Gateways

Revised: February 7, 2017

Gateways provide a number of methods for connecting a network of collaboration endpoints to the Public Switched Telephone Network (PSTN), a legacy PBX, or external systems. Voice and video gateways range from entry-level and standalone platforms to high-end, feature-rich integrated routers, chassis-based systems, and virtualized applications.

This chapter explains important factors to consider when selecting a Cisco gateway to provide the appropriate protocol and feature support for your voice and video network. The main topics discussed in this chapter include:

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	Revision Date
Toll fraud mitigation	Dial Plan Protection and Toll Fraud Mitigation, page 5-25	February 7, 2017
Multiple Expressway-Es and GeoDNS	Scaling the Expressway Solution, page 5-27	February 7, 2017
Cisco Expressway	Cisco Expressway, page 5-17	June 14, 2016
Encryption	Authentication and Encryption, page 5-24	January 19, 2016
Dial plan protection and Call Processing Language (CPL)	Dial Plan Protection and Toll Fraud Mitigation, page 5-25	January 19, 2016
Minor updates for Cisco Analog Voice Gateways VG204XM and VG300 Series; Cisco Analog Telephone Adapter (ATA) 190; and Cisco Integrated Services Routers (ISRs) 3900E, 4300 Series, and 4400 Series	Various sections of this chapter	June 15, 2015
Cisco Unified Border Element	Cisco Unified Border Element, page 5-16	June 15, 2015
Cisco Expressway	Cisco Expressway, page 5-17	June 15, 2015

Types of Cisco Gateways

Until approximately 2006, the only way for an enterprise to connect its internal voice and video network to services outside the enterprise was by means of TDM or serial gateways to the traditional PSTN. Cisco offers a full range of TDM and serial gateways with analog and digital connections to the PSTN as well as to PBXs and external systems. TDM connectivity covers a wide variety of low-density analog (FXS and FXO), low density digital (BRI), and high-density digital (T1, E1, and T3) interface choices.

Starting around 2006, new voice and video service options to an enterprise became available from service providers, often as SIP trunk services. Using a SIP trunk for connecting to PSTN and other destinations outside the enterprise involves an IP-to-IP connection at the edge of the enterprise's network. The same functions traditionally fulfilled by a TDM or serial gateway are still needed at this interconnect point, including demarcation, call admission control, quality of service, troubleshooting boundary, security checks, and so forth. For voice and video SIP trunk connections, the Cisco Unified Border Element and the Cisco Expressway Series fulfill these functions as an interconnection point between the enterprise and the service provider network.

This chapter discusses in detail Cisco TDM and Serial gateway platforms and Cisco Expressway. Cisco Unified Border Element is also discussed briefly.

Cisco TDM and Serial Gateways

Cisco gateways enable voice and video endpoints to communicate with external telecommunications devices. There are two types of Cisco TDM gateways, analog and digital. Both types support voice calls, but only digital gateways support video.

Cisco Analog Gateways

There are two categories of Cisco analog gateways, station gateways and trunk gateways.

- Analog station gateways

  Analog station gateways connect Unified CM to Plain Old Telephone Service (POTS) analog telephones, interactive voice response (IVR) systems, fax machines, and voice mail systems. Station gateways provide Foreign Exchange Station (FXS) ports.

- Analog trunk gateways

  Analog trunk gateways connect Unified CM to PSTN central office (CO) or PBX trunks. Analog trunk gateways provide Foreign Exchange Office (FXO) ports for access to the PSTN, PBXs, or key systems, and E&M (recEive and transMit, or ear and mouth) ports for analog trunk connection to a legacy PBX. Analog Direct Inward Dialing (DID) and Centralized Automatic Message Accounting (CAMA) are also available for PSTN connectivity.

Cisco analog gateways are available on the following products and series:

- Cisco Analog Voice Gateways VG204XM and VG300 Series (VG310, VG320, VG350) all support SCCP.
- Cisco Integrated Services Routers Generation 2 (ISR G2) 2900, 3900, 3900E, and 4000 Series (4300 and 4400) with appropriate PVDMs and service modules or cards. PVDM4s utilized by ISR 4000 Series do not support video today.
- Cisco Analog Telephone Adapter (ATA) 190 (SIP only) provides a replacement for the ATA188.

Cisco Digital Trunk Gateways

A Cisco digital trunk gateway connects Unified CM to the PSTN or to a PBX via digital trunks such as Primary Rate Interface (PRI), Basic Rate Interface (BRI), serial interfaces (V.35, RS-449, and EIA-530), or T1 Channel Associated Signaling (CAS). Digital T1 PRI and BRI trunks can be used for both video and audio-only calls.

Cisco digital trunk gateways are available on the following products and series:

- Cisco Integrated Services Routers Generation 2 (ISR G2) 1900, 2900, 3900, 3900E, 4300, and 4400 Series with appropriate PVDMs and service modules or cards
- Cisco TelePresence ISDN GW 3241 and MSE 8321
- Cisco TelePresence Serial GW 3340 and MSE 8330

Cisco TelePresence ISDN Link

The Cisco TelePresence ISDN Link is a compact appliance for in-room ISDN and external network connectivity supporting Cisco TelePresence EX, MX, SX, and C Series endpoints. While traditional voice and video gateways are shared resources that provide connectivity between the IP network and the PSTN for many endpoints, each Cisco ISDN Link is paired with a single Cisco endpoint. For more information, refer to Cisco TelePresence ISDN Link documentation available at

http://www.cisco.com/en/US/products/ps12504/tsd_products_support_series_home.html

TDM Gateway Selection

When selecting a gateway for your voice and video network, consider the following factors:

- Gateway Protocols for Call Control, page 5-4
- Core Feature Requirements, page 5-6

Gateway Protocols for Call Control

Cisco Unified Communications Manager (Unified CM) supports the following IP protocols for gateways:

- Session Initiation Protocol (SIP)
- H.323
- Media Gateway Control Protocol (MGCP)
- Skinny Client Control Protocol (SCCP)

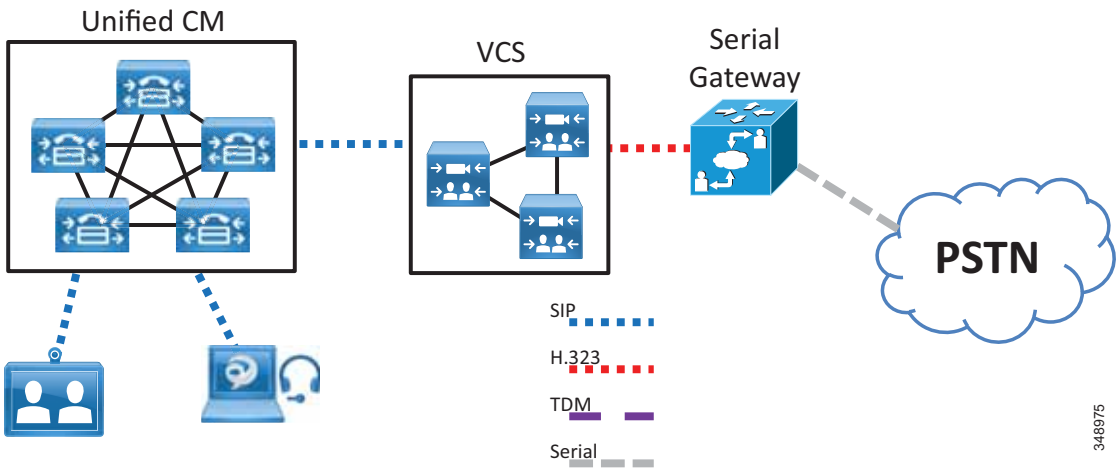
Cisco Expressway Series and Cisco TelePresence Video Communication Server (VCS) support the following IP protocols for gateways:

- Session Initiation Protocol (SIP)
- H.323

SIP is the recommended call signaling protocol because it aligns with the overall Cisco Collaboration solution and the direction of new voice and video products. However, protocol selection might depend on site-specific requirements and the current installed base of equipment. Existing deployments might be limited by the gateway hardware or require a different signaling protocol for a specific feature.

For example, placement of certain Cisco video gateways within the network depends upon the existing call control architecture. Both the Cisco ISDN and serial gateways are optimized for video calls and were initially designed to work with the Cisco VCS. The Cisco TelePresence Serial Gateway 8330 and 3340 platforms are recommended to register with a Cisco VCS using H.323, as shown in Figure 5-1.

Figure 5-1 Cisco TelePresence Serial Gateway Registered to Cisco VCS



The Cisco TelePresence ISDN Gateway 8321 and 3241 support SIP beginning with version 2.2. The Cisco 8321 and 3241 gateways can either register to VCS using H.323 (as shown in Figure 5-2) or trunk directly to Unified CM using SIP (as shown in Figure 5-3).

Figure 5-2 Cisco TelePresence ISDN Gateway Trunked to Cisco VCS

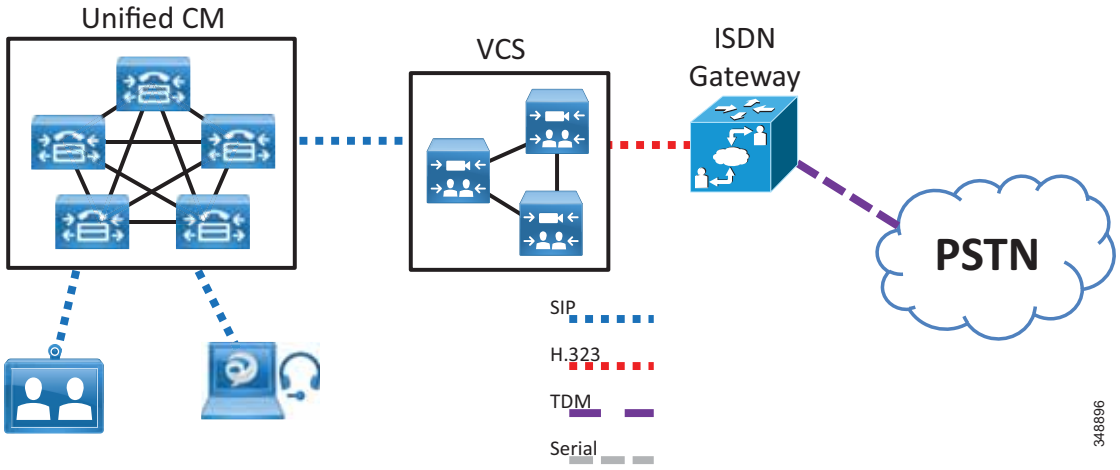
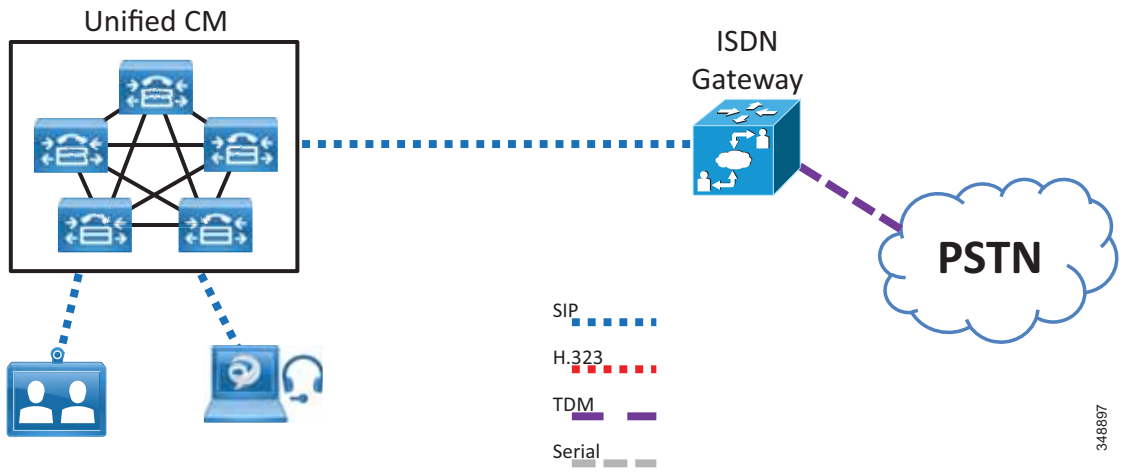


Figure 5-3 Cisco TelePresence ISDN Gateway Registered to Cisco Unified CM



In addition, the Unified CM deployment model being used can influence gateway protocol selection. (Refer to the chapter on Collaboration Deployment Models, page 10-1.)

Core Feature Requirements

Gateways used by voice and video endpoints must meet the following core feature requirements:

- DTMF Relay, page 5-6
- Supplementary Services, page 5-7

  Supplementary services are basic telephony functions such as hold, transfer, and conferencing.
- Unified CM Redundancy, page 5-10

  Cisco Unified Communications is based on a distributed model for high availability. Unified CM clusters provide for Unified CM redundancy. The gateways must support the ability to “re-home” to a secondary Unified CM in the event that a primary Unified CM fails. Some gateways may register to a Cisco VCS, in which case the gateway must support the ability to "re-home" to a secondary Cisco VCS if the primary fails.

Refer to the gateway product documentation to verify that any gateway you select for an enterprise deployment can support the preceding core requirements. Additionally, every collaboration implementation has its own site-specific feature requirements, such as analog or digital access, DID, and capacity requirements.

DTMF Relay

Dual-Tone Multifrequency (DTMF) is a signaling method that uses specific pairs of frequencies within the voice band for signals. A 64 kbps pulse code modulation (PCM) voice channel can carry these signals without difficulty. However, when using a low bite-rate codec for voice compression, the potential exists for DTMF signal loss or distortion. An out-of-band signaling method for carrying DTMF tones across an IP infrastructure provides an elegant solution for these codec-induced symptoms.

SCCP Gateways

The Cisco VG300 Series carries DTMF signals out-of-band using Transmission Control Protocol (TCP) port 2002. Out-of-band DTMF is the default gateway configuration mode for the VG310, VG320, and VG350.

H.323 Gateways

H.323 gateways, such as the Cisco 4000 Series products, can communicate with Unified CM using the enhanced H.245 capability for exchanging DTMF signals out-of-band. This capability is enabled through the command line interface (CLI) of the 4000 Series gateway and the **dtmf-relay** command available in its dial-peers.

MGCP Gateway

Cisco IOS-based platforms can use MGCP for Unified CM communication. Within the MGCP protocol is the concept of *packages*. The MGCP gateway loads the DTMF package upon start-up. The MGCP gateway sends *symbols* over the control channel to represent any DTMF tones it receives. Unified CM then interprets these signals and passes on the DTMF signals, out-of-band, to the signaling endpoint.

The method used for DTMF can be configured using the gateway CLI command:

mgcp dtmf-relay voip codec all mode {*DTMF method*}

Note An MGCP gateway cannot be forced to advertise only in-band DTMF. On enabling in-band DTMF relay, the MGCP gateway will advertise both in-band and out-of-band (OOB) DTMF methods. Unified CM determines which method should be selected and informs the gateway using MGCP signaling. If both the endpoints are MGCP, there is no ability to invoke in-band for DTMF relay because after enabling in-band DTMF, both sides will advertise in-band and OOB DTMF methods to Unified CM. Unified CM will always select OOB if in-band and OOB capabilities are supported by the endpoints.

SIP Gateway

Cisco IOS and ISDN gateways can use SIP for Unified CM communication. They support various methods for DTMF, but only the following methods can be used to communicate with Unified CM:

- Named Telephony Events (NTE), or RFC 2833
- Unsolicited SIP Notify (UN) (Cisco IOS gateways only)
- Key Press Markup Language (KPML)

The method used for DTMF can be configured in Cisco IOS using the gateway CLI command **dtmf-relay** under the respective **dial-peer**. The Cisco ISDN gateways support RFC 2833 and KPML for DTMF.

For more details on DTMF method selection, see the section on Calls over SIP Trunks, page 7-9.

Supplementary Services

Supplementary services provide user functions such as hold, transfer, and conferencing. These are considered basic telephony features and are more common in voice calls than in video calls.

SCCP Gateways

The Cisco SCCP gateways provide full supplementary service support. The SCCP gateways use the Gateway-to-Unified CM signaling channel and SCCP to exchange call control parameters.

H.323 Gateways

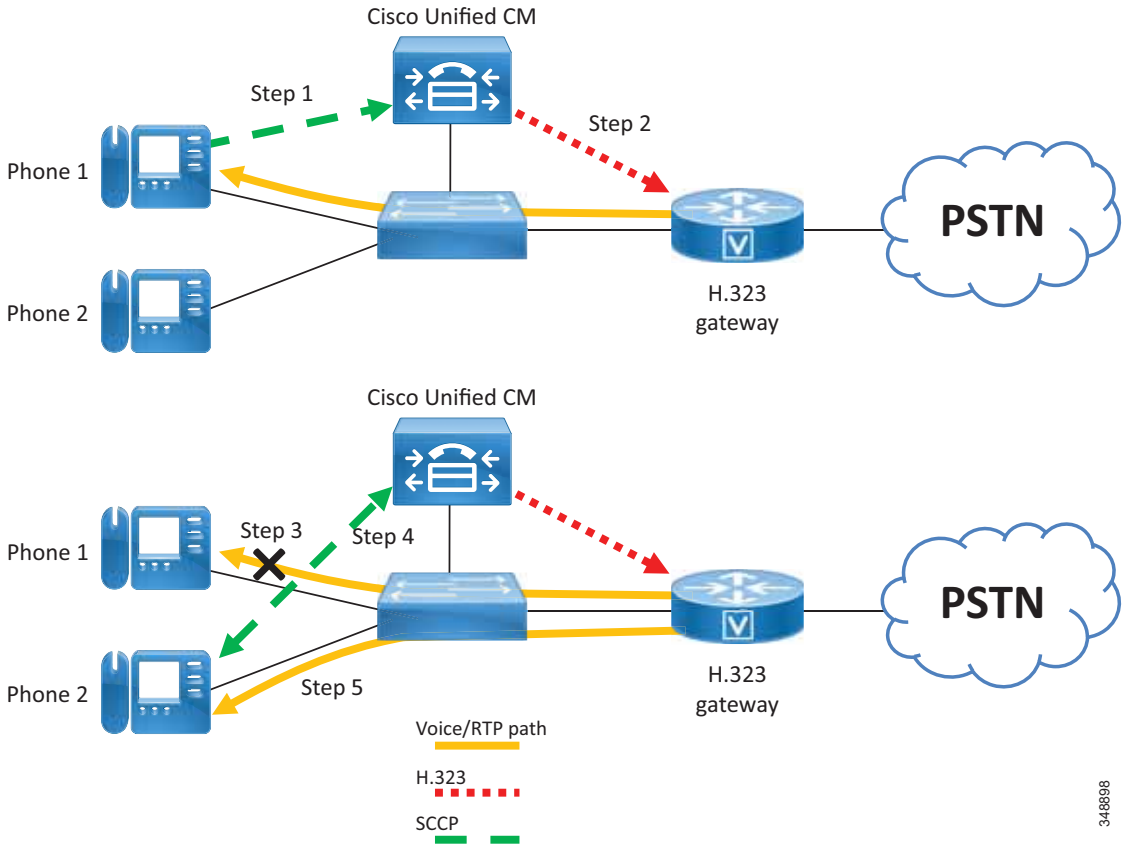
H.323v2 implements Open/Close LogicalChannel and the emptyCapabilitySet features. The use of H.323v2 by H.323 gateways eliminates the requirement for an MTP to provide supplementary services. A transcoder is allocated dynamically only if required during a call to provide access to G.711-only devices while still maintaining a G.729 stream across the WAN.

Once an H.323v2 call is set up between a Cisco IOS gateway and an IP endpoint, using the Unified CM as an H.323 proxy, the endpoint can request to modify the bearer connection. Because the Real-Time Transport Protocol (RTP) stream is directly connected to the endpoint from the Cisco IOS gateway, a supported media codec can be negotiated.

Figure 5-4 and the following steps illustrate a call transfer between two IP phones:

1. If IP Phone 1 wishes to transfer the call from the Cisco IOS gateway to Phone 2, it issues a transfer request to Unified CM using SCCP.
2. Unified CM translates this request into an H.323v2 CloseLogicalChannel request to the Cisco IOS gateway for the appropriate SessionID.
3. The Cisco IOS gateway closes the RTP channel to Phone 1.
4. Unified CM issues a request to Phone 2, using SCCP, to set up an RTP connection to the Cisco IOS gateway. At the same time, Unified CM issues an OpenLogicalChannel request to the Cisco IOS gateway with the new destination parameters, but using the same SessionID.
5. After the Cisco IOS gateway acknowledges the request, an RTP voice bearer channel is established between Phone 2 and the Cisco IOS gateway.

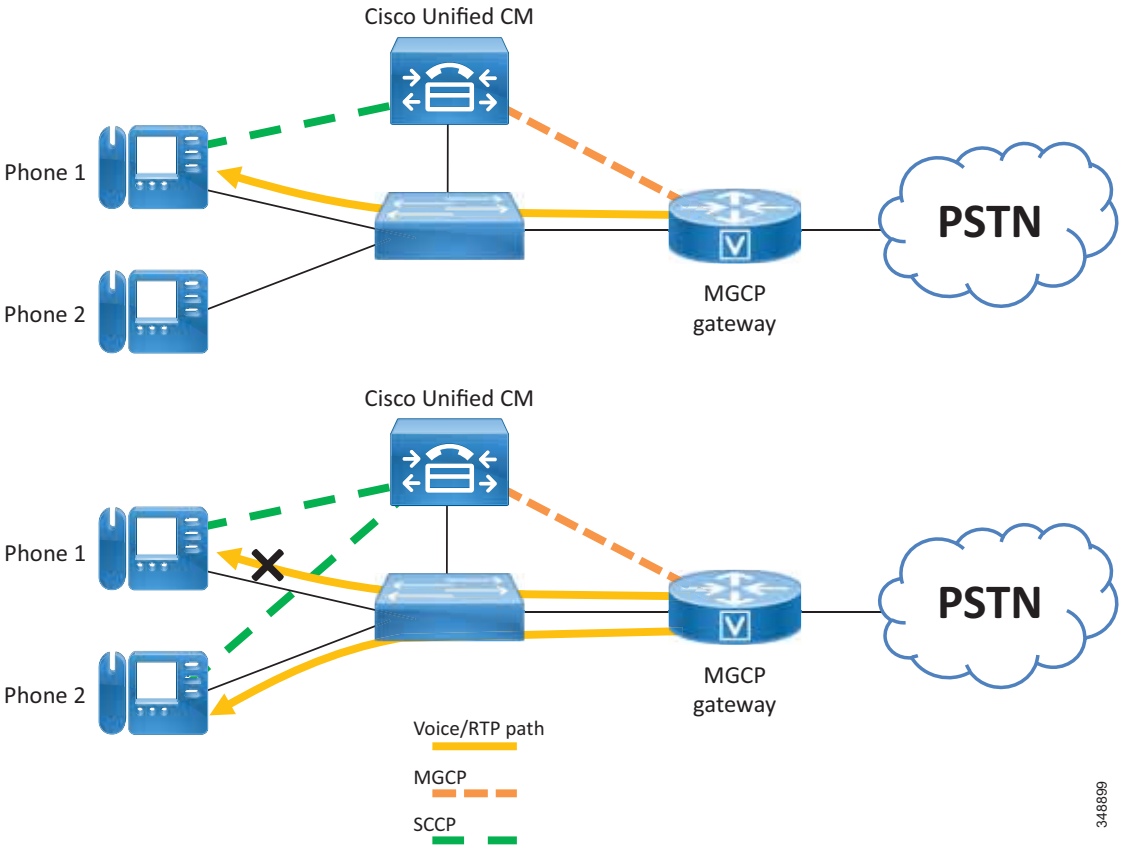
Figure 5-4 H.323 Gateway Supplementary Service Support



MGCP Gateway

The MGCP gateways provide full support for the hold, transfer, and conference features through the MGCP protocol. Because MGCP is a master/slave protocol with Unified CM controlling all session intelligence, Unified CM can easily manipulate MGCP gateway voice connections. If an IP telephony endpoint (for example, an IP phone) needs to modify the session (for example, transfer the call to another endpoint), the endpoint would notify Unified CM using SCCP. Unified CM then informs the MGCP gateway, using the MGCP User Datagram Protocol (UDP) control connection, to terminate the current RTP stream associated with the Session ID and to start a new media session with the new endpoint information. Figure 5-5 illustrates the protocols exchanged between the MGCP gateway, endpoints, and Unified CM.

Figure 5-5 MGCP Gateway Supplementary Service Support



SIP Gateway

The Unified CM SIP trunk interface to Cisco SIP gateways supports supplementary services such as hold, blind transfer, and attended transfer. The support for supplementary services is achieved via SIP methods such as INVITE and REFER. The corresponding SIP gateway must also support these methods in order for supplementary services to work. For more details, refer to the following documentation:

- *Cisco Unified Communications Manager System Guide*

  http://www.cisco.com/en/US/products/sw/voicesw/ps556/prod_maintenance_guides_list.html
- *Cisco IOS SIP Configuration Guide*

  http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/voice/sip/configuration/15-mt/sip-config-15-mt-book.html
- Cisco TelePresence ISDN Gateway documentation

  http://www.cisco.com/en/US/products/ps11448/tsd_products_support_series_home.html

Unified CM Redundancy

An integral piece of the collaboration solution architecture is the provisioning of low-cost, distributed PC-based systems to replace expensive and proprietary legacy PBX systems. This distributed design lends itself to the robust fault tolerant architecture of clustered Unified CMs. Even in its most simplistic form (a two-system cluster), a secondary Unified CM should be able to pick up control of all gateways initially managed by the primary Unified CM.

SCCP Gateways

Upon boot-up, the Cisco VG310, VG320, and VG350 gateways are provisioned with Unified CM server information. When these gateways initialize, a list of Unified CMs is downloaded to the gateways. This list is prioritized into a primary Unified CM and secondary Unified CM. In the event that the primary Unified CM becomes unreachable, the gateway registers with the secondary Unified CM.

H.323 VoIP Call Preservation for WAN Link Failures

H.323 call preservation enhancements for WAN link failures sustain connectivity for H.323 topologies where signaling is handled by an entity that is different from the other endpoint, such as a gatekeeper that provides routed signaling or a call agent (such as Cisco Unified CM) that brokers signaling between the two connected parties. Call preservation is useful when a gateway and the other endpoint are located at the same site but the call agent is remote and therefore more likely to experience connectivity failures.

H.323 call preservation covers the following types of failures and connections.

Failure Types:

- WAN failures that include WAN links flapping or degraded WAN links.
- Cisco Unified CM software failure, such as when the ccm.exe service crashes on a Unified CM server.
- LAN connectivity failure, except when a failure occurs at the local branch.

Connection Types:

- Calls between two Cisco Unified CM controlled endpoints under the following conditions:
  - During Unified CM reloads.
  - When a Transmission Control Protocol (TCP) connection used for signaling H.225.0 or H.245 messages between one or both endpoints and Unified CM is lost or flapping.
  - Between endpoints that are registered to different Unified CMs in a cluster, and the TCP connection between the two Unified CMs is lost.
  - Between IP phones and the PSTN at the same site.
- Calls between a Cisco IOS gateway and an endpoint controlled by a softswitch, where the signaling (H.225.0, H.245 or both) flows between the gateway and the softswitch and media flows between the gateway and the endpoint:
  - When the softswitch reloads.
  - When the H.225.0 or H.245 TCP connection between the gateway and the softswitch is lost, and the softswitch does not clear the call on the endpoint.
  - When the H.225.0 or H.245 TCP connection between softswitch and the endpoint is lost, and the softswitch does not clear the call on the gateway.
- Call flows involving a Cisco Unified Border Element running in media flow-around mode that reload or lose connection with the rest of the network.

Note that, after the media is preserved, the call is torn down later when either one of the parties hangs up or media inactivity is detected. In cases where there is a machine-generated media stream, such as music streaming from a media server, the media inactivity detection will not work and then the call might hang. Cisco Unified CM addresses such conditions by indicating to the gateway that such calls should not be preserved, but third-party devices or the Cisco Unified Border Element would not do this.

Flapping is defined for this feature as the repeated and temporary loss of IP connectivity, which can be caused by WAN or LAN failures. H.323 calls between a Cisco IOS gateway and Cisco Unified CM may be torn down when flapping occurs. When Unified CM detects that the TCP connection is lost, it clears the call and closes the TCP sockets used for the call by sending a TCP FIN, without sending an H.225.0 Release Complete or H.245 End Session message. This is called *quiet clearing*. The TCP FIN sent from Unified CM could reach the gateway if the network comes up for a short duration, and the gateway will tear down the call. Even if the TCP FIN does not reach the gateway, the TCP keepalives sent from the gateway could reach Unified CM when the network comes up. Unified CM will send TCP RST messages in response to the keepalives because it has already closed the TCP connection. The gateway will tear down H.323 calls if it receives the RST message.

Configuration of H.323 call preservation enhancements for WAN link failures involves configuring the **call preserve** command. If you are using Cisco Unified CM, you must enable the Allow Peer to Preserve H.323 Calls parameter from the Service Parameters window.

The **call preserve** command causes the gateway to ignore socket closure or socket errors on H.225.0 or H.245 connections for active calls, thus allowing the socket to be closed without tearing down calls using those connections.

MGCP Gateway

MGCP gateways also have the ability to fail over to a secondary Unified CM in the event of communication loss with the primary Unified CM. When the failover occurs, active calls are preserved.

Within the MGCP gateway configuration file, the primary Unified CM is identified using the **call-agent <hostname>** command, and a list of secondary Unified CM is added using the **ccm-manager redundant-host** command. Keepalives with the primary Unified CM are through the MGCP

application-level keepalive mechanism, whereby the MGCP gateway sends an empty MGCP notify (NTFY) message to Unified CM and waits for an acknowledgement. Keepalive with the backup Unified CMs is through the TCP keepalive mechanism.

If the primary Unified CM becomes available at a later time, the MGCP gateway can "re-home," or switch back to the original Unified CM. This re-homing can occur either immediately, after a configurable amount of time, or only when all connected sessions have been released.

SIP Gateway

Redundancy with Cisco IOS SIP gateways can be achieved similarly to H.323. If the SIP gateway cannot establish a connection to the primary Unified CM, it tries a second Unified CM defined under another dial-peer statement with a higher preference.

By default the Cisco IOS SIP gateway transmits the SIP INVITE request 6 times to the Unified CM IP address configured under the dial-peer. If the SIP gateway does not receive a response from that Unified CM, it will try to contact the Unified CM configured under the other dial-peer with a higher preference.

Cisco IOS SIP gateways wait for the SIP 100 response to an INVITE for a period of 500 ms. By default, it can take up to 3 seconds for the Cisco IOS SIP gateway to reach the backup Unified CM. You can change the SIP INVITE retry attempts under the **sip-ua** configuration by using the command **retry invite** <*number*>. You can also change the period that the Cisco IOS SIP gateway waits for a SIP 100 response to a SIP INVITE request by using the command **timers trying** <*time*> under the **sip-ua** configuration.

One other way to speed up the failover to the backup Unified CM is to configure the command **monitor probe icmp-ping** under the **dial-peer** statement. If Unified CM does not respond to an Internet Control Message Protocol (ICMP) echo message (ping), the dial-peer will be shut down. This command is useful only when the Unified CM is not reachable. ICMP echo messages are sent every 10 seconds.

The Cisco ISDN Gateway can connect to Unified CM via SIP trunk starting with Unified CM release 9.0 and ISDN Gateway release 2.2 and later. The ISDN Gateway SIP configuration consists of entering an IP address, hostname, DNS A record, or DNS SRV record for outbound SIP connections. Redundancy can be achieved by utilizing DNS SRV records with appropriate weight and priority so that, if the primary Unified CM fails, the ISDN Gateway will send outbound SIP calls to the secondary Unified CM.

Gateways for Video Telephony

Video gateways terminate video calls into an IP telephony network or the PSTN. Video gateways are different from voice gateways because they have to interact with the ISDN or serial links that support video and convert that call to a video call on the IP network using protocols such as H.323 or SIP. Enterprises can consider separate gateways for voice calls and video calls, or they can have integrated gateways that route both voice and video calls.

The following key considerations can help you decide if you need separate gateways for voice and video or an integrated gateway:

- Dial plan — If the enterprise has the flexibility of a separate dial plan for video users, it can use separate video gateways that allow it to keep existing enterprise dial plans.
- Video users — If the enterprise has a large number of users who primarily use voice rather than video, then Cisco recommends using separate video gateways to service the video call users.
- Locations — If the enterprise has a large number of distributed locations with video users at many locations, then Cisco recommends using an integrated gateway to reduce total cost of ownership (TCO).

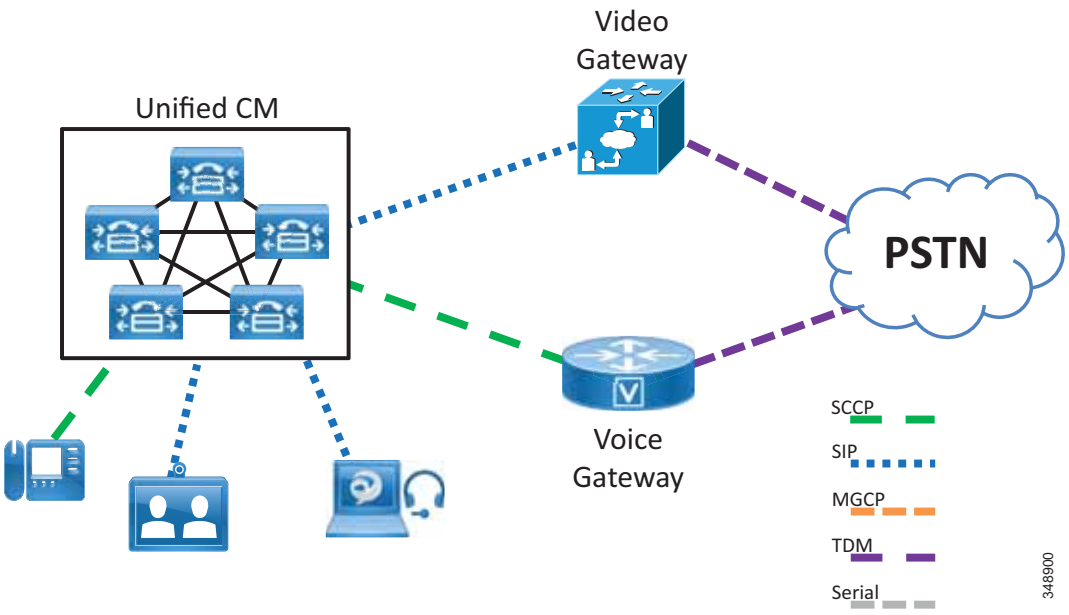
- Additional video capabilities such as video IVR, auto attendant, and bonding across trunks — Dedicated video gateways support advanced features that integrated gateways do not support.
- Protocol — Gateway protocol can be an important factor to align with enterprise policies and standards.
- Device management — Ease of maintenance, management, and troubleshooting can be an important factor. Dedicated gateways provide a better user interface (GUI) for management and configuration, while integrated gateways can provide better troubleshooting. However, these factors are dependent on the respective products.

Dedicated Video Gateways

Enterprises that have an extensive voice infrastructure with voice gateways can add dedicated video gateways so that users can make video calls through them to the PSTN. The Cisco ISDN Gateway and Serial Gateways are examples of dedicated video gateways. Although these products support audio-only calls, they were designed specifically with video users in mind. They support a wide range of video-centric protocols and features.

Figure 5-6 shows an enterprise deployment that can use existing protocols for its voice gateways and add video gateways so that Unified CM users can make voice and video calls to the PSTN.

Figure 5-6 Unified CM System with Separate PSTN Lines for Voice and IP Video Telephony



The Cisco video gateways, while excellent for video calls, do not support all of the telephony features that Cisco voice gateways offer. Cisco video gateways have the following characteristics:

- The Serial Gateway supports only H.323 for IP connectivity.
- The ISDN Gateway supports H.323 and SIP (starting with release 2.2) for IP connectivity.
- They support T1/E1-PRI, BRI, V.35, RS-449, and EIA-530.
- They support H.261, H.263, H.263+, and H.264 video codecs.

- They support G.711, G.722, G.722.1, and G.728; but they do not support G.729 audio.
- They support H.320, H.233, H.234, H.235 (AES), H.239, H.221, FTP, RTP, HTTP, HTTPS, DHCP, SNMP, and NTP.

As a result of these differences in the products, the Cisco TDM and Serial Gateways are not recommended as replacements for Cisco voice gateways. IP Telephony customers who want to add video to their communications environment should deploy both types of gateways and use the Cisco voice gateways for all voice calls and use the Cisco video gateways for video calls only. Customers might also have to procure separate circuits for voice and video from their PSTN service provider, depending on which model of Cisco gateway is deployed.

Also consider how calls will be routed across the IP network to a remote gateway for the purpose of providing toll bypass, and how calls will be re-routed over the PSTN in the event that the IP network is unavailable or does not have enough bandwidth to complete the call. More specifically, do you want to invoke automated alternate routing (AAR) for video calls?

Integrated Video Gateways

Although not recommended, enterprises may consider an integrated device for voice and video gateway functionality. This provides the enterprise the advantages of managing fewer devices and keeping the dial plan simple. The gateway processes the call as a voice call if it is voice and as a video call if it is video.

Cisco IOS, ISDN, and Serial Video gateway have the following characteristics:

- Provide H.323 and SIP support (except Serial Gateway, which is H.323 only)
- Supports H.261, H.263, H.263+, and H.264 video codec
- Provides extensive called and calling transformation capabilities
- Provides extensive logging and troubleshooting capabilities

The following considerations apply for deploying Cisco IOS, ISDN, and Serial Video gateways:

- Consider the capacity needed on PSTN links for additional video calls.
- Consider the need of devices to use content sharing such as Binary Floor Control Protocol (BFCP), and the additional bandwidth that will be used on the IP network.
- Consider if users need features such as far-end camera control or DTMF that is used for conferences that the gateway needs to support.

Configuring Video Gateways in Unified CM

You can configure a Cisco TelePresence ISDN Gateway in either of the following ways:

- Configure a SIP trunk pointing to the ISDN gateway (as shown in Figure 5-3), and add appropriate Unified CM route patterns pointing to the SIP trunk.
- Configure a SIP trunk from Unified CM to Cisco VCS. Have the ISDN gateway (or Serial gateway in this case) register to the VCS using H.323 (as shown in Figure 5-2.

The Cisco TelePresence Serial Gateway cannot be trunked directly to Unified CM. It must register to Cisco VCS, which in turn has a SIP trunk to Unified CM.

Either way, the goal is have all inbound calls received by the gateways sent to Unified CM so that Unified CM can decide how to route the calls. See the chapter on Cisco Unified CM Trunks, page 6-1, for more details on how to configure the SIP trunk between Unified CM and VCS.

Call Signaling Timers

Due to the delay inherent in H.320 bonding, video calls can take longer to complete than voice calls. Several timers in Unified CM are tuned, by default, to make voice calls process as fast as possible, and they can cause video calls to fail. Therefore, you must modify the following timers from their default values in order to support H.320 gateway calls:

- H.245TCSTimeout
- Media Exchange Interface Capability Timer
- Media Exchange Timer

Cisco recommends that you increase each of these timers to 25 by modifying them under the Service Parameters in Unified CM Administration. Note that these are cluster-wide service parameters, so they will affect calls to all types of devices, including voice calls to existing Cisco voice gateways.

Bearer Capabilities of Cisco IOS Voice Gateways

H.323 calls use the H.225/Q.931 Bearer Capabilities Information Element (bearer-caps) to indicate what type of call is being made. A voice-only call has its bearer-caps set to **speech** or **3.1 KHz Audio**, while a video call has its bearer-caps set to **Unrestricted Digital Information**. Some devices do not support Unrestricted Digital Information bearer-caps. Calls to these devices might fail if Unified CM attempts the call as a H.323 video call.

Unified CM decides which bearer-caps to set, based on the following factors:

- Whether the calling and/or called devices are video-capable
- Whether the region in Unified CM is configured to allow video for calls between those devices

Unified CM supports retrying the video call as audio, and this feature can be enabled through configuration. When Unified CM makes a video call with bearer-caps set to **Unrestricted Digital** and the call fails, Unified CM then retries the same call as an audio call with the bearer-caps set to **speech**.

When using H.323, Cisco IOS gateways can service calls as voice or video, based on the bearer capabilities it receives in the call setup. When using SIP, the gateway translates the ISDN capabilities into SDP for call negotiations.

If the Cisco voice gateway uses MGCP to communicate with Unified CM, the problem will not occur because Unified CM does not support video on its MGCP protocol stack and because, in MGCP mode, Unified CM has complete control over the D-Channel signaling to the PSTN.

IP Gateways

The Cisco IP gateways Include:

Cisco Unified Border Element

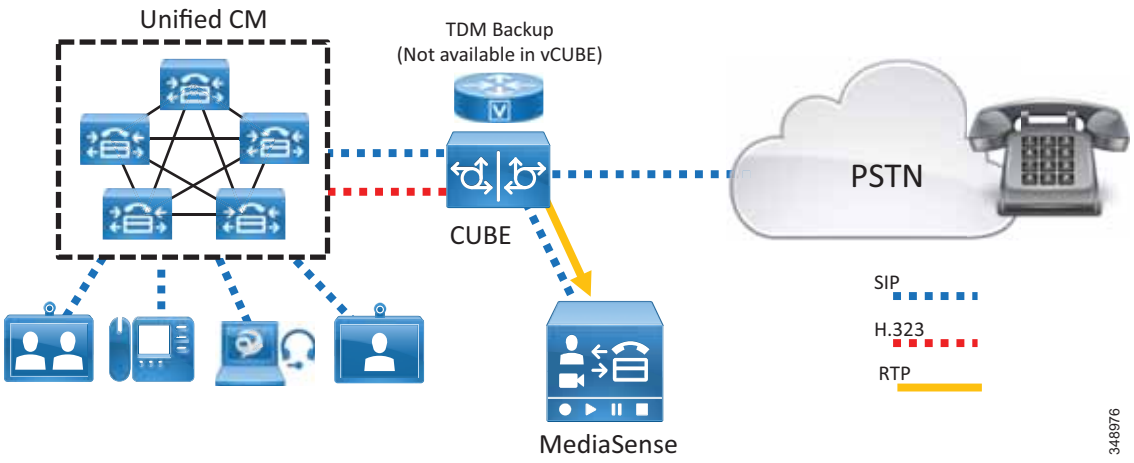
Innovations in collaboration services have delivered significant improvements in employee productivity, and enterprises are widely deploying IP-based Unified Communications, for both internal calling within the enterprise and external PSTN access. This has resulted in significant migration from TDM-based circuits, by both enterprises and telephony service providers, to IP-based trunks for Unified Communications. At the heart of IP-based telephony trunks lies the Session Initiation Protocol (SIP), which is an industry standard communications protocol based on RFC 3261 and is widely used for controlling multimedia communication sessions and applications such as voice, video, unified messaging, voicemail, and conferencing.

These PSTN SIP trunks terminate on a session Border Controller (SBC) at the enterprise, which serves as a demarcation point between the enterprise and the service provider IP networks, similar to how firewalls separate two data networks. The Cisco Unified Border Element (CUBE) Enterprise is Cisco's SBC offering, and it enables rich multimedia communications for enterprises by providing:

- Session Control — Call admission control, trunk routing, QoS, statistics, billing, redundancy, scalability, voice quality monitoring
- Security — Encryption, authentication, registration, SIP protection, voice policy, toll fraud prevention, telephony denial of service (TDoS) attack protection
- Interworking — Various SIP and H323 stack interoperability, sip normalization, dtmf, Transcoding, Transrating, Codec Filtering
- Demarcation — Fault isolation, topology and address hiding, L5/L7 protocol demarcation, network border

CUBE provides essential capabilities that ensure interoperability, security, and service assurance when carrying IP traffic via SIP trunks across various enterprises and service provider networks. It is a Back-to-Back User Agent (B2BUA) and is part of the Cisco IOS infrastructure on Cisco ISR G2 800 Series platforms, Cisco IOS-XE for the ASR 1000 Series, Cisco ISR 4000 Series, and CUBE on the Cisco Cloud Services Router (CSR) 1000V Series (virtual CUBE, or vCUBE). Figure 5-7 illustrates the enterprise CUBE deployment.

Figure 5-7 Cisco Unified Border Element Deployment



For more information about Cisco Unified Border Element, refer to the documentation at

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

Cisco Expressway

Use of the Internet for collaboration services continues to increase in popularity and is quickly replacing existing legacy ISDN video systems and gateways. The two primary protocols leveraged for Internet based collaboration services are SIP and H.323. The Internet is also used to connect remote and mobile users to voice, video, IM and presence, and content sharing services without the use of a virtual private network (VPN).

The Expressway-C and Expressway-E pair performs the following functions:

- Mobile and remote access, as well as business-to-business services, can be enabled as part of the same Cisco Expressway-C and Expressway-E solution pair.
- Interworking — The capability to interconnect H.323-to-SIP calls for voice, video, and content sharing.
- Boundary communication services — While Expressway-C sits in the corporate network, Expressway-E is in the enterprise DMZ and provides a distinct connection point for communication services between the enterprise network and the Internet.
- Security — The capability to provide authentication and encryption for both mobile and remote access and business-to-business communications.

Expressway-C and Expressway-E are designed to work together to form a firewall traversal solution that is the core component for business-to-business communications over the Internet. Expressway-C sits on the inside (trusted side) of the enterprise network and serves the role of providing a secure, trusted, and standards-based way of connecting to Expressway-E. It acts as a traversal client to all devices behind it. This solves the problem for devices using a large number of media ports by multiplexing all media to a very small number of ports opened for outbound communications. It provides an authenticated and trusted connection from inside the enterprise to outside by sending a keep-alive for the traversal zone from Expressway-C to Expressway-E. Additionally, it provides a single point of contact for all Internet communications, thus minimizing the security risk.

Real-time and near real-time communication protocols such as SIP, H.323, and XMPP do not address the need to communicate with devices that might be behind a firewall. Typical communications using these protocols include the device IP address in the signaling and media, which becomes the payload of the TCP and UDP packets, respectively. When these devices are on the same internally routable network, they can successfully communicate directly with each other. The signaling IP address carried in the payload of the TCP packet is routable back to the initiating device, and vice versa. However, when the initiating device is on a different network behind a public or network edge firewall, two problems are encountered. The first problem is that the receiving device, after decoding the packet, will respond to the internal IP address carried in the payload. This IP address is typically a non-routable RFC 1918 address and will never reach the return destination. The second problem is that, even if the return IP address is routable, the media (which is RTP/UDP) is blocked by the external firewall. This applies to both business-to-business and mobile and remote access communications.

Expressway-E sits at the network edge in the DMZ. It serves the role of solving both the signaling and media routing problems for SIP, H323, and XMPP, while maintaining standards interoperability. Expressway-E changes the appropriate headers and IP addresses to process the media and signaling on behalf of the endpoints, devices, and application servers that are inside the network.

Expressway-C and Expressway-E Deployment for Business-to-Business Communications

The standard deployment of the Cisco Expressway Series involves deploying at least one Expressway-C and Expressway-E pair for business-to-business communications. Both Expressway-C and Expressway-E should be deployed in a cluster to provide better resiliency. The number of servers for each cluster depends on the number of concurrent calls. (For details, see the chapter on Collaboration Solution Sizing Guidance, page 25-1.)

Frequently, multiple pairs of Expressway-C and Expressway-E are deployed for geographic coverage and scale, providing access to multiple instances of collaboration services. Unified CM is connected to Expressway-C through a SIP trunk for unified business communications access over the Internet. Based on the enterprise security policy, a number of different deployment models can be implemented. In this document we focus on a DMZ deployment of Expressway-E with dual network interfaces because it is the most common and secure deployment model. For additional deployment models, refer to latest version of the *Cisco Expressway Basic Configuration Deployment Guide*, available at

http://www.cisco.com/c/en/us/support/unified-communications/expressway-series/products-installation-and-configuration-guides-list.html

Expressway-C and Expressway-E provide firewall traversal capabilities. Firewall traversal works as follows:

1. Expressway-E is the traversal server installed within the enterprise DMZ, and Expressway-C is the traversal client installed inside the enterprise network.
2. Expressway-C initiates traversal connections outbound through the firewall to specific ports on Expressway-E, with secure login credentials. If the firewall allows outbound connections, as it does in the vast majority of cases, no additional ports are required to be opened in the enterprise firewall.

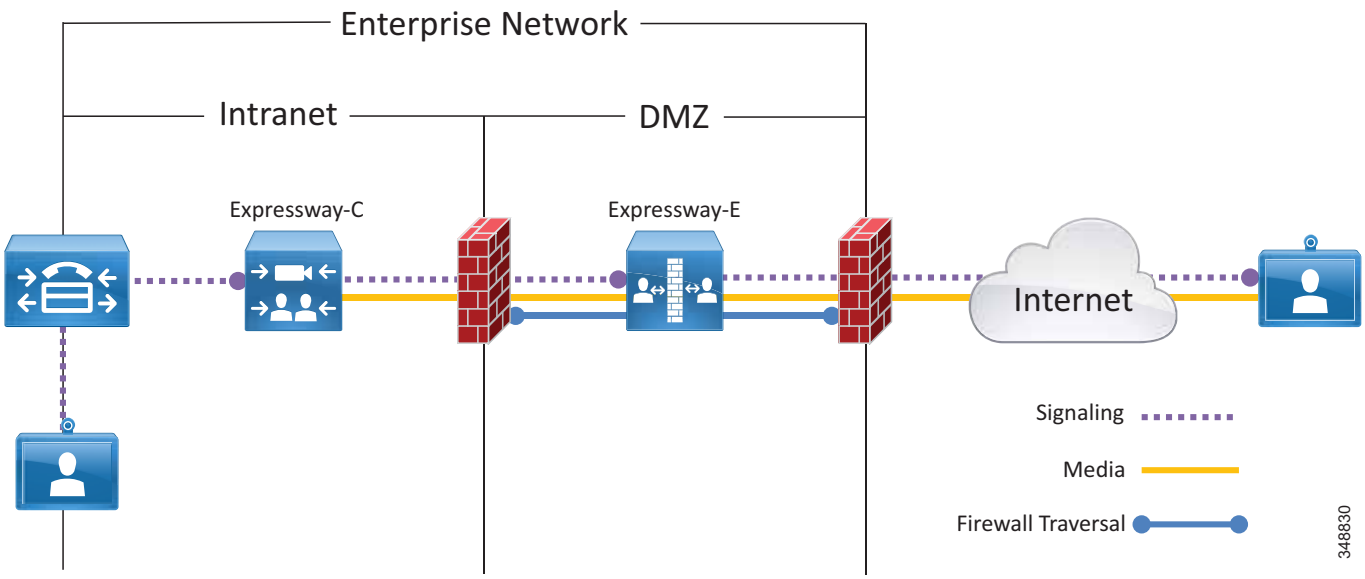
   For port details, refer to the latest version of the *Unified Communications Mobile and Remote Access via Cisco Expressway Deployment Guide*, which includes all ports used by Expressway in business-to-business and mobile and remote access scenarios. This guide is available at

   http://www.cisco.com/c/en/us/support/unified-communications/expressway-series/products-installation-and-configuration-guides-list.html
3. Once the connection has been established, Expressway-C sends periodic keep-alive packets to Expressway-E to maintain the connection.

4. When Expressway-E receives an incoming call or other collaboration service request, it issues an incoming request to Expressway-C.
5. Expressway-C then routes the request to Unified CM or other collaboration service applications.
6. The connection is established, and application traffic (including voice and video media) traverses the firewall securely over an existing traversal connection.

For firewall traversal to work, a traversal client zone has to be configured on Expressway-C and a traversal server zone has to be configured on Expressway-E. Figure 5-8 summarizes the firewall traversal process in a dual-interface deployment scenario for Expressway-E.

Figure 5-8 Firewall Traversal in a Dual-Interface Deployment



In the dual-interface deployment scenario, Expressway-E sits in the DMZ between two firewalls: the Internet firewall provides for NAT services toward the Internet, and the intranet firewall provides access to the corporate trusted network.

Expressway-E has two LAN interfaces: one toward the Internet firewall (also called the external interface) and the other toward the intranet firewall (also called the internal interface). In order to route packets to the external or internal interface, you create static routes on Expressway-E. The easiest way to create the static routes is by setting the Expressway-E default gateway equal to the default gateway for the external LAN interface, and by creating static routes for every internal network. In this way, internal traffic will be sent to the internal interface, and all traffic not matching the network range configured in the static routes will be sent to the Internet.

There is no need for the external interface to be assigned a public IP address because the address can be translated statically by NAT. In this case, the public IP address has to be configured on Expressway-E itself. The Expressway-E external interface can be statically translated by NAT, but the Expressway-E internal interface can be statically translated by NAT only if the Expressway is not clustered. The Expressway-C interface can be translated by NAT.

A connection from the Internet for business-to-business communications between Expressway-C and back-end application services may or may not be encrypted, based on the configuration and dictated by the corporate policies. Note that in this case the communication will be encrypted end-to-end only if both

the corporate and the remote business-to-business party supports encryption with public certificates. In all other cases, the video call will be sent unencrypted, or it will be dropped based on Expressway-E configuration policies.

Business-to-Business Call Flow

Business-to-business communications require the ability to look up the domains of remote organizations for the purpose of URI routing. This is done by creating a DNS zone on Expressway-E. This zone should be configured with the default settings. Both SIP and H.323 are set by default. Expressway-C and Expressway-E use the protocol that was used to initiate the call, and they automatically try the other protocol when SIP-to-H.323 gateway interworking is enabled on Expressway.

SIP-to-H.323 interworking should be set to **On** for Expressway-E. If a call is receive as an H.323 call, this allows Expressway-E to interwork the call to SIP and use native SIP for the rest of the call legs to Unified CM. Likewise, an outbound call to an H.323 system will remain a SIP call until it reaches Expressway-E, where it will be interworked to H.323.

In order to receive business-to-business communications over the Internet, External SIP and H.323 DNS records are required. These records allow other organizations to resolve the domain of the URI to the Expressway-E that is offering that call service. Cisco's validated design includes the SIP and SIPS SRV records and the H.323 SRV record for business-to-business communications. The H.323 SRV record is not necessary for Expressway-E because this record is used by an endpoint to find its gatekeeper for registration.

Table 5-2 shows the DNS SRV records used for resolving the domain of the URI.

Table 5-2 DNS SRV Records for Resolving the URI Domain

Type of Communication	Domain	Port	Protocol
SIP business-to-business	_sips._tcp.domain	5061	TLS
	_sip._tcp.domain	5060	TCP
	_sip._udp.domain	5060	UDP
H.323 business-to-business	_h323ls._udp.domain	1719	RAS
	_h323cs._tcp.domain	1720	H.225
Mobile and remote access	_collab-edge._tls.domain	8443	Jabber login
	_xmpp-server._tcp.domain	5269	XMPP Federation

For more information about configuring a DNS zone on Expressway-E, refer to latest version of the *Cisco Expressway Basic Configuration Deployment Guide*, available at

http://www.cisco.com/c/en/us/support/unified-communications/expressway-series/products-installation-and-configuration-guides-list.html

Outbound calls use a SIP Route Pattern on Cisco Unified CM set to "*". Any SIP URI that does not find a match inside the local Unified CM cluster or ILS table will be sent through this SIP Route Pattern, according to the routing rules logic defined in the chapter on Dial Plan, page 14-1. Configure this SIP Route Pattern to have a Route List to the Expressway-C cluster as a target.

Configure Expressway-C to have two rules for business-to-business communications:

- Send any SIP URI with the domain portion matching the domain of the company to Cisco Unified Communications Manager.
- Send any SIP URI with the domain portion matching any other domain to Expressway-E.

On Expressway-E configure two rules for business-to-business communications:

- Send any SIP URI with the domain portion matching the domain of the company to the Expressway-C cluster.
- Send any SIP URI with the domain portion matching any other domain to the DNS Zone that is used for DNS SRV resolution.

When a user dials a string followed by an external domain from an endpoint connected to Unified CM, the SIP Route Pattern will be matched. Unified CM will send the call to Expressway-C, and Expressway-C to will send it Expressway-E. Expressway-E will perform a DNS SRV lookup on a public DNS. The DNS will resolve the SRV record, and Expressway-E will be able to direct the call to the unknown remote edge.

Inbound calls will be received by the Expressway-E on the Default Zone, and based on the search rules specified above, Expressway-E will send the call to the Expressway-C, which will send it to Cisco Unified CM.

Note that any Cisco endpoint connected to Cisco Unified CM, regardless from model type or voice/video capabilities, will be reachable.

If the endpoint does not have any associated SIP URI, it will be reachable through the string <*DN*>@<*domain*>, where <*DN*> is the directory number configured on Cisco Unified CM and <*domain*> is the company SIP domain.

In case the device has a corresponding alphanumeric SIP URI associated with its DN, the same device can also be reached by dialing the alphanumeric SIP URI.

IP-Based Dialing for Business-to-Business Calls

IP-based dialing is a feature well known and used in most scenarios, especially when dealing with H.323 endpoints. The Cisco Collaboration Architecture uses SIP URIs and does not need IP-based dialing. However, when interacting with endpoints in other organizations that are capable of making and receiving calls using IP addresses only, the Cisco Collaboration Architecture allows IP-based dialing for both inbound and outbound calls.

Outbound Calls

Outbound IP dialing is supported on Expressway-E and Expressway-C, but it does not have full native support on Cisco Unified CM. However, it is possible to set up Unified CM to have IP-based dialing, as described here.

Instead of dialing the IP address alone, users on Cisco Unified CM can dial a SIP URI-based IP address as shown in this example: 10.10.10.10@ip, where "@ip" is literal and could be replaced with "external", "offsite" or other meaningful terms.

Unified CM will match a SIP route pattern configured to route the "ip" fictional domain to Expressway-C. Expressway-C strips off the domain "@ip" and sends the call to Expressway-E, which is also configured for IP address dialing.

Calls to unknown IP addresses on Expressway -E should be set to **Direct**. Since IP-based address dialing is mostly configured in H.323 endpoints when no call control is deployed, this allows Expressway-E to send H.323 calls directly to an endpoint at a public IP address. The call will remain a SIP call until interworked on Expressway-E.

Alternatively, instead of having to append the fictional domain, users might replace the dots with a star character, as in this example: 10*10*10*10.

Unified CM will match a Route Pattern defined as !*!*!*! and send the call to Expressway-C, which will replace the "star" character with a dot. In this case, the search rule will match the following regex expression: (\d\d?\d?)(*)(\d\d?\d?)(*)(\d\d?\d?)(*)(\d\d?\d?), and will have \1\.\3\.\5\.\7 as the replacement string.

Inbound Calls

IP-based inbound calls make use of a fallback alias configured in Expressway-E. When a user on the Internet dials the IP address of the Expressway-E external LAN interface, Expressway-E receives the call and sends the call to the alias configured in the fallback alias setting. As an example, if the fallback alias is configured to send the call to conference number 80044123 or to the conference alias meet@example.com, the inbound call will be sent to the TelePresence Server in charge of such conferences.

If the static mapping between the IP address and the fallback alias is too limited, it is possible to set the fallback alias to the pilot number of Cisco Unity Connection. In this way it is possible to use the Unity Connection auto-attendant feature to specify the final destination through DTMF, or by speech recognition if Unity Connection is enabled to support this feature. If Unity Connection is used as an auto-attendant feature for external endpoints dialing the IP address of the Expressway-E, remember to set the Rerouting Calling Search Space on the Unified CM trunk configuration for Unity Connection.

High Availability for Expressway-C and Expressway-E

We recommend deploying Expressway-C and Expressway-E in clusters. Each cluster can have up to six Expressway nodes and a maximum of N+2 physical redundancy. All nodes are active in the cluster. For details about cluster configuration, refer to the latest version of the *Cisco Expressway Cluster Creation and Maintenance Deployment Guide*, available at

http://www.cisco.com/c/en/us/support/unified-communications/expressway-series/products-installation-and-configuration-guides-list.html

Expressway clusters provide configuration redundancy. The first node configured in the cluster is the master. Configuration is done in the master and automatically replicated to the other nodes. Expressway clusters provide call license sharing and resilience. All rich-media session licenses are shared equally across nodes in the cluster. Call licenses are contributed by the licenses configured on each node.

Expressway-C and Expressway-E deployed as virtual machines support VMware VMotion. VMware. VMotion enables the live migration of running virtual machines from one physical server to another. When moving a virtual machine, Expressway-C and Expressway-E servers will maintain active calls when handling signaling only or when handling both signaling and media. This provides high availability for the Expressway nodes as well as call resilience across Cisco Unified Computing System (UCS) hosts.

The following rules apply to Expressway clustering:

- Expressway-C and Expressway-E node types cannot be mixed in the same cluster.
- All nodes in a cluster must have identical configurations.
- Configuration changes should be made only on the master node, and this will overwrite the configuration on the other nodes in the cluster when replication occurs.
- If a node becomes unavailable, the licenses it contributed to the cluster will become unavailable after 2 weeks.
- Deploy an equal number of nodes in Expressway-C and Expressway-E clusters.
- Deploy the same OVA template throughout the cluster.

- All nodes in a cluster need to be within 30 ms maximum round-trip time to all other cluster nodes. Clustering over the WAN is therefore not recommended due to latency constraints.
- You must use the same cluster preshared key for all nodes within the same cluster.
- If mobile and remote access and business-to-business communications are enabled on the same Expressway-C and Expressway-E pairs, the SIP port number used on the SIP trunk between Unified CM and Expressway-C must be changed from the default 5060 or 5061 (for example, use 5560 and 5561).
- A DNS SRV record must be available for the cluster and must contain A or AAAA records for each node of the cluster.

Since Expressway-C is deployed in the internal network and Expressway-E is in the DMZ, Expressway-C has to be connected to Expressway-E through a traversal zone for business-to-business calls. Mobile and remote access requires a separate traversal zone, referred to as **Unified Communication traversal zone**. The traversal server and traversal client zones include all the nodes of Expressway-C and Expressway-E, so that if one of the nodes is not reachable, another node of the cluster will be reached instead.

The traversal client zone configured on Expressway-C should contain the fully qualified domain names of all of the cluster nodes of the corresponding Expressway-E cluster. Likewise, the traversal server zone should connect to all Expressway-C cluster nodes. This is achieved by including, in the subject alternative names of the Expressway-C certificate, the FQDNs of the Expressway-C cluster nodes and by setting the TLS verify subject name equal to the FQDN of the Expressway-C cluster. This creates a mesh configuration of cluster nodes across the traversal zone and provides continuous and high availability of the traversal zone until the last cluster node is unavailable.

Expressway-C connects to Unified CM via a neighbor zone for routing inbound and outbound business-to-business calls. Unified CM also trunks to Expressway-C. For high availability, the fully qualified domain names of each Expressway-C cluster node should be listed in the trunk configuration on Unified CM. If Unified CM is clustered, the fully qualified domain name (FQDN) of each member of the cluster should be listed in the neighbor zone profile of Expressway-C.

A meshed trunk configuration is created here as well. Unified CM will check the status of the nodes in the trunk configuration via a SIP OPTIONS Ping. If a node is not available, Unified CM will take that node out of service and will not route calls to it. Expressway-C will also check the status of the trunk from Unified CM via a SIP OPTIONS Ping. Calls will be routed only to nodes that are shown as active and available. This provides high availability for both sides of the trunk configuration.

DNS SRV records can add to availability of Expressway-E for inbound business-to-business traffic. For high availability, all nodes in the cluster should be listed with the same priority and weight in the SRV record. This allows all nodes to be returned in the DNS query. A DNS SRV record helps to minimize the time spent by a client on lookups because a DNS response can contain all of the nodes listed in the SRV record. The far-end server or far-end endpoint will typically cache the DNS response and will try all nodes returned in the DNS query until a response is received. This provides the best chance for a successful call.

In addition, Expressway clusters support rich media license sharing across clusters. If a node is lost from the cluster, its call licenses will continue to be shared for the next 2 weeks.

Any one Expressway cannot process any more rich media licenses than its physical capacity, even though it can carry more licenses than its physical capacity.

Security for Expressway-C and Expressway-E

Security on Expressway-C and Expressway-E can be further partitioned into network level and application level. Network level security includes feature such as firewall rules and intrusion protection, while application level security includes authorization, authentication, and encryption.

Network Level Protection

Network level protection on Expressway-C and Expressway-E consists of two main components: firewall rules and intrusion protection. Firewall rules enable the ability to:

- Specify the source IP address subnet from which to allow or deny traffic.
- Choose whether to drop or reject denied traffic.
- Configure well known services such as SSH and HTTP/HTTPS, or specify customized rules based on transport protocols and port ranges.
- Configure different rules for the LAN 1 and LAN 2 interfaces on Expressway-E.

The Automated Intrusion Protection feature should be used to detect and block malicious traffic and to help protect the Expressway from dictionary-based attempts to breach login security. Automated Intrusion Protection works by parsing the system log files to detect repeated failures to access specific service categories such as SIP, SSH, and web/HTTPS. When the number of failures within a specified time reaches the configured threshold, the source host IP address (the intruder) and destination port are blocked for a specified period of time. The host address is automatically unblocked after that time period so as not to lock out any genuine hosts that might have been temporarily mis-configured.

Application Level Security

Application level security can be partitioned into:

- Authentication and Encryption, page 5-24
- Dial Plan Protection and Toll Fraud Mitigation, page 5-25

Authentication and Encryption

Securing business-to-business communications includes authentication, encryption, and authorization. Business-to-business communications use an authenticated traversal link by default. The traversal link can also benefit from the use of a Public Key Infrastructure (PKI) verified by a mutually authenticated transport layer security (MTLS) connection between Expressway-C and Expressway-E. If the business-to-business traversal link is deployed on the same Expressway-C and Expressway-E infrastructure as mobile and remote access, make sure that the traversal zone uses the FQDNs of the cluster nodes of Expressway-C and Expressway-E. This makes it straightforward to use certificates for each server to validate the offered certificate against its certificate trust for the traversal connection.

Signaling and media encryption is important for business-to-business calls, but it needs to be deployed carefully so as not to restrict or limit the ability to receive calls. There is a variety of older SIP and H.323 systems that you may be communicating with that do not support signaling or media encryption.

Based on zone configuration, encryption policies might be set as forced (**force encrypted**), desirable (**best effort**), not allowed (**force unencrypted**), or left to the endpoint decisions (**auto**).

If **force encrypted** is configured on a target zone and the Expressway is receiving a call for an endpoint on that remote zone, then Expressway will set up an encrypted call. If the remote party accepts only unencrypted calls, the call will be dropped. If the calling endpoint is using TCP and sending unencrypted media, and **force encrypted** is configured on the target zone, Expressway will terminate the call leg and set up another call leg to the destination with TLS and encryption.

When Expressway performs RTP to SRTP, it uses a back-to-back user agent (B2BUA) for business-to-business calls. The B2BUA terminates both signaling and media and sets up a new call leg to the destination. The B2BUA is engaged any time the media encryption mode is configured to a setting other than **auto**. Exception occurs only in the following scenario affecting Expressway-E: if the inbound zone and outbound zone are set to the same encryption media type and one of those zones is a Traversal Server zone, Expressway-E checks the value of the associated Traversal Client zone. If all three of these zones are set to the same value, the Expressway-E will not engage the B2BUA. In this case, B2BUA will be engaged only on Expressway-C. With **best effort**, if Expressway cannot set up an encrypted call, it will fall back to unencrypted.

Depending on the requirements, different media encryption policies might be configured. If a corporate enforcing policy is not in place, the recommendation is to set up zones with **auto** specified as the media encrypted mode. A setting of **auto** delegates the encryption decisions to endpoints, and Expressway does not perform any sort of RTP-to-SRTP conversion.

When the encryption policy is enforced on Expressway, the call will be divided into many call legs due to B2BUA engagement, as in the following scenario:

- Expressway-C neighbor zone to Unified CM set to **best effort**
- Expressway-E traversal server zone set to **best effort**
- Expressway-E DNS zone set to **auto**
- Calling endpoint on Unified CM configured for encryption, and Unified CM configured in mixed mode
- Called endpoint or system does not support encryption

If a Unified CM endpoint calls an unencrypted endpoint on the Internet, the call will consist of the following call legs:

1. Unified CM endpoint to Expressway-C B2BUA, encrypted
2. Expressway-C B2BUA to Expressway-E B2BUA, encrypted
3. Expressway-E B2BUA to the Internet, up to unknown remote edge or final destination, unencrypted
4. Remote edge to final destination, encrypted or unencrypted depending on called partner's settings

If call legs 1 through 3 are encrypted, the lock icon will display correctly. If one of these legs is not encrypted, the lock icon will not display. Note that the last call leg is under the control of another company, and as such does not influence the lock status.

Every company has the control of encryption up to the other company's edge, thus allowing an endpoint to establish an encrypted call from the endpoint to the remote edge. Encryption policy can protect media on the Internet if **force encrypted** is configured on Expressway; but once the call hits the remote edge, the call might be decrypted at the edge level before sending it to the called endpoint.

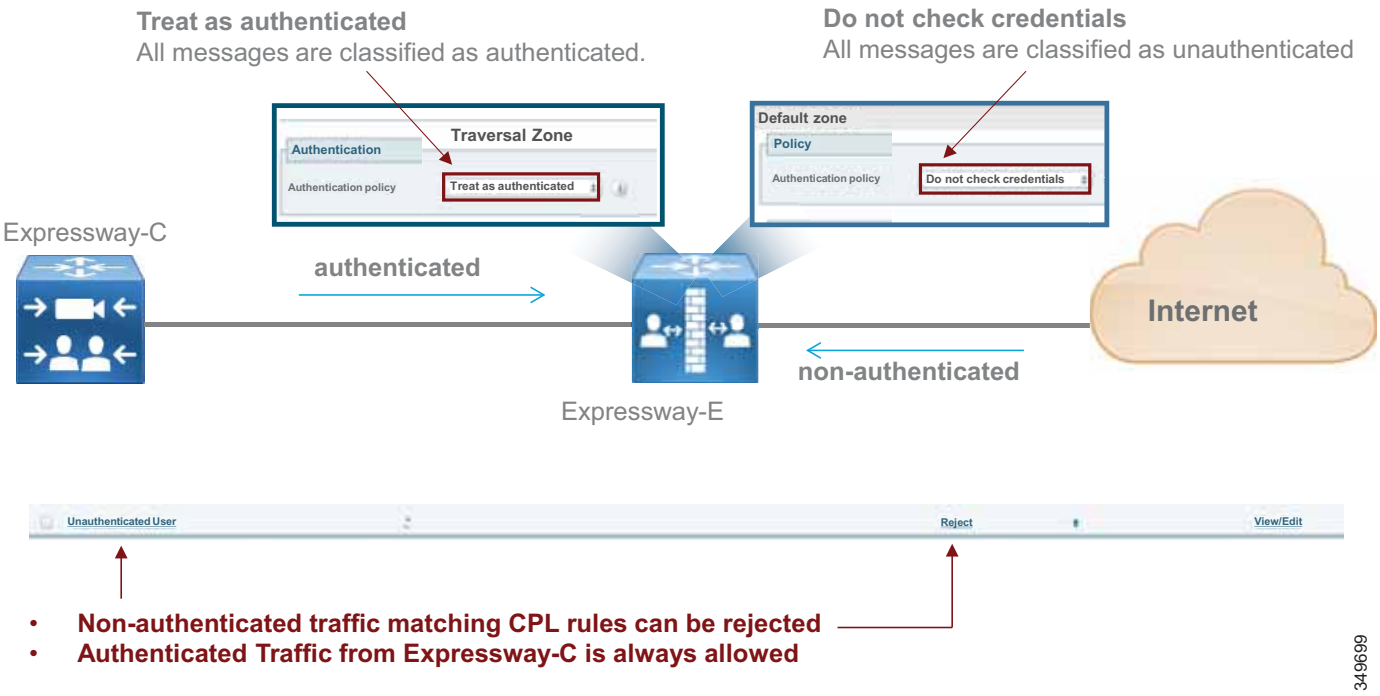
The only way to provide for true end-to-end encryption is by controlling the remote endpoint. This could be done by deploying Jabber Guest or giving the customer the ability to download Jabber and login through Mobile and Remote Access in the corporate Cisco Unified CM and IM and Presence servers.

Dial Plan Protection and Toll Fraud Mitigation

In order to block legal call attempts from unwanted users on the Internet, spam calls, and SIP or H.323 scans, Call Processing Language (CPL) rules can be used on Expressway-E. CPL rules can be applied to call attempts coming from the Internet only.

In order to do this, traffic coming from the traversal client zone can be set to **authenticated**, and traffic from the Internet can be set to **non-authenticated**. CPL rules can be applied to non-authenticated traffic only, bypassing checks for traffic from the internal network or from trusted neighbors on the Internet. Figure 5-9 illustrates this.

Figure 5-9 Zone Authentication Policy



CPL rules are processed using a top-down approach. Two sets of policies can be created:

- Allow-based policy

  An allow-based policy applies regular expressions (regex) to CPL in order allow calls only if they match the numeric range or the alphanumeric URI format internally configured. The last CPL rule will block all calls.

- Deny-based policy

  A deny-based policy denies calls to specific services such as gateways and voicemail, while allowing all the rest if the domain matches the corporate domain. A default CPL rule that blocks all calls is set as last rule.

As an example of a deny-based policy, consider a company where calls are allowed to a set of devices in the range 80XXXXXX only, and where gateway access and other services from external Internet destinations, here represented with 0 and 9, are forbidden. In this case the rules can be set as shown in Table 5-3.

Table 5-3 Example of Deny-Based Policy

Source Type	Destination	Action
Default Zone	**8[1-9]\d{6}@example\.com**	Reject
Default Zone	**[09]\d*@example\.com**	Reject
Default Zone	**\+\d*@example\.com**	Reject
Default Zone	**.*@example\.com**	Allow
Default Zone	**.***	Reject

In addition, it is possible to reject calls based on the calling ID. Unlike the PSTN, where Telecom providers preserve the calling numbers, the Internet is free and nobody is checking the identity of a user. Therefore, it is possible to reject incoming business-to-business calls if the calling alias contains the corporate domain or the IP address of the Expressway-E.

The example in Table 5-4 is based on Cisco Expressway release 8.9.

Table 5-4 Example of Deny-Based Policy Using Expressway-E IP Address

Source Type	Source Alias	Destination Alias	Action
Unauthenticated	**.*@10\.10\.10\.10.***	**.***	Reject
Unauthenticated	**.*@example\.com.***	**.***	Reject

10.10.10.10 in Table 5-4 represents the public address of Expressway-E. These rules can be added to the previous list just before the "allow" rule. In this way any call from the Internet containing the corporate domain or IP address will be rejected, thus mitigating identity theft.

Because the Default Zone is the target for business-to-business incoming calls, it has to be configured with an authentication policy set to "do not check credentials." In this way business-to-business calls will be considered unauthenticated and thus checked against the rules. Internal traffic coming from the Traversal Zone will bypass this check if that zone is configured with an authentication policy set to "treat as authenticated", as shown in Figure 5-9.

Scaling the Expressway Solution

When multiple Internet edges are deployed, it is important to set routing rules properly in order to send collaboration traffic to the nearest Internet edge.

Multiple Expressway-Es and GeoDNS

Scalability for business-to-business communications can be addressed by adding multiple Expressway-C and Expressway-E clusters, either in the same physical location or geographically dispersed. When multiple Expressway-C and Expressway-E pairs are deployed, Unified CM can direct an outbound call to the edge server that is nearest to the calling endpoint, thus minimizing internal WAN traffic. For large deployments it might be preferable to host business-to-business communications on Expressway-C and Expressway-E pairs separate from mobile and remote access. This allows the server resources to be dedicated to external Internet communications.

When two or more Internet edges are deployed, it is important to understand how to split the load between them. If the Internet edges are deployed in the same data center or in the same area, load balancing can occur at the DNS SRV level. As an example, if the enterprise network includes three Internet edges used for business-to-business communications, each one consisting of a cluster of two Expressway-E and Expressway-C nodes, the _sips._tcp.example.com and _sip._tcp.example.com records will include all six Expressway-E records at the same priority and weight. This distributes the registrations and calls equally across the various Expressway-E and Expressway-C clusters.

However, if the Expressway clusters are deployed across geographical regions, some intelligent mechanisms on top of the DNS SRV priority and weight record are needed to ensure that the endpoint uses the nearest Expressway-E cluster. As an example, if an enterprise has two Expressway clusters, one in the United States (US) and the other in Europe (EMEA), it is desirable for users located in the US to be directed to the Expressway-E cluster in the US while users in Europe are directed to the Expressway-E cluster in Europe. This is facilitated by implementing GeoDNS services. GeoDNS services are cost effective and easy to configure. With GeoDNS it is possible to route traffic based on different policies such as location (IP address routing), minimum latency, and others.

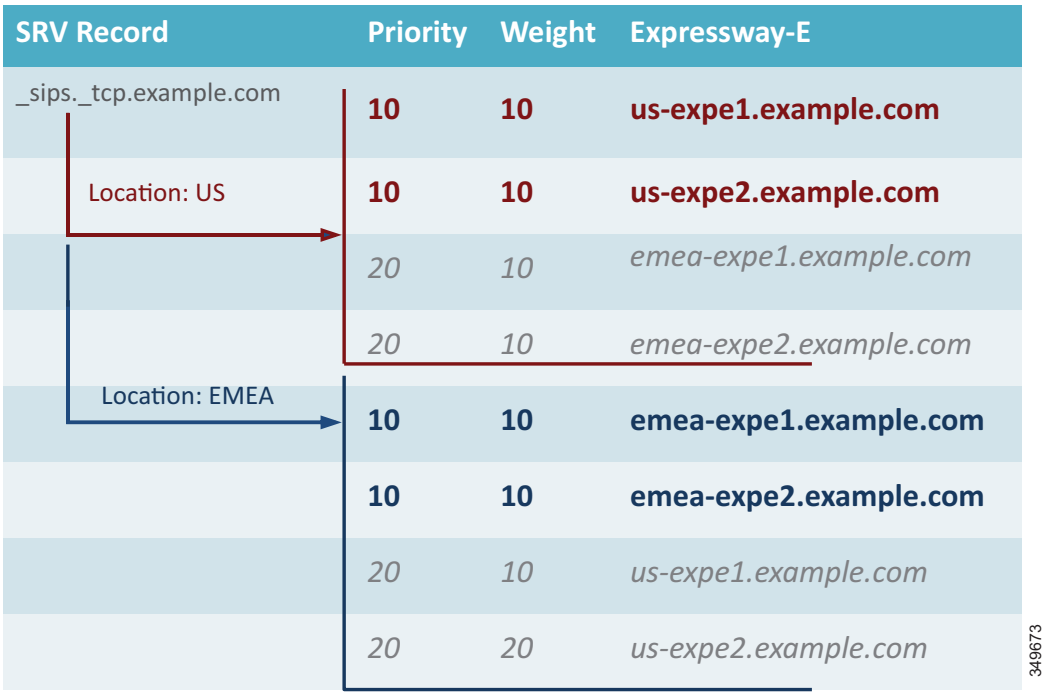
The following examples explain how to configure DNS for GeoDNS services.

In our example scenario, two Internet edge Expressway clusters are deployed, one in the US and one in Europe, each composed of two Expressway-C and Expressway-E servers. If the measured latency between the calling endpoint and the European edge is less than the latency between the endpoint and the US edge, or if the endpoint IP address matches the range for the US, the endpoint will be directed to the European edge for registration based on the configured policy (latency or IP address).

Although some GeoDNS providers support GeoDNS services on SRV records, many others allow GeoDNS for CNAME or A-records only. The recommendation is to implement GeoDNS services on SRV records because this allows for a simpler configuration and easy troubleshooting. A GeoDNS configuration for SRV records is shown in the following example.

If the calling user is in the US, the call will be sent to the US; but if the US data center is down, the call will be sent to EMEA. This configuration allows for geographic redundancy and is shown in Figure 5-10.

Figure 5-10 GeoDNS Configuration for SRV Records



However, if your GeoDNS provider allows you to specify GeoDNS services for CNAME records only and not for SRV records, the following example shows how to configure the GeoDNS if only CNAME is supported for GeoDNS services.

Following this scenario, a DNS SRV record resolves into a CNAME record which, in turn, resolves into an A-record. CNAME records can be assigned a geographic location. As an example, consider an Expressway-E cluster in the US and another Expressway-E cluster in EMEA. A SRV record *_sips._tcp.example.com* for SIP TLS and/or *_sip._tcp.example.com* is configured for business-to-business calls. This record resolves into *alias1.example.com*, a CNAME record.

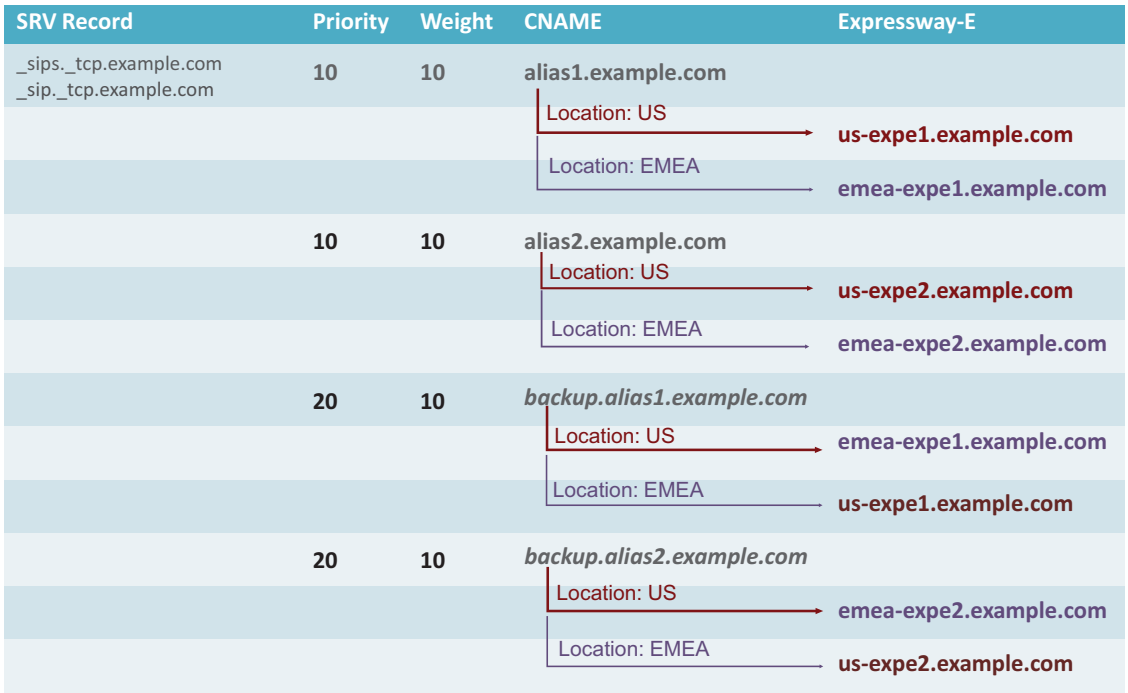
Based on the GeoDNS configuration, a label is applied to the CNAME record to identify the region where the record is active. In this case, the CNAME resolution will be an A-record for the US and another A-record for EMEA, with highest priority (10 in this example). This will address the first peer of the cluster in both regions.

The second CNAME record will be resolved into the second peer of US and EMEA clusters with highest priority. This needs to be repeated until all peers of the cluster are included.

In order to have geographic redundancy, backup CNAME aliases have to be created. In the example in Figure 5-11, *backup-alias1.example.com* resolves into the first EMEA Expressway peer for US users and into the first US Expressway peer for EMEA users, thus providing geographic redundancy for both regions. This backup alias process has to be repeated until all peers of the cluster are included. Those backup records will be used only if the first ones are not answering, because the DSN SRV is set to a lower priority (20 in the example).

Figure 5-11 shows the DNS record structure for GeoDNS services applied to CNAME records.

Figure 5-11 GeoDNS Record Structure for CNAME Records with Geographic Redundancy



Two Different Expressway Edges without GeoDNS

Though the recommendation is to adopt the GeoDNS approach, there might be cases where GeoDNS cannot be deployed; for example, in those cases where the DNS records are managed by a service provider that does not offer the GeoDNS services, or when the multiple edges are deployed in regions that are smaller than the capacity of GeoDNS to select between them. As an example, GeoDNS might be able to distinguish if the calling endpoint location is in California or in Pennsylvania, but it might be not be able to distinguish if the calling location endpoint is San Jose or San Diego. So GeoDNS could not be used if the two Expressway clusters are located in San Jose and in San Diego.

An alternative solution is designed to return the edge that is closest to the destination endpoint or device. This requires finding or knowing where the destination endpoint is located and then returning the appropriate edge. The benefit of this solution is to minimize the use of bandwidth on the customer network by delivering the shortest internal path to the endpoint.

In this scenario, business-to-business SRV records are set with the same priority and weight for all Expressway servers.

As an example, consider two Expressway-C and Expressway-E clusters in EMEA, and another two Expressway-C and Expressway-E clusters in APJC. The Unified CM inbound calling search space on the Expressway-C trunk in EMEA will contain the partition of the EMEA phones but not the partition of the APJC phones. Analogously, the inbound calling search space on the Expressway-C trunk in APJC will contain the partition of the APJC phones but not the partition of the EMEA phones. If a user on the Internet in EMEA calls a corporate endpoint in APJC, the call might be sent to either the EMEA or APJC Expressway cluster.

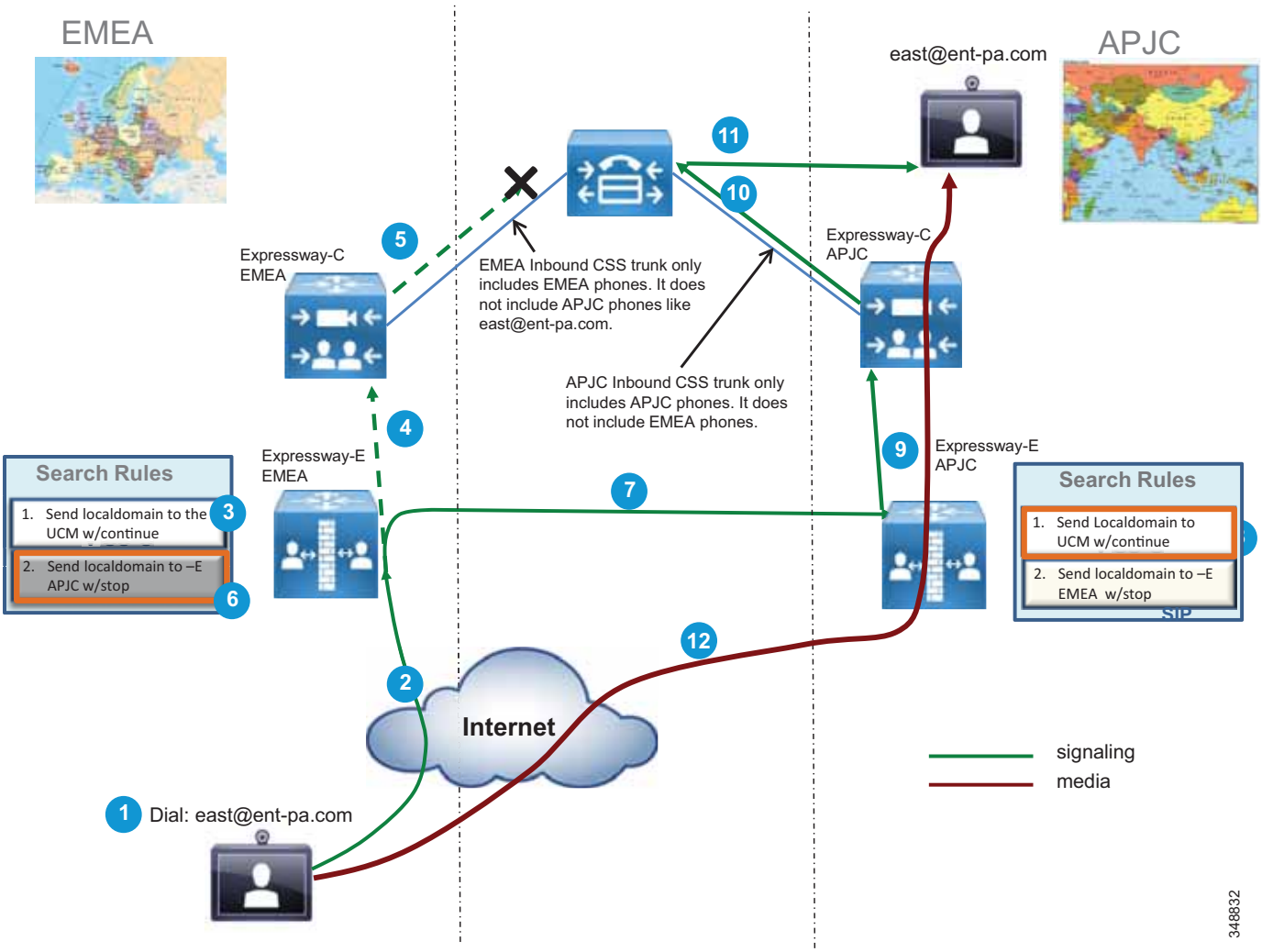
In this example, assume that the call is sent to the EMEA Expressway-E cluster. The EMEA Expressway-E and Expressway-C will try to send the call to the destination, but the inbound calling search space of the Expressway-C trunk will block the call. The EMEA Expressway-E will then forward the call to the APJC Expressway-E. This time the call will be delivered to the destination because the inbound calling search space of APJC Expressway-C contains the APJC endpoint's partition.

In order to allow the Expressway-E in EMEA to remove itself from the signaling and media path, it is important to make sure that there is no TCP-to-TLS or RTP-to-SRTP conversion on Expressway-E EMEA clusters, and to make sure that the call signaling optimization parameter is set to **On** in all Expressway-C and Expressway-E nodes.

Because this is not a deterministic process, in the case of three or more Expressway edges the searching mechanism would require too much time. Therefore, this configuration is recommended for no more than two Expressway edges. If more than two edges are required, the recommendation is to deploy a Directory Expressway architecture. Directory Expressway architecture is not covered in this document.

Figure 5-12 shows the Expressway edge design that enables selection of the edge closest to the destination endpoint.

Figure 5-12 Selection of the Expressway Cluster Closest to the Destination

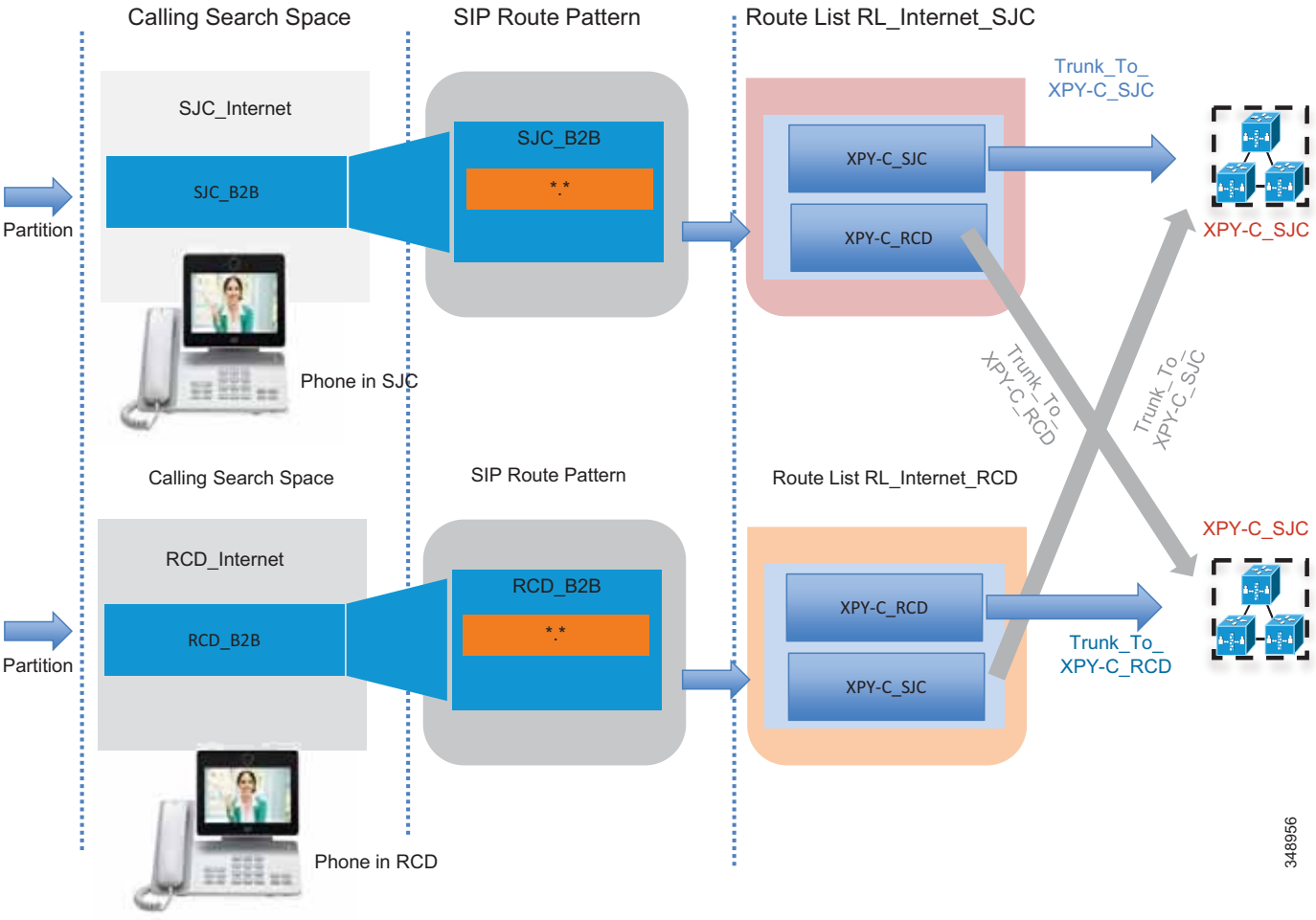


This architecture can scale to more than two sites, and it needs a central Expressway node called Directory Expressway. Directory Expressway is an Expressway acting as a transit node between Expressways in different regions. Directory Expressway architecture is not currently covered in this document.

Considerations for Outbound Calls

Outbound calls should be directed to the Expressway-C that is nearest to the calling endpoint. This can be achieved by using Cisco Unified CM mechanisms such as calling search spaces and partitions. Figure 5-13 shows the Unified CM configuration.

Figure 5-13 Unified CM Configuration to Direct Outbound Calls to the Nearest Expressway-C Cluster, and Use a Backup Cluster if the Nearest One is not Available



The Unified CM Local Route Group feature helps scale this solution when multiple sites access two or more Expressway-C clusters. This mechanism is also applied on ISDN gateways and Cisco Unified Border Element. A full description of the configuration is documented in the next two sections, since it also applies to Cisco Unified Border Element and voice gateways.

Best Practices for Gateways

This section addresses the following best practices with regard to gateways:

Tuning Gateway Gain Settings

Connecting a Cisco Unified Communications network to the PSTN through gateways requires that you properly address media quality issues arising from echo and signal degradation due to power loss, impedance mismatches, delay, and so forth. For this purpose, you must establish a Network Transmission Loss Plan (NTLP), which provides a complete picture of signal loss in all expected voice paths. Using this plan, you can identify locations where signal strength must be adjusted for optimum loudness and effective echo cancellation. Note that not all carriers use the same loss plan, and that the presence of cellular networks adds further complexity in creating the NTLP. Cisco does not recommend adjusting input gain and output attenuation on gateways without first completing such an NTLP. For more information, refer to *Echo Analysis for Voice Over IP*, available at

http://www.cisco.com/en/US/docs/ios/solutions_docs/voip_solutions/EA_ISD.pdf

Routing Inbound Calls from the PSTN

Use one of the following methods to route inbound calls from the PSTN:

- Assign a single directory number to each user for both video and voice calls. This method is not recommended because all calls would have to be received from the PSTN on a video gateway, including audio-only calls. This would waste valuable video gateway resources and be hard to scale.
- Assign at least two different directory numbers to each video-enabled device in the Unified CM cluster, with one line for audio and another line for video. With this method, the outside (PSTN) caller must dial the correct number to enable video.
- For video calls, have outside callers dial the main number of the video gateway. Cisco ISDN and Serial gateways offer an integrated auto-attendant that prompts the caller to enter the extension number of the party they are trying to reach. Unified CM will then recognize that it is a video call when ringing the destination device. This method relieves the caller from having to remember two different DID numbers for each called party, but it adds an extra step to dialing an inbound video call.

Note The outside video endpoints must support DTMF in order to enter the extension of the called party at the IVR prompt.

The following example illustrates the second method:

A user has a Cisco Unified IP Phone with video capabilities enabled. The extension of the IP Phone is 51212, and the fully qualified DID number is 1-408-555-1212. To reach the user from the PSTN for a voice-only call, people simply dial the DID number. The CO sends calls to that DID number through T1-PRI circuit(s) connected to a Cisco Voice Gateway. When the call is received by the gateway, Unified CM knows that the gateway is capable of audio only, so it negotiates only a single audio channel for that call. Conversely, for people to reach the user from the PSTN for a video call, they must dial the main number of the video gateway and then enter the user's extension. For example, they might dial 1-408-555-1000. The CO would send calls to that number through the T1-PRI circuit(s) connected to a Cisco ISDN video gateway. When the call is received by the gateway, an auto-attendant prompt asks the caller to enter the extension of the person they are trying to reach. When the caller enters the extension via DTMF tones, Unified CM knows that the gateway is capable of video, so it negotiates both audio and video channels for that call.

Gateway Digit Manipulation

The Cisco TelePresence ISDN Gateways 8321 and 3241 and the Cisco TelePresence Serial Gateways 8330 and 3340 all have capabilities for digit manipulation. It is possible to set up multiple dial plan rules on these video gateways. These rules match based on calling and/or called number and work in either the IP-to-PSTN or PSTN-to-IP direction. When an inbound call matches a configured dial plan rule, the ISDN or Serial gateway can take one of the following actions:

- Reject the call
- Enter the Auto Attendant
- Place a call to a number (or IP address, hostname, or URI in the case of a PSTN-to-IP call)

When the action is to place a call to a number, the original called number or parts of it can be used in the new number to call.

For more details, refer to the following documentation:

- Cisco TelePresence ISDN Gateway documentation, available at

  http://www.cisco.com/en/US/products/ps11448/tsd_products_support_series_home.html
- Cisco TelePresence Serial Gateway documentation, available at

  http://www.cisco.com/en/US/products/ps11605/tsd_products_support_series_home.html

Routing Outbound Calls to the PSTN

Use one of the following methods to route outbound calls to the PSTN:

- Assign different access codes (that is, different route patterns) for voice and video calls. For example, when the user dials 9 followed by the PSTN telephone number they are trying to reach, it could match a route pattern that directs the call out a voice gateway. Similarly, the digit 8 could be used for the route pattern that directs calls out a video gateway.
- Assign at least two different directory numbers on each video-enabled device in the Unified CM cluster, with one line for audio and another line for video. The two lines can then be given different calling search spaces. When users dial the access code (9, for example) on the first line, it could be directed out a voice gateway, while dialing the same access code on the second line could direct the call out a video gateway. This method alleviates the need for users to remember two different access

codes but requires them to press the correct line on their phones when placing calls. However not all Cisco video endpoints support multiple lines at this time, in which case prefixes would be the preferred method for routing outbound calls to the PSTN.

Video Gateway Call Bandwidth

The Cisco TelePresence ISDN Gateway dial plan rules can be configured so that calls with a certain prefix are limited to a maximum amount of bandwidth on the ISDN connection to the PSTN. This is useful to ensure that a single call cannot monopolize the entire PRI link. When you configure a service prefix in the gateway, you can choose one of the following maximum speeds:

- 128 kbps
- 192 kbps
- 256 kbps
- 320 kbps
- 384 kbps
- 512 kbps
- 768 kbps
- 1152 kbps
- 1472 kbps

Calls from an IP endpoint toward the PSTN can include the service prefix at the beginning of the called number in order for the gateway to decide which service to use for the call. Optionally, you can configure the default prefix to be used for calls that do not include a service prefix at the beginning of the number. This method can become quite complex because users will have to remember which prefix to dial for the speed of the call they wish to make, and you would have to configure multiple route patterns in Unified CM (one for each speed).

Note Two global settings on the Cisco TelePresence ISDN Gateway can be used to set a minimum or maximum bandwidth value for incoming and outgoing ISDN calls. The dial plan cannot override this value with a higher maximum bandwidth; however, a dial plan can impose a lower bandwidth for particular calls.

Automated Alternate Routing (AAR)

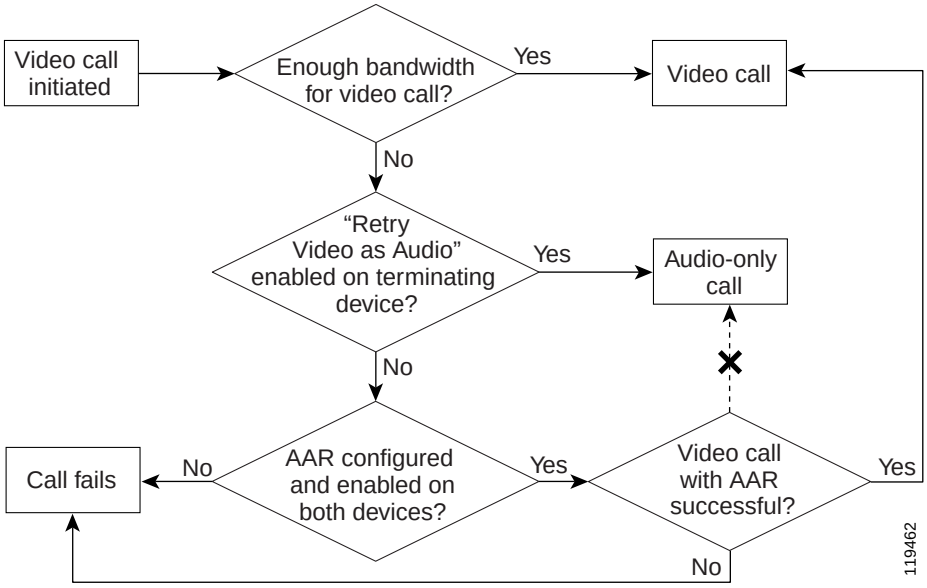
When the IP network does not have enough bandwidth available to process a call, Unified CM uses its call admission control mechanism to determine what to do with the call. Unified CM performs one of the following actions with the call, depending on how you have configured it:

- Fail the call, playing busy tone to the caller and displaying a Bandwidth Unavailable message on the caller's screen
- Retry a video call as an audio-only call
- Use automated alternate routing (AAR) to re-route the call over an alternative path, such as a PSTN gateway

The Retry Video Call as Audio option takes effect only on the terminating (called) device, thus allowing the flexibility for the calling device to have different options (retry or AAR) for different destinations.

If a video call fails due to bandwidth limitations but automated alternate routing (AAR) is enabled, Unified CM will attempt to reroute the failed call as a video call to the AAR destination. If AAR is not enabled, the failed call will result in a busy tone and an error message will be sent to the caller. (See Figure 5-14.)

Figure 5-14 Possible Scenarios for a Video Call



To provide AAR for voice or video calls, you must configure the calling and called devices as members of an AAR group and configure an External Phone Number Mask for the called device. The External Phone Number Mask designates the fully qualified E.164 address for the user's extension, and the AAR group indicates what digits should be prepended to the External Phone Number Mask of the called device in order for the call to route successfully over the PSTN.

For example, assume that user A is in the San Jose AAR group and user B is in the San Francisco AAR group. User B's extension is 51212, and the External Phone Number Mask is 6505551212. The AAR groups are configured to prepend 91 for calls between the San Jose and San Francisco AAR groups. Thus, if user A dials 51212 and there is not enough bandwidth available to process the call over the IP WAN between those two sites, Unified CM will take user B's External Phone Number Mask of 6505551212, prepend 91 to it, and generate a new call to 916505551212 using the AAR calling search space for user A.

By default, all video-capable devices in Unified CM have the Retry Video Call as Audio option enabled (checked). Therefore, to provide AAR for video calls, you must disable (uncheck) the Retry Video Call as Audio option. Additionally, if a call admission control policy based on Resource Reservation Protocol (RSVP) is being used between locations, the RSVP policy must be set to Mandatory for both the audio and video streams.

Furthermore, Unified CM looks at only the called device to determine whether the Retry Video Call as Audio option is enabled or disabled. So in the scenario above, user B's phone would have to have the Retry Video Call as Audio option disabled in order for the AAR process to take place.

Finally, devices can belong to only one AAR group. Because the AAR groups determine which digits to prepend, AAR groups also influence which gateway will be used for the rerouted call. Depending on your choice of configuration for outbound call routing to the PSTN, as discussed in the previous section,

video calls that are rerouted by AAR might go out a voice gateway instead of a video gateway. Therefore, carefully construct the AAR groups and the AAR calling search spaces to ensure that the correct digits are prepended and that the correct calling search space is used for AAR calls.

While these considerations can make AAR quite complex to configure in a large enterprise environment, AAR is easier to implement when the endpoints are strictly of one type or the other. When endpoints are capable of both audio and video calls (such as Cisco Unified IP Phone 9971 or a Cisco TelePresence System EX90), the configuration of AAR can quickly become unwieldy. Therefore, Cisco recommends that large enterprise customers who have a mixture of voice and video endpoints give careful thought to the importance of AAR for each user, and use AAR only for select video devices such as dedicated videoconference rooms or executive video systems. Table 5-5 lists scenarios when it is appropriate to use AAR with various device types.

Table 5-5 When to Use AAR with a Particular Device Type

Device Type	Device is used to call:	Enable AAR?	Comments
IP Phone	Other IP Phones and video-capable devices	Yes	Even when calling a video-capable device, the source device is capable of audio-only, thus AAR can be configured to route calls out a voice gateway.
Cisco Jabber or Cisco Unified IP Phone 9971	Other video-capable devices only	Yes	Because the device is used strictly for video calls, you can configure the AAR groups accordingly.
	IP Phones and other video-capable devices	No	It will be difficult to configure the AAR groups to route audio-only calls differently than video calls.
H.323 or SIP client	Other video-capable devices only	Yes	Because the device is used strictly for video calls, you can configure the AAR groups accordingly.
	IP Phones and other video-capable devices	No	It will be difficult to configure the AAR groups to route audio-only calls differently than video calls

Least-Cost Routing

Least-cost routing (LCR) and tail-end hop-off (TEHO) are very popular in VoIP networks and can be used successfully for video calls as well. In general, both terms refer to a way of configuring the call routing rules so that calls to a long-distance number are routed over the IP network to the gateway closest to the destination, in an effort to reduce toll charges. Unified CM supports this feature through its rich set of digit analysis and digit manipulation capabilities, including:

- Partitions and calling search spaces
- Translation patterns
- Route patterns and route filters
- Route lists and route groups

Configuring LCR for video calls is somewhat more complicated than for voice calls, for the following reasons:

- Video calls require their own dedicated gateways, as discussed previously in this chapter
- Video calls require much more bandwidth than voice calls

With respect to dedicated gateways, the logic behind why you might or might not decide to use LCR for video calls is very similar to that explained in the section on Automated Alternate Routing (AAR), page 5-35. Due to the need to have different types of gateways for voice and video, it can become quite complex to configure all the necessary partitions, calling search spaces, translation patterns, route patterns, route filters, route lists, and route groups needed for LCR to route voice calls out one gateway and video calls out another.

With respect to bandwidth requirements, the decision to use LCR depends on whether or not you have enough available bandwidth on your IP network to support LCR for video calls to/from a given location. If the current bandwidth is not sufficient, then you have to determine whether the benefits of video calls are worth the cost of either upgrading your IP network to make room for video calls or deploying local gateways and routing calls over the PSTN. For example, suppose you have a central site with a branch office connected to it via a 1.544-Mbps T1 circuit. The branch office has twenty video-enabled users in it. A 1.544-Mbps T1 circuit can handle at most about four 384-kbps video calls. Would it really make sense in this case to route video calls up to the central site in order to save on toll charges? Depending on the number of calls you want to support, you might have to upgrade your 1.544-Mbps T1 circuit to something faster. Is video an important enough application to justify the additional monthly charges for this upgrade? If not, it might make more sense to deploy a Cisco video gateway at the branch office and not bother with LCR. However, placing local Cisco video gateways at each branch office is not inexpensive either, so ultimately you must decide how important video-to-PSTN calls are to your business. If video is not critical, perhaps it is not worth upgrading the bandwidth or buying video gateways but, instead, using the Retry Video Call as Audio feature to reroute video calls as voice-only calls if they exceed the available bandwidth. Once a call is downgraded to voice-only, local gateway resources and bandwidth to perform LCR become more affordable and easier to configure.

Fax and Modem Support

For information on fax and modem support across Cisco gateways refer to the following documentation:

- The *Gateways* chapter of the *Cisco Unified Communications System 9.0 SRND*, available at http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/gateways.html
- *Fax, Modem, and Text Support over IP Configuration Guide*, available at http://www.cisco.com/en/US/docs/ios-xml/ios/voice/fax/configuration/15-mt/vf-15-mt-book.html