 3978864 KB  Allocated        : 0 KB
```

A second standard port class pool, pool 2, on port 3/1/1 is parented to mid pool 16. The two port class pools share the buffer allocation equivalent to two times that of mid pool 16:

```
*A:PE-1>config>qos>hs-port-pool-policy>std-port-class-pools# info
                -----------------------------------------------
                    class-pool 1
                        parent-mid-pool 16
                    exit
                    class-pool 2
                        parent-mid-pool 16
                    exit
*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 16" post-lines 33  |
            match "Pool" post-lines 1
Root Pool : 16
|  Total            : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|    | Total            : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|    | Total            : 1989432 KB  Allocated        : 0 KB
|    |--- Std Port Class Pool : 1
|    |     Total            : 1989432 KB  Allocated        : 0 KB
|    |--- Std Port Class Pool : 2
|    |     Total            : 1989432 KB  Allocated        : 0 KB
```

The proportion of buffers available to the port class pools can be modified by configuring their **allocation port-bw-weight**. If the **allocation port-bw-weight** of port class pool 2 is set to 2, the port class pools will be allocated buffers in a 2:1 ratio:

```
*A:PE-1>config>qos>hs-port-pool-policy>std-port-class-pools# info
                -----------------------------------------------
                    class-pool 1
```

```
                    parent-mid-pool 16
                exit
                class-pool 2
                    parent-mid-pool 16
                    allocation port-bw-weight 2
                exit

*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 16" post-lines 33 |
            match "Pool" post-lines 1
Root Pool : 16
|   Total              : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|     | Total              : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|     | Total              : 1989432 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |     Total              : 1326288 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 2
|     |     Total              : 2652576 KB  Allocated        : 0 KB
```

A third standard port class pool, pool 3, is parented to mid pool 16 with an **allocation port-bw-weight** of 2. Port class pools 1, 2, and 3 share the oversubscribed mid pool 16 size in a ratio of 1:2:2:

```
*A:PE-1>config>qos>hs-port-pool-policy>std-port-class-pools# info
----------------------------------------------
                class-pool 1
                    parent-mid-pool 16
                exit
                class-pool 2
                    parent-mid-pool 16
                    allocation port-bw-weight 2
                exit
                class-pool 3
                    parent-mid-pool 16
                    allocation port-bw-weight 2
                exit

*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 16" post-lines 40 |
            match "Pool" post-lines 1
Root Pool : 16
|   Total              : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|     | Total              : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|     | Total              : 1989432 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 1
|     |     Total              : 795772 KB   Allocated        : 0 KB
|     |--- Std Port Class Pool : 2
|     |     Total              : 1591544 KB  Allocated        : 0 KB
|     |--- Std Port Class Pool : 3
|     |     Total              : 1591544 KB  Allocated        : 0 KB
```

Standard port class pool 1, 2, and 3 are now configured with an allocation **explicit-percent** of 80%, 60%, and 60% respectively. This allocates these percentages of mid pool 16 real size to port class pools 1, 2, and 3, which oversubscribed the mid pool by 100%. The mid pool 16 oversubscription factor is not applied.

```
*A:PE-1>config>qos>hs-port-pool-policy>std-port-class-pools# info
----------------------------------------------
                class-pool 1
                    parent-mid-pool 16
```

```
                allocation explicit-percent 80.00
            exit
            class-pool 2
                parent-mid-pool 16
                allocation explicit-percent 60.00
            exit
            class-pool 3
                parent-mid-pool 16
                allocation explicit-percent 60.00
            exit

*A:PE-1# show hs-pools port 3/1/1 egress |
            match "Root Pool : 16" post-lines 40 |
            match "Pool" post-lines 1
Root Pool : 16
|  Total                 : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 15
|    | Total             : 1989432 KB  Allocated        : 0 KB
|--- Mid Pool : 16
|    | Total             : 1989432 KB  Allocated        : 0 KB
|    |--- Std Port Class Pool : 1
|    |      Total             : 1591544 KB  Allocated        : 0 KB
|    |--- Std Port Class Pool : 2
|    |      Total             : 1193658 KB  Allocated        : 0 KB
|    |--- Std Port Class Pool : 3
|    |      Total             : 1193658 KB  Allocated        : 0 KB
```

Shaping and Scheduling

HSQ Queue Groups

Each configuration uses the same four queues:

- Queue 7 at scheduling class 5
- Queue 6 at scheduling class 4
- Queues 1 and 2 in WRR group at scheduling class 1

The **low-burst-max-class** is configured to be class 1, to match the scheduling class of WRR group 1. This results in queues 1 and 2 being subject to the low burst limit threshold.

HSQ queue group queues can be attached to scheduling classes or to WRR groups, which can then be attached to a scheduling class. The eight queues and two WRR groups in an HSQ queue group can also be unattached (not attached to any scheduling class, or for queues, to any WRR group). This is configured using an **hs-attachment-policy**:

```
configure
    qos
        hs-attachment-policy <policy-name> [create]
            description <description-string>
            low-burst-max-class <class>
            queue <queue-id> sched-class <class-id>
            queue <queue-id> unattached
            queue <queue-id> wrr-group <wrr-group-id>
            wrr-group <group-id> sched-class <class-id>
```

Where (in order):

```
<policy-name>        : [32 chars max]
<description-string> : [80 chars max]
<class>              : [1..6]
<queue-id>           : [1..8]
<class-id>           : [1..6]
<wrr-group-id>       : [1..2]
<group-id>           : [1..2]
<class-id>           : [1..6]
```

When a queue or WRR group is unattached, the related queues discard all received packets.

When a queue is attached to a WRR group, the weight of that queue within the group is configured under the queue in the SAP egress QoS policy, network queue policy, and egress queue group template.

When a queue is attached to a WRR group, the following queue parameters are ignored. The corresponding configuration is applied to the entire WRR group in the SAP egress QoS policy for services, network queue policy for network interfaces, and egress queue group templates for both access and network egress queue group instances:

- **adaptation-rule**

- **hs-class-weight**

- **percent-rate**

- **rate**

A default **hs-attachment-policy** is created by the system and applied by default to all SAP egress QoS policy for services, to all network queue policy for network interfaces, and to all egress queue group templates for both access and network egress queue group instances. The default policy is not configurable.

```
*A:PE-1>config>qos# hs-attachment-policy "default"
*A:PE-1>config>qos>hs-attachment-policy# info detail
--------------------------------------------
            no description
            low-burst-max-class 6
            queue 1 wrr-group 1
            queue 2 wrr-group 1
            queue 3 wrr-group 1
            queue 4 sched-class 2
            queue 5 sched-class 3
            queue 6 sched-class 4
            queue 7 sched-class 5
            queue 8 sched-class 6
            wrr-group 1 sched-class 1
            wrr-group 2 unattached
```

> **Note:**
> Queues 1, 2, and 3 are attached by default to WRR group 1, so their configured rates are ignored.

To use a user-defined policy, a new HS attachment policy must be created and applied to the appropriate SAP egress QoS policy, network queue policy, or egress queue group template. A newly created policy has all queues and WRR groups unattached and the **low-burst-max-class** set to 6:

```
*A:PE-1# configure qos hs-attachment-policy hs-att-policy-new create
*A:PE-1>config>qos>hs-attachment-policy$ info detail
```

```
-----------------------------------------------
          no description
          low-burst-max-class 6
          queue 1 unattached
          queue 2 unattached
          queue 3 unattached
          queue 4 unattached
          queue 5 unattached
          queue 6 unattached
          queue 7 unattached
          queue 8 unattached
          wrr-group 1 unattached
          wrr-group 2 unattached
```

The hs-attachment-policy used for this example is as follows (queue 8 is reserved to be attached to scheduling class 6 for network protocol traffic, but is not used in this example):

```
*A:PE-1>config>qos# hs-attachment-policy "hs-att-pol-1"
*A:PE-1>config>qos>hs-attachment-policy# info detail
-----------------------------------------------
          no description
          low-burst-max-class 1
          queue 1 wrr-group 1
          queue 2 wrr-group 1
          queue 3 unattached
          queue 4 unattached
          queue 5 unattached
          queue 6 sched-class 4
          queue 7 sched-class 5
          queue 8 sched-class 6
          wrr-group 1 sched-class 1
          wrr-group 2 unattached
```

An HS attachment policy is shown as follows:

```
*A:PE-1# show qos hs-attachment-policy "hs-att-pol-1"


===============================================================================
HS Attachment Policy Information
===============================================================================
Policy Name          : hs-att-pol-1
Description          : (Not Specified)
Low Burst Max Class    : 1


-------------------------------------------------------------------------------
Queue                Scheduling Class            WRR Group
-------------------------------------------------------------------------------
1                    (Not-Applicable)            1
2                    (Not-Applicable)            1
3                    unattached                  unattached
4                    unattached                  unattached
5                    unattached                  unattached
6                    4                           (Not-Applicable)
7                    5                           (Not-Applicable)
8                    6                           (Not-Applicable)


-------------------------------------------------------------------------------
WRR Group            Scheduling Class
-------------------------------------------------------------------------------
1                    1
2                    unattached
===============================================================================
```

```
*A:PE-1#
```

It is also possible to show the associations for each policy:

```
*A:PE-1# show qos hs-attachment-policy "hs-att-pol-1" association

===============================================================================
HS Attachment Policy Information
===============================================================================
Policy Name            : hs-att-pol-1
Description            : (Not Specified)
Low Burst Max Class    : 1


-------------------------------------------------------------------------------
Associations
-------------------------------------------------------------------------------
Network-Queue Policy
-------------------------------------------------------------------------------
10

Sap-Egress Policy
-------------------------------------------------------------------------------
10
20

Egress Queue-Group Templates
-------------------------------------------------------------------------------
queue-group-1


-------------------------------------------------------------------------------
===============================================================================
```

HS Secondary Shapers

HS secondary shapers are only applicable to SAP egress in the context of this chapter (not to network egress or egress access and network queue group instances). However, because secondary shapers can also be used by subscribers, they are included with the generic configuration aspects.

Secondary shapers are aimed at providing QoS control for traffic forwarded to a specific downstream device, such as an access node. Multiple HS secondary shapers can be configured on an egress port.

An aggregate rate and per-scheduling class rates are configurable for each secondary shaper. A **low-burst-max-class** parameter is also available to provide granular control over the scheduling behavior of which queues (via a WRR group, if used) use the low burst limit threshold and which use the high burst limit threshold.

Secondary shapers are configured on each port under the **config>port>ethernet>egress** context:

```
configure
    port <port-id>
        ethernet
            egress
                hs-secondary-shaper <secondary-shaper-name>
                    description <description-string>
                    aggregate
                        low-burst-max-class <class>
                        rate <rate>
                    class <class-number>
                        rate <rate>
```

Where (in order):

```
<port-id>            : slot/mda/port
<secondary-shaper-*> : [32 chars max]
<description-string> : [80 chars max]
<class>              : [1..6]
<class-number>       : [1..6]
<rate>               : [1..100000000|max] Kbps
<rate>               : [1..100000000|max] Kbps
```

A default HS secondary shaper is applied to all egress HSQ ports with the rates set to **max** and the **low-burst-max-class** set to 6. It is possible to modify the configuration of the default HS secondary shaper.

```
*A:PE-1>config>port>ethernet>egress# hs-secondary-shaper "default"
*A:PE-1>config>port>ethernet>egress>hs-sec-shaper# info detail
---------------------------------------------
                        no description
                        aggregate
                            rate max
                            low-burst-max-class 6
                        exit
                        class 1
                            rate max
                        exit
                        class 2
                            rate max
                        exit
                        class 3
                            rate max
                        exit
                        class 4
                            rate max
                        exit
                        class 5
                            rate max
                        exit
                        class 6
                            rate max
                        exit
```

An HS secondary shaper is configured on port 3/1/2 with a rate of 100 Mb/s for scheduling class 1 and a **low-burst-max-class** set to 1:

```
*A:PE-1>config>port>ethernet>egress# hs-secondary-shaper "hs-sec-shaper-1"
*A:PE-1>config>port>ethernet>egress>hs-sec-shaper# info detail
---------------------------------------------
                        no description
                        aggregate
                            rate max
                            low-burst-max-class 1
                        exit
                        class 1
                            rate 100000
                        exit
                        class 2
                            rate max
                        exit
                        class 3
                            rate max
                        exit
                        class 4
                            rate max
```

```
                        exit
                        class 5
                            rate max
                        exit
                        class 6
                            rate max
                        exit
```

This is shown, in this case with its associations, as follows:

```
*A:PE-1# show port 3/1/2 hs-secondary-shaper "hs-sec-shaper-1" associations

===============================================================================
Ethernet Port 3/1/2 Egress HS Secondary Shaper Information
===============================================================================
Policy Name        : hs-sec-shaper-1
Description        : (Not Specified)
Rate               : max
Low Burst Max Class: 1


-------------------------------------------------------------------------------
Class                                   Rate
-------------------------------------------------------------------------------
1                                       100000 Kbps
2                                       max
3                                       max
4                                       max
5                                       max
6                                       max
-------------------------------------------------------------------------------


-------------------------------------------------------------------------------
Service Associations
-------------------------------------------------------------------------------
Service ID              Service Type            SAP
-------------------------------------------------------------------------------
1                       IES                     3/1/2:1
-------------------------------------------------------------------------------


-------------------------------------------------------------------------------
Subscriber Associations
-------------------------------------------------------------------------------
Subscriber ID
-------------------------------------------------------------------------------
No Subscriber Associations Found.
-------------------------------------------------------------------------------
===============================================================================
*A:PE-1#
```

All HS secondary shapers on a port can be shown using the same command, but omitting the shaper name and following parameter.

It is also possible to show the forwarding statistics related to an HS secondary shaper:

```
*A:PE-1# show port 3/1/2 hs-secondary-shaper "hs-sec-shaper-1" statistics

===============================================================================
Ethernet Port 3/1/2 Egress HS Secondary Shaper Information
===============================================================================
Policy Name        : hs-sec-shaper-1


-----------------------------------------------------------------------
Statistics Information
```

```
------------------------------------------------------------------------

------------------------------------------------------------------------
                        Packets                 Octets
Class 1
    Forwarded        : 4000                      592000

Class 2
    Forwarded        : 0                         0

Class 3
    Forwarded        : 0                         0

Class 4
    Forwarded        : 1000                      148000

Class 5
    Forwarded        : 1000                      148000

Class 6
    Forwarded        : 0                         0

Aggregate
    Forwarded        : 6000                      888000

------------------------------------------------------------------------
========================================================================
*A:PE-1#
```

These statistics are cleared using the following command:

```
clear port 3/1/2 hs-secondary-shaper "hs-sec-shaper-1" statistics
```

The SAP egress queues in an HSQ queue group can be associated with a secondary shaper by configuring an egress queue override under the SAP and specifying the name of the secondary shaper. In addition, when using PW-SAPs, an HS secondary shaper can be applied to the egress of a PW port to control the PW-SAP egress traffic over that PW. See the SAP Egress section for configuration details.

If the user configures an HS secondary shaper on a port, the system instantiates a default primary shaper for that secondary shaper (which is used by all HSQ queue groups sending traffic to the secondary shaper) when the first egress SAP or PW-SAP is associated with that HS secondary shaper.

The current traffic rates through the secondary shapers on a port are shown, as follows, where the default interval is 1 second:

```
show qos hs-scheduler-hierarchy port <port-id>
          [hs-secondary-shaper <shaper-name>] [interval <time-in-seconds>]
show qos hs-scheduler-hierarchy port <port-id>
          [interval <time-in-seconds>] hs-secondary-shapers
```

The output shows the current aggregate traffic rate and the current traffic rate for each scheduling class:

```
*A:PE-1# show qos hs-scheduler-hierarchy port 3/1/2
                    hs-secondary-shaper "hs-sec-shaper-1"

========================================================================
Hs Scheduler Hierarchy Information
========================================================================
Hs Sched Policy Name        : default

Port Max-Rate : 137 Mbps
```

```
Hs-Sec-Shaper:hs-sec-shaper-1 Agg-Rate : 57516 Kbps

Scheduler Priority 6
  Scheduler Class 6  Rate : 0 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 6 Rate : 0 Kbps

Scheduler Priority 5
  Scheduler Class 5  Rate : 22 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 5 Rate : 11422 Kbps

Scheduler Priority 4
  Scheduler Class 4  Rate : 45 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 4 Rate : 22797 Kbps

Scheduler Priority 3
  Scheduler Class 3  Rate : 0 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 3 Rate : 0 Kbps

Scheduler Priority 2
  Scheduler Class 2  Rate : 0 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 2 Rate : 0 Kbps

Scheduler Priority 1
  Scheduler Class 1  Rate : 69 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 1 Rate : 23296 Kbps
===============================================================================
*A:PE-1#
```

Ports

A single HS scheduler policy can be applied to an egress HSQ port to configure an aggregate rate (**max-rate**) and per-scheduling class rates on that port. In addition, contiguous scheduling classes can be configured with weights in a WRR group, which can also be configured with a rate. The WRR group is scheduled at the scheduling class of its highest member scheduling class. The **max-rate** caps the scheduling class and group rates if its rate is lower.

The rates configured within an HS scheduler policy are applicable to all types of traffic (SAP egress, network egress, and egress queue group instances) exiting that port.

An HS scheduler policy is configured as follows:

```
configure
    qos
        hs-scheduler-policy <policy-name> [create]
            description <description-string>
            group <group-id> rate <rate>
            max-rate <rate>
            scheduling-class <class-id> group <group-id> [weight <weight-in-group>]
            scheduling-class <class-id> rate <rate>
```

Where (in order):

```
<policy-name>        : [32 chars max]
<description-string> : [80 chars max]
<group-id>           : [1]
<rate>               : [1..100000|max] Mbps
<rate>               : [1..100000|max] Mbps
<class-id>           : [1..6]
<group-id>           : [1]
<weight-in-group>    : [1..127]
```

```
<rate>                : [1..100000|max] Mbps
```

Note:

The rates configured in an HS scheduler policy are in Mb/s (not kb/s).

A default HS scheduler policy is applied to all egress HSQ ports and all its rates are set to **max**. It is not possible to modify the default HS scheduler policy.

```
*A:PE-1>config>qos# hs-scheduler-policy "default"
*A:PE-1>config>qos>hs-scheduler-policy# info detail
----------------------------------------------
            description "Default hs scheduler QoS policy"
            max-rate max
            group 1 rate max
            scheduling-class 1 rate max
            scheduling-class 2 rate max
            scheduling-class 3 rate max
            scheduling-class 4 rate max
            scheduling-class 5 rate max
            scheduling-class 6 rate max
```

A newly created HS scheduler policy has the same configuration as the default policy.

An HS scheduler policy with a rate of 5 Gb/s for scheduling class 1 is applied to port 3/1/1:

```
*A:PE-1# configure qos
*A:PE-1>config>qos# hs-scheduler-policy "hs-sched-pol-1"
*A:PE-1>config>qos>hs-scheduler-policy# info detail
----------------------------------------------
            no description
            max-rate max
            group 1 rate max
            scheduling-class 1 rate 5000
            scheduling-class 2 rate max
            scheduling-class 3 rate max
            scheduling-class 4 rate max
            scheduling-class 5 rate max
            scheduling-class 6 rate max
----------------------------------------------
*A:PE-1>config>qos>hs-scheduler-policy# exit all
*A:PE-1# configure port 3/1/1 ethernet egress hs-scheduler-policy "hs-sched-pol-1"
*A:PE-1#
```

The preceding policy is shown as follows:

```
*A:PE-1# show qos hs-scheduler-policy "hs-sched-pol-1"

===============================================================================
HS Scheduler Policy Information
===============================================================================
Policy Name           : hs-sched-pol-1
Description           : (Not Specified)
Max Rate              : max

-------------------------------------------------------------------------------
Scheduling Class    Rate               Group              Weight in Group
-------------------------------------------------------------------------------
1                   5000 Mbps          0                  1
2                   max                0                  1
3                   max                0                  1
4                   max                0                  1
```

```
5                    max                0                    1
6                    max                0                    1
-------------------------------------------------------------------------
Group                Rate
-------------------------------------------------------------------------
1                    max
=========================================================================
*A:PE-1#
```

The ports associated with this policy are as follows:

```
*A:PE-1# show qos hs-scheduler-policy "hs-sched-pol-1" association

===============================================================================
HS Scheduler Policy Information
===============================================================================
Policy Name              : hs-sched-pol-1
Description              : (Not Specified)
Max Rate                 : max


    -------------------------------------------------------------------------
Port Ethernet Egress Associations
    -------------------------------------------------------------------------
3/1/1
    -------------------------------------------------------------------------
===============================================================================
*A:PE-1#
```

The current traffic rates through the port aggregate, scheduling class, and WRR group shapers are shown,
as follows, where the default interval is 1 second:

```
show qos hs-scheduler-hierarchy port <port-id> [interval <time-in-seconds>]
        queue-group <queue-group-name> instance <instance-id> {access|network}
show qos hs-scheduler-hierarchy sap <sap-id> egress [interval <time-in-seconds>]
show qos hs-scheduler-hierarchy subscriber <sub-ident> egress
                                                [interval <time-in-seconds>]
```

The output shows the current aggregate traffic rate and the current traffic rates for each scheduling class:

```
*A:PE-1# show qos hs-scheduler-hierarchy port 3/1/1

===============================================================================
Hs Scheduler Hierarchy Information
===============================================================================
Hs Sched Policy Name         : hs-sched-pol-1

Port Max-Rate : 659 Mbps

Scheduler Priority 6
  Scheduler Class 6  Rate : 0 Mbps

Scheduler Priority 5
  Scheduler Class 5  Rate : 127 Mbps

Scheduler Priority 4
  Scheduler Class 4  Rate : 254 Mbps

Scheduler Priority 3
  Scheduler Class 3  Rate : 0 Mbps

Scheduler Priority 2
  Scheduler Class 2  Rate : 0 Mbps
```

```
Scheduler Priority 1
  Scheduler Class 1  Rate : 277 Mbps
===============================================================================
*A:PE-1#
```

The HS scheduler policy parameters can be overridden under the port policy configuration:

```
configure
    port <port-id>
        ethernet
            egress
                hs-scheduler-overrides [create]
                    group <group-id> rate <rate>
                    max-rate <rate>
                    scheduling-class <class> rate <rate>
                    scheduling-class <class> weight <weight-in-group>
```

Where (in order):

```
<port-id>          : slot/mda/port
<group-id>         : [1..1]
<rate>             : [1..100000|max] Mbps
<rate>             : [1..100000|max] Mbps
<class>            : [1..6]
<rate>             : [1..100000|max] Mbps
<class>            : [1..6]
<weight-in-group>  : [1..127]
```

SAP Egress

A SAP configured on an HSQ IOM port uses an HSQ queue group for its egress queues. This occurs automatically and dedicates one HSQ queue group, so eight egress queues, to each SAP egress. Only the queues to be used need to be configured within the SAP egress QoS policy.

The operation of classification, policing, and marking within a SAP egress QoS policy when applied to a SAP on an HSQ port is unchanged. For example, it is possible to use egress policers and direct the post-policer traffic to either a local HSQ queue group queue or to an HSQ queue group queue in an access egress queue group instance.

Most of the commands in a SAP egress QoS policy apply to the HSQ egress SAPs, with the following exceptions (including their SAP egress related overrides) being ignored:

- HSMDA commands
- **parent-location sla**
- Policer commands
  - **policers-hqos-manageable**
  - **policer scheduler-parent**
- Queue related
  - **adaptation-rule cir <adaptation-rule>**
  - **adv-config-policy**
  - **avg-frame-overhead**

- – **burst-limit**

- – **cbs**

- – **drop tail**

- – **parent**

- – **percent-rate cir**

- – **percent-rate local-limit**

- – **pool**

- – **port-parent**

- – **rate cir**

- – **wred-queue**

- Subscriber commands

  - – **dynamic-policer** commands

  - – **sub-insert-shared-pccrule**

The following SAP commands are not configurable on HSQ SAPs:

- **"ingress qos shared-queuing**

- **"ingress qos multipoint-shared**

- **"egress agg-rate limit-unused-bandwidth**

- **"egress agg-rate queue-frame-based-accounting**

- **"multi-service-site**

As mentioned, an HS attachment policy is applied to the SAP egress QoS policy to define the attachment of the queues to scheduling classes or WRR groups, with the WRR group then being attached to a scheduling class:

```
configure
    qos
        sap-egress <policy-id>
            hs-attachment-policy <policy-name>
```

Where:

```
<policy-id>          : [1..65535]|<name:64 char max>
<policy-name>        : [32 chars max]
```

When queues are attached to one of the HSQ queue group WRR groups, the relative weight of each queue within the group is configured under the queue, with the default weight being 1:

```
configure
    qos
        sap-egress <policy-id>
            queue <queue-id>
                hs-wrr-weight <weight>
```

Where:

```
<policy-id>            : [1..65535]|<name:64 char max>
```

```
<queue-id>          : [1..8]
<weight>            : [1..127]
```

The rate-related configuration of the two WRR groups in the HSQ queue group is defined within the SAP egress QoS policy, as follows, with the defaults being **rate max** and **adaptation-rule closest**:

```
configure
    qos
        sap-egress <policy-id>
            hs-wrr-group <group-id>
                adaptation-rule [pir <adaptation-rule>]
                percent-rate <percent>
                rate <rate>
```

Where:

```
<policy-id>         : [1..65535]|<name:64 char max>
<group-id>          : [1..2]
<adaptation-rule>   : max|min|closest
<percent>           : [0.01..100.00]
<rate>              : [1..2000000000|max] Kbps
```

The **percent-rate** configured within the **hs-wrr-group** and under a queue is relative to the port rate, so is equivalent to the queue **port-limit** and includes both the **egress-rate** and HS scheduler policy **max-rate**, if configured.

The SAP egress queue default slope policy is _tmnx_hs_default_. A user-defined slope policy can be configured on a queue, as follows:

```
configure
    qos
        sap-egress <policy-id>
            queue <queue-id>
                hs-wred-queue [policy <slope-policy-name>]
```

Where:

```
<policy-id>         : [1..65535]|<name:64 char max>
<queue-id>          : [1..8]
<slope-policy-name> : [32 chars max]
```

The **highplus-slope** and **time-average-factor** in the applied slope policy are ignored on HSQ queue group queues.

By default, SAP egress HSQ queue group queues use buffers from the standard port class pools on their associated port. Each queue can be configured to use the port alternative class pools, as follows:

```
configure
    qos
        sap-egress <policy-id>
            queue <queue-id>
                hs-alt-port-class-pool
```

Where:

```
<policy-id>         : [1..65535]|<name:64 char max>
<queue-id>          : [1..8]
```

WRR group scheduling between queues and WRR groups in different HSQ queue groups is available at a primary shaper scheduling class. This is configured within a SAP egress QoS policy, as follows:

```
configure
    qos
        sap-egress <policy-id>
            hs-wrr-group <group-id>
                hs-class-weight <weight>
            queue <queue-id>
                hs-class-weight <weight>
```

Where (in order):

```
<policy-id>         : [1..65535]|<name:64 char max>
<group-id>          : [1..2]
<weight>            : 1|2|4|8
```

The **hs-class-weight** parameter under the **queue** or **hs-wrr-group** statement specifies the relative weight of the respective **queue** or **hs-wrr-group** for scheduling opportunities when their parent primary shaper scheduling class is serviced. By default, the **hs-class-weight** is 1.

> **Note:**
> This parameter should not be confused with the **hs-wrr-weight** parameter, which specifies the relative weights of different queues within the same HSQ queue group WRR group.

The HSQ queue group aggregate rate is applied to a SAP egress using the **agg-rate rate** command:

```
configure
    service
        {ipipe <service-id>|epipe <service-id>|vpls <service-id>|
         ies <service-id> interface <ip-int-name>|
         vprn <service-id> interface <ip-int-name>}
            sap
                egress
                    agg-rate
                        rate <kilobits-per-second>
```

Where:

```
<service-id>        : [1..2147483647]|<svc-name:64 char max>
<ip-int-name>       : [32 chars max]  (must start with a letter)
<kilobits-per-seco*> : [1..3200000000|max] Kbps
```

The following HSQ-specific overrides are available under a SAP egress corresponding to the preceding commands:

```
configure
    service
        {ipipe <service-id>|epipe <service-id>|vpls <service-id>|
         ies <service-id> interface <ip-int-name>|
         vprn <service-id> interface <ip-int-name>}
            sap
                egress
                    queue-override
                        hs-secondary-shaper <policy-name>
                        hs-wrr-group <group-id> [create]
                        hs-wrr-group <group-id> class-weight <weight>
                        hs-wrr-group <group-id> percent-rate <percent>
```

```
                        hs-wrr-group <group-id> rate <rate>
                        queue <queue-id> hs-class-weight <weight>
                        queue <queue-id> hs-wred-queue policy <slope-policy-name>
                        queue <queue-id> hs-wrr-weight <weight>
```

Where (in order):

```
<service-id>        : [1..2147483647]|<svc-name:64 char max>
<ip-int-name>       : [32 chars max]  (must start with a letter)
<policy-name>       : [32 chars max]
<group-id>          : [1..2]
<weight>            : 1|2|4|8
<percent>           : [0.01..100.00]
<rate>              : [1..2000000000|max] Kbps
<queue-id>          : [1..8]
<weight>            : 1|2|4|8
<slope-policy-name> : [32 chars max]
<weight>            : [1..127]
```

> **Note:**
> Queue depth monitoring is supported for SAP egress HSQ queue group queues. This is enabled by configuring the queue override **monitor-depth** command under SAP egress with the associated **show** command output displaying buffer occupancy in ranges of 10% of the queue depth for each configured queue.

An additional SAP egress override is provided to redirect the traffic from an HSQ queue group to a user-configured secondary shaper:

```
configure
    service
        {ipipe <service-id>|epipe <service-id>|vpls <service-id>|
         ies <service-id> interface <ip-int-name>|
         vprn <service-id> interface <ip-int-name>}
            sap
                egress
                    queue-override
                        hs-secondary-shaper <policy-name>
```

Where:

```
<service-id>        : [1..2147483647]|<svc-name:64 char max>
<ip-int-name>       : [32 chars max]  (must start with a letter)
<policy-name>       : [32 chars max]
```

When using pseudowire SAPs (PW-SAPs), an HS secondary shaper can be configured under the SDP binding to apply QoS control to the PW used by the SAPs, as follows:

```
configure
    service
        sdp <sdp-id>
            binding
                pw-port <pw-port-id>
                    egress
                        shaper
                        pw-sap-secondary-shaper <pw-sap-sec-shaper-name>
```

Where:

```
<sdp-id>            : [1..32767]
<pw-port-id>        : [1..32767]
<pw-sap-sec-shaper*> : [32 chars max]
```

When the first egress SAP or PW-SAP is associated with a user-configured HS secondary shaper, the system instantiates a default primary shaper for that secondary shaper, which is used by all HSQ queue groups sending traffic to that secondary shaper.

An IES interface SAP is configured with the HSQ queue group having an aggregate rate of 50 Mb/s (using the **agg-rate rate** command). The SAP egress traffic is directed to an HS secondary shaper, which is applied to port 3/1/2:

```
*A:PE-1>config>service>ies# info
--------------------------------------------
          description "HSQ egress SAP queues"
          interface "PE-1-IES-1" create
              address 192.168.11.1/30
              sap 3/1/2:1 create
                  egress
                      qos 10
                      queue-override
                          hs-secondary-shaper "hs-sec-shaper-1"
                      exit
                      agg-rate
                          rate 50000
                      exit
                  exit
              exit
          exit
          no shutdown
```

The SAP egress QoS policy contains the applied HS attachment policy described for the queue attachment. Queue 1 is configured with a WRR weight of 2. Rates of 20 Mb/s and 10 Mb/s are configured on queues 6 and 7, respectively. WRR group 1 is configured with a rate of 40 Mb/s. DSCP values are used to classify the egress traffic to the forwarding classes mapped to the queues:

```
*A:PE-1>config>qos# sap-egress 10
*A:PE-1>config>qos>sap-egress# info
--------------------------------------------
          hs-attachment-policy "hs-att-pol-1"
          queue 1 create
              hs-wrr-weight 2
          exit
          queue 2 create
          exit
          queue 6 create
              rate 20000
          exit
          queue 7 create
              rate 10000
          exit
          hs-wrr-group 1
              rate 40000
          exit
          fc af create
              queue 2
          exit
          fc ef create
              queue 6
```

```
            exit
            fc h1 create
                queue 7
            exit
            dscp cs1 fc "af"
            dscp be fc "be"
            dscp cs2 fc "ef"
            dscp cs3 fc "h1"
```

The queue information is shown as follows:

```
*A:PE-1# show hs-pools port 3/1/2 egress sap 3/1/2:1 |
              match "Queue Information" pre-lines 1 post-lines 40
-------------------------------------------------------------------------------
Queue Information
-------------------------------------------------------------------------------
Queue Name        : 1->3/1/2:1->1
FC Map            : be l2 l1 h2 nc
Admin PIR         : 40000              Oper PIR           : 0
Admin MBS         : 64 KB              Oper MBS           : 64 KB
HS Wrr Group      : 1
HS Wrr Class Weight: 1                 HS Wrr Weight      : 2
Depth             : 0
HS Class          : 1                  HS Alt Port Class Pool : No
HS Slope Policy   : _tmnx_hs_default

Queue Name        : 1->3/1/2:1->2
FC Map            : af
Admin PIR         : 40000              Oper PIR           : 0
Admin MBS         : 64 KB              Oper MBS           : 64 KB
HS Wrr Group      : 1
HS Wrr Class Weight: 1                 HS Wrr Weight      : 1
Depth             : 0
HS Class          : 1                  HS Alt Port Class Pool : No
HS Slope Policy   : _tmnx_hs_default

Queue Name        : 1->3/1/2:1->6
FC Map            : ef
Admin PIR         : 20000              Oper PIR           : 20000
Admin MBS         : 64 KB              Oper MBS           : 64 KB
HS Wrr Group      : (not-applicable)
HS Wrr Class Weight: 1                 HS Wrr Weight      : 0
Depth             : 0
HS Class          : 4                  HS Alt Port Class Pool : No
HS Slope Policy   : _tmnx_hs_default

Queue Name        : 1->3/1/2:1->7
FC Map            : h1
Admin PIR         : 10000              Oper PIR           : 10000
Admin MBS         : 64 KB              Oper MBS           : 64 KB
HS Wrr Group      : (not-applicable)
HS Wrr Class Weight: 1                 HS Wrr Weight      : 0
Depth             : 0
HS Class          : 5                  HS Alt Port Class Pool : No
HS Slope Policy   : _tmnx_hs_default
```

The current scheduler traffic rates, including the port and secondary shaper current aggregate traffic rate and current traffic rates for each scheduling class, together with queue current traffic rates on the SAP specified, are shown as follows:

```
*A:PE-1# show qos hs-scheduler-hierarchy sap 3/1/2:1 egress
```

```
===============================================================================
Hs Scheduler Hierarchy Information
===============================================================================
Hs Sched Policy Name         : default
PortId                       : 3/1/2

Port Max-Rate : 138 Mbps
Hs-Sec-Shaper:hs-sec-shaper-1 Agg-Rate : 57728 Kbps

Scheduler Priority 6
  Scheduler Class 6  Rate : 0 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 6 Rate : 0 Kbps
      sap-3/1/2:1->8          Rate : 0 Kbps

Scheduler Priority 5
  Scheduler Class 5  Rate : 22 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 5 Rate : 11454 Kbps
      sap-3/1/2:1->7          Rate : 10040 Kbps

Scheduler Priority 4
  Scheduler Class 4  Rate : 45 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 4 Rate : 22898 Kbps
      sap-3/1/2:1->6          Rate : 20080 Kbps

Scheduler Priority 3
  Scheduler Class 3  Rate : 0 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 3 Rate : 0 Kbps

Scheduler Priority 2
  Scheduler Class 2  Rate : 0 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 2 Rate : 0 Kbps

Scheduler Priority 1
  Scheduler Class 1  Rate : 69 Mbps
    Hs-Sec-Shaper:hs-sec-shaper-1 Class 1 Rate : 23375 Kbps
      sap-3/1/2:1->1  Group: 1 Rate : 13408 Kbps
      sap-3/1/2:1->2  Group: 1 Rate : 6684 Kbps
===============================================================================
===============================================================================
*A:PE-1#
```

The regular SAP **show** commands are supported with SAPs on an HSQ IOM; for example, the SAP
statistics:

```
*A:PE-1# show service id 1 sap 3/1/2:1 stats

===============================================================================
Service Access Points(SAP)
===============================================================================
Service Id        : 1
SAP               : 3/1/2:1                   Encap              : q-tag
Description       : (Not Specified)
Admin State       : Up                        Oper State         : Up
Flags             : None
Multi Svc Site    : None
Last Status Change : 09/28/2017 14:05:52
Last Mgmt Change   : 09/28/2017 14:00:29
-------------------------------------------------------------------------------
Sap per Queue stats
-------------------------------------------------------------------------------
                        Packets              Octets

Ingress Queue 1 (Unicast) (Priority)
```

```
Off. HiPrio            : 0                      0
Off. LowPrio           : 12000                  1536000
Dro. HiPrio            : 0                      0
Dro. LowPrio           : 0                      0
For. InProf            : 0                      0
For. OutProf           : 12000                  1536000

Egress Queue 1
For. In/InplusProf     : 0                      0
For. Out/ExcProf       : 2000                   256000
Dro. In/InplusProf     : 0                      0
Dro. Out/ExcProf       : 0                      0

Egress Queue 2
For. In/InplusProf     : 0                      0
For. Out/ExcProf       : 2000                   256000
Dro. In/InplusProf     : 0                      0
Dro. Out/ExcProf       : 0                      0

Egress Queue 6
For. In/InplusProf     : 0                      0
For. Out/ExcProf       : 1000                   128000
Dro. In/InplusProf     : 0                      0
Dro. Out/ExcProf       : 0                      0

Egress Queue 7
For. In/InplusProf     : 0                      0
For. Out/ExcProf       : 1000                   128000
Dro. In/InplusProf     : 0                      0
Dro. Out/ExcProf       : 0                      0
===============================================================================
*A:PE-1#
```

Network Egress

Each network egress port uses one HSQ queue group for its egress queues. This occurs automatically and dedicates one HSQ queue group, so eight egress queues, to each network egress port, which are used by all network interfaces configured on that port.

The HSQ-specific configuration of network egress queues on an HSQ IOM is similar to that for SAP egress.

The operation of classification, policing, and marking within the network QoS and network queue policies when applied to a network interface and egress HSQ port, respectively, is unchanged. Only the queues to be used need to be configured within the network queue policy.

Most of the commands in a network queue policy apply to the HSQ egress network interfaces, with the following exceptions being ignored:

- Queue commands
    - **adaptation-rule cir** *<adaptation-rule>*
    - **avg-frame-overhead**
    - **mbs**
    - **cbs**
    - **drop tail**
    - **port-parent**

- **pool**

- **rate cir**

The network queue policy **mbs** parameter is ignored, and replaced for HSQ queue group queues with the **hs-mbs** parameter. This is required to allow a more suitable default value to be assigned for the operation of HSQ queues. Both are configured as fractional percentages, with the default for the **mbs** parameter being 50% of the network egress pool, which is not used for HSQ queues, whereas the default for **hs-mbs** is 100% of one second of the queue PIR, converted to bytes. If the queue rate is **max**, the port rate is used (including the HS scheduler policy **max-rate** and the **egress-rate**, if configured on the port).

The network-queue policy has the same HS-specific configuration as in the SAP egress QoS policy, so is not repeated here, but includes:

- The application of an HS attachment policy to define the attachment of the queues to scheduling classes or WRR groups, with the WRR group then being attached to a scheduling class.

- The queue **hs-wrr-weight** to configure the relative weight of each queue within its parented WRR group.

- The rate-related configuration of the two WRR groups in the HSQ queue group.

- The slope policy configured on each HS WRED queue, again with the **highplus-slope** and **time-average-factor** being ignored.

- The use of the port alternative class pools by each queue.

- The HS class weight for WRR group scheduling between queues and WRR groups in different HSQ queue groups at a primary shaper scheduling class.

HS WRR group and queue rates are configured as a percentage of the port rate, which includes both the **egress-rate** and HS scheduler policy **max-rate**, if configured.

Secondary shapers and HSQ queue group aggregate rates are not applicable to network egress HSQ queue group queues.

The related network queue policy configuration is as follows:

```
configure
    qos
        network-queue <policy-name> [create]
            hs-attachment-policy <policy-name>
            hs-wrr-group <group-id>
                adaptation-rule [pir <adaptation-rule>]
                hs-class-weight <weight>
                rate <percent>
            queue <queue-id> [multipoint] [<queue-type>] [create]
                hs-alt-port-class-pool
                hs-class-weight <weight>
                hs-mbs <percent-of-queue-rate>
                hs-wred-queue [policy <slope-policy-name>]
                hs-wrr-weight <weight>
```

See the SAP Egress section for details of the preceding parameters.

A network interface is configured on port 3/1/1:1:

```
*A:PE-1>config>router# info | match "IP Configuration" pre-lines 1 post-lines 10
#-------------------------------------------------
echo "IP Configuration"
#-------------------------------------------------
        interface "PE-1-Network-1"
```

```
                address 192.168.10.1/30
                description "HSQ network egress queues"
                port 3/1/1:1
                qos 10
                no shutdown
            exit
```

The network QoS policy 10 applied to the interface only contains the necessary egress **dscp** statements to classify the egress traffic to the forwarding classes mapped to the queues.

The network queue policy configured on port 3/1/1 contains the applied HS attachment policy described for the queue attachment. Queue 1 is configured with a WRR weight of 2. Rates of 2% and 1% are configured on queues 6 and 7, respectively. WRR group 1 is configured with a rate of 2%:

```
*A:PE-1>config>qos>network-queue# info
----------------------------------------------
            hs-attachment-policy "hs-att-pol-1"
            queue 1 create
                hs-wrr-weight 2
            exit
            queue 2 create
            exit
            queue 6 create
                rate 2
            exit
            queue 7 create
                rate 1
            exit
            hs-wrr-group 1
                rate 2
            exit
```

The network queue information is shown as follows:

```
*A:PE-1# show hs-pools port 3/1/1 egress network-queues |
                        match "Queue Information" pre-lines 1 post-lines 40
-------------------------------------------------------------------------------
Queue Information
-------------------------------------------------------------------------------
Queue Name         : 1 Net=be Port=3/1/1
FC Map             : be l2 l1 h2 nc
Admin PIR          : 200000              Oper PIR           : 0
Admin MBS          : 25000000 B          Oper MBS           : 24416 KB
HS Wrr Group       : 1
HS Wrr Class Weight: 1                   HS Wrr Weight      : 2
Depth              : 0
HS Class           : 1                   HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default

Queue Name         : 2 Net=af Port=3/1/1
FC Map             : af
Admin PIR          : 200000              Oper PIR           : 0
Admin MBS          : 25000000 B          Oper MBS           : 24416 KB
HS Wrr Group       : 1
HS Wrr Class Weight: 1                   HS Wrr Weight      : 1
Depth              : 0
HS Class           : 1                   HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default

Queue Name         : 6 Net=ef Port=3/1/1
FC Map             : ef
Admin PIR          : 200000              Oper PIR           : 200000
```

```
Admin MBS          : 25000000 B        Oper MBS            : 24416 KB
HS Wrr Group       : (not-applicable)
HS Wrr Class Weight: 1                 HS Wrr Weight       : 0
Depth              : 0
HS Class           : 4                 HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default

Queue Name         : 7 Net=h1 Port=3/1/1
FC Map             : h1
Admin PIR          : 100000            Oper PIR            : 100000
Admin MBS          : 12500000 B        Oper MBS            : 12208 KB
HS Wrr Group       : (not-applicable)
HS Wrr Class Weight: 1                 HS Wrr Weight       : 0
Depth              : 0
HS Class           : 5                 HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default
```

The current port-based scheduler traffic rates for network egress are shown as follows:

```
*A:PE-1# show qos hs-scheduler-hierarchy port 3/1/1

===============================================================================
Hs Scheduler Hierarchy Information
===============================================================================
Hs Sched Policy Name       : hs-sched-pol-1

Port Max-Rate : 660 Mbps

Scheduler Priority 6
  Scheduler Class 6  Rate : 0 Mbps

Scheduler Priority 5
  Scheduler Class 5  Rate : 127 Mbps

Scheduler Priority 4
  Scheduler Class 4  Rate : 255 Mbps

Scheduler Priority 3
  Scheduler Class 3  Rate : 0 Mbps

Scheduler Priority 2
  Scheduler Class 2  Rate : 0 Mbps

Scheduler Priority 1
  Scheduler Class 1  Rate : 277 Mbps
===============================================================================
*A:PE-1#
```

The regular **show** commands are supported with network interfaces on an HSQ IOM; for example, the port queue statistics:

```
*A:PE-1# show port 3/1/1 detail |
    match expression "Ethernet Interface|Egress Queue" pre-lines 1 post-lines 6
===============================================================================
Ethernet Interface
===============================================================================
Description        : 10-Gig Ethernet
Interface          : 3/1/1             Oper Speed       : 10 Gbps
Link-level         : Ethernet          Config Speed     : N/A
Admin State        : up                Oper Duplex      : full
Oper State         : up                Config Duplex    : N/A
```

```
Egress Queue  1            Packets                Octets
    In/Inplus Prof fwded  :    0                      0
    In/Inplus Prof dropped:    0                      0
    Out/Exc Prof fwded    :    2000                   256000
    Out/Exc Prof dropped  :    0                      0
Egress Queue  2            Packets                Octets
    In/Inplus Prof fwded  :    0                      0
    In/Inplus Prof dropped:    0                      0
    Out/Exc Prof fwded    :    2000                   256000
    Out/Exc Prof dropped  :    0                      0
Egress Queue  6            Packets                Octets
    In/Inplus Prof fwded  :    1000                   128000
    In/Inplus Prof dropped:    0                      0
    Out/Exc Prof fwded    :    0                      0
    Out/Exc Prof dropped  :    0                      0
Egress Queue  7            Packets                Octets
    In/Inplus Prof fwded  :    1000                   128000
    In/Inplus Prof dropped:    0                      0
    Out/Exc Prof fwded    :    0                      0
    Out/Exc Prof dropped  :    0                      0
===============================================================================
*A:PE-1#
```

Access and Network Egress Queue Groups

Each access and network egress queue group instance configured on an HSQ IOM uses an HSQ queue group for its queues. This occurs automatically and dedicates one HSQ queue group, so eight egress queues, to each egress queue group instance. Only the queues to be used need to be configured within the egress queue group template.

The operation of classification, policing, and marking related to egress queue group instances on an HSQ port is unchanged.

Most of the commands in an egress queue group template apply to the HSQ egress queue group instances, with the following exceptions (including their related overrides) being ignored:

- HSMDA commands
- **queues-hqos-manageable**
- Queue related
  - **adaptation-rule cir <adaptation-rule>**
  - **adv-config-policy**
  - **avg-frame-overhead**
  - **burst-limit**
  - **cbs**
  - **drop tail**
  - **dynamic-mbs**
  - **parent**
  - **percent-rate cir**
  - **pool**
  - **port-parent**

- **rate cir**

- **wred-queue**

The following commands are not configurable under port access and network egress queue group instances:

- **egress agg-rate limit-unused-bandwidth**

- **egress agg-rate queue-frame-based-accounting**

The configuration of egress queue groups using HSQ queue groups is unchanged. The egress queue group template is applied under **config>port>ethernet>access>egress** or **config>port>ethernet>network>egr** to create the queue group instances, and traffic is redirected to these instances in either a SAP egress QoS policy or network QoS policy.

The system-created egress queue group instances each use an HSQ queue group; for example, the post-policer access egress *policer-output-queues* queue groups.

The egress queue group template has the same HS-specific configuration as in the SAP egress QoS policy, so is not repeated here, but includes:

- The application of an HS attachment policy to define the attachment of the queues to scheduling classes or WRR groups, with the WRR group then being attached to a scheduling class.

- The queue **hs-wrr-weight** to configure the relative weight of each queue within its parented WRR group.

- The rate-related configuration of the two WRR groups in the HSQ queue group.

- The slope policy configured on each HS WRED queue, again with the **highplus-slope** and **time-average-factor** being ignored.

- The use of the port alternative class pools by each queue.

- The HS class weight for WRR group scheduling between queues and WRR groups in different HSQ queue groups at a primary shaper scheduling class.

Secondary shapers are not applicable to both access and network egress queue group instance HSQ queue groups.

The related egress queue group template configuration syntax is as follows:

```
configure
    qos
        queue-group-templates
            egress
                queue-group <queue-group-name> [create]
                    hs-attachment-policy <policy-name>
                    hs-wrr-group <group-id>
                        adaptation-rule [pir <adaptation-rule>]
                        hs-class-weight <weight>
                        percent-rate <percent>
                        rate <rate>
                    queue <queue-id> [queue-type] [create]
                        hs-alt-port-class-pool
                        hs-class-weight <weight>
                        hs-wred-queue [policy <slope-policy-name>]
                        hs-wrr-weight <weight>
```

See the SAP Egress section for details of the preceding parameters.

The **percent-rate** configured within the **hs-wrr-group** and under a queue is relative to the port rate, and includes both the **egress-rate** and HS scheduler policy **max-rate**, if configured.

The HSQ queue group aggregate rate is applied to an egress queue group instance using the **agg-rate rate** command under the application of the queue group template on the port:

```
configure
    port <port-id>
        ethernet
            access
                egress
                    queue-group <queue-group-name> [instance <instance-id>]
                        agg-rate
                            rate <kilobits-per-second>
            network
                egress
                    queue-group <queue-group-name> [instance <instance-id>]
                        agg-rate
                            rate <kilobits-per-second>
```

Where:

```
<port-id>             : slot/mda/port
<queue-group-name>    : [32 chars max]
<instance-id>         : [1..65535]
<kilobits-per-seco*>  : [1..3200000000|max] Kbps
```

When using HSQ queue groups with access or network egress queue group instances on 100G ports, the **hs-turbo** parameter can be configured under the port queue group instance to allow the corresponding HSQ queue group queues to achieve a higher throughput. The default is **no hs-turbo**. The **hs-turbo** parameter is not applicable to 10G ports and is ignored when configured under a queue group instance on a 10G port.

```
configure
    port <port-id>
        ethernet
            access
                egress
                    queue-group <queue-group-name> [instance <instance-id>]
                        hs-turbo
            network
                egress
                    queue-group <queue-group-name> [instance <instance-id>]
                        hs-turbo
```

Where:

```
<port-id>             : slot/mda/port
<queue-group-name>    : [32 chars max]
<instance-id>         : [1..65535]
```

> **Note:**
> Queue depth monitoring is supported for access and network egress queue groups HSQ queues. This is enabled by configuring the queue override **monitor-depth** command under the queue group instance with the associated output showing buffer occupancy in ranges of 10% of the queue depth for each configured queue.

An egress queue group template is configured containing the applied HS attachment policy described for the queue attachment. Queue 1 is configured with a WRR weight of 2. Rates of 20 Mb/s and 10 Mb/s are configured on queues 6 and 7, respectively. WRR group 1 is configured with a rate of 40 Mb/s:

```
A:PE-1# configure qos queue-group-templates egress
A:PE-1>cfg>qos>qgrps>egr# info
----------------------------------------------
                queue-group "queue-group-1" create
                    hs-attachment-policy "hs-att-pol-1"
                    queue 1 best-effort create
                        hs-wrr-weight 2
                    exit
                    queue 2 best-effort create
                    exit
                    queue 6 best-effort create
                        rate 20000
                    exit
                    queue 7 best-effort create
                        rate 10000
                    exit
                    hs-wrr-group 1
                        rate 40000
                    exit
                exit
```

The queue group template is applied to the network port 3/1/1 and access port 3/1/2, each with an aggregate rate of 100 Mb/s:

```
A:PE-1# configure port 3/1/1
A:PE-1>config>port# info
----------------------------------------------
        ethernet
            network
                egress
                    queue-group "queue-group-1" instance 1 create
                        agg-rate
                            rate 100000
                        exit
                    exit
                exit
            exit
        exit
        no shutdown
----------------------------------------------
A:PE-1>config>port# exit all
A:PE-1# configure port 3/1/2
A:PE-1>config>port# info
----------------------------------------------
        ethernet
            access
                egress
                    queue-group "queue-group-1" instance 1 create
                        agg-rate
                            rate 100000
                        exit
                    exit
                exit
            exit
        exit
        no shutdown
```

The configured aggregate rates are shown as follows:

```
*A:PE-1# show port 3/1/[1,2] queue-group queue-group-1 instance 1 |
                             match "Ethernet port" pre-lines 1 post-line 7
===============================================================================
Ethernet port 3/1/1 Network Egress queue-group
===============================================================================
Group Name       : queue-group-1      Instance-Id  : 1
Description      : (Not Specified)
Sched Policy     : None               Acct Pol     : None
Collect Stats    : disabled           Agg. Limit   : 100000
Limit Unused BW  : Disabled
HS Turbo Queues  : Disabled
===============================================================================
Ethernet port 3/1/2 Access Egress queue-group
===============================================================================
Group Name       : queue-group-1      Instance-Id  : 1
Description      : (Not Specified)
Sched Policy     : None               Acct Pol     : None
Collect Stats    : disabled           Agg. Limit   : 100000
Limit Unused BW  : Disabled
HS Turbo Queues  : Disabled
*A:PE-1#
```

An IES interface SAP is configured on port 3/1/2 with a SAP egress QoS policy redirecting the traffic to the access egress queue group instance, and a network interface is configured on port 3/1/1 with a network QoS policy redirecting the traffic to the network egress queue group instance.

The queue information of the access egress queue group HSQ queue group queues is shown as follows:

```
A:PE-1# show hs-pools port 3/1/2 egress queue-group "queue-group-1" |
                         match "Queue Information" pre-lines 1 post-lines 40
-------------------------------------------------------------------------------
Queue Information
-------------------------------------------------------------------------------
Queue Name         : accQGrp->queue-group-1:1(3/1/2)->1
FC Map             : not-applicable
Admin PIR          : 40000             Oper PIR            : 0
Admin MBS          : 64 KB             Oper MBS           : 64 KB
HS Wrr Group       : 1
HS Wrr Class Weight: 1                 HS Wrr Weight       : 2
Depth              : 0
HS Class           : 1                 HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default

Queue Name         : accQGrp->queue-group-1:1(3/1/2)->2
FC Map             : not-applicable
Admin PIR          : 40000             Oper PIR            : 0
Admin MBS          : 64 KB             Oper MBS           : 64 KB
HS Wrr Group       : 1
HS Wrr Class Weight: 1                 HS Wrr Weight       : 1
Depth              : 0
HS Class           : 1                 HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default

Queue Name         : accQGrp->queue-group-1:1(3/1/2)->6
FC Map             : not-applicable
Admin PIR          : 20000             Oper PIR            : 20000
Admin MBS          : 64 KB             Oper MBS           : 64 KB
HS Wrr Group       : (not-applicable)
HS Wrr Class Weight: 1                 HS Wrr Weight       : 0
Depth              : 0
HS Class           : 4                 HS Alt Port Class Pool : No
```

```
HS Slope Policy    : _tmnx_hs_default

Queue Name         : accQGrp->queue-group-1:1(3/1/2)->7
FC Map             : not-applicable
Admin PIR          : 10000              Oper PIR           : 10000
Admin MBS          : 64 KB              Oper MBS           : 64 KB
HS Wrr Group       : (not-applicable)
HS Wrr Class Weight: 1                  HS Wrr Weight      : 0
Depth              : 0
HS Class           : 5                  HS Alt Port Class Pool : No
HS Slope Policy    : _tmnx_hs_default
```

The equivalent information can be displayed for the network egress queue group HSQ queue group queues by replacing 3/1/2 by 3/1/1.

The current port scheduler aggregate traffic rate and the current traffic rates for each scheduling class, together with current queue traffic rates in the specified access or network egress queue group instance, are shown as follows:

```
*A:PE-1# show qos hs-scheduler-hierarchy port 3/1/2
                queue-group "queue-group-1" instance 1 access

===============================================================================
Hs Scheduler Hierarchy Information
===============================================================================
Hs Sched Policy Name       : default

Port Max-Rate : 138 Mbps

Scheduler Priority 6
  Scheduler Class 6  Rate : 0 Mbps
     Queue 8           Rate : 0 Kbps

Scheduler Priority 5
  Scheduler Class 5  Rate : 22 Mbps
     Queue 7           Rate : 10062 Kbps

Scheduler Priority 4
  Scheduler Class 4  Rate : 45 Mbps
     Queue 6           Rate : 20124 Kbps

Scheduler Priority 3
  Scheduler Class 3  Rate : 0 Mbps

Scheduler Priority 2
  Scheduler Class 2  Rate : 0 Mbps

Scheduler Priority 1
  Scheduler Class 1  Rate : 69 Mbps
     Queue 1           Rate : 26837 Kbps
     Queue 2           Rate : 13397 Kbps
===============================================================================
*A:PE-1#
```

Conclusion

The HSQ IOM provides high scale QoS in terms of the number of ingress policers and egress queues supported. It supports six scheduling classes across multiple hierarchical levels of hardware egress shaping encompassing HSQ queue groups, primary shapers, secondary shapers, and port schedulers. A

flexible buffer allocation mechanism permits both buffer isolation and buffer oversubscription for the queue buffer allocation.

Pseudowire QoS

This chapter describes pseudowire QoS configurations.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

The information and configuration is based on Release 11.0.R4. There are no specific prerequisites for this configuration.

Overview

A pseudowire (PW) provides a virtual connection across an IP or MPLS network between services configured on provider edge (PE) devices. In SR OS Release 10.0.R1, and later, it is possible to provide specific QoS to either a single pseudowire or a multiple pseudowires. This is supported for the following:

- SDP
    - MPLS
    - GRE
- Epipe
    - Including vc-switching and dynamic MS-PW
    - PBB-epipe
    - BGP-VPWS (Release 11.0.R1 and later)
- VPLS
    - Mesh and spoke SDP
    - LDP signaled pseudowires
    - BGP-AD signaled pseudowires
    - I-VPLS, B-VPLS
    - R-VPLS
    - BGP-VPLS
- Spoke termination on IES/VPRN (both Epipe and Ipipe)
- Apipe (from R10.0.R4)

- Cpipe (from R10.0.R4)
- Fpipe (from R10.0.R4)
- Ipipe (from R10.0.R4)

It is supported at ingress on both Ethernet and POS/TDM ports and only on Ethernet ports at egress.

Bandwidth control is achieved using queue-groups which are implemented per flexpath (FP) at the ingress and per port at the egress (these being relative to the data path through the system), as shown in Figure 14: Ingress PW QoS and Figure 15: Egress PW QoS, respectively.

Figure 14: Ingress PW QoS

© 2026 Nokia.
Use subject to Terms available at: www.nokia.com/terms.
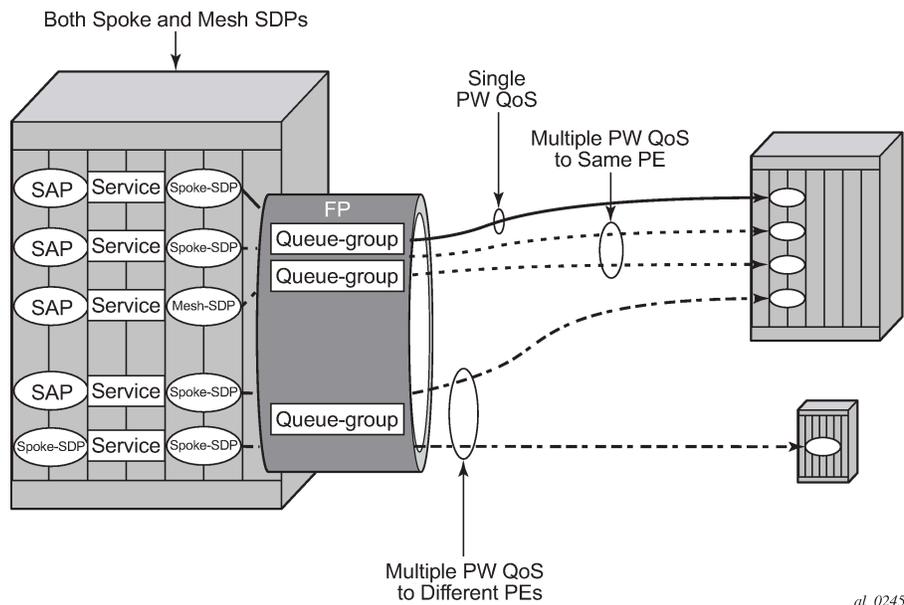
Figure 15: Egress PW QoS
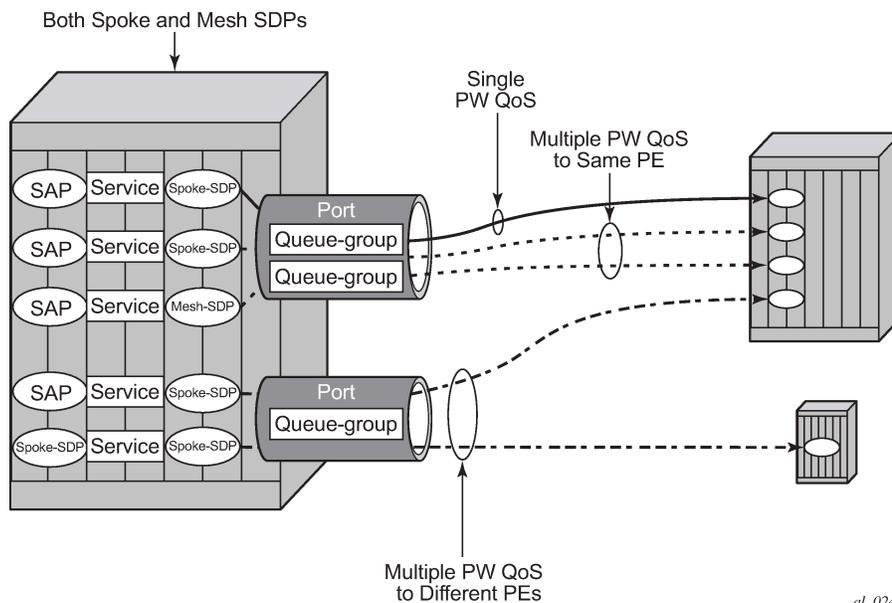
Bandwidth control is applied independently for ingress and egress, and can be set up for a single pseudowire or for multiple pseudowires where the remote services are located on a single PE or on multiple PEs.

It is possible to benefit from Hierarchical QoS which can be configured under the queue-groups, but this is beyond the scope of this chapter.

The ingress and egress classification and egress marking is configured by applying a network QoS policy to each pseudowire.

Ingress QoS

Ingress QoS is achieved using a queue group which is applied to an ingress FP on a card. Queue groups applied to an FP can only contain policers, not queues. The network QoS policy applied to the pseudowire redirects forwarding classes (FCs) to the individual queue group (unicast or multipoint) policers. The actual queue group to be used is defined separately to the network QoS policy, thereby allowing the network QoS policies to be independent from the queue groups used and therefore both are reusable.

Ingress bandwidth control does not take into account the outer Ethernet header, the MPLS labels/control word or GRE headers, or the FCS of the incoming frame. The configuration allows an offset to be added or subtracted from the received frame size in order to change the actual length used for the bandwidth control. For example: if the same ingress rate is configured on a pseudowire (without a control word) and a dot1q SAP, what packet-byte-offset needs to be used on the pseudowire in order to achieve the same throughput as on the SAP?

- SAP — The following shows the bytes in the frame that are used by default on a policer for the rate at a SAP ingress.
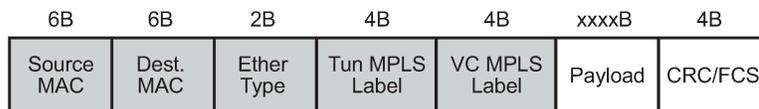
6B	6B	4B	2B	xxxxB	4B
Source MAC	Dest. MAC	802.1Q	Ether Type	Payload	CRC/FCS
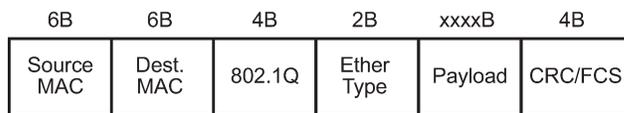
al_0247

- VPLS Pseudowire — For a tagged (**vc-type vlan**) pseudowire, it would be necessary to add 4 bytes using the packet-byte-offset applied to the ingress policer in order to achieve the same throughput as on the SAP. This compensates for the omission of the FCS that is included on the SAP and so needs to be added.

6B	6B	2B	4B	4B	6B	6B	4B	2B	xxxxB	4B
Source MAC	Dest. MAC	Ether Type	Tun MPLS Label	VC MPLS Label	Source MAC	Dest. MAC	802.1Q	Ether Type	Payload	CRC/FCS
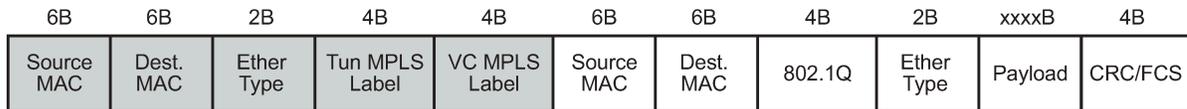
al_0248

- VPRN Pseudowire — For an Ipipe (**vc-type** ipipe) pseudowire, it would be necessary to add 22 bytes using the packet-byte-offset to the ingress policer to achieve the same throughput as on the SAP. This compensates for the omission of the source and destination MAC addresses (12 bytes), Ether type (2 bytes), VLAN tag (4 bytes) and the FCS (4 bytes) that are included on the SAP and so needs to be added.

6B	6B	2B	4B	4B	xxxxB	4B
Source MAC	Dest. MAC	Ether Type	Tun MPLS Label	VC MPLS Label	Payload	CRC/FCS

al_0249

The ingress classification is configured in the ingress section of the network QoS policy and is based on the outer encapsulation header only, the outer Ethernet header (dot1p/DE), MPLS labels (EXP), or GRE headers (DSCP). At an egress LER, the ler-use-dscp is applicable only to IES and VPRN pseudowires.

Egress QoS

Egress QoS is achieved using a queue group which is applied to an egress port. Queue groups applied to a port can contain both policers and queues. The network QoS policy applied to the pseudowire redirects forwarding classes (FCs) to the individual queue group policers/queues. The actual queue group to be used is defined separately to the network QoS policy, thereby allowing the network QoS policies to be independent from the queue groups used and therefore both are reusable.

Egress bandwidth control takes into account the outer Ethernet header, MPLS labels/control word, or GRE headers, and the FCS of the outgoing frame. The configuration allows an offset to be added or subtracted from the sent frame size in order to affect the actual length used for the bandwidth control.
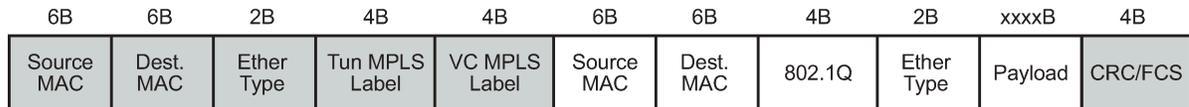
For example, if the same egress rate is configured on a pseudowire (without a control word) and a dot1q SAP, what packet-byte-offset needs to be used on the pseudowire in order to achieve the same throughput as on the SAP?

- SAP — The following shows the bytes in the frame that are used by default on a policer/queue at a SAP egress.
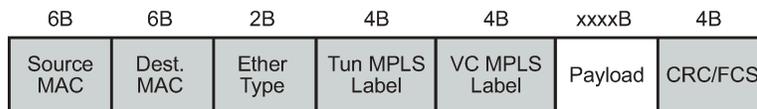
6B	6B	4B	2B	xxxxB	4B
Source MAC	Dest. MAC	802.1Q	Ether Type	Payload	CRC/FCS
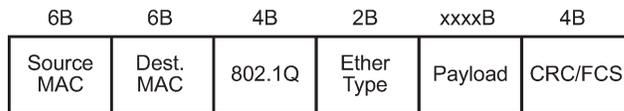
al_0250

- VPLS Pseudowire — For a tagged (**vc-type vlan**) pseudowire, it would be necessary to subtract 22 bytes using the packet-byte-offset applied to the egress policer/queue applied to achieve the same throughput as on the SAP. This compensates for the MPLS header (source and destination MAC addresses (12 bytes), Ether type (2 bytes), two labels (8 bytes)) that is not included on the SAP and needs to be subtracted.

6B	6B	2B	4B	4B	6B	6B	4B	2B	xxxxB	4B
Source MAC	Dest. MAC	Ether Type	Tun MPLS Label	VC MPLS Label	Source MAC	Dest. MAC	802.1Q	Ether Type	Payload	CRC/FCS

al_0251

- VPRN Pseudowire — For an Ipipe (**vc-type ipipe**) pseudowire, it would be necessary to subtract 4 bytes using the packet-byte-offset applied to the egress policer/queue applied to achieve the same throughput as on the SAP. This compensates for the MPLS header (source and destination MAC addresses (12 bytes), Ether type (2 bytes), two labels (8 bytes)) that is not included on the SAP so is subtracted, and the source and destination MAC addresses (12 bytes), dot1q header (4 bytes) and Ether type (2 bytes) of the SAP frame which needs to be added. This results in subtracting 4 bytes.

6B	6B	2B	4B	4B	xxxxB	4B
Source MAC	Dest. MAC	Ether Type	Tun MPLS Label	VC MPLS Label	Payload	CRC/FCS

al_0252

The egress classification and marking is configured in the egress section of the network QoS policy. DSCP/prec egress reclassification is supported in SR OS Release 10.0.R4, and later, for IES and VPRN spoke SDPs. The egress marking affects the outer encapsulation header, the outer Ethernet header (dot1p/DE), MPLS labels (EXP) or GRE headers (DSCP).

Configuration

The configuration of pseudowire QoS is described using an Epipe pseudowire. The topology is shown in .

Figure 16: Example Epipe Pseudowire Topology
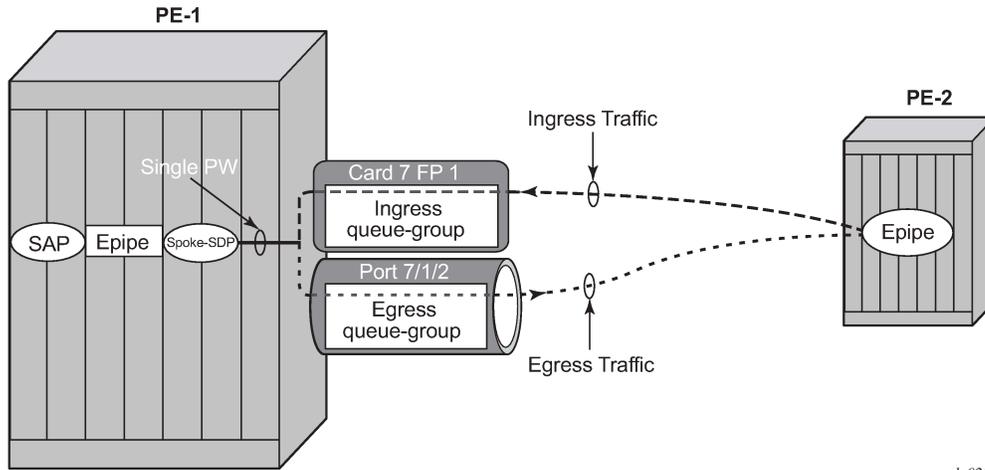
The following prerequisite configuration is assumed to be in place:

- Hardware provisioning

- IP address and routing

- MPLS protocols

- SDP

- Epipe service, including the SAP

- SAP QoS policies

Traffic is sent across a virtual leased line between PE-1 and PE-2 using Epipes with a pseudowire configured as a spoke SDP on each PE. The QoS is applied to the pseudowire at the ingress and egress of PE-1.

The following configuration is required for applying pseudowire QoS:

- Create the ingress and egress queue groups. These contain the ingress policer and egress policer/ queue definitions.

- Create an instance of the ingress queue group on the ingress FP and instance of the egress queue group on the port that will be used for the pseudowire traffic.

- Create a network QoS policy to redirect the traffic to the ingress and egress queue groups, and to perform the ingress classification and egress marking.

- Apply the network QoS policy, together with the reference to the ingress and egress queue group instances, to the spoke SDP representing the pseudowire.

The traffic consists of two bidirectional flows, one in FC BE and one in FC EF. At the ingress of the pseudowire, each FC is assigned to its own policer, whereas at the egress of the pseudowire, FC BE is assigned to a queue and FC EF is assigned to a policer.

Although this example makes use of both ingress and egress queue groups, the focus is pseudowire QoS, so the full details of queue group configuration are not covered.

Create the Ingress and Egress Queue Groups

Queue groups are created using templates, which are separate for ingress and egress. The following shows the queue group templates configured.

```
configure qos
        queue-group-templates
            ingress
                queue-group "ingress-queue-group" create
                    policer 1 create
                        rate 6000
                        packet-byte-offset add 4
                    exit
                    policer 2 create
                        rate 4000
                        packet-byte-offset add 4
                    exit
                exit
            exit
            egress
                queue-group "egress-queue-group" create
                    queue 1 best-effort create
                        rate 6000
                        xp-specific
                            packet-byte-offset subtract 22
                        exit
                    exit
                    policer 1 create
                        rate 4000
                        packet-byte-offset subtract 22
                    exit
                exit
            exit
    exit
```

The ingress queue group has two policers associated with it; policer 1 will be used for the FC BE traffic and policer 2 will be used for the FC EF traffic. The configuration of policers in an ingress queue group is the same as that in a sap-ingress QoS policy, with the exception that the percent-rate is not supported within the queue group.

In order to achieve the same ingress throughput as that when applying the same rates to policers on a dot1q tagged SAP, the packet-byte-offset adds 4 bytes to the packet length for both policers.

The egress queue group has one queue (queue 1) that will be used for the FC BE traffic and one policer (policer 1) that will be used for the FC EF traffic. The configuration of policers in an egress queue group is the same as that in a sap-egress QoS policy, with the exception that the percent-rate is not supported within the queue group. The configuration of queues in an egress queue group is the same as in a sap-egress QoS policy, with the exception that the avg-frame-overhead is not supported within the queue group.

In order to achieve the same egress throughput as that when applying the same rates to policers/queues on a dot1q tagged SAP, the packet-byte-offset subtracts 22 bytes from the packet length for both the policer and queue.

Rates have been configured such that the ingress and egress capacity of the BE traffic is 6Mb/s and 4Mb/s for the EF traffic.

Create the Ingress FP and Egress Port Queue Group Instances

The queue group templates are then applied as individual instances to the ingress FP and egress port; using instances allows the reuse of the same template.

Following is the ingress FP configuration. From a QoS perspective, it is also possible to configure a policer-control-policy under the ingress queue group in order to perform hierarchical policing. In Release 11.0.R4, and later, the configuration supports overrides for both the policer-control-policy parameters and some of the queue group policer parameters.

```
configure
    card 7
        card-type imm5-10gb-xfp
        mda 1
            no shutdown
        exit
        fp 1
            ingress
                network
                    queue-group "ingress-queue-group" instance 1 create
                    exit
                exit
            exit
        exit
        no shutdown
    exit
```

Following is the egress port configuration. From a QoS perspective, it is also possible to configure under the egress queue group a policer-control-policy in order to perform hierarchical policing, a scheduler-policy in order to perform hierarchical shaping and overrides for some of the queue group queue parameters.

```
configure
    port 7/1/2
        ethernet
            network
                egress
                    queue-group "egress-queue-group" instance 1 create
                    exit
                exit
            exit
        exit
        no shutdown
    exit
```

If there are redundant network interfaces over which the pseudowire traffic can enter or exit the system, it is necessary to configure any ingress FP and egress port queue groups consistently across all possible interfaces to be used by the pseudowire to ensure the QoS is always applied. If a queue group configuration was omitted, the pseudowire would not be subject to the QoS defined in that queue group.

If a LAG is used, the system only allows the egress port queue group to be added or removed from the LAG primary port, thereby keeping the LAG configuration consistent. However, this is not possible at the ingress as the queue-group is applied at the FP, so it is necessary to ensure that the ingress queue group is applied consistently on all FPs corresponding to the configured LAG.

Create the Network QoS Policy

A network QoS policy is created to redirect ingress and egress traffic to the respective queue groups, and perform ingress classification (in this example).

The redirection to the queue group policer/queue is performed per FC.

At ingress, traffic can be redirected to policers (being the same or different policers) based on the traffic type. Unicast traffic is redirected to a policer specified by the policer command and will use the ingress shared policer-output-queues to access the switch fabric. All multipoint traffic is redirected to the policer specified by the multicast-policer command (for example, with a pseudowire configured in a VPLS service, all broadcast, unknown, and multicast traffic will use this policer). The multipoint traffic accesses the switch fabric using the Ingress Multicast Path Management queues. It is possible to individually redirect one traffic type (unicast or multipoint) within an FC to a queue group policer while allowing the other traffic type to use default network queues.

At egress, traffic can be redirected to a queue or to a policer. The policed traffic will exit the egress port using one of the default network queues (with the queue chosen by FC assignment) or optionally can use a queue in the egress queue group if configured in the port-redirect-group command following the policer parameter.

Any FC not redirected to a queue-group, will continue to use the regular default network ingress and egress queues.

The syntax for the FC redirection is as follows.

```
config# qos
  network <network-policy-id> [create]
      ingress
          fc <fc-name>
              fp-redirect-group multicast-policer <policer-id>
              fp-redirect-group policer <policer-id>
      egress
          fc <fc-name>
              port-redirect-group {queue <queue-id>|
policer <policer-id> [queue <queue-id>]}
```

The required commands are shown below.

```
configure qos
        network 10 create
            ingress
                lsp-exp 5 fc ef profile in
                fc be
                    fp-redirect-group policer 1
                exit
                fc ef
                    fp-redirect-group policer 2
                exit
            exit
            egress
                fc be
                    port-redirect-group queue 1
                exit
                fc ef
                    port-redirect-group policer 1
                exit
            exit
        exit
```

At ingress, the FC BE and FC EF traffic are redirected to the two policers in the queue-group applied to the FP. At egress, the two FCs are redirected to the queue and policer in the queue group applied to the egress port.

The ingress classification required here is for the traffic which is received with exp=5 to be in FC EF.

Apply Network QoS Policy with Queue Group Instances to the Spoke SDP

To apply the QoS to the pseudowire, the following commands can be used, dependent on the service type.

```
config# service {apipe|cpipe|epipe|fpipe|ipipe} <service-id>
  spoke-sdp <sdp-id:vc-id>
    ingress
      qos <network-policy-id> fp-redirect-group <queue-group-name>
                              instance <instance-id>
    egress
      qos <network-policy-id> port-redirect-group <queue-group-name>
                              instance <instance-id>
```

```
config# service {ies|vprn} <service-id>
  interface <ip-int-name>
    spoke-sdp <sdp-id:vc-id>
      ingress
        qos <network-policy-id> fp-redirect-group <queue-group-name>
                                instance <instance-id>
      egress
        qos <network-policy-id> port-redirect-group <queue-group-name>
                                instance <instance-id>
```

```
config# service vpls <service-id>
  {spoke-sdp|mesh-sdp} <sdp-id:vc-id>
    ingress
      qos <network-policy-id> fp-redirect-group <queue-group-name>
                              instance <instance-id>
    egress
      qos <network-policy-id> port-redirect-group <queue-group-name>
                              instance <instance-id>
```

For services using BGP auto-discovery to signal the pseudowire, the QoS configuration is included in the pseudowire template.

```
config# service pw-template <policy-id>
    ingress
      qos <network-policy-id> fp-redirect-group <queue-group-name>
                              instance <instance-id>
    egress
      qos <network-policy-id> port-redirect-group <queue-group-name>
                              instance <instance-id>
```

To propagate changes in a pw-template to existing BGP-AD pseudowires, it is necessary to use the following command:

```
tools perform service eval-pw-template policy-id
```

The allow-service-impact parameter is not required for changing the ingress or egress QoS definition as these do not affect the operational state of the pseudowire.

QoS applied directly to a pseudowire, using the preceding commands, takes precedence over any QoS applied to the network interface (using a network QoS policy with or without queue group redirection).

Each time a pseudowire uses a network egress port queue group, an FP resource is allocated. This only requires that the pseudowire egress QoS is configured with a port-redirect-group, and will occur even if there are no FCs redirected using a port-redirect-group within the configured network QoS policy. The resources used can be seen using the **tools dump system-resources** command and is listed under Egr Network Queue Group Mappings which is part of the total for the "Dynamic Service Entries ".

As an Epipe is used in this example, QoS is configured directly under a spoke SDP.

```
configure service
        epipe 1 customer 1 create
            spoke-sdp 1:1 vc-type vlan create
                ingress
                    qos 10 fp-redirect-group "ingress-queue-group" instance 1
                exit
                egress
                    qos 10 port-redirect-group "egress-queue-group" instance 1
                exit
                no shutdown
            exit
            no shutdown
        exit
```

The created network QoS policy is applied at both ingress and egress, with the ingress referencing the ingress queue group instance applied to the FP and the egress referencing the egress queue group instance applied to the port.

Show Output

The configured ingress queue group can be shown, including the details of the configured policers and where it is applied, as follows.

```
*A:PE-1# show qos queue-group "ingress-queue-group" ingress detail
===============================================================================
QoS Queue-Group Ingress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name     : ingress-queue-group
Description    : (Not Specified)
-------------------------------------------------------------------------------
---snip---
===============================================================================
Queue Group FP Maps
===============================================================================
Card Num       Fp Num              Instance            Type
-------------------------------------------------------------------------------
7              1                   1                   Network
-------------------------------------------------------------------------------
Entries found: 1
-------------------------------------------------------------------------------
===============================================================================
Queue Group Policer
===============================================================================
Policer Id     : 1
Description    : (Not Specified)
```

```
PIR Adptn      : closest                CIR Adptn     : closest
Parent         : none                   Level         : 1
Weight         : 1                      Adv. Cfg Plcy: none
Admin PIR      : 6000                   Admin CIR     : 0
CBS            : def                    MBS           : def
Hi Prio Only   : def                    Pkt Offset    : 4
Profile Capped : Disabled
StatMode       : minimal
===============================================================================
Policer Id     : 2
Description    : (Not Specified)
PIR Adptn      : closest                CIR Adptn     : closest
Parent         : none                   Level         : 1
Weight         : 1                      Adv. Cfg Plcy: none
Admin PIR      : 4000                   Admin CIR     : 0
CBS            : def                    MBS           : def
Hi Prio Only   : def                    Pkt Offset    : 4
Profile Capped : Disabled
StatMode       : minimal
```

Similar information can be shown for the egress queue group, including the details of the configured queue and policer and again where it is applied.

```
*A:PE-1# show qos queue-group "egress-queue-group" egress detail
===============================================================================
QoS Queue-Group Egress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name     : egress-queue-group
Description    : (Not Specified)


-------------------------------------------------------------------------------
Q   CIR Admin PIR Admin CBS        HiPrio PIR Lvl/Wt    Parent     BurstLimit(B)
    CIR Rule  PIR Rule  MBS               CIR Lvl/Wt    Wred-Queue    Slope
    Named-Buffer Pool                     Adv Config Policy Name
-------------------------------------------------------------------------------
1   0       6000     def       def    1/1          None       default
    closest closest  def              0/1          disabled   default
    (not-assigned)                    (not-assigned)
---snip---
===============================================================================
Queue Group Ports (network)
===============================================================================
Port    Sched Pol    Policer-Ctrl-Pol  Acctg Pol Stats Description QGrp-Instance
-------------------------------------------------------------------------------
7/1/2                                   No                    1
-------------------------------------------------------------------------------
---snip---
===============================================================================
Queue Group Policer
===============================================================================
Policer Id     : 1
Description    : (Not Specified)
PIR Adptn      : closest                CIR Adptn     : closest
Parent         : none                   Level         : 1
Weight         : 1                      Adv. Cfg Plcy: none
Admin PIR      : 4000                   Admin CIR     : 0
CBS            : def                    MBS           : def
Hi Prio Only   : def                    Pkt Offset    : -22
Profile Capped : Disabled
StatMode       : minimal
```

```
---snip---
```

The following command shows where the ingress queue group has been applied.

```
*A:PE-1# show qos queue-group ingress association
===============================================================================
QoS Queue-Group Ingress
===============================================================================
---snip---
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name     : ingress-queue-group
Description    : (Not Specified)
---snip---
===============================================================================
Queue Group FP Maps
===============================================================================
Card Num       Fp Num              Instance            Type
-------------------------------------------------------------------------------
7              1                   1                   Network
-------------------------------------------------------------------------------
Entries found: 1
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

The following command shows where the egress queue group has been applied.

```
*A:PE-1# show qos queue-group egress association
===============================================================================
QoS Queue-Group Egress
===============================================================================
-------------------------------------------------------------------------------
QoS Queue Group
-------------------------------------------------------------------------------
Group-Name     : egress-queue-group
Description    : (Not Specified)
---snip---
===============================================================================
Queue Group Ports (network)
===============================================================================
Port   Sched Pol    Policer-Ctrl-Pol  Acctg Pol Stats Description QGrp-Instance
-------------------------------------------------------------------------------
7/1/2                                           No              1
-------------------------------------------------------------------------------
---snip---
===============================================================================
```

The following command shows the ingress queue group applied to the FP on card 7.

```
*A:PE-1# show card 7 fp 1 ingress queue-group "ingress-queue-group" instance 1
mode network
===============================================================================
Card:7  Net.QGrp: ingress-queue-group  Instance: 1
===============================================================================
Group Name    : ingress-queue-group
Description   : (Not Specified)
Pol Ctl Pol   : None                      Acct Pol      : None
Collect Stats : disabled
```

The following command show the details of the policers in the ingress FP queue group.

```
*A:PE-1# show qos policer card 7 fp 1 queue-group "ingress-queue-group" instance 1 network
 detail
===============================================================================
Policer Info (Net-FPQG-1-ingress-queue-group:1->1), Slot 7
===============================================================================
Policer Name       : Net-FPQG-1-ingress-queue-group:1->1
Direction          : Ingress         Fwding Plane     : 1
Depth PIR          : 0 Bytes         Depth CIR        : 0 Bytes
Depth FIR          : 0 Bytes
MBS                : 7680 B          CBS              : 0 KB
Hi Prio Only       : 768 B           Pkt Byte Offset  : 4
Admin PIR          : 6000 Kbps       Admin CIR        : 0 Kbps
Oper PIR           : 6000 Kbps       Oper CIR         : 0 Kbps
Oper FIR           : 6000 Kbps
Stat Mode          : minimal
PIR Adaption       : closest         CIR Adaption     : closest
Adv.Cfg Plcy       : None            Profile Capped   : disabled
Parent Arbiter Name: (Not Specified)
-------------------------------------------------------------------------------
Arbiter Member Information
-------------------------------------------------------------------------------
Offered Rate       : 0 Kbps
Level              : 0               Weight           : 0
Parent PIR         : 0 Kbps          Parent FIR       : 0 Kbps
Consumed           : 0 Kbps
-------------------------------------------------------------------------------
===============================================================================
Policer Info (Net-FPQG-1-ingress-queue-group:1->2), Slot 7
===============================================================================
Policer Name       : Net-FPQG-1-ingress-queue-group:1->2
Direction          : Ingress         Fwding Plane     : 1
Depth PIR          : 0 Bytes         Depth CIR        : 0 Bytes
Depth FIR          : 0 Bytes
MBS                : 5 KB            CBS              : 0 KB
Hi Prio Only       : 512 B           Pkt Byte Offset  : 4
Admin PIR          : 4000 Kbps       Admin CIR        : 0 Kbps
Oper PIR           : 4000 Kbps       Oper CIR         : 0 Kbps
Oper FIR           : 4000 Kbps
Stat Mode          : minimal
PIR Adaption       : closest         CIR Adaption     : closest
Adv.Cfg Plcy       : None            Profile Capped   : disabled
Parent Arbiter Name: (Not Specified)
-------------------------------------------------------------------------------
Arbiter Member Information
-------------------------------------------------------------------------------
Offered Rate       : 0 Kbps
Level              : 0               Weight           : 0
Parent PIR         : 0 Kbps          Parent FIR       : 0 Kbps
Consumed           : 0 Kbps
-------------------------------------------------------------------------------
===============================================================================
Network Interface Association
-------------------------------------------------------------------------------
No Association Found.
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
SDP Association
-------------------------------------------------------------------------------
Policer Info (1->1:1->10), Slot 7
-------------------------------------------------------------------------------
SDP Association Count : 1
```

```
--------------------------------------------------------------------------------
```

The details of the queue and policer in the egress queue group applied to port 7/1/2 can also be shown as follows.

```
*A:PE-1# show port 7/1/2 queue-group egress "egress-queue-group" network instance 1
===============================================================================
Ethernet port 7/1/2 Network Egress queue-group
===============================================================================
Group Name    : egress-queue-group Instance-Id   : 1
Description   : (Not Specified)
Sched Policy  : None               Acct Pol      : None
Collect Stats : disabled           Agg. Limit    : -1

Queues
-------------------------------------------------------------------------------
Queue-Group   : egress-queue-group Instance-Id   : 1      Queue-Id   : 1
Description   : (Not Specified)
Admin PIR     : 6000*              Admin CIR     : 0*
PIR Rule      : closest*           CIR Rule      : closest*
CBS           : def*               MBS           : def*
Hi Prio       : def*

Policers
-------------------------------------------------------------------------------
Queue-Group   : egress-queue-group Instance-Id   : 1      Policer-Id   : 1
Description   : (Not Specified)
Admin PIR     : 4000*              Admin CIR     : 0*
PIR Rule      : closest*           CIR Rule      : closest*
CBS           : def*               MBS           : def*
Hi Prio       : def*

* means the value is inherited
```

The network QoS policy can be shown with the details of the configured FC redirection and ingress classification used on the pseudowire, as follows.

```
*A:PE-1# show qos network 10 detail
===============================================================================
QoS Network Policy
===============================================================================
-------------------------------------------------------------------------------
Network Policy (10)
-------------------------------------------------------------------------------
Policy-id      : 10                  Remark          : False
Forward Class  : be                  Profile         : Out
LER Use DSCP   : False
Description    : (Not Specified)
---snip---
-------------------------------------------------------------------------------
LSP EXP Bit Map                      Forwarding Class          Profile
-------------------------------------------------------------------------------
5                                    ef                        In
---snip---
-------------------------------------------------------------------------------
Egress Forwarding Class Mapping
-------------------------------------------------------------------------------
FC Value       : 0                   FC Name         : be
- DSCP Mapping
Out-of-Profile : be                  In-Profile      : be
---snip---
DE Mark        :   None
```

```
Redirect Grp Q   :  1                       Redirect Grp Plcr:  None

---snip---
FC Value         : 5                        FC Name          : ef
---snip---
DE Mark          :  None
Redirect Grp Q   :  None                    Redirect Grp Plcr:  1


-------------------------------------------------------------------------------
Ingress Forwarding Class Mapping
-------------------------------------------------------------------------------
FC Value               : 0                  FC Name                : be
Redirect UniCast Plcr  : 1                  Redirect MultiCast Plcr : None

---snip---
FC Value               : 5                  FC Name                : ef
Redirect UniCast Plcr  : 2                  Redirect MultiCast Plcr : None
---snip---
```

The details of the configuration of the pseudowire QoS can be seen when showing the details of the SDP within the Epipe service, as follows.

```
*A:PE-1# show service id 1 sdp 1:1 detail
===============================================================================
Service Destination Point (Sdp Id : 1:1) Details
===============================================================================
-------------------------------------------------------------------------------
 Sdp Id 1:1  -(192.0.2.2)
-------------------------------------------------------------------------------
Description      : (Not Specified)
SDP Id           : 1:1                      Type             : Spoke
Spoke Descr      : (Not Specified)
VC Type          : VLAN                     VC Tag           : 0
Admin Path MTU   : 0                        Oper Path MTU    : 9190
Delivery         : MPLS
Far End          : 192.0.2.2
Tunnel Far End   : 192.0.2.2                LSP Types        : LDP
Hash Label       : Disabled                 Hash Lbl Sig Cap  : Disabled
Oper Hash Label  : Disabled

Admin State      : Up                       Oper State       : Up
---snip---
Ingress Qos Policy : 10                      Egress Qos Policy : 10
Ingress FP QGrp    : ingress-queue-group     Egress Port QGrp  : egress-queue*
Ing FP QGrp Inst   : 1                       Egr Port QGrp Inst: 1
```

The usage of the "Egr Network Queue Group Mappings" out of the total number of "Dynamic Service Entries" can be seen with the following command. Only one egress QoS pseudowire redirection has been configured.

```
*A:PE-1# tools dump system-resources
Resource Manager info at 005 m 07/31/13 13:11:03.355:

Hardware Resource Usage for Slot #7, CardType imm5-10gb-xfp, Cmplx #0:
                                | Total    | Allocated |   Free
   ----------------------------+----------+----------+------------
---snip---
       Dynamic Service Entries |     65535|         1|      65534
             Subscriber Hosts |          |         0|
           Encap Group Members |          |         0|
Egr Network Queue Group Mappings |        |         1|
```

It is possible to show the statistics on the ingress FP queue group used by the pseudowire.

```
*A:PE-1# show card 7 fp 1 ingress queue-group "ingress-queue-group" instance 1 mode network
 statistics

===============================================================================
Card:7  Net.QGrp: ingress-queue-group  Instance: 1
===============================================================================
Group Name    : ingress-queue-group
Description   : (Not Specified)
Pol Ctl Pol   : None                      Acct Pol       : None
Collect Stats : disabled
-------------------------------------------------------------------------------
Statistics
-------------------------------------------------------------------------------
                        Packets                   Octets

Ing. Policer:  1  Grp: ingress-queue-group (Stats mode: minimal)
Off. All             :          184275               23587200
Dro. All             :           36801                4710528
For. All             :          147474               18876672

Ing. Policer:  2  Grp: ingress-queue-group (Stats mode: minimal)
Off. All             :          184274               23587072
Dro. All             :           85955               11002240
For. All             :           98319               12584832
```

Similar statistics can be shown for the egress port queue group used by the pseudowire.

```
*A:PE-1# show port 7/1/2 queue-group egress "egress-queue-group" network statistics instance 1
-------------------------------------------------------------------------------
Ethernet port 7/1/2 Network Egress queue-group
-------------------------------------------------------------------------------
                        Packets                   Octets
Egress Queue:  1    Group: egress-queue-group    Instance-Id:  1
In Profile forwarded  : 0                         0
In Profile dropped    : 0                         0
Out Profile forwarded : 150989                    19326592
Out Profile dropped   : 37123                     4751744

Egress Policer:  1  Group: egress-queue-group   Instance-Id: 1
Stats mode: minimal
Off. All             : 188421                     24117888
Dro. All             : 87894                      11250432
For. All             : 100527                     12867456
```

Monitor commands are available to see the statistics (and rates on egress port queue group). As an example, the following shows the utilization on the queue and policer in the egress queue-group.

```
*A:PE-1# monitor port 7/1/2 queue-group "egress-queue-group" instance 1 egress network egress-
queue 1 repeat 1 rate
===============================================================================
Monitor Port Queue-Group Egress Network Queue Statistics
===============================================================================
-------------------------------------------------------------------------------
At time t = 0 sec (Base Statistics)
-------------------------------------------------------------------------------
                        Packets                   Octets

In Profile forwarded  : 0                         0
In Profile dropped    : 0                         0
Out Profile forwarded : 299113                    38286464
```

```
Out Profile dropped   : 74155                    9491840
--------------------------------------------------------------------------
At time t = 11 sec (Mode: Rate)
--------------------------------------------------------------------------
                        Packets/sec           Octets/sec            % Port
                                                                    Util.

In Profile forwarded  : 0                     0                     0.00
In Profile dropped    : 0                     0                     0.00
Out Profile forwarded : 5863                  750436                0.06
Out Profile dropped   : 1466                  187609                0.01
==========================================================================
```

```
*A:PE-1# monitor port 7/1/2 queue-group "egress-queue-group" instance 1 egress network policer
 1 repeat 1 rate
==========================================================================
Monitor Port Queue-Group Egress Network Policer Statistics
==========================================================================
--------------------------------------------------------------------------
At time t = 0 sec (Base Statistics)
--------------------------------------------------------------------------
                        Packets               Octets

Off. All              : 454750                58208000
Dro. All              : 212181                27159168
For. All              : 242569                31048832

--------------------------------------------------------------------------
At time t = 11 sec (Mode: Rate)
--------------------------------------------------------------------------
                        Packets/sec           Octets/sec            % Port
                                                                    Util.

Off. All              : 7326                  937716                0.07
Dro. All              : 3419                  437609                0.03
For. All              : 3907                  500108                0.04
==========================================================================
*A:PE-1#
```

As mentioned, the egress policer (FC EF) traffic exits the egress port by default using the related network queue on the port, as follows.

```
*A:PE-1# show port 7/1/2 detail
==========================================================================
Ethernet Interface
==========================================================================
Description         : 10-Gig Ethernet
Interface           : 7/1/2              Oper Speed       : 10 Gbps
Link-level          : Ethernet           Config Speed     : N/A
Admin State         : up                 Oper Duplex      : full
Oper State          : up                 Config Duplex    : N/A
---snip---
==========================================================================
Queue Statistics
==========================================================================
--------------------------------------------------------------------------
---snip---
Egress Queue  6              Packets               Octets
    In Profile forwarded  :    0                     0
    In Profile dropped    :    0                     0
    Out Profile forwarded :    102381                15357150
    Out Profile dropped   :    0                     0
```

The throughput achieved using the preceding configuration can be verified in the traffic generator output. Port 202/1 is connected to PE-1 and port 204/1 is connected to PE-2.

Port	Tx Test Packets	Rx Test Packets	Tx Test Octets	Rx Test Octets	Tx Test Throughput (Mb/s)	Rx Test Throughput (Mb/s)	Rx Packet Loss	Average Latency (us)
All Ports	29296	19531	3749888	2499968	29.999	20.000	n/a	15512.18
202/1	14648	9765	1874944	1249920	15.000	9.999	n/a	39.28
202/1->204/1, BE traffic	7324	5860	937472	750080	7.500	6.001	1464	51609.56
202/1->204/1, EF traffic	7324	3906	937472	499968	7.500	4.000	3418	39.13
204/1	14648	9766	1874944	1250048	15.000	10.000	n/a	30983.50
204/1->202/1, BE traffic	7324	5859	937472	749952	7.500	6.000	1465	39.28
204/1->202/1, EF traffic'	7324	3906	937472	499968	7.500	4.000	3418	39.27

Conclusion

This example has shown the configuration and monitoring of pseudowire QoS, providing a powerful QoS solution for pseudowire applications. QoS can be applied independently to the ingress and/or egress of a single pseudowire or multiple pseudowires.

QoS Architecture and Basic Operation

This chapter provides information about QoS architecture and basic operation.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

The chapter was initially written for SR OS Release 9.0.R3. The CLI in the current edition corresponds to SR OS Release 14.0.R2.

Overview

The SR-series platforms provide an extensive Quality of Service (QoS) capability for service provider solutions. QoS is a system behavior to provide different traffic with different amounts of resources, including buffer memory and queue serving time.

By allocating system resources with certain degrees of guarantee, the bandwidth can be used more efficiently and more controllably. Lack of buffer memory leads to packet drop, while a smaller amount of queue serving time normally means longer delay for the packet and may cause buffer memory to be completely consumed and eventually also lead to packet drop.

In a system, such as the SR-series platform, different types of traffic contend for the same resources at several major points, such as the ingress to the switch fabric and the egress out of a physical port. In a multi-node network, QoS is achieved on hop by hop basis. Thus, QoS needs to be configured individually but with the consistency across the whole network.

This chapter is focused on the configuration of the basic QoS, namely the use of queues to shape traffic at the ingress and egress of the system. More sophisticated aspects will be referenced where appropriate but their details are beyond the scope of this chapter. Other topics not included are Hierarchical QoS scheduling, egress port-scheduler, queue-groups, named buffer pools, WRED-per-queue, LAGs, high scale MDA, QoS for ATM/FR and Enhanced Subscriber Management.

QoS Components

QoS consists of four main components:

- Classification
- Buffering (enqueuing)

- Scheduling (dequeuing)

- Remarking

These are also the fundamental building blocks of the QoS implementation in the SR-series platforms. Ingress packets, classified by various rules, belong to one of eight Forwarding Classes (FC). An FC can be viewed a set of packets which are treated in a similar manner within the system (have the same priority level and scheduling behavior). Each FC has a queue associated with it and each queue has a set of parameters controlling how buffer memory is allocated for packets; if a packet is enqueued (placed on the queue) a scheduler will control the way the packet gets dequeued (removed from the queue) relative to other queues. When a packet exits an egress port, a remarking action can be taken to guarantee the next downstream device will properly handle the different types of traffic.

Configuration

Policies

QoS policies are used to control how traffic is handled at distinct points in the service delivery model within the device. There are different types of QoS policies catering to the different QoS needs at each point. QoS policies only take effect when applied to a relevant entity (Service Access Point (SAP) or network port/interface) so by default can be seen as templates with each application instantiating a new set of related resources.

The following QoS policies are discussed:

- Ingress/egress QoS Policies — For classification, queue attributes and remarking.

- Slope policies — Define the RED slope definitions.

- Scheduler policies — Determine how queues are scheduled (only the default scheduling is included here).

Access, Network, and Hybrid Ports

The system has two different types of interfaces: access and network.

- A network interface will classify packets received from the network core at ingress and remark packets sent to the core at egress. Aggregated differentiated service QoS is performed on network ports, aggregating traffic from multiple services into a set of queues per FC.

- An access interface connects to one or more customer devices; it receives customer packets, classifies them into different FCs at ingress and remarks packets according to FCs at egress. Since an access interface needs application awareness, it has many more rules to classify the ingress packets. Access and network also differ in how buffer memory is handled, as will be made clear when discussing the buffer management. Here the QoS is performed per SAP.

Access interfaces (SAPs) are configured on access ports and network interfaces are configured on network ports. A third type of port is available, the hybrid port, which supports both access and network interfaces on the same port.

Hybrid ports are only supported on Ethernet ports and optionally with a single-chassis LAG. They must be configured to use VLANs (either single (dot1q encapsulation) or double (QinQ encapsulation) tagging) with each VLAN mapping to either the access or network part of the port. This allows the classification to

associate incoming traffic with the correct port type and service. Port based traffic, such as LACP, CCM and EFM, uses a system queue on an access port, but the default network queues on a network or hybrid port.

Customer traffic follows the path shown below:

[service ingress → network egress]   →   [network ingress → network egress]   →   [network ingress → service egress]

ingress PE   transit P   egress PE

The network administrator needs to make sure that QoS is configured correctly at each point using the appropriate QoS policies (Figure 17: Service and Network QoS Policies).

Figure 17: Service and Network QoS Policies

Service Ingress QoS Policy

The SAP ingress policies are created in the **qos** context of the CLI and require a unique identifier (from 1 to 65535). The default **sap-ingress** policy has identifier 1.

Classification

Services can be delineated at the SAP ingress by

- A physical port (null encapsulated) or
- An encapsulation on the physical port, for example a VLAN ID on an Ethernet port or a DLCI on a Frame Relay port.

The following configuration is an example of an IES service created with an IP interface on VLAN 2 of port 3/2/10 (IOM 3, MDA 2, port 10) and has SAP ingress QoS policy 10 applied.

```
configure
    service
        ies 1 customer 1 create
            interface "int-access" create
                address 192.168.1.1/30
                sap 1/1/1:1 create
                    ingress
                        qos 10
                    exit
                exit
            exit
            no shutdown
        exit
```

As traffic enters the port, the service can be identified by the VLAN tag (and unwanted packets dropped). The ingress service QoS policy applied to the SAP maps traffic to FCs, and thus to queues, and sets the enqueuing priority. Mapping flows to FCs is controlled by comparing each packet to the match criteria in the QoS policy. The match criteria that can be used in ingress QoS policies can be combinations of those listed in Table 7: SAP Ingress Classification Match Criteria. When a packet matches two criteria (802.1p priority and DSCP) it is the lowest precedence value that is used to map the packet to the FC.

Table 7: SAP Ingress Classification Match Criteria

Match Precedence	Match Criteria		
1	IPv4 fields match criteria: • Destination IP address/prefix including prefix list • Destination port/range • DSCP value • IP fragment • Protocol type (TCP, UDP, etc.) • Source port/range • Source IP address/prefix including prefix list	IPv6 fields match criteria: • Destination IP address/prefix • Destination port/range • DSCP value • Next header • Source port/range • Source IP address/prefix	MAC fields match criteria: • Frame type [802dot3|802dot2-llc|802dot2-snap|ethernetII|atm] • ATM VCI value • IEEE 802.1p value/mask • Source MAC address/mask • Destination MAC address/mask • EtherType value • IEEE 802.2 LLC SSAP value/mask • IEEE 802.2 LLC DSAP value/mask • IEEE 802.3 LLC SNAP OUI zero or non-zero value • IEEE 802.3 LLC SNAP PID value
	Note: For an ingress QoS policy, either IP match criteria or MAC match criteria can be defined, not both.		
2	DSCP		

Match Precedence	Match Criteria
3	IP precedence
4	LSP EXP
5	IEEE 802.1p priority and/or Drop Eligibility Indicator (DEI)
6	Default forwarding class for non-matching traffic

It is possible to match MAC criteria on VPLS/Epipe SAPs and IP criteria on IP interface SAPs. However, it is also possible to classify on MAC criteria on an IP interface SAP and conversely to classify on IP criteria on VPLS/Epipe SAPs. When MPLS labeled traffic is received on a VPLS/Epipe SAP, it is possible to match on either of the LSP EXP bits (outer label) or the MAC criteria.

A SAP can be configured to have no VLAN tag (null encapsulated), one VLAN tag (dot1q encapsulated) or two VLAN tags (QinQ encapsulated). The configuration allows the selection of which VLAN tag to match against for the 802.1p bits, using the keyword **match-qinq-dot1p** with the keyword **top** or **bottom**.

The following example configuration shows match QinQ traffic with dot1p value 1 in the top VLAN tag entering the QinQ SAP in Epipe service 1 and assign it to FC *af* using queue 2.

```
configure
    qos
        sap-ingress 10 create
            queue 2 create
            exit
            fc "af" create
                queue 2
            exit
            dot1p 1 fc "af"
        exit
```

```
configure
    service
        epipe 2 customer 1 create
            sap 1/1/2:1.2 create
                ingress
                    qos 10
                exit
                ingress
                    match-qinq-dot1p top
                exit
            exit
            no shutdown
        exit
```

The classification of traffic using the default, **top** and **bottom** keyword parameters is summarized in Table 8: QinQ Dot1p Bit Classification . A TopQ SAP is a QinQ SAP where only the outer (top) VLAN tag is explicitly specified (sap 1/1/1:10.* or sap 1/1/1:10.0).

Table 8: QinQ Dot1p Bit Classification

Port/SAP Type	Existing Packet Tags	Pbits Used for Match		
		Default	**Match Top**	**Match Bottom**
Null	None	None	None	None
Null	Dot1P (VLAN-ID 0)	Dot1P PBits	Dot1P PBits	Dot1P PBits
Null	Dot1Q	Dot1Q PBits	Dot1Q PBits	Dot1Q PBits
Null	TopQ BottomQ	TopQ PBits	TopQ PBits	BottomQ PBits
Null	TopQ (No BottomQ)	TopQ PBits	TopQ PBits	TopQ PBits
Dot1Q	None (Default SAP)	None	None	None
Dot1Q	Dot1P (Default SAP VLAN-ID 0)	Dot1P PBits	Dot1P PBits	Dot1P PBits
Dot1Q	Dot1Q	Dot1Q PBits	Dot1Q PBits	Dot1Q PBits
QinQ/TopQ	TopQ	TopQ PBits	TopQ PBits	TopQ PBits
QinQ/TopQ	TopQ BottomQ	TopQ PBits	TopQ PBits	BottomQ PBits
QinQ/QinQ	TopQ BottomQ	BottomQ PBits	TopQ PBits	BottomQ Pbits

The Drop Eligibility Indicator (DEI) bit (IEEE 802.1ad-2005 and IEEE 802.1ah (PBB)) can be used to indicate the in/out profile state of the packet, this will be covered later in the discussion on profile mode.

Ingress traffic with a local destination (for example, OSPF hellos) is classified by the system automatically and uses a set of dedicated system queues.

After the traffic has been classified, the next step is to assign it to a given FC. There are 8 pre-defined FCs within the system which are shown in Table 10: Queue Priority vs. Profile Mode (the FC identifiers are keywords and do not have a fixed relationfship with the associated Differentiated Services Code Points (DSCP)).

Table 9: Forwarding Classes

FC Identifier	FC Name	Default Scheduling Priority
NC	Network Control	Expedited
H1	High-1	Expedited
EF	Expedited	Expedited
H2	High-2	Expedited
L1	Low-1	Best Effort

FC Identifier	FC Name	Default Scheduling Priority
AF	Assured	Best Effort
L2	Low-2	Best Effort
BE	Best Effort	Best Effort

When an FC is configured for a classification, it must first be created in the configuration. One of the FCs can be also configured to be the default in case there is no explicit classification match and by default this FC is *be*.

Normally, once traffic is assigned to an FC at the ingress it remains in that FC throughout its time within the system. Re-classification of IP traffic at a SAP egress is possible, but is beyond the scope of this chapter. The FC used at egress can also be specified to be different than that used at ingress by configuring **egress-fc** under the FC configuration in the SAP ingress policy.

Packets also have a state of being in-profile or out-of-profile which represents their drop precedence within the system, therefore there can be up to 8 distinct per hop behavior (PHB) classes with two drop precedences.

Buffering (Enqueuing)

Once a packet is assigned to a certain forwarding class, it will try to get a buffer in order to be enqueued. Whether the packet can get a buffer is determined by the instantaneous buffer utilization and several attributes of the queue (such as Maximum Burst Size (MBS), Committed Burst Size (CBS) and high-prio-only) that will be discussed in more detail later in this chapter. If a packet cannot get a buffer for whatever reason, the packet will get dropped immediately.

As traffic is classified at the SAP ingress it is also assigned an enqueuing priority, which can be high or low. This governs the likelihood of a packet being accepted into a buffer and so onto a queue, and is managed using the queue's high-prio-only parameter and the buffer pools weighted random early detection (WRED) slope policies. Traffic having a high enqueuing priority has more chance of getting a buffer than traffic with low enqueuing priority. The enqueuing priority is specified with the classification using the parameter **priority**, and a default enqueuing priority can be configured, its default being low.

Enqueuing priority is a property of a packet and should not to be confused with scheduling priority, expedited or best-effort, which is a property of a queue.

The following configuration shows an example where all packets with dot1p value 3 are classified as *ef* and have their enqueuing priority set to high, while all other packets are classified as *af* with a low enqueuing priority.

```
configure
    qos
        sap-ingress 10 create
            fc "af" create
            exit
            fc "ef" create
            exit
            dot1p 3 fc "ef" priority high
            default-fc "af"
            default-priority low # this is the default
        exit
```

Each forwarding class is associated with at most one unicast queue. In the case of a VPLS service, each FC can also be assigned a single multipoint queue at ingress, or for more granular control, separate queues for broadcast, multicast and unknown traffic. Since each queue maintains forward/drop statistics, it allows the network operator to easily track unicast, broadcast, multicast and unknown traffic load per forwarding class. Separate multicast queues can also be assigned for IES/VPRN services which have IP multicast enabled.

This results in an Epipe SAP having up to 8 ingress queues, an IES/VPRN SAP having up to 16 ingress queues and a VPLS SAP having up to 32 ingress queues. Each queue has a locally significant (to the policy) identifier, which can be from 1 to 32.

The default SAP ingress QoS policy (id=1) has two queues; queue 1 for unicast traffic and queue 11 for multipoint traffic, and is assigned to every ingress SAP at service creation time. Equally, when a new (non-default) SAP ingress policy is created, queue 1 and queue 11 are automatically created with all FCs assigned to both by default. Additional queues must be created before being assigned to a FC, with multipoint queues requiring the **multipoint** keyword. When a SAP ingress policy is applied to a SAP, physical hardware queues on the FP are allocated for each queue with a FC assigned (if no QoS policy is explicitly configured, the default policy is applied). Multipoint queues within the SAP ingress policy are ignored when applied to an Epipe SAP or an IES/VPRN SAP which is not configured for IP multicast.

The mechanism described here uses a separate set of queues per SAP. For cases where per-SAP queuing is not required it is possible to use port based queues, known as **queue-groups**, which reduces the number of queues required, as described in chapter *FP and Port Queue Groups*.

Scheduling (Dequeuing)

A queue has a priority which affects the relative scheduling of packets from it compared to other queues. There are two queue priorities: expedited and best-effort, with expedited being the higher. When creating a queue, one of these priorities can be configured, thereby explicitly setting the queue's priority. Alternatively, the default is auto-expedite in which case the queue's priority is governed by the FCs assigned to it, as shown in Table 9: Forwarding Classes. If there is a mix of expedited and best-effort FCs assigned, the queue is deemed to be best-effort.

The following configuration displays an example that ensures that EF traffic is treated as expedited by assigning it to new unicast and multicast queues.

```
configure
    qos
        sap-ingress 10 create
            queue 3 expedite create
            exit
            queue 13 multipoint expedite create
            exit
            fc "ef" create
                queue 3
                multicast-queue 13
            exit
            default-fc "ef"
        exit
```

Once a packet gets a buffer and is queued, it will wait to be served and sent through the switch fabric to its destination port by the hardware scheduler. There are two scheduler priorities: expedited or best-effort, corresponding to the queue's priority. The expedited hardware schedulers are used to enforce priority access to internal switch fabric destinations with expedited queues having a higher preference than best-

effort queues. Queues of the same priority get equally serviced in round robin fashion by the associated scheduler.

When a queue gets its turn to be serviced, the scheduler will use the operational Peak Information Rate (PIR) and Committed Information Rate (CIR) attributes of the queue to determine what to do with the packet.

- The scheduler does not allow queues to exceed their configured PIR. If the packet arrival rate for a given queue is higher than the rate at which it is drained, the queue will fill. If the queue size (in bytes or Kbytes) reaches its defined MBS all subsequent packets will be discarded, this is known as tail drop.

- If the dequeue rate is below or equal to the operational CIR, the packet will be forwarded and marked as **in-profile**.

- If the dequeue rate is below or equal to the operational PIR but higher than the CIR, the packet will be forwarded but marked as **out-of-profile**.

Out-of-profile packets have a higher probability of being dropped when there is congestion somewhere in the downstream path. Packets that are marked as out-of-profile will also be treated differently at the network egress and service egress.

These marking actions are known as color marking (green for in-profile and yellow for out-of-profile). Using the default queue setting of **priority-mode**, as described above, the in/out-of-profile state of a packet is determined from the queue scheduling state (within CIR or above CIR, as described later) at the time that the packet is dequeued. An alternative queue mode is **profile-mode**.

Profile Mode

A queue is created with profile mode when the aim is that the in/out-of-profile state of packets is determined by the QoS bits of the incoming packets, this is known as color-aware (as opposed to color-unaware for priority mode).

As part of the classification, the profile state of the packets is explicitly configured. To provide granular control, it is possible to configure FC sub-classes with each having a different profile state, while inheriting the other parameters from their parent FC (for example the queue, in order to avoid out of order packets). The FC subclasses are named *fc.sub-class*, where *sub-class* is a text string up to 29 characters (though normally the words *in* and *out* are used for clarity). Any traffic classified without an explicit profile state is treated as if the queue were in priority mode.

When using the profile mode, the DEI in the Ethernet header can be used to classify a packet as in-profile (DEI=0) or out-of-profile (DEI=1).

The following configuration shows traffic with dot1p 3 is set to in-profile, dot1p 2 to out-of-profile and the profile state of dot1p 0 depends on the scheduling state of the queue.

```
configure
    qos
        sap-ingress 20 create
            queue 2 profile-mode create
            exit
            fc "af" create
                queue 2
            exit
            fc "af.in" create
                profile in
            exit
            fc "af.out" create
                profile out
```

```
                exit
                dot1p 0 fc "af"
                dot1p 2 fc "af.out"
                dot1p 3 fc "af.in"
            exit
```

The difference between a queue configured in priority (default) and profile mode is summarized in Table 10: Queue Priority vs. Profile Mode (within/above CIR is described later).

Table 10: Queue Priority vs. Profile Mode

	Priority Mode	**Profile Mode**
Packet In-Profile/ Out-of-Profile state	Determined by state of the queue at scheduling time. Within CIR – In Profile Above CIR – Out Profile	Explicitly stated in FC or subclass classification. If not, then defaults to state of the queue at scheduling time
Packet High/Low Enqueuing Priority	Explicitly stated in FC classification. If not, then defaults to Low priority	Always follows state of in-profile/out-of-profile determined above In-profile = High Priority Out-Profile = Low Priority If not set = High Priority

Remarking

Remarking at the service ingress is possible when using an IES or VPRN service. The DSCP/precedence field can be remarked for in-profile (**in-remark**) and out-of-profile (**out-remark**) traffic as defined above for queues in either priority mode or profile mode. If configured for other services, the remarking is ignored. If remarking is performed at the service ingress, then the traffic is not subject to any egress remarking on the same system.

The following configuration displays an example classifying traffic to 10.0.0.0/8 as FC *ef* in-profile and remark its DSCP to *ef*.

```
configure
    qos
        sap-ingress 30 create
            queue 2 profile-mode create
            exit
            fc "ef" create
                queue 2
                in-remark dscp ef
                profile in
            exit
            ip-criteria
                entry 10 create
                    match
                        dst-ip 10.0.0.0/8
                    exit
                    action fc "ef"
                exit
            exit
```

```
        exit
```

Service Egress QoS Policy

The service egress uses a SAP egress QoS policy to define how FCs map to queues and how a packet of an FC is remarked. SAP egress policies are created in the CLI qos context and require a unique identifier (from 1 to 65535). The default SAP egress policy has identifier 1.

Once a service packet is delivered to the egress SAP, it has following attributes:

- Forwarding class, determined from classification at the ingress of the node.
- In/out-of-profile state from the service ingress or network ingress.

Similar to the service ingress enqueuing process, it is possible that a packet cannot get a buffer and thus gets dropped. Once on an egress queue, a packet is scheduled from the queue based on priority of the queue (expedited or best-effort) and the scheduling state with respect to the CIR/PIR rates (the profile state of the packet [in/out] is not modified here). Egress queues do not have a priority/profile mode and have no concept of multipoint.

Only one queue exists in the default SAP egress QoS policy (id=1) and also when a new **sap-egress** policy is created, this being queue 1 which is used for both unicast traffic and multipoint traffic. All FCs are assigned to this queue unless otherwise explicitly configured to a different configured queue. When a SAP egress policy is applied to a SAP, physical hardware queues on the FP are allocated for each queue with FC assigned (if no QoS policy is explicitly configured, the default policy is applied).

As mentioned earlier, re-classification at a SAP egress is possible based on the packet's dot1p, DSCP or precedence values or using IP or IPv6 criteria matching, similar to the functionality at SAP ingress.

SAP egress also supports two additional profiles, inplus-profile and exceed-profile. Both can be assigned to a packet using egress reclassification and the exceed-profile can be assigned to a packet in an egress policer configured with the **enable-exceed-pir** command.

Traffic originated by the system (known as self generated traffic) has its FC and marking configured under router/sgt-qos (for the base routing) or under service/vprn/sgt-qos (for a VPRN service). This is beyond the scope of this chapter.

Remarking

At the service egress, the dot1p/DEI can be remarked for any service per FC with separate marking for in/out/exceed profile if required (inplus-profile packets are marked with the same value as in-profile packets and exceed-profile packets are marked with the same value as out-of-profile packets if not explicitly configured). The DEI bit can also be forced to a specific value (using the **de-mark force** command). When no **dot1p/de-mark** is configured, the ingress dot1p/DEI is preserved; if the ingress was untagged, the dot1p/DEI bit is set to 0.

The following configuration shows a remark example with different FCs with different dot1p values. FC *af* also differentiates between in/out-of-profile and then remarks the DEI bit accordingly based on the packet's profile.

```
configure
    qos
        sap-egress 10 create
            queue 1 create
                rate 20000
```

```
                exit
                queue 2 create
                    rate 10000 cir 5000
                exit
                queue 3 create
                    rate 2000 cir 2000
                exit
                fc af create
                    queue 2
                    dot1p in-profile 3 out-profile 2
                    de-mark
                exit
                fc be create
                    queue 1
                    dot1p 0
                exit
                fc ef create
                    queue 3
                    dot1p 5
                exit
            exit
```

If QinQ encapsulation is used, the default is to remark both tags in the same way. However, it is also possible to remark only the top tag using the **qinq-mark-top-only** parameter configured under the SAP egress.

The following configuration shows a remark example with only the dot1p/DEI bits in top tag of a QinQ SAP.

```
configure
    service
        vpls 3 customer 1 create
            sap 1/1/2:2.2 create
                egress
                    qos 20
                    qinq-mark-top-only
                exit
            exit
            no shutdown
        exit
```

For IES and VPRN services, the DSCP/precedence field can also be remarked based on the in/out/exceed-profile state (with inplus-profile packets marked with the same values as in-profile packets, and exceed-profile packets marked with the same value as out-of-profile packets if not explicitly configured) of the packets (and only if no ingress remarking was performed).

The following configuration shows DSCP values for FC *af* based on in/out-of-profile traffic.

```
configure
    qos
        sap-egress 20 create
            queue 2 create
            exit
            fc af create
                queue 2
                dscp in-profile af41 out-profile af43
            exit
        exit
```

Network Ports

The QoS policies relating to the network ports are divided into a network and a network-queue policy. The network policy covers the ingress and egress classification into FCs and the egress remarking based on FCs, while the network-queue policy covers the queues parameters and the FC to queue mapping. The logic behind this is that there is only one set of queues provisioned on a network port, whereas the use of these queues is configured per network IP interface. This in turn determines where the two policies can be applied. Network ports are used for IP routing and switching, and for GRE/MPLS tunneling.

Network QoS Policy

The network QoS policy has an ingress section and an egress section. It is created in the **qos** context of the CLI and requires a unique identifier (from 1 to 65535). The default network policy has identifier 1. Network QoS policies are applied to IP interfaces configured on a network port.

The following configuration shows an example to apply different network QoS policies to two network interfaces.

```
configure
    router
        interface "int-PE-1-PE-2"
            address 192.168.12.1/30
            port 1/1/3:1
            qos 28
        exit
        interface "int-PE-1-PE-3"
            address 192.168.13.1/30
            port 1/1/4
            qos 18
        exit
```

Classification

Classification is available at both ingress and egress.

The ingress section defines the classification rules for IP/MPLS packets received on a network IP interface. The rules for classifying traffic are based on the incoming QoS bits (Dot1p, DSCP, EXP [MPLS experimental bits]). The order in which classification occurs relative to these fields is:

1. IPv4 and IPv6 match criteria for IP packets

2. EXP (for MPLS packets) or DSCP (for IP packets) Dot1p/DEI bit (network ports do not support QinQ encapsulation)

3. default action (default = fc be profile out)

The configuration specifies the QoS bits to match against the incoming traffic together with the FC and profile (in/out) to be used (it is analogous to the SAP profile-mode in that the profile of the traffic is determined from the incoming traffic, rather than the CIR configured on the queue). A **default-action** keyword configures a default FC and profile state.

The IPv4 and IPv6 criteria matching only applies to the outer IP header of *non-tunneled* traffic, except for traffic received on an RFC 6037 MVPN tunnel for which classification on the outer IP header only is supported, and is only supported on network interfaces.

For tunneled traffic (GRE or MPLS), the match is based on the outer encapsulation header unless the keyword **ler-use-dscp** is configured. In this case, traffic received on the router terminating the tunnel that is to be routed to the base router or a VPRN destination is classified based on the encapsulated packet DSCP value (assuming it is an IP packet) rather than its EXP bits.

The ability exists for an egress LER to signal an implicit-null label (numeric value 3). This informs the previous hop to send MPLS packets without an outer label and so is known as penultimate hop popping (PHP). This can result in MPLS traffic being received at the termination of an LSP without any MPLS labels. In general, this would only be the case for IP encapsulated traffic, in which case the egress LER would need to classify the incoming traffic using IP criteria.

The egress section also defines the classification rules based on both the DSCP and precedence values in a packet to re-assign the packet's FC and profile (inplus/in/out/exceed).

Remarking

The egress section of the network policy defines the remarking of the egress packets, there is no remarking possible at the network ingress. The egress remarking is configured per FC and can set the related dot1p/DEI (explicitly or dependent on in/out-of-profile), DSCP (dependent on in/out-of-profile) and EXP (dependent on in/out-of-profile; inplus-profile packets are marked with the same values as in-profile packets and exceed-profile packets are marked with the same value as out-of-profile packets).

The traffic exiting a network port is either tunneled (in GRE or MPLS) or IP routed.

For tunneled traffic exiting a network port, the remarking applies to the DSCP/EXP bits in any tunnel encapsulation headers (GRE/MPLS) pushed onto the packet by this system, together with the associated dot1p/DEI bits if the traffic has an outer VLAN tag. For MPLS tunnels, the EXP bits in the entire label stack are remarked.

> **Note:**
> Strictly speaking this is marking (as opposed to remarking) as the action is adding QoS information rather than changing it.
>
> A new outer encapsulation header is pushed onto traffic at each MPLS transit label switched router as part of the label swap operation.

For VPLS/Epipe services no additional remarking is possible. However, for IES/VPRN/base-routing traffic, the remarking capabilities at the network egress are different at the first network egress (egress on the system on which the traffic entered by a SAP ingress) and subsequent network egress in the network (egress on the systems on which the traffic entered through another network interface).

At the first network egress, the DSCP of the routed/tunneled IP packet can be remarked, but this is dependent on two configuration settings:

- The trusted state of the ingress (service/network) interface and

- The **remarking** keyword in the network QoS policy at the network egress. The configuration combinations are summarized in Table 11: Network QoS Policy DSCP Remarking .

This is in addition to the remarking of any encapsulation headers and, as stated earlier, is not performed if the traffic was remarked at the service ingress.

For traffic exiting a subsequent network egress in the network, only the IP routed traffic can be remarked, again this is dependent on the ingress trusted state and egress remarking parameter.

There is one addition to the above to handle the marking for IP-VPN Option-B in order to remark the EXP, DSCP and dot1p/DEI bits at a network egress, this being **remarking force**. Without this, only the EXP and

dot1p/DEI bits are remarked. This does not apply to label switched path traffic switched at a label switched router.

Table 11: Network QoS Policy DSCP Remarking

Ingress	Trusted State	Remarking Configuration	Marking Performed
IES	Untrusted (default)	remarking	Yes
		no remarking (default)	Yes
	Trusted	remarking	Yes
		no remarking (default)	No
Network	Untrusted	remarking	Yes
		no remarking (default)	Yes
	Trusted (default)	remarking	Yes
		no remarking (default)	No
VPRN	Untrusted	remarking	Yes
		no remarking (default)	Yes
	Trusted (default)	remarking	No
		no remarking (default)	No

The following configuration shows a ingress network classification for DSCP EF explicitly, with a default action for the remainder of the traffic and use the DSCP from the encapsulated IP packet if terminating a tunnel. Remark the DSCP values for FC *af* and *ef* and remark all traffic (except incoming VPRN traffic) at the egress. Apply this policy to a network interface.

```
configure
    qos
        network 20 create
            ingress
                default-action fc af profile out
                ler-use-dscp
                dscp ef fc ef profile in
            exit
            egress
                remarking
                fc af
                    no dscp-in-profile
                    dscp-out-profile af13
                    lsp-exp-in-profile 6
                    lsp-exp-out-profile 5
                exit
                fc ef
                    dscp-in-profile af41
                exit
            exit
        exit
```

```
configure
    router
        interface "int-PE-1-PE-4"
            address 192.168.14.1/30
            port 1/2/1
            qos 20
        exit
```

The following configuration shows the trusted IES interface.

```
configure
    service
        ies 1 customer 1 create
            interface "int-access" create
                address 192.168.1.1/30
                tos-marking-state trusted
                sap 1/1/1:1 create
                exit
            exit
            no shutdown
        exit
```

The network QoS ingress and egress sections also contain the configuration for the use of FP-based policers and port-based queues by queue-groups which are out of scope of this chapter.

Network Queue Policy

The network queue QoS policy defines the queues and their parameters together with the FC to queue mapping. The policies are named, with the default policy having the name *default*. Network queues policies are applied under **config>card>mda>network>ingress** for the network ingress queues though only one policy is supported per MDA, so when a new policy is applied under one MDA, it is automatically applied under the other MDA on the same FP. At egress, network queue policies are applied under Ethernet: **config>port>ethernet>network**, POS: **config>port>sonet-sdh>path>network**, TDM: **config>port>tdm>e3|ds3>network** for the egress.

The following configuration shows an ingress and egress network-queue policy.

```
configure
    card 1
        card-type iom3-xp
        mda 1
            mda-type m4-10gb-xp-xfp
            network
                ingress
                    queue-policy "network-queue-1"
                exit
            exit
            no shutdown
        exit

configure
    port 1/1/3
        ethernet
            encap-type dot1q
            network
                queue-policy "network-queue-1"
            exit
        exit
        no shutdown
```

```
        exit
```

Up to 16 queues can be configured in a network-queue policy, each with a queue-type of best-effort, expedite, or auto-expedite. A new network-queue policy contains two queues, queue 1 for unicast traffic and queue 9 for multipoint traffic and by default all FCs are mapped to these queues. There is no differentiation for broadcast, multicast and unknown traffic. If the policy is applied to the egress, then any multipoint queues are ignored. As there are 8 FCs, there would be up to 8 unicast queues and 8 multipoint queues, resulting in 16 ingress queues and 8 egress queues. Normally, the network queue configuration is symmetric (the same queues/FC-mapping at the ingress and egress).

The following configuration defines a network-queue policy with FC *af* and *ef* assigned to queues 2 and 3 for unicast traffic, and queue 9 for multipoint traffic.

```
configure qos
        network-queue "network-queue-1" create
            queue 1 create
                mbs 50
                high-prio-only 10
            exit
            queue 2 create
            exit
            queue 3 create
            exit
            queue 9 multipoint create
                mbs 50
                high-prio-only 10
            exit
            fc af create
                multicast-queue 9
                queue 2
            exit
            fc ef create
                multicast-queue 9
                queue 3
            exit
        exit
```

Summary of Network Policies

Figure 18: Visualization of Default Network Policies displays the default network policies with respect to classification, FC to queue mapping and remarking.

Figure 18: Visualization of Default Network Policies
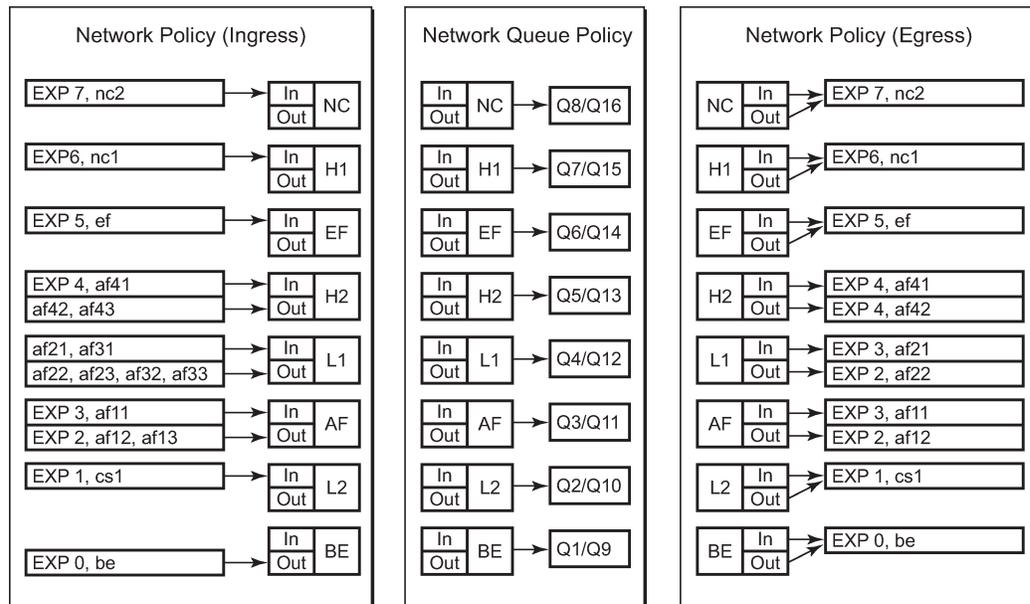
Queue Management

The policies described so far define queues but not the characteristics of those queues which determine how they behave. This section describes the detailed configuration associated with these queues. There are two aspects:

- Enqueuing packets onto a queue

  – buffer pools

  – queue sizing

  – Weighted Random Early Detection (WRED)

- Dequeuing packets from a queue

  – queue rates

  – scheduling

Enqueuing Packets: Buffer Pools

The packet buffer space is divided equally between ingress and egress. For line cards using a 50G FP2 for both ingress and egress traffic, the proportion of ingress versus egress buffer space can be modified using the following command:

```
configure
    card <slot-number>
        ingress-buffer-allocation <percentage>
```

The ingress buffer allocation percentage can be configured from 20% up to 80%.

Beyond that, by default there is one pool for network ingress per FP2/IOM, with one pool per access ingress port and one pool per access/network egress port. This is shown in Figure 19: Default Buffer Pools. This segregation provides isolation against buffer starvation between the separate pools. An additional ingress pool exists for managed multicast traffic (the multicast path management pool) but this is beyond the scope of this chapter.

The buffer management can be modified using named buffer pools and/or WRED-per-queue pools which are out of scope of this chapter.

Figure 19: Default Buffer Pools
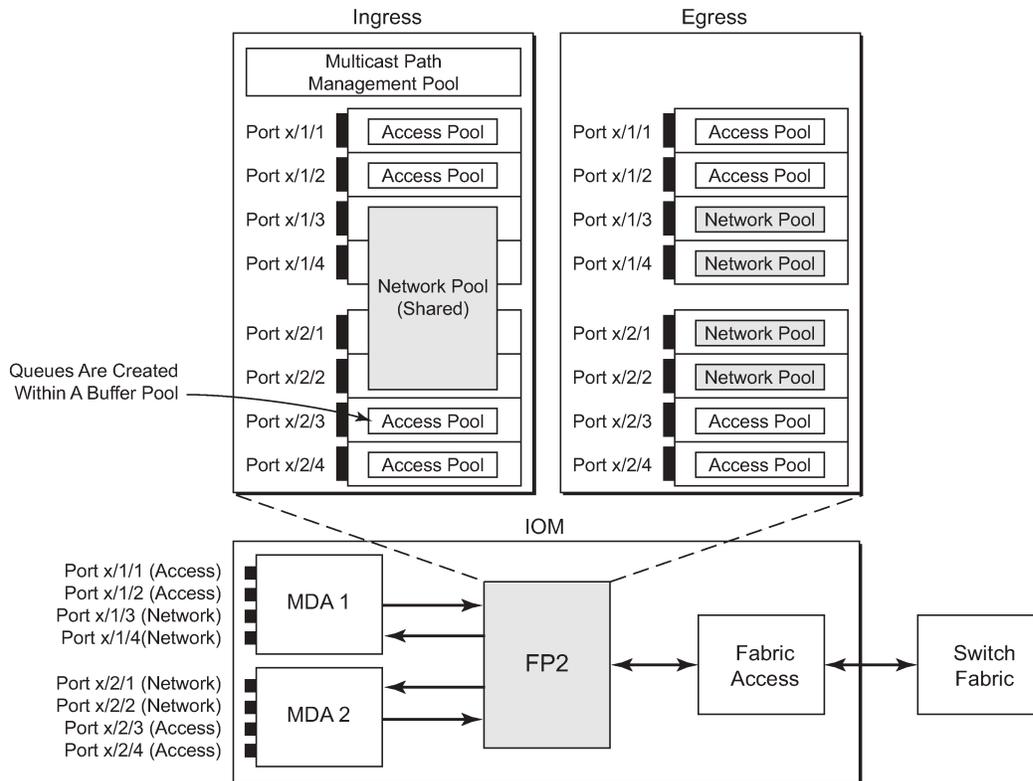
The size of the pools is based on the MDA type and the speed/type (access or network) of each port. Buffer space is allocated in proportion to the active bandwidth of each port, which is dependent on:

- The actual speed of the port
- Bandwidth for configured channels only (on channelized cards)
- Zero for ports without queues configured

This calculation can be tuned separately for ingress and egress, without modifying the actual port speed, using the port/modify-buffer-allocation-rate. Changing the port's ingress-rate or egress-rate will also modify its buffer sizes.

The following configuration changes the relative size for the ingress/egress buffer space on port 1/1/10 to 50% of the default.

```
configure
    port 1/1/3
```

```
        modify-buffer-allocation-rate
            ing-percentage-of-rate 50
            egr-percentage-of-rate 50
        exit
```

Each of the buffer pools created is further divided into a section of reserved buffers and another of shared buffers, see Figure 21: Ingress Buffer Pools and Queue Sizing. The amount of reserved buffers is calculated differently for network and access pools. For network pools, the default is approximately the sum of the CBS (committed burst size) values defined for all of the queues within the pool. The reserved buffer size can also be statically configured to a percentage of the full pool size (ingress: **config>card>mda>network>ingress>pool**; egress: **config>port>network>egress>pool**). For access pools, the default reserved buffer size is 30% of the full pool size and can be set statically to an explicit value (ingress: **config>port>access>ingress>pool**; egress: **config>port>access>egress>pool**).

The following configuration sets the reserved buffer size to 50% of the egress pool space on a network port.

```
configure
    port 1/1/3
        network
            egress
                pool
                        resv-cbs 50
                exit
            exit
        exit
```

On an access port, the reserved buffer size is set to 50% of the egress pool space, as follows:

```
configure
    port 1/1/1
        access
            egress
                pool
                        resv-cbs 50
                exit
            exit
        exit
```

Both the total buffer and the reserved buffer sizes are allocated in blocks (discrete values of Kbytes). The pool sizes can be seen using the **show pools** command.

It is possible to configure alarms to be triggered when the usage of the reserved buffers in the buffer pools reaches a certain percentage. Two alarm percentages are configurable, amber and red, **amber-alarm-threshold** *<percentage>* and **red-alarm-threshold** *<percentage>*. The percentage range is 1 — 1000.

- The percentage for the red must be at least as large as that for the amber.

- The alarms are cleared when the reserved CBS drops below the related threshold.

- When the amber alarm is enabled, dynamic reserved buffer sizing can be used; after the amber alarm is triggered the reserved buffer size is increased or decreased depending on the CBS usage. This requires a non-default resv-cbs to be configured together with a step and max value for the amber-alarm-action parameters. As the reserved CBS usage increases above the amber alarm percentage, the reserved buffer size is increased in increments defined by the step, up to a maximum of the max. If the CBS usage decreases, the reserved buffer size is reduced in steps down to its configured size.

- As the reserved buffer size changes, alarms will continue to be triggered at the same color (amber or red) indicating the new reserved buffer size. The pool sizing is checked at intervals, so it can take up to one minute for the alarms and pool re-sizing to occur.

The following displays a configuration for access ingress and egress pools.

```
configure
    port 1/1/1
        access
            ingress
                pool
                    amber-alarm-threshold 25
                    red-alarm-threshold 50
                    resv-cbs 20 amber-alarm-action step 5 max 50
                exit
            exit
            egress
                pool
                    amber-alarm-threshold 25
                    red-alarm-threshold 25
                    resv-cbs 20 amber-alarm-action step 5 max 50
                exit
            exit
        exit
```

The following is an example alarm that is triggered when the amber percentage has been exceeded and the reserved buffer size has increased from 20% to 25%:

```
82 2016/04/25 14:21:52.42 UTC MINOR: PORT #2050 Base Resv CBS Alarm
"Amber Alarm: CBS over Amber threshold: ObjType=port Owner=1/1/1 Type=accessIngress
 Pool=default NamedPoolPolicy= ResvSize=672 SumOfQ ResvSize=138 Old ResvCBS=20
New ResvCBS=25 Old ResvSize=528"
```

When a port is configured to be a hybrid port, its buffer space is divided into an access portion and a network portion. The split by default is 50:50 but it can be configured on a per port basis.

```
configure port 1/1/1
        ethernet
            mode hybrid
            encap-type dot1q
        exit
        hybrid-buffer-allocation
            ing-weight access 70 network 30
            egr-weight access 70 network 30
        exit
```

Enqueuing Packets: Queue Sizing

Queue sizes change dynamically when packets are added to a queue faster than they are removed, without any traffic the queue depth is zero. When packets arrive for a queue there will be request for buffer memory which will result in buffers being allocated dynamically from the buffer pool that the queue belongs to.

A queue has four buffer size related attributes: MBS, CBS, high-prio-only, and hi-low-prio-only, which affect packets only during the enqueuing process.

- Maximum Burst Size (MBS) defines the maximum buffer size that a queue can use. If the actual queue depth is larger than the MBS, any incoming packet will not be able to get a buffer and the packet will be

dropped. This is defined in bytes or Kbytes for access queues with a configurable non-zero minimum of 1byte or a default (without configuring the MBS) of the maximum between 10ms of the PIR or 64Kbytes. A value of zero will cause all packets to be dropped. MBS is a fractional percentage (xx.xx%) of pool size for network queues with defaults varying dependent on the queue (see default network-queue policy for default values). The MBS setting is the main factor determining the packet latency through a system when packets experience congestion. Queues within an egress queue group can have their MBS configured with as target packet queue delay in milliseconds.

- Committed Burst Size (CBS) defines the maximum guaranteed buffer size for an incoming packet. This buffer space is effectively reserved for this queue as long as the CBS is not oversubscribed (such the sum of the CBS for all queues using this pool does not exceed its reserved buffer pool size). For access queues, the CBS is defined in Kbytes with a configurable non-zero minimum of 6Kbytes or a default (without configuring the CBS) of the maximum between 10ms of the CIR or 6Kbytes. It is a fractional percentage (xx.xx%) of pool size for network queues with defaults varying dependent on the queue (see default network-queue policy for default values). Regardless of what is configured, the CBS attained normally will not be larger than the MBS. One case where CBS could be configured larger than MBS is for queues on LAGs, because in some cases the CBS is shared among the LAG ports (LAG QoS is not covered in this chapter). If the MBS and CBS values are configured to be equal (or nearly equal) this will result in the CBS being slightly higher than the value configured.

- High-prio-only. As a queue can accept both high and low enqueuing priority packets, a high enqueuing priority packet should have a higher probability to get a buffer. High-prio-only is a way to achieve this. Within the MBS, high-prio-only defines that a certain amount of buffer space will be exclusively available for high enqueuing priority packets. At network ingress and all egress buffering, highpriority corresponds to in-profile and low priority to out-of-profile. At service ingress, enqueuing priority is part of the classification. The high-prio-only is defined as a percentage of the MBS, with the default being 10%. A queue being used only for low priority/out-of-profile packets would normally have this set to zero. The high-prio-only could be considered to be an MBS for low enqueuing/out-of-profile packets.

- Hi-low-prio-only. There is an additional threshold, hi-low-prio-only, at egress which is equivalent to an MBS for exceed-profile packet. When the queue depth is beyond the hi-low-prio-only depth, the exceed-profile packets are dropped. The hi-low-prio-only is defined as a percentage of the MBS, with the default being 20%.

As with the buffer pools, the MBS, CBS, high-prio-only, and hi-low-prio-only values attained are based on a number of discrete values (not always an increment of 3Kbytes). The values for these parameters can be seen using the **show pools** command.

The MBS changes dynamically for queues in an egress queue group when H-QoS is used and the queue command **dynamic-mbs** is configured. This results in the MBS being modified in the ratio of the operational PIR to the admin PIR which derives an operational MBS. The ratio also affects the high-prio-only and hi-low-prio-only drop tails, and the WRED slopes if a slope policy is applied ot the queue. The configured CBS is used as the minimum operational MBS and the maximum MBS is capped by the maximum admin MBS (1GB).

As packets are added to a queue they will use the available CBS space, in which case they are placed in the reserved portion in the buffer pool. Once the CBS is exhausted, packets use the shared buffer pool space up to the hi-low-prio-only threshold (for exceed-profile packets), the high-prio-only threshold (for out-of-profile packets), or the maximum MBS size (for inplus-profile and in-profile packets).

The following configuration shows a queue with a specific MBS, CBS and disable high-prio-only.

```
configure
    qos
        sap-ingress 10 create
            queue 1 create
```

```
                mbs 10000
                cbs 100
                high-prio-only 0
            exit
        exit
```

Queue depth monitoring aims to give more visibility to the operator of the queue depths when traffic is bursty. It is a polling mechanism that is by default disabled. Queue depth monitoring can be enabled as a queue override on a service SAP or on a queue group.

The following command enables queue depth monitoring on SAP 1/1/1:11 in VPLS 10:

```
configure service vpls 10 sap 1/1/1:11 ingress queue-override queue 1 create monitor-depth
```

The result of the queue depth monitoring is presented in the form of occupancy ranges of 10% on the queue depth for each configured queue with the percentage of polls seen in each occupancy range, as follows:

```
*A:PE-1# show service id 10 sap 1/1/1:11 queue-depth
===============================================================================
Queue Depth Information (Ingress SAP)
===============================================================================
No Matching Entries
===============================================================================


===============================================================================
Queue Depth Information (Egress SAP)
===============================================================================
-------------------------------------------------------------------------------
Name                         : 10->1/1/1:11->1
MBS                          : Def

-------------------------------------------------------------------------------
Queue Depths (percentage)
-------------------------------------------------------------------------------
0%-10% 11%-20% 21%-30% 31%-40% 41%-50% 51%-60% 61%-70% 71%-80% 81%-90% 91%-100%
-------------------------------------------------------------------------------
68.21  3.64    3.43    3.47    3.86    3.22    3.86    2.78    3.78    3.66
-------------------------------------------------------------------------------
Average Elapsed Time       : 0d 00:39:11
Wghtd Avg PollEgr Interval: 100 ms
-------------------------------------------------------------------------------
*A:PE-1#
```

Enqueuing Packets: Weighted Random Early Detection (WRED)

In order to gracefully manage the use of the shared portion of the buffer pool, WRED can be configured on that part of the pool, and therefore applies to all queues in the shared pool as it fills. WRED is a congestion avoidance mechanism designed for TCP traffic. This chapter will only focus on the configuration of WRED. WRED-per-queue is an option to have WRED apply on a per egress queue basis, but is not covered here.

WRED is configured by a slope-policy which contains two WRED slope definitions, a high-slope which applies WRED to high enqueuing priority/in-profile packets and a low-slope which applies WRED to low enqueuing priority/out-of-profile packets. Both have the standard WRED parameters: start average (start-avg), maximum average (max-avg) and maximum probability (max-prob), and can be enabled or disabled individually. WRED slope policies also contain definitions for two slopes which are only applicable to access and network egress; a highplus-slope which applies WRED to inplus-profile packets and an

exceed-slope which applies WRED to exceed-profile packets. The WRED slope characteristics are shown in Figure 20: WRED Slope Characteristics.

Figure 20: WRED Slope Characteristics
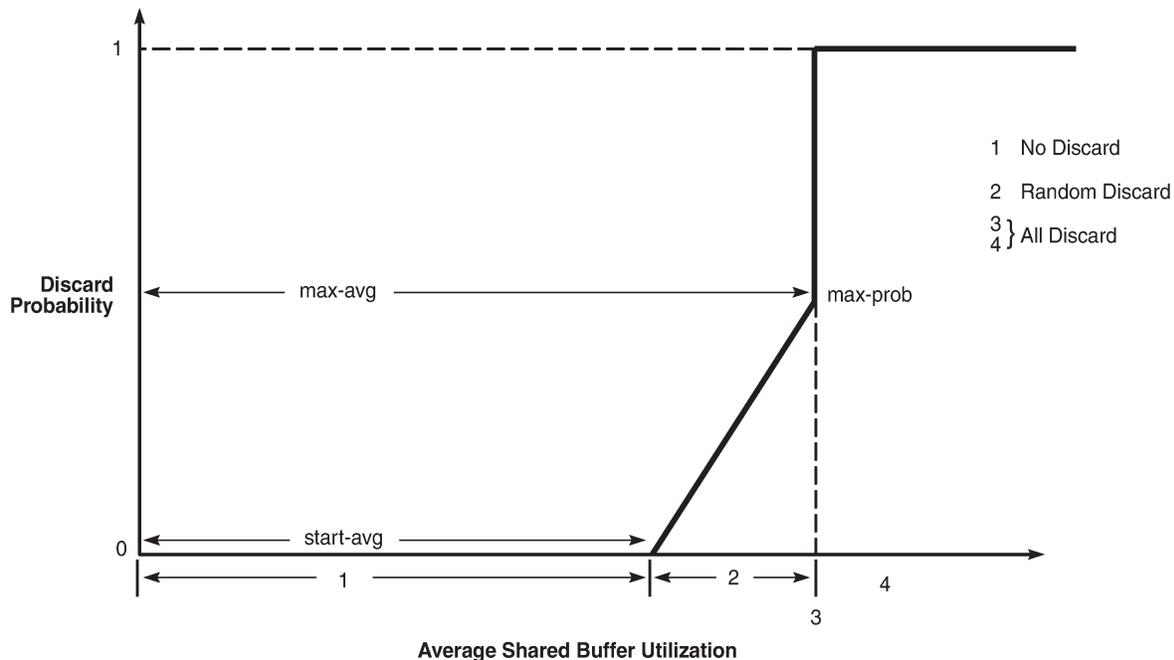
26132

A time-average-factor parameter can be configured per slope-policy which determines the sensitivity of the WRED algorithm to shared buffer utilization fluctuations (the smaller the value makes the average buffer utilization more reactive to changes in the instantaneous buffer utilization). The slope-policy is applied on a network port under **config>card>mda>network>ingress>pool** and **config>port>network>egress>pool** and on an access port under **config>port>access>ingress>pool** and **config>port>access>egress>pool**.

WRED is usually configured for assured and best-effort service traffic with premium traffic not typically being subject to WRED as it is always given preferential treatment and should never be dropped.

The following configuration defines a WRED slope policy and applies it to an ingress access port (the highplus and exceed slopes are ignored at ingress).

```
configure
    qos
        slope-policy "slope1" create
            exceed-slope
                shutdown
                start-avg 30
                max-avg 55
                max-prob 80
            exit
            high-slope
                start-avg 80
                max-avg 100
                max-prob 100
                no shutdown
            exit
```

```
            highplus-slope
                shutdown
                start-avg 85
                max-avg 100
                max-prob 80
            exit
            low-slope
                max-avg 100
                start-avg 80
                max-prob 100
                no shutdown
            exit
            time-average-factor 12
        exit
```
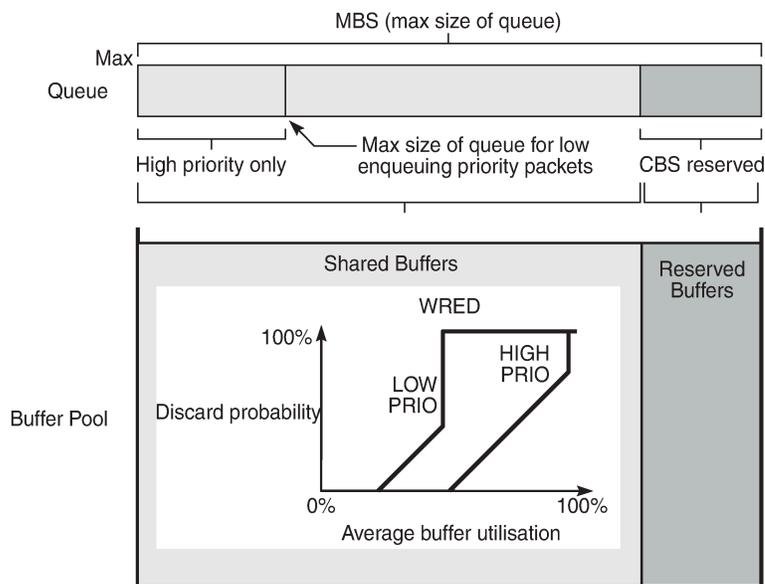
```
configure
    port 1/1/1
        access
            ingress
                pool
                    slope-policy "slope1"
                exit
            exit
        exit
```

The queue sizing parameters and buffer pools layout for ingress is shown in Figure 21: Ingress Buffer Pools and Queue Sizing.

Figure 21: Ingress Buffer Pools and Queue Sizing

25946

Figure 22: Egress Buffer Pools and Queue Sizing shows the queue sizing parameters and buffer pools layout for egress.

Figure 22: Egress Buffer Pools and Queue Sizing
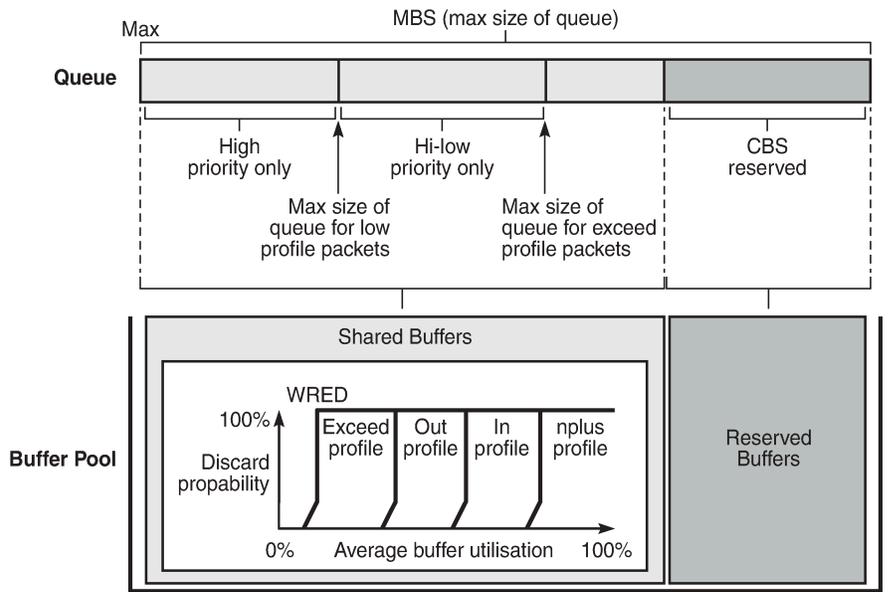
Dequeuing Packets: Queue Rates

A queue has two rate attributes: PIR and CIR. These affect packets only during the dequeue process.

- PIR — If the instantaneous dequeue rate of a queue reaches this rate, the queue is no longer served. Excess packets will be discarded eventually when the queue reaches its MBS/high-prio-only/hi-low-prio-only sizes. The PIR for access ports can be set in Kb/s with a default of **max** or as a percentage (see below). For network ports, the PIR is set as a percentage of the sum of the capacities of network and hybrid ports on that FP (taking into account any ingress-rate configuration) for ingress queues and of the port speed for egress queues, both with a default of 100%.

- CIR — The CIR is used to determine whether an ingress packet is in-profile or out-of-profile at the SAP ingress. It is also used by the scheduler in that queues operating within their CIRs will be served ahead of queues operating above their CIRs. The CIR for access ports can be set in Kb/s with a default of zero or as a percentage (see below). For network ports, it is set as a percentage of the sum of the capacities of network and hybrid ports on that FP (taking into account any ingress-rate configuration) for ingress queues and of the port speed for egress queues, with defaults varying dependent on the queue.

A percentage rate can be used in the sap-ingress and sap-egress policies, and can be defined relative to the local-limit (the parent scheduler rate) or the port-limit (the rate of the port on which the SAP is configured, including any egress-rate configured). The parameters rate and percent-rate are mutually exclusive and will overwrite each other when configured in the same policy. The following example shows a percent-rate configured as a port-limit.

```
configure
    qos
        sap-egress 10 create
            queue 1 create
                percent-rate 50.00 cir 10.00 port-limit
```

```
            exit
```

The PIR and CIR rates are shown in Figure 23: Scheduling (Dequeuing Packets from the Queue).

The queues operate at discrete rates supported by the hardware. If a configured rate does not match exactly one of the hardware rates an adaptation rule can be configured to control whether the rate is rounded up or down or set to the closest attainable value. The actual rate used can be seen under the operational PIR/CIR (O.PIR/O.CIR) in the **show pools** command output.

The following configuration shows a queue with a PIR, CIR and adaptation rule.

```
configure
    qos
        sap-ingress 20 create
            queue 2 profile-mode create
                adaptation-rule pir max cir min
                rate 10000 cir 5000
            exit
```

By default, the rates apply to packet bytes based on packet accounting, which for Ethernet includes the Layer 2 frame plus the FCS. An alternative is frame accounting which adds the Ethernet inter-frame gap, preamble and start frame delimiter.

Dequeuing Packets: Scheduling

Once a packet is placed on a queue, it is always dequeued from the queue by a scheduler. The scheduling order of the queues dynamically changes depending on whether a queue is currently operating below or above its CIR, with expedited queues being serviced before best-effort queues. This results in a default scheduling order of (in strict priority).

1. Expedited queues operating below CIR

2. Best-effort queues operating below CIR

3. Expedited queues operating above CIR

4. Best-effort queues operating above CIR

This is displayed in Figure 23: Scheduling (Dequeuing Packets from the Queue).

The scheduling operation can be modified using hierarchical QoS (with a scheduler-policy or port-scheduler-policy) which is out of scope of this chapter.
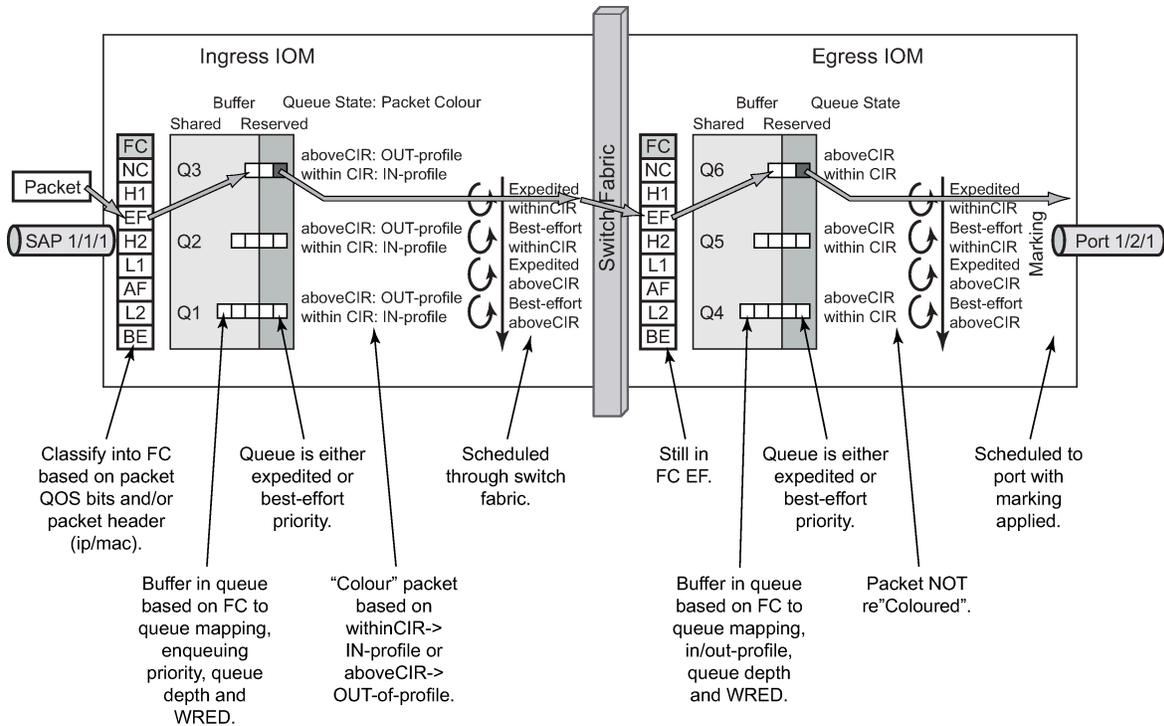
Figure 23: Scheduling (Dequeuing Packets from the Queue)
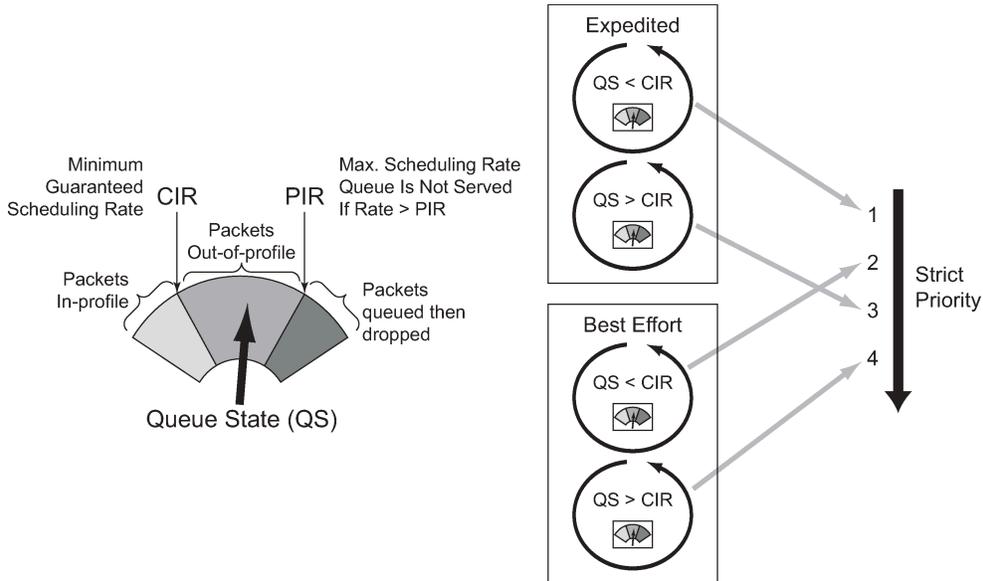
The overall QoS actions at both the ingress and egress IOMs are shown in Figure 24: IOM QoS Overview.

Figure 24: IOM QoS Overview

Show Output

The following displays **show** command output for:

- SAP queue statistics

- port queue statistics

- per-port aggregate egress-queue statistics monitoring

- access-ingress pools

The **show pools** command output for network-ingress and network/access-egress is similar to that of access-ingress and is not included here.

SAP Queue Statistics

The following output shows an example of the ingress and egress statistics on a SAP for an IES service (without multicast enabled, therefore no ingress multicast queue). There are two ingress queues, one being in priority mode and the other in profile mode. An explanation of the statistics is given for each entry.

Figure 25: Ingress and Egress SAP Queue Statistics for an IES Service
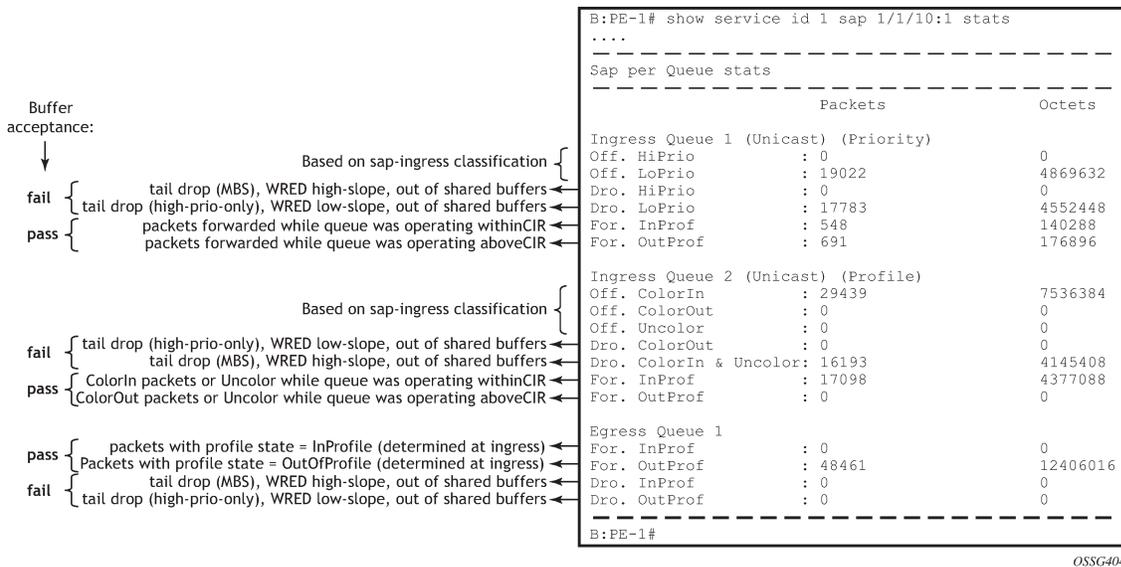
Port Queue Statistics

This output shows an example of the ingress and egress network port statistics. There are two unicast ingress queues (1 and 2) and one multicast ingress queue (9) with two egress queues. An explanation of the statistics is given for each entry.

Figure 26: Ingress and Egress Network Port Queue Statistics
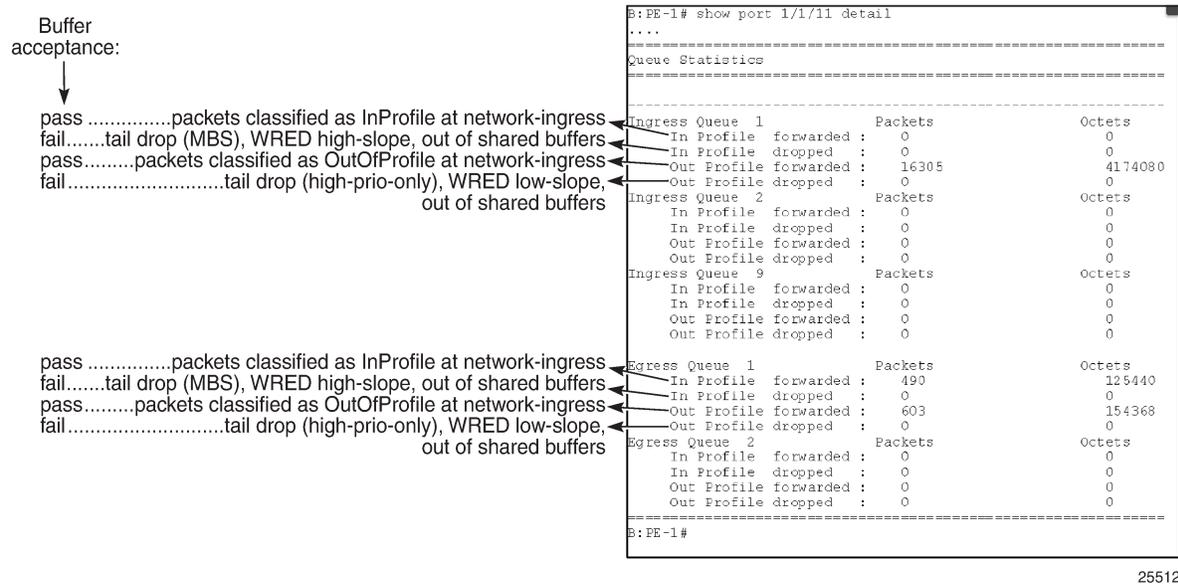
Per-Port Aggregate Egress-Queue Statistics Monitoring

Per-port aggregate egress-queue statistics show the number of forwarded and the number of dropped packets for in-profile and out-of-profile packets. All egress queues on the port are monitored: SAP egress, network egress, subscriber egress, egress queue group queues, system queues.

Per-port aggregate egress-queue statistics monitoring is enabled with the following command:

```
configure port 1/1/1 monitor-agg-egress-queue-stats
```

The collected statistics can be displayed as follows:

```
*A:PE-1# show port 1/1/1 statistics egress-aggregate

===============================================================================
Port 1/1/1 Egress Aggregate Statistics on Slot 1
===============================================================================
                    Forwarded               Dropped                  Total
-------------------------------------------------------------------------------
PacketsIn                   0                     0                      0
PacketsOut            5251690                     0                5251690
OctetsIn                    0                     0                      0
OctetsOut          357114942                     0              357114942
===============================================================================
*A:PE-1#
```

Access-Ingress Pools

This output shows an example of the default pools output for access-ingress. It includes the pools sizes, WRED information and queue parameters for each queue in the pool.

For this particular output, queue 2 on SAP 5/1/1:1 is being over-loaded which is causing its queue depth to be 67087296 bytes, made up of 64509 Kbytes from the shared pool (in use) and 1008 Kbytes from the reserved pool (in use). The output shows the pool total in usage as 65517 Kbytes, which is the sum of the shared and reserved pool in use. Sometimes the sum and total could be different by the size of one buffer, however, this is due to the dynamics of the buffer allocation which uses a 'sliding-window' mechanism and may therefore not always be perfectly aligned.

It can be seen that the high, low, and exceed WRED slopes are enabled and their instantaneous drop probability is shown 100% and their max averages are 64512 Kbytes, 46080 Kbytes, and 27648 Kbytes, respectively – this shows that the reserved portion of the buffer pool on this port is exhausted causing WRED to drop the packets for this queue.

The admin and operational PIR on the overloaded queues is 10Mb/s with CIR values of zero.

```
*A:PE-1# show pools 5/1/1 access-ingress


===============================================================================
Pool Information
===============================================================================
Port               : 5/1/1
Application        : Acc-Ing       Pool Name            : default
CLI Config. Resv CBS : 30%(default)
Resv CBS Step      : 0%            Resv CBS Max         : 0%
Amber Alarm Threshold: 0%          Red Alarm Threshold  : 0%
-------------------------------------------------------------------------------
Utilization                  State      Start-Avg    Max-Avg    Max-Prob
-------------------------------------------------------------------------------
High-Slope                   Up            70%         70%        100%
Low-Slope                    Up            50%         50%         80%
Exceed-Slope                 Up            30%         30%         80%
Time Avg Factor    : 12
Pool Total         : 132096 KB
Pool Shared        : 92160 KB      Pool Resv            : 39936 KB

High Slope Start Avg : 64500 KB    High Slope Max Avg   : 64512 KB
Low Slope Start Avg  : 46068 KB    Low Slope Max Avg    : 46080 KB
Excd Slope Start Avg : 27636 KB    Excd Slope Max Avg   : 27648 KB


       -----------------------------------------------------------------


       -----------------------------------------------------------------
Current Resv CBS  Provisioned   Rising       Falling       Alarm
%age              all Queues    Alarm Thd    Alarm Thd     Color
-------------------------------------------------------------------------------
30%               1020 KB       NA           NA            Green
Pool Total In Use    : 65517 KB
Pool Shared In Use   : 64509 KB      Pool Resv In Use     : 1008 KB
WA Shared In Use     : 64509 KB

Hi-Slope Drop Prob   : 100           Lo-Slope Drop Prob   : 100
Excd-Slope Drop Prob : 100


===============================================================================
Queue : 1->5/1/1:1->1
===============================================================================
FC Map             : be l2 af l1 h2 h1 nc
Tap                : 5/1
Admin PIR          : 10000000      Oper PIR        : Max
Admin CIR          : 0             Oper CIR        : 0
Admin MBS          : 12288 KB      Oper MBS        : 12288 KB
Hi Prio Only       : 1344 KB       Hi Low Prio Only : 2496 KB
CBS                : 12 KB         Depth           : 0
```

```
Slope            : not-applicable
===============================================================================

===============================================================================
Queue : 1->5/1/1:1->2
===============================================================================
FC Map           : ef
Tap              : 5/1
Admin PIR        : 10000              Oper PIR         : 10000
Admin CIR        : 0                  Oper CIR         : 0
Admin MBS        : 132096 KB          Oper MBS         : 132096 KB
Hi Prio Only     : 0 KB               Hi Low Prio Only : 27648 KB
CBS              : 1008 KB            Depth            : 67087296 B
Slope            : not-applicable
===============================================================================

===============================================================================
Queue : 28->5/1/1:28->1
===============================================================================
FC Map           : be l2 af l1 h2 ef h1 nc
Tap              : 5/1
Admin PIR        : 10000000           Oper PIR         : Max
Admin CIR        : 0                  Oper CIR         : 0
Admin MBS        : 12288 KB           Oper MBS         : 12288 KB
Hi Prio Only     : 1344 KB            Hi Low Prio Only : 2496 KB
CBS              : 0 KB               Depth            : 0
Slope            : not-applicable
===============================================================================

===============================================================================
Queue : 28->5/1/1:28->11
===============================================================================
FC Map           : be l2 af l1 h2 ef h1 nc
Tap              : MCast
Admin PIR        : 10000000           Oper PIR         : Max
Admin CIR        : 0                  Oper CIR         : 0
Admin MBS        : 12288 KB           Oper MBS         : 12288 KB
Hi Prio Only     : 1344 KB            Hi Low Prio Only : 2496 KB
CBS              : 0 KB               Depth            : 0
Slope            : not-applicable
===============================================================================
===============================================================================

*A:PE-1#
```

Conclusion

This chapter has described the basic QoS functionality available on the Nokia SR-series platforms, specifically focused on the FP2 chipset. This comprises of the use of queues to shape traffic at the ingress and egress of the system and the classification, buffering, scheduling and remarking of traffic on access, network, and hybrid ports.

Customer document and product support

Customer documentation
Customer documentation welcome page

Technical support
Product support portal

Documentation feedback
Customer documentation feedback