Cisco Unified Videoconferencing Solution Reference Network Design (SRND)

October 30, 2009

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Customer Order Number: OL-14569-03

Cisco Unified Videoconferencing Solution Reference Network Design (SRND)

CONTENTS

Preface

Last revised on: October 30, 2009

This design guide describes a traditional H.323 video architecture using Cisco IOS Gatekeeper functionality for call control.

There are three main video call control architectures:

- H323 video using a gatekeeper with a separate dial plan and no linkage to Cisco Unified Communications Manager (Unified CM).
- H323 video using a gatekeeper with call control linkage to a Cisco Unified Communications domain such as Unified CM, each with separate dial plans.
- H323 video using Unified CM, with a gatekeeper serving as an aggregation point to assist Unified CM with registration services. In this architecture, Unified CM provides the dial plan, call control, call routing, and call admission control.

Cisco recommends the third option listed above, the Cisco Unified Communications architecture for video. This architecture is based on Cisco Unified Communications Manager (Unified CM), which provides unified call control, unified dial plan, unified policy, unified call admission control, and protocol interoperability for both voice and video.

However, some video designs might require a standalone gatekeeper-based solution, as described in the first option listed above. Using a standalone gatekeeper architecture for H323 video deployments is an interim step to a comprehensive Cisco Unified Communications architecture utilizing Unified CM.

This document explicitly addresses only the standalone gatekeeper-based video architecture (the first option above), which incorporates a distinct and separate dial plan, call control, and routing architecture handled by the gatekeeper. All scenarios presented in this document address only H323 support, but it is also possible that Unified CM or some other call control entity might be hosting video applications based on Session Initiation Protocol (SIP) or Skinny Client Control Protocol (SCCP).

The Cisco Unified Communications architecture incorporates Unified CM as a core element that ties many applications together, including video. Unified CM provides the functionality to bring desktop video, traditional room-based video, and new rich-media desktop applications sharing for meetings under one call control domain. It also provides users with dial plan flexibility and ease of use. For details on how to integrate H323 video and support for other protocols under the call control domain of Unified CM, refer to the *Cisco Unified Communications SRND Based on Cisco Unified Communications Manager*, which is available online at

http://www.cisco.com/go/ucsrnd

New or Changed Information for This Release

Note Unless stated otherwise, the information in this document applies to Cisco Unified Videoconferencing 5.7 and subsequent releases. Any differences between the various releases of Cisco Unified Videoconferencing are specifically noted in the text.

The following chapters are either new in the current release of this document, or they contain information that has changed significantly from previous releases of this document.

- Campus Infrastructure, page 3-1
- WAN QoS, page 5-1
- Dial Plan Architecture, page 6-1
- Call Routing, page 7-1
- Cisco Video Infrastructure Components, page 8-1
- Interworking with Session Initiation Protocol (SIP), page A-1

Within each chapter, new and revised information is listed in a section titled *What's New in This Chapter*.

Revision History

This document replaces the *Cisco IP Videoconferencing Solution Reference Network Design Guide*, dated July of 2002.

This document may be updated at any time without notice. You can obtain the latest version of this document online at

http://www.cisco.com/go/ucsrnd

Visit this Cisco.com website periodically and check for documentation updates by comparing the revision date on the front title page of your copy with the revision date of the online document.

The following table lists the revision history for this document.

Revision Date	Comments
October 30, 2009	Updated content as indicated in New or Changed Information for This Release, page viii.
November 15, 2007	Document was updated and published under the new title of *Cisco Unified Videoconferencing Solution Reference Network Design (SRND)*.
July, 2002	Initial release of this document under the old title of *Cisco IP Videoconferencing Solution Reference Network Design Guide*.

Scope of This Document

This document describes the products and features used to build an H.323-based Cisco Unified Videoconferencing system, and it gives recommendations on how to combine those elements into an effective solution for your enterprise. However, this document does not contain specific implementation or configuration details for the products and features. For details about a particular product or feature, refer to the specific product documentation available online at

http://www.cisco.com

Note Unless stated otherwise, the solution designs presented in this document require the minimum software releases listed in Table 1, and the information presented here applies only to those releases.

Table 1 Cisco Unified Videoconferencing Minimum Software Releases

Platform	Minimum Required Software Release
Cisco Unified Videoconferencing 3515-MCU	v5.7
Cisco Unified Videoconferencing 3522-BRI	v5.6
Cisco Unified Videoconferencing 3527-PRI	v5.6
Cisco Unified Videoconferencing 3545-MCU	v5.7
Cisco Unified Videoconferencing Manager	v5.7
Cisco Unified Border Element	Cisco IOS 12.4(22)T

Obtaining Documentation and Submitting a Service Request

For information on obtaining documentation, submitting a service request, and gathering additional information, see the monthly *What's New in Cisco Product Documentation*, which also lists all new and revised Cisco technical documentation, at:

http://www.cisco.com/en/US/docs/general/whatsnew/whatsnew.html

Subscribe to the *What's New in Cisco Product Documentation* as a Really Simple Syndication (RSS) feed and set content to be delivered directly to your desktop using a reader application. The RSS feeds are a free service and Cisco currently supports RSS version 2.0.

CHAPTER 1

Introduction

Last revised on: October 30, 2009

This chapter provides an overview of the H.323 standard and the video infrastructure components used to build an H.323 videoconferencing network. It describes the basics of the H.323 video standard and infrastructure components used throughout this guide.

H.323 Basics

The H.323 standard provides a foundation for audio, video, and data communications across Internet Protocol (IP) networks. H.323 is an umbrella recommendation from the International Telecommunications Union (ITU) that sets standards for multimedia communications over local area networks (LANs). The H.323 standard is part of a larger range of videoconferencing standards (H.32*x*) for videoconferencing over various network media. For example, H.320 supports videoconferencing over Integrated Services Digital Network (ISDN), H.321 supports videoconferencing over Asynchronous Transfer Mode (ATM), H.324 supports videoconferencing over standard Plain Old Telephone Service (POTS) lines, and H.323 supports videoconferencing over IP LANs.

The H.323 specification consists of multiple protocols, including:

- H.245 — Provides control signaling used to exchange end-to-end control messages. These control messages carry information relating to:
  - Capabilities exchange
  - Opening and closing of logical channels used to carry media streams
  - Flow control messages
  - General commands and indications
- H.225 — Provides registration, admission, and status (RAS), which is the protocol used between H.323 devices and the gatekeeper for device registration. The RAS protocol is used to perform registration, admission control, bandwidth utilization updates, status, and disengagement procedures between H.323 devices and the gatekeeper. H.225 is also used during call setup to open a call signaling channel using standard Q.931 messaging protocol.
- H.235 — Provides security by passing tokens in H.225 and H.245 call signaling. It is commonly used for authentication and authorizing calls and for establishing encrypted channels for media of the calls.
- H.239 — Provides data sharing capabilities between video endpoints that can be used to share a PC desktop.

- H.460.17, H.460.18, and H.460.19 — Is used for NAT traversal between endpoints and gatekeepers. H.460.17 provides tunneling of call control signaling in Q.931 over TCP with the gatekeeper. H.460.18 uses extra signaling in RAS with service control indication (SCI) and service control response (SCR) messages, so that inside devices open pinholes from inside for calls using RAS Facility messages. H.460.19 uses empty RTP packets to open pinholes for incoming media.

Table 1-1 lists some of the standards supported by the H.323 specification.

Table 1-1 Protocols Supported by the H.323 Standard

Standard	Supported Functions
H.225	RAS, call setup, and tear down (Q.931 call establishment)
H.235	Security and encryption for H.323
H.245	Call control signaling
T.120	Data sharing for H.320 calls
H.239	Desktop sharing with H.323 calls
H.281	Far End Camera Control (FECC) for H.320 calls
Annex Q	FECC for H.323 calls
H.261 H.263 H.264	Video codecs
G.711 G.722 G.723 G.728 G.729 iLBC AAC-LC AAC-LD	Audio codecs

Videoconferencing with H.323

Historically, videoconferencing was done primarily over ISDN and time division multiplexed (TDM) networks using standard H.320. Running interactive video over data networks was not an option due to video's shared media characteristics, connectionless nature, and lack of guaranteed data flows. With the introduction of switched LAN networks, high-end routers, and Layer 2 and Layer 3 quality of service (QoS), delivering interactive video over IP is now a reality. Today there is a large installed base of H.320 networks that incur large monthly access and switched usage charges. With the current advances to the IP networks, it is now possible to run interactive video over an IP network, thus reducing cost with converged voice, video, and data traffic over a common path. H.323 builds on top of existing IP data networks, scaling to larger deployments and providing greater features. The data sharing capability, remote camera control and enhanced high resolution and high fidelity codecs provide much better video conferencing experience with the endpoints available today. The adoption of videoconferencing to save travel time and costs has attributed to an increase in deployments of video endpoints and videoconferencing devices.

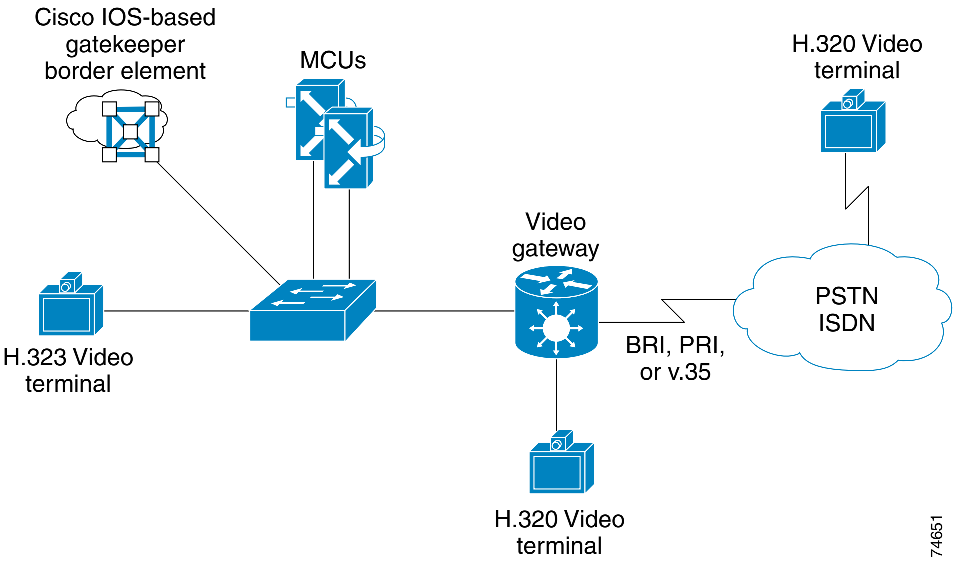
H.323 Videoconferencing Components

Five components make up an H.323 videoconferencing network:

- Video Terminal, page 1-4
- Gatekeeper, page 1-4
- Gateway, page 1-5
- Multipoint Control Unit (MCU), page 1-6
- Border Element, page 1-7

Cisco offers product solutions for all the above components except video terminals, which are covered in detail in *Chapter 8, Video Infrastructure*. Figure 1-1 illustrates a typical H.323 videoconferencing network.

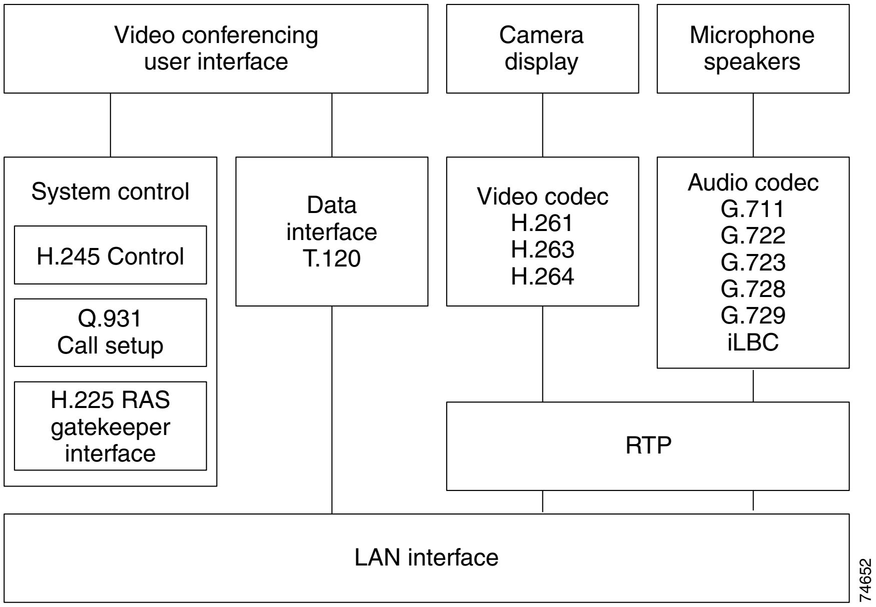
Figure 1-1 H.323 Videoconferencing Infrastructure Components



Video Terminal

Video terminals come in many forms. Some can be connected directly to the ISDN PSTN video network, while some include video systems installed on PCs as standalone desktop terminals and group-focused shared conference room devices with Ethernet for network connectivity. Figure 1-2 illustrates the functional components in an H.323 video terminal.

Figure 1-2 Functional Components of a Video Terminal


Video conferencing user interface
Camera display
Microphone speakers
System control
H.245 Control
Q.931 Call setup
H.225 RAS gatekeeper interface
Data interface T.120
Video codec H.261 H.263 H.264
Audio codec G.711 G.722 G.723 G.728 G.729 iLBC
RTP
LAN interface
74652

Note Some video endpoints support streaming with Real Time Streaming Protocol (RTSP), which can enable a larger number or participants to view the call.

Gatekeeper

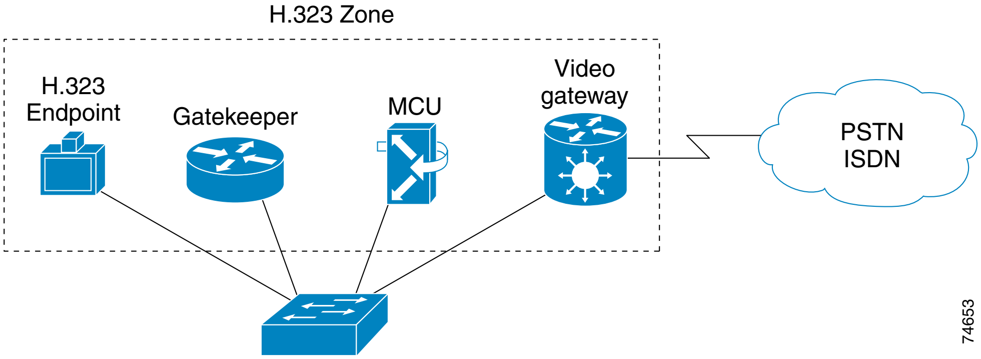
The gatekeeper is one of the most important components of an H.323 videoconferencing network. Although the H.323 standard lists the gatekeeper as an optional device, you cannot build a scalable video network without the application controls the gatekeeper provides. Each video infrastructure component registers with the gatekeeper. The gatekeeper performs all address resolution, bandwidth management, admission control, zone management, and intra-zone and inter-zone call routing.

A zone is a logical grouping of H.323 infrastructure components registered to, and managed by, a single gatekeeper. Zones are not dependent on physical network topology or IP subnets. Zones may span one or more network segments or IP subnets, and they are simply a logical grouping of devices. As such, zones can be defined based on geographical proximity, bandwidth availability, or other criteria. A *via-zone* is another type of zone that contains a Cisco Unified Border Element so that the gatekeeper can include it in the call, depending on the configuration.

The most fundamental function of a gatekeeper is to provide address resolution, thus allowing terminals, gateways, and Multipoint Control Units (MCUs) to be addressed using the international E.164 address standard and/or an H.323 alias. Each endpoint that is registered to a gatekeeper must be assigned a unique E.164 address (numeric identifier). As a result, zone prefixes are used in the H.323 video network to identify zones, similar to the use of area codes in telephony systems.

Throughout this document are example topologies that are based on single-zone and multi-zone configurations. For example, Figure 1-3 illustrates a single zone.

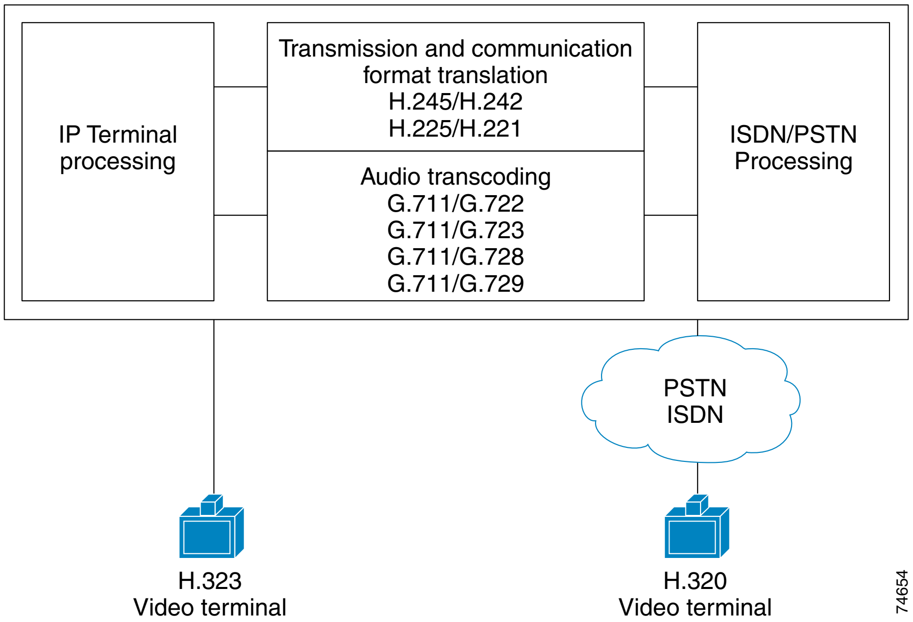
Figure 1-3 Single H.323 Zone



Gateway

Gateways provide interoperability between H.323 elements and an installed base of H.320 units. The H.323 gateway allows H.323 video terminals to communicate with other H.32*x* video terminals, such as H.320 and H.321 video terminals. Video gateways perform translation between different protocols, audio encoding formats, and video encoding formats that may be used by the various H.32*x* standards. For example, the ISDN H.320 standard uses the H.221 protocol for signaling, while the H.323 standard uses H.225. The gateway must translate between these two protocols to allow devices of different network media and protocols to communicate with each other. ISDN gateways can also support Interactive Voice Response (IVR), Direct Inward Dialing (DID), or TCS4 (ISDN H.320-based dialing) for video calls. Figure 1-4 illustrates the role of a gateway in an H.323 video network.

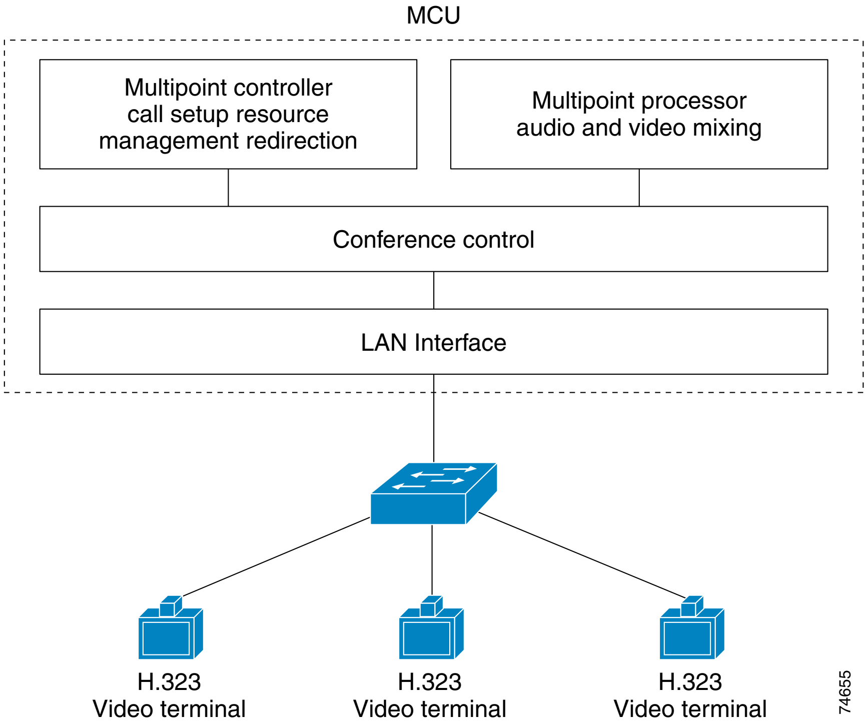
Figure 1-4 Functional Components of an H.323 Video Gateway



Multipoint Control Unit (MCU)

Video terminals are generally point-to-point devices, allowing only two participants per conversation. A multipoint control unit (MCU) allows video conferences to be extended to three or more participants, and some video terminals also support multipoint calls. An MCU consists of a multipoint controller (MC) and a multipoint processor (MP). The MC manages all call setup control functions and conference resources as well as the opening and closing of media streams. The MP processes audio and video media streams only. Cisco MCUs can be stacked to allow more conferences or cascaded to allow larger conferences. Stacking and cascading are covered in detail in *Chapter 8, Video Infrastructure*. Figure 1-5 illustrates the function of an MCU.

Figure 1-5 Functional Components of an MCU

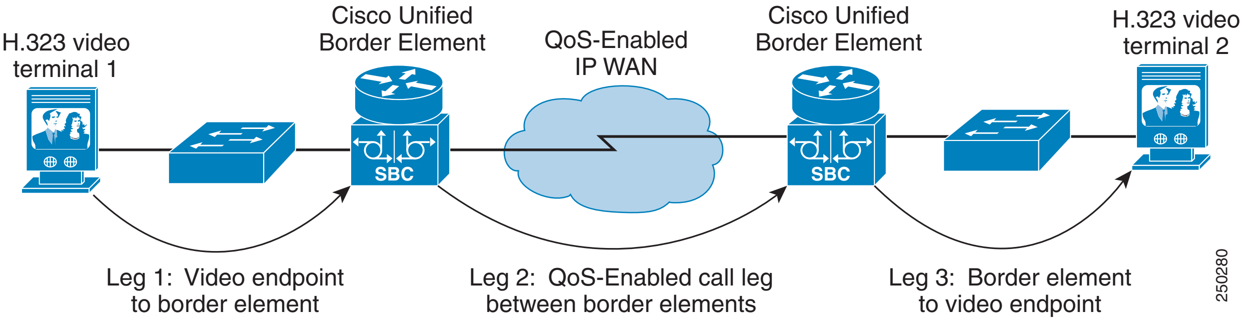


Border Element

The border element is a device that is used in the periphery of the network to separate two different networks, and it serves as a demarcation device. The border element can use H.323 or other protocols and is most commonly used for topology hiding and interworking. The border element can terminate H.323 calls from a local LAN or zone and establish sessions with H.323 endpoints located in other LANs or zones. In doing so, it provides network administrators with the ability to set and enforce quality of service (QoS) on inter-zone segments. The border element also provides a method of identifying H.323 videoconferencing connections for tunneling through firewalls and Network Address Translation (NAT) environments. Figure 1-6 illustrates a call passing through the border element over a WAN link.

The Cisco border element product is the Cisco Unified Border Element. (The Cisco Unified Border Element was previously named the IP-to-IP Gateway, which was a successor to the Cisco Multimedia Conference Manager (MCM) proxy.)

Figure 1-6 A Call Passing Through the Border Element over a WAN Link


H.323 video terminal 1
Cisco Unified Border Element
SBC
QoS-Enabled IP WAN
Cisco Unified Border Element
SBC
H.323 video terminal 2
Leg 1: Video endpoint to border element
Leg 2: QoS-Enabled call leg between border elements
Leg 3: Border element to video endpoint
250280

Note Cisco Unified Communications Manager (Unified CM) supports voice and video and can be extended to support multimedia conferencing with Cisco Unified MeetingPlace solutions. These systems can be integrated with Cisco Unified Videoconferencing devices using gatekeepers and border element devices with H.323 protocol. For information on integrating Cisco Unified Video Gateways, MCUs, and Cisco Unified MeetingPlace with Unified CM, refer to the *Cisco Unified Communications SRND Based on Cisco Unified Communications Manager* available at http://www.cisco.com/go/ucsrnd.

CHAPTER 2

Deployment Models

Last revised on: October 30, 2009

This chapter introduces four basic design models used to deploy Cisco Unified Videoconferencing solutions:

- Campus Single Zone, page 2-2
- Campus Multi Zone, page 2-4
- WAN Single Zone, page 2-5
- WAN Multi Zone, page 2-7

This chapter provides basic design criteria and guidelines for selecting the correct deployment model. Subsequent chapters of this design guide describe in more detail each of the basic models introduced here.

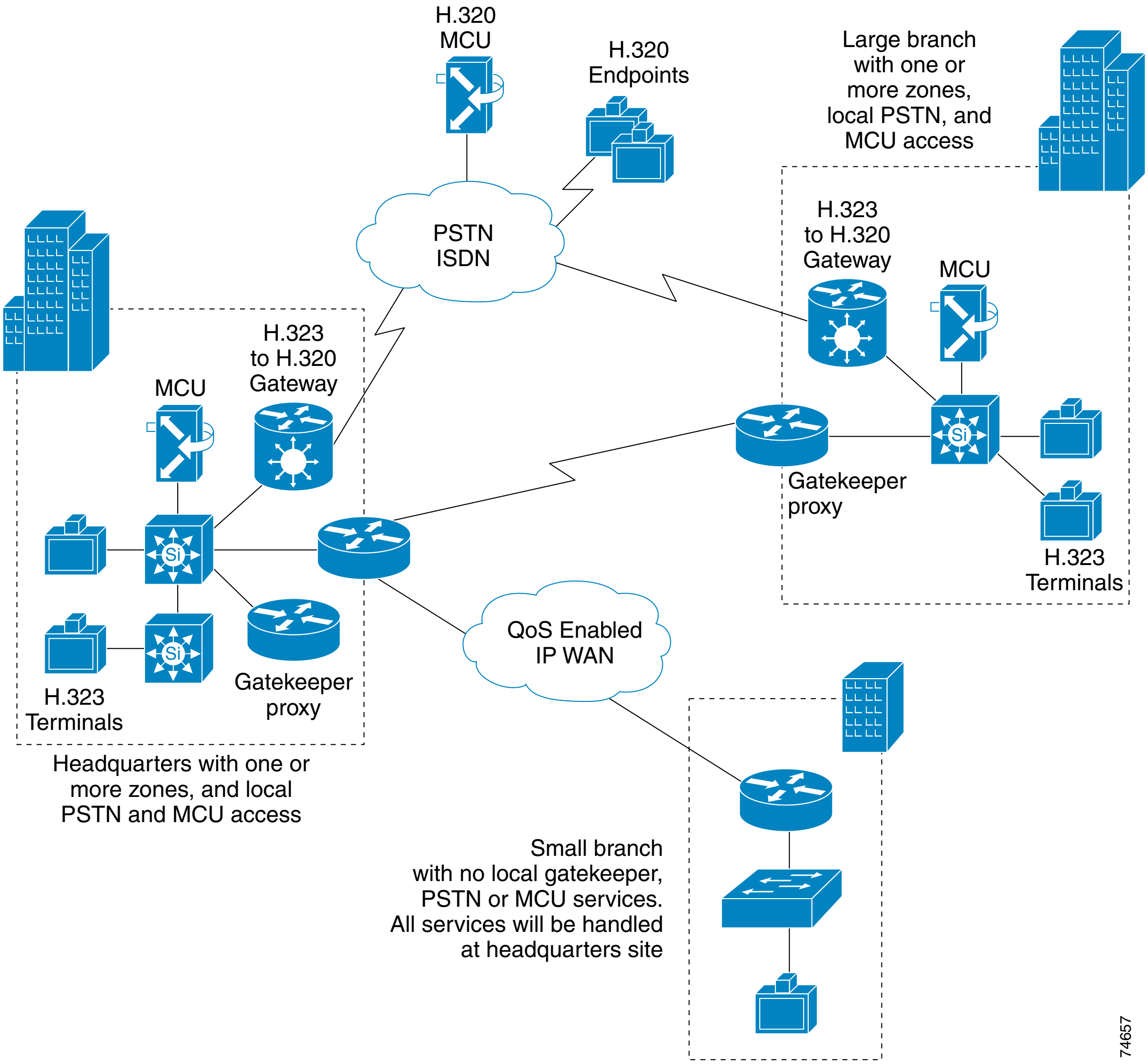
Composite Deployment Model

Figure 2-1 illustrates a composite topology that encompasses all of the deployment models discussed in this guide. All designs discussed in this chapter are supported with currently shipping products.

The overall goals of a Cisco-based H.323 videoconferencing solution are as follows:

- Provide end-to-end IP video connectivity across the corporate infrastructure, with *business quality* transmission. Business quality video is defined as 30 frames per second operation with a minimum of Common Intermediate Format (CIF) resolution. Typically, this level of quality requires 384 kbps of application bandwidth for most video terminals.
- Provide quality of service (QoS) — high availability with low latency and jitter (delay variability).
- Reduce Integrated Services Digital Network (ISDN) costs by eliminating the need for ISDN attachments directly to video terminals.
- Allow Public Switched Telephone Network (PSTN) access to legacy H.320 systems through shared gateway resources.
- Support multipoint calling through Multipoint Control Units (MCUs).
- Conserve WAN bandwidth by distributing MCU and gateway resources across the IP infrastructure.
- Lower total cost of ownership for the video network by utilizing the existing IP infrastructure.
- Support manageability of multiple H.323 elements in a distributed network topology.

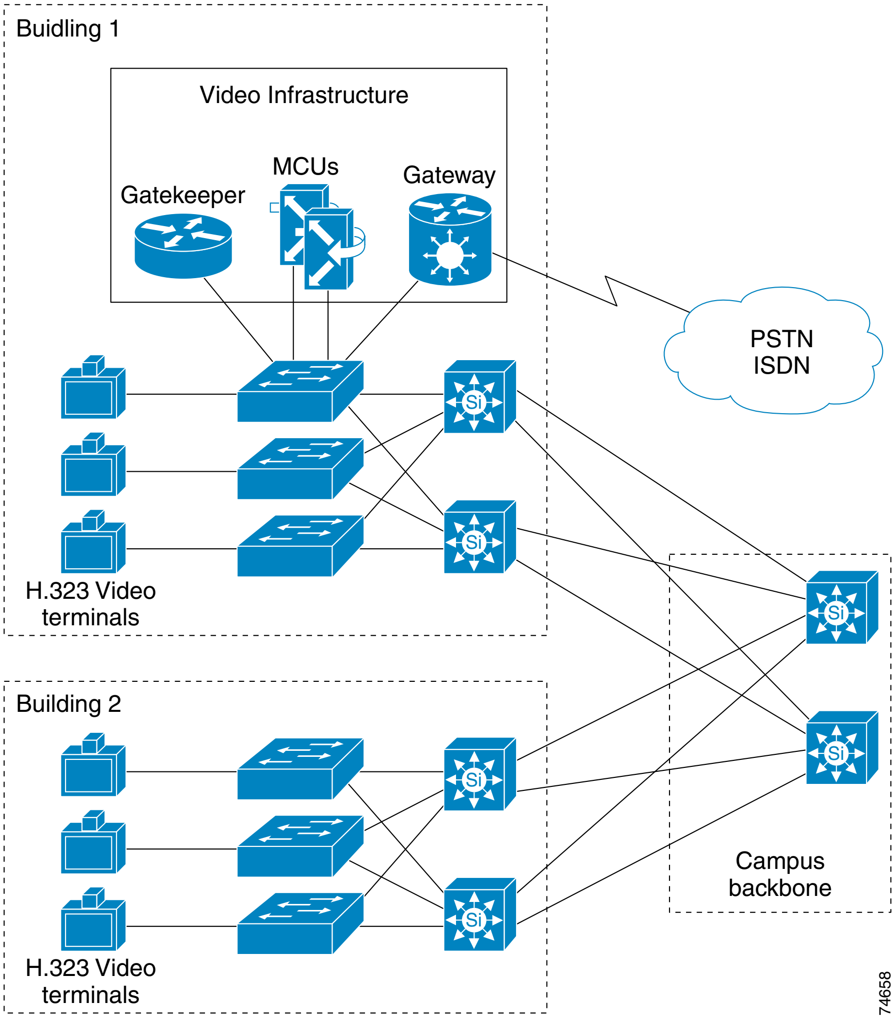
Figure 2-1 Composite Deployment Model



Campus Single Zone

Figure 2-2 illustrates an H.323 network in a campus environment configured with a single zone. This is the most basic design model to implement and is used in pilot installs and smaller video environments.

Figure 2-2 Campus Single Zone



The campus single-zone deployment model has the following design characteristics:

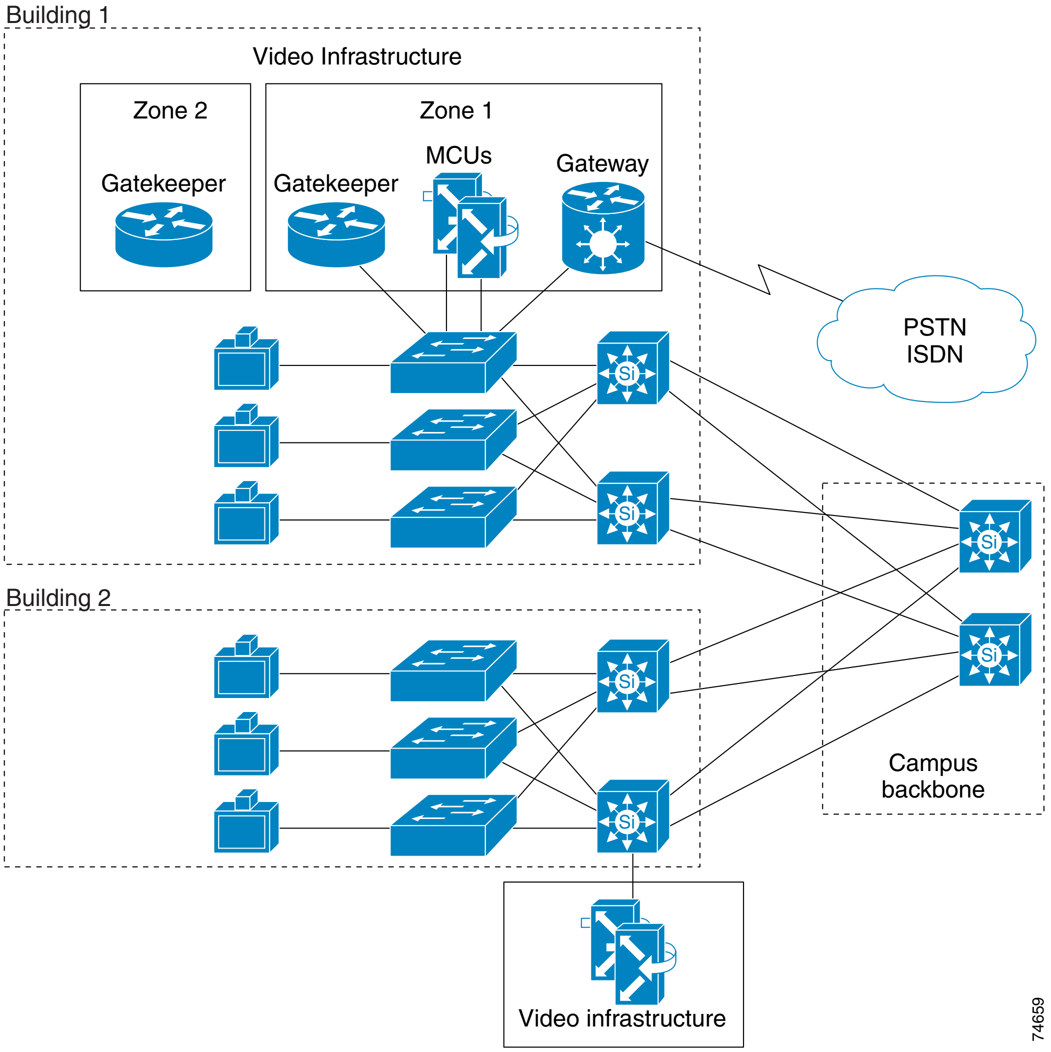
- A single gatekeeper supporting a single zone for H.323 video.
- All H.323 video users registered with the single gatekeeper. (For additional scalability details, refer to the Cisco IOS Gatekeeper information at www.cisco.com/go/cube.)
- Optional PSTN access available through the Cisco Unified Videoconferencing 3500 Series gateways.
- Optional multipoint conferencing available through the Cisco Unified Videoconferencing 3500 Series MCU.
- Zone bandwidth managed by the configured gatekeeper.
- All gateway and MCU services registered and managed by a single gatekeeper.
- Call routing between endpoints using fully qualified E.164 addresses or H.323-ID.

Campus Multi Zone

Figure 2-3 illustrates a multi-zone H.323 video network in a campus environment. This model is most often implemented in an enterprise campus network. Depending on business function, administrators may choose to create different zones for security reasons. For example, company executives may be registered in a single zone that is separate from other users to allow administrators to limit access to those video terminals. In addition, as a video network grows, a single zone may not be manageable because of the number of users or the ability to manage network resources.

Note Multiple zones can be configured on a single router. If you configure multiple local zones on a single router, and MCUs and/or gateways are registered with the zones, you must add hopoff statements for each service prefix. If hopoffs are not added for each service prefix, the video terminal will not be able to access MCUs or gateways outside its local zone. See Routing Inter-Zone Calls Using Hopoff Statements, page 7-8, for more information.

Figure 2-3 Campus Multi Zone



The campus multi-zone deployment model has the following design characteristics:

- Multiple gatekeepers supporting multiple zones for H.323 video.
- H.323 endpoints register with one of the multiple gatekeepers. (For additional scalability details, refer to the Cisco IOS Gatekeeper information at www.cisco.com/go/cube.)
- Bandwidth management for each zone and between zones is controlled by configured gatekeepers.
- Optional PSTN access available through Cisco Unified Videoconferencing 3500 Series gateways.
- Gateway and MCU services are registered and managed across multiple gatekeepers.
- Gateway and MCU services may be distributed throughout the campus.
- H.323 users and services are segmented for security, bandwidth control, and resource allocation.
- Intra-zone and inter-zone call routing using fully qualified E.164 address or H.323-ID.

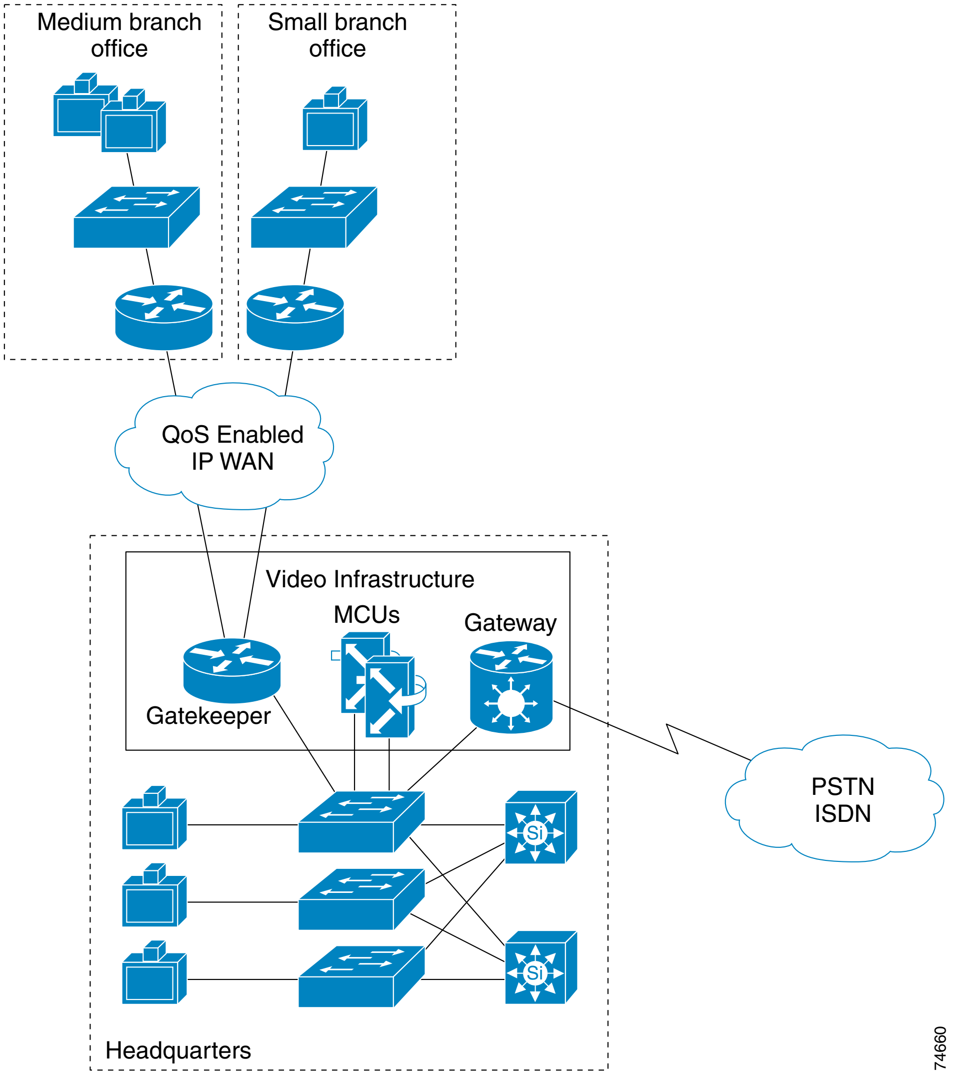
WAN Single Zone

Figure 2-4 illustrates a single-zone H.323 video network in a WAN environment. This deployment model is used when remote sites have a small number of video endpoints, usually no more than one or two at each remote site on a T1 WAN link. From a management or economic standpoint, it might not make sense to create a zone at each remote site for one or two video terminals. Call admission control across the WAN is not usually an issue with only one or a few video terminals at each remote site, but it is an issue when the number of simultaneous calls across the WAN from remote endpoints exceeds the provisioned video bandwidth.

In the absence of a gatekeeper, implement quality of service on the WAN ports by using one of the following methods:

- Priority queuing on traffic classification IP Precedence 4, or Differentiated Services Code Point (DSCP) AF41
- Access control list (ACL) for each video terminal at the remote site, to direct the video streams to the appropriate priority queue

Figure 2-4 WAN Single Zone



The WAN single-zone deployment model has the following design characteristics:

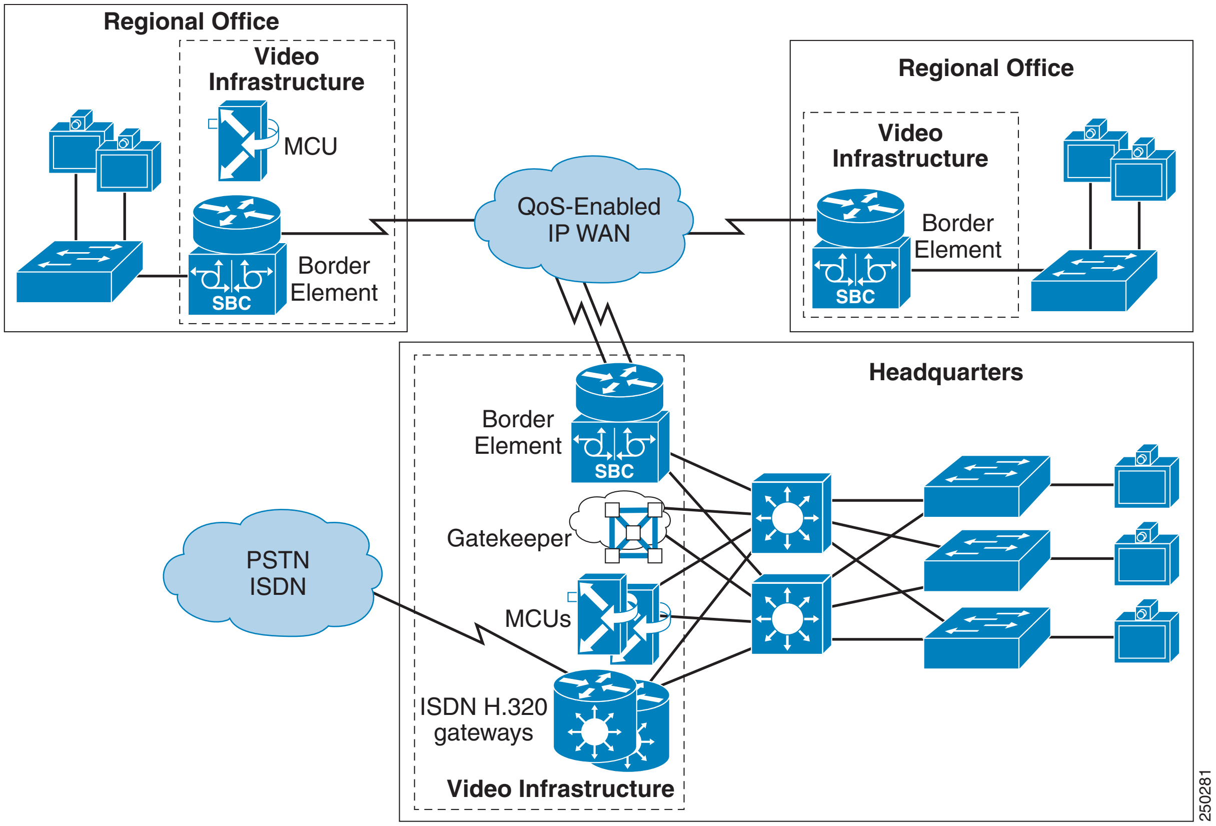
- A single gatekeeper supporting a single zone for H.323 video.
- All H.323 video users registered with the single gatekeeper. (For additional scalability details, refer to the Cisco IOS Gatekeeper information at www.cisco.com/go/cube.)
- Optional PSTN access available through Cisco Unified Videoconferencing 3500 Series gateways.
- Optional multipoint conferencing available through the Cisco Unified Videoconferencing 3500 Series MCU.
- H.323 video bandwidth managed by a single gatekeeper.
- All gateway and MCU services registered and managed by a single gatekeeper.
- WAN QoS, with priority queuing by means of traffic classification or ACL entries.
- Call routing between endpoints using fully qualified E.164 addresses or H.323-ID.

WAN Multi Zone

Figure 2-5 illustrates a multi-zone H.323 network in a WAN environment. This deployment model is used in large enterprise, government, and educational networks. QoS can be implemented using either the border element and priority queuing (PQ) features in Cisco IOS software, traffic classification by the video terminals, or Layer 3 switches in conjunction with priority queuing on the WAN ports of the routers.

Creating multiple zones in a WAN environment allows administrators to manage network resources and assure video quality across low-speed WAN links. Call admission control is very important in a large WAN environment. With multiple zones enabled, the gatekeeper can manage the total amount of H.323 video bandwidth allowed across a particular network link. For example, you could limit the total H.323 video bandwidth across a T1 WAN link to 768 kbps, and the gatekeeper would then reject any call request that exceeds this limit of 768 kbps.

Figure 2-5 WAN Multi Zone



The WAN multi-zone deployment model has the following design characteristics:

- Multiple gatekeepers supporting multiple zones for H.323 video.
- H.323 endpoints and services registered with the assigned gatekeeper, usually at the local site.
- Optional PSTN access available through Cisco video gateways.
- Bandwidth management available in each zone and across the WAN, using the gatekeeper at each site.

- Distributed services available at larger branch sites to conserve bandwidth.
- Inter-zone and intra-zone call routing using fully qualified E.164 addresses or H.323-ID.
- Priority queuing (PQ) based on traffic classification implemented on the WAN ports, or a border element at each site with PQ on the WAN.

CHAPTER 3

Campus Infrastructure

Last revised on: October 30, 2009

This chapter provides guidelines for deploying H.323 videoconferencing with Quality of Service (QoS) on a campus network using one of the following basic H.323 video designs:

- Single-Zone Campus, page 3-2
- Multi-Zone Campus, page 3-3

What's New in This Chapter

Table 3-1 lists the topics that are new in this chapter or that have changed significantly from previous releases of this document.

Table 3-1 New or Changed Information Since the Previous Release of This Document

New or Revised Topic	Described in:
Traffic classification	Traffic Classification Types, page 3-4

Network Infrastructure

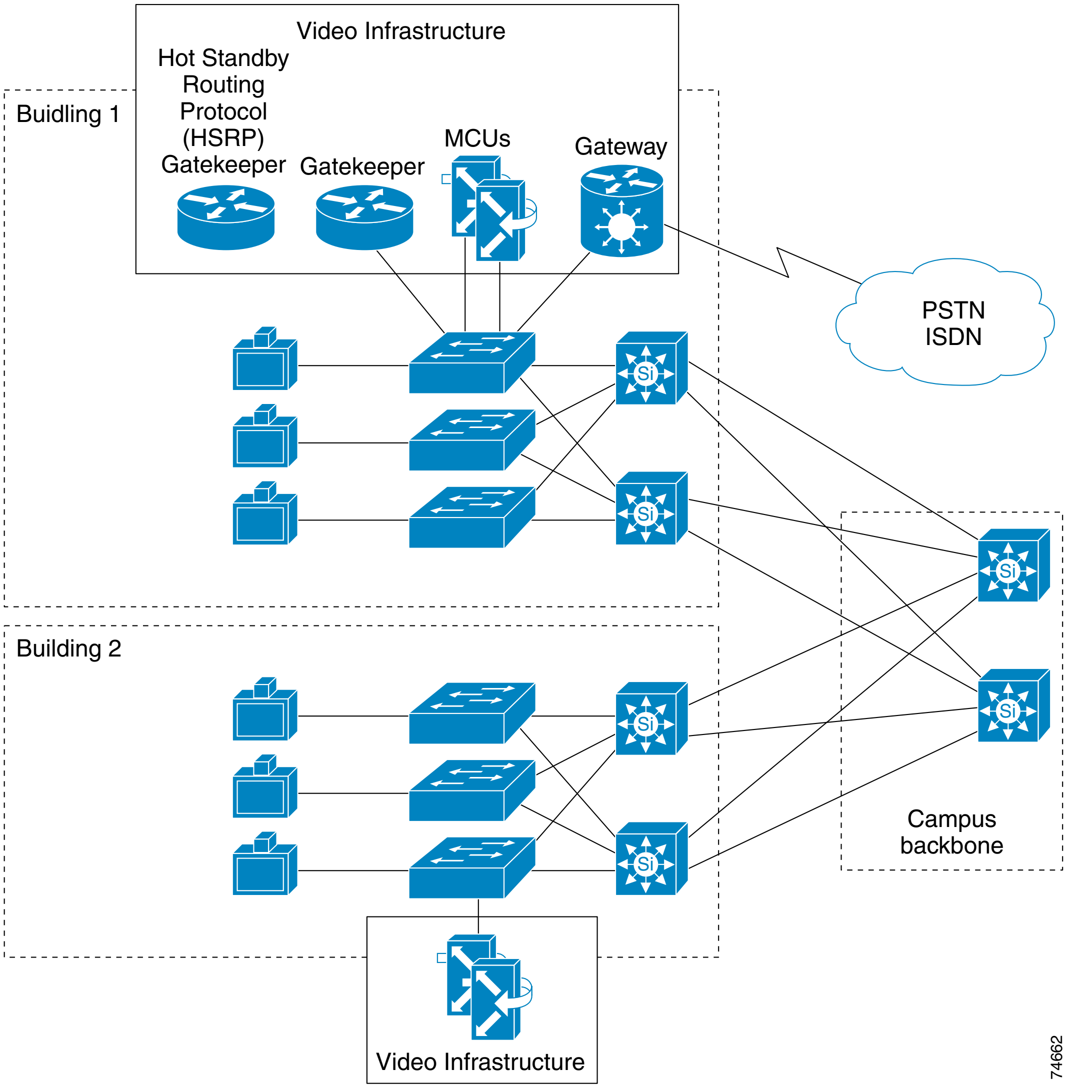
Building an end-to-end H.323 video network requires an infrastructure based on Layer 2 and Layer 3 switches and routers. It is important to have all H.323 video endpoints, gateways, and multipoint control units (MCUs) connected to a dedicated 10/100/1000 switched Ethernet port. Cisco recommends using a 100/1000-Mbps full duplex connection to the Cisco gatekeeper to ensure adequate bandwidth on all router platforms. Some endpoints, however, do not support 100/1000-Mbps full duplex. For example, older Polycom ViewStations and the Cisco Unified Videoconferencing 3530 both support 10-Mbps half duplex only.

Note Cisco recommends that you set all switch ports attached to H.323 video devices to 10/100/1000 Mbps full duplex whenever possible. If the video unit supports only 10 Mbps, configure the switch port for 10 Mbps half duplex.

Single-Zone Campus

Figure 3-1 illustrates an H.323 single-zone campus network.

Figure 3-1 Single-Zone Campus



Single-zone campus networks are usually used in pilot deployments or in campuses with a small number of video terminals or endpoints. The single-zone campus deployment allows an administrator to deploy H.323 video on the campus while keeping management overhead to a minimum. There is only one gatekeeper to manage, and the dial plan is very simple with no inter-zone call routing.

It is important to consider multi-zone dial plans when deploying a single-zone model. If you deploy a single-zone dial plan but need to upgrade to a multi-zone model in the future, you will have to change the entire dial plan. Therefore, to simplify future network scaling, Cisco recommends that you use a multi-zone dial plan even for a single-zone campus.

In summary, a single-zone campus model consists of:

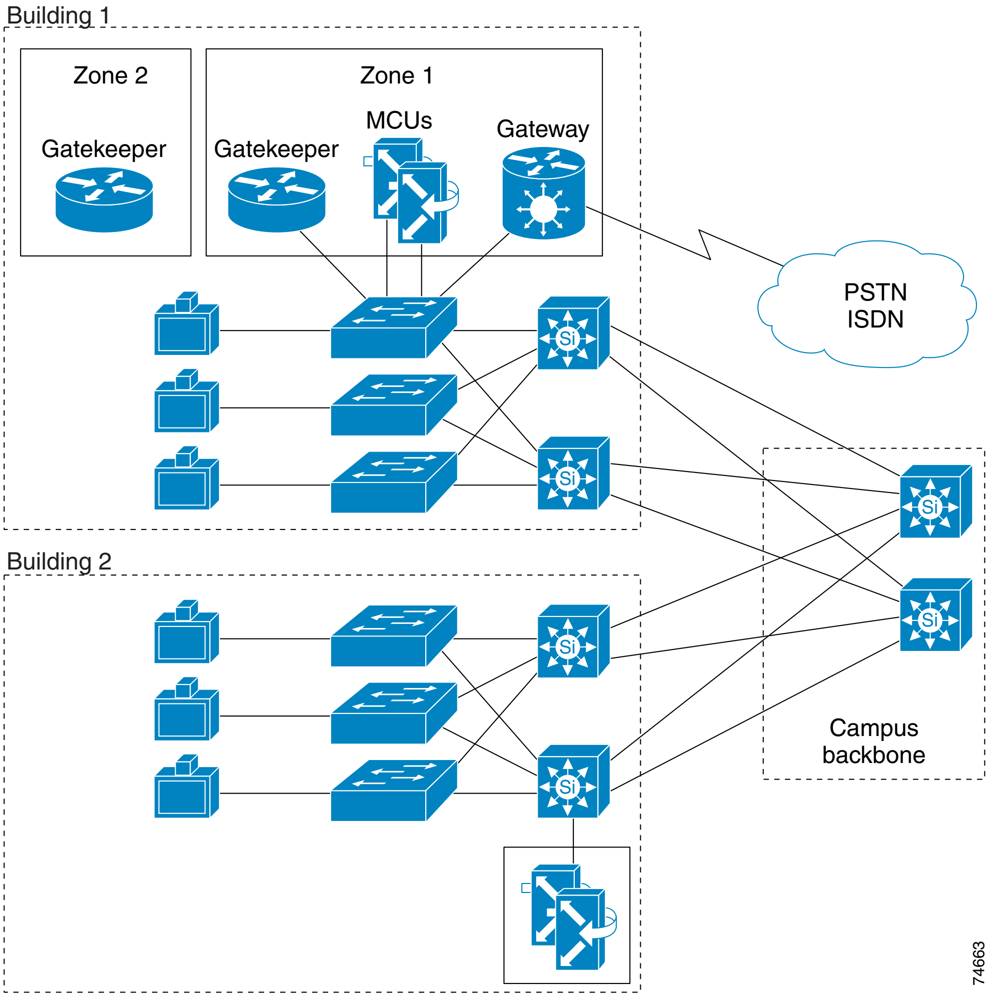
- Campus environment

- Pilot environments
- Small number of video endpoints
- No bandwidth limitations

Multi-Zone Campus

Figure 3-2 illustrates an H.323 multi-zone campus network.

Figure 3-2 Multi-Zone Campus



Multi-zone campus networks are common in large campus environments. Creating multiple zones allows administrators to segment user groups for security, better management of the H.323 video network, and bandwidth control in and between zones. For example, company executives may be registered in a single zone containing their own gateway and MCU resources.

In campuses with a large number of video terminals, it is important to control the amount of video bandwidth on the network. With a single zone, bandwidth management capabilities are very limited. Creating multiple logical zones on the campus allows an administrator to manage bandwidth within and between zones.

Physical placement of gatekeepers, MCUs, and gateways depends on customer preference and network configuration. Some deployments locate all of the gatekeepers, MCUs, and gateways in a single data center, while others may decide to distribute the equipment throughout the campus.

In summary, the multi-zone campus model consists of:

- Campus environment
- Large numbers of video terminals
- Users segmented into separate video domains
- Restricted access for some users

Note Multiple zones can be configured on a single gatekeeper. If you configure multiple local zones on a single gatekeeper, you must add hopoff commands for each service prefix registered. If hopoffs are not added for each service prefix, the video terminal will not be able to access MCUs or gateways outside its local zone. See Routing Inter-Zone Calls Using Hopoff Statements, page 7-8 for more information.

Quality of Service

In a converged environment, voice, video and data traffic all travel over a single transport infrastructure. Not all traffic types should be treated equally. Data traffic is bursty, loss tolerant, and not sensitive to delay. Video traffic, on the other hand, is bursty, has very little tolerance for loss, and is latency sensitive. The challenge is to provide the required level of service for all three traffic types.

Running both video and data on a common network requires the proper QoS tools to ensure that the delay and loss parameters of video traffic are satisfied in the face of unpredictable data flows. Some of these tools may be available as a feature in some video terminals (for example, Polycom, Tandberg, and Sony), switches, and routers.

Traffic Classification Types

The first step in preserving video quality on a data network is to classify video traffic as high priority and allow it to travel through the network before lower priority traffic. Data traffic can be classified into various data classes with data queues without adversely affecting its performance because of its characteristics as provided by the Transfer Control Protocol (TCP), which handles flow control and error correction. For video, classify traffic at Layer 2 and Layer 3 as follows:

- At Layer 2, use the three bits in the 802.1Qp field, referred to as class of service (CoS), which is part of the 802.1Q tag.
- At Layer 3, use the three bits of the Differentiated Services Code Point (DSCP) field in the type of service (ToS) byte of the IP header.

Traffic classification is the first step toward achieving QoS. Ideally, you should perform this step as close to the source as possible. However, you can also set this field within the Cisco Unified Border Element using a Cisco IOS feature. For H.323 signaling and RTP media, you can use access control lists to classify videoconferencing traffic by transport type and port ranges. Table 3-2 lists the recommended traffic classifications for various applications.

Table 3-2 Traffic Classification Guidelines for Various Types of Network Traffic

	Layer-3 Classification			**Layer-2 Classification**
Application	**IP Precedence (IPP)**	**Per-Hop Behavior (PHB)**	**Differentiated Services Code Point (DSCP)**	**Class of Service (CoS)**
Routing	6	CS6	48	6
Voice Real-Time Transport Protocol (RTP)	5	EF	46	5
Videoconferencing	4	AF41	34	4
Streaming video	4	CS4	32	4
Call signaling[1]	3	CS3 (currently) AF31 (previously)	24 (currently) 26 (previously)	3
Transactional data	2	AF21	18	2
Network management	2	CS2	16	2
Scavenger	1	CS1	8	1
Best effort	0	0	0	0

[1] The recommended DSCP/PHB marking for call control signaling traffic has been changed from 26/AF31 to 24/CS3. A marking migration is planned within Cisco to reflect this change, however many products still mark signaling traffic as 26/AF31. Therefore, in the interim, Cisco recommends that both AF31 and CS3 be reserved for call signaling.

Trust Boundaries

The concept of trust is an important and integral part of deploying QoS. Once the end devices have set ToS values, the switch has the option of trusting them or not. If the switch trusts the ToS values, it does not need to do any reclassification; if it does not trust the values, then it must reclassify the traffic for appropriate QoS.

The notion of trusting or not trusting forms the basis for the trust boundary. Ideally, traffic classification should be done as close to the source as possible. If the end device is capable of performing traffic classification, then the trust boundary for the network is at the access layer in the wiring closet. If the device is not capable of performing traffic classification, or if the wiring closet switch does not trust the classification done by the end device, the trust boundary should shift to other devices.

Shifting of the trust boundary depends on the capabilities of the switch in the wiring closet. If the switch can reclassify the packets, then the trust boundary remains in the wiring closet. If the switch cannot perform this function, then the task falls to other devices in the network going toward the backbone. In this case, reclassification occurs at the distribution layer, which means that the trust boundary has shifted to the distribution layer. For this shift to occur, there must be a high-end switch in the distribution layer with features to support traffic reclassification. If possible, try to avoid performing traffic reclassification in the core of the network.

In summary, try to maintain the trust boundary in the wiring closet. If necessary, move it down to the distribution layer on a case-by-case basis, but avoid moving it to the core of the network. This advice conforms to the general guidelines for keeping the trust boundary as close to the source as possible.

Note This discussion assumes a three-tier network model, which has proven to be a scalable architecture. If the network is small and the logical functions of the distribution layer and core layer happen to be in the same device, then the trust boundary can reside in the core layer if it has to move from the wiring closet. For detailed configuration information, refer to the *Enterprise QoS Solution Reference Network Design Guide*, available at http://www.cisco.com/go/designzone.

QoS Features Summary

Table 3-3 shows supported QoS features on each switch platform.

Table 3-3 Supported QoS Features by Switch Platform

Platform	Auto QoS	Reclassify CoS/DSCP	Congestion Avoidance	Priority Queues	Multiple Queues	Traffic Management	Policing
Catalyst 2960	Yes	Yes	Yes (WTD[1])	Yes	4 egress queues/port	SRR[2]	Yes
Catalyst 3560	Yes	Yes	Yes (WTD[1])	Yes	4 egress queues/port	SRR[2]	Yes
Catalyst 3760	Yes	Yes	Yes (WTD[1])	Yes	4 egress queues/port	SRR[2]	Yes (64 policy rates or individual port)
Catalyst 4006 or 450*x* with Supervisor Engine IV	Yes	Yes	Yes (DBL[3] and QoS sharing on non-blocking Gb ports)	Yes	4 egress queues/port	SRR[2]	Yes (64 policy rates or individual port)
Catalyst 4006 or 45*xx* with Supervisor Engine V	Yes	Yes	Yes (DBL[3] and QoS sharing)	Yes	4 egress queues/port	SRR[2]	Yes (64 policy rates or individual port)
Catalyst 6000 with Policy Feature Card (PFC3)	Yes	Yes	Yes (scheduling and QoS sharing)	Yes	4 egress queues/port	SRR[2] or DWRR[4]	Yes (64 policy rates or individual port)

[1] Weighted Tail Drop (WTD)

[2] Shaped Round Robin (SRR)

[3] Dynamic Buffer Limiting (DBL)

[4] Deficit Weighted Round Robin (DWRR)

In summary, follow these recommendations for QoS deployment:

- Create a trust boundary at the network edge in the wiring closet. Enable the trust boundary on ports on the wiring closet switch where video terminals have the ability to set IP precedence. A rule of thumb is to trust the classification from conference room systems and *not* trust classification from desktop video units.
- Reclassify ToS at the edge if devices (both room systems and desktop units) cannot be trusted.
- Shift the trust boundary to the distribution layer and reclassify ToS there if reclassification is not possible at the edge.
- Use a priority queue for delay-sensitive video traffic.

CHAPTER 4

WAN Infrastructure

Last revised on: October 30, 2009

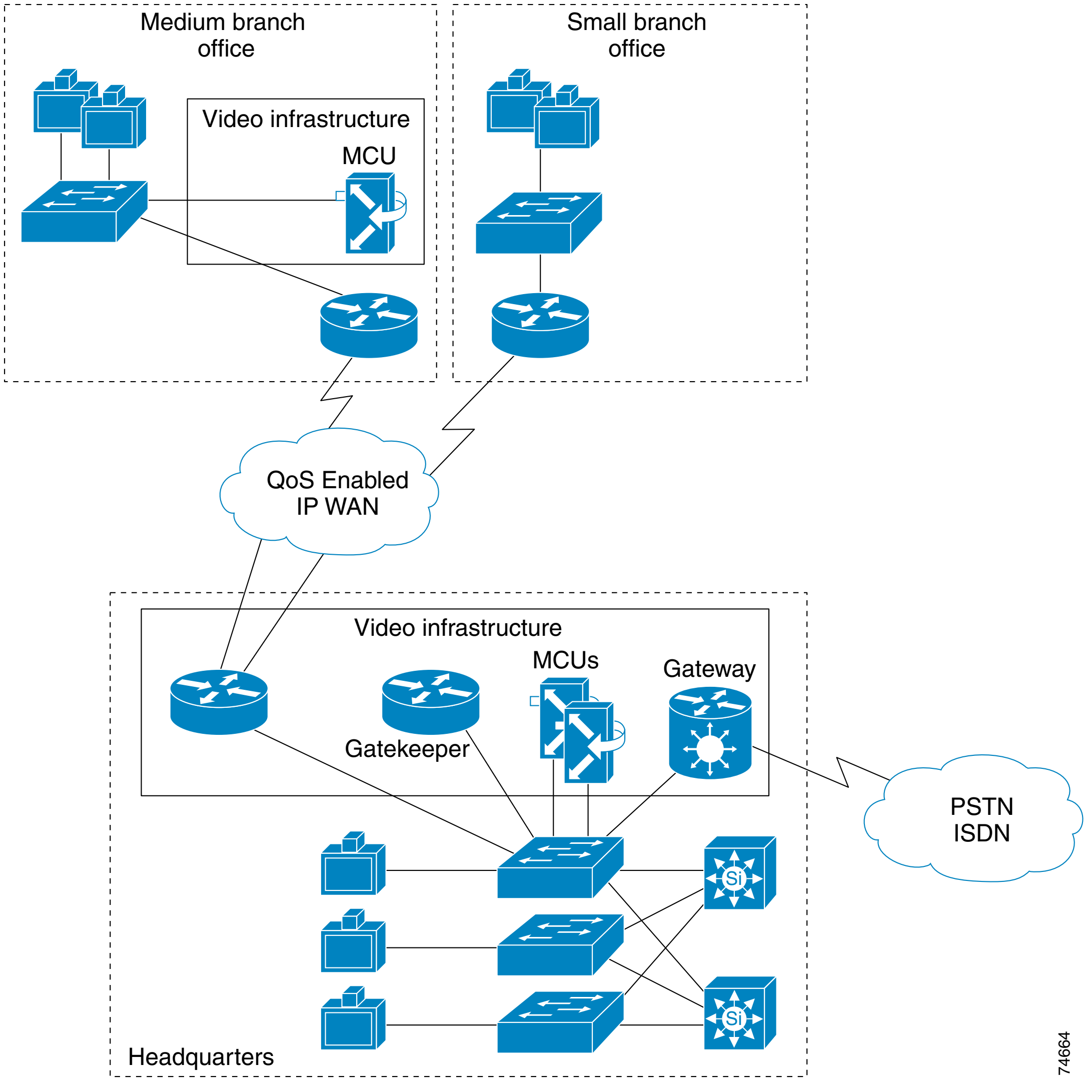
This chapter provides guidelines for deploying H.323 video across an IP WAN, and it describes IP WAN infrastructure design considerations for:

- Single-Zone WAN, page 4-2
- Multi-Zone WAN, page 4-4

Single-Zone WAN

Figure 4-1 illustrates a single-zone WAN network.

Figure 4-1 Single-Zone WAN



A single-zone WAN model consists of the WAN environment and less than three videoconferencing terminals per remote site. (This limit is based on a T1 WAN link.) Cisco recommends that you configure a gatekeeper and a zone for a remote site with one or two video terminals, but this configuration is not strictly required.

Due to the limited number of endpoints and traffic classification options, you can achieve quality of service (QoS) and call admission control by following these basic rules:

- The total data rate of the video terminals plus 20% should not exceed 33% of the WAN link capacity.
- The priority queue must be provisioned for the maximum data rate of the video terminals plus 20%.

  For example, assume a site has a link capacity of 1.544 Mbps and contains two video terminals that support a maximum data rate of 256 kbps each. Therefore, the required queue size for the two video terminals is (256+256)x120% = 614 kbps. Provisioning the priority queue for 614 kbps allows both video terminals to be in a call across the WAN at the same time, without the possibility of overrunning the priority queue. If we add a third video terminal in this example, we would need to add a gatekeeper and create a zone to provide call admission control.

The key elements for successful deployment of videoconferencing in a single-zone WAN environment are:

- Traffic Classification, page 4-3
- Call Admission Control, page 4-4
- Provisioning, page 4-4
- Priority Queuing on the WAN, page 4-4
- Entrance Criteria, page 4-4

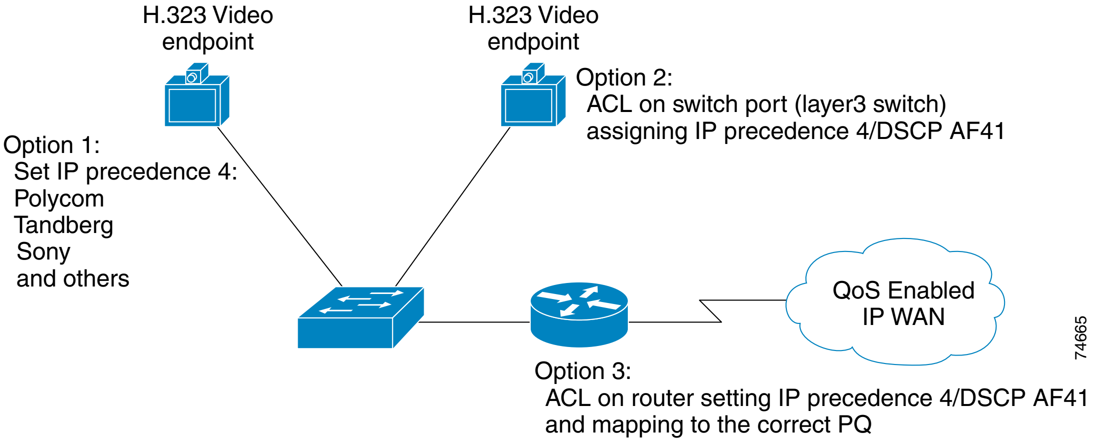
Traffic Classification

Classify traffic at one of the following places:

- Video endpoint (Polycom, Tandberg, Sony, and others); IP Precedence 4 or DSCP AF41
- Switch port (Layer 3 switch required); IP Precedence 4 or DSCP AF41 (recommended)
- Router (ACL entry); IP Precedence 4 or DSCP AF41

Figure 4-2 illustrates these three options for traffic classification.

Figure 4-2 Traffic Classification Options for Single-Zone WAN


H.323 Video endpoint
H.323 Video endpoint
Option 2:
ACL on switch port (layer3 switch)
assigning IP precedence 4/DSCP AF41
Option 1:
Set IP precedence 4:
Polycom
Tandberg
Sony
and others
QoS Enabled IP WAN
74665
Option 3:
ACL on router setting IP precedence 4/DSCP AF41
and mapping to the correct PQ

Call Admission Control

For remote sites that do not have a gatekeeper to enforce call admission control, provision the priority queue and limit the number of video terminals at each site. The number of video terminals multiplied by the maximum call data rate, must not exceed the capacity of the priority queue. Cisco recommends that you use a gatekeeper and zones for remote sites with more than two video terminals. You can install a gatekeeper at each remote site with more than two video terminals, or you can install one gatekeeper at the central site and define a separate zone for each remote site.

Note This recommendation is based on a T1 WAN link.

Provisioning

Provision WAN queues according to the following equation:

Priority queue size = (Number of users) x (Maximum data rate of video terminals) x 120%

The priority queue must be provisioned to handle the maximum data rate used by any of the video terminals, otherwise the priority queue has the potential to become oversubscribed. Add 20% to the maximum data rate of the video terminals to allow for IP and transport overhead. The priority queue for the link should not exceed 33% of the link capacity. Refer to the WAN QoS chapter for more information.

Priority Queuing on the WAN

Configure multiple queues for the WAN ports on routers. Videoconferencing traffic goes into a priority queue (PQ) that services IP Precedence 4 or DSCP AF41. Class-based weighted fair queuing (CBWFQ) is *not* recommended for interactive video.

Entrance Criteria

In the single-zone WAN model, use access control lists (ACLs) to access configured priority queues at remote sites. ACLs ensure that only traffic from the video terminals has access to the configure PQ. The small number of video terminals at remote sites makes ACL entries a viable option.

Central sites that have either Layer 3 switches or video terminals capable of setting IP Precedence, should set the entrance criteria for the PQ to any packets with IP Precedence 4 or DSCP AF41. This method, however, is not as secure as the ACL option but works properly if the trust boundaries are configured correctly.   This method can also be used at remote sites if ACLs are not acceptable.

Multi-Zone WAN

A multi-zone WAN model consists of the WAN environment and three or more videoconferencing terminals per remote site. (This model is based on a T1 WAN link.) Multi-zone WAN deployments are found in large enterprises and state-based distance-learning networks. Remote sites containing three or more video terminals are managed by either a centralized or local gatekeeper. The gatekeeper manages bandwidth within the local zone and across the WAN between zones.

The gatekeeper manages bandwidth only in a hub-and-spoke environment with bandwidth controls. An intermediate gatekeeper is not aware of a call passing through its zone. Only the originating zone gatekeeper and terminating zone gatekeeper are aware of the active call. Resource Reservation Protocol (RSVP) can be used in conjunction with Differentiated Services Code Point (DSCP) to scale larger than hub-and-spoke environments. For this configuration, you need to consider the impact of other applications such as IP telephony that monitor only the number of calls or bandwidth usage for call admission control. See the section on Interworking with Session Initiation Protocol (SIP), page A-1, for more information.

Figure 4-3 illustrates a multi-zone WAN network.

Figure 4-3 Multi-Zone WAN

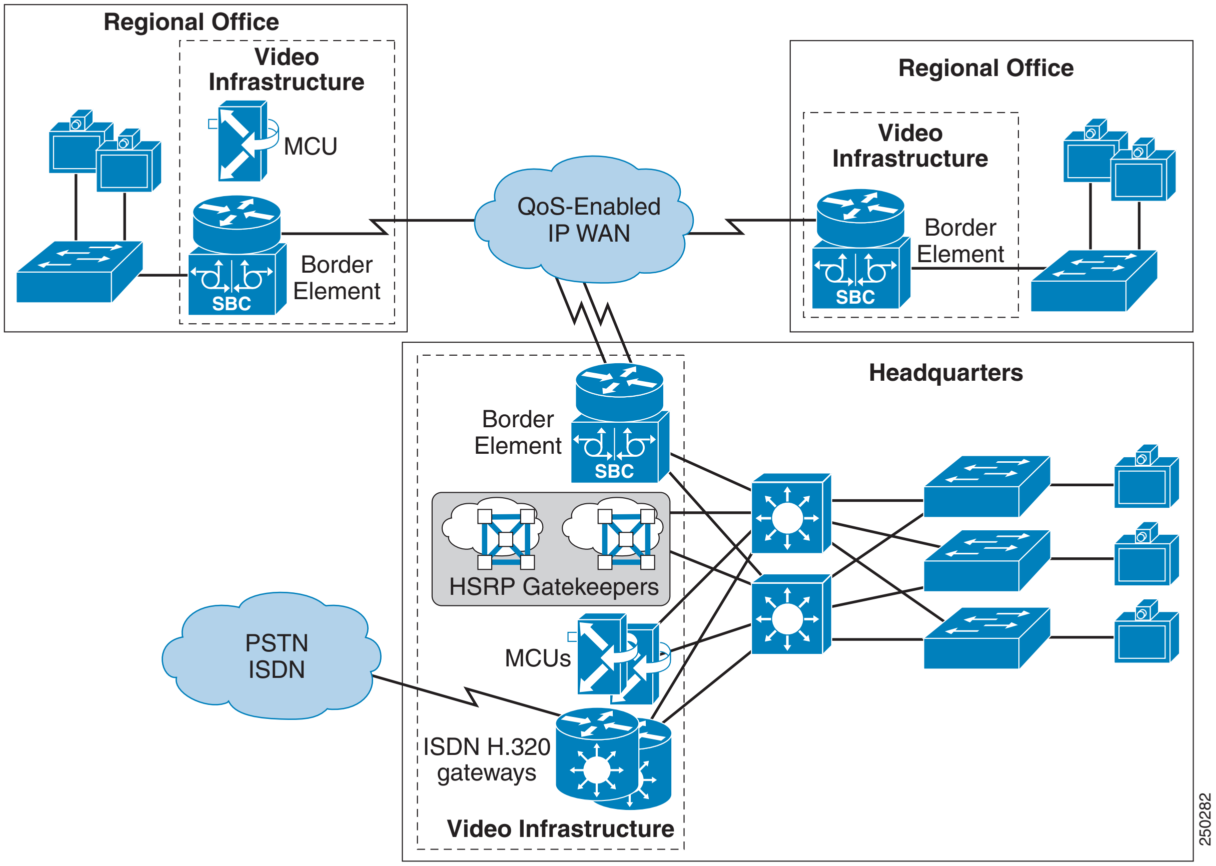


Figure 4-3 shows each remote site running the border element on the WAN router, and dedicated routers with Hot Standby Routing Protocol (HSRP) for the gatekeeper at the central site. You can deploy border elements with the gatekeeper enabled to support a large number of video endpoints and have flexibility of registration and call resolution in the event of a WAN failure.

The deployment guidelines for a multi-zone WAN environment are similar to those for a single-zone WAN. The biggest difference is the ability to control bandwidth in the multi-zone WAN through an added classification point (gatekeeper and zone). The key elements for successful deployment of videoconferencing in a multi-zone WAN environment are:

- Traffic Classification, page 4-6
- Bandwidth Control and Call Admission Control, page 4-6

- Provisioning, page 4-7
- Priority Queuing on the WAN, page 4-7
- Entrance Criteria, page 4-7

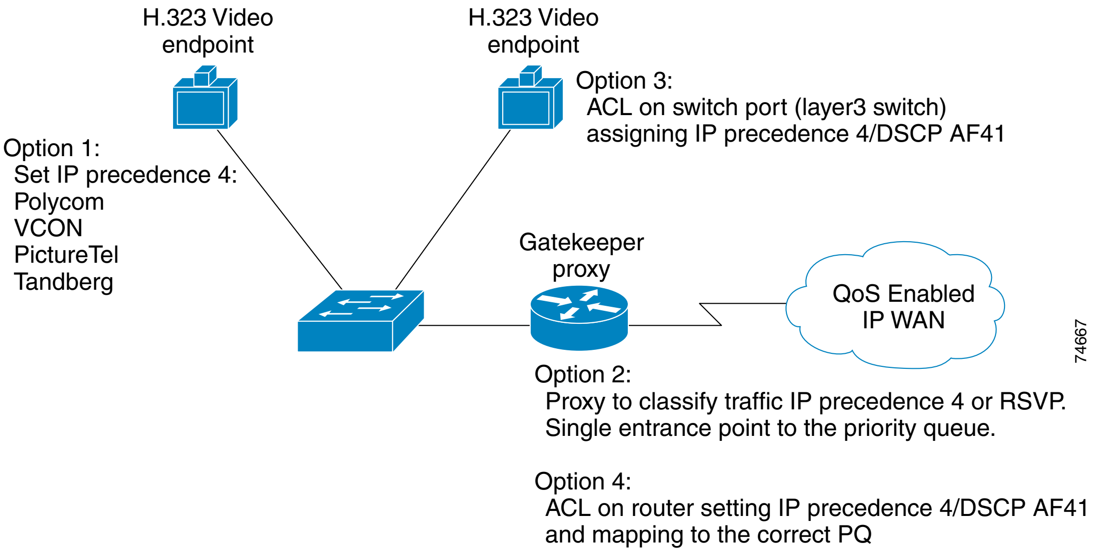
Traffic Classification

Classify traffic at one of the following places:

- Video endpoint (Polycom, Tandberg, Sony, and others); IP Precedence 4 or DSCP AF41
- Border Element classification; IP Precedence 4 or RSVP (recommended for traffic reclassification)
- Switch port (Layer 3 switch required); IP Precedence 4 or DSCP AF41 (recommended classification method for all video endpoints)
- Router (ACL entry); IP Precedence 4 or DSCP AF41 (Due to the larger number of video terminals at each site, this option is not typically used.)

Figure 4-4 illustrates classification options for a multi-zone WAN model.

Figure 4-4 Traffic Classification Options for Multi-Zone WAN


H.323 Video endpoint
H.323 Video endpoint
Option 3:
ACL on switch port (layer3 switch)
assigning IP precedence 4/DSCP AF41
Option 1:
Set IP precedence 4:
Polycom
VCON
PictureTel
Tandberg
Gatekeeper proxy
QoS Enabled IP WAN
74667
Option 2:
Proxy to classify traffic IP precedence 4 or RSVP.
Single entrance point to the priority queue.
Option 4:
ACL on router setting IP precedence 4/DSCP AF41
and mapping to the correct PQ

Bandwidth Control and Call Admission Control

Because each remote site in a multi-zone WAN has its own gatekeeper and zone, bandwidth control between zones is possible. By configuring the *remote* bandwidth in each remote gatekeeper, administrators can limit the amount of available bandwidth for calls to and from the WAN. Use the global **bandwidth remote** command at remote sites to control video calls across WAN links. For more information on the gatekeeper and bandwidth control, refer to the chapter on Cisco Video Infrastructure Components.

Provisioning

Provision WAN queues based on the bandwidth limits set in the gatekeeper, and do not provision more than 33% of the link capacity for voice and video applications. Cisco recommends that you provision the Priority Queue (voice and video traffic combined) to use no more than 33% of the link capacity.

Priority Queuing on the WAN

Configure multiple queues for WAN ports on routers. Videoconferencing traffic goes into a PQ that services the devices. If the border element is used, it can mark streams with IP Precedence 4 or DSCP AF41.

Entrance Criteria

Using the border element allows administrators to limit access to the priority queue by configuring an ACL on the WAN router. Only video calls authenticated by the gatekeeper go through the border element if configured appropriately. The ACL allows only packets received from the border element to access the configured priority queue. The ACL prevents unauthorized users from installing a video terminal on their desk, making video calls using IP addresses, and accessing the priority queue. By restricting access to the priority queue, the configured ACL ensures that unauthorized users cannot oversubscribe the priority queue. Rogue users are serviced out of the default queue, thus ensuring video quality for authorized video terminals.

If the border element is not used, the entrance criteria for the priority queue should be any packets with IP Precedence set to 4 or DSCP AF41. ACLs can be used if static addressing is deployed for the video endpoints. It is important to configure trust boundaries properly to prevent unauthorized traffic from accessing the priority queue.

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.

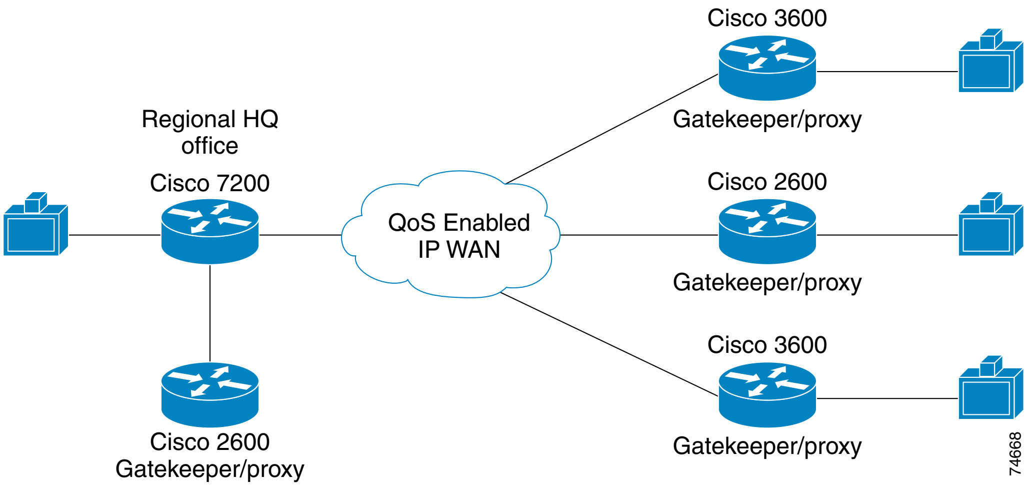
Table 5-1 New or Changed Information Since the Previous Release of This Document

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.

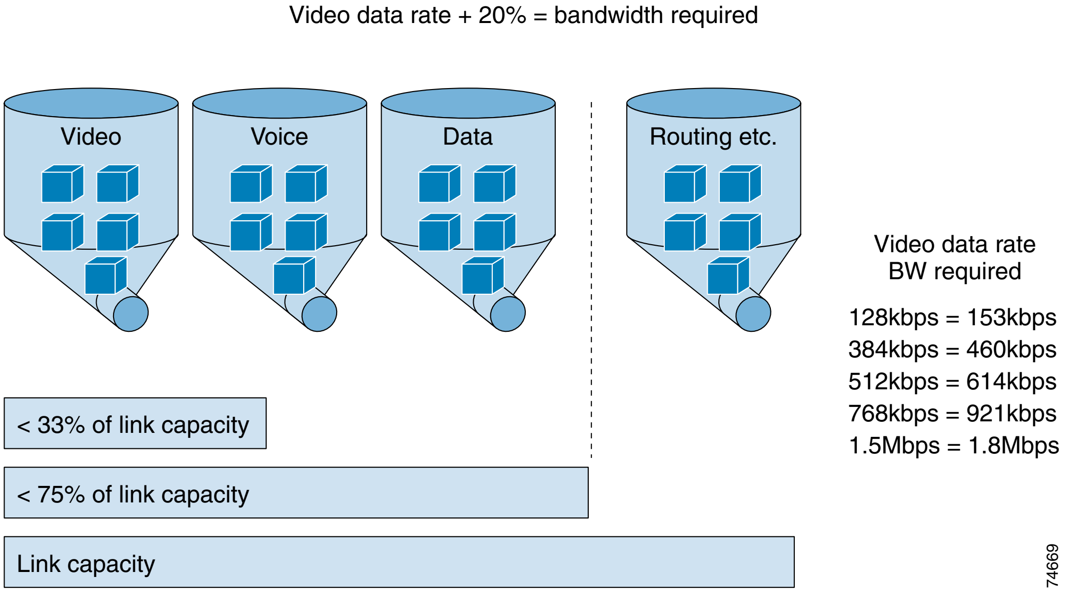
Figure 5-1 Enterprise WAN Model



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

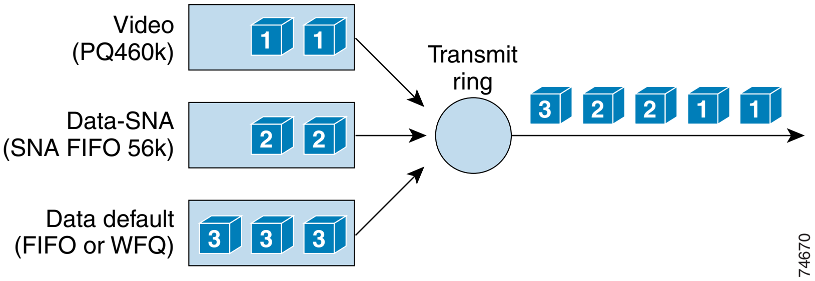


Table 5-2 gives the minimum bandwidth requirements for video and data networks. Note these values are minimum, and any network should be engineered with adequate capacity for all the applications that will use it.

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.

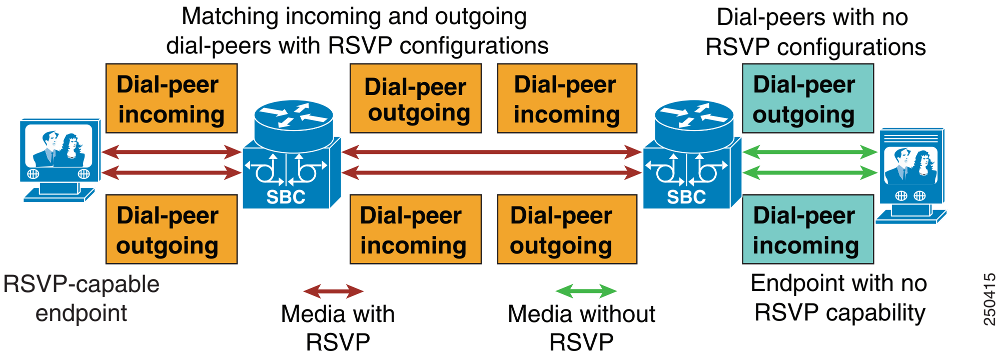
There are two options available for implementing RSVP:

- 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

CHAPTER 6

Dial Plan Architecture

Last revised on: October 30, 2009

This chapter defines and explains the key elements in designing a dial plan for an H.323 network. An H.323 video dial plan is a numbering scheme that allows H.323 video endpoints to dial other video endpoints or video services (multipoint control unit or gateway). This chapter discusses each of these components in the context of a single-zone or multi-zone scenario.

This chapter contains the following sections:

- Dial Plan Components, page 6-2
- Service Prefix Design, page 6-3
- Single-Zone Dial Plan, page 6-5
- Zone Prefix Design, page 6-7
- Multi-Zone Dial Plan, page 6-8
- Dial Plan for External IP Connectivity, page 6-11

What's New in This Chapter

Table 6-1 lists the topics that are new in this chapter or that have changed significantly from previous releases of this document.

Table 6-1 New or Changed Information Since the Previous Release of This Document

New or Revised Topic	Described in:
Dial plan for connecting to external entities	Dial Plan for External IP Connectivity, page 6-11

Dial Plan Components

A well designed dial plan is a key component of a successful H.323 video network, and it is one of the first things you need to consider when designing an H.323 video network. Without a well constructed dial plan, it is impossible to scale the network.

H.323 dial plans consist of four key elements:

- E.164 address

  An E.164 address is a numeric identifier defined on each H.323 video endpoint, just as E.164 is used in telephony systems.

- H.323-ID

  An H.323-ID is an alphanumeric identifier defined on each H.323 video endpoint, and it can be used to dial the H.323 endpoint. An alias may also be used to refer to an H.323-ID. For example, email addresses are often used as H.323-IDs. H.323-IDs cannot be used to dial to the PSTN or to a Cisco IP/VC 3510 multipoint control unit (MCU).

- Zone prefix

  A zone prefix is a numeric prefix that identifies a zone. Zone prefixes are used for inter-zone call routing, the same way an area code is used in a telephony system. Each zone in an H.323 network has one unique zone prefix. Area codes are often used as zone prefixes in H.323 networks.

- Service or technology prefix

  A service prefix is a numeric prefix used in an H.323 dialing string to access a defined service on an MCU or gateway. The Cisco gatekeeper refers to the service prefix as a technology prefix, which is also used by H.323 voice gateways. (This document refers to these prefixes as *service prefixes*.) Service prefixes are used on video gateways and MCUs to define parameter settings and to route calls for the device. On a Cisco Unified Videoconferencing 352x video gateway, a service prefix defines the type of call being made (voice or video) and the data rate of the call. On a Cisco Unified Videoconferencing 351x and 3545 MCUs, service prefixes define the data rate of the call, number of participants, and picture format. When an MCU or video gateway registers with the gatekeeper, it registers all defined service prefixes. When an H.323 endpoint uses a video gateway or MCU, the dial string must start with the service prefix followed by the PSTN number being dialed (in the case of a gateway call) or the conference ID being created or joined (in the case of an MCU call).

Table 6-2 shows the correlation between components of a video and IP telephony dial string.

Table 6-2 Correlations Between Video and IP Telephony Dial Stings

Video Dial String	IP Telephony Dial String
Service prefix	Technology prefix
Zone prefix	Area code
E.164 address	Local exchange Unit ID

Service Prefix Design

Service prefixes are a very important part of the dial plan. Inter-zone and intra-zone calls to an MCU or gateway are routed using the service prefix. The single-zone and multi-zone models are very similar, and both are discussed in this section, with minor differences between them noted.

It is important to keep dial strings intuitive. For example, the models in this section use dial strings that are very similar to telephony dial strings. Dial strings are reviewed in the chapter on Call Routing.

In a single-zone network, Cisco recommends that you reserve a block of numbers for service prefixes, such as 8* for MCUs and 9* for gateways.

Note The asterisk is a wildcard that represents any dialed digits. For example, the string 8* represents any dialed string beginning with the digit 8 followed by any number. Users do not dial the asterisk (*) when placing a call.

Cisco also recommends that you add the local area code to the service prefixes of MCUs. For example, a San Jose MCU might have a service prefix of 40880. Gateway prefixes should remain 9* to keep dial strings consistent with telephony dial plans. This service prefix structure also allows an easy migration to a multi-zone dial plan.

E.164 addresses must not overlap with service prefixes. For example, if an MCU registers with a service prefix of 40880* and a video terminal registers with 4088011212, all calls made to the video terminal would be routed to the MCU instead.

In a multi-zone network, service prefixes need to route between zones. Therefore, all service prefixes must be unique across all zones. All inter-zone or intra-zone calls are routed based on the service prefix. Cisco recommends that you design service prefixes in a multi-zone network to allow user dial strings to be consistent. To achieve this consistency, use the different approaches outlined in the following sections for service prefixes on MCUs and gateways. Service prefixes, E.164 addresses, and zone prefixes must not overlap, or call routing issues will arise.

MCU Service Prefixes

MCUs must be accessible from any H.323 endpoint on the network, which means that all service prefixes in all zones must be unique. In order to accommodate unique service prefixes without reserving large blocks of numbers, Cisco recommend that you design the MCU service prefixes to be a combination of the zone prefix and a service number. This design allows all the service prefixes for MCUs to be consistent in all zones.

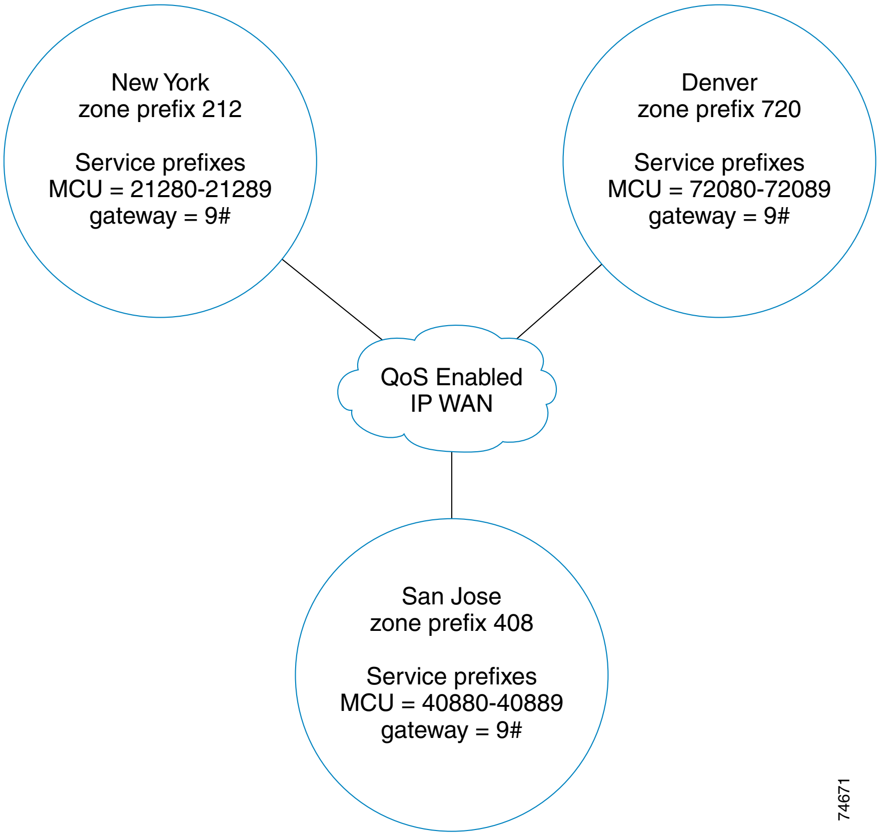
For example, if the reserved block of numbers is 8*, the service prefix for a 384 kbps call with five users could be 40880 in the 408 zone and 41580 in the 415 zone. The dial string for a 384 kbps conference call in zone 408 would be 40880<*conference ID*>. This design eliminates the need for hopoff entries, and it associates the service with the zone in which the service resides. (For more information on hopoffs, see the Call Routing chapter.)

Gateway Service Prefixes

Gateway services in a multi-zone network are similar to those in the single-zone model. Reserve a block of numbers for gateway services. In zones that contain gateways, off net calls always use the local gateway. For zones without a gateway, add a hopoff entry or use location request (LRQ) forwarding to route the call to a zone containing a gateway. (See the Call Routing chapter for more information regarding hopoffs, LRQ forwarding, and directory gatekeeper.)

For example, if the reserved block of numbers is 9*, the gateway service configured for all outbound calls could be 9#. Configure these service prefixes on all gateways in all zones. In zones that have a zone prefix starting with 9, ensure that the zone prefix and gateway service prefixes do not overlap. For example, if the zone prefix is 916, a gateway service prefix of 9 cannot be used in that zone, otherwise all calls in the zone would be routed to the gateway. To avoid this problem, Cisco recommends that you configure the gateway service prefixes to include a # sign, such as 9#. Figure 6-1 illustrates service prefix design in a multi-zone network.

Figure 6-1 Service Prefix Design in Multi-Zone Networks



Note Use a # in the service prefix for gateways to ensure that calls coming in from the PSTN network do not have the ability to hair-pin back onto the PSTN through the gateway. When a user dials a # from the PSTN to the gateway, the # is treated as a delimiter and the call fails.

Single-Zone Dial Plan

Dial plans for single-zone networks are straightforward. There are a few rules that you must follow to ensure that call routing in a single zone works properly. When developing a dial plan for a single-zone network, use the following components and guidelines:

- Incoming PSTN call routing method

  As a general rule, the incoming PSTN routing method is a good place to start when designing a dial plan because it determines the number strings and the E.164 numbering structure used in the dial plan. If Direct Inward Dialing (DID) is used, each H.323 endpoint is assigned a valid E.164 directory number (DN). If interactive voice response (IVR) or TCS4 is used, the administrator can choose the E.164 number structure. Cisco recommends using 10-digit numbers for E.164 addresses because 10-digit numbers allow for an easy migration to a multi-zone dial plan. (Endpoints should be configured with a local extension, and that extension plus the zone prefix together should consist of 10 digits.) Incoming PSTN routing methods, DID, IVR, and TCS4 are covered in detail in the Call Routing chapter.

- Service prefixes

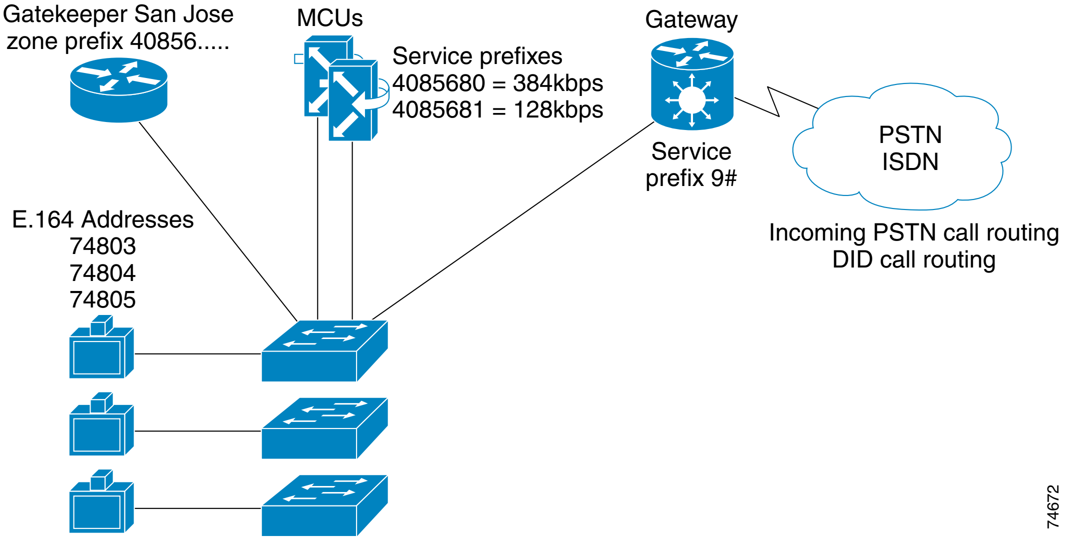
  Services prefixes must not overlap with E.164 addresses; therefore, it is a good idea to reserve a block of numbers for service prefixes. In Figure 6-2, the reserved block of numbers is 8* for MCUs, the zone prefix is 40856, and the two service prefixes for the MCU are 4085680 and 4085681. Gateway services are 9# and do not include the zone prefix.

- H.323-IDs

  H.323-IDs are alphanumeric strings used to identify an H.323 terminal. H.323-IDs are often email addresses of individual users or conference room names for room systems. Using H.323-IDs to place calls is intuitive, as long as the user-to-endpoint mapping is static. Some H.323 room systems are used in multiple conference rooms, and naming these units can be a challenge.

Figure 6-2 illustrates a single-zone design for a campus network.

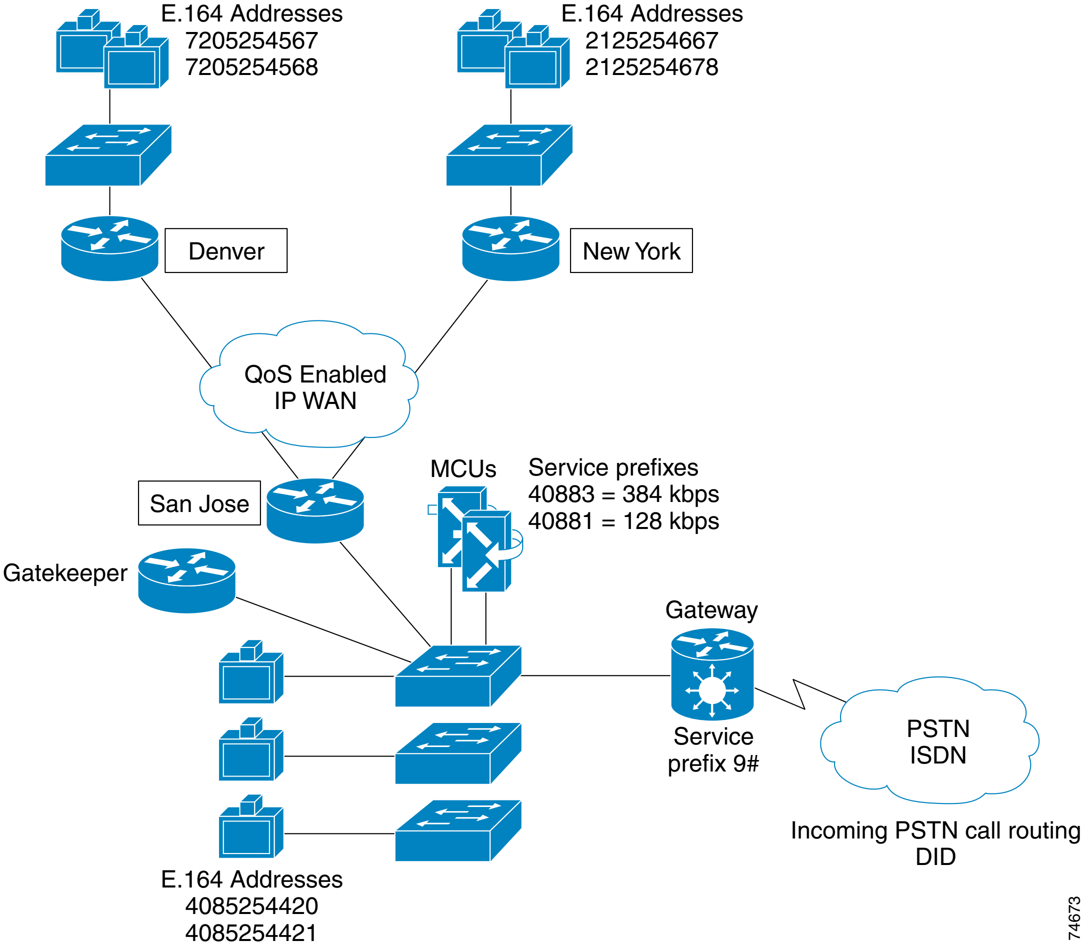
Figure 6-2 Single-Zone Configuration for a Campus Network



When creating a dial plan for a single zone in a WAN environment, it is always a good idea to use a numbering scheme that allows an easy migration to a multi-zone dial plan. For this purpose, Cisco recommends that you use fully qualified E.164 addresses for the video terminals.

Figure 6-3 illustrates a single-zone WAN dial plan. All video terminals, gateways, and MCUs register in one zone and are routed according to the E.164 address, H.323-ID, or service prefix registered by each device.

Figure 6-3 Single-Zone WAN Dial Plan



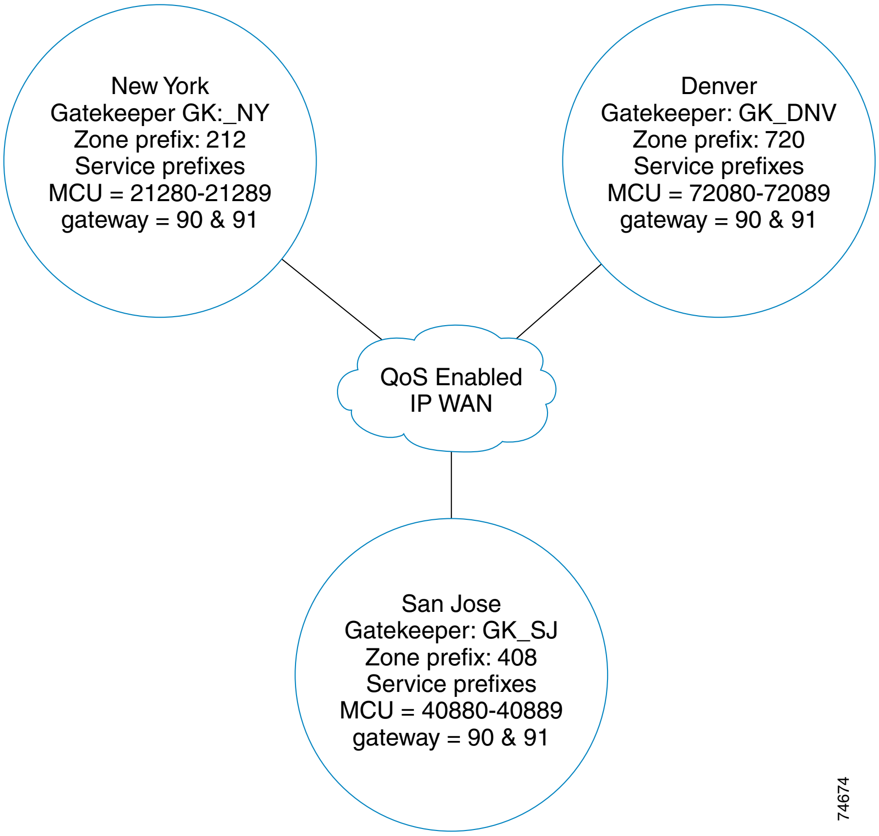
Zone Prefix Design

Zone prefixes are used in an H.323 video network to allow inter-zone call routing between H.323 endpoints, in the same way an area code is used in the PSTN. Each zone on the network must have a unique zone prefix that is used to identify the zone. Cisco recommends using the local area code for the zone prefix. For example, in Figure 6-4 there are three zones: San Jose campus zone 408*, New York 212*, and Denver 720*.

Zone prefixes can be configured with a wildcard (408*) or with dots (408…….). Cisco recommends that you use the dot method when configuring zone prefixes because this method lets you specify the exact number of digits to match, whereas the wildcard matches any number of digits. Zone prefixes can vary in length, and using more digits in the zone prefix reduces the number of available terminal addresses.

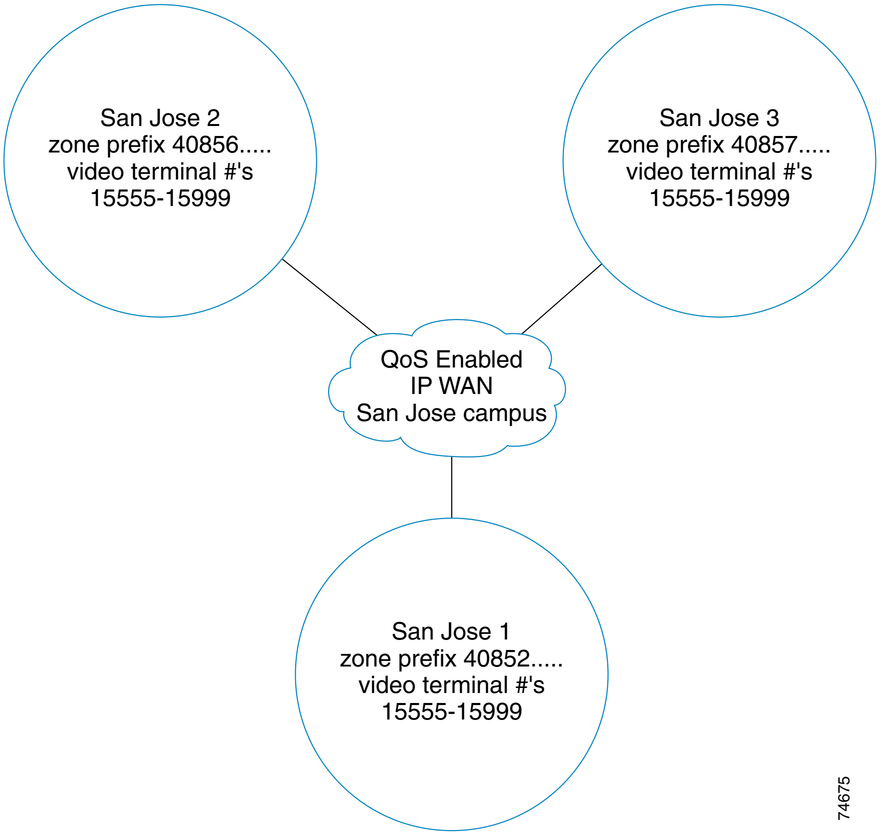
Note Never use a wildcard for the zone prefix of a Directory gatekeeper zone prefix. Doing so would cause all calls, including local calls, to be forwarded to the Directory gatekeeper. Instead, use the dot (.) method to specify the zone prefix of the Directory gatekeeper. For more information on using the Directory gatekeeper, refer to the chapter on Call Routing.

Figure 6-4 Example Network with Unique Zone Prefixes



Large sites that need more than one zone can still use the local area code and expand the zone prefix to include some of the E.164 address. For example, the San Jose campus in Figure 6-5 has three zones: one configured as (40852…..), the second as (40856…..), and the third as (40857…..). Video terminals can then register with five-digit extensions, allowing extension-based dialing within the local zone, but 10-digit dialing is still required between zones.

Figure 6-5 Using a Single Area Code for Multiple Zones



Multi-Zone Dial Plan

Dial plans for multi-zone networks have the added complexity of zone prefixes and inter-zone call routing. When developing a dial plan for a multi-zone network, consider the following components and guidelines:

- Incoming PSTN Call Routing

  Again, it is a good idea to start with the incoming PSTN routing method when developing the dial plan because the routing method determines which E.164 numbering structure is used in the dial plan. Unless there is at least one gateway in each PSTN area code, direct inward dialing (DID) is not recommended for use as the primary incoming PSTN routing method in a multi-zone network because the DID number is within one area code but the remote zone prefix may be in a different area code.

Rather than configuring your remote zone prefixes to match the area code, which would confuse the dial plan, Cisco recommends that you place a gateway in each area code. It is important that you order enough DID numbers for all zones located in the area code serviced by the gateway. Because the Cisco gatekeeper does not support digit manipulation, it is very difficult to route incoming DID calls between zones. There are cases in a multi-zone network where you might want to use a mix of incoming call routing methods; for example, you could use DID for endpoints but use IVR for MCU meet-me conferences.

If interactive voice response (IVR) or TCS4 is used, the administrator can choose the E.164 number structure. Cisco recommends using 10-digit numbers for E.164 addresses because 10-digit numbers allow for an easy migration to a multi-zone dial plan. (Endpoints should be configured with a local extension, and that extension plus the zone prefix together should consist of 10 digits.) Incoming PSTN routing methods, DID, IVR, and TCS4 are covered in detail in the Call Routing chapter.

- Service Prefixes

  Service prefixes must not overlap with E.164 addresses; therefore, it is a good idea to reserve a block of numbers for service prefixes. When a range of numbers is reserved for MCUs (for example, 8*), append the zone prefix to the reserved number to create a unique service prefix. For example, if the zone is 408 and the reserved block of numbers is 8*, the first service prefix might be 40880. In this case, an H.323 endpoint may not register with an E.164 address that starts with 40880-40889. If an MCU registers with a service prefix of 40880 in the zone and an H.323 endpoint registers with 4088012, all calls to 4088012 would be routed to the MCU.

- Zone Prefixes

  Zone prefixes are also very important in the development of the dial plan. Zone prefixes are much like area codes in a telephony system. Cisco recommends that you use local area codes for zone prefixes because area codes are unique, already defined, and people are familiar with them. Again, it is up to the administrator to choose the zone prefixes, but it is also important that the prefixes be intuitive and capable of growing with the network. Zone prefixes must not overlap with service prefixes, otherwise call routing issues will arise. (If you use a zone prefix plus a service prefix for MCUs, overlap with MCUs will not be an issue.) See Service Prefix Design, page 6-3, for details.

- H.323-IDs

  H.323-IDs are alphanumeric strings used to identify an H.323 terminal. H.323-IDs are often email addresses of individual users or conference room names for room systems. Using H.323-IDs to place calls is intuitive, as long as the user-to-endpoint mapping is static. Some H.323 room systems are used in multiple conference rooms, and naming these units can be a challenge.

If IVR is the chosen method for incoming PSTN call routing, observe the following guidelines:

- All systems dialing in from the PSTN must support Dual Tone Multi-Frequency (DTMF).
- Implement a private numbering plan.

If DID is the chosen method for incoming PSTN call routing, observe the following guidelines:

- Gateways must reside in each area code for zone prefix consistency.
- Use IVR to route MCU calls.

Figure 6-6 illustrates a multi-zone design using IVR, and Figure 6-7 illustrates a multi-zone design using DID.

Figure 6-6 Using IVR in a Multi-Zone Configuration

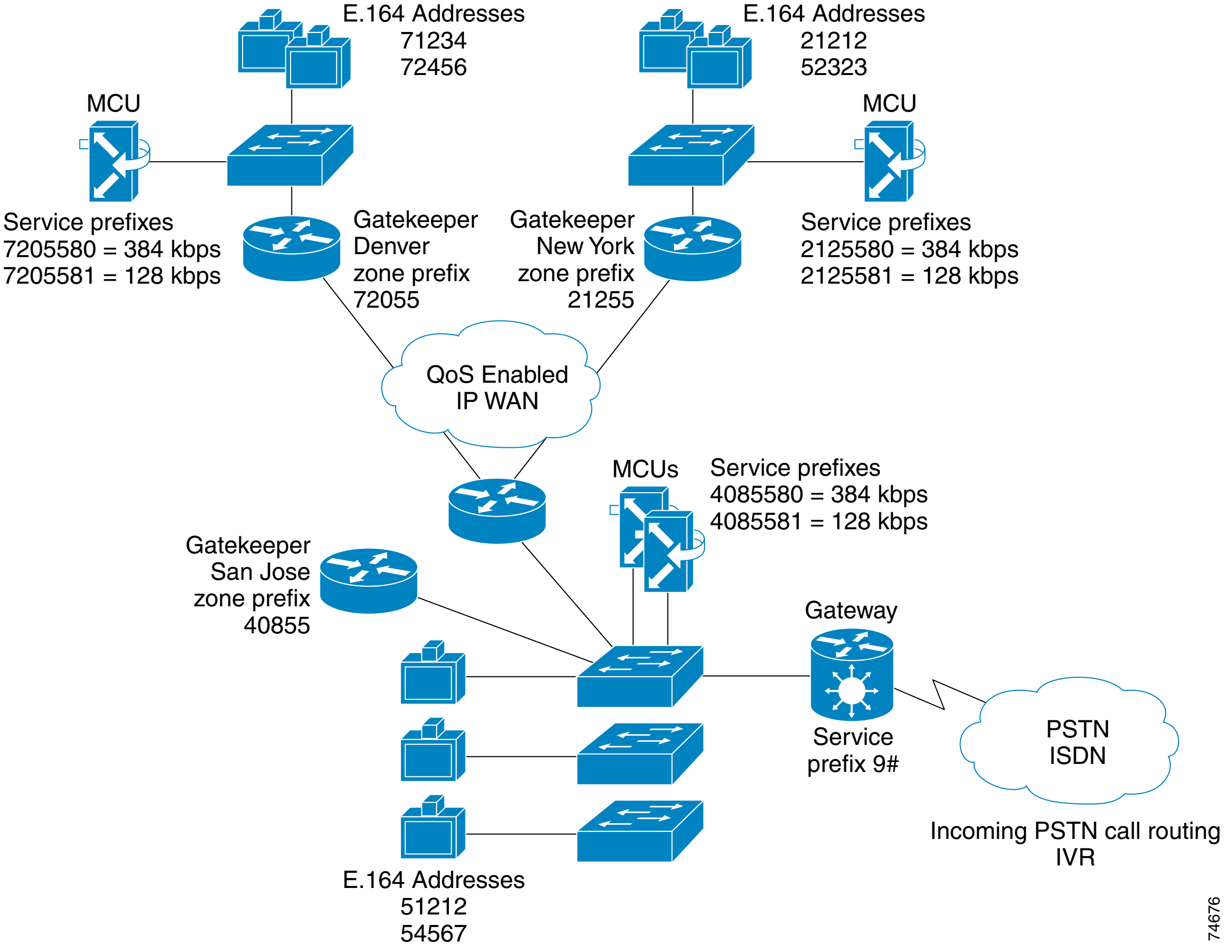
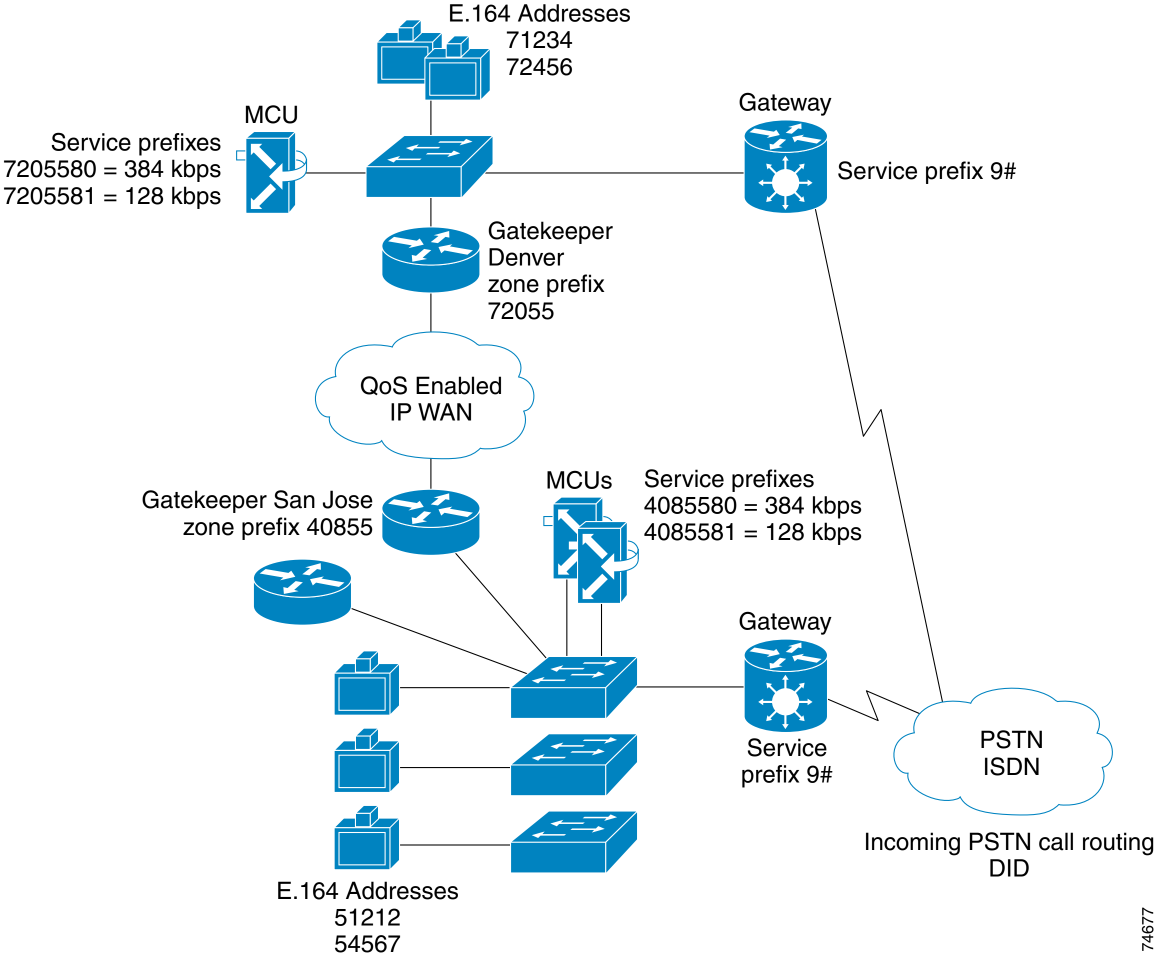


Figure 6-7 Using DID in a Multi-Zone Configuration



Note In Figure 6-7, the IVR is still enabled for access from the PSTN to the MCUs.

Dial Plan for External IP Connectivity

The enterprise dial plan will likely need to reach external entities such as vendors, suppliers, or other businesses. Connecting to them can be done using the PSTN, as discussed above, but connectivity on the IP network presents challenges for the dial plan. Cisco recommends that enterprises use E.164 addresses for the dial plan between organizations. This provides a non-overlapping unique E.164 address that can be configured easily in each enterprise. Enterprises then can use number translations to normalize the addresses to its internal dial plan.

When connecting numerous independent external units such as sales personnel or small remote locations over untrusted networks, Cisco strongly recommends making them part of the internal network for registration so that their dial plan can be controlled by the enterprise.

CHAPTER 7

Call Routing

Last revised on: October 30, 2009

This chapter describes various call routing methods used by Cisco gatekeeper and Cisco Unified Videoconferencing equipment in an H.323 video network. Calls can be routed to and from many types of devices in a variety of ways.

What's New in This Chapter

Table 7-1 lists the topics that are new in this chapter or that have changed significantly from previous releases of this document.

Table 7-1 New or Changed Information Since the Previous Release of This Document

New or Revised Topic	Described in:
Routing calls through untrusted networks	Routing Calls Through Untrusted Networks, page 7-17
Routing calls to other networks	Routing Calls to Other Networks, page 7-17

Call Routing Scenarios

There are four possible call routing scenarios in an H.323 network:

- H.323 endpoint to H.323 endpoint using the E.164 address

  Routing calls between H.323 endpoints is the simplest type of call routing in an H.323 network. To dial within a single zone, the endpoint initiating the call enters the E.164 address of the endpoint being called. (In most cases, the E.164 address is a video terminal extension). If the call is an inter-zone call, the initiator must enter the zone prefix and terminal extension. Using this type of dial string is similar to dialing outside the local area code in a telephony system. In multi-zone networks, service prefixes for Multipoint Control Units (MCUs) should contain the zone prefix.

- H.323 endpoint to H.323 endpoint using H.323-ID

  To use the H.323-ID to route calls between H.323 endpoints, the calling station must dial the H.323-ID of the video terminal being called. H.323-IDs are supported only for calls from video terminal to video terminal or from video terminal to Video Terminal Adapter (VTA). When using a VTA, exercise care in addressing because some H.320 units cannot send alphanumeric strings. In these cases, E.164 addresses are the only usable route table mechanism. Between zones, Domain

Name Service (DNS) may be used to reach the H.323-ID of an endpoint registered to a remote gatekeeper. To use DNS, the calling station dials *H.323-ID@Domain*, which allows the gatekeepers to resolve the remote zone destination using DNS.

- H.323 endpoint to an H.323 service (gateway or MCU)

  Routing calls from an H.323 endpoint to a service is also simple. In a single zone, an H.323 endpoint dials the service prefix followed by either the conference ID (for an MCU call) or the Integrated Services Digital Network (ISDN) telephone number of the H.320 endpoint. The service prefix can also route inter-zone calls to services, but in this case the service prefix contains the zone prefix for the MCUs, and the inter-zone calls use hopoffs for gateways.

- Incoming PSTN to H.323 endpoint or service

  You can use any of the following methods to route calls from the Public Switched Telephone Network (PSTN) to H.323 endpoints or services:

  – Multiple Subscriber Number (MSN) and Direct Inward Dialing (DID)

  – Interactive Voice Response (IVR)

  – TCS4

  – Default extension

  For more details on these routing methods, see Routing PSTN Calls to H.323, page 7-4.

Example

Figure 7-1 illustrates a multi-zone network with video terminals and services in each zone. The following dial strings apply to the scenarios in Figure 7-1:

- H.323 endpoint to H.323 endpoint:

  Intra-zone call, User1 to User2 — User1 dials 4085558072

  Inter-zone call, User1 to User3 — User1 dials 7207125543

- H.323 endpoint to H.323 endpoint using H323-ID:

  Intra-zone call, User1 to User2 — User1 dials User2@cisco.com

  Inter-zone call, User1 to User3 — User1 dials User3@cisco.com

- H.323 endpoint to service:

  Intra-zone call, User1 to H.320 system — User1 dials 9#12125551212

  Inter-zone call, User3 to H.320 system — User3 dials 9#12125551212

  Gateway calls always use the local gateway if one is present.

- PSTN endpoint to H.323 endpoint or service using IVR (see Routing PSTN Calls to H.323, page 7-4):

  Intra-zone call, H.320 system to User1 — H.320 system dials 4085552000, IVR answers, and the H.320 system enters 40855558071. Or, if DID is enabled to User1, H.320 system dials 4085558071 directly.

  Intra-zone call, H.320 system to 408, and MCU conferences 40880123 — H.320 system dials 4085552000, IVR answers, and the H.320 system enters 40880123. (For DID and IVR deployments, IVR should be used for routing calls to an MCU conference.)

  Intra-zone call, H.320 system to User3 — H.320 system dials 4085552000 followed by 7207125543. (If there is a gateway in the 720 area code, DID could be enabled to User3 and IVR could be used to reach User3 in the 720 zone instead of having to dial the 408 gateway.)

Inter-zone call, H.320 system to MCU, with conference to 72080111 — H.320 system dials 408555200 followed by 72080111. (For DID and IVR deployments, IVR should be used for routing calls to an MCU conference.)

Figure 7-1 Call Scenarios in an Example Multi-Zone Network

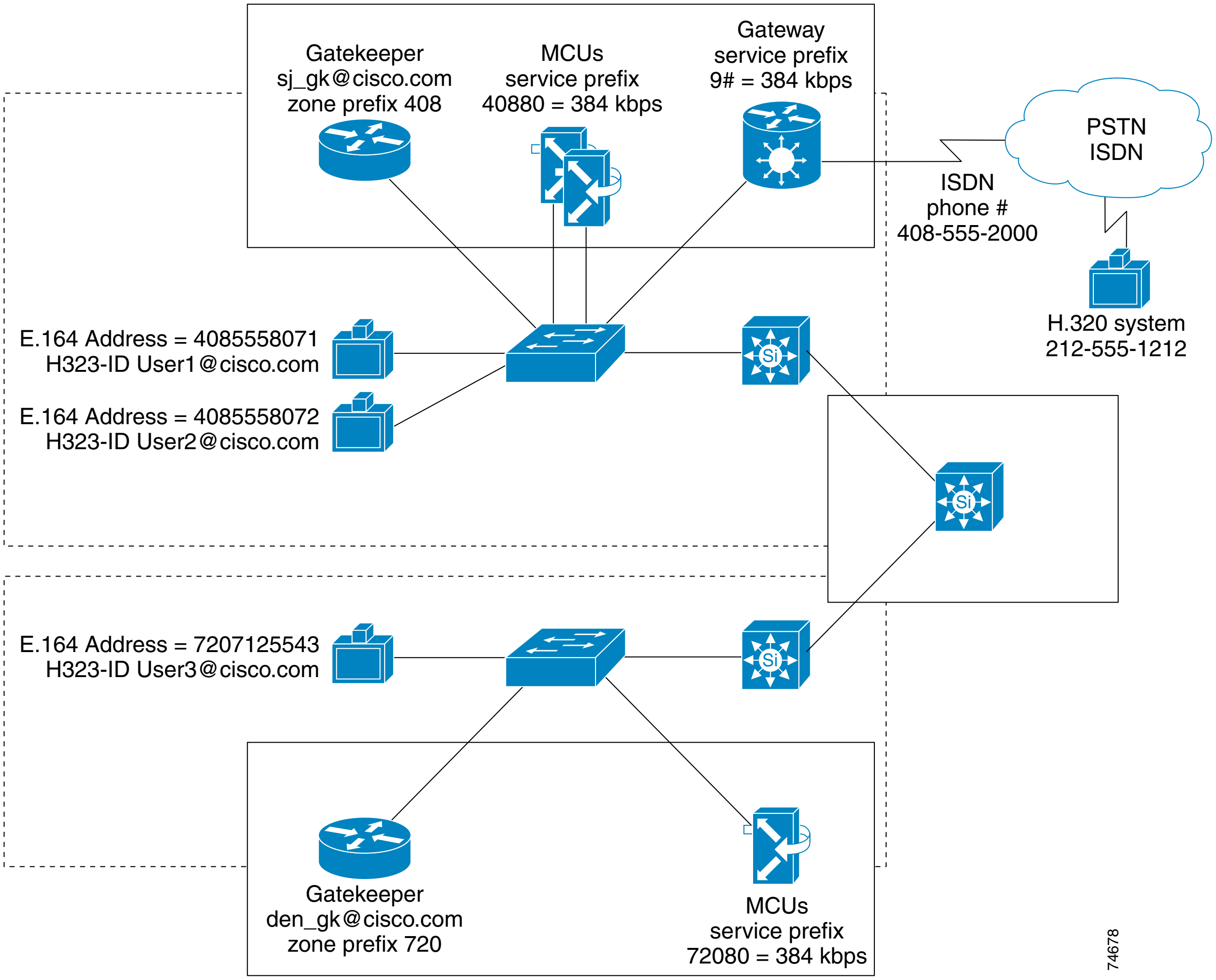


Table 7-2 shows the dial strings for the intra-zone call types and Table 7-3 shows the dial strings for the inter-zone call types in Figure 7-1.

Table 7-2 Dial Strings for Intra-Zone Calls

Call from:	Call to:	Dial String
H.323 Endpoint	H.323 Endpoint	*<E.164 address>* or *<H.323-ID>*
H.323 Endpoint	Service	*<Service Prefix> <Conference ID or PSTN E.164 address>*
Service	H.323 Endpoint	*<E.164 address>*
Service	Service	*<Service Prefix> <Conference ID or PSTN E.164 address>*

Table 7-3 Dial Strings for Inter-Zone Calls

Call from:	Call to:	Dial Sting
H.323 Endpoint	H.323 Endpoint	*<Zone prefix + E.164 address>* or *<H.323-ID>*
H.323 Endpoint	Service	*<Zone Prefix and/or Service Prefix> <Conference ID or PSTN E.164 address>*
Service	H.323 Endpoint	*<Zone Prefix and/or E.164 address>*
Service	Service	*<Zone Prefix and/or Service Prefix> <Conference ID or PSTN E.164 address>*

Routing PSTN Calls to H.323

There are several methods for routing calls from the PSTN to H.323 endpoints and services:

- Multiple Subscriber Numbering (MSN) with Direct Inward Dialing (DID), page 7-4
- Interactive Voice Response (IVR), page 7-4
- TCS4, page 7-4
- Default Extension, page 7-5

An H.323 video network can use one or more of these available routing methods, and each routing method has advantages over the others in different situations.

Multiple Subscriber Numbering (MSN) with Direct Inward Dialing (DID)

Multiple Subscriber Numbering (MSN) is a group of phone numbers assigned to a single ISDN Basic Rate Interface (BRI) line. MSN is not available in most regions of the United States, Canada, or South America, but it is widespread in Europe.

Direct Inward Dialing (DID) is supported on Primary Rate Interface (PRI) lines. DID allows multiple directory numbers to be assigned to a single PRI line. DID is supported throughout the United States and Europe.

Interactive Voice Response (IVR)

IVR is a widely deployed automated call answering system that responds with a voice menu, allowing the H.320 endpoint to access H.323 endpoints by entering an extension from a keypad. When an incoming call arrives, the IVR answers the call and asks for the extension. The caller enters an E.164 address, and the call is transferred to the appropriate H.323 endpoint. Using IVR requires the calling H.320 endpoint to support Dual Tone Multi-Frequency (DTMF). Most legacy conference room systems support DTMF.

TCS4

TCS4 is a special method for routing incoming H.320 video calls by using extensions. TCS4 allows direct extension dialing to an H.323 endpoint on the LAN, which register to the gatekeeper with an E.164 address. When an H.320 endpoint dials a gateway's phone number followed by a TCS4 delimiter and the E.164 address, the call is routed directly to the corresponding H.323 endpoint. TCS4 is new, and only some of the H.320 endpoints permit the user to enter a TCS4 extension when dialing. Due to the limited support for the TCS4 standard in H.320 devices, TCS4 is not frequently used for incoming call routing and, therefore, DID or IVR are typically better choices.

Default Extension

Specifying a default extension in the gateway forces all calls received by the video gateway to be routed directly to a default E.164 address. A default extension can also be used in conjunction with any of the routing methods mentioned previously. If the call cannot be routed by one of the previous methods, the call is then forwarded to the default E.164 address.

Routing Inbound PSTN Calls in a Single-Zone Network

You can use any of the routing methods to route calls from the PSTN to H.323 endpoints and services in a single-zone network. Each method offers the following functions and numbering structures:

- DID

  Using DID in a single-zone network allows administrators to order blocks of DID numbers and assign each H.323 endpoint a DID number to be used as its E.164 address. This method allows H.320 users and H.323 users to dial the same number to access an H.323 endpoint. (This method assumes that, in most cases, the carrier sends 10 digits.) DID can also be used for MCU conferences; however, in order to route calls to an MCU service in a zone, the zone prefix, the service prefix, and the conference ID combined must match one of the DID numbers associated with the ISDN line. This method disables the use of ad-hoc conference IDs created by the users on the MCU, but it may be preferable over using IVR to reach these conferences. This method does, however, require that the conference ID match the statically registered directory number. DID call routing is very desirable because the dial strings are exactly the same as those used in telephony systems, but routing H.323 service prefixes can become complex when using DID call routing.

- IVR

  IVR allows administrators to define the dial plan. E.164 addresses can be four-digit extensions or 10-digit directory numbers. (The video terminal extension and the zone prefix combined should be 10 digits). When routing incoming PSTN calls with IVR, the call initiator must dial the directory number of the PRI gateway and enter the E.164 address or service prefix dial string after the IVR has answered. IVR requires DTMF support on the dialing endpoint, but some older H.320 systems do not support DTMF.

- TCS4

  When using TCS4 to route incoming calls, the administrator again defines the numbering plan. When using TCS4, the initiator dials the directory number of the gateway, a TCS4 delimiter, and the E.164 address or service. The delimiter must be configured in each video gateway, and the options are # or *. Using TCS4 requires the dialing endpoint to support TCS4. TCS4 is not a commonly used routing method at the present time.

- Default Extension

  A default extension is usually used in special cases such as call center applications or to route calls to a single H.323 video terminal.

Note All of these dial-in methods are mutually exclusive, and you can implement multiple incoming routing methods on the same gateway. If an incoming PSTN call arrives at a gateway supporting all of the routing methods, the gateway first tries to resolve the address using the routing methods in this order: DID, IVR, TCS4, and default extension. (A DID environment is a typical example that would use a gateway supporting multiple incoming call routing methods.) Cisco recommends that you do not assign a DID number for ad-hoc MCU calls; instead, use IVR to route incoming calls to an MCU and use DID to route incoming calls to video terminals.

Table 7-4 summarizes partner product capabilities as they relate to interoperability with the Cisco Unified Videoconferencing gateways.

Note The information included in Table 7-4 is subject to change, and you should contact the product manufacturer directly for updated information.

Table 7-4 Partner Product Capabilities

Partner and Product	IP Bitrate	ISDN, PRI, or Serial Bitrate	Software Version
Polycom			
PVX			8.02
Viewstation FX	Up to 2 Mbps	Up to 512 kbps	6.05
Viewstation EX	Up to 2 Mbps	Up to 512 kbps	6.05
VS4000	Up to 2 Mbps	Up to 512 kbps	6.05
VSX3000	Up to 2 Mbps	Up to 512 kbps	8.7
VSX5000	Up to 768 kbps	Up to 512 kbps	8.7
VSX6000	Up to 768 kbps	Up to 512 kbps	8.7
VSX7000	Up to 2 Mbps	Up to 512 kbps	8.7
VSX7000e	Up to 2 Mbps	Up to 512 kbps	8.7
VSX7000s	Up to 2 Mbps	Up to 512 kbps	8.7
VSX8000	Up to 2 Mbps	Up to 512 kbps	8.7
Sony			
PCS-1	Up to 2 Mbps	Up to 768 kbps with additional hardware	3.41
PCS-TL30	Up to 2 Mbps	Up to 768 kbps with additional hardware	1.3
PCS-TL50	Up to 2 Mbps	Up to 768 kbps with additional hardware	2.41
PCS-G50	Up to 4 Mbps	Up to 2 Mbps with additional hardware	2.51
PCS-G70	Up to 4 Mbps	Up to 2 Mbps with additional hardware	2.51
Tandberg			
75MXP	Up to 2 Mbps	Up to 512 kbps	F6.1
85MXP	Up to 2 Mbps	Up to 512 kbps	F6.1
95MXP	Up to 2 Mbps	Up to 512 kbps	F6.1
150MXP	Up to 512 kbps		F6.1
770MXP	Up to 2 Mbps	Up to 512 kbps	F6.1
880MXP	Up to 2 Mbps	Up to 512 kbps	F6.1
990MXP	Up to 2 Mbps	Up to 512 kbps	F6.1

Table 7-4 Partner Product Capabilities (continued)

Partner and Product	IP Bitrate	ISDN, PRI, or Serial Bitrate	Software Version
1000MXP	Up to 768 kbps	Up to 384 kbps	F6.1
1700MXP	Up to 2 Mbps		F6.1
Aethra			
Vega Pro S	Up to 768 kbps	Up to 128 kbps	
Vega X3	Up to 2 Mbps with asymmetric rates	Up to 512 kbps	
Vega X5	Up to 4 Mbps with asymmetric rates	Up to 2 Mbps with additional hardware	
Vega X7	Up to 4 Mbps with asymmetric rates	Up to 512 kbps	
Nova Entry X50	Up to 4 Mbps with asymmetric rates	Up to 768 kbps with additional hardware	
Dual Nova X50	Up to 4 Mbps with asymmetric rates	Up to 768 kbps with additional hardware	
Supernova X150	Up to 4 Mbps with asymmetric rates	Up to 2 Mbps with additional hardware	
Supernova Xline	Up to 4 Mbps with asymmetric rates	Up to 2 Mbps with additional hardware	

Routing Inbound PSTN Calls in a Multi-Zone Network

Call routing in a multi-zone network becomes more complicated due to the use of zone prefixes and inter-zone routing of service prefixes. For example, the executive staff of a company can be assigned to a single zone to keep the dial strings simple, and DID can be implemented in the executive zone. Other zones on the network might use IVR due to the lack of video gateway services in every zone. By using the dial plans outlined in this document, you can keep the dial strings consistent across all zones.

You can use any of the following routing methods to route calls from the PSTN to H.323 endpoints and services in a multi-zone network:

- DID

  If you use DID to route calls from the PSTN to the H.323 endpoints and services, each E.164 address and service is a valid DID number associated with a PRI line attached to a Cisco IP/VC gateway. In order to use DID in a multi-zone network where zones may reside in different geographic regions, PSTN area codes and "data" boundaries require a video gateway in each area code.

- IVR

  IVR allows administrators to define the number structure of the dial plan. E.164 addresses can be four-digit extensions or 10-digit directory numbers. (The video terminal extension and the zone prefix combined should be 10 digits). IVR is the easiest method for routing incoming PSTN calls in a multi-zone network. When using IVR, calls terminate at the gateway, and then the caller enters the E.164 address or service prefix. If the call is in the local zone, only the E.164 address or service prefix is needed. If the call is going to another zone, the caller enters the zone prefix followed by the E.164 address. Services hosted on a remote MCU are dialed the same way (that is, zone prefix + service prefix + conference ID). IVR requires DTMF support from the dialing endpoint, but some older H.320 systems do not support DTMF.

- TCS4

  When using TCS4, the dial string from the H.320 endpoint contains the ISDN directory number for the gateway followed by a TCS4 delimiter and the E.164 address of the H.323 endpoint. If the incoming call is destined for an H.323 endpoint outside of the local zone, the zone prefix must be added to the dial string. Using TCS4 requires the dialing endpoint to support TCS4. Because TCS4 is not a commonly used routing method, Cisco recommends IVR instead.

- Default Extension

  A default extension is usually used in special cases such as call center applications or to route calls to a single H.323 endpoint.

Routing Inter-Zone Calls Using Hopoff Statements

You can add hopoff statements in the gatekeeper to route calls between zones without using a zone prefix. Hopoffs are used for routing gateway services because the service has no association with the zone where the gateway resides. To create a common dial plan, strategically deploy gateways in major sites and use hopoff statements in all smaller *stub* zones that do not contain a gateway. Then users, regardless of what zone they are in, can dial a common service prefix to access the outside world.

Use of the hopoff statement eliminates the need for users in a stub zone to dial the zone prefix of the zone that contains the gateway. Hopoffs override the gatekeeper parse order and direct calls with the defined service to a specific zone. Use the following command syntax to configure hopoff statements in the gatekeeper:

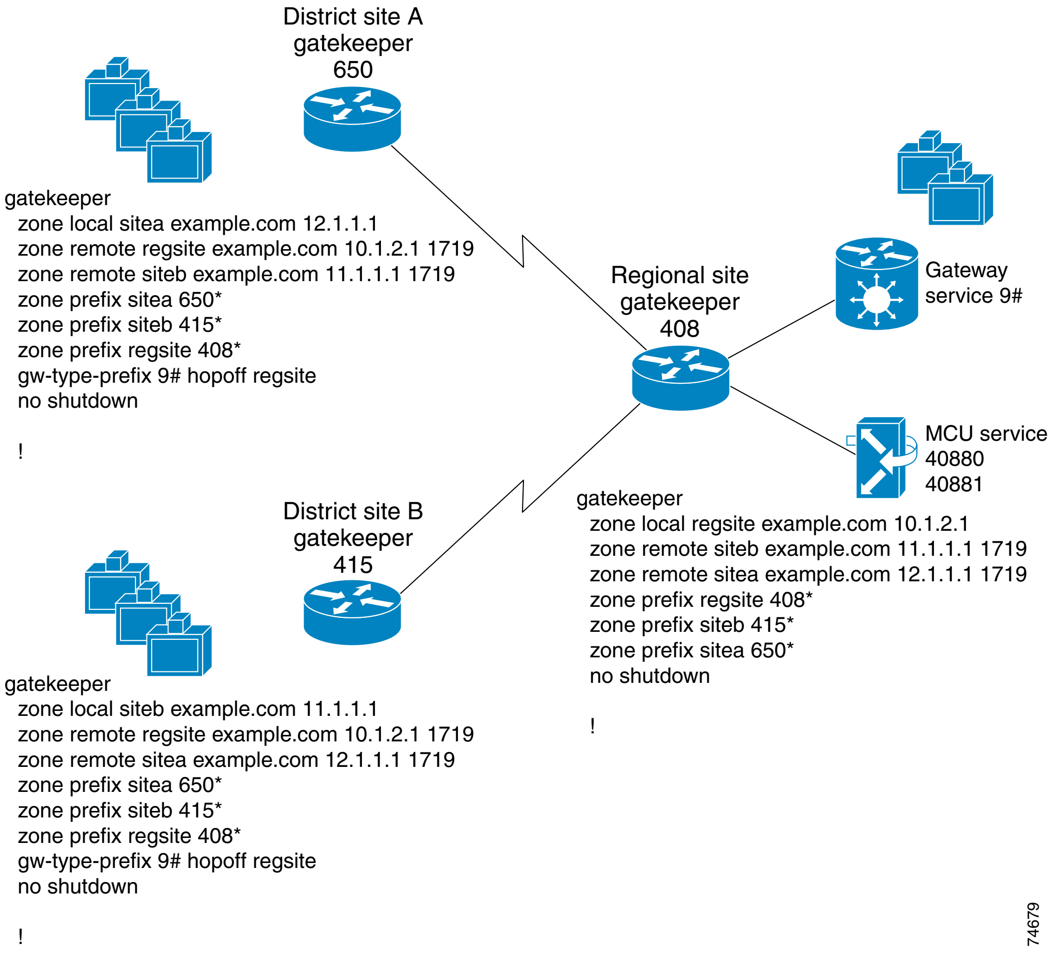
```
gw-type-prefix <prefix #> hopoff  <gatekeeper name>
```

MCUs do not require hopoff statements because the zone prefix is always embedded in the service prefix.

Note When creating multiple zones on a single router and registering MCUs or gateways in any of the zones, enter a hopoff command for each service prefix. Routing of service prefixes between local zones also requires a hopoff.

In Figure 7-2, District Site A and District Site B have hopoffs configured to forward all gateway calls (service prefix 9#) to the regional site. These hopoff statements forward calls matching 9#* to the regional site.

Figure 7-2 Inter-Zone Routing with Hopoff Statements Configured



```
gatekeeper
  zone local sitea example.com 12.1.1.1
  zone remote regsite example.com 10.1.2.1 1719
  zone remote siteb example.com 11.1.1.1 1719
  zone prefix sitea 650*
  zone prefix siteb 415*
  zone prefix regsite 408*
  gw-type-prefix 9# hopoff regsite
  no shutdown

!
```

```
gatekeeper
  zone local regsite example.com 10.1.2.1
  zone remote siteb example.com 11.1.1.1 1719
  zone remote sitea example.com 12.1.1.1 1719
  zone prefix regsite 408*
  zone prefix siteb 415*
  zone prefix sitea 650*
  no shutdown

!
```

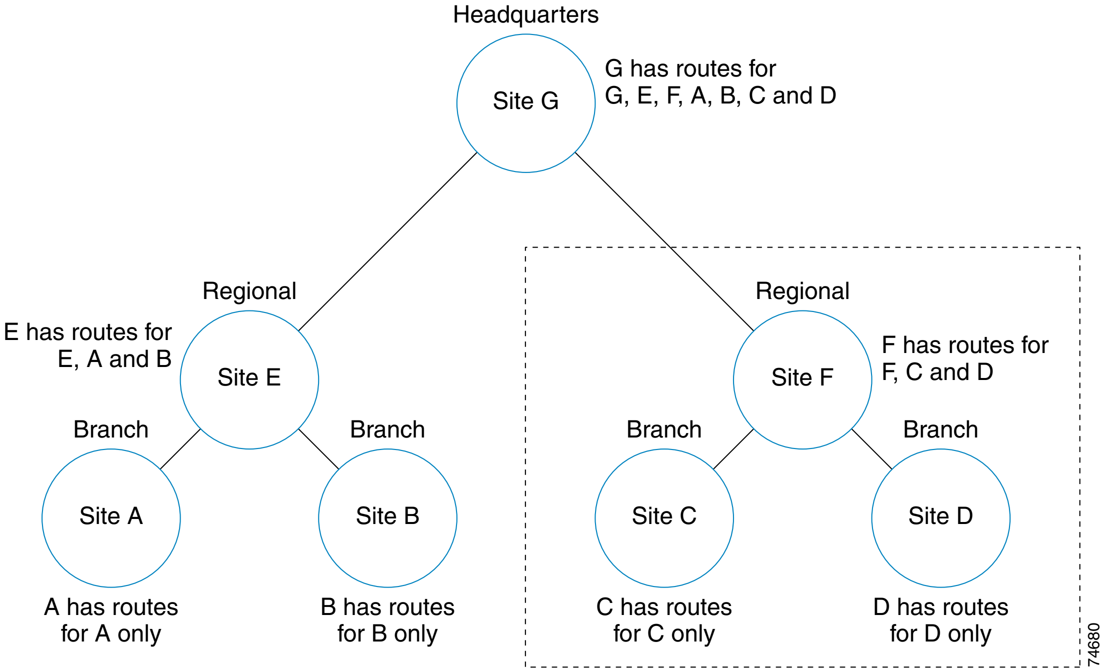
```
gatekeeper
  zone local siteb example.com 11.1.1.1
  zone remote regsite example.com 10.1.2.1 1719
  zone remote sitea example.com 12.1.1.1 1719
  zone prefix sitea 650*
  zone prefix siteb 415*
  zone prefix regsite 408*
  gw-type-prefix 9# hopoff regsite
  no shutdown

!
```

Routing Inter-Zone Calls Using a Directory Gatekeeper

Currently, there is no gatekeeper protocol that allows gatekeepers to update each other with routing information. This limitation implies a full-mesh topology, where every gatekeeper must be statically configured to know about every other gatekeeper to which it is going to route calls. In effect, all gatekeepers must be known to each other. This poses scalability problems when a new zone or service is added because the administrator must add an entry in every gatekeeper for the new zone or service.

By using a directory gatekeeper and Location Request (LRQ) forwarding, a hierarchical gatekeeper design can limit the administrative overhead in a large multi-zone network. LRQ forwarding allows an administrator to create a directory gatekeeper that maintains all zone prefixes for the network or subset of the network. Call admission control is addressed by local gatekeepers, and the directory gatekeeper plays no role in it. In Figure 7-3, sites A, B, C, and D are configured to forward all LRQs that cannot be resolved locally to directory gatekeeper sites (E and F).

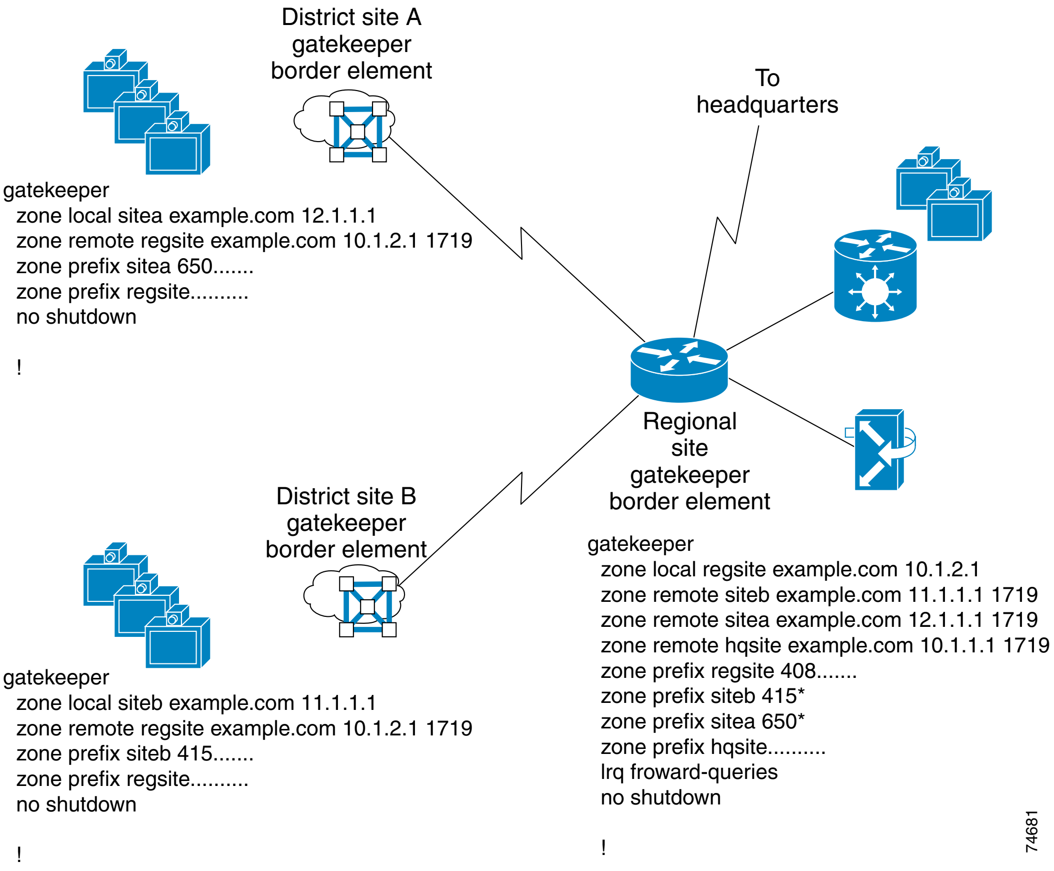
Figure 7-3 Inter-Zone Routing with Directory Gatekeepers



In Figure 7-4 there are three zones, two district zones and a regional zone that has a connection back to headquarters. Each district zone contains information about its local zone only. The command line **zone prefix regsite ....** routes any call placed with 10 digits, but not matched in the local zone, to the regional site.

The regional site contains the routing information for its own zone as well as the two district zones associated with it. Zone prefix and hopoff statements are added to the regional site as the zones are added to the network. There is also a **zone prefix hqsite ....** entry in the regional gatekeeper that forwards any 10-digit call with no match to the headquarters gatekeeper. If LRQs are going to be forwarded past the directory gatekeeper, an **lrq forward-queries** entry must be added to the gatekeeper, otherwise LRQs will not be forwarded past the directory gatekeeper. (LRQ forwarding has a maximum limit of seven hops.) This model can be expanded in a large network to make an H.323 network more manageable.

Figure 7-4 Directory Gatekeeper Example



```
gatekeeper
 zone local sitea example.com 12.1.1.1
 zone remote regsite example.com 10.1.2.1 1719
 zone prefix sitea 650.......
 zone prefix regsite..........
 no shutdown

 !
```

```
gatekeeper
 zone local siteb example.com 11.1.1.1
 zone remote regsite example.com 10.1.2.1 1719
 zone prefix siteb 415.......
 zone prefix regsite..........
 no shutdown

 !
```

```
gatekeeper
 zone local regsite example.com 10.1.2.1
 zone remote siteb example.com 11.1.1.1 1719
 zone remote sitea example.com 12.1.1.1 1719
 zone remote hqsite example.com 10.1.1.1 1719
 zone prefix regsite 408.......
 zone prefix siteb 415*
 zone prefix sitea 650*
 zone prefix hqsite..........
 lrq froward-queries
 no shutdown

 !
```

Note When configuring the directory gatekeeper, do not use the wildcard (*) as the directory gatekeeper zone prefix, otherwise calls will not be routed properly. For example, the command **zone prefix regsite *** will route all calls, even local ones, to the directory gatekeeper.

In Figure 7-4 the directory gatekeeper entry is **zone prefix regsite ....**, which allows any 10-digit dial string that is not matched locally to be forwarded to the directory gatekeeper. If there is a need for users to dial 11- or 12-digit dial strings, you can enter multiple zone prefix entries for the directory gatekeeper. Deployments that support international locations are more likely to require multiple zone prefix entries for the directory gatekeeper.

If a root zone contains a video gateway, and multiple directory gatekeeper zone prefixes are configured, you might have to add a hopoff to the configuration. If any of the directory gatekeeper zone prefix lengths match the dial string minus the service prefix, the call is forwarded to the directory gatekeeper. For example, if a local gateway service prefix is 9#, PSTN calls will be either nine digits (local calls) or 12 digits (long distance) including the service prefix.

When the gatekeeper starts to parse the dial string, it strips the service prefix and starts looking for a match. In the preceding example, local PSTN calls are parsed on seven digits and long distance PSTN calls are parsed on 11 digits. If the gatekeeper configuration contains a directory gatekeeper entry with seven dots or 11 dots, a hopoff is needed. The same rule applies to MCUs, but in most cases MCU calls are parsed on five digits or less, while most directory gatekeeper zone prefix entries are matched on 10 digits or more.

Example 7-1 illustrates the configuration of a root zone containing multiple directory gatekeeper zone prefix entries and a hopoff for 9#. The reason for the hopoff is to eliminate long distance calls (which are parsed on 11 digits) from matching the DGK zone prefix entry with 11 dots. Figure 7-5 and Figure 7-6 illustrate the parse order for Admission Requests (ARQs) and via-zone processing in the Cisco gatekeeper.

Example 7-1 Configuration of Root Zone with Multiple Director Gatekeepers

```
gatekeeper
 zone local HKG cisco.com 10.1.3.1
 zone remote APAC_DGK cisco.com 10.1.2.1
 zone prefix HKG 852.......
 zone prefix APAC_DGK ..........
 zone prefix APAC_DGK ...........    (This entry matches long distance PSTN calls to a gateway)
 zone prefix APAC_DGK ............
 gw-type-prefix 9#* hopoff HKG    (This entry keeps all 9# dial strings in the HGK zone)
 no use-proxy HKG default inbound-to terminal
 no use-proxy HKG default outbound-from terminal
 bandwidth remote 1000
 no shutdown
```

Note The functionality illustrated in Figure 7-5 was first introduced in Cisco IOS Release IOS 12.3(15)T and continues to be enhanced in subsequent releases such as Cisco IOS Release 12.4(15)T and later.

Figure 7-5 Gatekeeper Address Resolution for ARQ

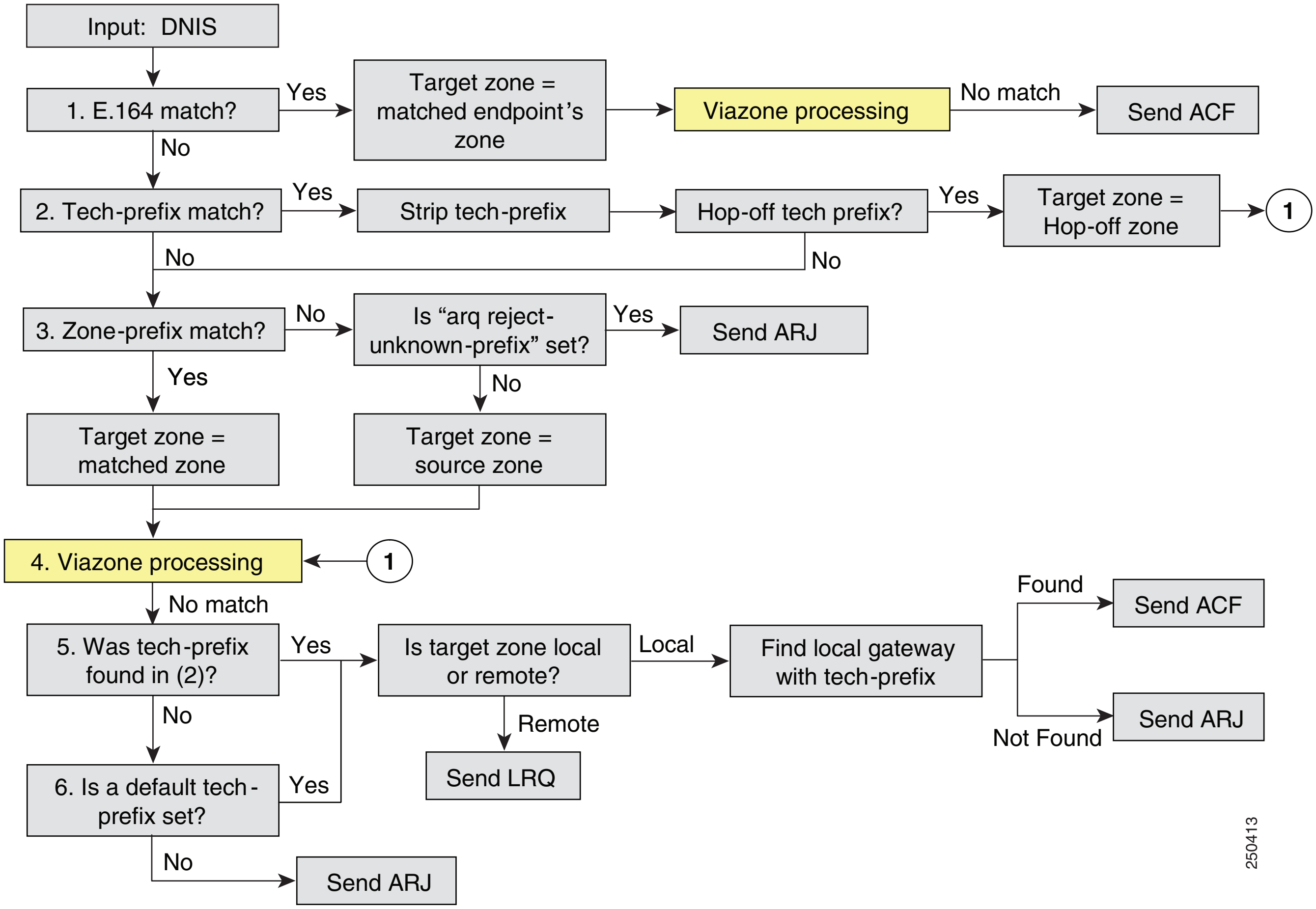
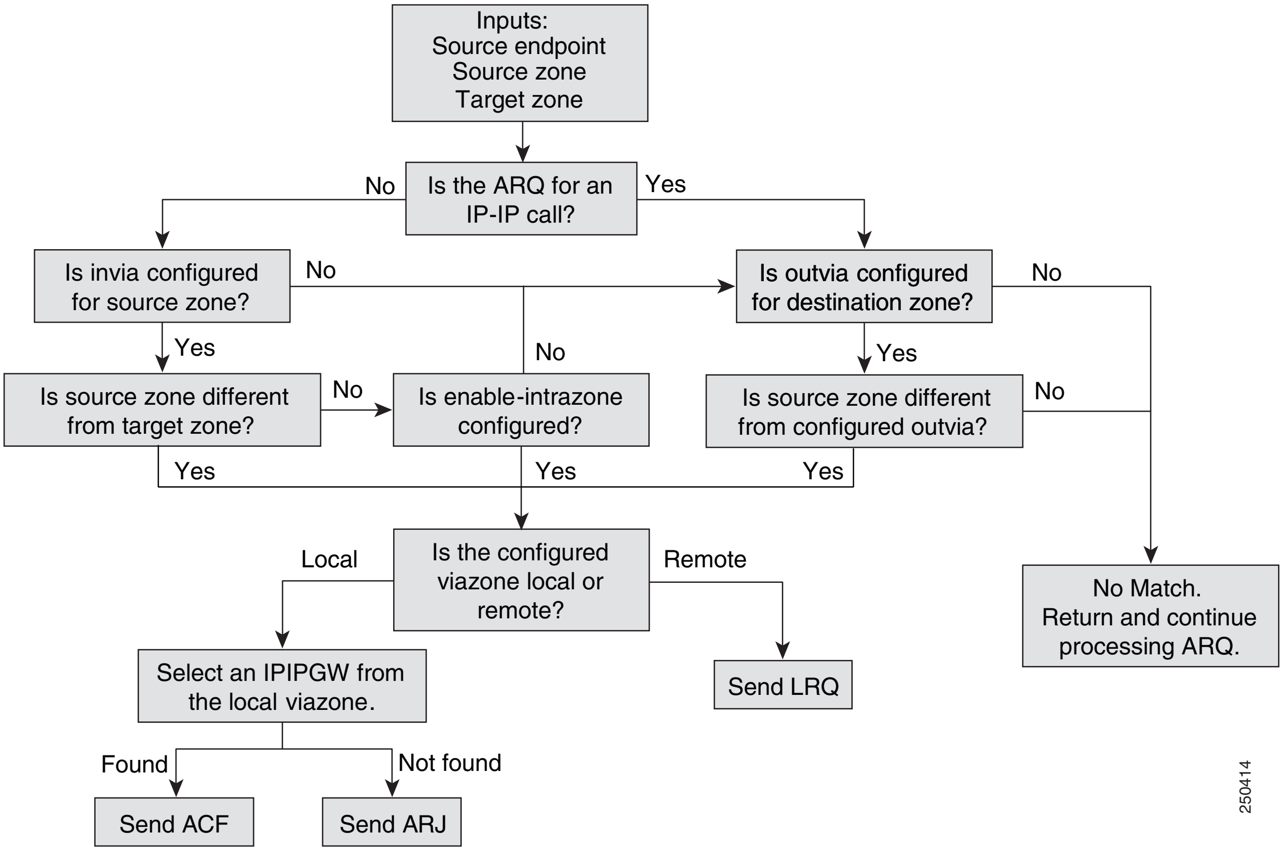

Input: DNIS
1. E.164 match?
Yes
Target zone = matched endpoint's zone
Viazone processing
No match
Send ACF
No
2. Tech-prefix match?
Yes
Strip tech-prefix
Hop-off tech prefix?
Yes
Target zone = Hop-off zone
1
No
No
3. Zone-prefix match?
No
Is "arq reject-unknown-prefix" set?
Yes
Send ARJ
Yes
No
Target zone = matched zone
Target zone = source zone
4. Viazone processing
1
No match
5. Was tech-prefix found in (2)?
Yes
Is target zone local or remote?
Local
Find local gateway with tech-prefix
Found
Send ACF
Not Found
Send ARJ
Remote
Send LRQ
No
6. Is a default tech-prefix set?
Yes
No
Send ARJ
250413

Figure 7-6 Via-Zone Processing


Inputs:
Source endpoint
Source zone
Target zone
Is the ARQ for an IP-IP call?
No
Yes
Is invia configured for source zone?
No
Is outvia configured for destination zone?
No
Yes
No
Yes
Is source zone different from target zone?
No
Is enable-intrazone configured?
Is source zone different from configured outvia?
No
Yes
Yes
Yes
Is the configured viazone local or remote?
Local
Remote
No Match. Return and continue processing ARQ.
Select an IPIPGW from the local viazone.
Send LRQ
Found
Not found
Send ACF
Send ARJ
250414

Routing Using the Border Element

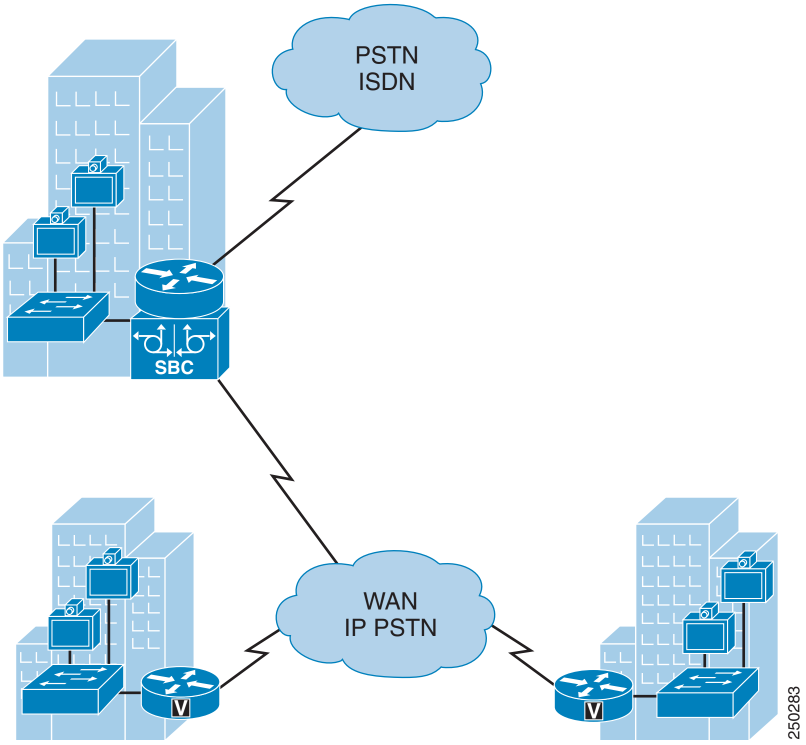
Border elements can be deployed in a single gatekeeper zone or in multiple zones, as described in the following sections.

Calls with a Single Zone

Gatekeeper deployments of one local zone can have the border element inserted for all calls or for calls to and from remote zones. Although inserting border elements for all calls can have advantages, most scenarios that need border element are for isolating calls to and from remote zones.

Figure 7-7 has a regional location with the gatekeeper and the border element. The local sites register the video endpoints to the gatekeeper and use gatekeeper bandwidth for call admission control. The border element is inserted into the calls for calls exiting the region to the IP PSTN, represented in the gatekeeper configuration as the remote zone. (Example 7-2 lists the configuration details.)

Figure 7-7 Single Zone for Calls



Example 7-2 Configuration of a Single Zone

```
dial-peer voice 1 voip
  destination-pattern 91……….
  session target ras
  incoming called-number 91……….
  codec transparent

gatekeeper
  zone local regsite example.com 10.1.1.1
  zone local vzone example.com enable-intrazone
  zone remote pstnzone example.com 15.1.1.1 1719 invia vzone outvia vzone
  no zone subnet vzone default enable
  zone subnet vzone 10.1.1.1/32 enable
  zone prefix pstnzone 91..........
  no use-proxy regsite default inbound-to terminal
  no use-proxy regsite default outbound-from terminal
  no use-proxy vzone default inbound-to terminal
  no use-proxy vzone default outbound-from terminal
  no shutdown
```

Note Deployments may use an external NAT and Firewall device to ensure security and IP address hiding. The border element can provide IP address hiding because it can have an outside address. However, firewalls need to be aware of the protocol (H.323v4/5) to provide greater security by dynamically opening holes for the calls.

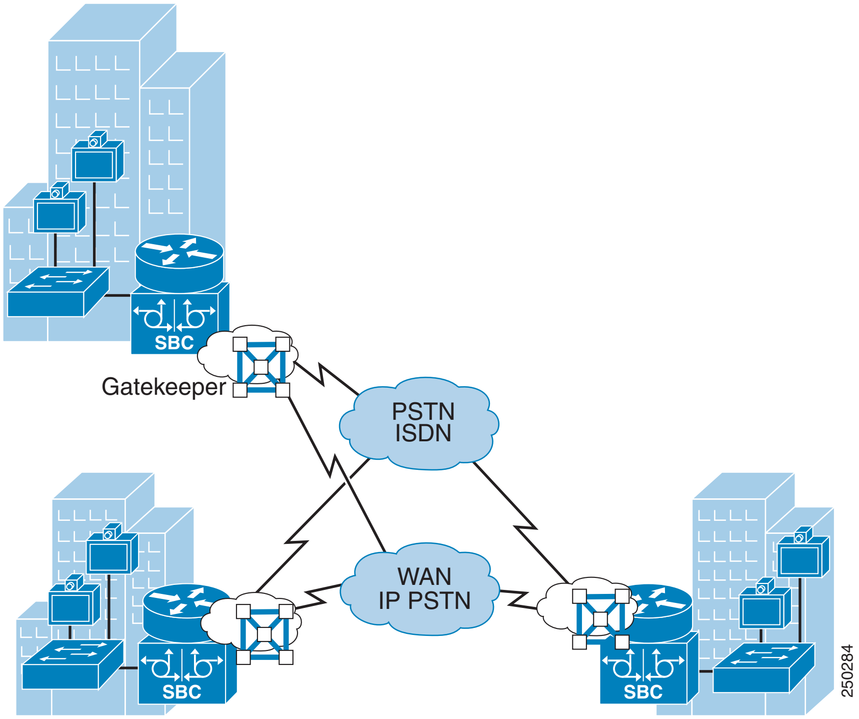
Calls with Multiple Zones

Deployments with multiple locations are more scalable with a gatekeeper and border element per location. With a border element per location, RSVP configurations are more manageable, and this also helps keep the dial plan distributed. Each gatekeeper resolves calls within its location or sends calls to other locations. The border element can be inserted into calls originated by or sent from that gatekeeper to other locations, or for calls coming into that location. For larger deployments with multiple gatekeepers, a directory gatekeeper can keep the dial plan simple.

The border element must be inserted when calls traverse the management domains or when connecting to the IP PSTN from an inside network. There would then be a single trusted device for incoming and outgoing calls through an application-aware firewall.

As illustrated in Figure 7-8, multiple locations can each have a border element in multiple respective zones on the gatekeeper. The respective border elements are chosen for calls to and from the locations. Media bypass can be used for intra-site calls because every call will involve the border element. (Example 7-3 lists the configuration details.)

Figure 7-8 Multiple Zones for Calls



Example 7-3 Configuration of Multiple Zones

```
interface FastEthernet0/0
  ip address 10.1.1.1 255.255.255.0
  h323-gateway voip interface
  h323-gateway voip id vzone ipaddr 10.1.1.1 1719
  h323-gateway voip h323-id regsite-cube
  h323-gateway voip tech-prefix 1#
  h323-gateway voip bind srcaddr 10.1.1.1
```

```
dial-peer voice 1000 voip
  destination-pattern .T
  session target ras
  incoming called-number .T
  codec transparent

gatekeeper
  zone local regsite example.com 10.1.1.1
  zone local vzone example.com enable-intrazone
  zone remote pstnzone example.com 14.1.1.1 1719 invia vzone outvia vzone
  zone remote vsitea example.com 11.1.1.1 1719 invia vzone outvia vzone
  zone remote vsiteb example.com 12.1.1.1 1719 invia vzone outvia vzone
  no zone subnet regsite default enable
  zone subnet regsite 10.1.1.10/32 enable
  zone subnet regsite 10.1.1.11/32 enable
  no zone subnet vzone default enable
  zone subnet vzone 10.1.1.1/32 enable
  zone prefix vsitea 415.......
  zone prefix vsiteb 650.......
  gw-type-prefix 1#* default-technology
  no use-proxy regsite default inbound-to terminal
  no use-proxy regsite default outbound-from terminal
  no use-proxy vzone default inbound-to terminal
  no use-proxy vzone default outbound-from terminal
  no shutdown
```

Note This approach can be used to migrate existing Cisco Multimedia Conference Manager (MCM) Proxy users to the present Cisco Unified Border Element. In addition, this approach allows you to configure call admission control as needed.

Routing Calls to Other Networks

Calls to other unified communications network can be done using their trunk interfaces. The gatekeeper plays a key role to route calls to and from such networks. Unified communications networks can be peers or remote gatekeepers to this video network, or they can register as gateway devices to the gatekeeper. Call admission control by the gatekeeper is done using bandwidth configurations. The dial plan and call routing on the gatekeeper can be done with the zone prefix commands.

Some unified communications systems need both the signaling and media for the calls to be from a single device, or they need fixed devices to provide the call traffic. In such cases the gatekeeper can be used as the signaling entity, and the border element can do additional call signaling and media. Appropriate devices must be chosen to support the scale according to device performance and capabilities.

Routing Calls Through Untrusted Networks

When enterprises need to communicate across untrusted networks, they use firewalls as security devices to prevent unauthorized access to internal networks. Firewalls can do protocol inspection of VoIP protocols to provide a similar level of security for calls. Firewalls can inspect VoIP protocols, H.323, or SIP signaling messages and allow the appropriate traffic for the media to go between the two entities. The border element adds value here by being the one trusted device that the firewall can inspect for call signaling traffic to and from the internal network. This not only reduces the firewall configuration complexity but also allows the border element to do the topology hiding for the internal network.

Calls then can be routed through external public gatekeepers or directly to other enterprises based on the dial plan configuration. External untrusted endpoints must register with an enterprise gatekeeper so that the registration and dial plan can be managed by the enterprise. Gatekeepers then can restrict unauthorized registrations; the firewall can provide protocol inspection and security against unauthorized calls; and the border element can provide number manipulation, load balancing, call admission control, and RSVP if configured.

CHAPTER 8

Cisco Video Infrastructure Components

Last revised on: October 30, 2009

This chapter describes the Cisco Unified Videoconferencing infrastructure components and related network design considerations. This chapter focuses on traditional (room-based) IP videoconferencing. Therefore, it is specifically intended for those environments that have yet to migrate to an end-to-end Cisco Unified Communications Manager Video Telephony model or Cisco Unified MeetingPlace Video Integration.

For information on integrating the Cisco Unified Videoconferencing infrastructure components with Cisco Unified Communications Manager, refer to the *IP Video Telephony* section in the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Note MCU and gateway features may change with new software releases, and those changes might not be represented in this document. Refer to the latest product documentation and release notes for details.

What's New in This Chapter

Table 8-1 lists the topics that are new in this chapter or that have changed significantly from previous releases of this document.

Table 8-1 New or Changed Information Since the Previous Release of This Document

New or Revised Topic	Described in:
Collaboration	Collaboration for Desktop Sharing, page 8-21
Connecting to external networks	Connecting the Video Network with External Untrusted Networks, page 8-27
Integration with other collaboration applications	Integration and Interoperability, page 8-25
Multipoint conferencing	Multipoint Conferencing, page 8-2
Recording conferences	Conference Recording, page 8-24
Scheduling video conferences with Cisco WebEx	Video Scheduling with Cisco WebEx, page 8-8
Streaming conference video	Conference Streaming, page 8-23

Design and Deployment Overview

Video infrastructure design is a very important element in an H.323, Session Initiation Protocol (SIP), or Skinny Client Control Protocol (SCCP) videoconferencing network. In the H.320 circuit-switched network, Multipoint Control Units (MCUs) and H.320 endpoints are connected directly to the PSTN network. In the past, an MCU could have multiple PRI connections into the switched network. The switched network supplied a dedicated transport with guaranteed bandwidth and predictable delay. Now that video is being moved onto IP networks that share bandwidth with data, placement of video infrastructure components becomes very important. Installing a central MCU and/or gateway in an IP environment does not always work. Bandwidth in an IP network is not dedicated to each video device on the network, therefore it is important to design the network accordingly.

The design of systems with Cisco Unified Videoconferencing components is best illustrated by reviewing design scenarios step-by-step. This method provides an understanding of how best to deploy Cisco Unified Videoconferencing for your particular environment. The scenarios in this chapter are designed to simplify the end-user experience and conserve bandwidth. Therefore, no matter what stage your videoconferencing network is in today, you can benefit from deploying one or more of the videoconferencing scenarios described in this chapter.

Multipoint Conferencing

Whenever three or more parties join a videoconference, a Multipoint Control Unit (MCU) is required. The MCU can mix all video streams together and transmit a composite image of all participants back to the originating sources. This composite view (called *continuous presence*) is necessary for all participants to see each other simultaneously. The continuous presence view can display from 2 to 32 windows (participants) in a variety of different layouts. Each layout offers the ability to make one of the windows voice-activated, which is useful if there are more participants in the conference than there are windows to display them all in the composite view. More than 32 participants can be in a single videoconference, and the MCU will allow the last active sites to be displayed. Moreover, as participants join or leave the conference, the Cisco Unified Videoconferencing MCU can automatically adjust the continuous presence view by dynamically changing the layout.

Continuous presence can have multiple participants using different video and audio codecs. In this situation, the MCU must reconstitute all the video streams into a single image while at the same time preserving the video and audio codec of each participant. In addition, the MCU can save network bandwidth by avoiding the need for the sources to transmit large video streams to all sources. Although MCUs can be implemented as either hardware or software, hardware-based MCUs guarantee superior video quality by performing advanced transcoding, transrating, and composition features.

The Cisco Unified Videoconferencing solution consists of the following main components:

- Cisco IOS Gatekeeper

  This component provides address resolution and H.323 audio and video setup and tear down.

- Cisco Unified Videoconferencing MCU

  This component performs call signaling and multipoint processing of all audio. In addition, the MCU hosts the web interface and controls one or more Enhanced Media Processors (EMPs).

- EMP

  The EMP is a dedicated multipoint processor for video. The EMP cannot function without a controlling MCU.

- H.320 Gateway

  This component bridges ISDN video endpoints into the IP H.323 infrastructure. There are several ways to use the ISDN gateways in a video solution.

- Desktop Streaming Server

  The Desktop streaming server streams the conference to user desktop PCs or laptops. The Desktop server provides the collaboration features for video conferences by allowing users to share desktop or laptop screens and by providing enhanced roster and conference moderation functionality. It also enables conferences to be streamed using Real Time Streaming Protocol (RTSP), to be viewed through a media player or browser plug-in as a webcast.

- Desktop Recording Server

  The Recording server can provide recording of the videoconference and desktop sharing. It can also be a repository for the recorded meetings to be viewed later.

- Management and Integration Server

  The Management server can be the central entity for configuration and management of the various video devices and the videoconferencing elements. Thus, it can be used to configure, manage, and monitor devices such as MCUs and gatekeepers, among other. The server can also perform conference scheduling, be a gatekeeper, and provide cascading logic for video conferences between different locations to optimize WAN utilization if possible.

  The Integration servers allow the enterprise calendaring systems to assign and reserve resources such as MCU ports for the conference, thus ensuring that sufficient resources are available during the conference.

The recommended multipoint hardware platforms are the Cisco Unified Videoconferencing 5000 Series, 3545 and 3515 MCUs. With the Cisco Unified Videoconferencing 3545 and 3515 MCUs, video processing is done by the EMP module, and the MCU performs audio processing and conference control.

The Cisco Unified Videoconferencing 3515 MCU is designed to be a self-contained unit for smaller deployments. Therefore, it has an integrated EMP along with an integrated MCU inside a fixed chassis that cannot be upgraded in the field. By contrast, the Cisco Unified Videoconferencing 5200 or 3545 System is modular and is designed for maximum deployment flexibility. Therefore, MCUs, EMPs, or H.320 gateway modules can be inserted into the Cisco Unified Videoconferencing 5200 or 3545 chassis in order to process audio and video, and this system can be upgraded in the field.

The Cisco Unified Videoconferencing 5000 Series and 3500 Series MCUs support H.323 (standard definition and high definition), SCCP, and SIP endpoint interoperability. The encoder-per-port hardware architecture allows any supported connection speed from 64 kbps up to 2 Mbps, with any supported codec and at any standard-definition resolution. Supported protocols include H.261, H.263, and H.264 video codecs, as well as G.711, G.722, G.722.1, G.723.1, G.728, and G.729A audio codecs.

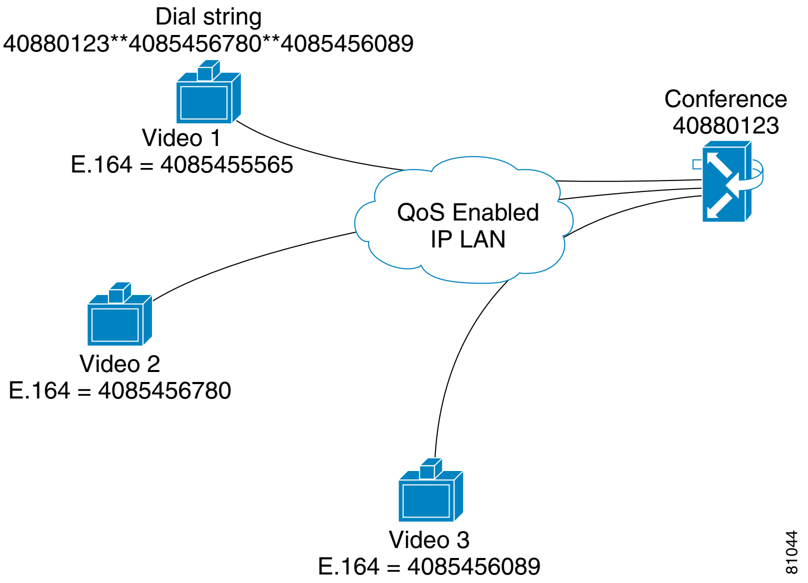
High-definition video calls provide better video to users. Using traditional or desktop-based high-definition endpoints can enhance the conference experience. High-definition video also requires that the network must be able to process the greater volume of traffic for the calls and must provide Quality of Service (QoS), traffic classification, and call admission control where network resources are not sufficient. For additional information on multipoint conferencing, continuous presence, and voice activation, refer to the *IP Video Telephony* section in the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Initiating a Call

To initiate a multipoint call using a Cisco Unified Videoconferencing MCU, the endpoint dials the appropriate service prefix followed by a conference ID. (The conference ID can be up to 9 digits long.) If a service on an MCU is 40880 for a 384-kbps call, the user might dial 4088011223, the call would be routed to the MCU using the 40880 service prefix, and the MCU would initiate an ad hoc conference with an ID of 4088011223. Users can also initiate a call and invite the other participants by dialing the conference ID, the invite string **, and the E.164 address of the another participant. Figure 8-1 shows the dial sequence of an MCU call with Video 1 initiating an MCU call to 40880123 and inviting Video 2 and Video 3.

Figure 8-1 Example Dial Sequence for Initiating a Call (MCU Invite Call)

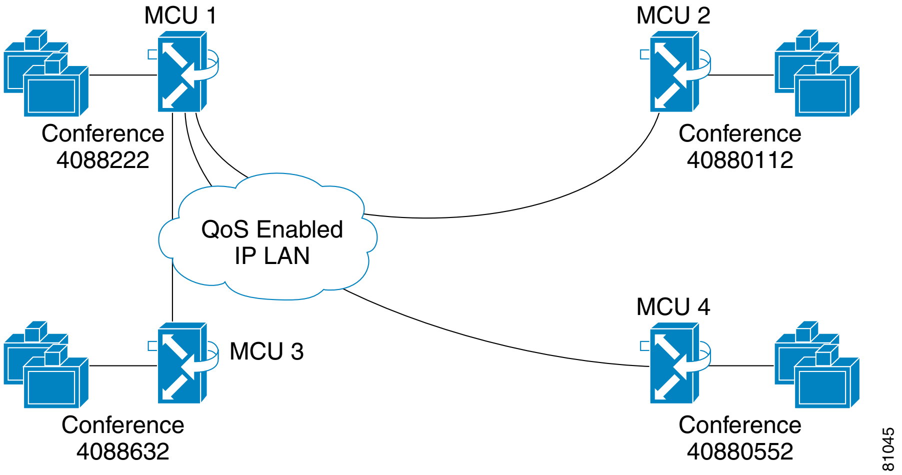


In addition, it is possible for an H.323 endpoint to dial the MCU by IP address. The MCU will answer the call with an audio IVR and an on-screen menu.

Cascading MCUs

Cascading MCUs allows larger conferences to be created by combining resources from multiple MCU blades or units. Cascading is also used in distributed MCU environments to save bandwidth on low-speed WAN links (see Distributed MCUs, page 8-5). Cisco Unified Videoconferencing MCUs support cascading. An administrator can cascade MCUs by inviting conference calls on different MCUs to join in a single combined call. In Figure 8-2, a conference was started on each of the four MCUs. To cascade the MCUs in this example, an administrator accessed the web interface on MCU 1 and invited conferences 4088112 on MCU 2, 4088632 on MCU 3, and 4088552 on MCU 4.

Figure 8-2 Cascaded MCU Conference



Note There is no physical connection made between the cascaded devices. Cascading occurs over the LAN or WAN, allowing MCUs to be distributed across a network. An MCU can invite H.323 endpoints, H.320 endpoints through a gateway, or other MCUs through the web interface on the MCU.

Distributed MCUs

With the ability to cascade multipoint conferences, administrators can build an H.323 video network with distributed MCU services. A distributed MCU architecture saves WAN bandwidth when a conference includes multiple participants on two or more campuses connected by a WAN. By distributing MCUs across the network, it is possible to have multipoint conferences across WAN links without limiting the number of users at remote sites.

Centrally located MCU services require all conference participants to place a call across the WAN to the MCU, while distributed MCUs allow users to call to their local MCU and join the other MCUs into a cascaded conference across the WAN. Figure 8-3 illustrates centralized MCU services, with three users from Campus B joining a conference hosted at Campus A. In this model, all three calls must traverse the WAN link.

Figure 8-3 Centralized MCU Services, with All Calls Traversing the WAN

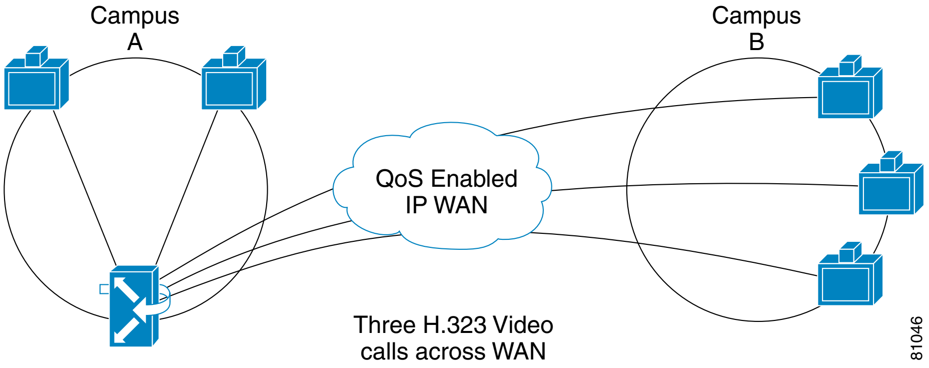
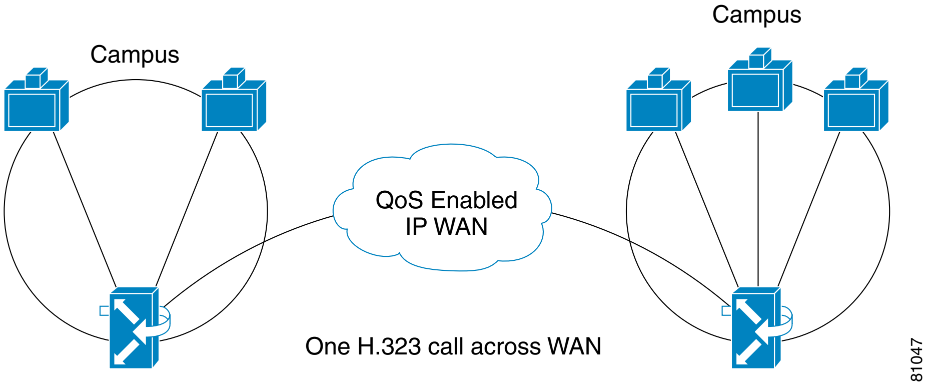


Figure 8-4 illustrates a distributed MCU model with two video terminals at Campus A calling into a local MCU, and three video terminals at Campus B calling into a local MCU. In this model, there is just a single call cascading the two conferences across the WAN. In most distributed networks, a Cisco Unified Videoconferencing 5000 or 3545 System will be located at the headquarters site and a Cisco Unified Videoconferencing 3515 MCU will be located at each large remote site.

Figure 8-4 Distributed MCU Services, with a Single Call Traversing the WAN

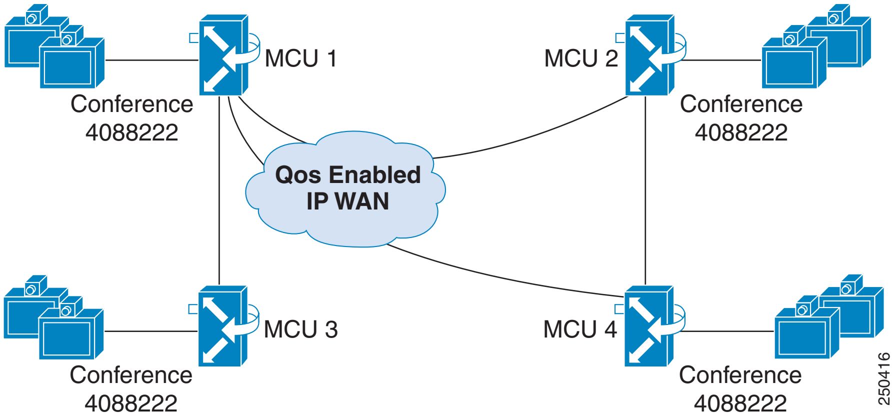


Dynamic Cascading of MCUs

Dynamic cascading improves upon normal cascading by allowing conferences to be distributed automatically to local MCUs, with no administrator or participant intervention. By using single number reach, dynamic cascading hides the complexity from participants and automatically combines distributed MCUs from multiple locations. Dynamic cascading works through automatic scheduling and intelligent assignment of distributed MCU resources. Participants join a single combined call that reduces the network bandwidth consumption to only a single video call across the WAN.

As the conference grows in size, dynamic cascading will intelligently expand the conference as long as sufficient network resources are available. The Cisco Unified Videoconferencing 5000 Series or 3500 Series MCUs support dynamic cascading. In Figure 8-5, a conference was automatically started on each of the four MCUs. The same number (4088222) was dialed by all eight participants, but dynamic cascading automatically assigned the local terminals to the respective local MCUs. The option is available for the MCUs to dial out directly to the terminals, thereby simplifying the user experience.

Figure 8-5 Dynamically Cascaded MCU Conference



Note Dynamic cascading requires a Cisco IOS Gatekeeper and a videoconferencing scheduler using either Cisco Unified MeetingPlace Video Integration or Cisco Unified Videoconferencing Manager.

Note When deployments have a combination of Cisco Unified Videoconferencing 5000 Series or 3500 Series MCUs, cascading between the two series types of MCUs must be done manually per conference. However, dynamic cascading is supported between MCUs of the same series type.

Scheduled Conferencing

Scheduled videoconferencing provides the ability to reserve resources in advance, as well as to simplify the videoconferencing experience for end users. Administrators can grant users the ability to schedule conferences through email integration or a web interface. The scheduling interface hides all videoconferencing complexity from the end user.

There are three applications used to provide scheduling videoconferencing in a Cisco Unified Videoconferencing environment:

- Cisco Unified MeetingPlace with Video Integration
- Cisco Unified Videoconferencing Manager
- Cisco WebEx with Cisco Unified Videoconferencing integration

Video Scheduling with Cisco Unified MeetingPlace

Cisco Unified MeetingPlace is the recommended method for scheduling video conferences. The participants schedule video resources through email (Microsoft Outlook or Lotus Notes) or a simple-to-use web interface. Web conferencing can easily be added to provide integrated voice, video, and web conferencing solutions that fit any enterprise.

For additional information on Cisco Unified MeetingPlace Video Integration, including detailed video call flows, refer to the *Cisco Unified MeetingPlace* section of the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Video Scheduling with Cisco WebEx

Enterprises that use Cisco WebEx for collaboration can use the WebEx scheduling features for their conferences. WebEx scheduling is done through the cloud services provided by WebEx. WebEx scheduling provides integration for various email systems through plug-ins. The plug-ins enable users to schedule conferences using the enterprise calendaring systems and to choose users and resources that then get updated in the WebEx cloud through the users' WebEx accounts.

WebEx integrations with the MCU and with Cisco Unified Videoconferencing Desktop allows the enterprise to provide the needed QoS for the scheduled calls and for integration with user endpoints. For additional details on the integration, see the section on Integration and Interoperability, page 8-25.

Video Scheduling with Cisco Unified Videoconferencing Manager

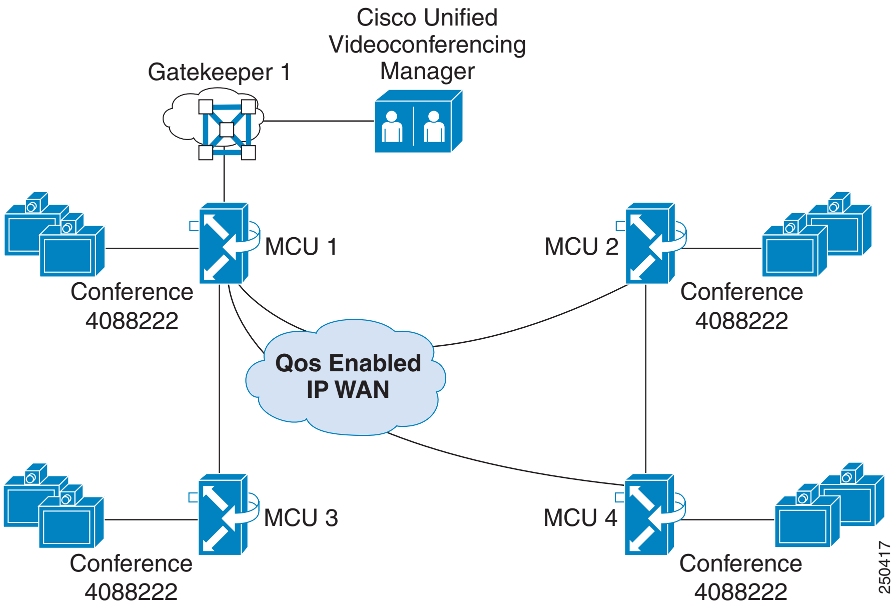
For environments without Cisco Unified MeetingPlace, scheduling is accomplished with Cisco Unified Videoconferencing Manager, which is a standalone videoconferencing option only. It enables users to schedule voice and video conferences. Cisco Unified Videoconferencing Manager can be used with the Outlook integration option or with web-based scheduling. Moreover, synchronization with LDAP can minimize user setup and maintenance as well as providing authentication. In addition, LDAP can be used to integrate video terminals into the LDAP directory and associate Class of Service policies to users.

Class of Service (CoS) is accomplished by restricting users to specific meeting types. For example, administrators can grant users the ability to schedule meetings with only standard-rate video service (384 kbps video streams). Likewise, administrators can grant another set of users the ability to schedule meetings with high-definition video service (up to 2 Mbps).

Note Cisco Unified Videoconferencing Manager can provide call detail records (CDRs) for all videoconferencing calls.

Figure 8-6 illustrates Cisco Unified Videoconferencing Manager together with a Cisco IOS Gatekeeper. Using Microsoft Outlook, a participant in this example has scheduled the video terminals into a conference. When the meeting starts, Cisco Unified Videoconferencing Manager automatically out-dials all terminals into the call. All of this occurs without any user intervention aside from using Outlook to schedule the video conference. Optionally, Cisco Unified Videoconferencing Manager can also dynamically cascade all terminals. Thus, bandwidth consumption on the network is minimized to a single call.

Figure 8-6 Cisco Unified Videoconferencing Manager



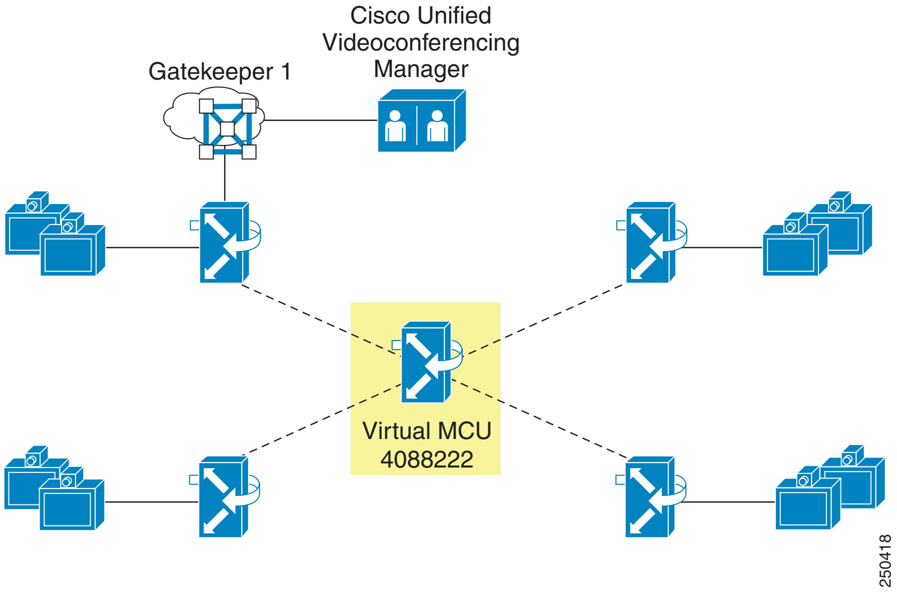
Note Video media streams do not flow through Cisco Unified Videoconferencing Manager. Video media terminates on the MCUs.

Virtual MCU

Cisco Unified Videoconferencing Manager has the capability to allow all MCU resources to appear as a single MCU. This virtual MCU concept provides a single conference ID across multiple MCUs. Therefore, this reduces the complexity of scheduling because participants can dial a single number and automatically be associated to the closest MCU. Likewise, the system can dial out to the participants. The virtual MCU provides the logic for dynamic cascading.

In Figure 8-7, the same conference call illustrated in Figure 8-6 is now depicted as a purely logical entity. The four distributed MCUs become a single virtual MCU when using Cisco Unified Videoconferencing Manager for scheduling the video conference. Participants can dial the number 4088222 located on the meeting invitation, and Cisco Unified Videoconferencing Manager will assign video terminals to local MCU resources. The virtual MCU has intelligent topology awareness of local resources, so it can avoid unnecessary bandwidth congestion over the WAN.

Figure 8-7 Virtual MCU Created by Cisco Unified Videoconferencing Manager



Note To enable dynamic cascading, Video terminals and MCUs must be associated to a location in the Resource Manager component of Cisco Unified Videoconferencing Manager.

Cisco Unified Videoconferencing Manager Components

Cisco Unified Videoconferencing Manager contains two major software components:

- Resource Manager — A scheduling component
- Network Manager — A network management component

Thus, Cisco Unified Videoconferencing Manager combines a scheduling application with a network management application.

Resource Manager Component of Cisco Unified Videoconferencing Manager

The Resource Manager component of Cisco Unified Videoconferencing Manager controls all aspects of videoconferencing and interfaces with MCUs and the gatekeeper. Resource Manager is the main control point for all incoming and outgoing video calls, and it makes all decisions regarding resource utilization and cascading of resources. In addition to basic MCU cascading to increase conference size, Resource Manager can intelligently select which MCU to use based on internally defined locations for each participant. Multisite video meetings can result in cascading MCUs local to user groups, thus creating one link across a WAN between cascaded MCUs.

The Resource Manager component of Cisco Unified Videoconferencing Manager has the following characteristics:

- Resource Manager resides on a separate Windows-based server and cannot reside on the same server with other Cisco Unified Communications components.
- Resource Manager sits between the MCUs and all other components.

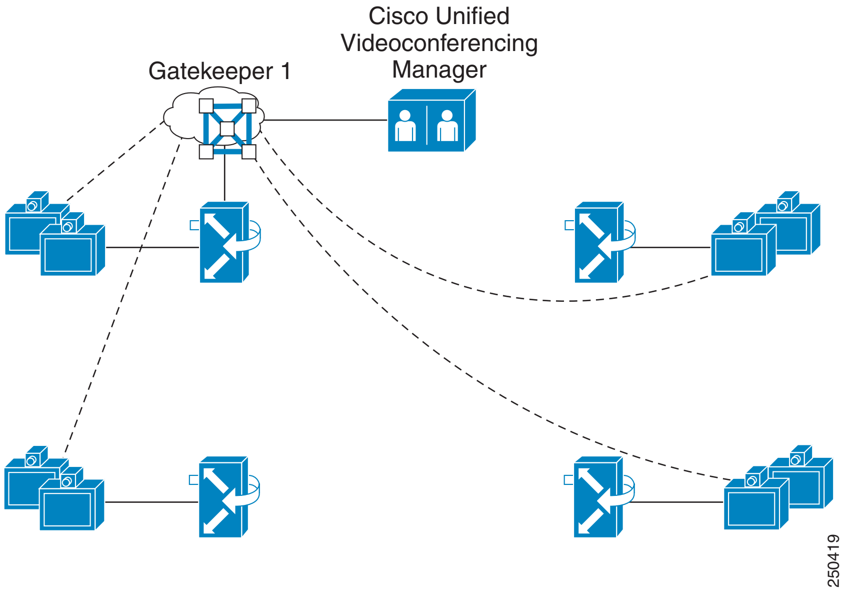
- Resource Manager contains a special integrated gatekeeper to which only the MCUs register. The MCUs must register to Resource Manager.
- All routing decisions for MCU selection and cascading are made by Resource Manager.
- Resource Manager terminates and authenticates all inbound calls, and only media is sent to the MCUs.
- Resource Manager originates and maintains all outbound calls, and only media is sent to the MCUs.
- Resource Manager requires the use of an external gatekeeper for outbound calls.
- Resource Manager does not register to the gatekeeper it uses for outbound calls.
- Resource Manager does not support a direct SIP trunk connection from Cisco Unified Communications Manager.
- Resource Manager does not support a direct H.323 gatekeeper-controlled trunk connection from Cisco Unified Communications Manager.
- All scheduled video meetings are replicated in Resource Manager, with appropriate resources reserved.
- Participant information and real-time status are relayed from the MCUs to Resource Manager.
- Other than media stream termination, only Resource Manager communicates with the MCUs.
- Video Terminals can be defined in Resource Manager and reserved within Cisco Unified Videoconferencing Manager when users create a meeting.
- Any endpoint (H.323, SCCP, or SIP) can be defined in Resource Manager as a Video Terminal.
- MCU selection can be impacted by assigning locations to MCUs and Video Terminals defined in Resource Manager.

Note The MCUs register to the Resource Manager internal gatekeeper as an MCU instead of as a gateway. The parameter is located on the MCU under Advanced H.323 Settings.

Video Endpoint and MCU Registration

When using Cisco Unified Videoconferencing Manager, all H.323 video endpoints should be configured to register to the Cisco IOS gatekeeper. Figure 8-8 shows all H.323 video endpoints configured with the Cisco IOS gatekeeper as their designated gatekeeper.

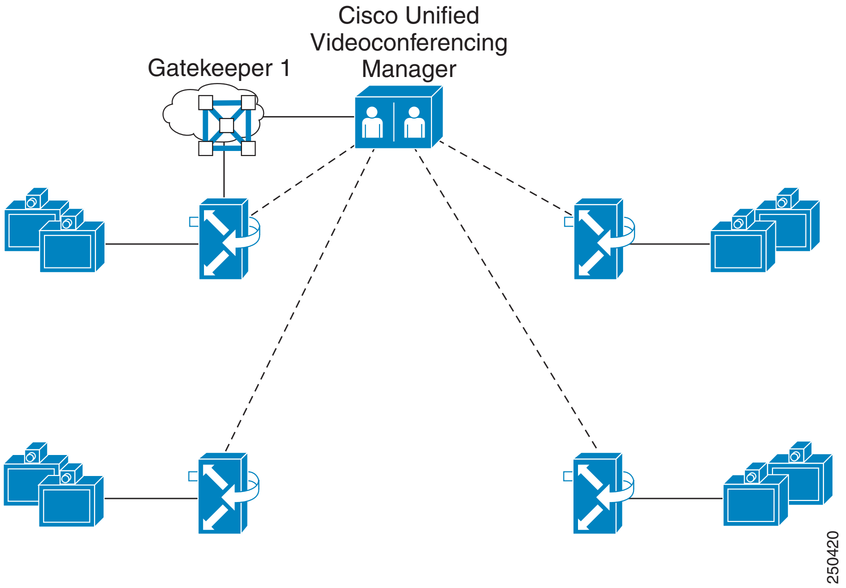
Figure 8-8 Video Endpoint H.323 Registration to a Cisco IOS Gatekeeper



You must register Cisco Unified Videoconferencing MCUs and Cisco Unified Videoconferencing Gateways with the Cisco Unified Videoconferencing Manager internal gatekeeper, which is included with the product. This preserves the virtual MCU features and conference setup, and it allows for conference control via a web interface during a meeting.

Figure 8-9 shows all Cisco Unified Videoconferencing MCUs configured with the Resource Manager internal gatekeeper as their designated H.323 gatekeeper.

Figure 8-9 MCU H.323 Registration to Cisco Unified Videoconferencing Manager



For integration of video endpoints or Cisco Unified Videoconferencing MCUs with Cisco Unified Communications Manager, refer to the *IP Video Telephony* section in the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Gatekeeper Integration with Resource Manager

The Resource Manager internal gatekeeper and the Cisco IOS Gatekeeper are configured as H.323 gatekeeper neighbors.

Example 8-1 illustrates the minimal configuration necessary for the Cisco IOS gatekeeper to integrate with the Cisco Unified Videoconferencing Manager. The Resource Manager internal gatekeeper is a remote zone from the perspective of the Cisco IOS gatekeeper.

Example 8-1 Cisco IOS Gatekeeper Minimal Configuration with Cisco Unified Videoconferencing Manager

```
gatekeeper
   zone local HQ cisco.com 10.1.1.1
   zone remote CUVM cisco.com 10.2.2.1 1719
   zone prefix CUVM 83*                  (This entry matches service prefixes configured 
on the MCUs)
   gw-type-prefix 1#* default-technology
   lrq forward-queries add-hop-count
   no use-proxy HQ remote-zone CUVM inbound-to terminal
   no use-proxy HQ remote-zone CUVM outbound-from terminal
   no use-proxy HQ default inbound-to terminal
   no use-proxy HQ default outbound-from terminal
   no shutdown
```

When integrating Resource Manager into an infrastructure gatekeeper environment with Cisco Unified Communications Manager (Unified CM), the deployment options discussed in this section will bypass Unified CM when one H.323 endpoint calls another H.323 endpoint. Having the endpoints controlled by Unified CM is a deployment option not covered here due to the focus on environments that have yet to deploy Cisco Unified CM. For information on integrating Cisco Unified CM with video endpoints, refer to the *IP Video Telephony* section in the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Resource Manager Redundancy Considerations

Resource Manager is limited to server component redundancy, and it does not have any software-level redundancy. The server on which Video Administration is deployed should contain redundant components to minimize the risk of downtime. Cisco Unified Videoconferencing Manager is deployed on an MCS-7825 server regardless of the total number of endpoints, MCUs, or ISDN gateways under control.

Redundancy of the Cisco IOS Gatekeeper via Hot Standby Router Protocol (HSRP) provides for outbound calling redundancy from Resource Manager. If the primary gatekeeper fails, call resolution requests are automatically sent to the backup gatekeeper. Resource Manager does not need to know about the backup gatekeeper or primary gatekeeper state to accomplish this failover.

Gatekeeper Clustering and Alternate Gatekeeper

With the gatekeeper clustering and alternate gatekeeper methods, an issue arises with outbound calls from Resource Manager. Outbound calls from Resource Manager are sent to the Cisco IOS Gatekeeper without any registration to that gatekeeper. This prevents Resource Manager from being aware of the gatekeeper state and the existence of an alternate gatekeeper.

Network Manager Component of Cisco Unified Videoconferencing Manager

Cisco Unified Videoconferencing Manager can provide network management of MCUs and video endpoints. The Network Manager component of Cisco Unified Videoconferencing Manager queries the Cisco IOS Gatekeeper to discover all video endpoints on the network. This allows organizations to have a global view of MCUs, gatekeepers, and video endpoints. This video infrastructure can change dynamically yet be discovered automatically. Network Manager can also provide software upgrades to MCUs and many third-party endpoints.

In addition, the Network Manager component can provide real-time status of MCU health, alarms, events, active calls, and active conferences. Moreover, recipients can be alerted to events through email, and SNMP traps can be forwarded to other management systems.

Cisco Unified Videoconferencing Manager Deployment Considerations

There are three general types of deployments for Cisco Unified Videoconferencing Manager:

- Scheduling mode

  This mode allows scheduling integration with Microsoft Outlook and simplifies all aspects of video conferencing setup. This mode can be further categorized as with or without dynamic cascading.

- Network Management mode

  The network management features of Cisco Unified Videoconferencing Manager can be used solely for endpoint and MCU management without any participation in scheduling or dynamic cascading.

- Scheduling mode with Network Management mode

  This method utilizes all of the components of the Cisco Unified Videoconferencing Manager.

Video Gateways

The Cisco Unified Videoconferencing 3545 Dual PRI, 3527 Single PRI, and 3522 4-Port BRI Gateways give enterprises the ability to connect ISDN-based H.320 systems with IP-based H.323 videoconference endpoints. These gateways provide translation services between H.320 and H.323 networks to convert multimedia information between circuit-switched ISDN and IP networks. The gateway also supports G.711, G.722, G.722.1, G.723.1, and G.728 audio codecs. In addition, voice transcoding between IP and the Public Switched Telephone Network (PSTN) is supported using G.711, G.723, or G.728. These systems enable users to videoconference with others users via the LAN or the PSTN, regardless of location. The Cisco 3500 Series Gateways also provide in-band DTMF conversion to out-of-band DTMF.

The three video gateway models provide the following features:

- Cisco Unified Videoconferencing 3522 BRI Gateway

  This gateway can be configured with two or four BRI ports. When the gateway is equipped with BRI ports, it can support calls up to 384 kbps on aggregated channels.

- Cisco Unified Videoconferencing 3527 PRI Gateway

  This gateway is a self-contained system that supports a high volume of calls over a single high-speed ISDN PRI connection, allowing dynamic allocation of its 23 B channels.

- Cisco Unified Videoconferencing 3545 Gateway Modules

  These gateway modules are either a two-port PRI T1/E1 module that supports a high volume of calls over multiple high-speed ISDN PRI connections, or a four-port serial module for IP connectivity to older ISDN H.320 videoconferencing endpoints.

Table 8-2, Table 8-3, and Table 8-4 list the maximum number of calls supported per platform.

Table 8-2 Cisco Unified Videoconferencing 3522 BRI Gateway Call Handling Capacity

Number of Calls	Capacity
1	384 kbps or 412 kbps
2	256 kbps
4	128 kbps
8	64 kbps

Table 8-3 Cisco Unified Videoconferencing 3527 PRI Gateway Call Handling Capacity

Call Type	Maximum Number of Calls Using One E1 PRI Line	Maximum Number of Calls Using One T1 PRI Line
Voice (64 kbps)	30	23
2B video (128 kbps)	15	11
6B video (384 kbps)	5	3
12B video (768 kbps)	2	1

Table 8-4 Cisco Unified Videoconferencing 3545 PRI Gateway Call Handling Capacity

Call Type	Maximum Number of Calls Using One E1 PRI Line	Maximum Number of Calls Using One T1 PRI Line	Maximum Number of Calls Using Two E1 PRI Lines	Maximum Number of Calls Using Two T1 PRI Lines
Voice (64 kbps)	30	23	60	46
2B video (128 kbps)	15	11	30	23
6B video (384 kbps)	5	3	10	7
12B video (768 kbps)	2	1	5	3

Outbound Dialing Service Prefixes

Video gateways must be configured with service prefixes to define the speed of outgoing calls and calling routing to the video gateway. In telephony systems, dialing 9 to access an outside line is very common. In order to keep dialing strings consistent with existing voice dial plans, Cisco recommends using 9# for video gateway service prefixes. Using the # in the service prefix ensures that ISDN users do not access the IVR and hairpin the call back out the ISDN network. The # is used as a delimiter by the gateway and prevents hair pinning from the ISDN network.

From the users' perspective, they will have to dial the service prefix, which in this case is equivalent to an access code, followed by the ISDN number of the H.320 video unit. For this configuration, a local (intra-zone) gateway is used whenever one is present. In zones that do not contain a gateway, the administrator should assign a gateway in another zone as the primary gateway for the local zone. Configure location request (LRQ) forwarding or a static hopoff statement to route all calls to a zone with a gateway for PSTN access.

Network Load Balancing

The Cisco Unified Videoconferencing Gateway supports the network load balancing (LAN to PSTN) feature. This feature allows users to build a pool of gateways for PSTN access. Network load balancing creates a larger number of access lines serviced by a single set of service prefixes.

Gatekeepers can perform load balancing on the network using feedback from the gateway in the form of Resource Availability Indication (RAI) messages that inform the gatekeeper of gateway resource availability. If the gateway is unavailable, the gatekeeper performs line hunting operations to route the call to an alternate gateway. When you set the gateway for RAI and Resource Availability Confirmation (RAC), it sends periodic RAI messages that inform the gatekeeper of the current resource availability in the gateway. The gatekeeper responds with Resource Available Confirmation (RAC) messages to acknowledge receipt of the RAI messages.

To implement network load balancing, you configure multiple gateways with identical service prefixes and register them with the same gatekeeper. Outbound PSTN calls are sent to the gateways based on resource availability, using RAI and RAC. In the gateway configuration, you set utilization parameters based on gateway resource percentages.

The main gateway configuration parameters for line hunting are:

- Utilization (percent load) for sending RAI ON message
- Utilization (percent load) for sending RAI OFF message

The RAI ON message tells the gatekeeper that resources are running low on the gateway that sent the message, and the gatekeeper should not forward any more calls to that gateway. (The default for sending a RAI ON is 80% load.) The RAI OFF message tells the gatekeeper that there are enough available resources on the gateway, and calls can be forwarded to the gateway again. (The default for sending a RAI off is 60% load.) Periodic RAI messages are sent from the gateway to the gatekeeper when one of the above thresholds is not achieved in a specified period of time (The default period for these messages is 30 seconds.)

Cisco Unified Border Element

The Cisco Unified Border Element is a Cisco IOS software component that incorporates the gatekeeper and border element functions for an H.323 video network. The Cisco IOS gatekeeper provides endpoint registration and call resolution for large H.323 video networks. The border element provides topology hiding, network isolation, call admission control and Quality of Service (QoS). The border element is increasingly used for dynamic call admission control with RSVP devices or on behalf of devices that do not support RSVP. Deployments with a border element simplify Network Address Translation (NAT) and firewall integration.

Gatekeeper

The Cisco gatekeeper performs all call routing and address registration (RAS) for all H.323 video components. The gatekeeper is one of the most important components in an H.323 network because it is the central management device for the H.323 video network and it performs functions required for a successful H.323 video deployment. Some of the most commonly used functions of the Cisco IOS gatekeeper include:

- H.323 component registration and call routing

  The gatekeeper registers the IP address, E.164 address, H.323-ID, device type, and signaling ports for all the video infrastructure components. This registration allows the gatekeeper to provide call routing for all devices that are registered with the it.

- Bandwidth management

  Managing video bandwidth on IP networks is an essential feature of any gatekeeper. By setting the following bandwidth parameters, you can configure the Cisco gatekeeper to manage the bandwidth in a zone, between zones, or per call.

  – Inter-zone: Total bandwidth allowed from a local or default zone to and from all other zones.

  – Remote: Total bandwidth allowed from all local zones to and from all remote zones.

  – Session: Bandwidth allowed per session in a zone.

  – Total: Total bandwidth allowed in a zone.

  – Resource Availability Indicator (RAI): Endpoints supporting RAI update the gatekeeper of availability or unavailability.

  – Circuit-id: Calls with this circuit-id call identifier are limited by the gatekeeper to max-calls per endpoint.

  – Max-calls: Maximum number of calls that the gatekeeper will permit per endpoint.

- Authentication, authorization, and accounting (AAA) support

  The Cisco gatekeeper works in conjunction with Remote Access Dial-In User Service (RADIUS) and Terminal Access Controller Access Control System (TACACS) servers to provide authentication of devices and accounting via call detail recording (CDR).

- Via-zone

  The gatekeeper can include the border element in the call flow by using the via-zone with **invia** and **outvia** configurations to include the border element for incoming calls or outgoing calls.

The Cisco gatekeeper also supports features such as the following, which enable users to build reliable and scalable H.323 networks:

- Hot Standby Router Protocol (HSRP) enables administrators to build a standby gatekeeper that becomes active if the primary gatekeeper fails.
- Gatekeeper Update Protocol (GUP) enables multiple gatekeeper entities to behave as a single gatekeeper cluster, providing greater resilience for endpoint registrations and calls. (Limits on scalability must be taken into consideration.)
- Directory Gatekeeper, or Location Request (LRQ) forwarding, enables administrators to build large multi-tier networks, minimizing the configuration required in the lower-tier gatekeepers. When a call is made in a lower-tier zone and a match is not found, the call is automatically forwarded up to the directory gatekeeper for resolution. (For more information on directory gatekeepers, refer to the Call Routing chapter.) Figure 8-10 on page 8-20 illustrates a network configured with two regional directory gatekeepers, one at Site E and another at Site F.

Redundancy

A video network needs redundancy for the endpoints to register and for call resolution. Failure in call control devices can impact the ability to do videoconferencing. Any of the following methods can provide redundancy:

- Hot Standby Router Protocol (HSRP)

  HSRP can be used to provide redundancy at the IP layer for the gatekeeper. Thus, endpoints can use one IP address for the gatekeeper address. Failures of devices providing hot standby for the gatekeeper IP address do not impact the endpoints drastically. The blackout period after the failure of the primary device and until the endpoints are fully registered and functional with the standby device, depends on the device time-outs and registration retries. This method is recommended for the directory gatekeeper because its main function is call resolution and it might not have any endpoints registered. (For more information on directory gatekeepers, see Routing Inter-Zone Calls Using a Directory Gatekeeper, page 7-9.)

- Alternate gatekeeper

  One of the simplest methods of providing redundancy is to configure the endpoints with a list of gatekeepers in order of preference, so that if the first gatekeeper in the list fails, the devices register with the next gatekeeper in the list. However, not all endpoints have this support. Cisco Unified Communications Manager, Cisco Unified Border Element, and Cisco Gateways support alternate gatekeeper functionality.

- Gatekeeper Update Protocol (Cisco GUP)

  The Cisco Gatekeeper Update Protocol (GUP) enables gatekeepers in the cluster to update information on registrations and calls in the cluster. Alternate gatekeeper support is needed in the endpoints. When an endpoint registers or unregisters with its primary gatekeeper, that gatekeeper updates the remaining gatekeepers in the cluster about the change. Similar information is shared when the gatekeeper resolves a call. If any gatekeeper in the cluster fails, the endpoint registers to the alternate gatekeeper that the primary gatekeeper had informed when the endpoint initially registered. With GUP, if a gatekeeper failure occurs, active calls are not disconnected. GUP is recommended where the number of endpoints is not very large. However, most video endpoints do not support alternate gatekeeper, therefore they cannot support the use of GUP.

Scalability

Call routing scalability can be achieved by using the hierarchical model. Common regions or areas are grouped together to be serviced by access gatekeepers which then point to a directory gatekeeper. The directory gatekeeper may not have any endpoints registered to it and may have a role of routing calls between access gatekeepers or to hop-off zones. The access gatekeeper may have endpoints registered to it or, in larger deployments, may have an additional level of gatekeepers to which it can route calls.

HSRP

Hot Standby Router Protocol (HSRP) enables a set of routers with the Cisco Unified Border Element to work together as a single virtual gatekeeper or border element. You can implement this feature by creating a *phantom* router that has its own IP and MAC addresses.

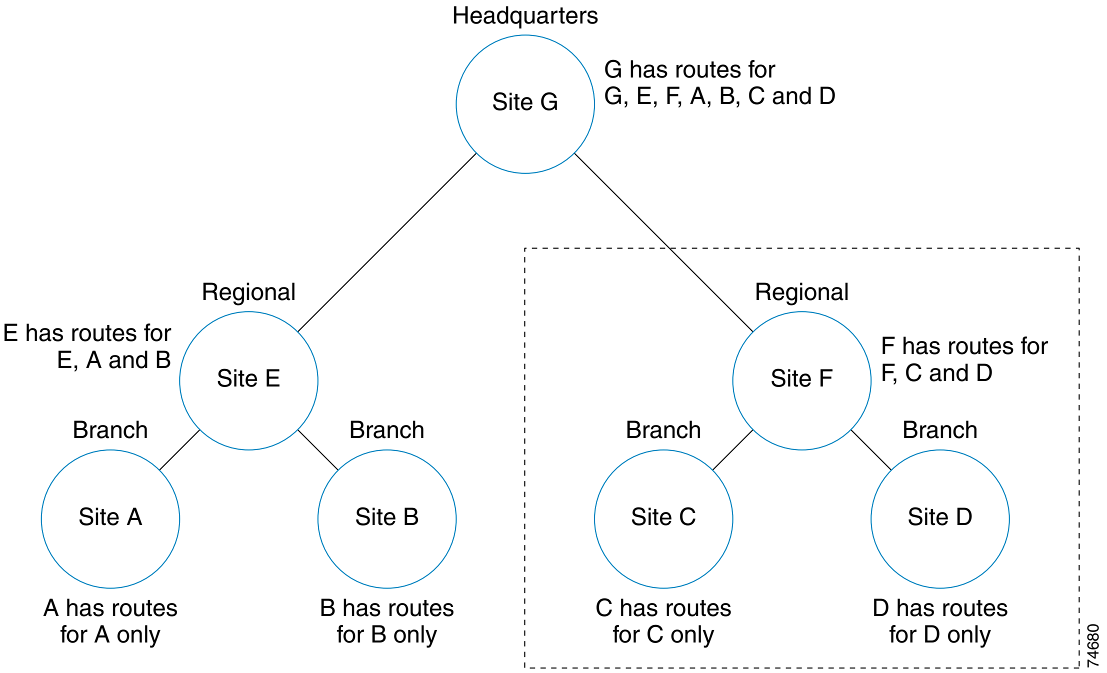
Based on the priority given by the network administrator, one of the HSRP gatekeepers in each group is selected to be active and the other to be standby. The gatekeeper with the highest priority serves as the active gatekeeper. The active gatekeeper does the work for the HSRP phantom. If an end node sends a packet to the phantom's MAC address, the active gatekeeper receives that packet and processes it. If an end node sends an Address Resolution Protocol (ARP) request for the phantom's IP address, the active gatekeeper replies with the phantom's MAC address.

The HSRP gatekeepers (both active and standby) watch for *hello* packets to monitor the status of each other. The gatekeeper group learns the hello and hold timers, as well as the standby address to be shared, from the active gatekeeper. If the active gatekeeper becomes unavailable for any reasons (such as power failure, scheduled maintenance, or failure to respond to three successive hello packets), the standby gatekeeper assumes the active role transparently within a few seconds. Because the new active gatekeeper assumes both the IP and MAC addresses of the phantom, video terminal registrations time out, and the terminals re-register with their same IP address to the newly active gatekeeper.

Note When configuring gatekeepers and border elements on routers supporting HSRP, configure the border elements to register with the virtual, or phantom, IP address of the gatekeeper pair. This configuration enables both border elements to register with the active gatekeeper so that video calls are load-balanced between the two devices. If the primary router fails, the border element on the standby router registers with the now active gatekeeper, and calls are forward through it. The border element configured on the primary router will not re-register with the standby router if the primary router fails.

Note Gatekeeper clustering is not supported in a videoconferencing environment. For clustering to work, video endpoints would have to support alternate gatekeepers, but currently there are no video terminals with this support.

Figure 8-10 Network with Two Directory Gatekeepers



Border Element

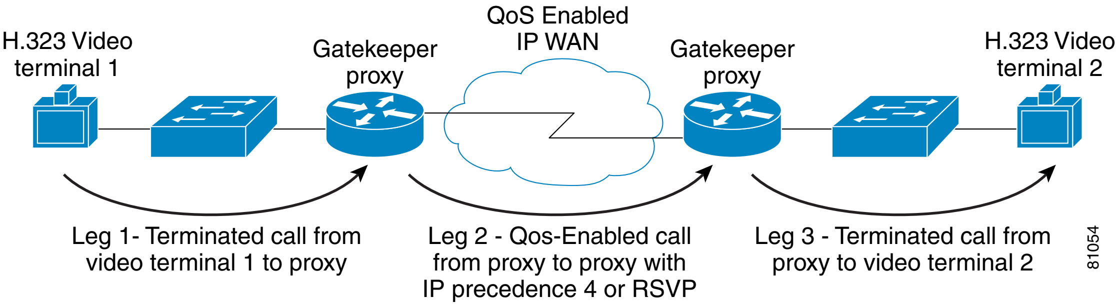
The Cisco Unified Border Element has the ability to proxy call signaling and media by termination and re-origination. The border element can register with the gatekeeper and be inserted in the call flows. The Cisco Unified Border Element has the following functionality:

- Network isolation is achieved by re-origination of signaling and media for the call. You can use multiple interfaces to provide enhanced isolation.
- Classification of signaling, audio, and video traffic can be done with IP Precedence, DSCP, and Resource Reservation Protocol (RSVP).
- Call admission control with the help of the Resource Availability Indicator (RAI) as a gateway function is widely used. In addition, call treatment based on CPU, memory, total calls, circuit ID capacity, and gatekeeper bandwidth can be used in various combinations to provide efficient call admission control. RSVP can be used for dynamic call admission control.
- Topology hiding is used for inter-working various types of signaling and protocol.
- The border element can provide ease of deployment with firewalls and NAT because it can be a single trusted entity in either the inside network or the DMZ. With its protocol awareness for NAT and firewalls, the border element can also reduce management overhead.

Note Because the gatekeeper is an optional device, the border element can also be used as a standalone device. The border element will have call routing management overhead in this case.

Figure 8-11 illustrates a call across a WAN link through border elements.

Figure 8-11 Call Across a WAN Through Border Elements


H.323 Video terminal 1
Gatekeeper proxy
QoS Enabled IP WAN
Gatekeeper proxy
H.323 Video terminal 2
Leg 1- Terminated call from video terminal 1 to proxy
Leg 2 - Qos-Enabled call from proxy to proxy with IP precedence 4 or RSVP
Leg 3 - Terminated call from proxy to video terminal 2
81054

Note Single-legged Cisco Unified Border Element is supported in Cisco IOS Release 12.3(15)T or later.

Collaboration for Desktop Sharing

Deployments that want to leverage collaboration in their video networks can use any of the following methods to share user desktop screens.

H.239-Based

With this method, an additional video channel is established by the endpoints. The desktop screen is sent using this additional video channel during the call. In this method, the endpoint capabilities determine the screen sharing resolution and characteristics. All participants must support this mechanism. Audio and video for the conference is through regular channels by the endpoints, as in point-to-point calls.

The following design considerations apply to this method:

- Endpoints must support H.239.
- Intermediate devices such as call agents, border elements, NAT devices, or firewalls must support H.239.
- Additional WAN bandwidth might be needed and must be provisioned and accounted for.

Server and Client Communication

This method is commonly used by collaboration systems. Cisco Unified MeetingPlace, WebEx, and other systems use a similar method. In this method, a user-side client is used that can run on a PC or laptop and that makes desktop screen sharing possible. The server is the application server providing the services to the clients. Users that do not have access to a PC cannot see the screen being shared. Audio and video for the conference is through regular channels by the endpoints, as in point-to-point calls.

The following design considerations apply to this method:

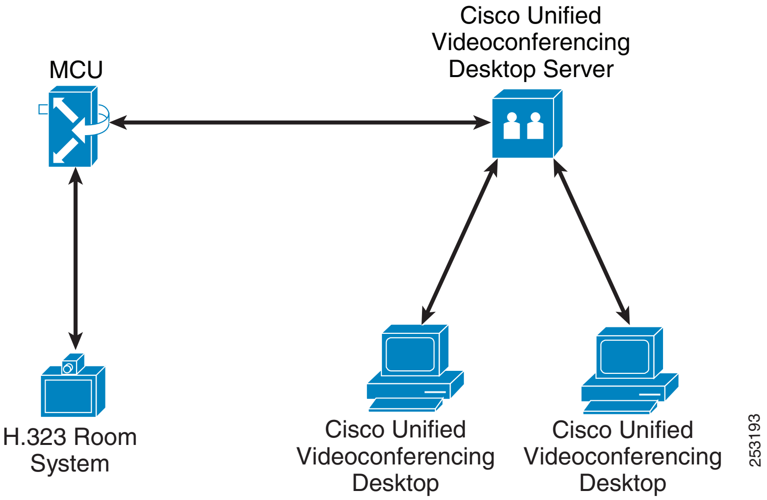
- Deployments need to consider the operating system and browsers for the user client on the desktop or the laptops.
- Consider the way the client software can be deployed for internal and external devices. Most systems will check installations or upgrades before users join the collaboration sessions through browsers.
- Intermediate devices such as NAT devices or firewalls must support traversal for the user-side client, if applicable.

- Quality of Service classification and call admission control are required for calls.
- Additional WAN bandwidth might be needed and should be accounted for and provisioned to accommodate the collaboration traffic.

Cisco Unified Videoconferencing Desktop

The Cisco Unified Videoconferencing Desktop solution is similar to the server and client communication. (See Figure 8-12.)The client component takes care of the video, audio, and desktop sharing capability. The server component is the application server for collaboration. In addition to this, it interacts with the MCU that hosts the video conference and enables traditional endpoints in the conference to view the content shared through the desktop. On the client side it uses a web protocol such as HTTP/HTTPs to tunnel the data, and on the endpoint side it uses H.239 for screen sharing. In this way, the Cisco Unified Videoconferencing Desktop provides a combination of H.239 and server/client communication for the collaboration solution.

Figure 8-12 Cisco Unified Videoconferencing Desktop


Cisco Unified Videoconferencing Desktop Server
MCU
H.323 Room System
Cisco Unified Videoconferencing Desktop
Cisco Unified Videoconferencing Desktop
253193

The following design considerations apply to this method:

- Video endpoints must support H.239.
- Intermediate devices such as call agents, border elements, NAT devices, or firewalls must support H.239.
- Deployments need to consider the operating system and browsers for the user client on the desktop or the laptops.
- Consider the way the client software can be deployed for internal and external devices. Most systems will check installations or upgrades before users join the collaboration sessions through browsers.
- Quality of Service classification and call admission control are required for calls.
- Intermediate devices such as NAT devices or firewalls must support traversal for the Cisco Unified Videoconferencing Desktop client, as applicable.
- Additional WAN bandwidth might be needed and should be accounted for and provisioned to accommodate the collaboration traffic.

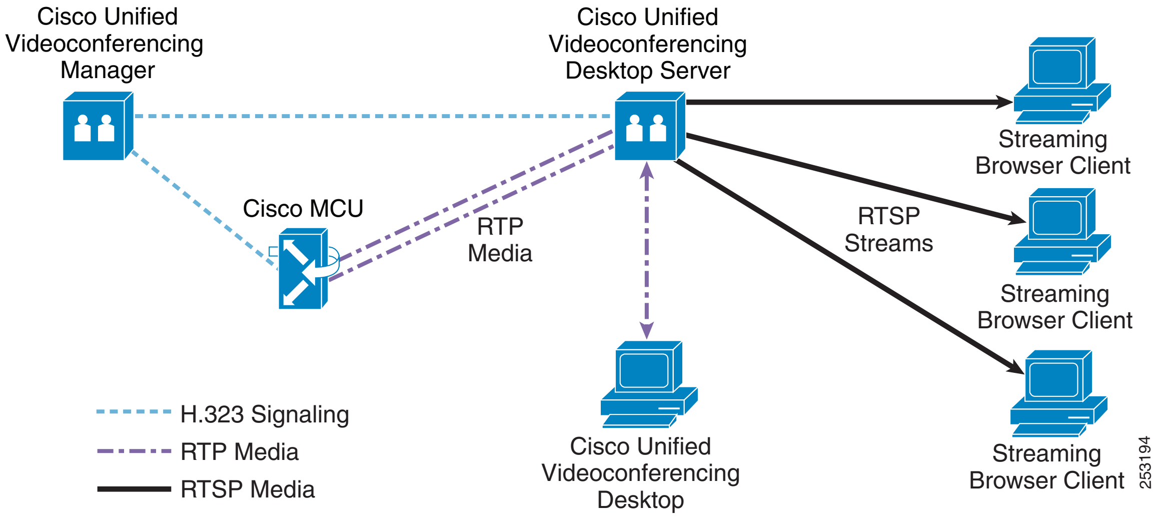
- Cisco recommends co-locating the desktop server and the conference MCU because the Cisco Unified Videoconferencing Desktop server initiates a video channel to the MCU for every desktop client.

Conference Streaming

Some conferences such as training sessions, large team meetings, or panel discussions have the key people such as the trainer, team leader, or moderators and panel numbers participate interactively in the conference, while the large majority of users are viewing participants only. Rather than have every participant join the conference with their video endpoints, streaming such conferences is a more efficient and scalable way to present the conference to viewers. (See Figure 8-13.) This not only optimizes the conferencing resources, but it also allows users to view the conference from simple devices such as media players that support streaming. For such solutions Real Time Streaming Protocol (RTSP) is commonly used.

The Cisco Unified Videoconferencing Desktop server supports streaming of live conferences. The participants who will be transmitting in the conference can join the conference through conventional endpoints that support interactive calls. The server uses the active conference from the MCU and converts it into an RTSP stream. The viewers can also view the conference using common players such as Quicktime player, VLC media player, Windows Media player, Real media player, or any other players that support RTSP.

Figure 8-13 Conference Streaming



The following design considerations apply to conference streaming:

- The streaming media might lag behind actual conference.
- Quality of Service classification and policing traffic are required.
- Unicast streaming may be used with a small number of viewers.
- Multicast streaming can be used if there is a need to conserve WAN bandwidth or if there is a large number of users viewing conferences.
- Additional WAN bandwidth may be needed and should be accounted for and provisioned to accommodate the streaming traffic.

Conference Recording

Recording the conference provides users with the capability to view the conference at a later time. Recording servers are separate application servers that store the recording and make it available for later viewing. The Cisco Unified Videoconferencing Desktop server supports recording of conferences. Conferences can be recorded by selecting that option when scheduling the conferences or by the moderator through the user desktop.

The following design considerations apply to conference recording:

- The recording server should be deployed on a separate server. Co-hosting with other Cisco Unified Videoconferencing server should be done only after carefully considering the server capacity and scalability.
- Recordings must be stored and managed on the desktop recording server. Access can be through the Recording access webpage. Security of these recordings depends on the operating system of the recording server.
- The recording server capacity should be planned based on the type of recordings, the frequency of recording, and how long the recordings will be available to users before the system archives them.
- Enterprises need to define a conference recording storage and retention policy.

Conference recording is dependent on the recording bitrate configured on the recording server. Table 8-5 provides brief guidelines on the storage needed for recording meetings.

Table 8-5 Storage Guidelines for Recording Servers

Recording Bitrate (kbps)	Amount of Data Recorded (MBytes/min)[1]
256	2.3
384	3.5
512	4.7
768	7.0
1024	9.2

[1] These values include a factor of 20% for overhead.

Use the following general formula to calculate the minimum server storage capacity needed for recording conferences:

Minimum Storage Capacity = [Recorded Mbytes/min (User corresponding values from Table 8-5 based on the recording bitrate)] ∗ [Recording time per day in minutes] ∗ [Number of days] ∗ [Number of simultaneous conferences]

For example, if the recording rate is 256 kbps and you want to be able to record 15 full days (24 hours) of 5 simultaneous conferences, the minimum required storage capacity would be:

Minimum Storage Capacity = 2.3 MB/min ∗ [60 min/hr ∗ 24 hr/day] ∗ [15 days] ∗ [5 conferences]

Minimum Storage Capacity = 248.4 GBytes

Integration and Interoperability

The Cisco Unified Videoconferencing MCU can be used for various system integration needs and to fulfill interoperability requirements with various systems, as described in the following sections.

Telepresence Integration

The Cisco Unified Videoconferencing 5000 Series or 3500 Series MCUs provide Cisco Telepresence Multipoint Switch integration to traditional videoconferencing networks. This allows H.323 terminals, desktop telephony, and other videoconferencing devices to participate in a conference with Cisco Telepresence systems. The Cisco Telepresence Multipoint Switch extends the conference to the Unified Videoconferencing MCU. This enables traditional and video telephony devices to join into the MCU conference, while the Cisco Telepresence system is conferenced through the Cisco Telepresence Multipoint Switch as a single conference.

For additional details, refer to the Cisco Telepresence Multipoint Switch documentation at

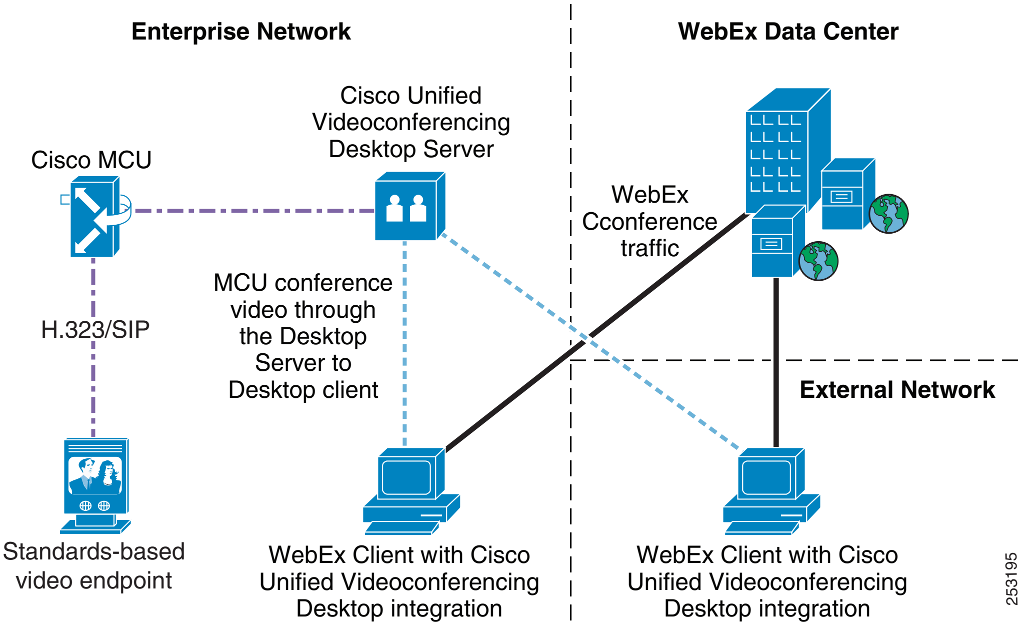
http://www.cisco.com/en/US/products/ps7315/tsd_products_support_series_home.html

WebEx Integration

Through the deployment of the entire Cisco Unified Videoconferencing portfolio, especially the Desktop Streaming Server, WebEx clients on WebEx site versions T27 or later can be configured to use the Advanced Video plug-in. Advanced Video enables on-premise room systems, video telephony, and even telepresence systems to participate in a video conference with webcams both on-premise and over the Internet.

Integrating the Cisco Unified Videoconferencing MCU of the enterprise with WebEx leverages the Cisco Unified Videoconferencing Desktop solution to bring in the video of traditional desktop systems and traditional videoconferencing systems and presenting it in a WebEx collaboration session. In this way, the Native Video support in WebEx gets replaced with the one from the enterprise MCU. (See Figure 8-14.)

Figure 8-14 WebEx Integration


Enterprise Network
WebEx Data Center
Cisco Unified Videoconferencing Desktop Server
Cisco MCU
WebEx Cconference traffic
H.323/SIP
MCU conference video through the Desktop Server to Desktop client
External Network
Standards-based video endpoint
WebEx Client with Cisco Unified Videoconferencing Desktop integration
WebEx Client with Cisco Unified Videoconferencing Desktop integration
253195

Microsoft Office Communicator

Enterprises that deploy Microsoft Office Communicator need integration for desktop video. The desktop client uses Cisco Unified Videoconferencing Desktop server integration with Microsoft Office Communicator to display video from the MCU in Microsoft Office Communicator, while using video from the desktop cameras in the conference as user video. Scheduling conferences is done through Outlook Plug-in integration.

Lotus Sametime

Enterprises that deploy Lotus Sametime need integration for desktop video. The desktop client uses Cisco Unified Videoconferencing Desktop server integration with Sametime to display video from the MCU, while using video from the desktop cameras in the conference as user video. Scheduling conferences is done through Lotus Notes Plug-in integration.

For more details on the integrations, refer to the latest Install and Upgrade Guides for the Cisco Unified Videoconferencing Desktop, available at

http://www.cisco.com/en/US/products/ps7088/prod_installation_guides_list.html

Connecting the Video Network with External Untrusted Networks

Enterprises want to leverage the benefits of videoconferencing not only for their internal users but also with external enterprises and users. The following design considerations can help identify the best approach.

External Connectivity

External connectivity to the enterprise network is a key requirement. The external devices can also be devices at the enterprise partners, vendors, or suppliers within a closed user group (CUG), but they might be untrusted by the enterprise. In some cases the external devices may be devices on the networks of other enterprises or on the untrusted Internet.

Note Quality of Service for calls might require special consideration because QoS on the Internet is best-effort and cannot be guaranteed. Enterprise QoS can address quality within the enterprise network. For additional details, refer to the *Enterprise QoS Solution Reference Network Design Guide*, available at http://www.cisco.com/go/designzone.

Dial Plan

A key consideration for external connectivity is how the enterprise routes the calls. Does the enterprise act as a peer with other enterprises to route calls between themselves through call agents, gatekeepers, or SIP proxies, based on a protocol of choice?

Enterprises can also assign E.164 addresses to external endpoints so that calls can be routed only within the enterprise network to those external endpoints, thus controlling the dial plan to them. In this case, you must ensure that external networks cannot use the enterprise-assigned dial plan to reach these external devices. Enterprises can enforce rules or registration to prevent unauthorized access.

Enterprises that wish to support external devices that dial by IP address for conferences can use mechanisms that direct callers to a fixed enterprise number and present them to an auto-attendant that provides the endpoint information to select and join conferences using DTMF digits. Such deployments can leverage the Cisco Unified Border Element, which can service calls that do not have a called number in the call setup and can present that call with an auto-attendant number as the called number to an MCU with auto-attendant service.

External Endpoints or Devices

Identifying the type of endpoint can significantly change the way connectivity to external networks can be achieved.

Traditional endpoints or desktop endpoints support standards-based protocol such as H.323 or SIP and can register with the gatekeeper or SIP proxy as applicable. These endpoints then can register with internal trusted devices. Cisco recommends firewalls that support application-level gateway inspection. Topology hiding devices can provide a more efficient method for deployment and management.

Desktops or laptops with no clients might not support standards-based protocols but might use tunneling through HTTP/HTTPs to the communication servers. To service such clients, the Cisco Unified Videoconferencing Server can be deployed in the DMZ network. Users can access the server using the Microsoft Internet Explorer (IE) browser and can authenticate through credentials and the conference personal identification number (PIN) to participate in the conference. The lightweight software for the desktop or the PC can be downloaded and installed if needed.

Firewalls and Network Address Translation (NAT)

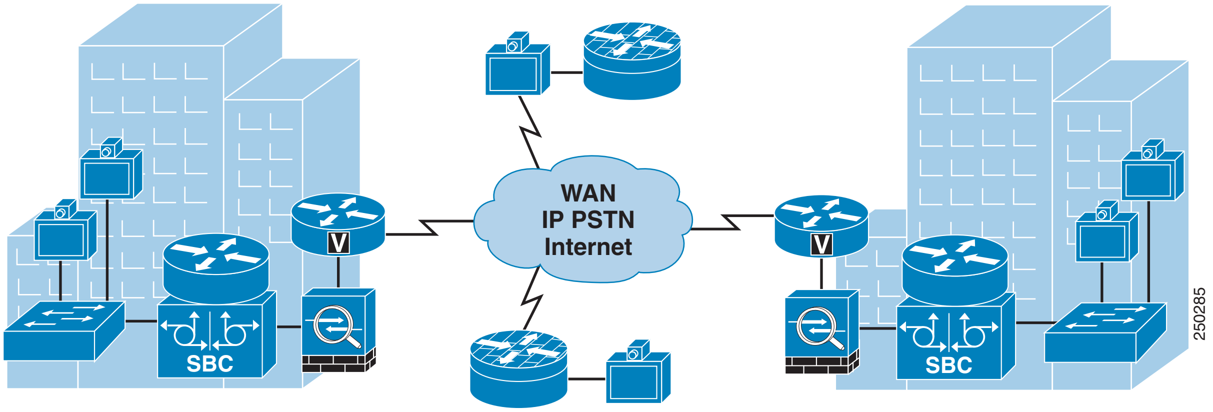
Firewalls provide security for unauthorized access to internal trusted networks. Cisco recommends that the most secure way to prevent unauthorized access for VoIP traffic is by inspection of the VoIP registrations and calls. Enterprises may need to deploy application-aware firewalls. These firewalls can then inspect on VoIP protocols too.

H.323 uses multiple signaling ports. An application-aware firewall can inspect the H.323 messages and open respective ports for the call to proceed. For an H.323 call to take place, it must first open an H.225 connection on TCP port 1720 using Q.931 signaling. Next, the H.245 management session is established. While this session can take place on a separate channel from the H.225 setup, it can also be done using H.245 tunneling, which takes the H.245 messages and embeds them in the Q.931 messages in the previously established H.225 channel. Next, the H.245 session opens dynamically assigned ports for the UDP-based RTP and RTCP video and audio data streams. The port numbers can range from 1024 to 65535. Because the port numbers are not known in advance, and because it would defeat the purpose of a firewall to open all these ports, a firewall must be able to snoop the H.323 data stream in order to open the additional ports needed for the call. This snooping is also known as *stateful inspection*. Firewalls that support H.323 message inspection in order to open just the needed ports per call are termed *application-aware firewalls*.

An additional problem encountered with most firewalls is the use of Network Address Translation (NAT). Within H.323, the H.225 and H.245 signaling channels make heavy use of the embedded IP address. For example, assume a terminal has a private address of 10.1.1.125, which gets translated to 206.165.202.125 when it tries to place a call to an H.323 terminal with an IP address of 206.165.201.78 on the outside network. The terminal on the outside still receives the private address within the H.225 signaling stream. Because this is a non-routable address, an attempt to make a connection back will fail. One way to work around this problem is to use an H.323-aware NAT firewall, which can rewrite the addresses in the signaling payload.

Using the border element in NAT or firewall environments allows administrators to target a single IP address to terminate all H.323 video calls. All incoming and outgoing video calls that access the public network will use the border element. With the Cisco ASA Firewall, administrators can enable H.323 fix-up and allow UDP port 1720 traffic to access the IP address of the border element. Without the border element, administrators would have to configure UDP port 1720 to all videoconferencing devices and have static NAT for each device, which may not be scalable with a large number of video endpoints in the network. If you use a Cisco IP/VC 352*x* or 3540 gateway, port 1820 must be configured for the videoconferencing devices. Figure 8-15 illustrates the call flow in a network with NAT and a firewall.

Figure 8-15 Call Flow with NAT and a Firewall



When using the gatekeeper on the border element, you can have external devices on the outside untrusted network register to the gatekeeper through the UDP 1718 or 1719 port pinhole to the gatekeeper for RAS registrations. The application-aware firewall inspects RAS signaling and opens pinholes for H.225 call signaling based on the embedded ports in the RAS messages. Additional inspection of H.225 can open ports dynamically for H.245 communication and then RTP media, respectively, through the application-aware firewall. Organizations can use external gatekeepers to ensure E.164 number resolution.

CHAPTER 9

Multi-Zone WAN Case Study

Last revised on: October 30, 2009

This chapter provides an example of a typical WAN multi-zone model deployed in an enterprise environment.

In this case study, the enterprise is a health care provider with locations spread across the United States. Five locations currently use ISDN-based videoconferencing. The enterprise has a T1 to each site and would like to install new H.323 videoconferencing units and utilize their existing WAN bandwidth. Each site contains a minimum of three video units, and the enterprise has standardized on 384 kbps as their call data rate. The enterprise requires multipoint calls as well as the ability to call off-net to their clients.

Network Topology

Currently the enterprise in this example has five sites in the United States, consisting of Sacramento CA, Los Angeles CA, Dallas TX, Columbus OH, and Chicago IL. Each site connects back to Columbus, the headquarters, with a T1 link. The bandwidth utilization on all the connections is low. The enterprise has just upgraded their WAN routers at remote sites to Cisco 2851 routers to support voice, video, and data in the near future. The headquarters is connected with a T1 line to the internet. Currently, all videoconferencing units are directly connected to an IMUX with three BRI lines, allowing bonded 384-kbps calls. The Columbus site contains an H.320 multipoint control unit (MCU) with three PRI lines supporting multipoint calls among the sites. Some remote specialists use a public gatekeeper on the internet to reach the health care provider. Figure 9-1 illustrates the current IP network, and Figure 9-2 illustrates the current videoconferencing network.

Figure 9-1 Current IP Network

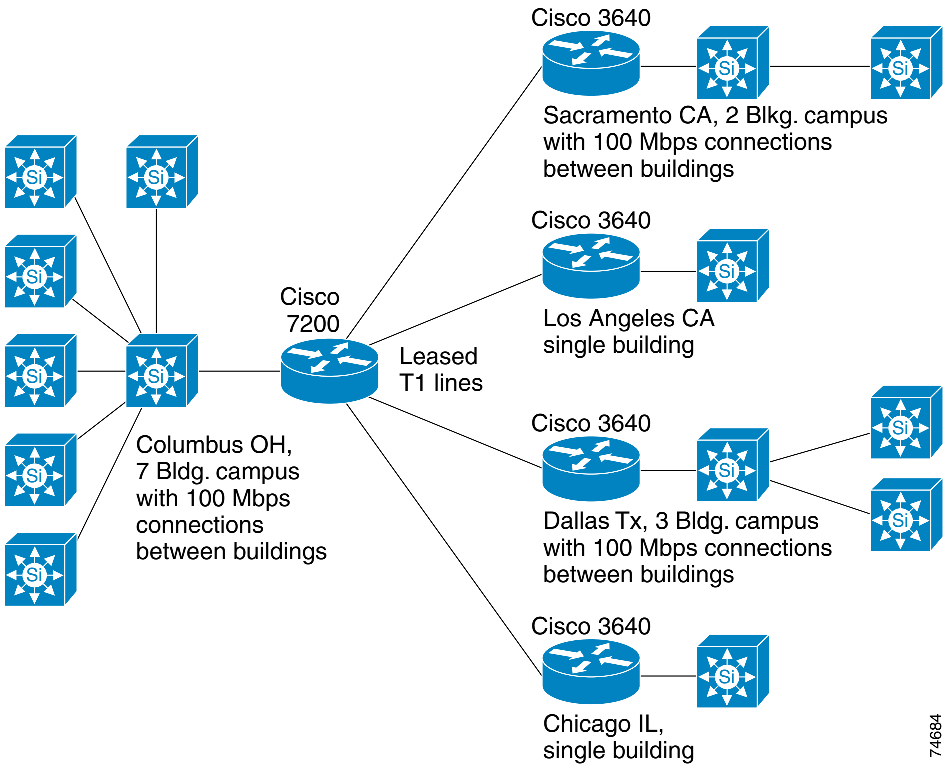
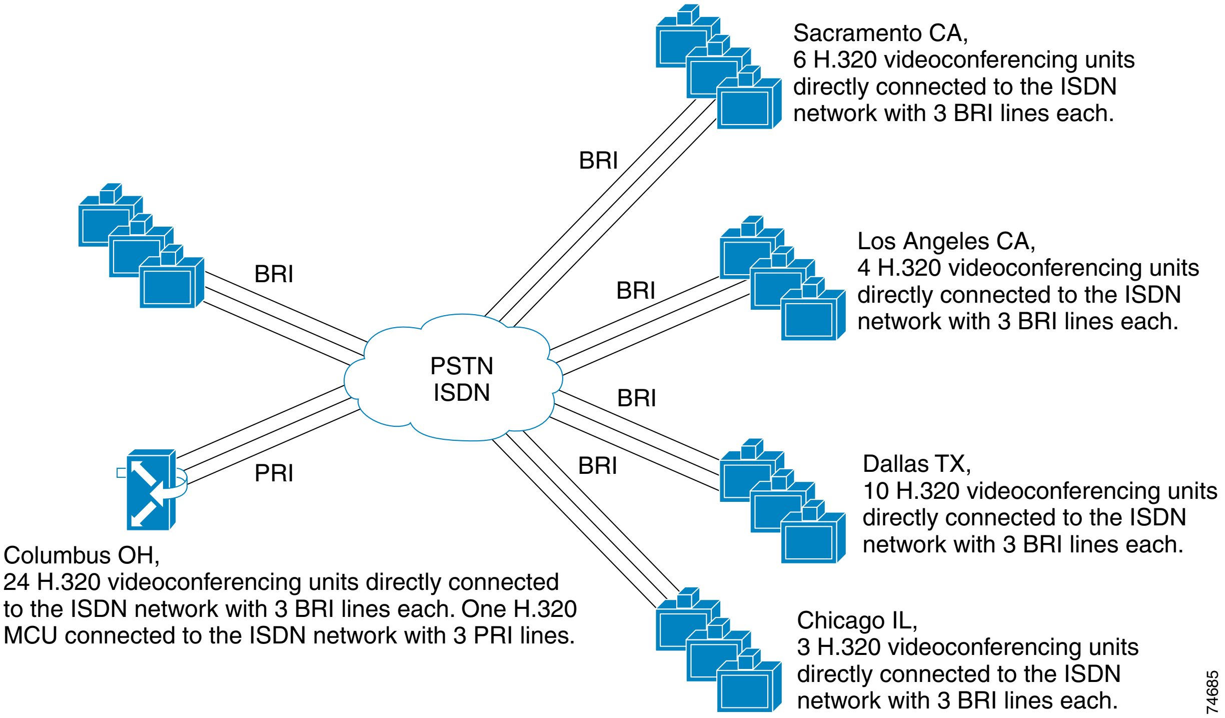

Cisco 3640
Sacramento CA, 2 Blkg. campus with 100 Mbps connections between buildings
Cisco 3640
Los Angeles CA single building
Cisco 7200
Leased T1 lines
Cisco 3640
Dallas Tx, 3 Bldg. campus with 100 Mbps connections between buildings
Cisco 3640
Chicago IL, single building
Columbus OH, 7 Bldg. campus with 100 Mbps connections between buildings
Si
74684

Figure 9-2 Current Videoconferencing Network


Sacramento CA, 6 H.320 videoconferencing units directly connected to the ISDN network with 3 BRI lines each.
BRI
Los Angeles CA, 4 H.320 videoconferencing units directly connected to the ISDN network with 3 BRI lines each.
BRI
BRI
PSTN ISDN
BRI
PRI
BRI
Dallas TX, 10 H.320 videoconferencing units directly connected to the ISDN network with 3 BRI lines each.
Columbus OH, 24 H.320 videoconferencing units directly connected to the ISDN network with 3 BRI lines each. One H.320 MCU connected to the ISDN network with 3 PRI lines.
Chicago IL, 3 H.320 videoconferencing units directly connected to the ISDN network with 3 BRI lines each.
74685

Network Design

The network outlined in the previous section is a classic WAN multi-zone model. There is sufficient WAN bandwidth, and each site contains three or more video terminals. In this network, a gatekeeper and border element are located at each site. Directory gatekeeper services are configured, and Hot Standby Routing Protocol (HSRP) is used for gatekeeper redundancy at the Columbus site. Quality of service (QoS) and call admission control (CAC) need to be configured in the network to ensure video quality. Firewall security is used to connect to the internet, and the border element is the trusted device through the firewall.

Quality of Service (QoS)

End-to-end QoS is a key factor in a successful deployment. The enterprise in this example has decided to use an H.323 video terminal that supports marking of IP Precedence. The Columbus, Sacramento, and Dallas sites have upgraded to Catalyst 6500 switches. In these three sites, LAN QoS will be configured; the remaining three sites will support LAN QoS when the switches at those sites are also upgraded. All video units will be connected to 10/100 Ethernet ports.

All video terminals, gateways, and MCUs are configured to mark IP Precedence 4. In Columbus, Sacramento, and Dallas, trust boundaries are set on the Catalyst 6500 switches. Video gateways and MCUs are also installed in Columbus, Sacramento, and Dallas. For gateways and MCUs that do not support IP Precedence, IP Precedence is marked and a trust boundary is set on the Catalyst 6500 ports to which the gateways and MCUs are connected. Gateways are also installed in Los Angeles and Chicago.

Priority queues are configured on all WAN routers and are provisioned for 920 kbps. This guarantees that bandwidth is available for two 384-kbps calls. An access list entry is also added on the WAN router to set the entrance criterion for the priority queue. Only video traffic received from the border element is admitted to the priority queue.

The gatekeeper at each site is configured to use the local border element for all inter-zone calls. The border element rewrites IP Precedence 4 and provides a single access point to the configured priority queue.

For more information regarding network QoS, refer to the *Enterprise QoS Solution Reference Network Design Guide*, available at http://www.cisco.com/go/designzone.

Call Admission Control

Call admission control (CAC) must be implemented for inter-zone calls, and it is also a good idea to configure CAC for intra-zone calls. Enabling CAC for inter-zone calls guarantees that the bandwidth limits provisioned on the priority queues are not exceeded. If the provisioned bandwidth for the priority queue on the WAN route is exceeded, all video calls in the queue will be affected. The gatekeeper at each site contains the following three bandwidth statements for CAC.

```
bandwidth total {default | zone <zone-name>} <bandwidth-size>
bandwidth remote 1536
bandwidth session default 768
```

It is important to note that the bandwidth is calculated in half-duplex mode, so the call data rate must be doubled. A 384-kbps call is represented as 768 kbps in the bandwidth statements. The **bandwidth total** command allows administrators to limit the bandwidth within a single local zone, or for all local zones by adding the **default** statement. The remote bandwidth (available bandwidth to and from any remote zone) is limited to 1536 kbps, or two 384-kbps calls. The bandwidth per session is limited to 768 kbps, or one 384-kbps call.

Figure 9-3 illustrates QoS and CAC points for Columbus, Figure 9-4 illustrates QoS and CAC points for Dallas and Sacramento, and Figure 9-5 illustrates QoS and CAC points for Los Angeles and Chicago.

Figure 9-3 QoS and CAC for Columbus

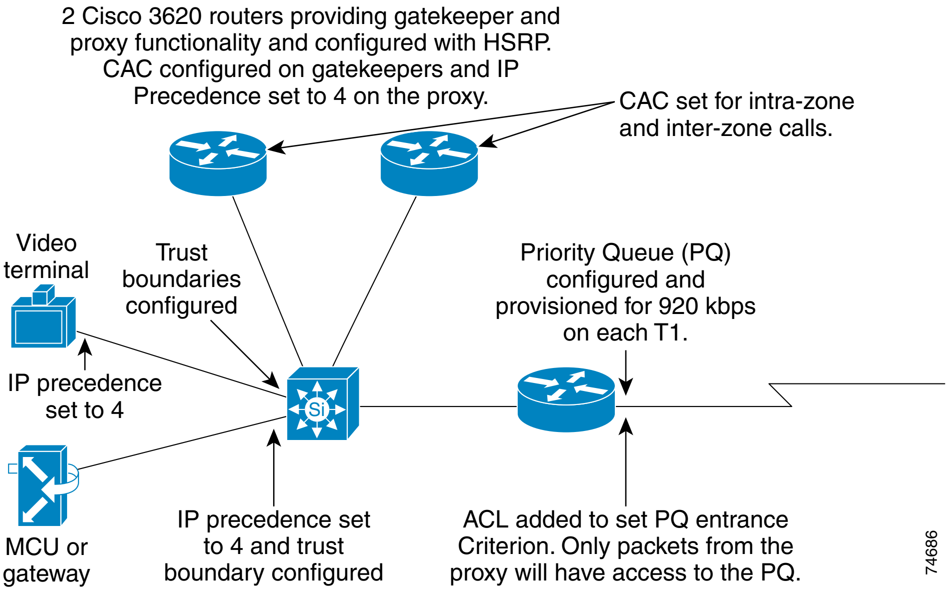

2 Cisco 3620 routers providing gatekeeper and proxy functionality and configured with HSRP. CAC configured on gatekeepers and IP Precedence set to 4 on the proxy.
CAC set for intra-zone and inter-zone calls.
Video terminal
Trust boundaries configured
Priority Queue (PQ) configured and provisioned for 920 kbps on each T1.
IP precedence set to 4
Si
MCU or gateway
IP precedence set to 4 and trust boundary configured
ACL added to set PQ entrance Criterion. Only packets from the proxy will have access to the PQ.
74686

Figure 9-4 QoS and CAC for Dallas and Sacramento

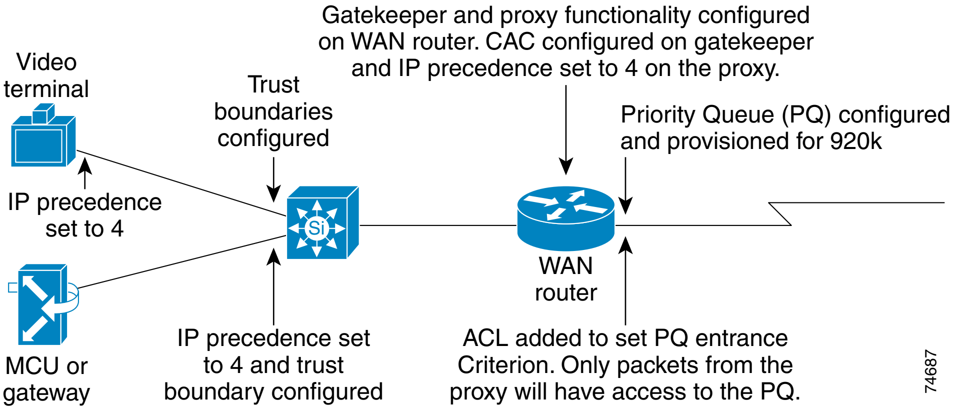
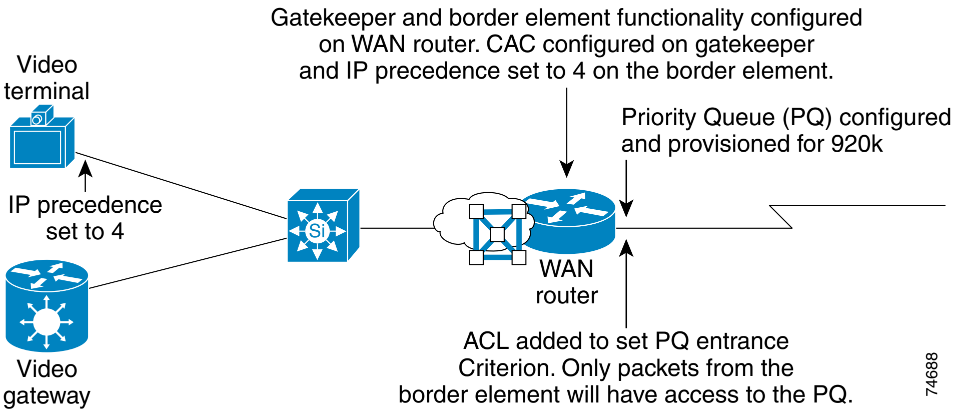

Gatekeeper and proxy functionality configured on WAN router. CAC configured on gatekeeper and IP precedence set to 4 on the proxy.
Video terminal
Trust boundaries configured
Priority Queue (PQ) configured and provisioned for 920k
IP precedence set to 4
Si
WAN router
MCU or gateway
IP precedence set to 4 and trust boundary configured
ACL added to set PQ entrance Criterion. Only packets from the proxy will have access to the PQ.
74687

Figure 9-5 QoS and CAC for Los Angeles and Chicago


Video terminal
IP precedence set to 4
Video gateway
Gatekeeper and border element functionality configured on WAN router. CAC configured on gatekeeper and IP precedence set to 4 on the border element.
Priority Queue (PQ) configured and provisioned for 920k
WAN router
ACL added to set PQ entrance Criterion. Only packets from the border element will have access to the PQ.
Si
74688

Dial Plan

When deciding on a dial plan, it is always a good idea to start with the incoming PSTN call routing. In our example, we have created five zones that all contain video gateways. DID is used to route incoming calls to video terminals. IVR is used to route calls from the video gateways in Columbus, Dallas, and Sacramento, to their local MCUs. If, for some reason, one or more of the zones in our example did not contain a gateway, IVR for routing all incoming PSTN calls would have been a better choice.

Zone Prefixes

The zone prefix for each zone is based on the local area code. Area codes are unique, and users are familiar with the numbering structure. In our configuration there is a single zone in each site, so the zone prefixes are based on area codes. If more than one zone were required in a single area code, longer zone prefixes could be used. (Refer to Zone Prefix Design, page 6-7.) The zone prefixes in this network are:

- Columbus = 614
- Sacramento = 916
- Dallas = 972
- Chicago = 847
- Los Angeles = 213

Service Prefixes

Service prefixes must be configured for all MCUs and video gateways. As described in the chapter on Dial Plan Architecture, it is a good idea to reserve a block of numbers for video gateways and a block of numbers for MCUs. In this case, the enterprise has chosen to standardize on 384-kbps calls; this makes service prefixes for gateways very simple. The obvious choice would be to use 9 for all PSTN calls, but that would cause routing problems in the Sacramento and Dallas zones. The Sacramento zone prefix is 916, and overlapping gateway service prefixes and zone prefixes will cause routing problems. There are two options; reserve another block of numbers other than 9*, or use a service prefix such as 9#. In this case, we have chosen 9# for PSTN access in all zones. Any time a user tries to access the WAN, the dial string will start with 9#.

For MCU service prefixes, 8* is reserved, and the zone prefix is appended to associate it with the zone in which the MCU resides. The MCU service prefix in the Sacramento zone is 9168*, and in Los Angeles it is 2128*. Table 9-1 lists the service prefixes chosen for different types of calls on the MCU. (These service prefixes are used in every zone, and the zone prefix is appended.)

Table 9-1 MCU Service Prefixes

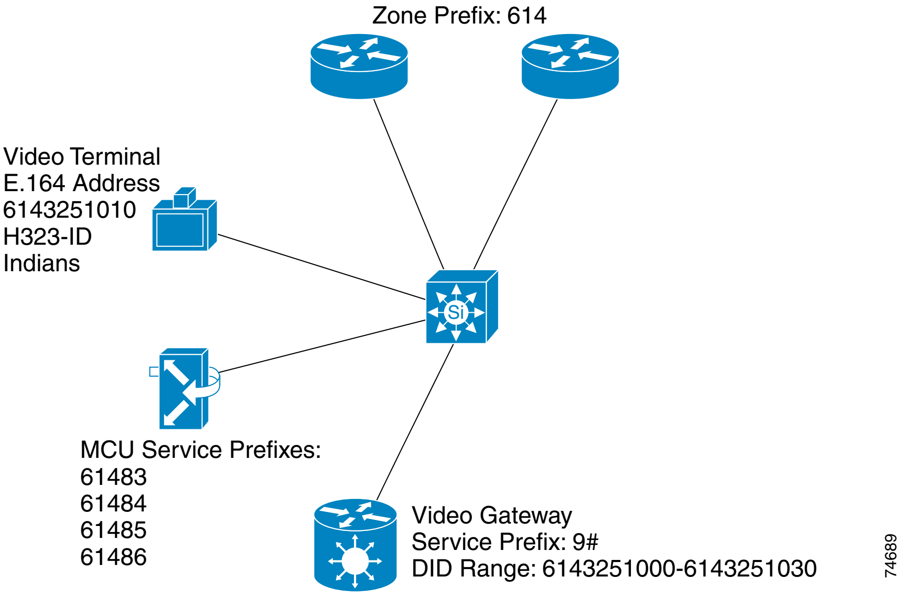
Service Prefix	Data Rate	Number of Parties	Video Format	Continuous Presence
83	384 kbps	3	H.261	No
84	384 kbps	4	H.261	Yes
85	384 kbps	5	H.261	No
89	384 kbps	9	H.261	No

E.164 Addresses and H.323-IDs

The carrier provides E.164 addresses for this enterprise. Because DID has been chosen for incoming PSTN routing method, the enterprise will order blocks of DID numbers with each PRI line. Each video terminal is assigned a valid DID number for its E.164 address. In Columbus, there are 24 video terminals, and thirty DID numbers are ordered with the PRI line for Columbus. (The extra six numbers are for expansion.) In Los Angeles and Chicago, DID numbers off the BRI lines will be used as E.164 addresses. (See Video Infrastructure, page 9-7, for the video components at each site.)

H.323-IDs are based on the name of the conference room where the video system resides. Since the video terminals may be moved from room to room, H.323-IDs will not be used for dialing. The enterprise is using a global address book that will display all of the IP video terminals on the network. Users can choose to dial from the address book or manually enter the E.164 address of the unit being called. Figure 9-6 illustrates the Dial plan in Columbus.

Figure 9-6 Dial Plan for Columbus



Video Infrastructure

When deciding on location and number of video components, it is important to understand the enterprise's needs. This enterprise made it clear that less than ten percent of all video calls were off-net calls. The number of video calls placed daily ranges from 10 to 15, and most calls are multipoint. For this reason, the enterprise decided to go with video gateways at each site and MCUs in Columbus, Dallas, and Sacramento. The following sections list the video components for each site. All IP video calls for the enterprise are received at the headquarters.

Columbus

- IP Video Terminals, 24

  The current 24 H.320 video systems are over three years old and will be replaced with new H.323 group systems that support IP Precedence.

- MCUs, 4

  The four MCUs will be configured in a stack, allowing one set of service prefixes to be shared by all four MCUs.

- Video Gateway PRI, 1

  A single PRI gateway will be installed with 30 DID numbers that will be assigned to the IP video terminals. All 10 digits will be passed to the gateway by the carrier, allowing each video terminal to register with a 10-digit fully qualified E.164 address. IVR will be enabled and used for PSTN access to MCU conferences. One DID number will have to be reserved for IVR calls.

Sacramento

- IP Video Terminals, 6

  The existing six H.320 video systems are over three years old and will be replaced with new H.323 group systems that support IP Precedence.

- MCU, 1

  A single MCU will be located on the Sacramento campus for local, on-site multipoint calls. The Sacramento campus is in the process of adding another building and possibly adding two or three additional IP video terminals. The MCU will also allow multiple video terminals to participate in an off-campus multipoint call while consuming the bandwidth of only a single call. This will be done by cascading a Sacramento MCU conference with a Columbus MCU conference.

- Video Gateway, 1

  A single PRI gateway will be installed with 10 DID numbers that will be assigned to the IP video terminals. All 10 digits will be passed to the gateway by the carrier, allowing each video terminal to register with a 10-digit fully qualified E.164 address. IVR will also be enabled and used for PSTN access to MCU conferences. One DID number will have to be reserved for IVR calls.

Dallas

- IP Video Terminals, 10

  The existing 10 H.320 video systems are over three years old and will be replaced with new H.323 group systems that support IP Precedence.

- MCU, 1

  A single MCU will be located on the Dallas campus for local, on-site multipoint calls. The MCU will also allow multiple video terminals to participate in an off-campus multipoint call while consuming the bandwidth of only a single call. This will be done by cascading a Dallas MCU conference with a Columbus MCU conference.

- Video Gateway, 1

  A single PRI gateway will be installed with 15 DID numbers that will be assigned to the IP video terminals. All 10 digits will be passed to the gateway by the carrier, allowing each video terminal to register with a 10-digit fully qualified E.164 address. IVR will also be enabled and used for PSTN access to MCU conferences. One DID number will have to be reserved for IVR calls.

Los Angeles

- IP Video Terminals, 4

  The existing four H.320 video systems are over three years old and will be replaced with new H.323 group systems that support IP Precedence.

- MCU, 0

  Los Angles will not have a local MCU.

- Video Gateway, 1

  A single BRI gateway will be installed with four BRI lines. Each video terminal will receive a DID number from one of the BRI lines. IVR will not be enabled on the gateway.

Chicago

- IP Video Terminals, 3

  The existing three H.320 video systems are over three years old and will be replaced with new H.323 group systems that support IP Precedence.

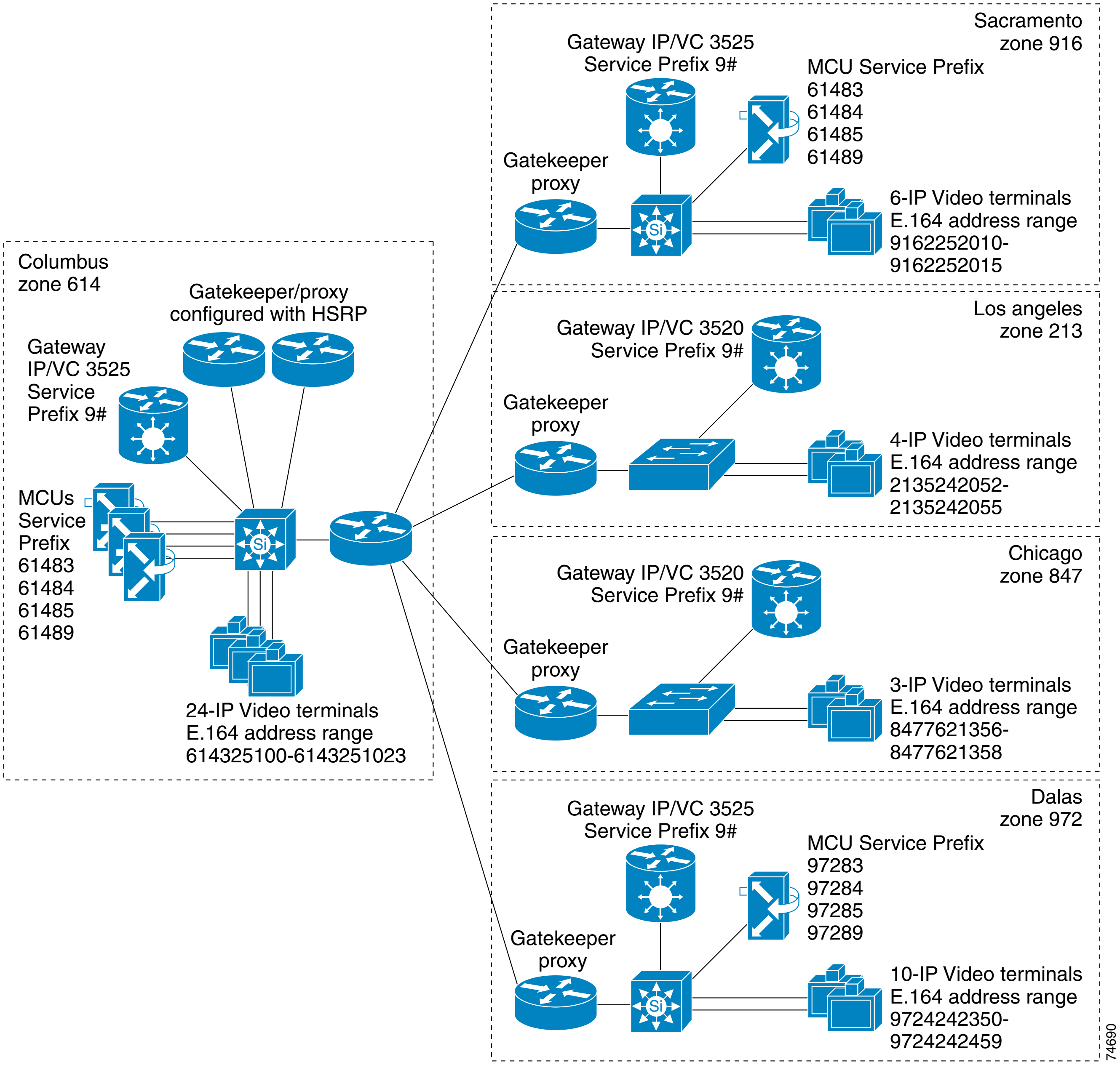
- MCU, 0

  Chicago will not have a local MCU.
- Video Gateway, 1

  A single BRI gateway will be installed with three BRI lines. Each video terminal will receive a DID number from one of the BRI lines. IVR will not be enabled on the gateway.

Figure 9-7 illustrates the video components and dial plan for the new IP video network.

Figure 9-7 Dial Plan for Example Video Network



APPENDIX A

Interworking with Session Initiation Protocol (SIP)

Last revised on: October 30, 2009

The most popular protocol used by enterprises for video conferencing is H.323. That protocol has been used for a long time, and capabilities such as H.235 for security and H.239 for data sharing make it popular. However, newer endpoints and systems support SIP too, which is seeing increased adoption.

Enterprises that want to use a common protocol for voice, video, instant messaging, and presence would like to standardized on SIP, and new deployments with call agents that have SIP servers providing call resolution find the support of SIP in video communication devices desirable. Enterprises that use Secure RTP (sRTP) for call encryption and do not need data sharing capabilities through the call, may prefer to deploy endpoints with SIP protocol.

A SIP network can consist of various SIP servers that provide the following functionality:

- Registration

  The SIP registrar server provides registration for the endpoints. Usually the SIP server that provides the dial plan capabilities would also support the registration functionality.

- Signal proxy

  SIP proxy servers proxy the call signaling through them. The proxy servers can support two modes:

  - Proxy the initial call signaling and then have the SIP servers control the call.
  - Proxy the call and the signaling on behalf of the SIP server. In this mode the SIP proxy operates in a record route mode.

- Location services.

  The location services inform the SIP registrar of a new location of an endpoint. In scalable networks, this is useful to associate properties of that location with the endpoint.

A typical SIP network consists of all these functions in a single server, or multiple servers can be used with the functionality distributed to each server according to its role.

Note Call agents such as Cisco Unified Communications Manager support registration of SIP endpoints and connectivity with SIP trunks. Cisco Unified Communications Manager supports additional protocols such as H.323 and Skinny Client Control Protocol (SCCP), and it can also be used to provide interworking between H.323 networks and SIP networks while supporting large-scale deployments.

Give adequate consideration to the MCUs, gateways, and other systems or servers that support SIP protocol to service all the deployment requirements, so that your system design can support all the required calls, conferences, and collaboration services needed by your enterprise.

GLOSSARY

Last revised on: October 30, 2009

Numerics

802.1P
802.1Q 802.1P and 802.1Q are the standards proposed by the inter-working task groups of the 802 standards committee. 802.1Q is the IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridge Local Area Network (VLAN). 802.1P is the IEEE Standard for Local and Metropolitan Area Networks – Supplement to Media Access Control (MAC) Bridges: Traffic Expending and Dynamic Multicasting Filtering.

A

ARQ Admission Request

B

Border Element A device that is used in the periphery of the network to separate two different networks and to serve as a demarcation device

BRI Basic Rate Interface

C

CAC Call admission control

Cascade The process of connecting two or more MCUs to create a larger conference.

Codec Coder-decoder, for digitizing voice and video. Compression algorithms can also be used during the digitizing process.

CoS Class of Service

cRTP Compressed Real-time Transport Protocol

D

DID Direct Inward Dialing

DSCP Differentiated Services Code Point is an Internet Engineering Task Force (IETF) standard that uses six bits in the ToS (Type of Service) field of the IPv4 header to specify class of service for each packet.

DTMF Dual Tone Multi-Frequency

E

E.164 Address format used for H.323 devices

EMP Enhanced Media Processor

G

Gatekeeper Used for H.323 registration, call routing, and admission control

G.711 G.711 pulse code modulation (PCM) encoding provides 64 kbps analog-to-digital conversion using mu-law or a-law.

H

H.261 Video codec protocol

H.263 Video codec protocol

H.323 Standard protocol for audio, video, and data communications across IP-based networks

H.323-ID Alphanumeric identifier assigned to an H.323 video terminal

hopoff Command added to a Cisco gatekeeper for static inter-zone routing

HSRP Hot Standby Routing Protocol

I

IMUX Inverse Multiplexer

IP Internet Protocol

IP Precedence IP Precedence uses the three precedence bits in the ToS (Type of Service) field of the IPv4 header to specify class of service for each packet.

ISDN	Integrated Services Digital Network
IVR	Interactive Voice Response

L

LAN	Local Area Network
LLQ	Low Latency Queuing is a QoS mechanism that ensures the timely queuing of critical, delay-sensitive traffic.
LRQ	Location Request

M

MC	Multipoint Controller
MCM	Multimedia Conference Manager
MCU	Multipoint Control Unit, used for video conferences containing more than two endpoints
MP	Multipoint Processor
MSN	Multiple Subscriber Number

P

PRI	Primary Rate Interface
Proxy	H.323-to-H.323 gateway used for assigning QoS and security access
PSTN	Public Switched Telephone Network

Q

QoS	Quality of Service

R

RAS	Registration, Admission, and Status protocol
RRQ	Registration Request
RSVP	Resource Reservation Protocol

RTCP	Real-time Control Protocol
RTP	Real-time Transport Protocol

S

SCCP	Skinny Client Control Protocol
Service Prefix	A digit string used to identify a service on an MCU or gateway
Stacking	Grouping MCUs to obtain a larger number of multipoint conferences

T

ToS	Type of Service

W

WAN	Wide Area Network
WRED	Weighted Random Early Detection
WRR	Weighted Round Robin

Z

Zone	A logical group of H.323 infrastructure components managed by a single gatekeeper
Zone Prefix	A digit string used to identify a group of H.323 devices

INDEX