SD-WAN QUALITY OF SERVICE

Supplemental Guide

June 2019

Table of Contents

Quality of Service

Quality of service (QoS) refers to the ability of a network to provide higher levels of service using traffic prioritization and control mechanisms. Applying the proper QoS policy is important for real-time traffic that has specific latency requirements, such as Skype, video conferencing, or any business-critical applications. To accurately measure QoS on a network, there are several aspects to consider, such as bit rate, throughput, path availability, delay, jitter, and loss. You can improve the last three—delay, jitter and loss—by using an appropriate scheduling algorithm on the egress interfaces of your network devices to deliver applications with higher requirements before applications with lower requirements.

There are two main strategies to consider when creating a QoS scheduling policy. The first method is to identify applications that are important to your business and give them a higher level of service using the QoS scheduling techniques described in this guide. The remaining applications stay in the best-effort queue to minimize the upfront configuration time and to lower the day-to-day operational effort of troubleshooting a complex QoS policy. If additional applications become important in the future, you identify them and add the new applications to your list of business-critical applications. This can be repeated as needed without requiring a comprehensive policy for all applications on your network. This strategy is normally used by organizations who do not have a corporate-wide QoS policy or are troubleshooting application performance problems across their WAN.

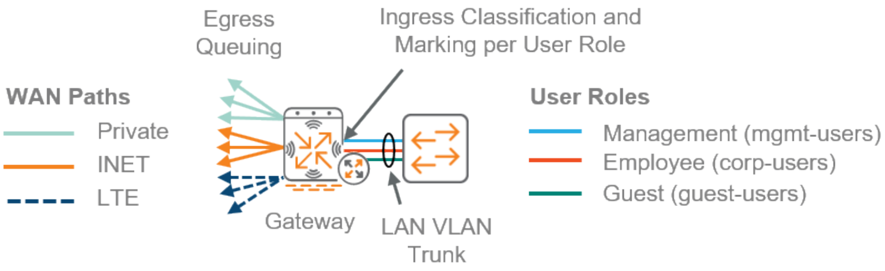
An example of this type of strategy is to prioritize your real-time voice and video applications, along with a few other key applications that require fast response times. This allows your employees to remain productive doing the network-based tasks that matter the most to your business. Real-time applications are always placed into a strict priority queue and business-critical applications are serviced by one or more premium queues that provide a higher level of service during times of congestion. The rest of the traffic is placed into a default queue with a lower level of service than the applications that are critical to running the business receive. If the higher priority applications are not using the bandwidth assigned, the default queue can use all available bandwidth.

The second strategy is to create a comprehensive QoS policy that identifies all traffic flows and applications. This is made possible using the Aruba deep packing inspection (DPI) engine, which can identify more than 3100 applications using well-known signatures and protocols. These applications are placed into pre-defined categories in the DPI engine for your convenience, but you may find it necessary to create your own custom groupings if they do not align with your organization’s needs. This strategy is best suited for organizations that have an existing QoS policy and want to use it with an SD-WAN solution. For a detailed explanation of an example QoS strategy using Aruba gateways, see “Comprehensive QoS Policy” in Appendix A.

CLASSIFICATION AND MARKING

No matter which strategy you choose, Aruba recommends you use the gateway as a QoS policy enforcement point for traffic over the WAN. This means you identify selected applications by user-role, subnet, or VLAN at the ingress of the gateway and place them into designated egress queues, as shown in Figure 1. Any applications that are not identified are placed into the default queue, giving them a best-effort level of service. East-west traffic that remains in the location is identified and marked when it passes through the gateway between the VLANs.

Figure 1 QoS policy enforcement point



Scheduling algorithms rely on classification markings to identify applications as they pass through a network device. Aruba recommends marking your applications with class of service (CoS) for queueing and optionally, differentiated service code point (DSCP) if your service provider (SP) honors the markings in their managed network. The goal of the QoS policy is to allow your critical applications to share the available WAN bandwidth with minimal system impact and engineering effort.

In a typical enterprise network, applications with similar characteristics are categorized based on how they operate over the network. These application categories are sent into different queues according to the types of applications in your organization. For example, if you are not using broadcast video or multimedia streaming applications for business purposes, there is no need to account for them in your QoS policy. Because the Aruba gateway supports four hardware queues, you can combine the real-time applications into one strict priority queue, put your critical applications and collaboration applications into deficit round robin (DRR) queues, and use the last queue for your default traffic. DRR is a packet-based scheduling algorithm that groups applications into classes and that shares the available capacity between them according to a percentage of the bandwidth, which is defined by the network administrator. Each DRR queue is given its fair share of the bandwidth during times of congestion, but all of them can use as much of the bandwidth as needed when there is no congestion.

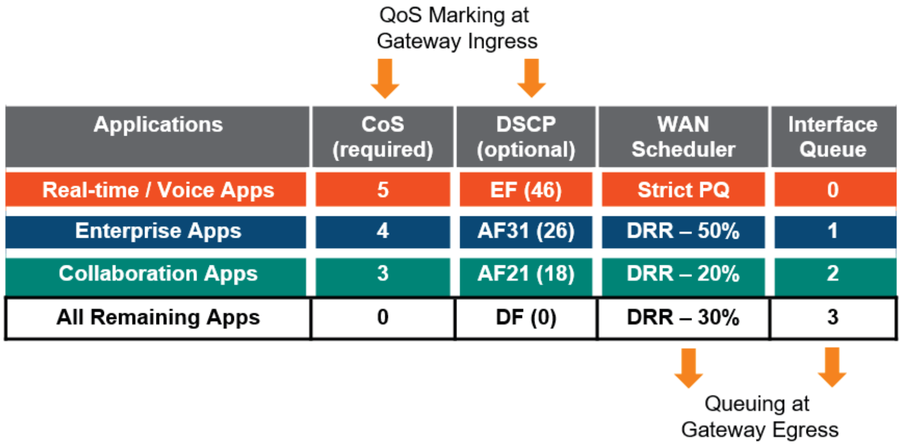
QOS FOR ARUBA GATEWAYS

One of the foundational elements of the Aruba SD-WAN solution is to provide a transport-independent WAN overlay network to ensure reachability regardless of the WAN circuit being used. This means your applications could be going through the Internet, where QoS is not honored, or they could also be going through a SP-managed WAN, where QoS is honored. The dynamic path selection policies take care of the optimal path selection, but QoS marking ensures business-critical or latency-sensitive applications get prioritized as they leave the gateway and pass through the SP-managed network.

Figure 2 shows an example of how you can mark applications in an Aruba gateway. The outbound interface requires the CoS values shown in the second column in order to queue applications through the four hardware queues. The optional DSCP values in the third column are used in the SP-managed networks that honor QoS markings. Service providers often accommodate at least four-classes in their MPLS network, which is a good reason to mark your applications according to the DSCP values agreed upon in your service level agreement.

The weighted values used in the DRR WAN scheduler queues need to be adjusted according to the volume of applications in each category on your network. This adjustment process is often done with trial and error as you learn how the QoS policy is affecting the applications in your environment.

Figure 2   Example QoS policy for Aruba gateways


QoS Marking at Gateway Ingress
Queuing at Gateway Egress

Applications	CoS (required)	DSCP (optional)	WAN Scheduler	Interface Queue
Real-time / Voice Apps	5	EF (46)	Strict PQ	0
Enterprise Apps	4	AF31 (26)	DRR – 50%	1
Collaboration Apps	3	AF21 (18)	DRR – 20%	2
All Remaining Apps	0	DF (0)	DRR – 30%	3

Deploying SD-WAN QoS

As your business needs evolve, so do the demands on QoS technologies to not only prioritize applications but to do it as simply as possible. The need to protect real-time and critical business applications remains extremely important on the WAN because access speeds are much lower than the LAN networks that feed them. However, in the world of SD-WAN networks with multiple options for active/active WAN links, it is not as important to identify all applications as it is to simply identify the ones that are critical to running your business.

Procedures

Aruba Central uses a two-level hierarchy for configuration tasks. A device's final configuration is a result of configuration that is applied at the group level, along with configuration that is applied at a device level. Parameters added at the device level override the configuration performed at the group level. Aruba recommends performing the bulk of the configuration at the group level and using device-level configurations only when specific overrides are needed.

The applications entering a gateway must be classified and marked on the ingress interface per user-role, subnet, or VLAN. This allows you to differentiate between voice applications for employees that you want to protect versus voice applications for guests which you want to give a less-than best effort service. The easiest way to classify large amounts of traffic is with the deep packet inspection (DPI) engine's application categories. Each of the 3000+ DPI applications identified are given an application category to allow you to group applications with similar characteristics for queuing on the outbound interface.

1.1 Configure Employee QoS Policy at the Group Level

In this procedure, you configure the employee QoS policy for the gateway at the group level.

The following table lists an example of combining a couple of the DPI application categories into the QoS markings used by the SP and the gateway. The CoS markings are required for the gateway hardware queuing. However, the DSCP markings are needed only if your MPLS provider offers prioritization for your traffic. You can easily add additional applications by using the techniques in this procedure.

The "All other" entry at the end of the QoS policy marks all application flows that are not recognized by the DPI engine into the best effort queue. This prevents end users who mark their own packets from getting higher priority access across your WAN.

Note The DSCP value in the tunnel header must match what your MPLS provider is expecting. The values shown below are examples of common settings in a 4-class MPLS environment.

Table 1 Example application category DSCP and CoS mapping for employees

Category group	Application category	DSCP	CoS
Real-time	Unified-communications	46	5
Transactional	Enterprise-apps	26	4
Collaboration	Collaboration	18	3
All other	Any	0	0

For an example of an application category mapping with all the Aruba categories, see "Application Category Mapping" in Appendix B.

Step 1: Log in to Aruba Central, and then in the upper left, in the Current App section, select the **Gateway Management** application.

Step 2: At the top of the page, click the **Filter Gateway Management** list, and then from the **Groups** list, select the gateway group.

Step 3: Navigate to **Security** > **Policies**, and at the bottom of the Policies table, click **+.**

Step 4: In the Add Policy window, enter your information, and then click **Save Settings**.

- Policy type—**Session** (list)

- Policy name—**corp-users-qos**

Step 5: In the Policies table, select the policy name you just created, and at the bottom of the Policy > [policy-name] table, click **+**.

Step 6: In the [policy-name] > New Application Rule table, enter your information, and then click **Save Settings**.

- Service/app—**App category** (list)
- App category—**unified-communications** (list)
- Action—**Permit** (list)
- DSCP—**46** (optional if your MPLS carrier honors DSCP markings)
- 802.1p priority—**5** (list)

Step 7: For each entry in the application category table above, repeat the two previous steps.

1.2 Configure Guest QoS Policy at the Group Level

In this procedure, you configure the guest QoS policy for the gateway at the group level.

All guest applications should be marked with a less-than best effort class to prevent them from interfering with the employee traffic on your network. The DSCP marking is not used on the Internet paths for guest applications, but the CoS marking puts the traffic into the default DRR queue with a limited amount of available bandwidth during times of congestion.

Table 2 Example application category DSCP and CoS mapping for guests

Category Group	Application Category	DSCP	CoS
All	Any	0	1

Step 1: In the Add Policy window, enter your information, and then click **Save Settings**.

- Policy type—**Session** (list)
- Policy name—**guest-users-qos**

Step 2: In the Policies table, select the policy name you just created, and at the bottom of the Policy > [policy-name] table, click **+**.

Step 3: In the [policy-name] > New Application Rule table, enter your information, and then click

Step 4: **Save Settings**.

- Service/app—**Any** (list)
- Action—**Permit** (list)
- DSCP—**0** (remark all guest traffic to default)
- 802.1p priority—**1** (list)

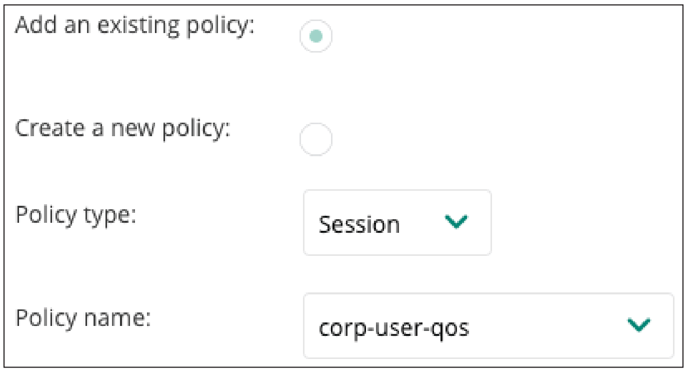
1.3 Apply QoS Policy to User-Role at the Group Level

In this procedure, you apply the QoS policy to an existing user-role at the group level. This user-role is assigned to all users as their initial role on a particular ingress LAN VLAN of the gateway.

Step 1: Navigate to **Security > Roles**, in the Roles table, click the existing user role, and then at the bottom of the <role name> table, click **+**.

Step 2: In the Add Policy dialog box, enter your information, and then click **Save Settings**.

- Add an existing policy—**Yes** (radio button)
- Policy type—**Session** (list)
- Policy name—**copr-user-qos** (list for previously created QoS policy)



Add an existing policy: ●

Create a new policy: ○

Policy type: Session

Policy name: corp-user-qos

Step 3: For each additional user-role and accompanying QoS policy, repeat this procedure.

1.4 Configure WAN Scheduler at the Group Level

In this procedure you, configure the WAN scheduler for the gateway at the group level.

After the applications have been marked with CoS, you assign them to the four hardware queues on the gateway. The table below shows how to map the CoS values into the interface queues. The DRR bandwidth percentage values are examples. For the best results in your environment, you should adjust the values to match your anticipated traffic volumes.

Table 3 Example CoS to interface queue mapping

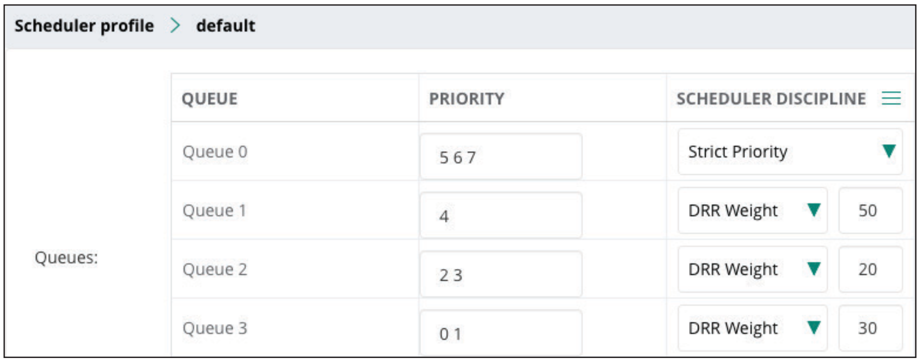
Application type	CoS	Scheduler discipline	Interface queue
Real-time	5	Strict PQ	0
Transactional	4	DRR - 50%	1
Collaboration	3	DRR - 20%	2
Best effort	0	DRR - 30%	3

Note The DRR percentage values must equal 100%.

Step 1: Navigate to **WAN > WAN Scheduler**, and then in the WAN scheduler profiles table click the **default** profile.

Step 2: In the Scheduler profile > [profile name] box, enter your information, and then click **Save Settings**.

- Queue 0 Priority—**5 6 7**
- Queue 0 Scheduler Discipline—**Strict Priority** (list)
- Queue 1 Priority—**4**
- Queue 1 Scheduler Discipline—**DRR Weight 50**
- Queue 2 Priority—**2 3**
- Queue 2 Scheduler Discipline—**DRR Weight 20**
- Queue 3 Priority—**0 1**
- Queue 3 Scheduler Discipline—**DRR Weight 30**



Scheduler profile > default

Queues:

QUEUE	PRIORITY	SCHEDULER DISCIPLINE	
Queue 0	5 6 7	Strict Priority	
Queue 1	4	DRR Weight	50
Queue 2	2 3	DRR Weight	20
Queue 3	0 1	DRR Weight	30

Procedures

Configuring Gateway QoS at the Device Level

Aruba Central uses a two-level hierarchy for configuration tasks. A device's final configuration is a result of configuration that is applied at the group level, along with configuration that is applied at a device level. Parameters added at the device level override the configuration performed at the group level. Aruba recommends performing the bulk of the configuration at the group level and using device-level configurations only when specific overrides are needed.

After you configure the QoS policy at the group level, it is time to configure traffic shaping for each gateway interface. The amount of outbound WAN bandwidth can be different for every interface on your gateways based on the specific service offerings at your branch locations. Although it is possible to configure the transmit rate at the group level if all your gateways have the exact same WAN bandwidth, in most cases it must be done at the device level.

2.1 Configure WAN Transmit Rate at the Device Level

In this procedure, you configure the WAN transmit rate of the gateway at the device level.

By default, the gateway delivers traffic to the WAN at the native speed of the interface. To prevent the gateway from overrunning the provider's service rate, a traffic shaping policy must be applied to each interface. Because Internet services are often asymmetrical, the transmit rate is defined as the upload or outbound bandwidth you purchased for the WAN service at each location.

Step 1: Log in to Aruba Central, and then in the upper left, in the Current App section, select the **Gateway Management** application.

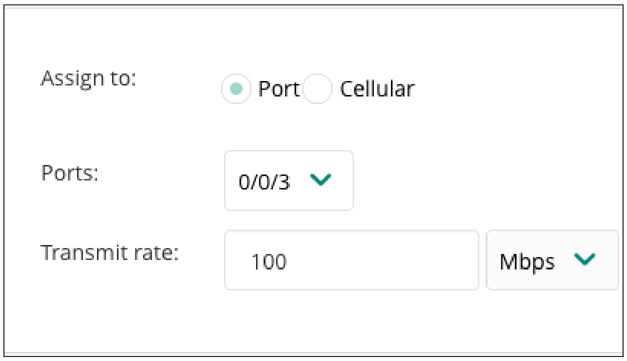
Step 2: At the top of the page, click the **Filter Gateway Management** list, and then from the **Gateways** list, select the individual gateway.

Step 3: Navigate to **WAN > WAN Scheduler**, and then in the WAN scheduler profiles table click the **default** profile.

Step 4: In the Scheduler profile section, at the bottom of the Assignments table, click **+**.

Step 5: In the Assign to page, enter your information, click **OK**, and then click **Save Settings**.

- Assign to—**Port** (radio button)
- Ports—**0/0/3** (list)
- Transmit Rate—**100 Mbps** (outbound service rate)



Assign to: ● Port ○ Cellular

Ports: 0/0/3

Transmit rate: 100 Mbps

Step 6: For each additional WAN interface, repeat the previous two steps.

Step 7: For each additional gateway, repeat this procedure.

Appendix A: QoS Design

A comprehensive QoS policy requires you to identify as much of the applications on your network as possible. After applications are identified, they must be marked using one of the techniques described below.

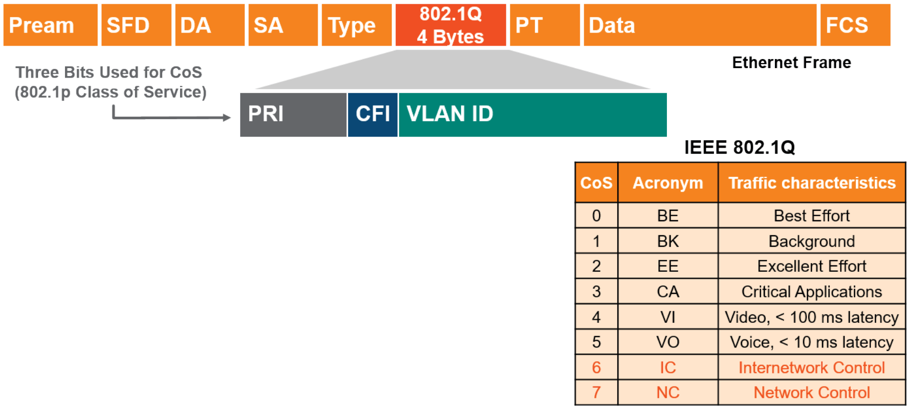
COMPREHENSIVE QOS POLICY

Scheduling algorithms rely on classification markings to identify applications as they pass through a network device. Aruba recommends marking your applications with CoS for queueing in the gateway and, optionally, DSCP if your service provider honors the markings in their managed network. The goal of the QoS policy is to allow your critical applications to share the available WAN bandwidth with minimal system impact and engineering effort.

CoS Marking

CoS marking uses the IEEE 802.1Q portion of the Ethernet header, which contains the 802.1p user priority field (CoS). CoS markings persist only until the next layer-3 device in the network is reached, but they are also used inside the Aruba gateway to categorize applications into the four hardware queues. The 3-bit CoS field allows for 8 values from 0 to 7 as shown in the diagram below.

Figure 3 *CoS marking*



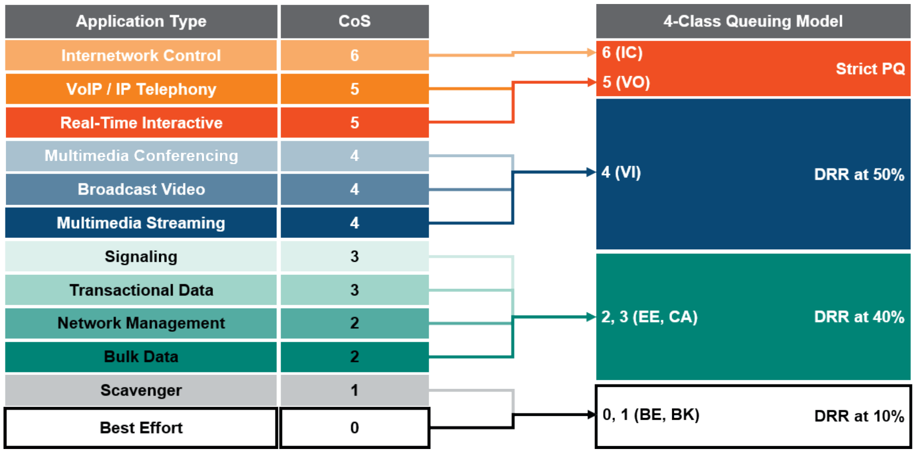
CoS	Acronym	Traffic characteristics
0	BE	Best Effort
1	BK	Background
2	EE	Excellent Effort
3	CA	Critical Applications
4	VI	Video, < 100 ms latency
5	VO	Voice, < 10 ms latency
6	IC	Internetwork Control
7	NC	Network Control

Voice (VO) and real-time applications are placed into a strict priority queue (PQ), along with Internetwork Control (IC) traffic. Voice and real-time applications require an external call admission control agent to prevent them from using all the available bandwidth on an egress interface. IC traffic is sourced from the gateway itself and consists of a very small amount of bandwidth when compared to the other flows on an egress interface.

Video (VI) and streaming applications are placed into their own DRR queue using an appropriate bandwidth percentage for the volume on your network. Signaling, network management, transactional and bulk applications are also placed into a DRR bandwidth percentage queue for the critical applications (CA) and excellent effort (EE) classes. Finally, default and scavenger applications are placed into their own DRR bandwidth percentage queue for the best effort (BE) and background (BK) classes.

Figure 4 shows an example 8-class CoS to 4-class queuing model. The DRR percentage values are examples that you should change to account for the anticipated traffic volume in your network.

Figure 4 Example 8-class CoS to 4-class queuing model



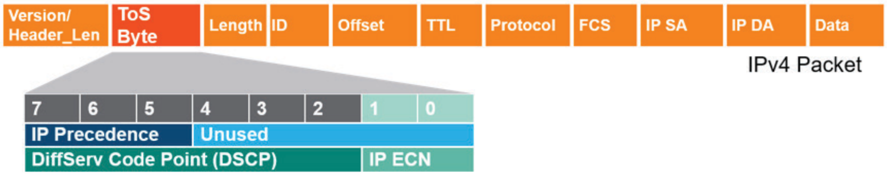
Application Type	CoS
Internetwork Control	6
VoIP / IP Telephony	5
Real-Time Interactive	5
Multimedia Conferencing	4
Broadcast Video	4
Multimedia Streaming	4
Signaling	3
Transactional Data	3
Network Management	2
Bulk Data	2
Scavenger	1
Best Effort	0

4-Class Queuing Model	
6 (IC) 5 (VO)	Strict PQ
4 (VI)	DRR at 50%
2, 3 (EE, CA)	DRR at 40%
0, 1 (BE, BK)	DRR at 10%

Layer-3 Marking

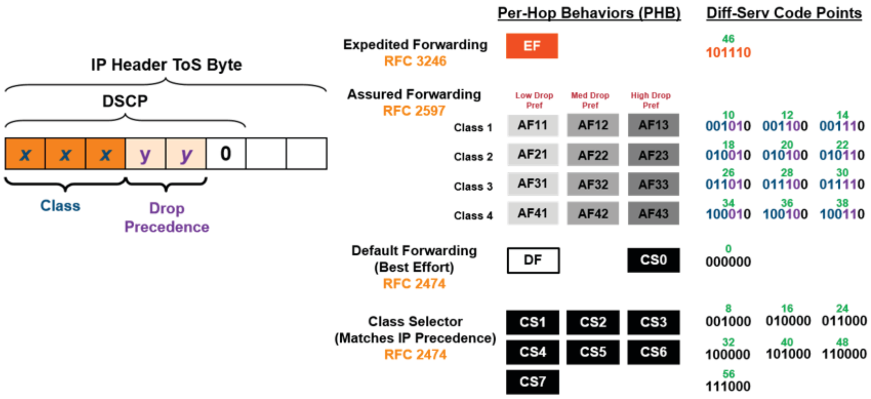
Layer-3 marking uses the IP type of service (ToS) byte with either the IP Precedence three most-significant bit values from 0 to 7 or the DSCP six most-significant bit values from 0 to 63. The DSCP values are more common because they provide a higher level of QoS granularity, but they are also backward-compatible to IP precedence because of their left-most placement in the ToS byte. Layer-3 markings are added in the standards-based IP header, so they remain with the packet as it travels across the network. When an additional IP header is added to a packet, like in the case of traffic in an IPSec tunnel, the inner header DSCP marking must be copied to the outer header to allow the SP to use the values.

Figure 5 Layer-3 marking



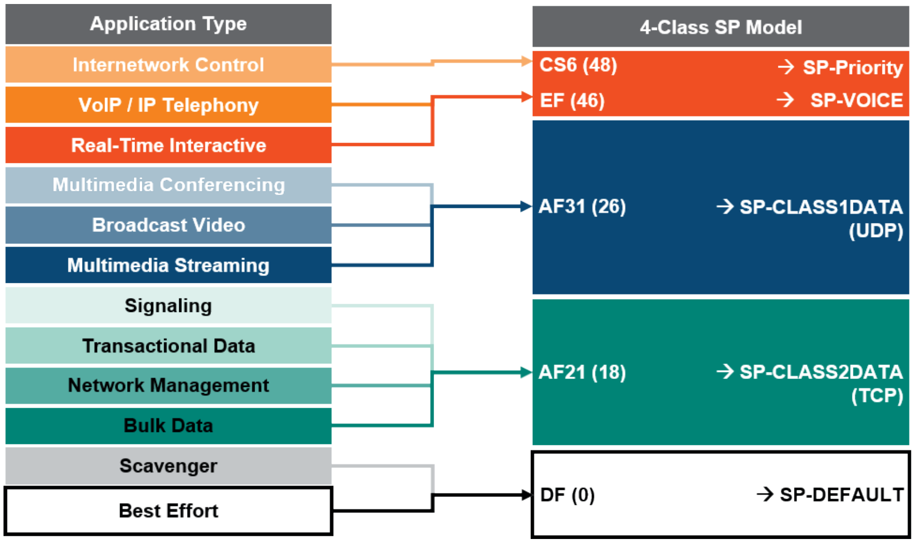
There are several RFCs associated with the DSCP values as they pertain to the per-hop behaviors (PHBs) of traffic as it passes through the various network devices along its path. Figure 6 shows the relationship between PHB and DSCP, along with their associated RFCs.

Figure 6 DSCP relationship with per-hop behaviors



As with the CoS markings discussed above, voice is marked with the highest priority using an Expedited Forwarding (EF) class. IC traffic is marked with a Class Selector (CS6) class for inclusion into the service providers special priority queues across their network. Multimedia applications, broadcast and real-time video are placed into an assured forwarding (AF31) class to give them a percentage of the available bandwidth as they cross the providers network. Signaling, network management, transactional and bulk applications are given an assured forwarding (AF21) class. Finally, default and scavenger applications are placed into the Default (DF) class to give them a reduced amount of bandwidth but not completely starve them during times of interface congestion. Figure 7 shows an example of application-type-to-4-class service provider mapping with DSCP.

Figure 7 Example application type to 4-class SP model with DSCP



The DSCP markings are done in the ToS byte of the IP header. When using SD-WAN, the IP packet is encapsulated in an IPSec outer header when the packets are sent into the overlay tunnel over the WAN interface. To accommodate the additional header, the Aruba gateway copies the DSCP from the inner header to the outer header in order to influence the per-hop behavior within the service provider network, as shown in Figure 8.

Figure 8 DSCP copied from inner IP header to IPSec tunnel header

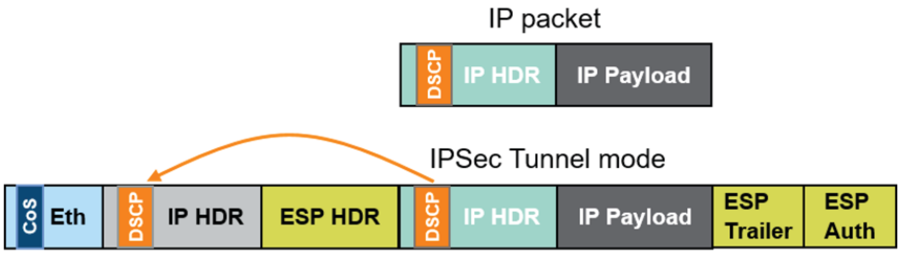
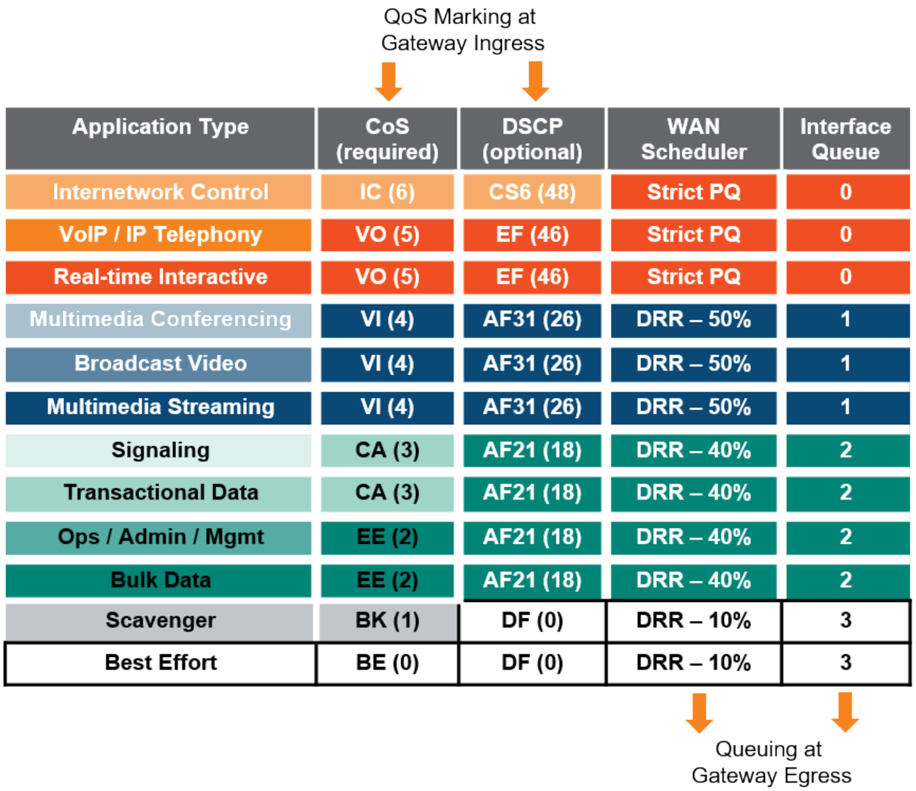


Figure 9 shows an example of how you can mark applications in an Aruba gateway. The outbound interface requires the CoS values shown in the second column in order to queue applications through the four hardware queues. The optional DSCP values in the third column are used in the SP-managed networks that honor QoS markings. Service providers often accommodate at least four-classes in their MPLS network, which is a good reason to mark your applications according to the DSCP values agreed upon in your service level agreement.

The weighted values used in the DRR WAN scheduler queues need to be adjusted according to the traffic volume of each category in your network. This process is often done with trial and error as you learn how the QoS policy is working in your environment.

Figure 9 Example QoS Policy for Aruba Gateways



QoS Marking at Gateway Ingress

Application Type	CoS (required)	DSCP (optional)	WAN Scheduler	Interface Queue
Internetwork Control	IC (6)	CS6 (48)	Strict PQ	0
VoIP / IP Telephony	VO (5)	EF (46)	Strict PQ	0
Real-time Interactive	VO (5)	EF (46)	Strict PQ	0
Multimedia Conferencing	VI (4)	AF31 (26)	DRR – 50%	1
Broadcast Video	VI (4)	AF31 (26)	DRR – 50%	1
Multimedia Streaming	VI (4)	AF31 (26)	DRR – 50%	1
Signaling	CA (3)	AF21 (18)	DRR – 40%	2
Transactional Data	CA (3)	AF21 (18)	DRR – 40%	2
Ops / Admin / Mgmt	EE (2)	AF21 (18)	DRR – 40%	2
Bulk Data	EE (2)	AF21 (18)	DRR – 40%	2
Scavenger	BK (1)	DF (0)	DRR – 10%	3
Best Effort	BE (0)	DF (0)	DRR – 10%	3

Queuing at Gateway Egress

Appendix B: QoS Deployment

A comprehensive QoS policy requires you to categorize the business relevant and scavenger class applications on your network. Using the Aruba DPI engine, the applications are grouped together into categories to help you identify the ones that have similar characteristics. After you sort the applications that are important for your business from the ones you do not care about, you combine them into groups for your queuing and marking policies.

APPLICATION CATEGORY MAPPING

This is an example showing all the Aruba application categories mapped to DSCP and CoS for employees.

Most managed-SPs have a special network control queue that they do not declare as part of their customer-facing model for internetwork control traffic. This traffic is marked with CS6 (48) to place it into the network control queue in the SP network. Because the network elements use IC traffic to ensure stability under congestion and when a device is oversubscribed, CS6 traffic must be preserved.

The “All other” entry at the end of the QoS policy marks all application flows that are not recognized by the DPI engine into the best effort queue. This prevents end users who mark their own packets from getting higher priority access across your WAN.

Note The DSCP value in the tunnel header must match what your MPLS provider is expecting. The values shown below are examples of common settings in a 4-class MPLS environment.

Table 4   Example application category DSCP and CoS mapping for employees

Category group	Application category	DSCP	CoS
Real-time	Unified-communications	46	5
Internetwork control	Network-service	48	6
	Tunneling	48	6
Streaming	Streaming	26	4
Transactional data	Antivirus	18	3
	Authentication	18	3
	Behavioral	18	3
	Collaboration	18	3
	Encrypted	18	3
	Enterprise-apps	18	3
	Instant-messaging	18	3
	Mobile	18	3
	Peer-to-peer	18	3
	Thin-client	18	3
Bulk Data	Cloud-file-storage	18	2
	IM-file-transfer	18	2
	Mail-protocols	18	2
	Webmail	18	2
Best Effort	Standard	0	0
	Web	0	0
Scavenger (do not care about)	Gaming	0	1
	Mobile-app-store	0	1
	Social-networking	0	1
All other	Any	0	0

What's New in This Version

The following changes have been made since Aruba last published this guide:

- This is a new supplemental guide.

You can use the feedback form to send suggestions and comments about this guide.

aruba

a Hewlett Packard
Enterprise company

B-000610A-19A-1 06/19
a00074797enw