CHAPTER 5

Quality of Service and Bandwidth Management

Cisco Unified Border Element (SP Edition) distributed model for the Cisco ASR 1000 Series Routers provides Quality of Service (QoS) and bandwidth management features to assure quality end-to-end connection for real-time voice, video, and multimedia traffic. The packet marked for higher priority is delivered faster than non-prioritized packets. The data border element (DBE) supports statistics collection and saves all QoS statistics, including packets transmitted per second and packets dropped for exceeding allocated bandwidth on a per-call or per-interface basis. The Cisco ASR 1000 Series Routers support QoS functions such as Low Latency Queueing (LLQ), Class-Based Weighted Fair Queueing (CBWFQ), and shaping at the subinterface level.

The DBE has different packages to enhance QoS and these packages are described in this chapter. For a complete description of commands used in this chapter, refer to the *Cisco Unified Border Element (SP Edition) Command Reference: Distributed Model* at http://www.cisco.com/en/US/docs/ios/sbc/command/reference/sbc_book.html.

Cisco Unified Border Element (SP Edition) was formerly known as Integrated Session Border Controller and may be commonly referred to in this document as the session border controller (SBC).

Contents

H.248 Traffic Management Package Support

The DBE supports the sustained data rate (tman/sdr), maximum burst size (tman/mbs), and policing (tman/pol) properties of the ETSI TS 102 333 Traffic Management (Tman) package.[1] Support of these tman properties allows additional pinhole programming in the Tman package to inform the DBE how to police media and signaling flows. These tman properties can be assigned to both media and signaling flows.

1. ETSI TS 102 333 version 1.1.2 Traffic Management Package

The DBE performs asymmetric policing. Asymmetric policing allows the MGC to impose different flow policing on traffic traveling in each direction on the same stream. For example, traffic traveling from the subscriber side to the DBE can be policed independently of traffic from the network core to the DBE.

Asymmetric policing is accomplished by allowing the tman/pol property to be specified separately for the two sides of a gate, which typically might be the access (subscriber) side and the backbone side. The tman/pol property can be specified as ON, OFF, or Absent on either the access side or the backbone side for either a media flow or signaling flow. Once tman/pol is specified as ON and both the tman/sdr and tman/mbs properties are present, the DBE polices traffic based on the values of the tman/sdr and tman/mbs parameters.

The supported tman properties have the following functions:

- The tman/sdr property defines the sustainable data rate in bytes per second that is permitted for the stream. It has a numerical value.
- The tman/mbs property defines maximum burst size in bytes for the stream. It has a numerical value.
- The tman/pol (policing) property can be set to ON or OFF or Absent.
  - When the tman/pol property is set to ON, policing is applied at the point of entry for traffic entering the media gateway (MG).

    When both the tman/sdr and tman/mbs properties are present (and the tm/pol property is ON), the DBE polices traffic based on the sdr and mbs Tman parameters.

    However, the absence of both tman/sdr and tman/mbs properties is permissible. In this case, the DBE polices traffic based on the Session Description Protocol (SDP).

Note If either the tman/sdr or tman/mbs property is present, then the other property must be present; that is, both the tman/sdr and tman/mbs properties must be present. In this case, the DBE polices traffic based on the sdr and mbs parameters.

  - When the tman/pol property is set to OFF, no policing is applied to traffic entering the media gateway.
  - If the tman/pol property is Absent, policing is done based on the SDP for the stream for a media flow. No policing is done for a signaling flow.

Note Absent means that the property has not been defined. If no tman properties (tman/pol, tman/sdr, and tman/mbs) have been defined, then the behavior for a media flow is to calculate the required bandwidth from the Session Description Protocol (SDP) in the local descriptor. For a signaling flow, the behavior is to perform no policing.

The tman properties (tman/pol, tman/sdr, and tman/mbs) are defined using Add and Modify requests, and they are returned on subsequent responses to Audit requests.

Note For additional information on RTP and RTCP streams and RTCP policing based on the Tman Package, see the "RTCP Policing" section on page 5-4.

The tman properties have the following caveat:

The DBE issues error 421 indicating "Unknown action or illegal combination of actions" for any programming containing other fields and programming that sets the tman/pol flag, but only specifies one of the tman/sdr or tman/mbs values.

Table 5-1 describes the asymmetric flow policing behavior of the two sides of a gate based on whether the tman/pol property is specified as ON, OFF, or Absent. Each side of the gate behaves independently of the other side. The Access Side might be the subscriber side to the DBE, and the backbone side might be the network core to the DBE.

Table 5-1 Asymmetric Flow Policing—Independent Behavior of Signaling and Media Flows on Two Sides of a Media Gateway

		Access Side (AC)		
	tman/pol Property	**Absent**	**ON**	**OFF**
Back Bone Side (BB)	**Absent**	Signaling: No policing Media: Policing per SDP	Signaling: Policing per Tman parameters on AC and no policing on BB Media: Policing per Tman parameters on AC and per SDP on BB	Signaling: No policing on AC and BB Media: No policing on AC and policing per SDP on BB
	ON	Signaling: Policing per Tman parameters on BB and no policing on AC Media: Policing per Tman parameters on BB and per SDP on AC	Signaling: Policing per Tman parameters on AC and on BB independently Media: Policing per Tman parameters on AC and on BB independently	Signaling: No policing on AC and policing per Tman parameters on BB Media: No policing on AC and policing per Tman parameters on BB
	OFF	Signaling: No policing on BB and AC Media: No policing on BB and policing per SDP on AC	Signaling: No policing on BB and policing per Tman parameters on AC Media: No policing on BB and policing per Tman parameters on AC	Signaling: No policing on BB and AC Media: No policing on BB and AC

DSCP Marking and IP Precedence Marking

The DBE supports marking of differentiated services code point (DSCP) bits and IP precedence marking for egress traffic and media relay. Using standard Router features, these markings can be used to prioritize packets for faster delivery or for lower risk of drop under congestion.

DSCP Re-Markings

For every media stream, the DBE receives a DSCP value to use in the Real-Time Transport Protocol (RTP) and RTP Control Protocol (RTCP) packets. The DBE receives these values at the call setup time on a per-flow basis and maintains the values as a part of the connection table entry. The DBE modifies the type of service (TOS) bits in the IP header for every outgoing packet and updates the checksum accordingly.

QoS Bandwidth Allocation

The DBE supports QoS bandwidth allocation. The DBE has the ability to limit excess traffic beyond the allocated bandwidth by performing session-based policing. For information on the different types of policing performed by the DBE, see the "H.248 Traffic Management Package Support" section on page 5-1, the "RTCP Policing" section on page 5-4, and the "Two-Rate Three-Color Policing and Marking" section on page 5-5.

RTCP Policing

An SBC session may be of two types, media or signaling. The media portion of the call comprises the Real-Time Transport Protocol (RTP) stream and, optionally, a RTP Control Protocol (RTCP) stream. Calls typically have both a RTP and RTCP stream.

Each session may be subject to policing. RTCP streams are not explicitly configured using the H.248 protocol and therefore cannot have their policing parameters set. Instead, the policing parameters for RTCP sessions are derived from the corresponding RTP flows based on the usage of the ETSI TS 102 333 Traffic Management (Tman) package.

For more information on the Tman package, see the "H.248 Traffic Management Package Support" section on page 5-1.

RTCP Policing Using Tman Package

The H.248 Tman package is used to set the sustained data rate (tman/sdr) and maximum burst size (tman/mbs) properties, and the policing type (tman/pol), on the RTP session. The tman/pol property can be specified as one of the following:

- ON—Policing is enabled on the RTP session.
- OFF—No policing is applied to incoming traffic.
- Absent—Property is not defined, but policing is enabled on the RTP session.

When tman/pol is ON, then tman/sdr and tman/mbs set the policing parameters for the RTP session. When tman/pol is OFF, then there is no policing of either RTP or RTCP sessions. When tman/pol is Absent or undefined, then default policing parameters for RTP sessions are derived from the codec embedded in the SDP string.

When policing is enabled on the RTP session, RTCP policing parameters for setting rate limits (sustained data rate) and maximum burst size are derived from RTP parameters by calculating 5 percent of the RTP sdr and 5 percent of the RTP mbs, with a minimum RTCP sdr and mbs. See Table 5-2 on page 5-5 for details on how policing parameters for RTP and RTCP sessions are derived.

Table 5-2 Policing Parameters for RTP and RTCP Sessions

	If tman/pol = ON, policing parameters are derived as follows:	If tman/pol = Absent (undefined), policing parameters are derived as follows:	If tman/pol = OFF, policing parameters are derived as follows:
For the RTP session	tman/sdr and tman/mbs set the policing parameters.	Default policing parameters are derived from the codec embedded in the SDP string.	There is no policing of RTP sessions.
For the RTCP session	• The sustained data rate for RTCP policing is 5 percent of the sustained data rate for the RTP session, or a minimum rate of 208 bytes per second, whichever is the greater. • The maximum burst size for RTCP policing is 5 percent of the RTP maximum burst size, or a minimum value of 1500 bytes, whichever is the greater.	• The sustained data rate for RTCP policing is 5 percent of the sustained data rate for the RTP session, or 208 bytes per second, whichever is the greater. • The maximum burst size for RTCP policing is 5 percent of the RTP maximum burst size, or 1500 bytes, whichever is the greater.	There is no policing of RTCP sessions.

RTCP Policing Not Using Tman Package

RTCP policing can be implemented by not setting any of the properties of the H.248 Tman package.

When the tman/sdr and tman/mbs properties are not specified, the RTCP rate limiting and maximum burst size are set at 5 percent of the codec (embedded in the SDP string) for the RTP stream.

Two-Rate Three-Color Policing and Marking

Traffic policing is a traffic regulation mechanism that is used to limit the rate of traffic streams. Policing allows you to control the maximum rate of traffic sent or received on an interface. When the traffic rate exceeds the configured maximum rate, policing drops or re-marks the excess traffic.

The ETSI TS 102 333 Traffic Management (Tman) package defined a number of properties to allow traffic policing to be explicitly enabled. However, because the current H.248 standard only supported specifying one rate with a traffic flow, the only action available in the H.248 standard for non-conforming packets had been to discard them. (See the "Enhanced Event Notification and Auditing" section on page 11-4).

The Two-Rate Three-Color Policing and Marking feature is an enhancement to how the DBE polices traffic flow by introducing two-rate policing and three-color marking.

The DBE previously supported three tman properties—policing type (pol), sustainable data rate (sdr), and maximum burst size (mbs). In supporting the Two-Rate Three-Color Policing and Marking feature, the DBE uses one additional property of the ETSI TS 102 333 Traffic Management (Tman) package:

peak data rate (pdr)—Defines the peak data rate in bytes per second that is permitted for the stream.

See the "H.248 Traffic Management Package Support" section on page 5-1 for more information on the tman properties.

Enabling Two-Rate Three-Color Policing and Marking

All of the following conditions must occur to enable the Two-Rate Three-Color Policing and Marking feature for a specific flow:

- The DSCP value is provisioned via diffserv package during call setup.
- The sdr and mbs are provisioned via the Tman package during call setup. (The mbs property is used for pdr policing as well and has an assumed minimal value of 1500 bytes.)
- Two DSCP values (control and marker DSCPs) and the pdr coefficient are configured via the **control-dscp marker-dscp pdr-coefficient** CLI.
- The **control-dscp** value configured must match the diffserv DSCP value for a specific flow to enable the Two-Rate Three-Color Policing and Marking feature.

If any one of the conditions is not met, this feature is not enabled for the flow.

Implementing Two-Rate Three-Color Policing and Marking

In the Two-Rate Three-Color Policing and Marking feature, only two rates—sdr and pdr—allow traffic to be policed into three categories of traffic, which are handled as follows:

- Traffic conforming to both sdr and pdr.

  These packets are colored using the DSCP value provisioned via the H.248 diffserv package; that is, H.248 passes the DSCP value. These packets are forwarded and the DSCP value comes from the H.248 diffserv package.

- Traffic not conforming to the lower sdr rate, but conforming to the higher pdr rate.

  These packets are colored with the marker DSCP value and pdr configured with the **control-dscp** *value1* **marker-dscp** *value2* **pdr-coefficient** *value3* command.

  - The **control-dscp** *value1* keyword enables the Two-Rate Three-Color Policing and Marking feature for a specific flow if *value1* matches the DSCP value for the flow in the diffserv package transmitted via H.248.
  - The **marker-dscp** *value2* keyword colors traffic packets with a DSCP *value2*. This traffic conforms to the peak data rate (pdr), but does not conform to the sustainable data rate (sdr).
  - The **pdr-coefficient** *value3* keyword applies the following formula to calculate the pdr value (which is not passed from H.248).

    The pdr-coefficient value is calculated as pdr = sdr * *value3* /100 (and pdr must be greater than sdr).

- Traffic not conforming to either of the sdr or pdr rates.

  These packets are dropped.

Note The Two-Rate Three-Color Policing and Marking feature is not enabled for a particular flow if the DSCP value set by H.248 for that flow does not match the configured **control-dscp** *value1*.

Traffic flows must have the Two-Rate Three-Color Policing and Marking feature enabled to use the three parameters configurable with the Two-Rate Three-Color CLI. Traffic flows that do not have the Two-Rate Three-Color Policing and Marking feature enabled are subject to normal Tman-based policing with pdr, sdr, and mbs parameters configured via H.248, regardless of the pdr coefficient configured via the Two-Rate Three-Color CLI.

The DBE supports dual token bucket policing to support this feature. The DBE uses a "token bucket algorithm" to manage the maximum rate of traffic. This algorithm is used to define the maximum rate of traffic allowed on an interface at a given moment in time. The token bucket algorithm processes as follows—each arriving packet's size (frame's size) is subtracted from the contents of the bucket. If the bucket does not have enough tokens for the arriving packet, the packet is dropped and no tokens are removed. The passage of time fills the bucket with tokens and the dispatching of a packet depletes the bucket.

DBE Restrictions

The following are DBE restrictions for the Two-Rate Three-Color Policing and Marking feature:

- The DSCP values configured via CLI are global. Specifically, the DSCP values are shared among all the terminations when this feature is enabled.
- The Two-Rate Three-Color CLI is only applicable to future new call flows and do not trigger any backtrack to established terminations.
- A media gateway controller (MGC), also called the SBE, is only able to use the Tman fields if the DBE supports the various features which make up Tman support.

Related Commands

The **control-dscp marker-dscp pdr-coefficient** command enables the Two-Rate Three-Color Policing and Marking feature, and configures differentiated services code point (DSCP) values and the peak data rate (pdr) coefficient for the feature on the data border element (DBE) for each affected flow.

The **show sbc dbe forwarder-stats** command output entries are added to report statistics of colored traffic.

For a description of the commands used, see the *Cisco Unified Border Element (SP Edition) Command Reference: Distributed Model* at http://www.cisco.com/en/US/docs/ios/sbc/command/reference/sbc_book.html.