CHAPTER **13**

Cisco Voice Messaging

Last revised on: June 4, 2010

This chapter describes the voice messaging solutions available in Cisco Unified Communications System Release 7.*x*. It includes the Cisco voice messaging products Cisco Unity, Cisco Unity Connection, and Cisco Unity Express, and it covers the design guidelines and best practices for deploying these products together with Cisco Unified Communications Manager (Unified CM).

Note   In previous versions of this Solution Reference Network Design (SRND) document, the information in this chapter was presented as two separate chapters on Cisco Unity and Cisco Unity Express.

Although this guide focuses on the messaging deployment scenarios with regard to Unified CM, Cisco Unified Communications Manager Express (Unified CME) is also noted where applicable, especially when used with Survivable Remote Site Telephony (SRST) fallback support in a centralized Unified CM deployment.

This chapter covers the following topics:

- Voice Messaging Portfolio, page 13-2
- Messaging Deployment Models, page 13-5
- Messaging and Unified CM Deployment Model Combinations, page 13-7
- Fax Deployment, page 13-26
- Voicemail Networking, page 13-28
- Best Practices for Voice Messaging, page 13-31

The chapter begins with a short description of each of the products in the Cisco messaging solutions portfolio and provides a simple overview of where each product fits in an enterprise Unified Communications solution. Next, messaging deployment models form the basis of discussion for voicemail integrations, which start with a definition of the various messaging deployment models and then explain how each of the messaging deployment models fits into the various Unified CM call processing deployment models. Cisco Unity and Unity Connection are discussed together in this section, while Cisco Unity Express has a dedicated section for its supported deployment models. Many system-level design considerations and best practices are explained in this section.

Next, supported fax integration types are described in the section on fax integration. This section is organized by product type. Finally, the section on best practices highlights some general best practices and design guidelines that were not mentioned previously yet are important aspects of the product that should be considered when deploying the respective messaging products in the solution.

Keep in mind that this is a high-level design discussion and is focused on how the voice messaging products fit into the Unified Communications System release with Unified CM. For detailed design guidelines for each product as well as interoperability information for third-party messaging and telephony systems, refer to the product-specific design guides on http://www.cisco.com.

What's New in This Chapter

Table 13-1 lists the topics that are new in this chapter or that have changed significantly from previous releases of this document.

Table 13-1   New or Changed Information Since the Previous Release of This Document

New or Revised Topic	Described in:
Best practices for Cisco Unity Express	Best Practices for Deploying Cisco Unity Express, page 13-41
Cisco Unified Messaging Gateway supported modules, capacities, and best practices	Voicemail Networking with Cisco Unified Messaging Gateway, page 13-30
Cisco voice messaging portfolio	Voice Messaging Portfolio, page 13-2
Deployment model details for Cisco Unity Express	Cisco Unity Express Deployment Models, page 13-19
Digital networking	Distributed Messaging with Centralized Call Processing, page 13-9
Enhanced Message Waiting Indicator (eMWI)	Enhanced Message Waiting Indicator (eMWI), page 13-35
Failover operation and deployment	Messaging Redundancy, page 13-16
Fax integration	Fax Deployment, page 13-26
LDAP integration	Distributed Messaging with Centralized Call Processing, page 13-9
Microsoft Exchange integration	Voice Messaging Portfolio, page 13-2

Voice Messaging Portfolio

The Cisco Unified Communications messaging portfolio consists of three main messaging products: Cisco Unity, Cisco Unity Connection, and Cisco Unity Express. Each product fits different requirements yet each one contains overlapping features and scalability with regard to the others. They also have the ability to interwork with one another using Voice Mail Networking and can also leverage the Cisco Unified Messaging Gateway to achieve this in a highly scalable fashion, as discussed later in this chapter.

When considering these products, it helps to think of the messaging types that the products apply to in order to understand the messaging options they include and to determine which options could fit your deployment requirements. The following definitions help define these messaging types:

- *Voicemail-only* refers to a telephony voicemail integration where there is no access to the voicemail via any messaging client.
- *Integrated messaging* refers to voicemail with telephony access as well as voicemail-only access via a messaging client.

- *Unified messaging* refers to voicemail with telephony access as well as voicemail, email, and fax access via a messaging client.

Based on the above messaging types and definitions, the three messaging product options are:

- Cisco Unity

    This solution option scales to meet the needs of large enterprise organizations and delivers powerful voice, integrated, and unified messaging options that integrate with Microsoft Exchange (including Exchange 2007) and Lotus Domino.

- Cisco Unity Connection

    This option combines integrated messaging, voice recognition, and call transfer rules into an easy-to-manage system for medium-sized businesses with up to 10,000 users, or it can network up to 5 systems to support larger-sized businesses with up to 50,000 users. For organizations with up to 500 users, Cisco Unity Connection is available as a single-server solution with Cisco Unified Communications Manager Business Edition.

- Cisco Unity Express

    This option provides cost-effective voice and integrated messaging, automated attendant, and interactive voice response (IVR) capabilities in certain Cisco Integrated Services Routers for small and medium-sized businesses and enterprise branch offices with up to 250 users.

For a complete comparison of product feature, refer to the *Cisco Unified Messaging Feature Comparison*, available on http://www.cisco.com.

Figure 13-1 depicts a product comparison with regard to scalability. The first line indicates the maximum supported users for a single server (or redundant/failover server pair). The second line refers to a fully meshed, digitally networked deployment of the individual products. Note that, when leveraging the Cisco Unified Messaging Gateway (UMG) to network the various products, the number of supported users increases significantly and is directly related to the number of networked nodes and maximum users supported by the UMG. See Voicemail Networking with Cisco Unified Messaging Gateway, page 13-4, for more details on scalability.

Figure 13-1   *Cisco Voice Messaging Solutions*
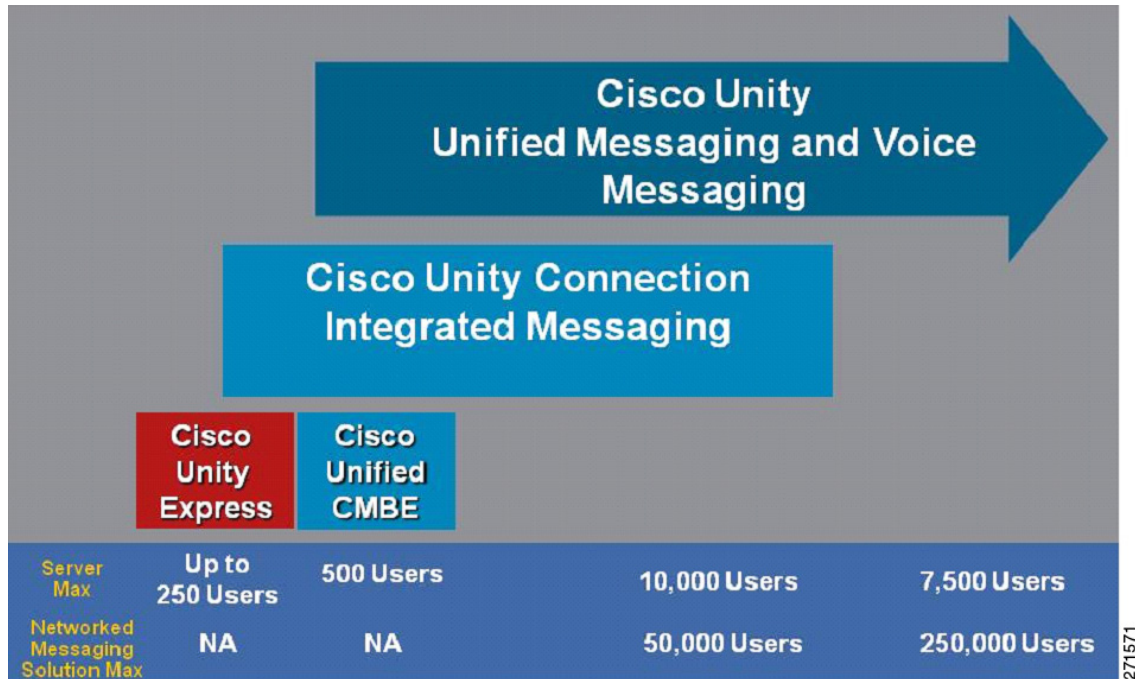
Voicemail Networking with Cisco Unified Messaging Gateway

Cisco Unified Messaging Gateway (UMG) enables Cisco's end-to-end networked voice messaging solution by providing intelligent voice message routing, management of system directories, messaging format, and delivery of a scalable voice messaging framework. Cisco UMG supports Cisco Unity, Cisco Unity Express, Cisco Unity Connection, and Avaya Interchange.

In Figure 13-1, Networked Messaging Solution Max refers to the maximum users supported in a full-mesh digital networking environment. A VPIM networking solution leveraging the Cisco Unified Messaging Gateway can support a maximum of 500,000 users or subscribers.

This chapter focuses on the design aspects of integrating Cisco Unity, Cisco Unity Connection, and Cisco Unity Express with Cisco Unified Communications Manager (Unified CM). This chapter focuses on the following components and versions:

- Cisco Unified Communications Manager (Unified CM) 7.*x*
- Cisco Unified Communications Manager Express (Unified CME) 4.*x*
- Cisco Unity 7.*x* and 4.*x*
- Cisco Unity Connection 7.*x*
- Cisco Unity Express 3.2

Cisco Unified CM 7.*x* provides functionality for Session Initiation Protocol (SIP) trunks, which support integration directly to Cisco Unity and Unity Connection without the need for a SIP proxy server.

For information on earlier releases of Cisco Unity, Unity Connection, Unity Express, and Unified CM or Unified CM Express, refer to the appropriate online product documentation available at http://www.cisco.com.

As mentioned, the design topics covered in this chapter apply to voicemail-only, unified messaging, and integrated messaging configurations. Additionally, this chapter discusses design issues for deploying Cisco Unity with Microsoft Exchange (2000, 2003, or 2007) or Lotus Domino message stores and Microsoft Windows (2000 or 2003). This chapter does not cover deployments or upgrades from Microsoft NT 4.0 and/or Exchange 5.5. Cisco Unity Connection and Unity Express have no dependencies on an external message store. Although Cisco Unity Connection 7.0 supports Exchange integration, it is not dependent upon Exchange integration as is Cisco Unity.

For additional design information about Cisco Unity or Cisco Unity Connection, including integrations with other non-Cisco messaging systems, refer to the *Design Guide for Cisco Unity* or the *Design Guide for Cisco Unity Connection*, respectively, available at http://www.cisco.com.

For additional design information about Cisco Unity Express, including integrations with other non-Cisco messaging systems, refer to the applicable product documentation, available at http://www.cisco.com.

For additional design information about Cisco Unified Messaging Gateway, including integrations with other non-Cisco messaging systems, refer to the Cisco Unified Messaging Gateway product documentation, available at http://www.cisco.com.

Messaging Deployment Models

This section summarizes the various messaging deployment models for Cisco Unity, Cisco Unity Connection, and Cisco Unity Express. For a complete discussion of the deployment models and design considerations specific to Cisco Unity, Unity Connection, and the various messaging components, refer to the *Design Guide for Cisco Unity* or the *Design Guide for Cisco Unity Connection*, respectively, available at http://www.cisco.com. For Cisco Unity Express, refer to the applicable product documentation available at http://www.cisco.com.

Cisco Unity and Unity Connection support three primary messaging deployment models:

- Single-site messaging
- Multisite WAN deployment with centralized messaging
- Multisite WAN deployment with distributed messaging

Cisco Unity Express also supports three primary messaging deployment models:

- Single-site messaging
- Multisite WAN deployment with distributed messaging
- Multisite WAN deployment with distributed messaging with Cisco Unified CME

Note   The Cisco Unity Express 3.2 supports centralized voice messaging for up to 10 Unified CMEs. For more information, refer to the Cisco Unified Communications Manager Express documentation on http://www.cisco.com.

Although the call processing deployment models for Cisco Unified CM and Unified CME are independent of the messaging deployment models for Cisco Unity, Unity Connection, and Unity Express, each has implications toward the other that must be considered.

In addition to the three messaging deployment models, Cisco Unity also supports messaging redundancy. (See Messaging Redundancy, page 13-16.) Starting with Cisco Unity Connection 7.0, messaging redundancy is also available in an active/active configuration. For more information, refer to the *Design Guide for Cisco Unity Connection* available on http://www.cisco.com.

All messaging deployment models support voicemail, integrated messaging, and unified messaging installations. The following definitions apply to these configurations:

- *Voicemail-only* refers to a telephony voicemail integrations where there is no access to the voicemail via any messaging client.

- *Integrated messaging* refers to voicemail with telephony access as well as voicemail-only access via a messaging client.

- *Unified messaging* refers to voicemail with telephony access as well as voicemail, email, and fax access via a messaging client.

Table 13-2 shows which Cisco products support these types of messaging.

Table 13-2   Supported Messaging Environments per Product

Messaging Type	Cisco Unity	Cisco Unity Connection	Cisco Unity Express
Voicemail-only	Yes	Yes	Yes
Integrated messaging	Yes	Yes	Yes
Unified messaging	Yes	No	No

Single-Site Messaging

In this model, the messaging systems and messaging infrastructure components are all located at the same site, on the same highly available LAN. The site can be either a single site or a campus site interconnected via high-speed metropolitan area networks (MANs). All clients of the messaging system are also located at the single (or campus) site. The key distinguishing feature of this model is that there are no remote clients.

Centralized Messaging

In this model, similar to the single-site model, all the messaging system and messaging infrastructure components are located at the same site. The site can be one physical site or a campus site interconnected via high-speed MANs. However, unlike the single-site model, centralized messaging clients can be located both locally and remotely.

Because messaging clients may be either local or remote from the messaging system, special design considerations apply to the following graphical user interface (GUI) clients: ViewMail for Outlook (VMO) and the use of the telephone record and playback (TRaP) and message streaming features. Remote clients should not use TRaP and should be configured to download messages before playback. Because different features and operations for local and remote clients can cause user confusion, TRaP should be disabled on the voice ports and GUI clients should be configured to download messages and not use TRaP, regardless of whether the client is local or remote. This also applies to ViewMail for Outlook (VMO) for Unity Connection IMAP clients, which is new in Cisco Unity Connection 7.0.

The Cisco Unity telephone user interface (TUI) operates the same way for both local and remote clients.

Distributed Messaging

A distributed messaging model consists of multiple single-site messaging systems distributed with a common messaging backbone. There can be multiple locations, each with its own messaging system and messaging infrastructure components. All client access is local to each messaging system, and the

messaging systems share a messaging backbone that spans all locations. Message delivery from the distributed messaging systems occurs via the messaging backbone through a full-mesh or hub-and-spoke type of message routing infrastructure. No messaging infrastructure components should be separated by a WAN from the messaging system they service.

Distributed messaging is essentially multiple, single-site messaging models with a common messaging backbone. The exception to this rule is the PBX-IP Media Gateway (PIMG) and T1-IP Media Gateway (TIMG) integrations. PIMG and TIMG integrations are not discussed in this design document. For further information regarding PIMG or TIMG, refer to the Cisco Unity integration guides available on http://www.cisco.com.

The distributed messaging model has the same design criteria as centralized messaging with regard to local and remote GUI clients, TRaP, and message downloads.

Messaging and Unified CM Deployment Model Combinations

This section discusses the design considerations for integrating the various messaging deployment models with the Unified CM call processing deployment models. Table 13-3 lists the various combinations of messaging and call processing deployment models supported by Cisco Unity, Unity Connection, and Unity Express.

Table 13-3   Supported Combinations of Messaging and Unified CM Call Processing Deployment Models

Model Type	Cisco Unity	Cisco Unity Connection	Cisco Unity Express
Single-site messaging and single-site call processing	Yes	Yes	Yes
Centralized messaging and centralized call processing	Yes	Yes	No[1]
Distributed messaging and centralized call processing	Yes	Yes	Yes
Centralized messaging and distributed call processing	Yes	Yes	No[1]
Distributed messaging and distributed call processing	Yes	Yes	Yes
Centralized messaging with cluster over the WAN	Yes	Yes	No
Distributed messaging with cluster over the WAN	Yes	Yes	Yes

1. Support for centralized voicemail messaging with Unified CME is available starting with Cisco Unity Express 3.2; however, this is not applicable to Unified CM call processing deployment models.

This section covers the following topics:

- Cisco Unity and Unity Connection messaging and Unified CM deployment models
- Cisco Unity Express deployment models

Each topic defines a messaging and Unified CM deployment model combination and then highlights each Cisco voicemail messaging product applicable to that model as well as the design considerations for that model combination. Not all combinations are discussed for each product. Some examples are provided, with best practices and design considerations for each product. The intention is to provide an understanding of the base messaging deployment models and the interaction with Unified CM without detailing all possibilities.

Refer to the *Design Guide for Cisco Unity* and the *Design Guide for Cisco Unity Connection*, available at http://www.cisco.com, for further details on site classification and a detailed analysis of supported combinations of messaging and call processing deployment models.

Cisco Unity and Unity Connection Messaging and Unified CM Deployment Models

This section discusses some of the various combinations of messaging and call processing deployment models for Cisco Unity and Unity Connection.

Centralized Messaging and Centralized Call Processing

In centralized messaging, the voice messaging server is located in the same site as the Unified CM cluster. With centralized call processing, subscribers may be located either remotely and/or locally to the cluster and messaging server(s). (See Figure 13-2.) When remote users access resources at the central site (such as voice ports, IP phones, or PSTN gateways, as in Tail-End Hop-Off (TEHO)), these calls are transparent to gatekeeper call admission control. Therefore, regions and locations must be configured in Unified CM for call admission control. (See Managing Bandwidth, page 13-31.) When making inter-region calls to IP phones or MGCP gateways, IP phones automatically select the inter-region codec that has been configured. With Cisco Unity messaging deployments, native transcoding should be disabled so that the voice ports use Unified CM transcoding resources for calls that transverse the WAN (inter-region). See the section on Native Transcoding Operation, page 13-32, for information on disabling this feature in Cisco Unity.

Figure 13-2   Centralized Messaging with Centralized Call Processing
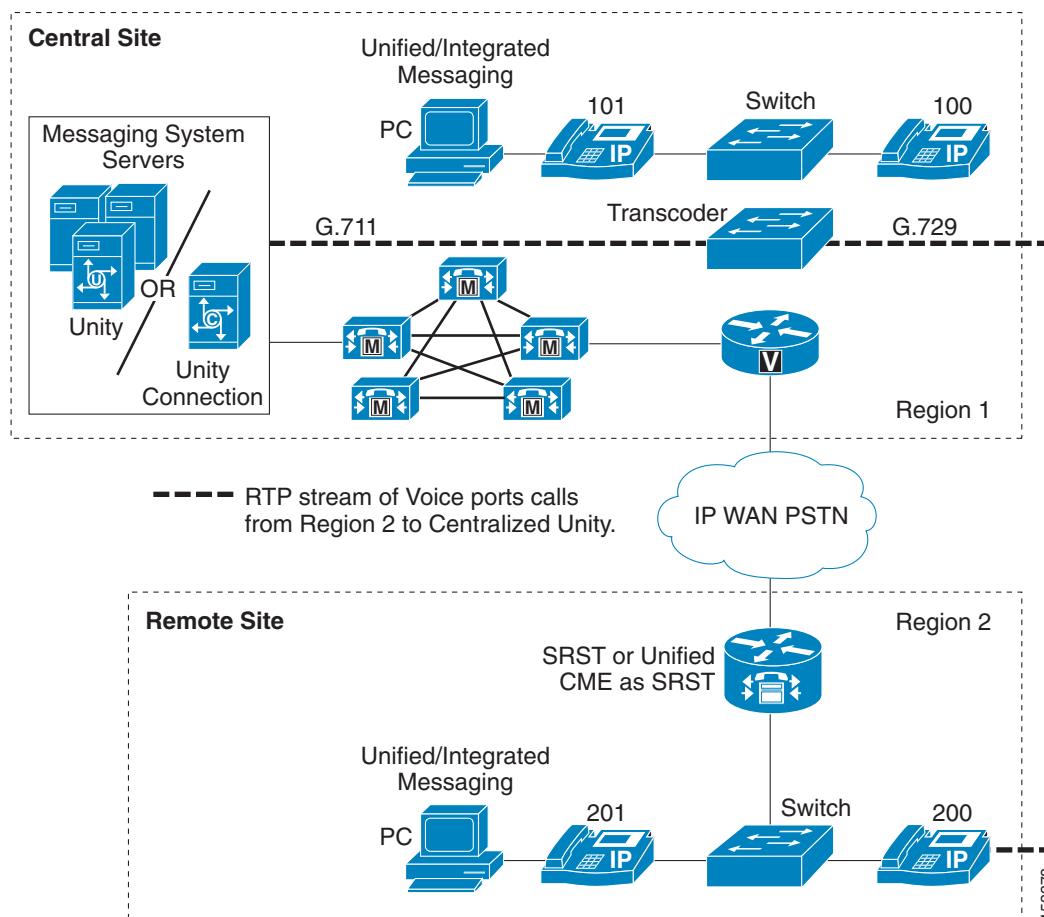
In Figure 13-2, regions 1 and 2 are configured to use G.711 for intra-region calls and G.729 for inter-region calls. Native transcoding has been disabled on the Cisco Unity or Unity Connection server.

As Figure 13-2 shows, when a call is made from extension 200 to the voicemail ports in Region 1, the inter-region G.729 codec is used at the endpoint but the RTP stream is transcoded to use G.711 on the voice ports. Native transcoding on the Cisco Unity or Unity Connection server has been disabled in this example. Unified CM transcoding resources must be located at the same site as the voicemail system.

Impact of Non-Delivery of RDNIS on Voicemail Calls Routed via AAR

In centralized messaging environments, automated alternate routing (AAR), a Unified CM feature, can route calls over the PSTN to the messaging store at the central site when the WAN is oversubscribed. However, when calls are rerouted over the PSTN, Redirected Dialed Number Information Service (RDNIS) can be affected. Incorrect RDNIS information can impact voicemail calls that are rerouted over the PSTN by AAR when Cisco Unity or Unity Connection is remote from its messaging clients. If the RDNIS information is not correct, the call will not reach the voicemail box of the dialed user but will instead receive the auto-attendant prompt, and the caller might be asked to re-enter the extension number of the party they wish to reach. This behavior is primarily an issue when the telephone carrier is unable to ensure RDNIS across the network. There are numerous reasons why the carrier might not be able to ensure that RDNIS is properly sent. Check with your carrier to determine if they provide guaranteed RDNIS deliver end-to-end for your circuits. The alternative to using AAR for oversubscribed WANs is simply to let callers hear reorder tone in an oversubscribed condition.

Distributed Messaging with Centralized Call Processing

Distributed messaging means that there are multiple messaging systems distributed within the telephony environment, and each messaging system services only local messaging clients. This model differs from centralized messaging, where clients are both local and remote from the messaging system.

Figure 13-3 illustrates the distributed messaging model with centralized call processing. As with other multisite call processing models, the use of regions and locations is required to manage WAN bandwidth. For this model, you should also disable native transcoding in Cisco Unity.

Note that Cisco Unified Communications Manager Express (Unified CME) in SRST mode is used for call processing backup of both IP phones and Cisco Unity or Unity Connection voicemail ports. Deployed at the remote site (for example, Region 2 in Figure 13-3), this fallback support provides backup call processing in the event that the phones lose connectivity with Unified CM, such as during a WAN failure, while simultaneously providing users at the remote site with access to the local Cisco Unity or Unity Connection server as well as MWI support during WAN failure. For further details on Unified CME in SRST mode, refer to the Unified CME product documentation on http://www.cisco.com.

Figure 13-3   Distributed Messaging with Centralized Call Processing
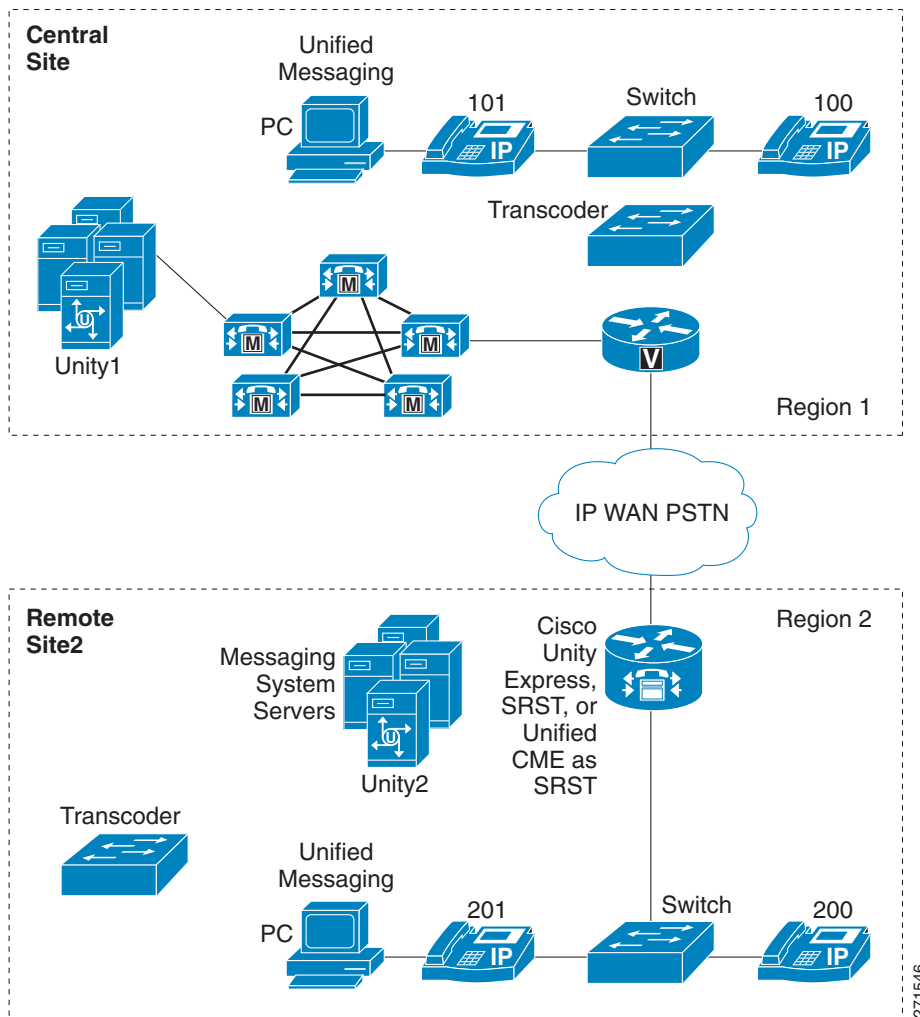
For the configuration in Figure 13-3, transcoder resources must be local to each Cisco Unity message system site. Regions 1 and 2 are configured to use G.711 for intra-region calls and G.729 for inter-region calls. Native transcoding has been disabled on the Cisco Unity servers.

Voice messaging ports for both Cisco Unity or Unity Connection servers must be assigned the appropriate region and location by means of calling search spaces and device pools configured on the Unified CM server. In addition, to associate telephony users with a specific group of voicemail ports, you must configure Unified CM voicemail profiles. For details on configuring calling search spaces, device pools, and voicemail profiles, refer to the applicable version of the *Cisco Unified Communications Manager Administration Guide*, available at http://www.cisco.com.

The messaging systems are "networked" together to present a single messaging system to both inside and outside users. For information about Cisco Unity Networking for the distributed Unity servers, refer to the *Networking in Cisco Unity Guide*, available at

http://www.cisco.com/en/US/products/sw/voicesw/ps2237/products_feature_guides_list.html

Also new to Cisco Unity Connection is support for digital networking, allowing multiple Cisco Unity Connection systems to be networked together. Up to 5 nodes (single or active/active pair) can be connected together to support up to 50,000 entities in the directory. Cisco Unity Connection can integrate

with a corporate directory such as Microsoft Active Directory to synchronize users and use digital networking simultaneously. In this configuration, each Cisco Unity Connection server or server pair will be able to synchronize up to 10,000 users from the corporate directory. Refer to the *Design Guide for Cisco Unity Connection*, available at http://www.cisco.com, for further information regarding digital networking or directory integration in Cisco Unity Connection.

Cisco Unity and Unity Connection with Unified CME in SRST Mode

Unified CME in SRST mode offers the possibility for both Cisco Unity and Unity Connection servers located in remote sites and registered with a Unified CM at the central site to fall-back to Unified CME in the remote location. When the WAN link is down and the phones fail-over to Unified CME in SRST mode, Cisco Unity and Unity Connection voicemail ports can also fail-over to Unified CME in SRST mode to provide the remote site users with access to their voicemail with MWI during the WAN outage.

This scenario requires the following:

- Cisco Unified CME 4.0 or later
- Cisco Unity 4.0(5) or later with TSP version 8.1(3) or later
- Cisco Unity Connection 2.*x* or later

Note   MWI has to be resynchronized from the Cisco Unity or Unity Connection server whenever a failover happens from Unified CM to Unified CME in SRST mode, or vice versa.

Combined Messaging Deployment Models

It is possible to combine messaging models in the same deployment, provided that the deployment adheres to all the guidelines listed in the preceding sections. Figure 13-4 shows a user environment in which both centralized and distributed messaging are employed simultaneously.
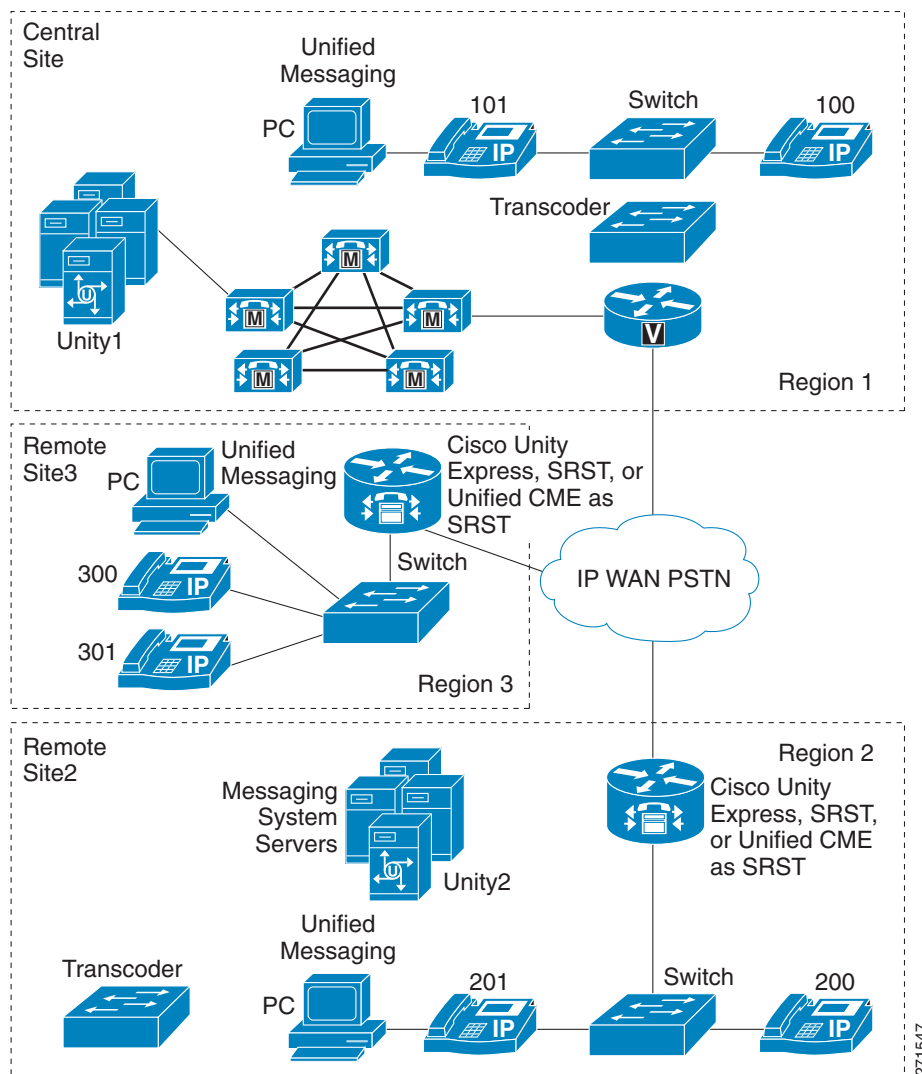
Figure 13-4   Combined Deployment Models

Figure 13-4 shows the combination of two messaging models. Regions 1 and 3 use centralized messaging with centralized call processing, while Region 2 uses distributed messaging with centralized call processing. All regions are configured to use G.711 for intra-region calls and G.729 for inter-region calls.

In Figure 13-4, centralized messaging and centralized call signaling are used between the Central Site and Site3. The messaging system at the Central Site provides messaging services for clients at both the Central Site and Site3. Site2 uses the distributed messaging model with centralized call processing. The messaging system (Unity2) located at Site2 provides messaging services for only those users located within Site2. In this deployment, both models adhere to their respective design guidelines as presented

in this chapter. Transcoding resources are located locally to each messaging system site, and they support clients who access messaging services from a remote site (relative to the messaging system), as in the case of a Site2 user leaving a message for a Central Site user.

In addition, Cisco Unified Communications Manager Express (Unified CME) in SRST mode is used for call processing backup of both IP phones and Cisco Unity or Unity Connection voicemail ports. Deployed at the remote site (for example, Region 2 in Figure 13-4), this fallback support provides backup call processing in the event that the phones lose connectivity with Unified CM, such as during a WAN failure, while simultaneously providing users at the remote site with access to the local Cisco Unity or Unity Connection server as well as MWI support during WAN failure. For further details on Unified CME in SRST mode, refer to the product documentation on http://www.cisco.com.

Centralized Messaging with Clustering Over the WAN

This section addresses Cisco Unity design issues for deploying centralized messaging with Unified CM clustering over the WAN with local failover. In the case of a WAN failure with this model, all remote messaging sites will lose voicemail capability until the WAN is restored. (See Figure 13-5.)

Clustering over the WAN supports local failover. With local failover, each site has a backup subscriber server physically located at the site. This section focuses on deploying Cisco Unity centralized messaging with local failover for clustering over the WAN.

For additional information, refer to the section on Clustering Over the IP WAN, page 2-21.

Figure 13-5   Cisco Unity Centralized Messaging and Clustering Over the WAN with Local Failover
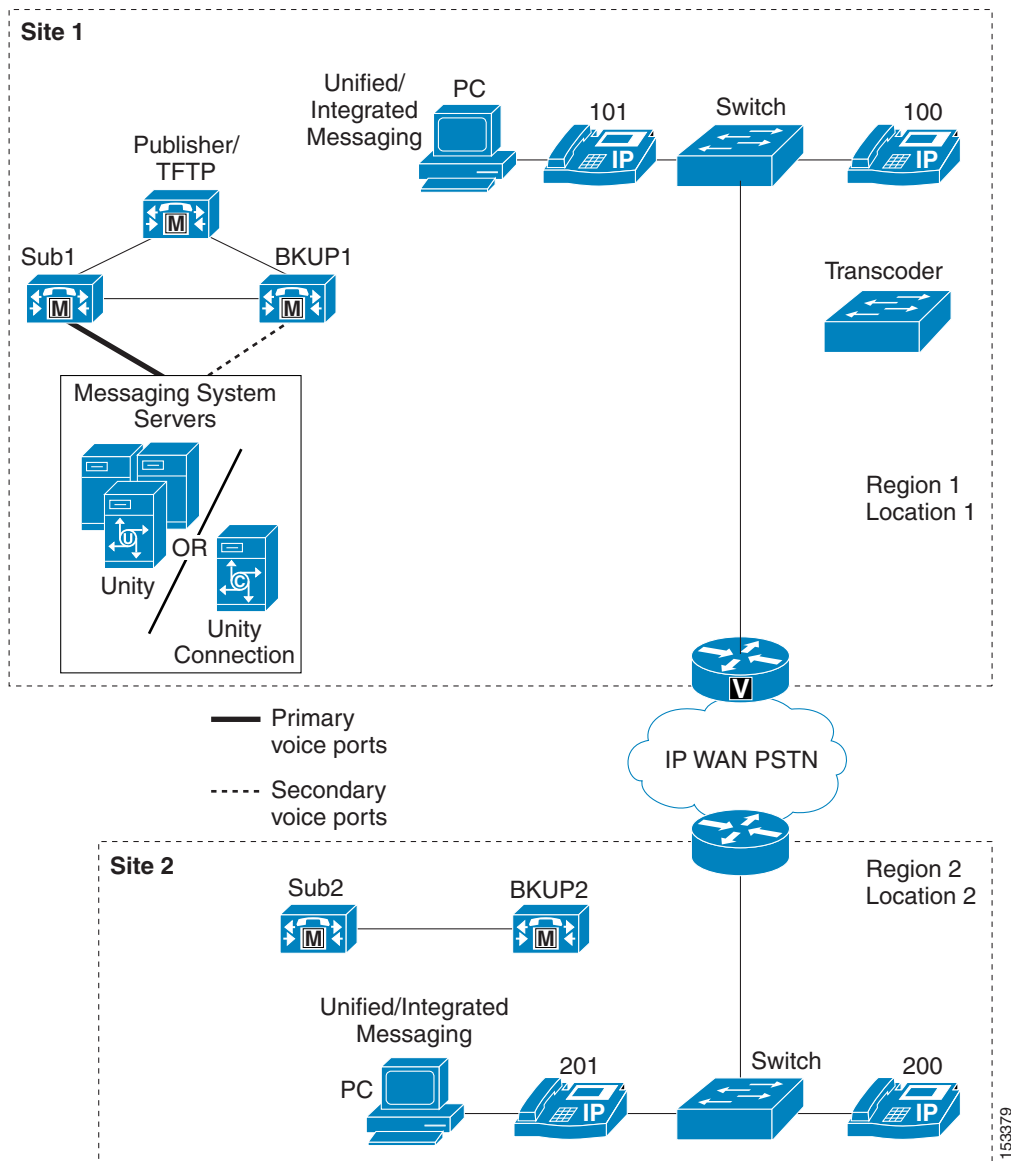

For minimum bandwidth requirements between clustered servers see the section on Local Failover Deployment Model, page 2-25.

Clustering over the WAN with Unified CM supports up to eight sites, as does Cisco Unity. The voicemail ports are configured only at the site where the Cisco Unity messaging system is located (see Figure 13-5). Voicemail ports do not register over the WAN to the remote site(s). Messaging clients at the other site(s) access all voicemail resources from the primary site. There is no benefit to configuring voice ports over the WAN to any of the remote sites because, in the event of a WAN failure, remote sites would lose access to the centralized messaging system. Because of bandwidth consideration, the voicemail ports should have TRaP disabled and all messaging clients should download voicemail messages to their local PCs (unified messaging only).

Distributed Messaging with Clustering Over the WAN

Local failover sites that also have Cisco Unity messaging server(s) deployed would have voice ports registered to the local Unified CM subscriber server(s), similar to the centralized messaging model. For information about configuring the voice ports, see Voice Port Integration with a Unified CM Cluster, page 13-38, and Voice Port Integration with Dedicated Unified CM Backup Servers, page 13-40.

Figure 13-6   Cisco Unity Distributed Messaging and Clustering over the WAN with Local Failover
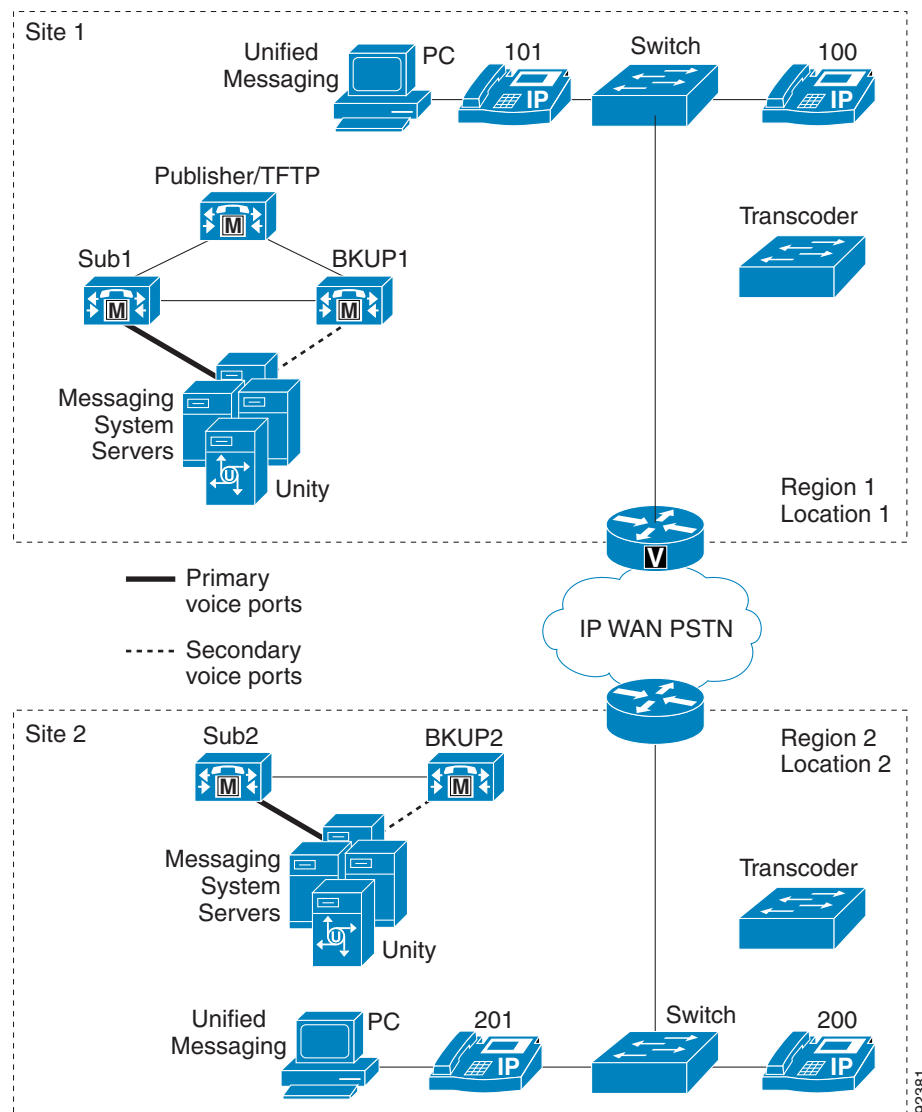
In a purely distributed messaging implementation with clustering over the WAN, each site in the cluster would have its own Cisco Unity messaging server with messaging infrastructure components. If not all of the sites have local Cisco Unity messaging systems but some sites have local messaging clients using a remote messaging server(s), this deployment would be a combination model with both distributed messaging and centralized messaging. (See Combined Messaging Deployment Models, page 13-12.) In the event of a WAN failure in this model, all remote sites that use centralized messaging will lose voicemail capability until the WAN is restored.

Each site that does not have a local messaging server must use a single messaging server for all of its messaging clients, but all such sites do not have to use the same messaging server. For example, suppose Site1 and Site2 each have a local messaging server. Site3 can then have all of its clients use (register with) the messaging server at Site2, while Site4 can have all of its clients use the messaging server at Site1. Transcoder resources are required at sites that have local Cisco Unity messaging server(s).

As with other distributed call processing deployments, calls going between these sites are transparent to gatekeeper call admission control, therefore you must configure regions and locations in Unified CM to provide call admission control. (See Managing Bandwidth, page 13-31.)

The distributed Cisco Unity servers may also be networked digitally. For more information on this topic, refer to the *Cisco Unity Networking Guide*, available at http://www.cisco.com. The networking guides are specific to the particular messaging store deployed.

Messaging Redundancy

Messaging redundancy is discussed in this section as it refers to Cisco Unity and Cisco Unity Connection. Cisco Unity Express does not support messaging redundancy.
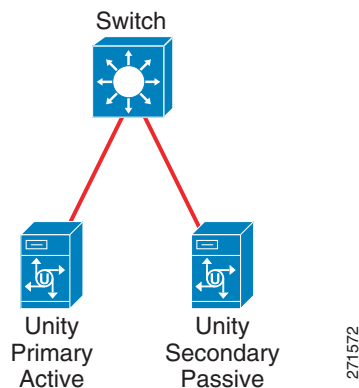
Cisco Unity

Cisco Unity supports two categories of redundancy. The first one is referred to simply as Unity Failover (local messaging failover), and it provides system malfunction failover. The second category is referred to as Standby Redundancy, and it provides disaster-recovery across geographic locations. See the *Cisco Unity Design Guide* for a comparison of Cisco Unity Failover and Standby Redundancy.

Cisco Unity failover consists of two servers, a primary and a secondary, configured as an active/passive redundant pair of servers, where the primary server is active and taking calls while the secondary is inactive and not taking calls. Any changes to subscriber or configuration data on the primary server are automatically replicated to the secondary server. If the primary server stops functioning for some reason, the secondary server automatically becomes the active server and starts taking calls while the primary server temporarily becomes inactive.

You can implement local messaging failover as illustrated in Figure 13-7. With local failover, both the primary and secondary Cisco Unity servers are located at the same site on the same highly available LAN.

*Figure 13-7   Local Failover of Cisco Unity Messaging*

Cisco Unity Standby Redundancy also provides disaster recovery across geographic locations. There are still two servers, a primary and a secondary, but they are installed in separate data centers, commonly in separate cities. If the data center in which the primary server is installed is affected by a natural disaster or other catastrophe, someone in (or with remote access to) the disaster-recovery facility manually activates the secondary server, and the secondary server begins taking calls. For further information on the requirements for Standby Redundancy or Failover (local messaging failover), refer to the *System Requirements for Cisco Unity Release 7.x*, available on http://www.cisco.com.

Cisco Unity Connection

Cisco Unity Connection supports messaging redundancy and load balancing in an active-active redundancy model consisting of two servers, a primary and a secondary, configured as an active/active redundant pair of servers, where both the primary and secondary servers actively accept calls as well as HTTP and IMAP requests. This functionality is new in Cisco Unity Connection release 7.0. For more information, refer to the *Design Guide for Cisco Unity Connection*, available at http://www.cisco.com.

Figure 13-8 illustrates Cisco Unity Connection active/active messaging redundancy.

*Figure 13-8   Redundancy of Cisco Unity Connection Messaging*
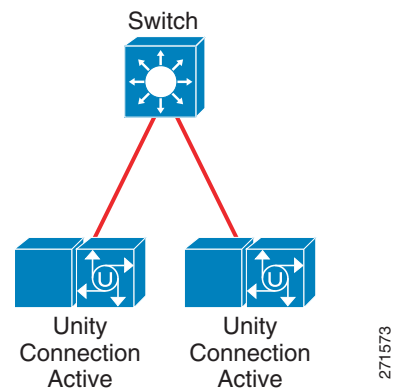
Both Cisco Unity and Cisco Unity Connection SIP trunk implementation requires call forking for messaging redundancy functionality. Currently, Unified CM does not support call forking on SIP trunks; therefore, Cisco Unity failover is not available when SIP trunks are used with Unified CM. However, when using SIP trunks with a Cisco Unity Connection server pair in active/active redundancy, Cisco recommends that you configure two separate SIP trunks, one to each server in the server pair, and add them to the same route group associated to the same route list. This configuration allows Unified CM to load-balance calls to the two servers.

Cisco Unity Failover and Clustering Over the WAN

When deploying Cisco Unity local failover with clustering over the WAN, apply the same design practices described in Centralized Messaging with Clustering Over the WAN, page 13-13, and Distributed Messaging with Clustering Over the WAN, page 13-15. The voice ports from the primary Cisco Unity server should not cross the WAN during normal operation.

Figure 13-9 depicts Cisco Unity local failover. Note that the primary and secondary Cisco Unity servers are both physically located at the same site. Cisco Unity failover supports up to the maximum number of remote sites available with clustering over the WAN for Unified CM.

*Figure 13-9*   *Cisco Unity Local Failover and Clustering Over the WAN*
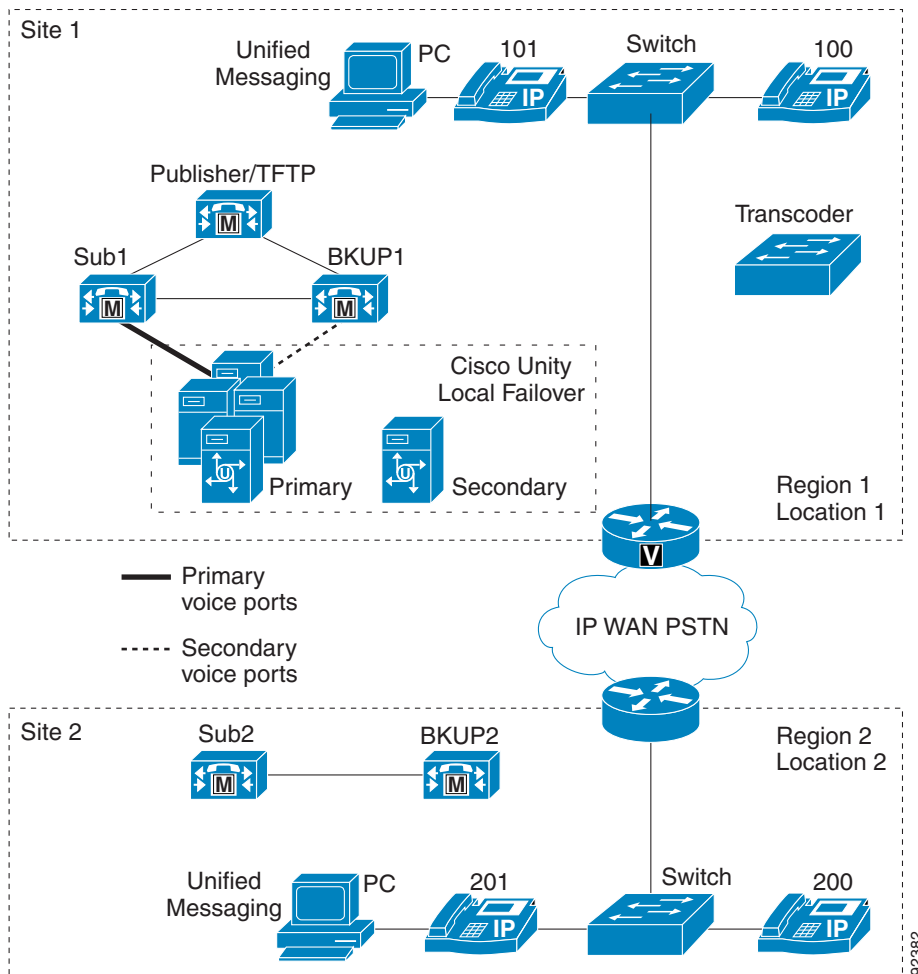
For information on configuring Cisco Unity failover, refer to the *Cisco Unity Failover Configuration and Administration Guide*, available at http://www.cisco.com.

As mentioned earlier, Cisco Unity failover can be configured to operate for high availability over a WAN; however, there are a number of requirements for deploying this. For example, if full redundancy at geographically separated data centers is important, there are certain requirements that must be met in order to ensure the successful operation of an installation in this configuration. These requirements are documented in the *System Requirements for Cisco Unity Release 7.x* available on http://www.cisco.com.

Centralized Messaging with Distributed Unified CM Clusters

Cisco Unity and Unity Connection can also be deployed in a centralized messaging configuration with multiple Unified CM clusters (see Figure 13-10). See the section on Telephony Integrations, page 13-35, for details on multiple integrations and MWI considerations with multiple Unified CM clusters.

Figure 13-10   Integrating Cisco Unity or Unity Connection with Multiple Unified CM Clusters
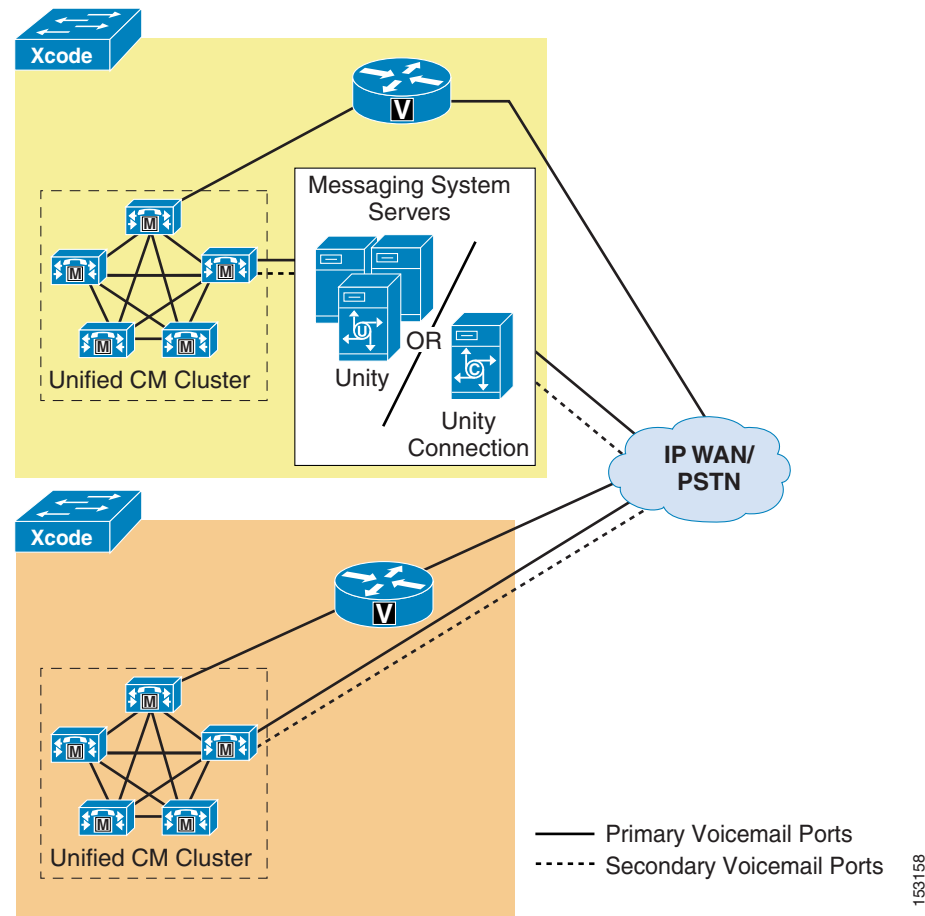
For the configuration in Figure 13-10, messaging clients at both Cluster 1 and Cluster 2 sites use the Cisco Unity or Unity Connection messaging infrastructure physically located at Cluster 1.

Cisco Unity Express Deployment Models

This section begins with a quick overview of Cisco Unity Express, covering product related information. Next the deployment models section presents three supported deployment models with Cisco Unity Express, focusing on distributed voice messaging with both centralized and distributed call processing followed by some deployment characteristics and design guidelines. Lastly, this section discusses the signaling call flows and the various protocols used between Cisco Unity Express and Unified CM as well as between Cisco Unity Express and Unified SRST or Unified CME in SRST mode.

Overview of Cisco Unity Express

Cisco Unity Express is delivered on a Cisco Unity Express Enhanced Network Module (NME-CUE) or Cisco Unity Express Advanced Integration Module (AIM-CUE), which directly integrate into the Cisco 2800 and 3800 Series Integrated Services Routers (ISRs) and Cisco Unified Communications 500 (UC 500) platforms. It also comes preloaded on factory-installed hardware on the Cisco 1861 ISR. In its various configurations, Cisco Unity Express supports from 8 to 250 users with up to 24 ports of concurrent voicemail and integrated messaging, automated attendant sessions, or optional IVR services. Cisco Unity Express 3.0 and later releases support the NME-CUE module and 24 ports.

Note   The Cisco Unity Express NM-CUE and NM-CUE-EC modules are End-of-Sale and End-of-Life and are replaced by the Cisco Unity Express Enhanced Network Module (NME-CUE). For more information, refer to the End-of-Sale and End-of-Life Announcement at http://www.cisco.com/en/US/prod/collateral/voicesw/ps6789/ps5745/ps5520/end_of_life_notice_c51-453045.html.

Note   Cisco Unified Communications 500 Series for Small Businesses supports Cisco Unified Communications Manager Express (Unified CME) and Cisco Unity Express, but it does not support SRST for Cisco Unified CM. Cisco Unified Communications 500 Series for Small Businesses supports a maximum of up to 5 sites.

Cisco Unity Express can be deployed as a voicemail solution with Cisco Unified Communications Manager, Unified SRST, or Cisco Unified Communications Manager Express (Unified CME). There is no method to convert or back-up and restore Cisco Unity Express from a Unified CME integration to a Unified CM integration, or vice versa. The Cisco Unity Express module must be re-imaged to work with Unified CME from a previous Unified CM installation, or vice versa. This means that you must re-apply the software and license, and all the configuration and data, which includes voicemail messages that are lost. In prior releases, Cisco Unity Express was limited to a co-resident deployment with Unified CME or a Survivable Remote Site Telephony (SRST) router. However, Cisco Unity Express and SRST or Unified CME can reside on two separate routers when deployed with Unified CM or Unified CME, respectively, with the H.323-to-SIP call routing capability introduced in Cisco IOS Release 12.3(11)T.

While Cisco Unity Express uses SIP to communicate with Cisco Unified Communications Manager Express (Unified CME), Cisco Unity Express uses JTAPI to connect to Cisco Unified Communications Manager (Unified CM). When connected with Unified CM, Cisco Unity Express auto-detects the version of Unified CM that is running and compares that with its own JTAPI library version. If Cisco Unity Express detects that the Unified CM release has changed, Cisco Unity Express automatically reboots and reconfigures itself with the correct JTAPI library for the version of Unified CM to which it is connected.

For details on interoperability of Unified CM and Unified CME, see Interoperability of Unified CM and Unified CM Express, page 8-36.

For additional information on supported deployment models with Unified CME, refer to the appropriate Cisco Unified Communications Manager Express design documentation available at http://www.cisco.com.

Deployment Models

Cisco Unity Express can be deployed as a single site or distributed voicemail and automated attendant (AA) solution for Cisco Unified Communications Manager (Unified CM) or Unified Communications Manager Express (Unified CME). However, Cisco Unity Express is supported with all of the Cisco Unified CM deployment models, including:

- Single-site deployments
- Multisite WAN deployments with centralized call processing
- Multisite WAN deployments with distributed call processing

Figure 13-11 shows a centralized call processing deployment incorporating Cisco Unity Express, and Figure 13-12 shows a distributed call processing deployment.

Cisco Unity Express sites controlled by Unified CME, as well as other sites controlled by Unified CM, can be interconnected with each other using H.323 or SIP trunking protocol. Although Cisco Unity Express can integrate with either Unified CM or Unified CME, it cannot integrate with both simultaneously.

Note Cisco Unity Express 3.2 supports a centralized deployment model with up to 10 Unified CMEs.

Figure 13-11   Cisco Unity Express in a Centralized Call Processing Deployment
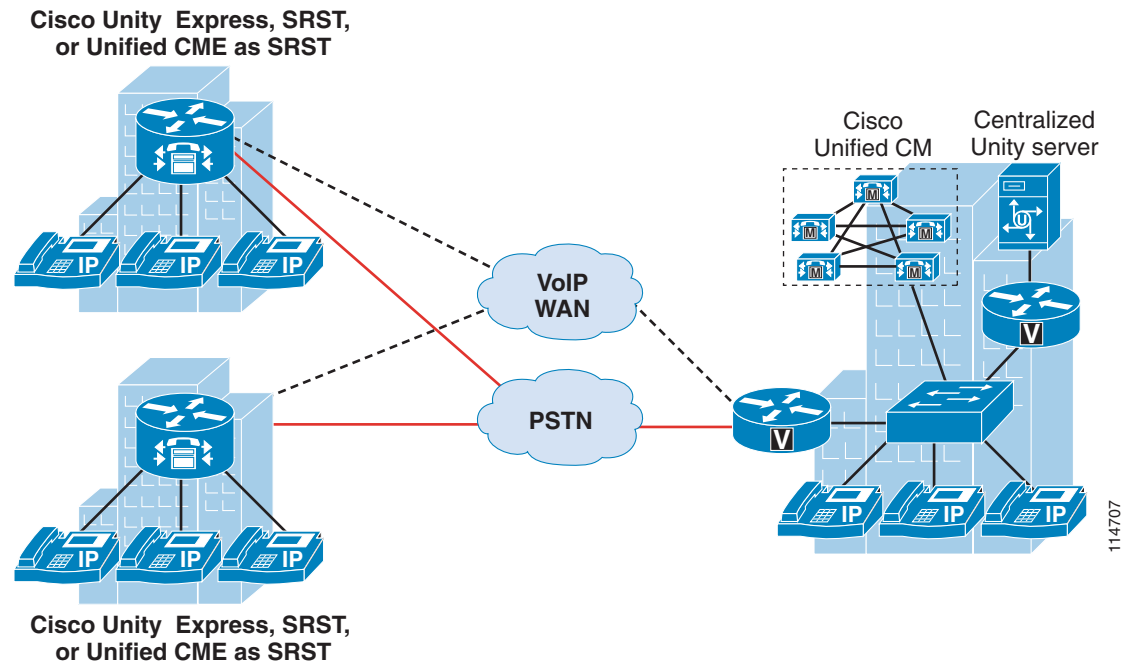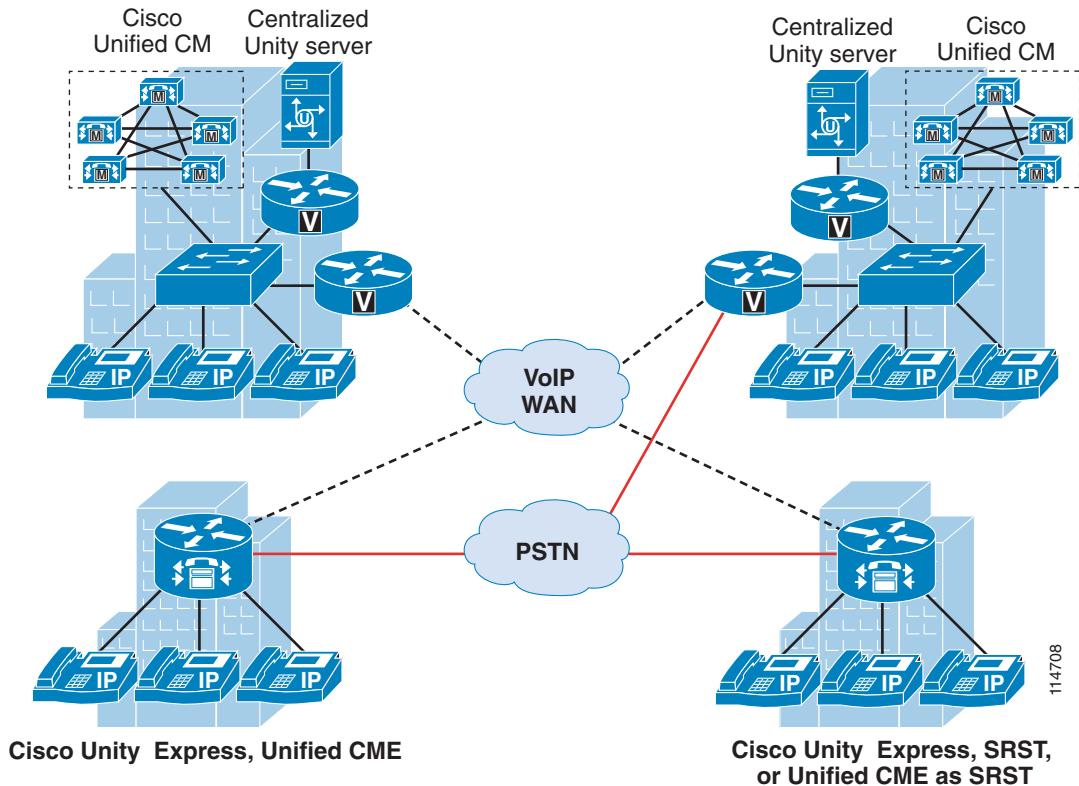
Figure 13-12   Cisco Unity Express in a Distributed Call Processing Deployment

Cisco Unity Express, Unified CME

**Cisco Unity Express, SRST,
or Unified CME as SRST**

The most likely deployment model to use Cisco Unity Express is the multisite WAN model with centralized call processing, where Cisco Unity Express provides distributed voicemail at the smaller remote offices and a central Cisco Unity system provides voicemail to the main campus and larger remote sites.

Use Cisco Unity Express as a distributed voicemail solution if any of the following conditions apply to your Unified CM network deployment:

- Survivability of voicemail and AA access must be ensured regardless of WAN availability.

- Available WAN bandwidth is insufficient to support voicemail calls traversing the WAN to a central voicemail server.

- There is limited geographic coverage of the AA or branch site PSTN phone numbers published to the local community, and these numbers cannot be dialed to reach a central AA server without incurring toll charges.

- The likelihood is high that a PSTN call into a branch office will be transferred from the branch AA to a local extension in the same office.

- Management philosophy allows remote locations to select their own voicemail and AA technology.

The following characteristics and guidelines apply to Cisco Unity Express in either a centralized or distributed Unified CM deployment:

- A single Cisco Unity Express can be integrated with a single Unified CM cluster.

- Cisco Unity Express integrates with Unified CM using a JTAPI application and Computer Telephony Integration (CTI) Quick Buffer Encoding (QBE) protocol. CTI ports and CTI route points control the Cisco Unity Express voicemail and automated attendant (AA) applications.

- Cisco Unity Express provides voicemail functionality to Cisco Unified IP Phones running Skinny Client Control Protocol (SCCP). Cisco Unity Express 2.3 and later releases also provide support for Session Initiation Protocol (SIP) IP phones with Unified CM 5.*x* and later releases.

- The following CTI route points are defined on Unified CM for Cisco Unity Express:

    - Automated attendant entry point (Cisco Unity Express can contain up to five distinct AAs and may therefore require up to five different route points.)

    - Voicemail pilot number

    - Greeting management system (GMS) pilot number (Optional; if the GMS is not used, then this route point need not be defined.)

- The following CTI ports are defined on Unified CM for Cisco Unity Express:

    - 12 or 25-mailbox system (4 ports)

    - 50-mailbox AIM-CUE system (6 ports)

    - 250-mailbox NME-CUE system (24 ports)

- Each Cisco Unity Express site has 250 mailboxes or fewer. For deployments that require more than 250 mailboxes, consider using Cisco Unity or other voicemail solutions.

- Each Cisco Unity Express mailbox can be associated with a maximum of two different extensions, if needed.

- The automated attendant function for any office deployed with Cisco Unity Express can be local to the office (using the AA application in Cisco Unity Express) or centralized (using Cisco Unity Express for voicemail only).

- Cisco Unity Express can be networked with other Cisco Unity Expresses or with Cisco Unity via Voice Profile for Internet Mail (VPIM) version 2. Thus, a Cisco Unity Express subscriber can send, receive, or forward messages to or from another remote Cisco Unity Express or Cisco Unity subscriber.

- Cisco Unity Express allows you to specify up to three Unified CMs for failover. If IP connectivity to all three Unified CMs is lost, Cisco Unity Express switches to Survivable Remote Site Telephony (SRST) call signaling, thus providing AA call answering service as well as mailbox access to IP phones and PSTN calls coming into the branch office.

- Cisco Unity Express automated attendant supports dial-by-extension and dial-by-name functions. The dial-by-extension operation enables a caller to transfer a call to any user endpoint in the network. The dial-by-name operation uses the directory database internal to Cisco Unity Express and does not interact with external LDAP or Active Directory databases.

- Centralized Cisco Unity Express with Unified CM is not supported.

- Cisco Unity Express is not supported in pure SIP networks that do not have either Cisco Unified CM or Unified CME controlling the SIP phones.

- Cisco Unity Express can be deployed on a separate Unified CME or SRST router or a separate PSTN gateway.

- When Cisco Unity Express is deployed on a router separate from Unified CME or SRST, configure the command **allow-connections h323 to sip** for H.323-to-SIP routing.

Figure 13-13 shows the protocols involved in the call flow between Unified CM and Cisco Unity Express.

Figure 13-13   *Protocols Used Between Cisco Unity Express and Unified CM*
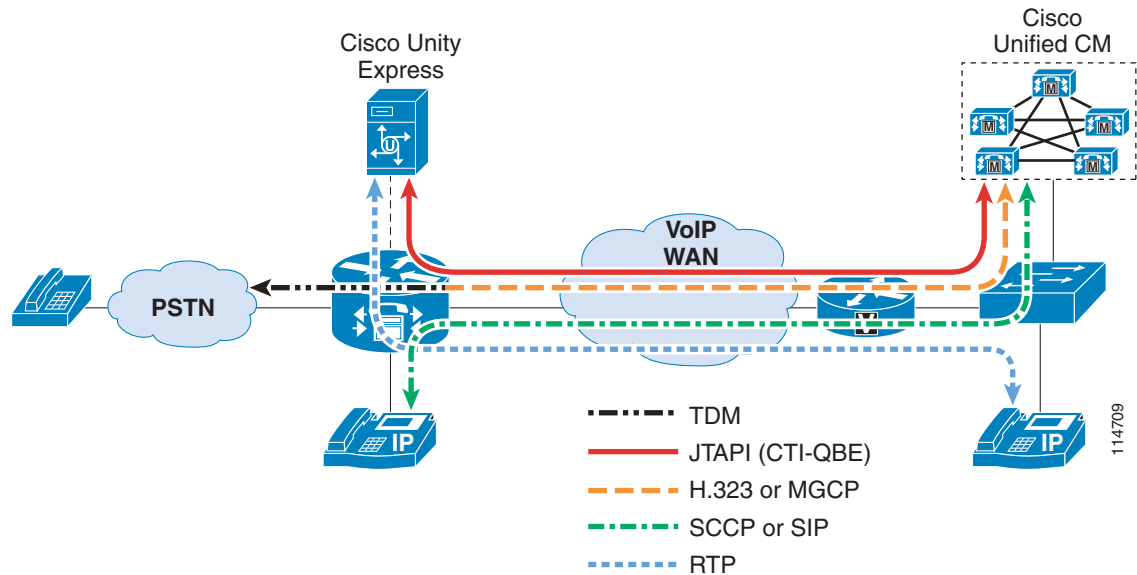

Figure 13-13 illustrates the following signaling and media flows:

- Phones are controlled via SCCP or SIP from Unified CM.

- Cisco Unity Express is controlled via JTAPI (CTI-QBE) from Unified CM.

- The Message Waiting Indicator (MWI) on the phone is affected by Cisco Unity Express communicating a change of mailbox content to Unified CM via CTI-QBE, and by Unified CM in turn sending a MWI message to the phone to change the state of the lamp.

- The voice gateway communicates via H.323, SIP, or MGCP to Unified CM.

- Real-Time Transport Protocol (RTP) stream flows carry the voice traffic between endpoints.

Figure 13-14 shows the protocols involved in the call flow between the router for SRST or Unified CME in SRST mode and Cisco Unity Express when the WAN link is down.

Figure 13-14   Protocols Used Between Cisco Unity Express and the Router for SRST or Unified CME in SRST Mode
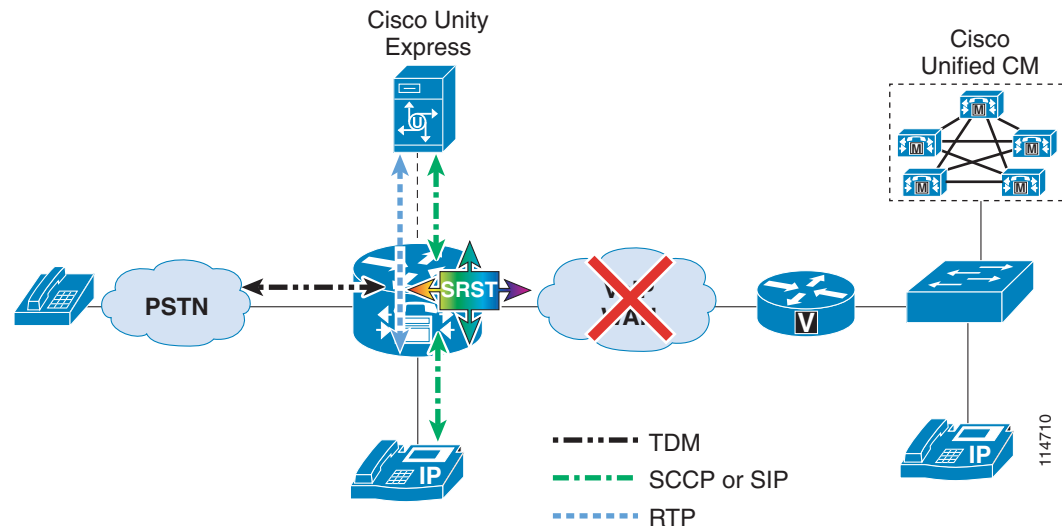

Figure 13-14 illustrates the following signaling and media flows:

- Phones are controlled via SCCP or SIP from the router for SRST or Unified CME in SRST mode.

- Cisco Unity Express communicates with the SRST router via an internal SIP interface.

- Although MWI changes are not supported in SRST mode with previous releases of Cisco Unity Express, voice messages can be sent and retrieved as during normal operation, but the MWI lamp state on the phone remains unchanged until the phone registers again with Unified CM. At that time, all MWI lamp states are automatically resynchronized with the current state of the users' Cisco Unity Express voicemail boxes. Cisco Unity Express 3.0 and later releases support MWI for SRST mode.

- Cisco Unity Express 2.3 and later releases now support SIP Subscriber/Notify and Unsolicited Notify to generate MWI notifications, in both Unified CME and SRST modes.

- RTP stream flows carry the voice traffic between endpoints.

- SRST subscribes to Cisco Unity Express for MWI for each of the ephone-dns registered to receive MWI notifications.

Note   Unified CM MWI (JTAPI) is independent of the SIP MWI methods.

Fax Deployment

Fax deployment is an important aspect of any Unified Communications implementation. This section is not intended to cover all aspects of fax deployment and integration but should serve as a starting point. For fax deployments with Cisco Unity and Unity Connection, references are provided to other documents where fax integration and deployment for these products are covered in detail. The main focus of this section is on fax deployment for Cisco Unity Express.

Cisco Unity and Unity Connection Fax Deployment

Cisco Unity 7.x supports a large number of fax integrations. Refer to the *Cisco Unity Install and Upgrade Guide*, available on http://www.cisco.com, for details on supported fax integrations.

Cisco Unity Connection 7.x supports Cisco Fax Server. For details, refer to the *Design Guide for Cisco Unity Connection*, available on http://www.cisco.com.

Cisco Unity Express Fax Deployment

Cisco Unity Express 3.0 and later releases support both inbound and outbound faxing. Outbound faxing enables faxes to be printed to the fax machine. Cisco Unity Express supports both a single DID number and separate DID numbers for voice and fax calls. The single DID approach allows each user to have one number for both voicemail and fax, while separate DID numbers provide each user with two numbers, one for fax and one for their voice extension. In both cases, the same mailbox is used for storing both voice and fax messages.

The fax support relies on T.37 fax support from the Cisco IOS Gateway. The T.37 store-and-forward Cisco IOS fax gateway uses the T.37 store-and-forward fax application, which consists of on-ramp and off-ramp processes. Cisco Unity Express relies on the fax detection application to support single DID functionality but requires that the on-ramp application be configured on the fax gateway when separate DIDs are used for fax and voice calls. The Cisco fax gateway, acting as the on-ramp gateway, receives faxes from end users, converts them into TIFF files, creates standard Multipurpose Internet Mail Extension (MIME) email messages with attached TIFF files, and forwards the fax mail messages to the designated SMTP server (which is Cisco Unity Express) for storage. The off-ramp gateway handles calls going out from the network to a fax machine or the PSTN by dialing the POTS to communicate with a remote fax machine (Group 3 fax device).

Cisco IOS fax gateways support the following fax capabilities:

- Fax pass-through and fax pass-through with upspeed
- Cisco Fax Relay
- T.38 Fax Relay
- T.37 Store-and-Forward Fax
- IVR applications for fax

On-ramp and off-ramp fax processes can reside on the same gateway or separate gateways.

Note   Cisco Unity Express supports only Cisco IOS fax gateways. Third-party fax gateways are not supported.

For more details on Cisco IOS fax gateway capabilities, refer to the related product documentation available at http://www.cisco.com.

Cisco Unity Express supports the following fax deployment scenarios:

- Cisco Unity Express, on-ramp, and off-ramp applications all reside on the same voice gateway.
- Cisco Unity Express and the on-ramp application reside on the same voice gateway, but the off-ramp application runs on a separate voice gateway.
- Cisco Unity Express and the off-ramp application reside on the same voice gateway, but the on-ramp application runs on a separate voice gateway.
- Cisco Unity Express, the on-ramp application, and the off-ramp application each runs on a separate voice gateway.
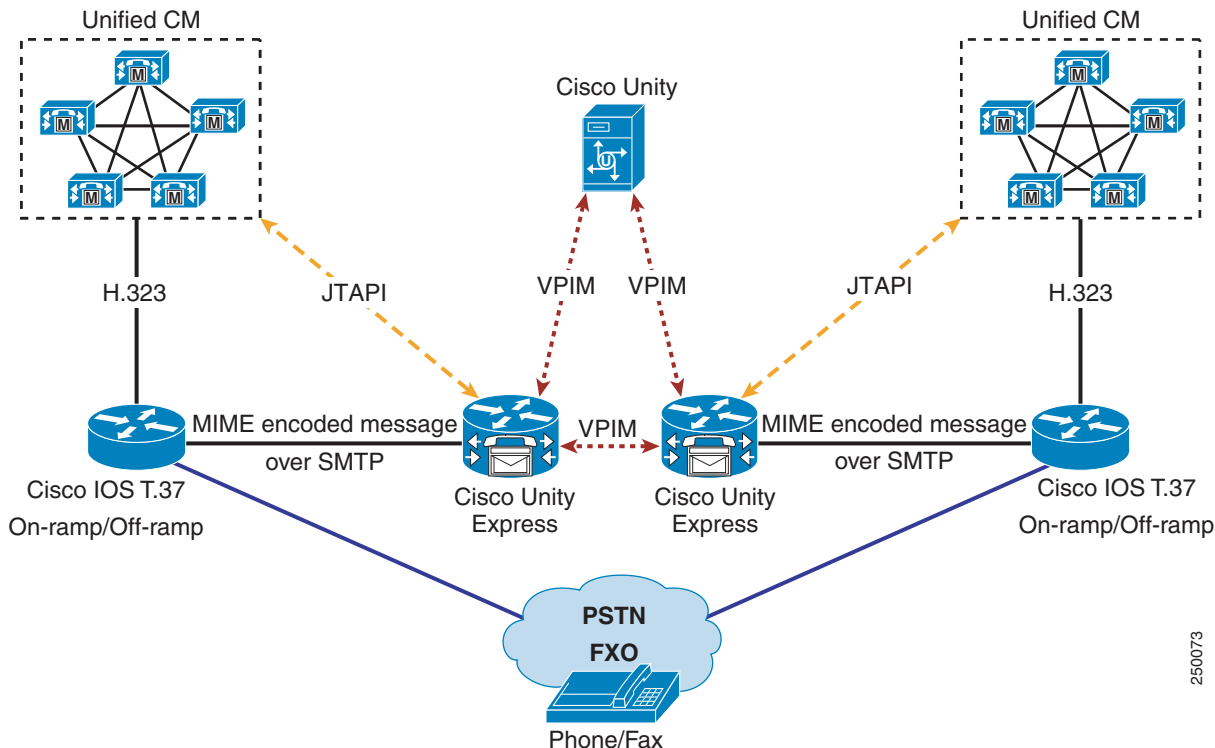
> **Note**   Two or more instances of Cisco Unity Express cannot integrate with the same gateway for inbound fax calls.

Figure 13-15 shows a sample fax deployment scenario for Cisco Unity Express with Unified CM.

Figure 13-15   Fax Deployment Scenario for Cisco Unity Express with Unified CM

If the call is a fax call, whether Cisco Unity Express is integrated with Unified CME or Unified CM, the fax application configured on the outbound Multimedia Mail over IP (MMoIP) dial peer converts the fax into an email message with TIFF attachment(s) and sends the message over SMTP to Cisco Unity Express.

Cisco Unity Express uses the fax detection application to support single DID functionality. The fax detection application has a limitation that causes the fax call to get disconnected and requires the fax to be resent if either of the following cases occurs:

- The fax call is picked up and disconnected before the application detects that it is the fax call.

- After picking up the call and hearing CNG (fax) tones, the user tries to transfer the call to the fax (MMoIP) dial peer.

Cisco recommends that you configure the on-ramp application on the fax gateway when you use separate DIDs for fax and voice calls.

The following characteristics and guidelines apply when deploying fax capability with Cisco Unity Express and Unified CM:

- Cisco Unity Express supports only analog fax machines to fax services.

- The administrator should enable fax on the mailboxes requiring the fax support.

- The fax detection application required to support single DID usage for fax and voicemail must run on the originating gateway.

- Cisco Unity Express sends VPIM-encoded voice messages or faxes over SMTP to another Cisco Unity Express or Cisco Unity server in the network.

- Regardless of whether Cisco Unity Express is integrated with Unified CME or Unified CM, Cisco Unity Express can interoperate with Cisco Unity to receive, send, and forward fax messages as well as voice messages from Cisco Unity, or vice versa, via VPIM.

- A fax message can be printed using the fax number configured at the system level. This number is displayed when the user is trying to print the fax using the Telephony User Interface (TUI) or Voice View Express (VVE), and it is changeable to allow users to print the fax messages at an alternative fax machine.

- Cisco Unity Express does not support broadcast fax messages.

- Cisco Unity Express sends a Delayed Delivery Record (DDR) or Non-Delivery Receipt (NDR) when fax message delivery is delayed or fails, respectively.

- Cisco Unity Express supports the same number of total TUI sessions as the total number of fax sessions.

- Cisco Unity Express integrates with one Cisco fax gateway for outbound fax and one for inbound fax calls. Inbound and outbound calls can use the same gateway or different gateways. Two or more instances of Cisco Unity Express cannot integrate with the same fax gateway for inbound fax calls.

- Assign a unique fax DID number to a user when separate DIDs are used for fax and voice.

- Cisco Unity Express supports mailbox login using the voice number but not the fax DID number.

- Cisco Unity Express supports the fax feature when Cisco Unity Express is in SRST mode. When in SRST mode, Cisco Unity Express fax functionality works between the users within the same SRST site without needing to go across the WAN.

Voicemail Networking

This section covers specific considerations for voicemail networking in Cisco Unity Express. Also included is a high-level overview of voicemail networking with the Cisco Unified Messaging Gateway. For information specific to voicemail networking in Cisco Unity and Unity Connection, refer to the *Design Guide for Cisco Unity* or the *Design Guide for Cisco Unity Connection*, respectively, available at http://www.cisco.com.

Voicemail networking is the ability to allow subscribers (voicemail users) to send, receive, reply to, and forward voicemail messages between systems such as Cisco Unity, Cisco Unity Connection, and Cisco Unity Express using an embedded Simple Mail Transfer Protocol (SMTP) server and a subset of the Voice Profile for Internet Mail (VPIM) version 2 protocol. All three voicemail messaging products support interoperability between one another using VPIM messaging.

Cisco Unity Express Voicemail Networking

Cisco Unity Express communicates with Cisco Unity 4.0.3 and later releases as well as Cisco Unity Connection 1.2 and 2.0. Cisco Unity Express uses SMTP (RFC 2821) for message delivery between remote sites. Cisco Unity Express voicemail networking provides the following capabilities:

- Subscribers can receive, send, and forward messages to or from another remote Cisco Unity Express or Cisco Unity for locations configured on the originating system.

- Subscribers can also reply to a remote message received from a remote system.

- Subscribers can be recipients of a distribution list or individual message originating from Cisco Unity.

The following characteristics and design considerations apply to Cisco Unity Express voicemail networking:

- Cisco Unity Express voicemail networking uses blind addressing with known location ID and user extension.

- Cisco Unity Express voicemail networking is able to define remote locations by means of a five-letter location abbreviation (used for confirmation of addressing only).

- Message encoding can be dynamic, and SMTP session negotiation determines either .wav or G.726 encoding format.

- Cisco Unity Express voicemail networking is able to exchange G.711 .wav file and G.726 Message Encoding.

- Use G.726 for sites with low bandwidth. G.726 forces 32 kbps Adaptive Differential Pulse Code Modulation (ADPCM) format regardless of SMTP negotiation.

- The sender's spoken name may or may not be contained in the VPIM message. If the spoken name is not included in a message, then the name confirmation will state only the extension and location of the sender when a VPIM message is played. Cisco Unity Express includes the sender's spoken name in the VPIM message if **send spoken name** is configured for a particular remote location.

- Cisco Unity Express voicemail networking notifies the sender if the message has not been sent for an hour. It also sends a Non-Delivery Receipt (NDR) if it fails to deliver for 6 hours or if the recipient's mailbox is full or non-existent.

- Cisco Unity Express voicemail networking uses a built-in SMTP server, which accepts only incoming sessions from locations defined in the remote location tables.

- Each Cisco Unity Express must have its own email domain or sub-domain name so that messages can eventually be routed to it.

- For interworking with Cisco Unity, configure Cisco Unity Express locations with a domain or sub-domain name.

- Cisco Unity Express tries every 15 minutes for 6 hours to deliver a message to a recipient. (Neither value is configurable.)

- Each Cisco Unity Express allows only 2 simultaneous SMTP send and receive sessions.

- Cisco Unity Express supports a total of 500 remote sites.

Voicemail Networking with Cisco Unified Messaging Gateway

The Cisco Unified Messaging Gateway is Linux-based software running on a Cisco Network Module (NME-UMG or NME-UMG-EC) in Cisco Integrated Services Routers (ISRs). The Unified Messaging Gateway can act as the central hub for Cisco Unity, Cisco Unity Connection, and Cisco Unity Express to network VPIM v2 voicemail systems in a hub-and-spoke or hierarchical structure. This approach dramatically reduces the VPIM connections between voicemail systems and simplifies the configuration effort on each system. Each voicemail system (Cisco Unity, Cisco Unity Connection, Cisco Unity Express, or Avaya Interchange) needs to configure only the connection between itself and the Cisco Unified Messaging Gateway. The Unified Messaging Gateway then handles message routing and delivery between the systems. This end-to-end message networking functionality is required by medium to larger distributed enterprises in order to migrate to the Cisco Unified Communications solution.

The Cisco Unified Messaging Gateway provides the following benefits:

- It enables intelligent routing for multiple autonomous voicemail networks with VPIM.
- It provides scalable voicemail networks and interoperability with third-party voicemail systems (such as Avaya Interchange) over VPIM networks.
- It simplifies voicemail VPIM network additions and expansion.

The Cisco Unified Messaging Gateway running on a Cisco Network Module NME-UMG can support up to 250 nodes and 12,500 subscribes, while the Unified Messaging Gateway running on an NME-UMG-EC module can support up to 1000 nodes and 50,000 subscribers. The number of subscribers is calculated based on 50 subscribers on any single Cisco Unity Express node that is registered with the Unified Messaging Gateway. The capacity of the Unified Messaging Gateway is tied to both the maximum number of nodes supported and the maximum number of subscribers supported, and increasing one quantity will cause a decrease in the other quantity. For example, if the network has Cisco Unity and/or Avaya endpoints with a large number of subscribers, the number of nodes that can register with the Unified Messaging Gateway will be significantly less than 250 for the NME-UMG or 1000 for the NME-UMG-EC.

Observe the following guidelines when deploying the Cisco Unified Messaging Gateway:

- Plan in advance for possible upgrades to the network if you anticipate more than 250 nodes because the NME-UMG module cannot be upgraded to the NME-UMG-EC module.
- Cisco Unity Express 3.1 and later releases register automatically and exchange directory information with the Cisco Unified Messaging Gateway, whereas Cisco Unity, Cisco Unity Connection, and Avaya Interchange must be provisioned manually on the Unified Messaging Gateway.
- Deploy a pair of Unified Messaging Gateways (primary and backup) for redundancy.
- Deploy up to 20 Unified Messaging Gateways (10 primary and 10 backup) for large deployments of up to 10,000 nodes.

Note   Voicemail networking with the Cisco Unified Messaging Gateway does not apply to deployments of Cisco Unified Communications 500 Series for Small Businesses because the Cisco Unified Communications 500 Series supports a maximum of only 5 sites in a distributed environment.

For information on the Cisco Unified Communications 500 Series for Small Business deployments, refer to the product documentation on http://www.cisco.com.

For a complete discussion of VPIM as a distributed messaging solution and for design guidelines on the Cisco Unified Messaging Gateway, refer to the Cisco Unified Messaging Gateway product documentation available on http://www.cisco.com.

Best Practices for Voice Messaging

This section discusses some general best practices and guidelines that were not mentioned previously yet are important aspects of the products and should be considered in the solution. They are separated into two groupings, Cisco Unity and Cisco Unity Connection in one grouping and Cisco Unity Express in another.

Best Practices for Deploying Cisco Unity and Cisco Unity Connection with Unified CM

This section applies to Cisco Unity and Unity Connection. For Cisco Unity Express, see Best Practices for Deploying Cisco Unity Express, page 13-41.

Managing Bandwidth

Unified CM provides a variety of features for managing bandwidth. Through the use of regions, locations, and even gatekeepers, Unified CM can ensure that the number of voice calls going over a WAN link does not oversubscribe the existing bandwidth and cause poor voice quality. Cisco Unity and Unity Connection rely on Unified CM to manage bandwidth and to route calls. If you deploy Cisco Unity or Unity Connection in an environment where calls or voice ports might cross WAN links, these calls will be transparent to gatekeeper-based call admission control. This situation occurs any time the Cisco Unity or Unity Connection server is servicing either distributed clients (distributed messaging or distributed call processing) or when Unified CM is remotely located (distributed messaging or centralized call processing). Unified CM provides regions and locations for call admission control.

Figure 13-16 uses a small centralized messaging and centralized call processing site to illustrate how regions and locations work together to manage available bandwidth. For a more detailed discussion of regions and locations, refer to the chapter on Call Admission Control, page 9-1.

Figure 13-16   Locations and Regions
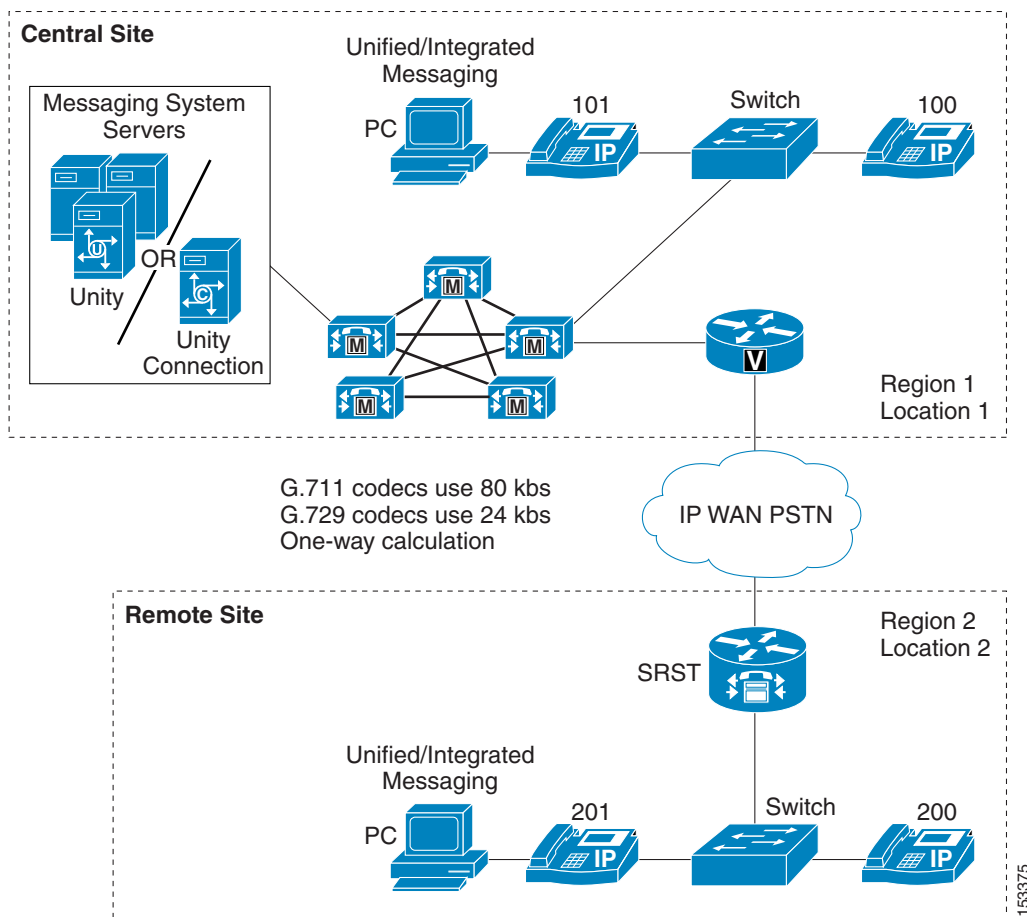
In Figure 13-16, regions 1 and 2 are configured to use G.711 for intra-region calls and G.729 for inter-region calls. Locations 1 and 2 are both set to 24 kbps. Location bandwidth is budgeted only in the case of inter-location calls.

An intra-region (G.711) call would not be budgeted against the available bandwidth for the location. For example, when extension 100 calls extension 101, this call is not budgeted against the 24 kbps of available bandwidth for Location 1. However, an inter-region call using G.729 is budgeted against both bandwidth allocations of 24 kbps for Location 1 and Location 2. For example, when extension 100 calls extension 200, this call would be connected but any additional (simultaneous) inter-region calls would receive reorder (busy) tone.

Native Transcoding Operation

In Cisco Unity and Unity Connection, native transcoding occurs when a call is negotiated between an IP endpoint and the Cisco Unity or Unity Connection server in one codec and is recorded or played out in another codec format. If a call is negotiated in G.729 and the system-wide recording format is done in G.711, then the server has to transcode that call natively. Cisco Unity and Unity Connection native transcoding does not use external hardware transcoders but instead uses the server's main CPU. This is what is meant by native transcoding.

Cisco Unity Operation

By default, Cisco Unity servers support G.711 and G.729 in the telephony integration via Skinny Client Control Protocol (SCCP) and/or SIP. Cisco Unity also has a default system-wide message recording format set to G.711. In order to disable native transcoding, Cisco recommends configuring the system recording format and the SCCP or SIP integration codec advertisement to coincide. For example, if you have a Cisco Unity implementation with Unified CM using SCCP Ports or a SIP trunk, you would configure the ports or trunk to advertise G.711 only by removing G.729 as one of the advertised codecs. Also, by leaving the default system-wide recording format set to G.711, any call that was negotiated with this system would effectively be in G.711 and the recording would also occur in that format, thus removing any need to transcode natively on the messaging server.

Disabling Native Transcoding in Cisco Unity

For SCCP integrations only in Cisco Unity, the native transcoding must be disabled (turned off) on the Cisco Unity server via a registry setting in order for Unified CM to assign hardware transcoders to voice port calls. The registry setting is called **Audio - Enable G.729a codec support**, and the tool for setting it is the Advanced Settings Tool available at http://www.CiscoUnityTools.com. (For information on disabling native transcoding on Cisco Unity when integrated via SIP, refer to the *Cisco Unified Communications Manager SIP Trunk Integration Guide for Cisco Unity* for your specific release of Cisco Unity, available at http://www.cisco.com.)

By default, there is no codec registry key, therefore native transcoding is enabled (on). The Advanced Settings Tool adds a new registry key that enables you to select either one of the two available codecs. Cisco Unity will then "advertise" to Unified CM that it supports only one codec. If a call terminating or originating on the voice ports is using a different codec than the type configured for the Cisco Unity server, Unified CM will assign an external transcoding resource for the call. For detailed information on how to configure transcoding resources on Unified CM, refer to the chapter on Media Resources, page 6-1.

✎
Note   Currently, the Cisco Unity TAPI Service Provider (TSP) for Unified CM supports only G.729 and G.711 mu-law (a-law is not supported) for the Skinny Client Control Protocol (SCCP) voice ports. A software Media Termination Point (MTP), such as Unified CM itself or an Integrated Services Router (ISR), is required for mu-law to a-law conversion.

When the Advanced Settings Tool is used to advertise only one codec, the Cisco Unity server system prompts must be the same as the codec used. By default, the system prompts are G.711. If the codec is set to G.711, the system prompts will work correctly. However, if G.729 is selected, you will have to change the system prompts. Refer to the *Cisco Unity System Administration Guide* on http://www.cisco.com for details on how to change the system prompts. If the system prompts are not the same as the codec selected by the registry, then the end users will hear unintelligible system prompts.

Cisco Unity Connection Operation

Cisco Unity Connection handles transcoding operations differently than Cisco Unity. In Cisco Unity Connection, a call in any codec format supported by Cisco Unity Connection SCCP or SIP signaling (G.711 mu-law, G.711 a-law, G.729, iLBC, and G.722) will always be transcoded to Linear PCM. From Linear PCM, the recording is encoded in the system-level recording format (Linear PCM, G.711 mu-law/a-law, G.729a, or G.726), which is set system-wide in the general configuration settings (G.711 mu-law is the default). As such the overall impact of transcoding in Cisco Unity Connection is quite

different from Cisco Unity. In the rest of this chapter, we refer to the codec negotiated between the calling device and Unity Connection as the "line codec," and we refer to the codec set in the system-level recording format as the "recording codec."

Because transcoding is inherent in every connection, there is little difference in system impact when the line codec differs from the recording codec. The exception to this is when using iLBC or G.722. G.722 and iLBC require more computation to transcode, therefore they have a higher system impact. G.722 and iLBC use approximately twice the amount of resources as G.711 mu-law. The subsequent impact this has is that a system can support only half as many G.722 or iLBC connections as it can G.711 mu-law connections.

As a general rule, Cisco recommends leaving the default codec as G.711. If the configuration is constrained by disk space, then a lower bit rate codec such as G.729a or G.726 can be configured as the recording format; however, keep in mind that the audio quality will not have the fidelity of G.711 audio. Also, if G.722 is used by devices on the line, then linear pulse code modulation (PCM) is an option to improve the audio quality of the recording. This will, however, increase the disk usage and impact disk space.

There are also a few reasons to change the recording codec or to choose to advertise only specific line codecs. Consider the following factors when deciding on the system-level recording format and the advertised codecs on the SCCP or SIP integration:

- Which codecs will be negotiated between the majority of the endpoints and Cisco Unity Connection? This will help you decide on which codecs need to be advertised by Cisco Unity Connection and which do not. You can then decide on when you need Unified CM to provide hardware transcoding resources in lieu of doing computationally significant native transcoding in Cisco Unity Connection, such as when requiring a large number of clients connected to Cisco Unity Connection using G.722 or iLBC.

- Which types of graphical user interface (GUI) clients (web browsers, email clients, media players, and so forth) will be fetching the recordings, and which codecs do the GUI clients support?

- What quality of the sound is produced by the selected codec? Some codecs are higher quality than others. For example, G.711 has a higher quality than G.729a, and it is a better choice if higher audio quality is necessary.

- How much disk space does the codec use per second of recording time?

Table 13-4 summarizes the characteristics of the codec formats supported by Cisco Unity Connection.

Table 13-4   Codec Characteristics

Recording Format (Codec)	Audio Quality	Supportability	Disk Space (Bandwidth)
Linear PCM	Highest	Widely supported	16 kbps
G.711 mu-law and a-law	Moderate	Widely supported	8 kbps
G.729a	Lowest	Poorly supported	1 kbps
G.726	Moderate	Moderately supported	3 kbps

To change how Cisco Unity Connection advertises the codecs that it supports on the SIP or SCCP ports, the configuration is done differently than for Cisco Unity. Refer to the *System Administration Guide for Cisco Unity Connection* for details on changing the codec advertised by Cisco Unity Connection. The choices for advertised codecs are G.711 mu-law, G.711 a-law, G.729, iLBC and G.722. There is also a list of preferences according to how they are ordered in the list (top-down). For SCCP integrations, the order of the codecs has no bearing because codecs are advertised and Unified CM negotiates the codec

based on the location of the port and device in the negotiated call. For SIP integrations, however, the order list is significant. If the codec is preferred, then Cisco Unity Connection will advertise that it supports both protocols but will prefer to use the one specified over the other.

For information on how to change the system-level recording format in Cisco Unity Connection Administration, refer to the *System Administration Guide for Cisco Unity Connection*.

Integration with Unified CM

Cisco Unified CM can integrate with both Cisco Unity and Unity Connection via SCCP or SIP. This section discusses some specifics of that integration regarding phones, SIP trunks, and voice ports.

Telephony Integrations

Cisco Unity supports multiple separate telephony integrations, which allow users to be associated with a particular telephony integration. The message waiting indication (MWI) ports are also associated with a particular integration; thus, MWI requests are made through the ports that are associated with that specific integration.

In Cisco Unity Connection this functionality is similar. Users are associated to a phone system that contains one or more port groups. The port groups are associated with MWI ports; thus, the MWI requests are made through the ports associated to that specific port group.

Cisco Unity telephony integrations are configured with the Cisco Unity Telephony Integration Manager (UTIM), while Cisco Unity Connection phone systems and port groups are configured with the System Administrator.

Cisco Unity and Unity Connection now support an unlimited number of telephony integrations and are limited only by the number of ports per system. These features function the same way for both SCCP and SIP integrations. For details, refer to the appropriate Cisco Unity or Cisco Unity Connection administration guides available at http://www.cisco.com.

In addition to the option of adding multiple clusters by adding additional integrations for each new Unified CM cluster in Cisco Unity, Unified CM supports Annex M.1, Message Tunneling for QSIG, which gives administrators the ability to enable QSIG on intercluster trunks (ICTs) between Unified CM clusters. When QSIG is enabled on ICTs, Cisco Unity needs to integrate with only one Unified CM cluster and designate ports only in this one cluster for turning MWIs on and off, even when supporting multiple clusters. The Annex M.1 feature in Unified CM allows for propagation of the MWI requests across the ICTs to the proper Unified CM cluster and phone within that cluster. All calls originating in other clusters can be forwarded to the Cisco Unity server integrated to that one cluster. There is no need to designate MWI ports on the other clusters when Annex M.1 is enabled on the ICT.

For more information on Annex M.1, refer to the protocol descriptions in the *Cisco Unified Communications Manager System Guide*, available at http://www.cisco.com.

Enhanced Message Waiting Indicator (eMWI)

Enhanced Message Waiting Indicator (eMWI) is an enhancement to traditional MWI, and it provides a visual indication of the number of voice messages. Traditional MWI works in a binary format by either enabling or disabling the message lamp on the phone whenever a new voice message arrives in or is deleted from a user's voicemail box. EMWI works with both Cisco Unity and Cisco Unity Connection and is supported on the Cisco Unified IP Phones 8900 and 9900 Series SIP phones.

> **Note** Cisco 8900 and 9900 Series SIP phones are supported in Cisco Unified Communications Manager 7.1.3 and later releases. For further details, refer to the *Cisco Unified Communications Manager Software Compatibility Matrix*, available at http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/compat/ccmcompmatr.html.

eMWI is a visual indication of unplayed messages in the user's voicemail box, with a colored indication depicting the status of the message. An unplayed message displays a red indication on the screen of the phone. eMWI is supported on Unified CM for both Cisco Unity and Cisco Unity Connection through SIP and SCCP integrations. eMWI does not function when the system is running in SRST or SRSV mode. In an integration with Cisco Unity Connection, only the messages stored on the Cisco Unity Connection servers will be indicated with eMWI, and any messages stored on an external IMAP server will not be indicated.

eMWI works in distributed call processing environments with Unified CM. In a system with distributed call processing and centralized voice messaging integration, where one cluster provides the connectivity to the voice messaging server through an intercluster trunk (H.323 or SIP), eMWI updates over the intercluster trunk are supported and are displayed on the end device. (See Figure 13-17.)

> **Note** eMWI also works in a distributed call processing environment with centralized messaging over an intercluster trunk (H.323 or SIP).

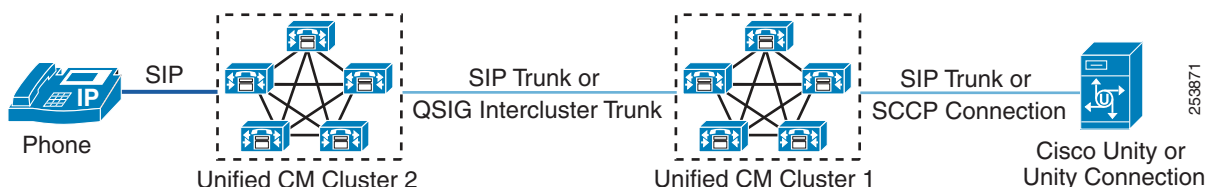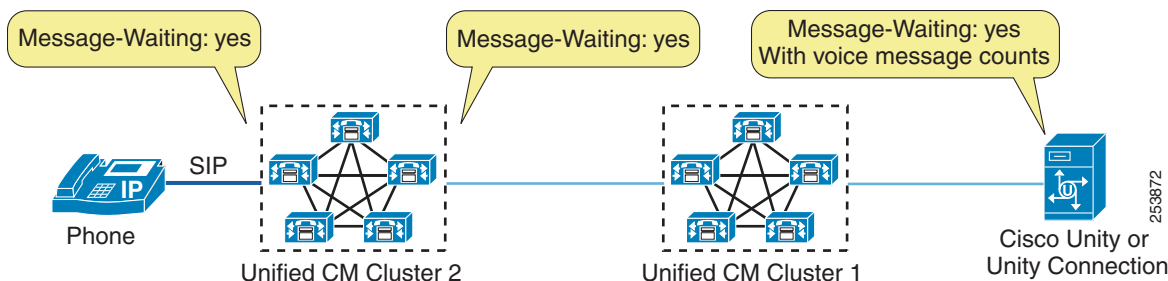*Figure 13-17   Enhanced Message Waiting Indicator (eMWI)*

Figure 13-18 illustrates eMWI over an intercluster trunk (H.323 or SIP) in a distributed call processing environment with centralized voice messaging.

Figure 13-18   eMWI with Distributed Call Processing and Centralized Voice Messaging

As shown in Figure 13-18, Cluster 2 and its voice messaging solution support eMWI, but Cluster 1 does not. If an eMWI update with a voice message count is sent from the voice messaging solution intended for the Cluster 2 phone, Cluster 1 will forward only a standard MWI to Cluster 2 without the voice message count.

The following guidelines apply to eMWI:

- All clusters should support eMWI. If an intermediate cluster does not support eMWI, then the terminating cluster will receive a standard MWI only without voicemail counts.

- Standard MWI does not generate much traffic because it sends only a change of lamp state (ON or OFF). However, enabling eMWI can increase the amount of traffic because it also sends message counts from the messaging system. The amount of traffic depends on the number of messages and change notifications.

Configuring a Unified CM SIP Trunk for Voicemail Integration

Cisco Unified CM provides SIP support for line-side devices as well as trunk-side SIP. These features enable Unified CM to integrate directly to Cisco Unity and Unity Connection via SIP without the need for a SIP proxy server.

However, while Unified CM allows direct integration to voicemail via SIP trunks, it does not currently have the same Voicemail Port Wizard that Skinny Client Control Protocol (SCCP) ports do. Some manual configuration is required. The following steps are required to configure a SIP trunk for a basic integration scenario.

Step 1   Create a SIP Profile in Unified CM Administration by selecting **Device > Device Setting > SIP Profile**.

There is nothing specific to voicemail integrations here, but for easy and consistency of administration functions you will want to create a SIP profile specific for your voicemail integration(s) and name it accordingly.

Step 2   Create a SIP Trunk Security Profile in Unified CM Administration by selecting **System > Security Profile > SIP Trunk Security Profile**.

The **Incoming Port #** under the SIP Trunk Security Profile is specific to voicemail, and it is the number of the SIP port number that the Cisco Unity or Unity Connection server uses to connect to the Unified CM cluster. Additionally, check (enable) **Accept Unsolicited Notifications** so that Cisco Unity and Unity Connection can notify Unified CM of Message Waiting Indicator (MWI) events.

> **Note**   MWI functions are handled via Unsolicited NOTIFY. There is no need to configure the MWI DNs in the SIP integration (MWI DNs are required for SCCP integrations).

Step 3   Create a SIP Trunk in Unified CM Administration by selecting **Device > Trunk [Add New]**. Under the trunk creation page, assign the previously configured SIP Profile and SIP Trunk Security Profile in the appropriate fields. Additionally, configure the destination address of the Unity or Unity Connection server and check (enable) **Unattended Port**. Also check (enable) **Redirecting Number IE Deliver - Outbound**.

> **Note**   One exception to the Unattended Port occurs when multiple Cisco Unity servers are attached to a single Unified CM cluster and live reply or cross-server logon are configured. In this scenario, you would want to leave Unattended Port uncheck (disabled) to allow calls transferred from voicemail ports on one server to be terminated on the other server. This exception currently applies to Unity only.

Step 4   Create a route pattern and specify the voicemail SIP trunk as the destination.

Step 5   Create a Voice Mail Pilot number that is the same number as the route pattern configured in Step 4.

Step 6   Create a Voice Mail Profile using the appropriate Voice Mail Pilot number.

Voice Port Integration with a Unified CM Cluster

When deploying Cisco Unity in a single-site messaging environment, integration with the Unified CM cluster occurs through the SCCP voice ports or SIP trunks. Design considerations must include proper deployment of the voice ports among the Unified CM subscribers so that, in the event of a subscriber failure (Unified CM failover), users and outside calls can continue to access voice messaging. (See Figure 13-19.)

Figure 13-19   Cisco Unity Server(s) Integrated with a Unified CM Cluster (No Dedicated Backup Servers)
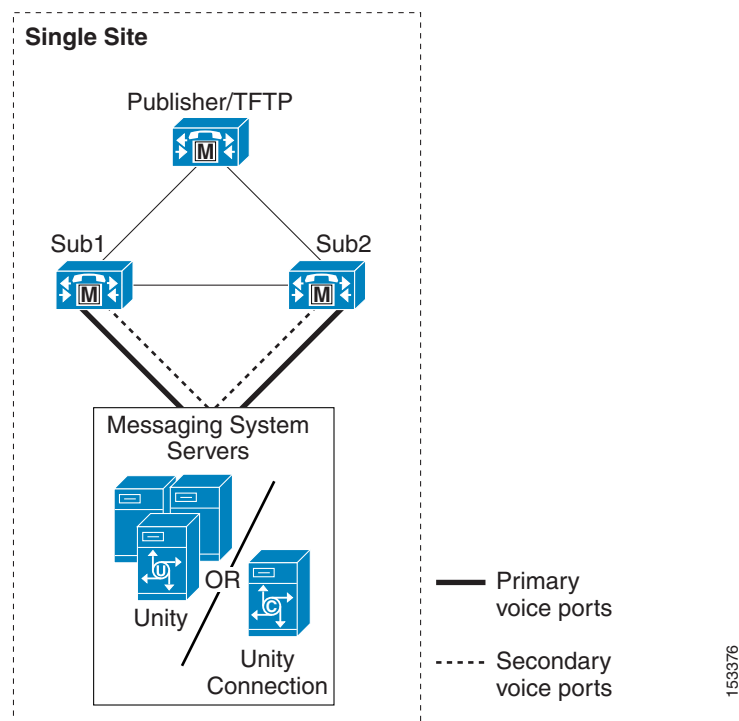
The Unified CM cluster in Figure 13-19 employs 1:1 server redundancy and 50/50 load balancing. During normal operations, each subscriber server is active and handles up to 50% of the total server call processing load. In the event of a subscriber server failure, the remaining subscriber server takes up the load of the failed server.

This configuration uses two groups of voicemail ports, with each group containing one-half of the total number of licensed voice ports. One group is configured so that its primary server is Sub1 and its secondary (backup) server is Sub2. The second group is configured so that Sub2 is the primary server and Sub1 is the backup.

Make sure that MWI-only ports or any other special ports are equally distributed between the two groups. During the configuration of the voice ports, pay special attention to the naming convention. When configuring the two groups of ports in the Cisco Unity Telephony Integration Manager (UTIM) utility, make sure that the device name prefix is unique for each group and that you use the same device

name when configuring the voicemail ports in Unified CM Administration. The device name prefix is unique for each group of ports in this example, with group Sub1 using CiscoUM1 as the device name prefix and Sub2 using CiscoUM2 in this example.

For additional design information on the ratio of inbound to outbound voicemail ports (for MWI, message notification, and TRaP), refer to the *Cisco Unity System Administration Guide* available at http://www.cisco.com.

> **Note**   The device name prefix is unique for each group of ports and must match the same naming convention for the voicemail ports configured in Unified CM Administration.

In Unified CM Administration, half of the ports in this example are configured to register using the unique device name prefix of CiscoUM1, and the other half are configured to register using the unique device prefix CiscoUM2. (See Table 13-5.) When the ports register with Unified CM, half will be registered with subscriber server Sub1, and the other half will be registered with Sub2, as shown in Table 13-5.

Table 13-5   Voicemail Port Configuration in Unified CM Administration

| Device Name	Description	Device Pool	SCCP Security Profile	Status	IP Address
CiscoUM1-VI1	Unity1	Default	Standard Profile	Registered with sub1	1.1.2.9
CiscoUM1-VI2	Unity1	Default	Standard Profile	Registered with sub1	1.1.2.9
CiscoUM1-VI3	Unity1	Default	Standard Profile	Registered with sub1	1.1.2.9
CiscoUM1-VI4	Unity1	Default	Standard Profile	Registered with sub1	1.1.2.9
CiscoUM2-VI1	Unity1	Default	Standard Profile	Registered with sub2	1.1.2.9
CiscoUM2-VI2	Unity1	Default	Standard Profile	Registered with sub2	1.1.2.9
CiscoUM2-VI3	Unity1	Default	Standard Profile	Registered with sub2	1.1.2.9
CiscoUM2-VI4	Unity1	Default	Standard Profile	Registered with sub2	1.1.2.9

> **Note**   The naming convention used for the voicemail ports in Unified CM Administration must match the device name prefix used in Cisco UTIM, otherwise the ports will fail to register.

Voice Port Integration with Dedicated Unified CM Backup Servers

This Unified CM cluster configuration allows each subscriber server to operate at a call processing load higher than 50%. Each primary subscriber server has either a dedicated or shared backup server. (See Figure 13-20.) During normal operation, the backup server processes no calls; but in the event of failure or maintenance of a Subscriber server, the backup server will then take the full load of that server.

Figure 13-20   Cisco Unity Server(s) Integrated with a Single Unified CM Cluster with Backup Subscriber Server(s)
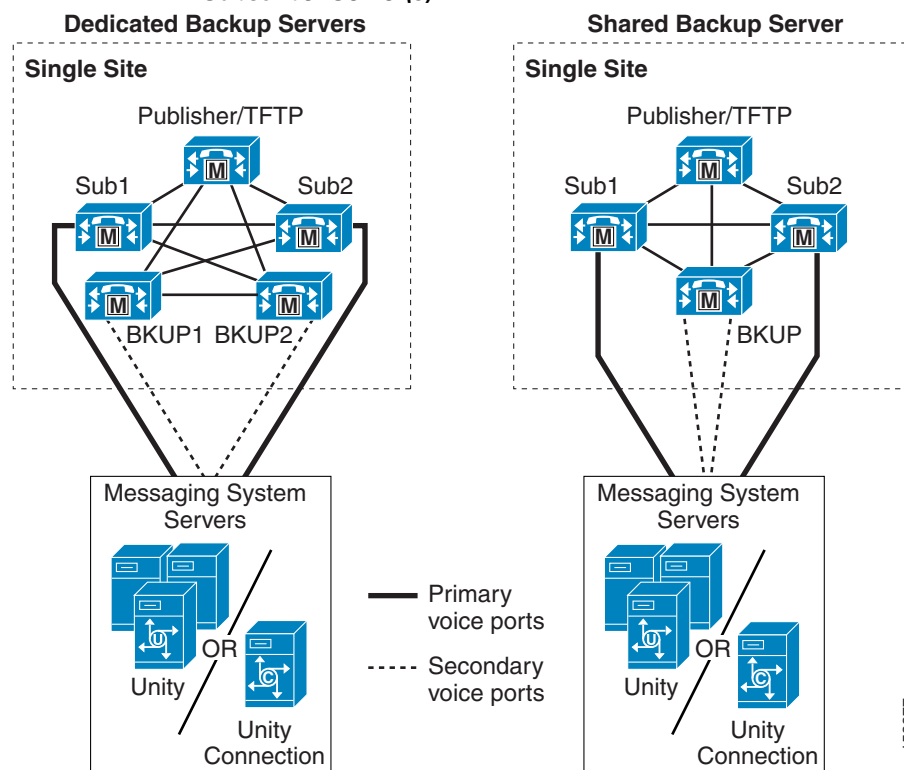
Configuration of the voicemail ports in this case is similar to the 50/50 load-balanced cluster. However, instead of configuring the voice ports to use the opposite subscriber server as the secondary server, the individual shared or dedicated backup server is used. In the Unified CM cluster with a shared backup server, both of the secondary ports for the subscriber servers are configured to use the single backup server.

The voice port names (device name prefix) must be unique for each Cisco UTIM group and must be the same as the device names used on the Unified CM server.

To configure the voicemail ports on Cisco Unity, use the UTIM tool. For Cisco Unity Connection, use the Telephony Integration section of the Unity Connection Administration console. For details, refer to the Cisco Unity or Cisco Unity Connection administration guides available at http://www.cisco.com.

Best Practices for Deploying Cisco Unity Express

When deploying Cisco Unity Express, use the following guidelines and best practices:

- Ensure that the IP phones having Cisco Unity Express as their voicemail destination are located on the same LAN segment as the router hosting Cisco Unity Express.

- If uninterrupted automated attendant (AA) and voicemail access is required for a site deployed with Cisco Unity Express, ensure that Cisco Unity Express, SRST, and the PSTN voice gateway are all located at the same physical site. Hot Standby Router Protocol (HSRP) or other redundant router configurations are not currently supported with Cisco Unity Express.

- Each mailbox can be associated with a primary extension number and a primary E.164 number. Typically, this number is the direct-inward-dial (DID) number that PSTN callers use. If the primary E.164 number is configured to any other number, use Cisco IOS translation patterns to match either the primary extension number or primary E.164 number so that the correct mailbox can be reached during SRST mode.

Voicemail Integration with Unified CM

- Each Cisco Unity Express site must be associated with a CTI route point for voicemail and one for AA (if licensed and purchased), and you must configure the same number of CTI route points as Cisco Unity Express ports licensed. Ensure that the number of sites with Cisco Unity Express does not exceed the CTI scalability guidelines presented in the chapter on Call Processing, page 8-1.

- Cisco Unity Express is associated with a JTAPI user on Unified CM. Although a single JTAPI user can be associated with multiple instances of Cisco Unity Express in a system, Cisco recommends associating each dedicated JTAPI user in Unified CM with a single Cisco Unity Express.

- If Unified CM is upgraded from a previous version, the password of the JTAPI user automatically gets reset on Unified CM. Therefore, after the upgrade, the administrator must make sure that the JTAPI password is synchronized between Cisco Unity Express and Unified CM so that Cisco Unity Express can register with Unified CM.

- The CTI ports and CTI route points can be defined in specific locations. Cisco recommends using locations-based call admission control between Unified CM and Cisco Unity Express. RSVP may also be used.

- Ensure proper Quality of Service (QoS) and bandwidth for signaling traffic that traverses the WAN between Cisco Unity Express and Unified CM. Provision 20 kbps of bandwidth for CTI-QBE signaling for each Cisco Unity Express site. See the chapter on Network Infrastructure, page 3-1, for more details.

- The CTI-QBE signaling packets from Unified CM to Cisco Unity Express are marked with a DSCP value of AF31 (0x68). Unified CM uses TCP port 2748 for CTI-QBE signaling.

- The Unified CM JTAPI library sets the proper IP Precedence bits in all outgoing QBE signaling packets. As a result, all signaling between Cisco Unity Express and Unified CM will have the proper QoS bits set.

Cisco Unity Express Codec and DTMF Support

Calls into Cisco Unity Express use G.711 only. Cisco recommends using a local transcoder to convert the G.729 calls traversing the WAN into G.711 calls. You can configure Unified CM regions with the G.711 voice codec for intra-region calls and the G.729 voice codec for inter-region calls.

If transcoding facilities are not available at the Cisco Unity Express site, provision enough bandwidth for the required number of G.711 voicemail calls over the WAN. Configure the Unified CM regions with the G.711 voice codec for calls between the IP phones and Cisco Unity Express devices (CTI ports and CTI route points).

Cisco Unity Express does not support in-band DTMF tones; it supports only DTMF relay. With Cisco Unity Express, DTMF is carried out-of-band via either the SIP or JTAPI call control channels. Cisco Unity Express 2.3 supports G.711 SIP calls with RFC 2833 into Cisco Unity Express.

JTAPI, SIP Trunk and SIP Phone Support

Cisco Unified CM 5.1 and later releases support SIP trunking protocol; however, Cisco Unity Express uses JTAPI to communicate with Unified CM. Cisco Unity Express supports both SCCP and SIP phones.

- Configure a SIP trunk for SRST and Unified CM for support of SIP phones (through JTAPI).

- Cisco Unity Express 3.0 supports G.729 SIP calls via a transcoder, with the ability added in Cisco IOS Release 12.3(11)XW for RFC 2833 to pass through a transcoder.

- Cisco Unity Express supports delayed media (no SDP in the INVITE message) for call setup in case of a slow-start call from Unified CM.

- Cisco Unity Express supports both blind and consultative transfer, but the default transfer mode is consultative transfer (semi-attended) using REFER in SIP calls. Use the Cisco Unity Express command line interface to explicitly change the transfer mode to consultative transfer using REFER or blind transfer using BYE/ALSO. If REFER is not supported by the remote end, BYE/ALSO will be used.

- Cisco Unity Express supports outcall for voice message notifications. It also supports consultative transfers. During both of these call setups, Cisco Unity Express can receive 3xx responses to the INVITE. Cisco Unity Express processes only 301 (Moved Permanently) and 302 (Moved Temporarily) responses to the INVITE. This requires the URL from the Contact header from the 3xx response to be used to send a new INVITE. 305 (Use Proxy) responses are not supported.

Note   Interoperability with Cisco Unified CM 7.x or 6.x requires a minimum of Cisco Unity Express release 3.2 or 3.0, respectively.

For more information about Cisco Unity Express, refer to the product documentation available at http://www.cisco.com http://www.cisco.com.