APPENDIX A

Platform-Specific Capabilities and Constraints

Platforms within the NG-WAN/MAN architecture perform QoS either in software or hardware. Furthermore, hardware QoS is hardware-specific and thus may vary significantly platform-to-platform, even from line card-to-line card. QoS functionality may vary in subtle ways, as may QoS command syntax. The following sections address the main platform-specific concerns and design recommendations for the NG-WAN/MAN. The platforms discussed include:

- Cisco 7200
- Cisco 7304
- Cisco 7600
- Cisco 12000 Gigabit Switch Router (GSR)

Cisco 7200 QoS Design

As previously mentioned, the Cisco 7200 series router performs QoS within Cisco IOS software. However because QoS is performed in software, QoS policies require marginal CPU processing to implement. The degree of impact is a function of the complexity of the policy, the traffic profile, the speeds, and the Network Processing Engine hardware. The rule of thumb to keep in mind is to design and test QoS policies such that when enabled the CPU does not exceed 75 percent utilization during normal operation. The first section examines configuring Uniform Mode MPLS DiffServ Tunneling on these platforms and the following sections examine designs for an 8-class QoS model and an 11-class QoS for these platforms.

Cisco 7200—Uniform Mode MPLS DiffServ Tunneling

Configuring uniform mode MPLS DiffServ Tunneling on the Cisco 7200 requires two parts, but the first is by default. Specifically, the mapping of IP Precedence to MPLS EXP is performed on Cisco 7200 PE routers (for customer-to-provider traffic) by default.

However for PE-to-CE egress traffic (exiting the MPLS VPN), additional configuration is required on the PE to achieve mapping of MPLS EXP to IP Precedence. This is because the final label is popped (and discarded) when it is received from the MPLS VPN cloud and therefore cannot be used as a match criterion for policies applied to the egress interface of the final PE router (facing the destination CE). The solution is to copy the final MPLS EXP bit values to a temporary placeholder on the PE ingress from the MPLS core (before the label is discarded) and then use these temporary placeholder values for setting the IP Precedence bits on egress to the CE.

Cisco IOS provides two such temporary placeholders, the QoS group and the discard class. For uniform mode scenarios, it is recommended to copy the MPLS EXP values to QoS group values on the ingress from the MPLS VPN cloud. (The discard class is recommended for use in pipe mode scenarios only.) QoS group values can then be copied to IP Precedence values (on egress to the customer CE).

The following is a sample Cisco 7200 uniform mode configuration.

```
!
policy-map MPLSEXP-TO-QOSGROUP
 class class-default
  set qos-group mpls experimental topmost           ! Copies EXP to QoS Group
!
policy-map QOSGROUP-TO-IPP
 class class-default
  set precedence qos-group           ! Copies QoS Group to IPP
!
!
interface GigabitEthernet1/0
 description GE TO MPLS VPN CORE          ! Link to/from MPLS VPN Core
 ip address 20.2.34.4 255.255.255.0
 ip vrf forwarding RED
 ip address 10.1.45.4 255.255.255.0
 service-policy input MPLSEXP-TO-QOSGROUP           ! MPLS EXP to QoS Group
 tag-switching ip
!
…
!
interface FastEthernet2/0
 description GE TO RED CE         ! Link to/from CE
 ip vrf forwarding RED
 ip address 10.1.45.4 255.255.255.0
 service-policy output QOSGROUP-TO-IPP           ! QoS Group to IPP
```

Cisco 7200—8-Class QoS Model

In the 8-class model is provisioned for the following application types:

- Voice
- Interactive-Video (video-conferencing)
- Network control (a combination of IP Routing and Network Management traffic)
- Call-Signaling
- Critical Data
- Bulk Data
- Best Effort
- Scavenger

This model takes advantage of the implicit policer function with Cisco IOS LLQ that allows you to time-division multiplex the LLQ. Essentially this functionality allows you to configure “dual-LLQs” even though only a single LLQ is operational. For example, assume you have configured one LLQ for Voice, set to 100 kbps, and another LLQ for Interactive-Video, set to 400 kbps. The software actually provisions a single LLQ for 500 kbps and allows only up to 100 kbps of Voice traffic and 400 kbps worth of Interactive-Video traffic into this queue on a first-in, first-out (FIFO) basis. If more than 100 kbps of voice is offered to this LLQ, it is dropped, and if more than 400 kbps of Interactive-Video is offered to it, it is also dropped. In this manner, both Voice and Interactive-Video benefit from Strict Priority servicing and, at the same time, data applications are protected from starvation (via the implicit policer). Following the LLQ queuing best-practice design principle presented previously in this chapter, Cisco recommends that the sum of the LLQs be less than 33 percent of a given link.

Note This implicit policing functionality of Cisco IOS LLQ should further impress on the network administrator the need to accurately provision Call Admission Control (CAC) to be in sync with the LLQ-provisioned bandwidth.

Voice is marked as EF, which is set by default on Cisco IP phones. When identified, VoIP is admitted into its own LLQ which, in this example, is set to 18 percent. CAC correspondingly should be assigned to this link by dividing the allocated bandwidth by the voice codec (including Layer 2 overhead) to determine how many calls can be permitted simultaneously over this link.

Interactive-Video (also known as IP videoconferencing [IP/VC]) is recommended to be marked AF41 (which can be marked down to AF42 or AF43, in the case of single- or dual-rate policing at the campus access edge). Interactive-Video is also assigned an LLQ under this dual-LLQ design. Cisco recommends overprovisioning the Interactive-Video LLQ by 20 percent of the IP/VC rate. This takes into account IP/UDP/RTP headers as well as Layer 2 overhead.

Additionally, Cisco IOS Software automatically includes a 200-ms burst parameter (defined in bytes) as part of the priority command. This default burst parameter has tested sufficient for protecting a single 384-kbps IP/VC stream; on higher speed links, the default burst parameter has shown to be insufficient for protecting multiple IP/VC streams. However multiple-stream IP/VC quality tested well with the burst set to 30,000 bytes (for example, priority 920 30000). The main point is that the default LLQ burst parameter might require tuning as multiple IP/VC streams are added.

Optionally DSCP-based WRED can be enabled on the Interactive-Video class, but testing has shown negligible performance difference in doing so. This is because congestion avoidance algorithms such as WRED are more effective on TCP-based flows than UDP-based flows, such as Interactive-Video.

A Network Control class is included within this 8-class model to protect network control plane traffic, specifically IP Routing (marked as CS6) and Network Management traffic (marked as CS2). As previously mentioned, Interior Gateway Protocol packets (such as RIP, EIGRP, OSPF, and IS-IS) are protected through the PAK_priority mechanism within the router. However EGP protocols such as BGP do not get PAK_priority treatment and might need explicit bandwidth guarantees to ensure that peering sessions do not reset during periods of congestion. Additionally administrators might want to protect network management access to devices during periods of congestion.

Call-Signaling traffic is also marked on the IP phones as CS3 (although some older versions of Cisco CallManager may mark Call-Signaling to the legacy value of AF31); Call-Signaling requires a moderate but dedicated bandwidth guarantee.

In these designs, WRED is not enabled on classes such as Network Control (IP Routing/Network Management) and Call-Signaling, because WRED would take effect only if such classes were filling their queues nearly to their limits. Such conditions would indicate an under-provisioning problem that would be better addressed by increasing the minimum bandwidth allocation for these classes rather than by enabling WRED.

The Critical Data class requires Transactional/Interactive Data traffic to be marked to AF21 (or AF22 or AF23 in the case of single- or dual-rate policers deployed within the campus). A bandwidth guarantee is made for this class, and DSCP-based WRED is enabled on this class to achieve the RFC 2597 Assured Forwarding Per-Hop Behavior.

The Bulk Data class requires Bulk Data to be marked to AF11 (or AF12 or AF13 in the case of single- or dual-rate policing deployed within the campus). Because TCP continually increases its window sizes, which is especially noticeable in long sessions such as large file transfers, constraining Bulk Data to its own class alleviates other data classes from being dominated by such large file transfers. A moderate bandwidth guarantee is made for this class and DSCP-based WRED is enabled on this class to achieve the RFC 2597 Assured Forwarding Per-Hop Behavior.

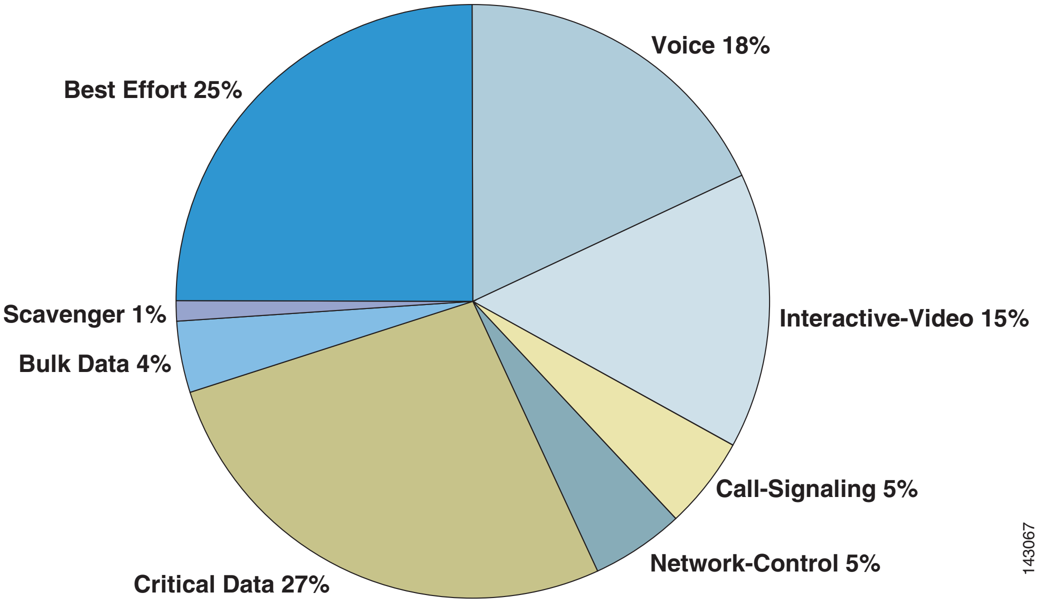
The Scavenger class constrains any traffic marked to DSCP CS1 to 1 percent of the link during periods of congestion; this allows class-default to use the remaining 25 percent. However to constrain Scavenger to 1 percent, an explicit bandwidth guarantee of 25 percent must be given to the Best Effort class. Otherwise if class-default is not explicitly assigned a minimum bandwidth guarantee, the Scavenger class can still rob it of bandwidth. This is because of the way the CBWFQ algorithm has been coded; if classes protected with a bandwidth statement are offered more traffic than their minimum bandwidth guarantee, the algorithm tries to protect such excess traffic at the direct expense of robbing bandwidth from class-default (if class-default is configured with fair-queue), unless class-default itself has a bandwidth statement that provides itself with a minimum bandwidth guarantee. However assigning a bandwidth statement to class-default on non-distributed platforms such as the Cisco 7200 currently precludes the enabling of fair queuing (fair-queue) on this class and forces FIFO queuing on class-default.

An additional implication of using a bandwidth statement on class-default is that even though 25 percent of the link is reserved explicitly for class-default, the parser does not attach the policy to a physical interface unless the **max-reserved-bandwidth 100** command is entered on the interface before the service-policy output statement. This is because the parser adds the sum of the bandwidth statements (regardless of whether one of these is applied to the class-default) and if the total is in excess of 75 percent of the link bandwidth, rejects the application of the policy to the interface.

Finally WRED can be enabled on the Best Effort class to provide congestion management on this default-class. Because all traffic assigned to the default class is to be marked to the same DSCP value of 0, it is superfluous to enable DSCP-based WRED on such a class; WRED (technically RED in this case because all the IP Precedence weights are the same) would suffice.

The Cisco 7200 8-Class QoS Model is shown in Figure A-1.

Figure A-1 Cisco 7200 8-Class QoS Model Example



The following configuration example shows the corresponding configuration for this Cisco 7200 8-Class Model. Keep in mind that this model is intended for links of speeds greater than 3 Mbps.

```
!
class-map match-all VOICE
 match ip dscp ef                 ! QoS Baseline marking for Voice
class-map match-all INTERACTIVE-VIDEO
 match ip dscp af41 af42 af43         ! QoS Baseline marking Interactive-Video
class-map match-any NETWORK-CONTROL
 match ip dscp cs6                ! QoS Baseline marking for IP Routing
 match ip dscp cs2                ! QoS Baseline marking for Network Mgmt
class-map match-all CALL-SIGNALING
 match ip dscp cs3                ! QoS Baseline marking for Call-Signaling
class-map match-all CRITICAL-DATA
 match ip dscp af21 af22 af23         ! QoS Baseline marking Transactional-Data
class-map match-all BULK-DATA
 match ip dscp af11 af12 af13         ! QoS Baseline marking for Bulk-Data
class-map match-all SCAVENGER
 match ip dscp cs1                ! QoS Baseline marking for Scavenger
!
policy-map WAN-EDGE
 class VOICE
  priority percent 18                 ! Voice gets LLQ - “dual-LLQ” design
 class INTERACTIVE-VIDEO
  priority percent 15                 ! Int-Video gets LLQ - “dual-LLQ” design
 class NETWORK-CONTROL
  bandwidth percent 5                 ! Routing and Network Mgmt gets CBWFQ
 class CALL-SIGNALING
  bandwidth percent 5                 ! Call-Signaling gets CBWFQ
 class CRITICAL-DATA
  bandwidth percent 27                ! Critical Data gets CBWFQ
  random-detect dscp-based            ! Critical Data also gets DSCP-based WRED
 class BULK-DATA
  bandwidth percent 4                 ! Bulk Data gets CBWFQ
  random-detect dscp-based            ! Bulk Data also gets DSCP-based WRED
 class SCAVENGER
  bandwidth percent 1                 ! Scavenger gets minimum CBWFQ
 class class-default
  bandwidth percent 25                ! Best Effort is protected with CBWFQ
  random-detect                   ! Best Effort also gets WRED (RED)
!
```

Cisco 7200—11-Class QoS Model

As mentioned previously, the 11-class QoS Baseline is a guiding model for addressing the QoS needs of today and the foreseeable future. The QoS Baseline is not a mandate dictating what enterprises must deploy today; instead this strategic document offers standards-based recommendations for marking and provisioning traffic classes that allow for greater interoperability and simplified future expansion.

Building on the previous 8-class model and as illustrated in Figure A-1, the Network Control class is subdivided into the IP Routing and Network Management classes.

Additionally the Critical Data class is subdivided into the Mission-Critical Data and Transactional Data classes.

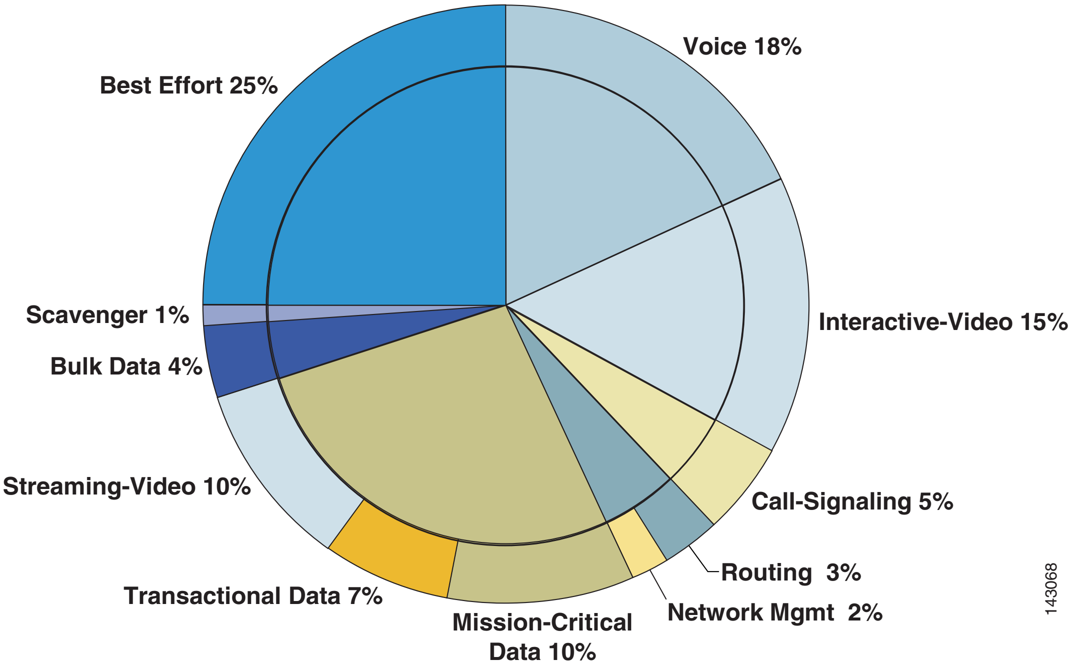
The Locally-Defined Mission-Critical Data class requires Mission-Critical Data traffic to be marked to AF31 (or AF32 or AF33 in the case of single- or dual-rate policers deployed within the campus). A bandwidth guarantee is made for this class and DSCP-based WRED is enabled on this class to achieve the RFC 2597 Assured Forwarding Per-Hop Behavior.

Note This recommendation assumes that Call-Signaling migration from AF31 to CS3 is complete within the enterprise; if not, then a temporary, non-standard DSCP value, such as 25, might be used for Mission-Critical Data marking. Also no AF-markdown PHB can be provisioned on this class during this interim.

Finally a new class is provisioned for Streaming Video. Testing has shown that there is a negligible difference in enabling WRED on this UDP-based traffic class, so although it remains an option, WRED is not enabled in these design examples.

The Cisco 7200 11-Class QoS Model is shown in Figure A-2. The inner circle shows how this model is backwards-compatible and consistent with the previous 8-Class model.

Figure A-2 Cisco 7200 11-Class QoS Model Example



The following configuration example shows the corresponding configuration for this Cisco 7200 11-Class Model. Keep in mind that this model is intended for links of speeds greater than 3 Mbps.

```
!
class-map match-all VOICE
 match ip dscp ef                  ! QoS Baseline marking for Voice
class-map match-all INTERACTIVE-VIDEO
 match ip dscp af41 af42 af43         ! QoS Baseline marking Interactive-Video
class-map match-all IP-ROUTING
 match ip dscp cs6                 ! QoS Baseline marking for IP Routing
class-map match-all NET-MGMT
 match ip dscp cs2                 ! QoS Baseline marking for Network Mgmt
class-map match-all CALL-SIGNALING
 match ip dscp cs3                 ! QoS Baseline marking for Call-Signaling
class-map match-all MISSION-CRITICAL-DATA
 match ip dscp af31 af32 af33         ! QoS Baseline marking Mission-Critical
class-map match-all TRANSACTIONAL-DATA
 match ip dscp af21 af22 af23         ! QoS Baseline marking Transactional-Data
class-map match-all BULK-DATA
 match ip dscp af11 af12 af13         ! QoS Baseline marking for Bulk-Data
```

```
class-map match-all STREAMING-VIDEO
 match ip dscp cs4                ! QoS Baseline marking Streaming-Video
class-map match-all SCAVENGER
 match ip dscp cs1                ! QoS Baseline marking for Scavenger
!
policy-map WAN-EDGE
 class VOICE
  priority percent 18                 ! Voice gets LLQ - “dual-LLQ” design
 class INTERACTIVE-VIDEO
  priority percent 15                 ! Int-Video gets LLQ - “dual-LLQ” design
 class IP-ROUTING
  bandwidth percent 3                 ! Routing gets CBWFQ
 class NET-MGMT
  bandwidth percent 2                 ! Network Management gets CBWFQ
 class CALL-SIGNALING
  bandwidth percent 5                 ! Call-Signaling gets CBWFQ
 class MISSION-CRITICAL-DATA
  bandwidth percent 10                ! Critical Data gets CBWFQ
  random-detect dscp-based            ! Critical Data also gets DSCP-based WRED
 class TRANSACTIONAL-DATA
  bandwidth percent 7                 ! Critical Data gets CBWFQ
  random-detect dscp-based            ! Critical Data also gets DSCP-based WRED
 class BULK-DATA
  bandwidth percent 4                 ! Bulk Data gets CBWFQ
  random-detect dscp-based            ! Bulk Data also gets DSCP-based WRED
 class STREAMING-VIDEO
  bandwidth percent 10                ! Streaming-video gets CBWFQ
 class SCAVENGER
  bandwidth percent 1                 ! Scavenger gets minimum CBWFQ
 class class-default
  bandwidth percent 25                ! Best Effort is protected with CBWFQ
  random-detect                   ! Best Effort also gets WRED (RED)
!
```

Cisco 7304 QoS Design

Cisco 7304 NSE-100 implements high-performance IP forwarding with services in the PXF processor on the NSE-100 forwarding engine. The following features are supported in the PXF path:

- Classification
- Marking
- Policing
- WRED
- LLQ
- CBWFQ
- Traffic Shaping

Currently, unlike regular IOS, NSE-100 PXF executes all MQC actions in a fixed order, regardless of the configured sequence. The general order of execution of QoS features is:

1. Input classification.
2. Only necessary if an input service-policy is present.
3. Input marking (**set**).
4. Input policing (**police**).

5. Output classification—The output classification is optimized in a way that it occurs if the packet was not already classified during input **or** if the packet header was modified during input processing (through marking/policing or some other feature, like NAT).
6. Output marking (**set**).
7. Output policing (**police**).
8. WRED (**random-detect**).
9. CBWFQ/LLQ/Traffic Shaping (bandwidth, priority, and shape).

It supports up to eight (8) different output queues per physical or logical interface. From these queues, two have special purposes:

- Crucial traffic queue—Dedicated for some types of internally (RP) generated vital traffic (mostly Layer 2 keep-alives).
- Default queue—Dedicated to traffic that does not match any user defined classes.

The other six (6) queues are the "user-defined queues," available for classes requiring queue related actions in the output service-policy.

The crucial traffic queue is a NSE-100-specific solution for handling those special types of internally-generated vital traffic that cannot be dropped. This queue has its own fixed parameters that cannot be configured by the user.

When no output service-policy is attached to an interface, only the crucial traffic and default queues are used. Those types of special locally-generated vital traffic go through the crucial traffic queue (usually very light traffic) and all other traffic goes through the default queue.

Classification

Classification on the Cisco 7304 NSE-100 can be applied based on:

- Layer 3 Criteria (IP packets only):
  - Access Control Lists (Turbo ACLs)
  - IP Precedence
  - IP DSCP
  - IP RTP port number/range
- MPLS Related Criteria (MPLS packets only):
  - MPLS experimental bits
- Internal settable variables (all packets):
  - QoS-Groups

Note When combining qos-group with other matching criteria within the same class-map, only the qos-group statements are considered; the other match statements are completely ignored.

Policing

On the NSE-100, the policing implementation is a single rate, 3-color policer. The traffic policer accepts the following parameters:

- **sustained rate** (bps)
- **normal burst** (bytes)
- **max burst** (bytes)

The term color refers to the following actions:

- **conform-action**—Traffic is less than the specified sustained rate
- **exceed-action**—Traffic exceeds the normal burst
- **violate-action**—Traffic exceeds the maximum burst

The result of these actions can be set to:

- **transmit**—Transmit the packet
- **drop**—Drop the packet

Or to remark the packets:

- **set-prec-transmit**—Rewrite the packet precedence and send it.
- **set-dscp-transmit**—Set the DSCP bits and send packet.
- **set-clp-transmit**—Set the ATM CLP bit and send packet.
- **set-frde-transmit**—Set the FR DE bit and send packet.
- **set-mpls-exp-imposition-transmit**—Set the MPLS experimental bit at imposition and send the packet.
- **set-qos-transmit**—Set the QoS group within the router and send packet.

Weighted Random Eary Detection (WRED)

WRED is enabled for congestion avoidance in a class through the **random-detect** command. Either the **bandwidth** or **shape** command must be already present in the class for IOS to allow the configuration of **random-detect**. With WRED configured, the queue size for that particular class is set to be the highest WRED maximum threshold x 2, rounded up to the next power of 2. The WRED maximum threshold ranges from 1 to 4096 packets and hence the queue size can be up to 8192 packets. By default, the max-thresholds for all drop profiles are set to 40, which gives a default queue size (rounding up to nearest power of 2) of 128 packets. The implementation on Cisco 7304 NSE-100 supports:

- IP precedence based WRED
- IP DSCP based WRED
- MPLS exp based WRED

This can give packets with low IP precedence, DSCP, or MPLS exp value a higher probability of being dropped than packets with high value. Up to 64 drop profiles are supported.

Note Discard-class based WRED is not supported.

Class-based Weighted Fair Queuing (CBWFQ)

CBWFQ is a congestion management implementation of WFQ. It provides support for configurable queuing for different traffic classes. A FIFO queue is allocated for each class containing queuing actions and traffic belonging to a class is directed to its proper queue. The minimum guaranteed bandwidth can be assigned in three different forms:

- A committed information rate in Kbps (**bandwidth** <kbps>)
- A percentage of the underlying link rate (**bandwidth percent** <percent>)
- A percentage of the bandwidth not allocated by the other classes (**bandwidth remaining percent** <percent>)

On the NSE-100, CBWFQ is implemented through the scheduler. While each form of the bandwidth command provides a means to allocate bandwidth to a traffic class, it is also the case that if a class does not use its allocated bandwidth (i.e., no traffic is offered to the class), then this excess bandwidth is available to other classes, which are then allowed to use this bandwidth and exceed their minimum allocation. On the NSE-100, by default, classes share excess bandwidth proportionally to the allocated bandwidth.

On the NSE-100, even though the CLI granularity is 1 Kbps, the actual bandwidth granularity is a factor of the link speed. More precisely, it is 1/65536 (1/64K) of the link speed. The value configured at CLI is internally rounded down to the closest multiple of 1/65536th of the link speed.

Examples:

- FastEthernet Link speed—100 Mbps => Granularity 100 Mbps / 65536 = ~ 1.53 Kbps
- GigEthernet Link speed—1 Gbps => Granularity 1 Gbps / 65536 = ~ 15.3 Kbps

Note The crucial traffic queue bandwidth is allocated after the bandwidth for the class queues. If, after the class queues bandwidths are allocated, there is still enough bandwidth left for the crucial traffic queue, it is simply allocated. On the other hand, if after allocating bandwidths for the class queues, there is not enough bandwidth left for the crucial traffic queue, all class queues bandwidths are internally and proportionally adjusted to the link speed minus the crucial traffic queue bandwidth.

Hierarchical Policies

A traffic policy is called hierarchical policy when it is defined using two or more policy-maps nested through the policy-map class sub-mode service-policy command. The operation of a hierarchical policy is recursive. When a hierarchical policy is used, the traffic matching a 1st-level policy-map class that has a child policy-map is subject to both the actions in this 1st-level class and the actions on the matching class on the child policy-map.

NSE-100 supports four different generic “flavors” of hierarchical traffic policy configurations, namely:

- Hierarchical traffic shaping for sub-interfaces
- Ingress hierarchical policing
- Queuing on parent, selective marking/policing on child
- Shaping on parent, selective marking/policing on child

The following configuration provides examples for both “flat” (core facing) as well as “hierarchical” (CE facing) policies on a 7304 NSE-100 PE:

```
class-map match-any realtime
```

```
  match mpls experimental topmost 5
class-map match-any realtime-2ce
  match ip precedence 5
class-map match-any bulk-data
  match mpls experimental topmost 1
class-map match-any bulk-data-2ce
  match ip precedence 1
class-map match-any bus-critical
  match mpls experimental topmost 3
class-map match-any trans-data
  match mpls experimental topmost 2
class-map match-any bus-critical-2ce
  match ip precedence 3
class-map match-any trans-data-2ce
  match ip precedence 2
class-map match-any control
  match mpls experimental topmost 6
  match mpls experimental topmost 7
class-map match-any video-2ce
  match ip precedence 4
class-map match-any video
  match mpls experimental topmost 4
class-map match-any control-2ce
  match ip precedence 7
  match ip precedence 6
!
policy-map q-core-out
  class realtime
    priority
   police cir 300000000
     conform-action transmit
     exceed-action drop
  class control
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 6 300 1500 1
    random-detect precedence 7 300 1500 1
  class bus-critical
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 3 300 1500 1
  class trans-data
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 2 300 1500 1
  class video
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 4 300 1500 1
  class bulk-data
    bandwidth remaining percent 7
    random-detect
    random-detect precedence 1 300 1500 1
  class class-default
    bandwidth remaining percent 36
    random-detect
    random-detect precedence 0 300 1500 1
policy-map q-2ce-out-1
  class control-2ce
    bandwidth percent 14
    random-detect
    random-detect precedence 6 300 1500 1
    random-detect precedence 7 300 1500 1
  class bus-critical-2ce
```

```
    bandwidth percent 14
    random-detect
    random-detect precedence 3 300 1500 1
  class trans-data-2ce
    bandwidth percent 14
    random-detect
    random-detect precedence 2 300 1500 1
  class video-2ce
    bandwidth percent 14
    random-detect
    random-detect precedence 4 300 1500 1
  class bulk-data-2ce
    bandwidth percent 7
    random-detect
    random-detect precedence 1 300 1500 1
  class class-default
    bandwidth percent 36
    random-detect
    random-detect precedence 0 300 1500 1
policy-map q-2ce-out-data
  class class-default
    shape average 650000000
   service-policy q-2ce-out-1
policy-map q-2ce-out-2
  class realtime-2ce
    priority
   police cir 300000000
     conform-action transmit
     exceed-action drop
  class control-2ce
    random-detect
    random-detect precedence 6 300 1500 1
    bandwidth percent 3
  class class-default
    random-detect
    bandwidth percent 2
policy-map q-2ce-out-voice
  class class-default
    shape average 350000000
   service-policy q-2ce-out-2
!
interface GigabitEthernet3/0/1
 description Core Facing
 ip address 125.1.103.22 255.255.255.252
 mpls ip
 service-policy output q-core-out
!
interface GigabitEthernet3/1/0.1
 description BLUE-DATA - To CE
 encapsulation dot1Q 175
 ip vrf forwarding blue-data
 ip address 125.1.103.49 255.255.255.252
 service-policy output q-2ce-out-data
!
interface GigabitEthernet3/1/0.2
 description BLUE-VOICE - To CE
 encapsulation dot1Q 176
 ip vrf forwarding blue-voice
 ip address 125.1.103.53 255.255.255.252
 service-policy output q-2ce-out-voice
```

Cisco 7600 QoS Design

The Cisco 7600 series routers perform QoS in hardware. Classification, marking, and policing are performed within the Policy Feature Card (PFC3B or PFC3BXL, hereafter referred to as PFC3B)) and queuing and dropping (tail-drop or WRED) is performed within line card hardware. As such there is no incremental CPU load when enabling such features, even at GE and 10GE speeds.

The following sections examine the configuration of Uniform Mode MPLS DiffServ Tunneling on the Cisco 7600, which is performed within the PFC3B, the trust states of various MPLS label pushing, swapping, and popping operations, and finally line card-specific queuing designs.

Cisco 7600—Uniform Mode MPLS DiffServ Tunneling

Configuring Uniform Mode MPLS DiffServ Tunneling on the Cisco 7600 consists of two main parts: trusting IPP or DSCP (recommended) on the PE-to-CE interface (applied in the ingress direction from the CE) and applying an **mpls propagate-cos command** to overwrite the IPP value with the last MPLS EXP value (applied in the egress direction towards the CE).

The following is an example of the Cisco 7600 Uniform Mode configuration.

```
C7600(config)# interface GE-WAN 3/1
C7600(config-if)# mls qos trust dscp
C7600(config-if)# interface GE-WAN 3/2.32
C7600(config-if)# mpls propagate-cos
```

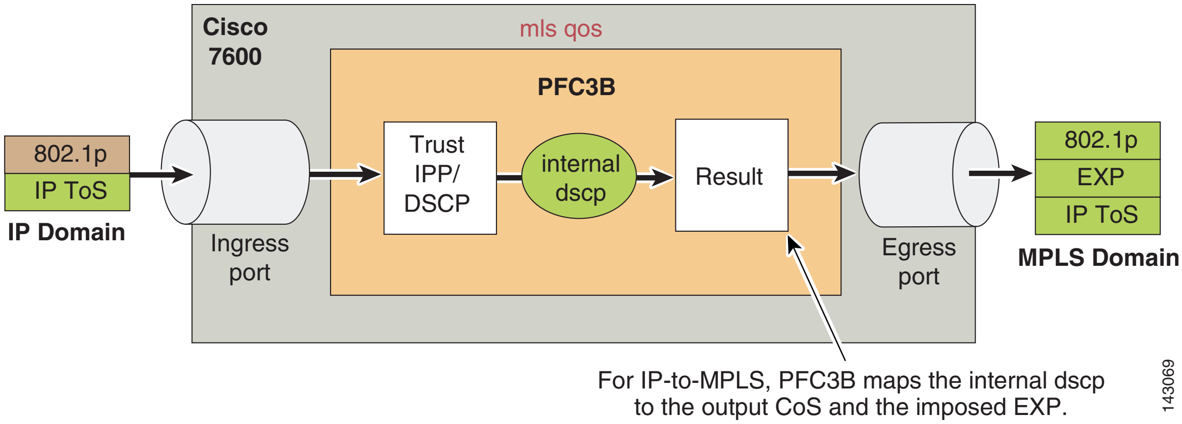
Cisco 7600—Trust States and Internal DSCP Generation

The Cisco 7600 performs queuing and dropping decisions based on the concept of an "internal DSCP." The Cisco 7600 generates an internal DSCP for all packets/frames, regardless of whether they are IP or otherwise. For example, this internal DSCP may be derived by setting a port state to **trust dscp**, in which case the internal DSCP is set to match the packet DSCP. Alternatively, in an 802.1Q/p environment, the port can be set to **trust cos**, in which case the internal DSCP is generated by accepting the CoS marking value and performing a conversion to DSCP by means of the CoS-to-DSCP mapping table.

In MPLS environments the following rules apply to trust and the generation of the internal DSCP (used for queuing and dropping):

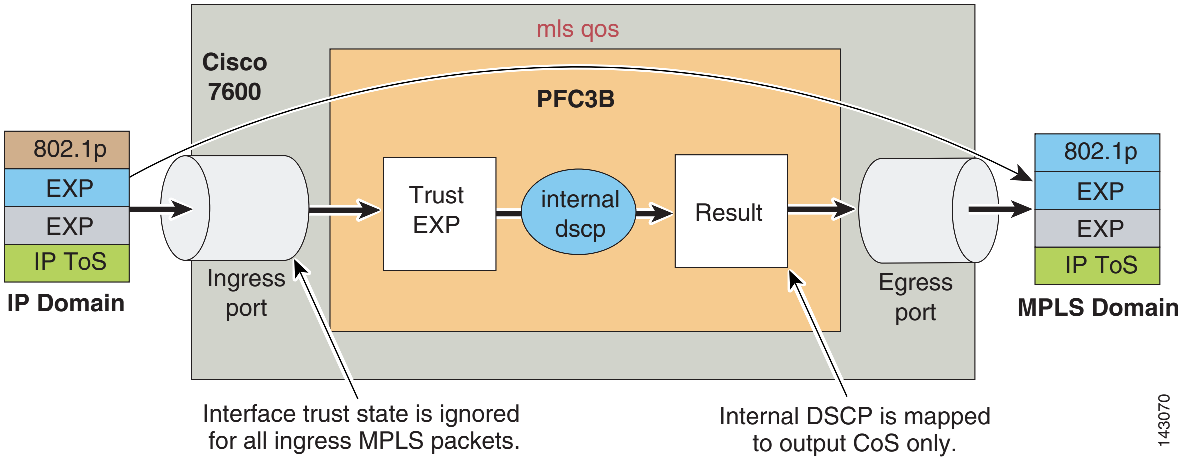
- When PFC3B receives an IP packet (IP-to-IP or IP-to-MPLS), it uses the input interface trust state and, if configured, the **policy-map trust** command. During MPLS label imposition, for packets received on an interface with trust IPP or trust dscp, PFC3B maps the IPP/DSCP to the internal DSCP. It then maps the internal DSCP to the imposed EXP and the output CoS. It always preserves the underlying IP ToS as shown in Figure A-3.

Figure A-3 Cisco 7600 MPLS Label Imposition (Pushing) Trust



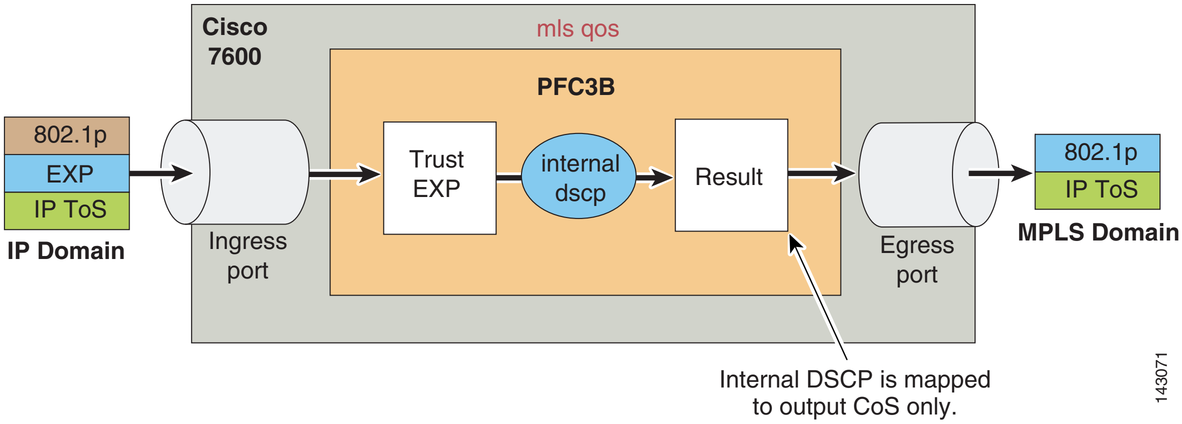
- When PFC3B receives an MPLS packet to be swapped (MPLS-to-MPLS), it trusts EXP; the interface trust state and the **policy-map trust** command have no effect. During swapping, PFC3B trusts the topmost EXP and maps it to the internal DSCP. During the swap, it copies EXP from the swapped-off label to the swapped-on label. After the swap, it maps the internal DSCP to the egress frame CoS as shown in Figure A-4.

Figure A-4 Cisco 7600 MPLS Label Switching (Swapping) Trust



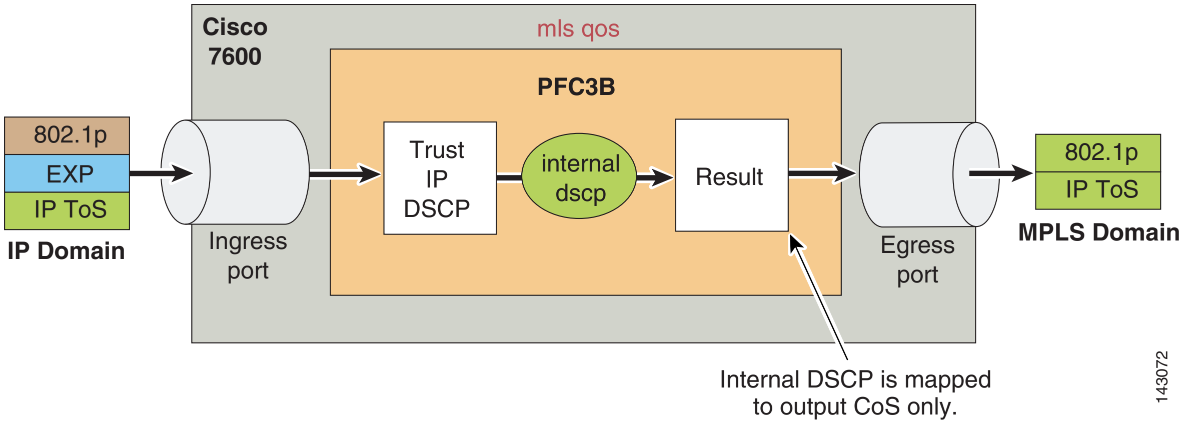
- When PFC3B receives an MPLS packet to be popped (MPLS-to-IP), trust depends on the type of label; however, in all cases, the interface trust state and the **policy-map trust** command have no effect.
- Non-aggregate label—PFC3B/PFC3BXL trusts EXP in the topmost label. PFC3B trusts EXP and maps it to the internal DSCP. By default, PFC3B discards the popped EXP and does not propagate it to the exposed IP ToS. After the pop, it maps the internal DSCP to the egress frame CoS as shown in Figure A-5.

Figure A-5 Cisco 7600 MPLS Label Disposition (Popping) Trust—Case 1 (Non-Aggregate Label)



- Aggregate label in VPN CAM—PFC3B/PFC3BXL trusts IP DSCP.
- Aggregate label not in VPN CAM—PFC3B/PFC3BXL trusts IP DSCP (after recirculation). PFC3B trusts the exposed IP ToS and maps it to the internal DSCP. By default, PFC3B discards the popped EXP and does not propagate it to the exposed IP ToS. After the pop, it maps the internal DSCP to the egress frame CoS as shown in Figure A-6.

Figure A-6 Cisco 7600 MPLS Label Disposition (Popping) Trust—Case 1 (Non-Aggregate Label)



For example, if an IP packet is received and the interface is set to trust dscp (recommended), then the internal DSCP is set to match the packet DSCP. When queued on egress, even though a CoS-to-queue mapping is used, this actually represents an internal DSCP-to-queue mapping, such that mapping CoS 2-to-queue 2 actually represents assigning any packets that have an internal DSCP value of 16 through 23 to queue 2. Continuing the example, if a MPLS packet is received and is being forwarded as an MPLS packet (swapping function of a P Cisco 7600 router), then the EXP value is trusted, the internal DSCP value is calculated from the EXP-to-DSCP mapping table, and the packet is queued based on the internal DSCP value (via the CoS-to-queue mapping table). Concluding the example, if a non-aggregate MPLS packet is received and is to be forwarded as an IP packet (popping function of a PE Cisco 7600 router), then the internal DSCP value is calculated by trusting the topmost label EXP value and applying the EXP-to-DSCP mapping, and finally the packet is queued based on the internal DSCP value (via the CoS-to-queue mapping table).

Cisco 7600—Queuing Design

Cisco 7600 line cards support both ingress and egress queuing; furthermore, these ingress and egress queuing structures vary by line card. Ingress congestion implies that the combined ingress rates of traffic exceed the switch fabric channel speed, and thus need to be queued simply to gain access to the switching fabric. On newer platforms, such as the Cisco 7600 Sup720, this means that a combined ingress rate of up to 40 Gbps per slot is required to create such an event. However to obviate such an extreme event, the Catalyst 7600 schedules ingress traffic through the receive queues based on CoS values. In the default configuration, the scheduler assigns all traffic with CoS 5 to the strict priority queue (if present); in the absence of a strict priority queue, the scheduler assigns all traffic to the standard queues. All other traffic is assigned to the standard queues (with higher CoS values being assigned preference over lower CoS values, wherever supported). Thus even in the highly unlikely event of ingress congestion, the default settings for the receive queues of the Cisco 7600 line cards are more than adequate to protect VoIP and network control traffic. Therefore the focus of this section is on Cisco 7600 egress/transmit queuing design recommendations.

The following five egress queuing structures are currently supported on Cisco 7600 Gigabit, Ten-Gigabit, or 10/100/1000 line cards:

- 1P2Q1T—Indicates one strict priority queue and two standard queues, each with one configurable WRED-drop threshold and one non-configurable (100 percent) tail-drop threshold.
- 1P2Q2T—Indicates one strict priority queue and two standard queues, each with two configurable WRED-drop thresholds.
- 1P3Q1T—Indicates one strict priority queue and three standard queues, each with one configurable WRED-drop or tail-drop threshold and one non-configurable (100 percent) tail-threshold.
- 1P3Q8T—Indicates one strict priority queue and three standard queues, each with eight configurable WRED-drop or tail-drop thresholds.
- 1P7Q8T—Indicates one strict priority queue and seven standard queues, each with eight configurable WRED-drop or tail-drop thresholds.

Table A-1 and Table A-2 summarize these queuing structures by line card.

T

Table A-1 Cisco 7600 Classic and CEF256 Line Cards and Queuing Structures

Classic/ CEF256 Ethernet Modules	Description	Rx Queuing	Tx Queuing	Buffer Size
WS-X6148-GE-TX	48-Port 10/100/1000 RJ-45 Module	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6148V-GE-TX	48-Port 10/100/1000 Inline Power RJ-45 Module	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6316-GE-TX	16-Port 1000TX GigabitEthernet RJ-45 Module	1P1Q4T	1P2Q2T	512KB per port
WS-X6408A-GBIC	8-Port GigabitEthernet Module (with enhanced QoS; requires GBICs)	1P1Q4T	1P2Q2T	512KB per port
WS-X6416-GBIC	16-Port GigabitEthernet Module (requires GBICs)	1P1Q4T	1P2Q2T	512KB per port
WS-X6416-GE-MT	16-Port GigabitEthernet MT-RJ Module	1P1Q4T	1P2Q2T	512KB per port
WS-X6501-10GEX4	10 GigabitEthernet Module	1P1Q8T	1P2Q1T	64MB per port
WS-X6502-10GE	10 GigabitEthernet Base Module (requires OIM)	1P1Q8T	1P2Q1T	64MB per port

Table A-1 Cisco 7600 Classic and CEF256 Line Cards and Queuing Structures (continued)

WS-X6516A-GBIC	GigabitEthernet Module (fabric-enabled; requires GBICs)	1P1Q4T	1P2Q2T	1MB per port
WS-X6516-GBIC	GigabitEthernet Module (fabric-enabled; requires GBICs)	1P1Q4T	1P2Q2T	512KB per port
WS-X6516-GE-TX	16-Port GigabitEthernet Copper Module; (crossbar-enabled)	1P1Q4T	1P2Q2T	512KB per port
WS-X6524-100FX-MM	24-Port 100FX MT-RJ Module (Fabric-Enabled)	1P1Q0T	1P3Q1T	1MB per port
WS-X6548-RJ-21	48-Port 10/100 RJ-21 Module (fabric- enabled)	1P1Q0T	1P3Q1T	1MB per port
WS-X6548-RJ-45	48-Port 10/100 RJ-45 Module (crossbar-enabled)	1P1Q0T	1P3Q1T	1MB per port
WS-X6548V-GE-TX	48-Port 10/100/1000 Inline Power RJ- 45 Module (fabric-enabled)	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6548-GE-TX	48-Port 10/100/1000 RJ-45 Module (fabric-enabled)	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6816-GBIC	16-Port GigabitEthernet Module (fabric-enabled; requires GBICs)	1P1Q4T	1P2Q2T	512KB per port

Table A-2 Cisco 7600 CEF720 Line Cards and Queuing Structures

C2 (xCEF720) Modules	Description	Rx-Queuing	Tx-Queuing	Buffer Size
WS-X6704-10GE	4-Port 10 GigabitEthernet Module	1Q8T (8Q8T with DFC3a)	1P7Q8T	16MB per port
WS-X6724-SFP	24-Port GigabitEthernet SFP Module	1Q8T (2Q8T with DFC3a)	1P3Q8T	1MB per port
WS-X6748-GE-TX	48-Port 10/100/1000 RJ-45 Module	1Q8T (2Q8T with DFC3a)	1P3Q8T	1MB per port
WS-X6748-SFP	48-Port GigabitEthernet SFP Module1	1Q8T (2Q8T with DFC3a)	1P3Q8T	1MB per port

Note For any newer line cards not on this list, the queuing structure can be ascertained by the **show queuing interface verification** command.

Cisco 7600 1P2Q1T 10GE Queuing Design

Under the 1P2Q1T queuing model, buffer space can be allocated as follows: 30 percent for Scavenger/Bulk plus Best Effort queue (Q1) and 40 percent for Q2, the Critical Data queue (assigning buffer space for Q3, the PQ in this model, is not supported on this line card).

The WRR weights for Q1 and Q2 (for dividing the remaining bandwidth, after the priority queue has been fully serviced) can be set to 30:70 respectively for Q1:Q2.

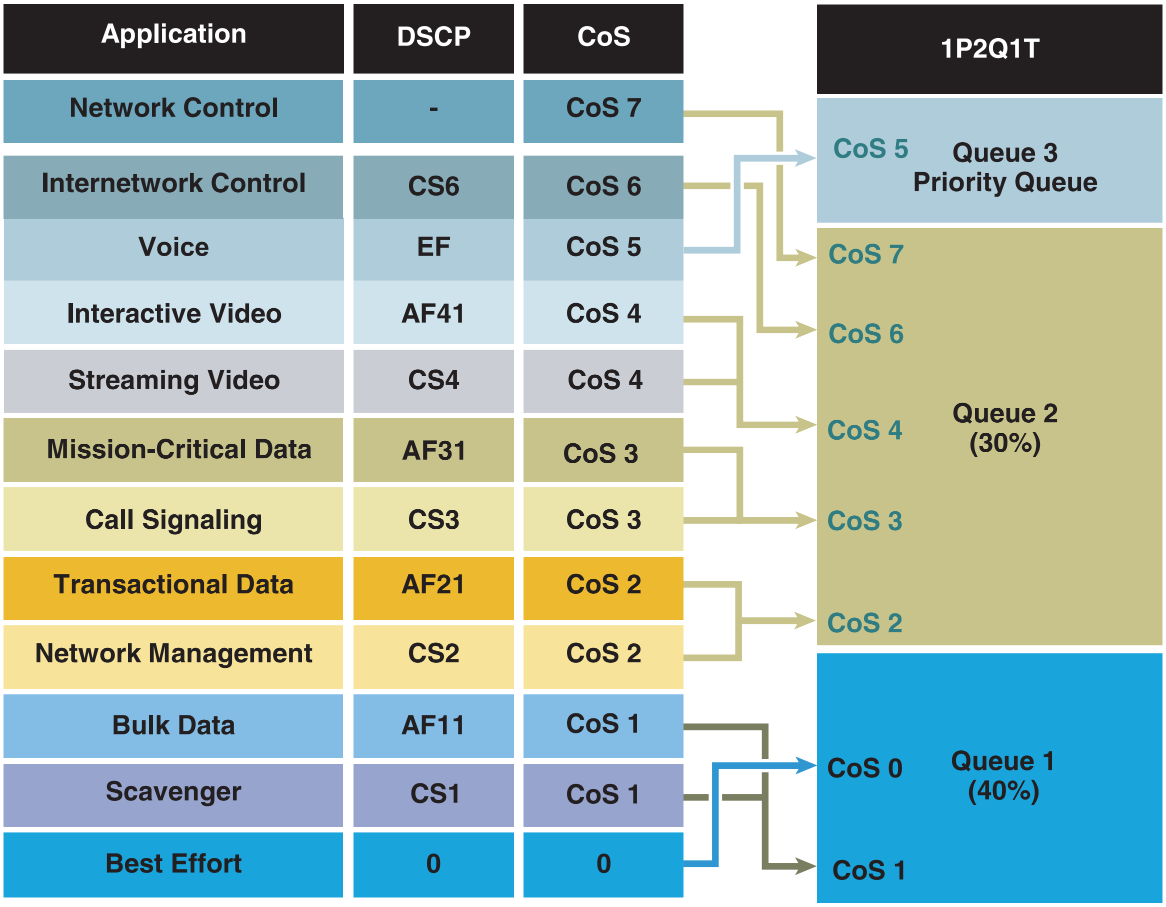
The Q1T1 WRED threshold can be set to 80:100 and the Q2T1 WRED threshold can be set to 80:100.

After these WRED thresholds have been altered, the following assignments can be made:

- CoS 1 (Scavenger/Bulk) and CoS 0 (Best Effort) to Q1T1.
- CoS 2 (Network Management and Transactional Data), CoS 3 (Call-Signaling and Mission-Critical Data), CoS 4 (Interactive and Streaming Video) ,and CoS 6 and 7 (Internetwork and Network Control) to Q2T1.
- CoS 5 (VoIP) to Q3 (the PQ).

These Cisco 7600 1P2Q1T queuing recommendations are illustrated in Figure A-7.

Figure A-7 Cisco 7600 1P2Q1T Queuing Model



The Cisco 7600 commands to configure 1P2Q1T queuing recommendations are shown in the following configuration example.

```
C7600(config)#interface TenGigabitEthernet1/1
C7600(config-if)# wrr-queue queue-limit 30 40
! Sets the buffer allocations to 30% for Q1 and 40% for Q2
C7600(config-if)# wrr-queue bandwidth 30 70
! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 1 80
! Sets Min WRED Threshold for Q1T1 to 80%
C7600(config-if)# wrr-queue random-detect max-threshold 1 100
! Sets Max WRED Threshold for Q1T1 to 100%
C7600(config-if)# wrr-queue random-detect min-threshold 2 80
! Sets Min WRED Threshold for Q2T1 to 80%
C7600(config-if)# wrr-queue random-detect max-threshold 2 100
! Sets Max WRED Threshold for Q2T1 to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue cos-map 1 1 1 0
! Assigns Scavenger/Bulk and Best Effort to Q1 WRED Threshold 1
C7600(config-if)# wrr-queue cos-map 2 1 2 3 4 6 7
! Assigns CoS 2,3,4,6 and 7 to Q2 WRED Threshold 1
C7600(config-if)# priority-queue cos-map 1 5
! Assigns VoIP to PQ (Q3)
C7600(config-if)#end
C7600(config-if)#
```

Cisco 7600 1P2Q2T GE Queuing Design

On the Cisco 7600, setting the size of the priority queue is not supported on any queuing structure with one exception: the 1P2Q2T structure, where the priority queue (Q3) is indirectly set to equal the Q2 size.

Under a 1P2Q2T model, buffer space can be allocated as follows: 40 percent for Q1 (the Scavenger/Bulk plus Best Effort queue) and 30 percent for Q2 (the Critical Data queue); therefore Q3 (the priority queue) is also indirectly set to 30 percent (to equal the size of Q2).

The WRR weights for Q1 and Q2 (for dividing the remaining bandwidth, after the priority queue has been fully serviced) remain at 30:70 respectively for Q1:Q2.

Under the 1P2Q2T model, each WRED threshold is defined with a lower and upper limit. Therefore the first WRED threshold for Q1 can be set to 40:80, so that Scavenger/Bulk Data traffic can be WRED-dropped if Q1 hits 40 percent and can be tail-dropped if Q1 exceeds 80 percent of its capacity; this prevents Scavenger/Bulk Data from drowning out Best Effort traffic in Q1. The second WRED threshold for Q1 can be set to 80:100 to provide congestion avoidance for Best Effort traffic.

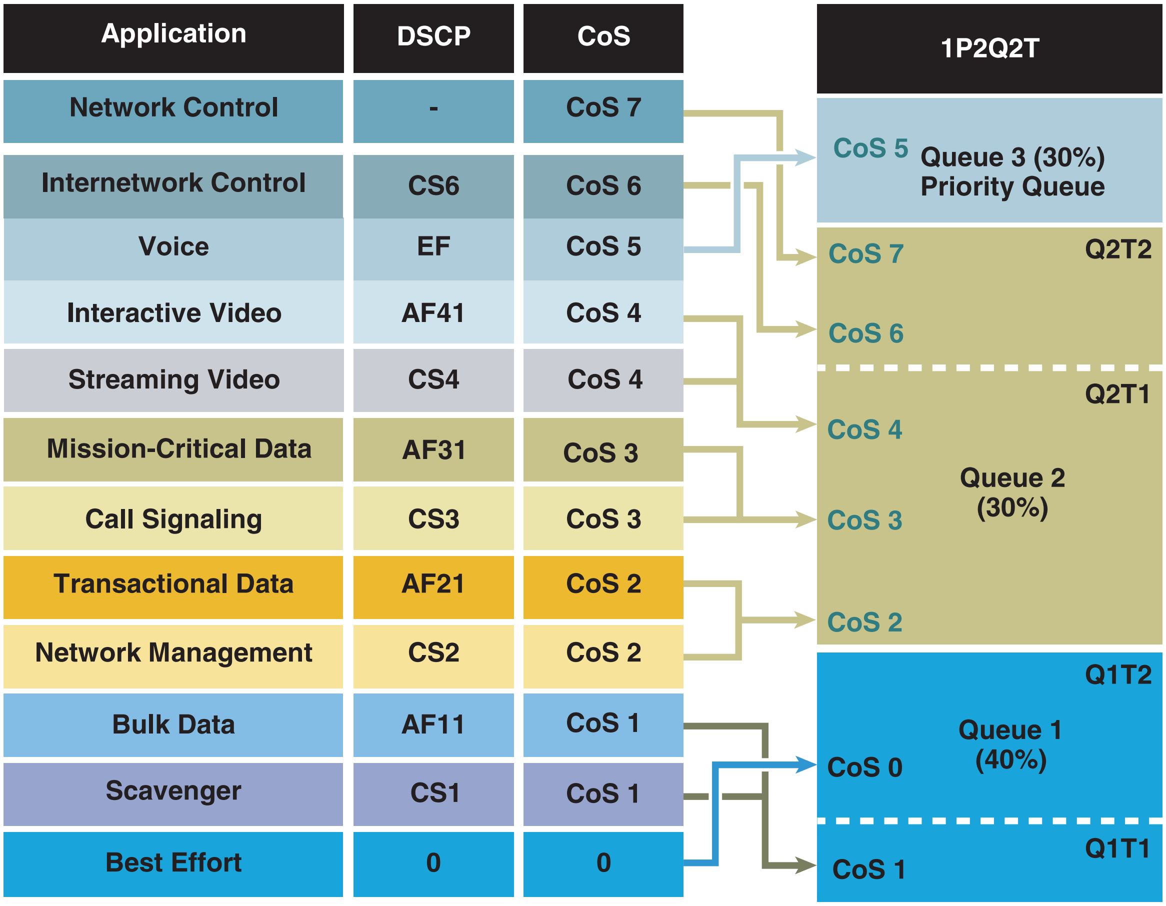
Similarly, the first WRED threshold of Q2 can be set to 70:80 and the second can be set to 80:100. In this manner, congestion avoidance is provided for all traffic types in Q2 and there is always room in the queue to service Network and Internetwork Control traffic.

After the queues have been defined as above, the following assignments can be made:

- CoS 1 (Scavenger/Bulk) to Q1T1
- CoS 0 (Best Effort) to Q1T2
- CoS 2 (Network Management and Transactional Data), CoS 3 (Call-Signaling and Mission-Critical Data), and CoS 4 (Interactive and Streaming Video) to Q2T1
- CoS 6 and 7 (Internetwork and Network Control) to Q2T2
- CoS 5 (VoIP) to Q3T1 (the PQ)

These 1P2Q2T queuing recommendations are illustrated in Figure A-8.

Figure A-8 Cisco 7600 1P2Q2T Queuing Model



The Cisco 7600 commands to configure 1P2Q2T queuing recommendations are shown in the following configuration example.

```
C7600(config)#interface range GigabitEthernet4/1 - 8
C7600(config-if-range)# wrr-queue queue-limit 40 30
! Sets the buffer allocations to 40% for Q1 and 30% for Q2
! Indirectly sets PQ (Q3) size to equal Q2 (which is set to 30%)
C7600(config-if-range)# wrr-queue bandwidth 30 70
! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 1 40 80
! Sets Min WRED Thresholds for Q1T1 and Q1T2 to 40 and 80
C7600(config-if-range)# wrr-queue random-detect max-threshold 1 80 100
! Sets Max WRED Thresholds for Q1T1 and Q1T2 to 80 and 100
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 2 70 80
! Sets Min WRED Thresholds for Q2T1 and Q2T2 to 70 and 80
C7600(config-if-range)# wrr-queue random-detect max-threshold 2 80 100
! Sets Max WRED Thresholds for Q2T1 and Q2T2 to 80 and 100
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
C7600(config-if-range)# wrr-queue cos-map 1 2 0
! Assigns Best Effort to Q1 WRED Threshold 2
C7600(config-if-range)# wrr-queue cos-map 2 1 2 3 4
! Assigns CoS 2,3,4 to Q2 WRED Threshold 1
```

```
C7600(config-if-range)# wrr-queue cos-map 2 2 6 7
! Assigns Network/Internetwork Control to Q2 WRED Threshold 2
C7600(config-if-range)#
C7600(config-if-range)# priority-queue cos-map 1 5
! Assigns VoIP to PQ
C7600(config-if-range)#end
C7600#
```

Cisco 7600 1P3Q1T GE Queuing Design

Tuning the transmit size ratios is not supported in the Cisco 7600 1P3Q1T queuing structure. Furthermore under this queuing model Q4 becomes the priority queue.

The WRR weights for the standard queues (Q1, Q2, Q3) for dividing the remaining bandwidth, after the priority queue has been fully serviced, can be set to 5:25:70 respectively for Q1:Q2:Q3.

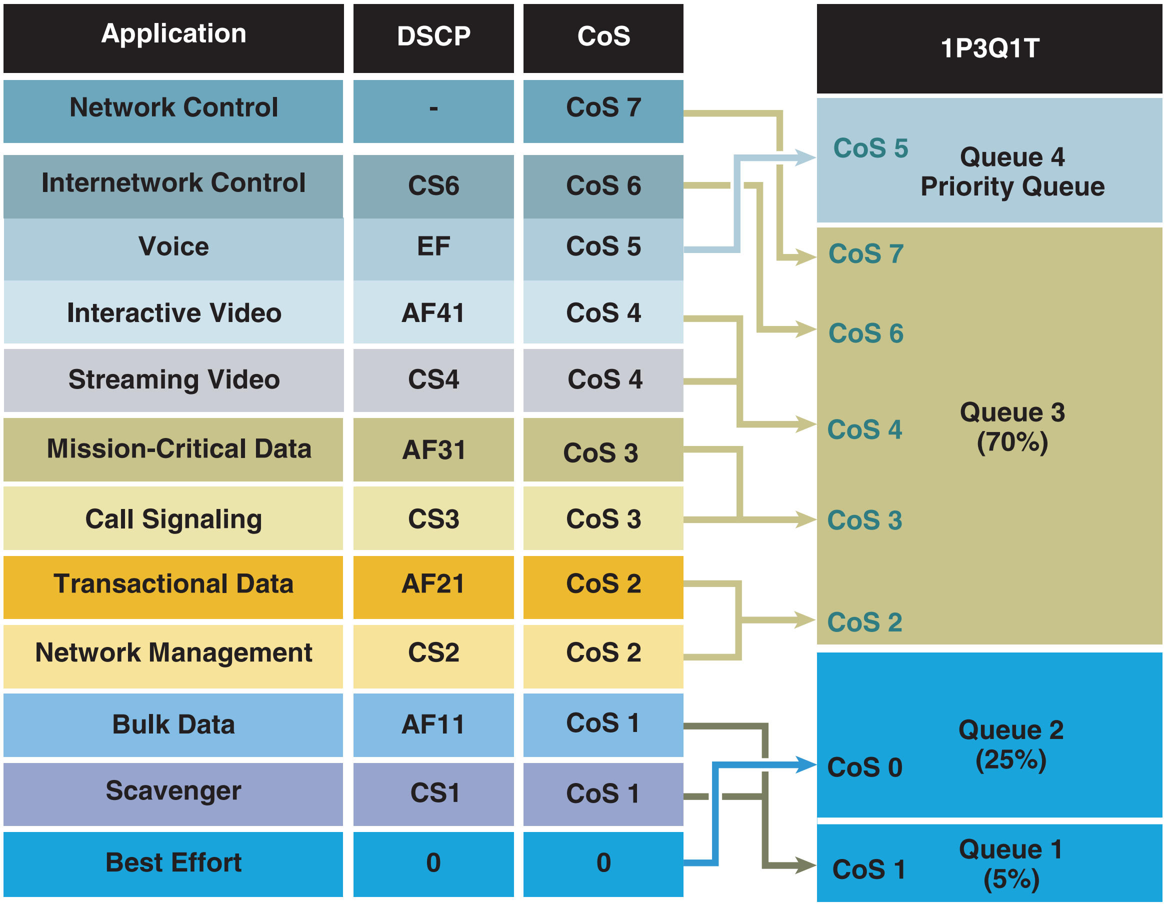
The Cisco 7600 1P3Q1T, 1P3Q8T, and 1P7Q8T queuing structures can be configured to use tail-drop or WRED. By default, WRED is disabled. Therefore, it is good practice to always explicitly enable WRED on a queue before setting WRED thresholds for these queuing structures. The WRED thresholds for all three preferential queues can be set to 80:100.

After the queues and thresholds have been defined as above, the following assignments can be made:

- CoS 1 (Scavenger/Bulk) to Q1T1
- CoS 0 (Best Effort) to Q2T1
- CoS 2 (Network Management and Transactional Data), CoS 3 (Call-Signaling and Mission-Critical Data), CoS 4 (Interactive and Streaming Video), and CoS 6 and 7 (Internetwork and Network Control) to Q3T1
- CoS 5 (VoIP) to Q4 (the PQ)

These 1P3Q1T queuing recommendations are illustrated in Figure A-9.

Figure A-9 Cisco 7600 1P3Q1T Queuing Model



Application	DSCP	CoS
Network Control	-	CoS 7
Internetwork Control	CS6	CoS 6
Voice	EF	CoS 5
Interactive Video	AF41	CoS 4
Streaming Video	CS4	CoS 4
Mission-Critical Data	AF31	CoS 3
Call Signaling	CS3	CoS 3
Transactional Data	AF21	CoS 2
Network Management	CS2	CoS 2
Bulk Data	AF11	CoS 1
Scavenger	CS1	CoS 1
Best Effort	0	0

The Cisco 7600 commands to configure 1P3Q1T queuing recommendations are shown in the following configuration example.

```
C7600(config)# interface range FastEthernet3/1 - 48
C7600(config-if)# wrr-queue bandwidth 5 25 70
! Sets the WRR weights for 5:25:70 (Q1:Q2:Q3) bandwidth servicing
C7600(config-if)#
C7600(config-if)#
C7600(config-if-range)# wrr-queue random-detect 1
! Enables WRED on Q1
C7600(config-if-range)# wrr-queue random-detect 2
! Enables WRED on Q2
C7600(config-if-range)# wrr-queue random-detect 3
! Enables WRED on Q3
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 1 80
! Sets Min WRED Threshold for Q1T1 to 80%
C7600(config-if)# wrr-queue random-detect max-threshold 1 100
! Sets Max WRED Threshold for Q1T1 to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 2 80
! Sets Min WRED Threshold for Q2T1 to 80%
C7600(config-if)# wrr-queue random-detect max-threshold 2 100
! Sets Max WRED Threshold for Q2T1 to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 3 80
```

```
! Sets Min WRED Threshold for Q3T1 to 80%
C7600(config-if)# wrr-queue random-detect max-threshold 3 100
! Sets Max WRED Threshold for Q3T1 to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1 (80:100)
C7600(config-if)# wrr-queue cos-map 2 1 0
! Assigns Best Effort to Q2 WRED Threshold 1 (80:100)
C7600(config-if)# wrr-queue cos-map 3 1 2 3 4 6 7
! Assigns CoS 2,3,4,6 and 7 to Q3 WRED Threshold 1 (80:100)
C7600(config-if)# priority-queue cos-map 1 5
! Assigns VoIP to PQ (Q4)
C7600(config-if)#end
C7600#
```

Cisco 7600 1P3Q8T GE Queuing Design

The Cisco 7600 1P3Q8T queuing structure is identical to the 1P3Q1T structure except that it has eight tunable WRED thresholds per queue (instead of one) and it also supports tuning the transmit size ratios.

Under a 1P3Q8T model buffer space can be allocated as follows: 5 percent for the Scavenger/Bulk queue (Q1), 25 percent for the Best Effort queue (Q2), 40 percent for the Critical Data queue (Q3), and 30 percent for the strict priority queue (Q4).

The WRR weights for the standard queues (Q1, Q2, Q3) for dividing the remaining bandwidth, after the priority queue has been fully serviced, can be set to 5:25:70 respectively for Q1:Q2:Q3.

The tunable WRED threshold for Q1 can be set to 80:100 to provide congestion avoidance to Scavenger/Bulk Data traffic. The WRED threshold for Q2 similarly can be set to 80:100 to provide congestion avoidance on all Best Effort flows.

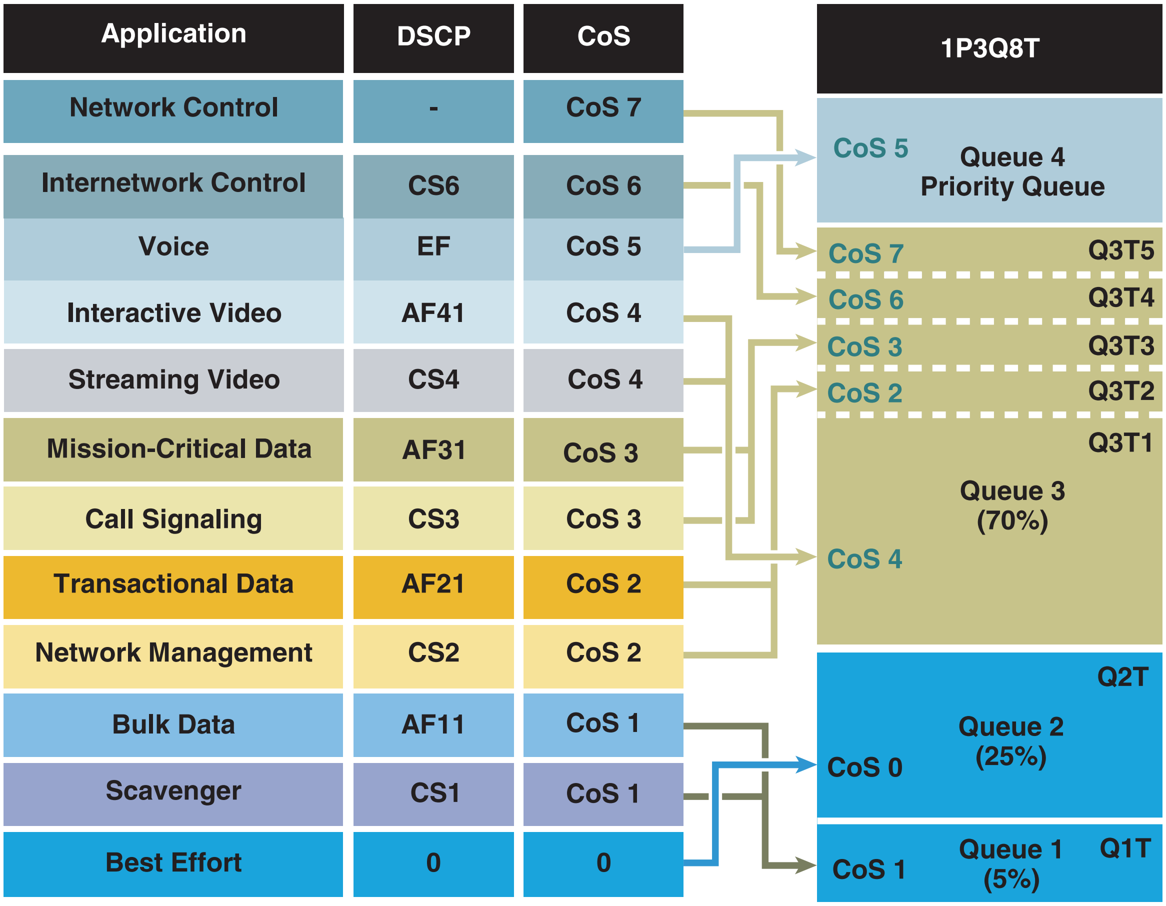
The 1P3Q8T queuing structure support for up to eight WRED thresholds per queue allows for additional QoS granularity for the applications sharing Q3. Because only five discrete CoS values are sharing this queue, only five of eight thresholds need to be defined for subqueue QoS. For example, Q3T1 can be set to 50:60, Q3T2 can be set to 60:70, Q3T3 can be set to 70:80, Q3T4 can be set to 80:90, and Q3T5 can be set to 90:100.

After the queues and thresholds have been defined as above, the following assignments can be made:

- CoS 1 (Scavenger/Bulk) to Q1T1
- CoS 0 (Best Effort) to Q2T1
- CoS 4 (Interactive and Streaming Video) to Q3T1
- CoS 2 (Network Management and Transactional Data) to Q3T2
- CoS 3 (Call-Signaling and Mission-Critical Data) to Q3T3
- CoS 6 (Internetwork Control) to Q3T4
- CoS 7 (Internetwork and Network Control) to Q3T5
- CoS 5 (VoIP) to Q4 (the PQ)

These Cisco 7600 1P3Q8T queuing recommendations are illustrated in Figure A-10.

Figure A-10 Cisco 7600 1P3Q8T Queuing Model



Application	DSCP	CoS
Network Control	-	CoS 7
Internetwork Control	CS6	CoS 6
Voice	EF	CoS 5
Interactive Video	AF41	CoS 4
Streaming Video	CS4	CoS 4
Mission-Critical Data	AF31	CoS 3
Call Signaling	CS3	CoS 3
Transactional Data	AF21	CoS 2
Network Management	CS2	CoS 2
Bulk Data	AF11	CoS 1
Scavenger	CS1	CoS 1
Best Effort	0	0

The Cisco 7600 commands to configure 1P3Q8T queuing recommendations are shown in the following configuration example.

```
C7600(config)# interface range GigabitEthernet1/1 - 48
C7600(config-if)# wrr-queue queue-limit 5 25 40
! Allocates 5% for Q1, 25% for Q2 and 40% for Q3
C7600(config-if)# wrr-queue bandwidth 5 25 70
! Sets the WRR weights for 5:25:70 (Q1:Q2:Q3) bandwidth servicing
C7600(config-if)#
C7600(config-if-range)# wrr-queue random-detect 1
! Enables WRED on Q1
C7600(config-if-range)# wrr-queue random-detect 2
! Enables WRED on Q2
C7600(config-if-range)# wrr-queue random-detect 3
! Enables WRED on Q3
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 1 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q1T1 to 80% and all others to 100%
C7600(config-if)# wrr-queue random-detect max-threshold 1 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q1T1 to 100% and all others to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 2 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q2T1 to 80% and all others to 100%
```

```
C7600(config-if)# wrr-queue random-detect max-threshold 2 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q2T1 to 100% and all others to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue random-detect min-threshold 3 50
60 70 80 90 100 100 100
! Sets Min WRED Threshold for Q3T1 to 50%, Q3T2 to 60%,
! Q3T3 to 70%, Q3T4 to 80%, Q3T5 to 90% and all others to 100%
C7600(config-if)# wrr-queue random-detect max-threshold 3 60
70 80 90 100 100 100 100
! Sets Max WRED Threshold for Q3T1 to 60%, Q3T2 to 70%,
! Q3T3 to 80%, Q3T4 to 90%, Q3T5 to 100% and all others to 100%
C7600(config-if)#
C7600(config-if)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
C7600(config-if)# wrr-queue cos-map 2 1 0
! Assigns Best Effort to Q2 WRED Threshold 1
C7600(config-if)# wrr-queue cos-map 3 1 4
! Assigns Video to Q3 WRED Threshold 1
C7600(config-if)# wrr-queue cos-map 3 2 2
! Assigns Net-Mgmt and Transactional Data to Q3 WRED T2
C7600(config-if)# wrr-queue cos-map 3 3 3
! Assigns call signaling and Mission-Critical Data to Q3 WRED T3
C7600(config-if)# wrr-queue cos-map 3 4 6
! Assigns Internetwork-Control (IP Routing) to Q3 WRED T4
C7600(config-if)# wrr-queue cos-map 3 5 7
! Assigns Network-Control (Spanning Tree) to Q3 WRED T5
C7600(config-if)# priority-queue cos-map 1 5
! Assigns VoIP to the PQ (Q4)
C7600(config-if)#end
C7600#
```

Cisco 7600 1P7Q8T 10GE Queuing Design

The Cisco 7600 1P7Q8T queuing structure adds four additional standard queues to the 1P3Q8T structure and moves the PQ from Q4 to Q8, but otherwise is identical.

Under a 1P7Q8T model, buffer space can be allocated as follows:

- 5 percent for the Scavenger/Bulk queue (Q1)
- 25 percent for the Best Effort queue (Q2)
- 10 percent for the Video queue (Q3)
- 10 percent for the Network-Management/Transactional Data queue (Q4)
- 10 percent for the Call-Signaling/Mission-Critical Data queue (Q5)
- 5 percent for the Internetwork-Control queue (Q6)
- 5 percent for the Network Control queue (Q7)
- 30 percent for the PQ (Q8)

The WRR weights for the standard queues (Q1 through Q7) for dividing the remaining bandwidth after the priority queue has been fully serviced can be set to 5:25:20:20:20:5:5 respectively for Q1 through Q7.

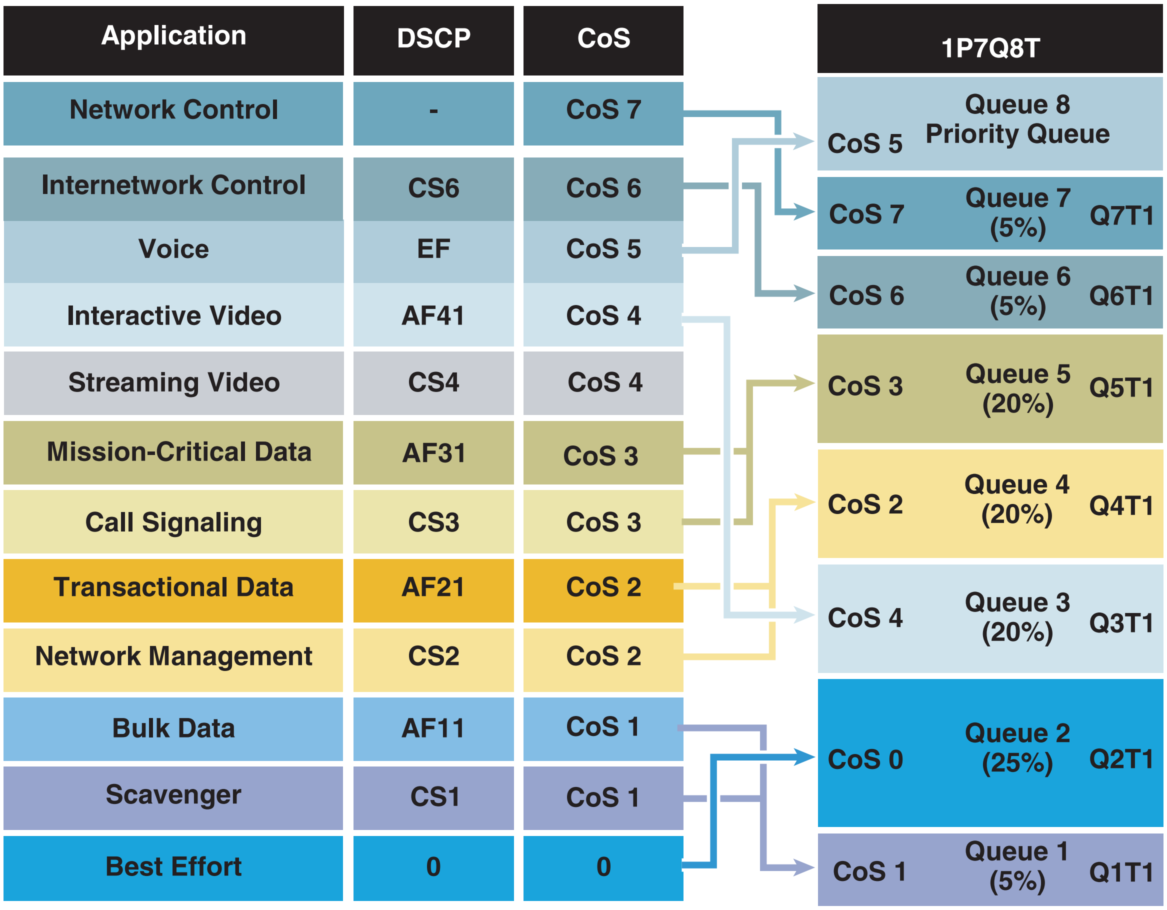
Because eight queues are available, each CoS value can be assigned to its own exclusive queue. WRED can be enabled on each queue to provide it with congestion avoidance by setting the first WRED threshold of each queue to 80:100. All other WRED thresholds can remain at 100:100.

After the queues and thresholds have been defined as above, the following assignments can be made:

- CoS 1 (Scavenger/Bulk) to Q1T1
- CoS 0 (Best Effort) to Q2T1
- CoS 4 (Interactive and Streaming Video) to Q3T1
- CoS 2 (Network Management and Transactional Data) to Q4T1
- CoS 3 (Call-Signaling and Mission-Critical Data) to Q5T1
- CoS 6 (Internetwork Control) to Q6T1
- CoS 7 (Internetwork and Network Control) to Q7T1
- CoS 5 (VoIP) to Q8 (the PQ)

These 1P7Q8T queuing recommendations are illustrated in Figure A-11.

Figure A-11 Cisco 7600 1P7Q8T Queuing Model



Application	DSCP	CoS
Network Control	-	CoS 7
Internetwork Control	CS6	CoS 6
Voice	EF	CoS 5
Interactive Video	AF41	CoS 4
Streaming Video	CS4	CoS 4
Mission-Critical Data	AF31	CoS 3
Call Signaling	CS3	CoS 3
Transactional Data	AF21	CoS 2
Network Management	CS2	CoS 2
Bulk Data	AF11	CoS 1
Scavenger	CS1	CoS 1
Best Effort	0	0

1P7Q8T		
CoS 5	Queue 8 Priority Queue	
CoS 7	Queue 7 (5%)	Q7T1
CoS 6	Queue 6 (5%)	Q6T1
CoS 3	Queue 5 (20%)	Q5T1
CoS 2	Queue 4 (20%)	Q4T1
CoS 4	Queue 3 (20%)	Q3T1
CoS 0	Queue 2 (25%)	Q2T1
CoS 1	Queue 1 (5%)	Q1T1

The Cisco 7600 commands configure 1P7Q8T queuing recommendations are shown in the following configuration example.

```
C7600(config)#interface range TenGigabitEthernet4/1 - 4
C7600(config-if-range)# wrr-queue queue-limit 5 25 10 10 10 5 5
! Allocates 5% to Q1, 25% to Q2, 10% to Q3, 10% to Q4,
! Allocates 10% to Q5, 5% to Q6 and 5% to Q7
C7600(config-if-range)# wrr-queue bandwidth 5 25 20 20 20 5 5
! Sets the WRR weights for 5:25:20:20:20:5:5 (Q1 through Q7)
C7600(config-if-range)#
```

```
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect 1
! Enables WRED on Q1
C7600(config-if-range)# wrr-queue random-detect 2
! Enables WRED on Q2
C7600(config-if-range)# wrr-queue random-detect 3
! Enables WRED on Q3
C7600(config-if-range)# wrr-queue random-detect 4
! Enables WRED on Q4
C7600(config-if-range)# wrr-queue random-detect 5
! Enables WRED on Q5
C7600(config-if-range)# wrr-queue random-detect 6
! Enables WRED on Q6
C7600(config-if-range)# wrr-queue random-detect 7
! Enables WRED on Q7
C7600(config-if-range)#
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 1 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q1T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 1 100 100 100 100 100 100
100 100
! Sets Max WRED Threshold for Q1T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 2 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q2T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 2 100 100 100 100 100 100
100 100
! Sets Max WRED Threshold for Q2T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 3 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q3T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 3 100 100 100 100 100 100
100 100
! Sets Max WRED Threshold for Q3T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 4 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q4T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 4 100 100 100 100 100 100
100 100
! Sets Max WRED Threshold for Q4T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 5 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q5T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 5 100 100 100 100 100 100
100 100
! Sets Max WRED Threshold for Q5T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 6 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q6T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 6 100 100 100 100 100 100
100 100
! Sets Max WRED Threshold for Q6T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue random-detect min-threshold 7 80 100 100 100 100 100 100
100
! Sets Min WRED Threshold for Q7T1 to 80% and all others to 100%
C7600(config-if-range)# wrr-queue random-detect max-threshold 7 100 100 100 100 100 100
100 100
```

```
! Sets Max WRED Threshold for Q7T1 to 100% and all others to 100%
C7600(config-if-range)#
C7600(config-if-range)#
C7600(config-if-range)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
C7600(config-if-range)# wrr-queue cos-map 2 1 0
! Assigns Best Effort to Q2 WRED Threshold 1
C7600(config-if-range)# wrr-queue cos-map 3 1 4
! Assigns Video to Q3 WRED Threshold 1
C7600(config-if-range)# wrr-queue cos-map 4 1 2
! Assigns Net-Mgmt and Transactional Data to Q4 WRED T1
C7600(config-if-range)# wrr-queue cos-map 5 1 3
! Assigns call signaling and Mission-Critical Data to Q5 WRED T1
C7600(config-if-range)# wrr-queue cos-map 6 1 6
! Assigns Internetwork-Control (IP Routing) to Q6 WRED T1
C7600(config-if-range)# wrr-queue cos-map 7 1 7
! Assigns Network-Control (Spanning Tree) to Q7 WRED T1
C7600(config-if-range)# priority-queue cos-map 1 5
! Assigns VoIP to the PQ (Q4)
C7600(config-if-range)#end
C7600-IOS#
```

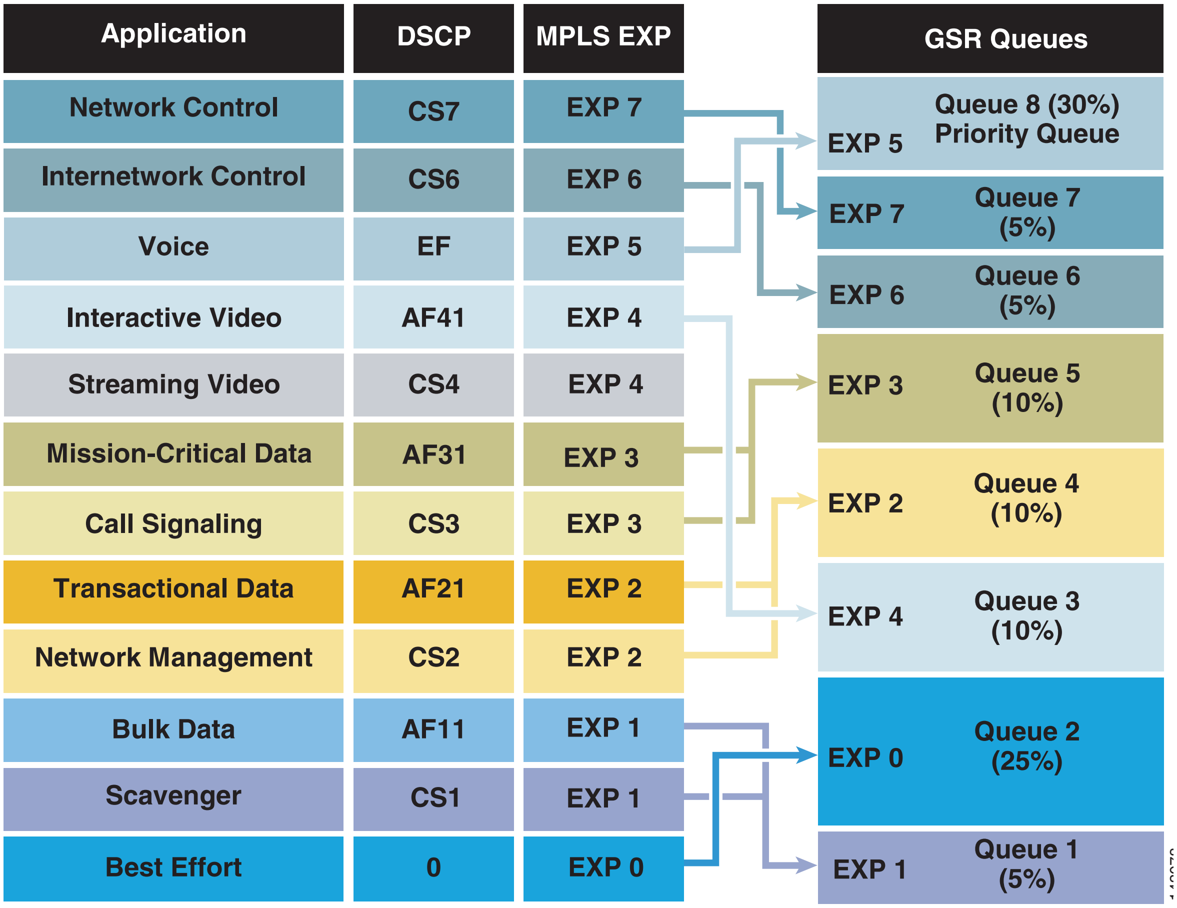
Cisco 12000 QoS Design

Cisco 12000 series routers are also an option as edge or core routers in the NG-WAN/MAN. These high performance routers have long been deployed in performance intensive networks of service providers and enterprises and offer extremely rich features with high performance. This section describes the specific QoS considerations when Cisco 12000 series routers is used in enterprise networks as the PE and the P routers using ISE (or Engine 5) Gigabit line cards.

Cisco 12000 GSR Edge Configuration

MPLS EXP is three bits in length and can support a maximum of eight traffic classes. As mentioned earlier, you assume the maximum of 8 classes in the core. The 11 enterprise traffic classes are mapped to the 8 core classes at the ingress PE. Although not more than three traffic classes are typical in a service provider core, an enterprise core can have up to 8 traffic classes to provide better control over the individual classes as illustrated in Figure A-12.

Figure A-12 Cisco 12000 8-Class Queuing Model



Application	DSCP	MPLS EXP
Network Control	CS7	EXP 7
Internetwork Control	CS6	EXP 6
Voice	EF	EXP 5
Interactive Video	AF41	EXP 4
Streaming Video	CS4	EXP 4
Mission-Critical Data	AF31	EXP 3
Call Signaling	CS3	EXP 3
Transactional Data	AF21	EXP 2
Network Management	CS2	EXP 2
Bulk Data	AF11	EXP 1
Scavenger	CS1	EXP 1
Best Effort	0	EXP 0

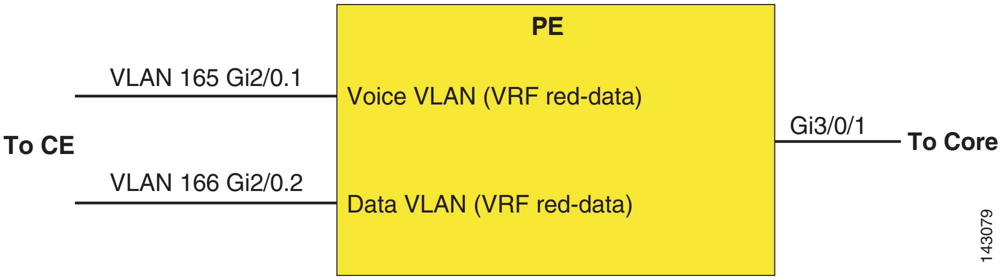
GSR Queues	
EXP 5	Queue 8 (30%) Priority Queue
EXP 7	Queue 7 (5%)
EXP 6	Queue 6 (5%)
EXP 3	Queue 5 (10%)
EXP 2	Queue 4 (10%)
EXP 4	Queue 3 (10%)
EXP 0	Queue 2 (25%)
EXP 1	Queue 1 (5%)

The QoS features of ISE (Engine 3) 4-port Gigabit Ethernet line card makes it very suitable for being used as an edge line card although it can also be used in the core as well. The following sections are based on using this line card as an edge line card. Engine 5 line cards, being QoS compatible with the Engine 3 line cards, can be used instead. Cisco IOS 12.0.30S2 or above is assumed.

The simplest QoS configuration at the ingress PE is to assume that all traffic classes are processed by the main interface without having any subinterface. In this case a single service policy is attached to the main interface (or sub-interface).

However some enterprises may need to segregate their Voice and other traffic into separate VRFs. In this case you can send the traffic via different VLANs terminating at separate subinterfaces. These subinterfaces are mapped to different VRFs. The service policy is still attached to the main interface as shown in Figure A-13.

Figure A-13 Sample PE Configuration with Separate VRFs for Voice and other Data Traffic



PE Config (CE Facing Configuration—Ingress QoS)

No specific ingress QoS is configured for policing or marking in this example. Because the PE router is not separating QoS domains with different marking policies, no packet remarking is necessary. Further it is assumed that the default mapping of IP Precedence (or DSCP) to MPLS EXP is used (although it is fine to have a policy map use a different mapping on ingress, if so desired).

An enterprise may not enforce any rate limits of different traffic classes. In case rate-limiting is a requirement, an appropriate service-policy using policing can be attached to the individual sub-interfaces.

PE Config (CE Facing Configuration—Egress QoS)

```
interface GigabitEthernet2/0
 description To DL2 - intf G5/2 - CE facing
 no ip address
 no ip directed-broadcast
 negotiation auto
 service-policy output q-2ce-out-parent
!
interface GigabitEthernet2/0.1
 description RED-DATA
 encapsulation dot1Q 165
 ip vrf forwarding red-data
 ip address 125.1.102.49 255.255.255.252
 no ip directed-broadcast
 ip pim sparse-mode
!
interface GigabitEthernet2/0.2
 description RED-VOICE
 encapsulation dot1Q 166
 ip vrf forwarding red-voice
 ip address 125.1.102.53 255.255.255.252
 no ip directed-broadcast
 ip pim sparse-mode

class-map match-all red-voice   <------- VLAN 166 carries Voice Traffic (plus routing
traffic)
  match vlan  166
class-map match-all red-data   < -------- VLAN 165 carries rest of the Traffic classes (+
routing)
  match vlan  165
```

```
class-map match-any realtime-2ce
  match qos-group 5
class-map match-any network-control-2ce
  match qos-group 7
class-map match-any bulk-data-2ce
  match qos-group 1
class-map match-any interwork-control-2ce
  match qos-group 6
  match  IP precedence 6
class-map match-any bus-critical-2ce
  match qos-group 3
class-map match-any trans-data-2ce
  match qos-group 2
class-map match-any video-2ce
  match qos-group 4



policy-map q-2ce-out-parent  <-------- Policy map attached to the main interface
  class red-data           <------- No Priority traffic, but seven classes of traffic,
plus OSPF
    shape average percent 50
    service-policy q-2ce-out-1
  class red-voice                      <----- Carries voice traffic + OSPF
    shape average percent 40
    service-policy q-2ce-out-2

The child policy for voice VRF:

policy-map q-2ce-out-2
  class realtime-2ce
    priority
   police cir percent 95 bc 500 ms conform-action transmit  exceed-action drop
  class interwork-control-2ce     <== OSPF
    bandwidth percent 3
    random-detect
    random-detect precedence 6 4720 packets 4721 packets 1
  class class-default
    bandwidth percent 2
    random-detect                        <===  There is no traffic in this class
    random-detect precedence 6 4720 packets 4721 packets 1
```

Note If a traffic class is configured with a “priority” command using MQC, important traffic marked with PAK_PRIORITY (for example, routing traffic) goes either to the class matching IP Precedence 6 if defined or to a class matching IP Precedence 7 if defined or to the “class-default.” Because in the above case a class with the “priority” command is defined, OSPF traffic to the CE goes to the interwork-control-2ce class that matches IP Precedence 6.

The child policy map for data VRF is as follows:

```
policy-map q-2ce-out-1           <--- DATA VLAN policy (no priority Q)
  class network-control-2ce
    bandwidth percent 7
    random-detect discard-class-based
    random-detect discard-class 7 625 packets 4721 packets 1
  class interwork-control-2ce
    bandwidth percent 7
    random-detect discard-class-based
    random-detect discard-class 6 625 packets 4721 packets 1
  class bus-critical-2ce
    bandwidth percent 14
    random-detect discard-class-based
    random-detect discard-class 3 625 packets 4721 packets 1
  class trans-data-2ce
    bandwidth percent 14
    random-detect discard-class-based
    random-detect discard-class 2 625 packets 4721 packets 1
  class video-2ce
    bandwidth percent 14
    random-detect discard-class-based
    random-detect discard-class 4 625 packets 4721 packets 1
  class bulk-data-2ce
    bandwidth percent 7
    random-detect discard-class-based
    random-detect discard-class 1 625 packets 4721 packets 1
  class class-default
    bandwidth percent 36
    random-detect discard-class-based
    random-detect discard-class 0 625 packets 4721 packets 1
```

Note If a traffic class is *not* configured with a “priority” command using MQC (as in the above case, routing traffic goes either to the class matching IP Precedence 6 if defined or to a class matching IP Precedence 7 if defined or to the priority queue. Because in the above case there is *not* defined a class with the “priority” command, OSPF traffic to the CE goes to the priority queue.

```
PE Config (P Facing Configuration – Egress QoS)
class-map match-any realtime
  match mpls experimental  5
class-map match-any bulk-data
  match mpls experimental  1
class-map match-any interwork-control
  match mpls experimental  6
class-map match-any network-control
  match mpls experimental  7
class-map match-any bus-critical
  match mpls experimental  3
class-map match-any trans-data
  match mpls experimental  2
class-map match-any video
  match mpls experimental  4

policy-map q-core-out
  class realtime
    priority
   police cir percent 30 bc 500 ms conform-action transmit  exceed-action drop
  class network-control
    bandwidth remaining percent 7
    random-detect
    random-detect precedence 7 625 packets 4721 packets 1
  class interwork-control
    bandwidth remaining percent 7
    random-detect
    random-detect precedence 6 625 packets 4721 packets 1
  class bus-critical
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 3 625 packets 4721 packets 1
  class trans-data
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 2 625 packets 4721 packets 1
  class video
    bandwidth remaining percent 14
    random-detect
    random-detect precedence 4 625 packets 4721 packets 1
  class bulk-data
    bandwidth remaining percent 7
    random-detect
    random-detect precedence 1 625 packets 4721 packets 1
  class class-default
    bandwidth remaining percent 36
    random-detect
    random-detect precedence 0 625 packets 4721 packets 1

interface GigabitEthernet3/3/0
 description To P1 - intf G4/0/2
 ip address 125.1.102.6 255.255.255.252
 no ip directed-broadcast
 ip pim sparse-mode
 load-interval 30
 negotiation auto
 tag-switching ip
 service-policy input egr-pe-in
 service-policy output q-core-out
```

PE Config (P Facing Configuration—Ingress QoS)

```
class-map match-any realtime
  match mpls experimental  5
class-map match-any bulk-data
  match mpls experimental  1
class-map match-any interwork-control
  match mpls experimental  6
class-map match-any network-control
  match mpls experimental  7
class-map match-any bus-critical
  match mpls experimental  3
class-map match-any trans-data
  match mpls experimental  2
class-map match-any video
  match mpls experimental  4

policy-map egr-pe-in
  class realtime
   set qos-group 5
   set discard-class 5
  class network-control
   set qos-group 7
   set discard-class 7
  class interwork-control
   set qos-group 6
   set discard-class 6
  class bus-critical
   set qos-group 3
   set discard-class 3
  class trans-data
   set qos-group 2
   set discard-class 2
  class video
   set qos-group 4
   set discard-class 4
  class bulk-data
   set qos-group 1
   set discard-class 1
```

The following are general notes on the ISE 4-Port GigabitEthernet line card:

- This line card supports only four traffic classes by default. However it can support up to eight traffic classes when you configure the following:

  ```
  hw-module slot <slot#> qos interface queues 8
  ```

- The MQC bandwidth, shape commands, and policers by default use Layer 3 packet size for bandwidth calculations. If you want to include Layer 2 header size in bandwidth calculations, use the following command:

  ```
  hw-module slot slot-number qos-account-layer2-encapsulation {arpa | dot1q | length}
  ```

Cisco 12000 GSR ToFab Queuing

Typically, no specific QoS configuration is necessary for incoming traffic at the ingress interface. However, in a GSR, “line card-to-fabric” queuing (or toFab queuing) should be configured to avoid congestion of traffic from the line cards to the switching fabric. Such congestion may happen when multiple ports on the ingress line card receive incoming traffic peaks at the exact same time and the sum of the peak bandwidths exceed the fabric bandwidth.

WRED is less important here because WRED works on sustained congestion, while the ToFab congestion is likely to be of instantaneous nature. ToFab queuing on a GSR is configured using legacy CLI as in the following configuration.

Although the following sample configuration is for three core classes, it can be expanded accordingly when eight core classes are used.

```
slot-table-cos SLOT_TABLE
 destination-slot all core_policy
!
rx-cos-slot all SLOT_TABLE
!
cos-queue-group core_policy
  precedence 0 queue 0 ! Traffic with IP Prec 0 or EXP 0 go to queue # 0 (best  effort)
  precedence 1 queue 1
  precedence 2 queue 1 ! Traffic with IP Prec(EXP) 2,3,6,1 go to queue # 1 (businessclass)
  precedence 3 queue 1
  precedence 6 queue 1
  precedence 5 queue low-latency ! Traffic with IP Precedence 5 or EXP 5 go to PQ (real 
time class)
  precedence 0 random-detect-label 1       ! For traffic with precedence 0, use WRED label 
number 1
precedence 1 random-detect-label 0       ! For traffic with precedence 0, use WRED label 
number 1
                                                                 ! out-of-contract 
traffic, so lower thresholds
  precedence 2 random-detect-label 1
  precedence 3 random-detect-label 1
  precedence 6 random-detect-label 1
random-detect-label 0   500 1012 1          ! WRED label # 0 defines the WRED thresholds

random-detect-label 1 1500 9692 1
  queue 0 1                                                  ! Queue weight
  queue 1 71                                                ! Queue weight
  queue low-latency strict                           ! Configures strict priority to real 
time traffic class
```

Note the following:

- The above configuration is GSR specific and in native GSR CLI.

- The toFab queuing uses MDRR algorithm where you assign relative weights to the non- priority queues. The weights are used to calculate the quantum of a queue (maximum number of bytes that can be output from the queue in one round-robin cycle). The ratios of the quanta among the queues determine how the excess bandwidth is allocated to these queues after the PQ has been serviced.

  The Quantum for a queue is calculated as follows:

  Quantum = MTU + (queue weight - 1) * 512
- Although the configuration specifies IP Precedence, it also matches EXP for MPLS packets.
- See the next section for WRED parameter calculation

WRED Tuning at the Edge and Core

In general, WRED tuning is a complex task and depends on several factors such as traffic load, traffic mix, the ratio of the offered load to the link capacity, and traffic behavior in the presence of congestion (for example, how do the TCP stacks react to traffic drops may vary among implementations). These factors vary from network to network and depend on the type of services offered as well as the properties of the applications such as TCP stacks the customers run. It is therefore very difficult to provide recommendations that work equally well in all networks.

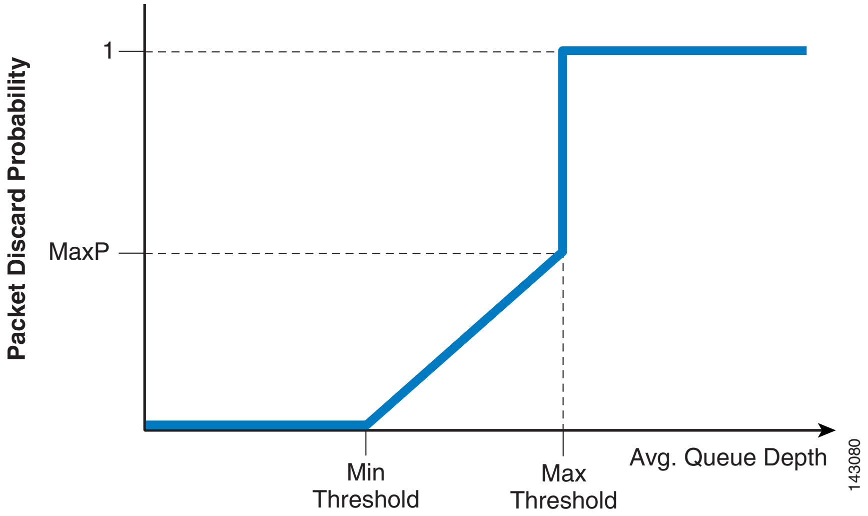
The recommendations given here should be taken as a starting point and should be further refined per real life testing and operational experience in the target network environment.

Without proper parameter setting, WRED may affect link utilization and latency. If the WRED thresholds are too large, a packet near the tail end of the queue can spend a longer time before being scheduled for transmission, thus increasing latency. On the other hand, having a small WRED queue may drop more packets, thus decreasing link utilization. WRED tuning involves setting WRED parameters such as the minimum, maximum thresholds, and drop probabilities so that the link utilization is maximized, while the mean WRED queue depth per class is minimized to decrease latency.

Typical WRED parameters that can be tuned are the minimum and maximum thresholds, the drop probability, and the exponential weighting constant. It is recommended to set the drop probability to 1 and to not change the default value of the exponential weighting constant. Tuning of the other parameters (min and max threshold) depends on the traffic volume, RTT (round trip time), and interface MTU (assumed to be 1500 for Ethernet).

Figure A-14 shows WRED parameters that can be tuned.

Figure A-14 WRED Parameters for Tuning



The WRED tuning goals and how they can be realized are described as follows for link speeds of 10 Mbps or higher:

- Min threshold should be high enough so that link utilization is maximized.

  Set min threshold = 0.15 * P, where

  P is the pipe size = RTT * BW / (MTU * 8)

  RTT = Round Trip Time ~ 100 ms (typical value used in USA)

  MTU = Maximum Transmission Unit of the interface (use 1500 as MTU in this formula, even if the actual MTU is configured on the interface is higher; for example, 4470)

- The difference between the min and the max threshold should be high enough to avoid TCP global synchronization.

  If the difference is small, then too many packets could be dropped in a very small time interval (which would be nearly similar to a tail-drop situation) leading to TCP global synchronization and consequent throughput reduction.

  The recommendation for max threshold is:

  Max threshold = 1 * P

The other rules that apply while calculating these parameters are as follows:

- The difference between the max and the min threshold should be a power of two (for GSR routers).
- Min and max threshold should be calculated based on the traffic volume on the link, rather than the full link speed. For example, assuming that traffic, on an average, would not exceed 50 percent of the link speed, the min and max thresholds would be half of the calculated values by the above formulas. Lower traffic volume may allow lower thresholds.
- Set the packet discard probability to 1; that is, drop all exceeding packets of a class after the WRED max threshold is exceeded.
- Do not change exponential weighting constant from its default.

With the above rules, min and max thresholds for an OC-48 link (2.5 Gbps) are calculated as follows:

P = 100ms * 2.5Gbps / (1500 * 8) = 20,000

Therefore, min threshold = 0.15 * 20,000 = 3000, and max threshold = 20000

Assuming 50 percent normal traffic load, min and max thresholds are adjusted to 1500 and 10000 respectively. However because their difference should be a power of 2, you further adjust the max threshold to 9692.

Similar calculations for OC-3, OC-12, and Gigabit Ethernet are summarized in Table A-3:

Table A-3 Minimum and Maximum Thresholds

Link Speed	Pipe Size (P)	Minimum Threshold	Maximum Threshold
OC-3	1292	97	609
OC-12	5184	389	2437
OC-48	20000	1500	9692
Gigabit		625	4176

Note that these values are to be taken as starting values. These assume 50 percent link rate utilization for the traffic class and can be further adjusted per the actual utilization for the classes and also per real life traffic pattern.