CHAPTER 5

WAN QoS

Last revised on: October 30, 2009

This chapter addresses quality of service (QoS) requirements for implementations of H.323 videoconferencing solutions over the enterprise WAN. By applying the prerequisite tools, you can achieve excellent video, voice, and data transmissions over an IP WAN, irrespective of media and even low data rates.

What's New in This Chapter

Table 5-1 lists the topics that are new in this chapter or that have changed significantly from previous releases of this document.

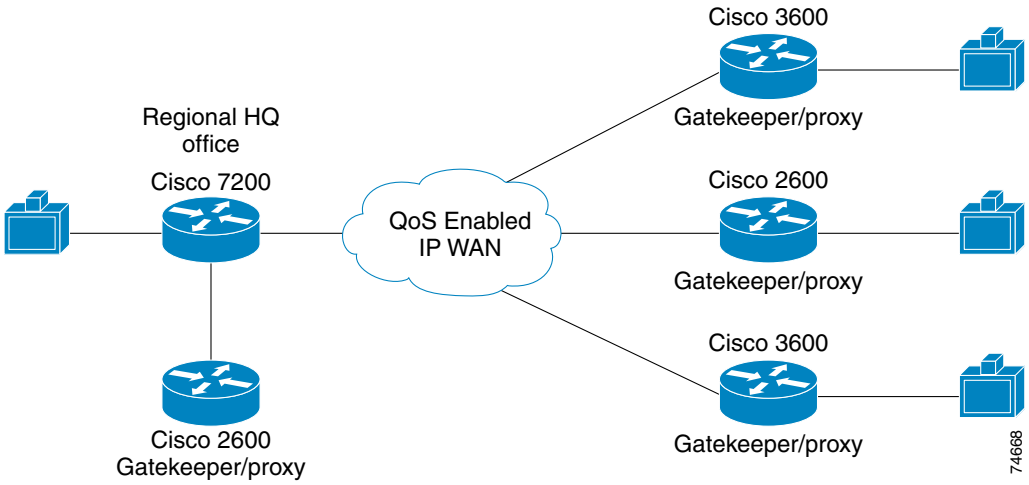
Table 5-1 New or Changed Information Since the Previous Release of This Document

New or Revised Topic	Described in:
Resource Reservation Protocol (RSVP)	Resource Reservation Protocol (RSVP), page 5-6

WAN QoS Model

Figure 5-1 illustrates the typical hub-and-spoke topology of the enterprise WAN model described in this chapter.

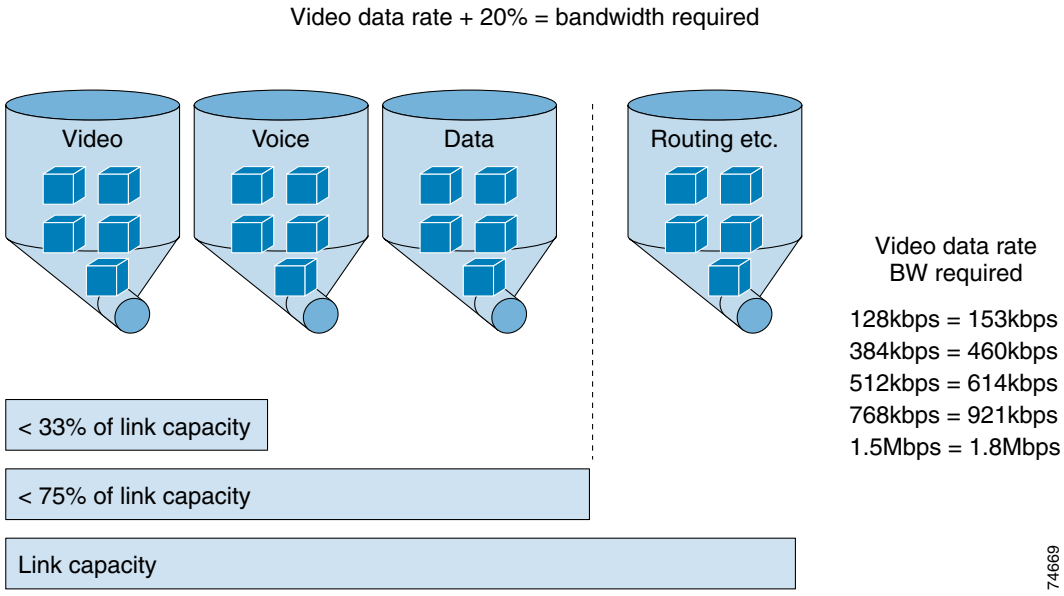
Figure 5-1 Enterprise WAN Model



Capacity Planning

Before placing video traffic on a network, ensure that adequate bandwidth exists for all required applications. First, calculate the minimum bandwidth requirements for each major application (for example, voice, video, and data). This sum represents the minimum bandwidth requirement for any given link, and it should consume no more than 75% of the total bandwidth available on that link. This 75% rule assumes that some bandwidth is required for overhead traffic, such as routing and Layer 2 keepalives, as well as additional applications such as email and HyperText Transfer Protocol (HTTP) traffic. Figure 5-2 illustrates capacity planning on a converged network.

Figure 5-2 Capacity Planning on a Data, Voice, and Video Network



QoS Tools

This section discusses the tools used to implement QoS for H.323 videoconferencing over an enterprise WAN. These tools include:

- Traffic Classification, page 5-3
- Border Element Usage, page 5-4
- Traffic Prioritization, page 5-4

This section concludes with a summary of best practices for each of the applicable data link protocols.

Traffic Classification

Before traffic can be handled according to its unique requirements, it must be identified or labeled. There are numerous classification techniques, including Layer 3 schemes such as IP Precedence or Differentiated Services Code Point (DSCP).

In many cases, traffic classification is done at the edge of the network by the video terminal or an Ethernet switch such as the Catalyst 6000. In these cases, the trust boundary is extended to the edge of the enterprise network and resides in the access or distribution layer. For a more detailed discussion of trust boundaries, see Trust Boundaries, page 3-5.

In some cases, however, the ability to classify and define a trust boundary at the edge of the network might not exist, such as in a branch with Ethernet switches and video endpoints that cannot classify traffic. In this situation, you can implement the trust boundary and classification on the router itself by using ACL entries for small sites without a gatekeeper or by using the Border Element in larger branch sites that contain a gatekeeper.

Border Element Usage

In the multi-zone WAN model, Cisco recommends that you use the Border Element whenever possible. The Border Element allows the classification or reclassification of video streams with IP Precedence or Resource Reservation Protocol (RSVP). The Border Element also provides a single access point for the priority queue to keep unauthorized video streams from oversubscribing the priority queue. Video terminals must be registered with the gatekeeper to obtain access to the Border Element. The gatekeeper is configured for a maximum video bandwidth allowed outside its local zone. This maximum bandwidth should match the amount of bandwidth provisioned for the priority queue to ensure proper queuing functionality.

Traffic Prioritization

In choosing from among the many available prioritization schemes, the major factors to consider include the type of traffic being put on the network and the wide area media being traversed. For multi-service traffic over an IP WAN, Cisco recommends low-latency queuing for the WAN. This allows up to 64 traffic classes, with the ability to use multiple queues for different traffic types, such as priority queuing behavior for videoconferencing and voice, a minimum bandwidth for Systems Network Architecture (SNA) data and market data feeds, and weighted fair queuing for other types of traffic.

Figure 5-3 shows this prioritization scheme as follows:

- Video traffic is placed into a queue with priority queuing (PQ) capabilities and is allocated a bandwidth of 460 kbps. The entrance criterion for this queue could be any video stream received from the specific IP address of a Border Element or any traffic with IP Precedence set to 4. Traffic in excess of 460 kbps would be dropped if the interface becomes congested. Therefore, an admission control mechanism (such as gatekeeper bandwidth limits) must be used to ensure that this limit is not exceeded.
- SNA traffic is placed into a queue that has a specified bandwidth of 56 kbps. Queuing operation within this class is first-in-first-out (FIFO) with a maximum allocated bandwidth of 56 kbps. Traffic in this class that exceeds 56 kbps is placed in the default queue. The entrance criterion for this queue could be Transmission Control Protocol (TCP) port numbers, Layer 3 address, IP Precedence, or DSCP.
- All remaining traffic can be placed in a default queue. If you specify a bandwidth, the queuing operation is FIFO. Alternatively, if you specify the keyword **fair**, the queuing operation is weighted fair queuing (WFQ).

Figure 5-3 illustrates optimized queuing for videoconferencing on the WAN.

Figure 5-3 Optimized Queuing

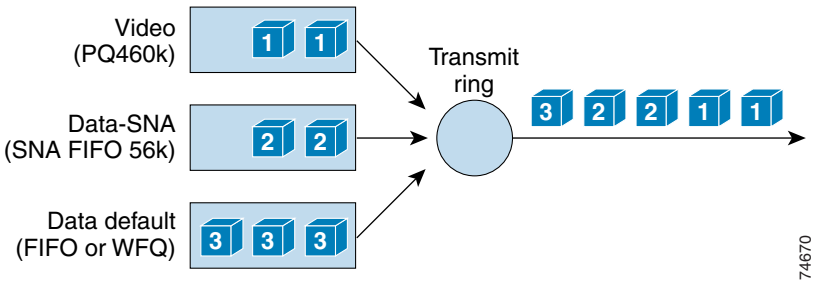


Table 5-2 gives the minimum bandwidth requirements for video and data networks. Note these values are minimum, and any network should be engineered with adequate capacity for all the applications that will use it.

Table 5-2 Minimum Bandwidth Requirements

Traffic Type	Leased Lines	Frame Relay	ATM	ATM Over Frame Relay
Video + Data Maximum video data rates up to 384 kbps	768 kbps	768 kbps	768 kbps	768 kbps
Video + Data Maximum video data rates > 384 kbps	1.544 Mbps	1.544 Mbps	1.544 Mbps	1.544 Mbps

Best Practices

Compressed Real-time Transport Protocol (cRPT) is not recommended for use with IP videoconferencing. Best practices for cRTP are as follows:

- Use cRTP only with low bit rate voice codecs such as G.729. If G.711 is used as the audio codec for a voice or videoconferencing call, the statistical throughput gains achieved with cRTP are not significant enough to merit its use.
- Use cRTP only when low bit rate voice is a significant percentage of the offered load. In general, this feature is beneficial only when low bit rate voice is greater than 30% of the offered load to a circuit.
- cRTP can affect forwarding performance, and Cisco recommends that you monitor CPU utilization when this feature is enabled.

Call Admission Control

Call admission control, or bandwidth control, is required to ensure that the network resources are not oversubscribed. Calls that exceed the specified bandwidth limit are rejected to ensure video quality.

The following three methods can be used to provide call admission control for video calls over the WAN.

Limiting the Number of Video Terminals

Limiting the number of video terminals for call admission control is necessary only in the single-zone WAN model. With no gatekeeper at the remote sites in this model, the only way to control the amount of bandwidth used for video across the WAN is to limit the number of video terminals at the remote sites. The priority queue at each site must then be provisioned for the maximum possible data rate of all the video endpoints at any given site. See Single-Zone WAN, page 4-2, for more information on this call admission control scheme.

Gatekeeper Call Admission Control

This method of call admission control is available only in the multi-zone WAN model. The gatekeeper allows administrators to set bandwidth limits for inter-zone calls, intra-zone calls, or sessions. This scheme allows administrators to set an inter-zone or remote bandwidth limit, provision a priority queue for the same amount of bandwidth, and ensure the integrity of that queue. Currently, gatekeeper call admission control is limited to hub-and-spoke configurations. See Multi-Zone WAN, page 4-4, for more information on this call admission control scheme.

Resource Reservation Protocol (RSVP)

Enterprises that deploy large-scale IP videoconferencing networks using Cisco Unified Communications solutions based on Cisco Unified Border Elements and Cisco Unified Videoconferencing products face significant limitations in call admission control if they employ thin links between fat links for end-to-end calls. The limitations of the gatekeeper bandwidth controls are especially significant. Currently, bandwidth management is limited to hub-and-spoke configurations, which do not allow video networks to scale adequately. With the Unified Border Element and Resource Reservation Protocol (RSVP), reservation requests can be made across the network on a per-call basis. By using RSVP for call admission control on a hop-by-hop basis, you can scale IP videoconferencing networks to meet the needs of most enterprises and allow video networks to scale larger than a hub-and-spoke environment.

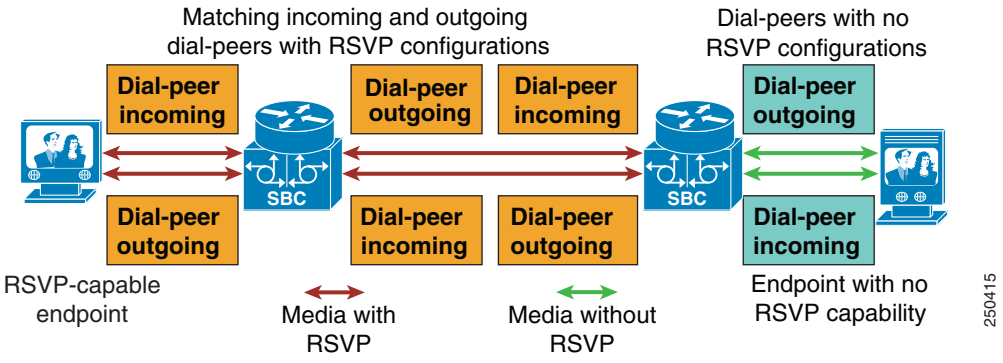
There are two options available for implementing RSVP:

- Use RSVP for call admission control and queuing.
- Use of RSVP for call admission control, and use Differentiated Services (DiffServ) and Cisco modular QoS to service packet flows.

By decoupling the RSVP setup request from the servicing of media flows, network administrators can scale call admission control without maintaining RSVP state for every video call across the entire network.

In Figure 5-4, the two border elements are involved in a video call where one endpoint is RSVP-capable and the other is not. RSVP is enabled for the respective incoming and outgoing dial-peers between the RSVP-enabled border elements. RSVP is also enabled on the border element that is nearest to the endpoint with RSVP support. The border element that is nearest the endpoint with no RSVP support does not have RSVP configured for the dial-peers for this endpoint. The RSVP policy of QoS can be chosen based on preference.

Figure 5-4 Use of Border Elements Between Endpoints with Different RSVP Capabilities



Note On the border element where RSVP can be enabled for video calls, RSVP is associated with its respective dial-peer configuration. Ensure that the incoming and outgoing dial-peers that align with their respective devices either support RSVP or do not have RSVP configured.

For detailed configuration information, refer to the Cisco Unified Border Element documentation available at

http://www.cisco.com/en/US/products/sw/voicesw/ps5640/tsd_products_support_series_home.html