



IP Telephony Deployment Models

Last Updated: January 8, 2009

Sections in this chapter address the following topics:

- [Single Site, page 2-1](#)
- [Multisite Implementation with Distributed Call Processing, page 2-3](#)
- [Design Considerations for Section 508 Conformance, page 2-6](#)



Note

For additional information, see the [“Related Documents and References” section on page xii](#).

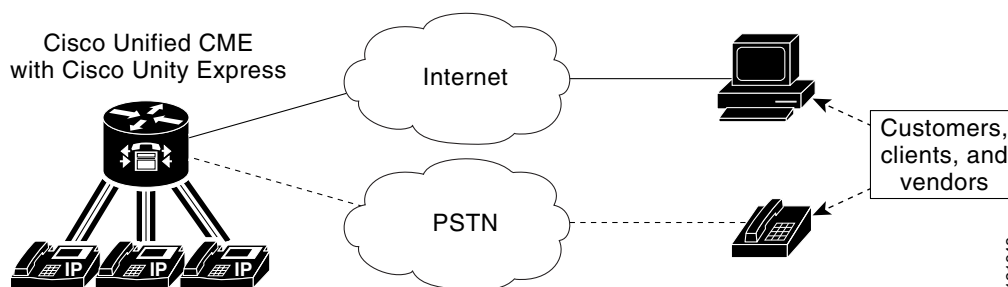
Single Site

The single-site model for IP telephony consists of a call processing agent located at a single site, or campus, with no telephony services provided over an IP WAN. An SMB would typically deploy the single-site model over single site router running at a single location. In this model, calls beyond the LAN use the PSTN.

The single-site model has the following design characteristics:

- Single Cisco Unified CallManager Express (Cisco Unified CME) router or dual Cisco Unified CME router for redundancy
- Maximum of 240 IP phones
- PSTN for all external calls
- Digital signal processor (DSP) resources for transcoding and PSTN termination.
- Voicemail or unified messaging with AA components
- Only G.711 codecs for all IP phone calls (80 kbps of IP bandwidth per call, uncompressed)
- Capability to integrate with legacy private branch exchange (PBX) and voicemail systems

[Figure 2-1](#) illustrates an example of the model for an IP telephony network within a single campus or site.

Figure 2-1 Single-Site Model

Benefits of the Single-Site Model

A single infrastructure for a converged network solution provides significant cost benefits and enables IP telephony to take advantage of the many IP-based applications in the enterprise. Single-site deployment also allows each site to be completely self-contained. There is no dependency for service in the event of an IP WAN failure or insufficient bandwidth, and there is no loss of call processing service or functionality.

In summary, the main benefits of the single-site model are:

- Common infrastructure for a converged solution
- Ease of deployment
- No transcoding resources required, due to the use of only G.711 codecs
- Simplified dial plan

Best Practices for the Single-Site Model

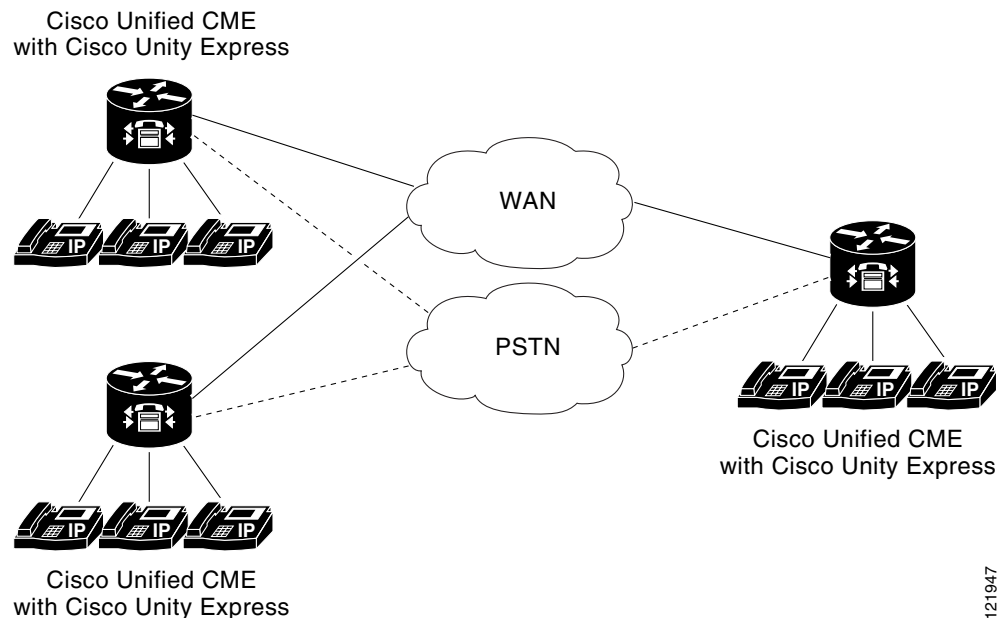
Follow these guidelines and best practices when implementing the single-site model:

- Provide a highly available, fault-tolerant infrastructure based on a common infrastructure philosophy. A sound infrastructure is essential for easier migration to IP telephony, integration with applications such as video streaming and video conferencing, and expansion of your IP telephony deployment across the WAN or to multiple Cisco Unified CallManager clusters.
- Know the calling patterns for your enterprise. Use the single-site model if most of the calls from your enterprise are within the same site or to PSTN users outside your enterprise.
- Use G.711 codecs for all endpoints. This practice eliminates the consumption of digital signal processor (DSP) resources for transcoding, and those resources can be allocated to other functions such as conferencing and Media Termination Points (MTPs).
- Implement the recommended network infrastructure for high availability, connectivity options for phones (in-line power), Quality of Service (QoS) mechanisms, and security. See [Chapter 3, "Network Infrastructure."](#)

Multisite Implementation with Distributed Call Processing

The multisite WAN model with distributed call processing consists of multiple independent sites, each with its own call processing agent connected to a PSTN or an IP WAN (or combination including both) that carries voice traffic between the distributed sites. [Figure 2-2](#) illustrates a typical distributed call processing deployment, featuring an IP WAN interconnection and a backup PSTN connection.

Figure 2-2 A Distributed Call Processing Deployment featuring Cisco Unified CME



Each site in a distributed call processing model can be one of the following:

- A single site with its own call processing agent, which can be either:
 - Cisco Unified CME
 - Cisco Unified CallManager
 - Other IP PBX
- A centralized call manager and all of its associated remote sites
- A legacy PBX with Voice over IP (VoIP) gateway
- One or more gatekeepers can be deployed to help with dial plan management.



Note

See the following Cisco Unified CME documentation:

http://www.cisco.com/en/US/docs/voice_ip_comm/cucme/admin/configuration/guide/cmeadm.html

An IP WAN interconnects all the distributed call processing sites. Typically, the PSTN serves as a backup connection between the sites in case the IP WAN connection fails or does not have any more available bandwidth. A site connected only through the PSTN is a standalone site and is not covered by the distributed call processing model. (See [Single Site, page 2-1](#).)

Connectivity options for the IP WAN include:

- Leased lines
- Frame Relay
- Asynchronous Transfer Mode (ATM)
- ATM and Frame Relay Service Inter-Working (SIW)
- Multiprotocol Label Switching (MPLS) Virtual Private Network (VPN)
- Voice and Video Enabled IP Security Protocol (IP Sec) VPN (V3PN)

Benefits of the Distributed Call Processing Model

The multisite WAN model with distributed call processing provides the following benefits:

- PSTN call cost savings when using the IP WAN for calls between sites
- Use of the IP WAN to bypass toll charges by routing calls through remote site gateways, closer to the PSTN number dialed. This practice is known as tail-end hop-off (TEHO).
- Maximum utilization of available bandwidth by allowing voice traffic to share the IP WAN with other types of traffic.
- No loss of functionality during IP WAN failure because there is a call processing agent at each site.
- Scalability to hundreds of sites.

Best Practices for the Distributed Call Processing Model

A multi-site deployment with distributed call processing has many of the same requirements as a single site or a multi-site deployment with centralized call processing. Follow the best practices from these other models in addition to the ones listed here for the distributed call processing model.

Gatekeeper or Session Initiation Protocol (SIP) proxy servers are among the key elements in the multi-site WAN model with distributed call processing. They each provide dial plan resolution, with the gatekeeper also providing call admission control. A gatekeeper is an H.323 device that provides call admission control and E.164 dial plan resolution.

The following best practices apply to the use of a gatekeeper:

- Use a Cisco IOS gatekeeper to provide call admission control into and out of each site.
- To provide high availability of the gatekeeper, use Hot Standby Router Protocol (HSRP) gatekeeper pairs, gatekeeper clustering, and alternate gatekeeper support. In addition, use multiple gatekeepers to provide redundancy within the network.
- Size the platforms appropriately to ensure that performance and capacity requirements can be met.
- Use only one type of codec on the WAN because the H.323 specification does not allow for Layer 2, IP, User Data Protocol (UDP), or Real-time Transport Protocol (RTP) header overhead in the bandwidth request. (Header overhead is allowed only in the payload or encoded voice part of the packet.) Using one type of codec on the WAN simplifies capacity planning by eliminating the need to over-provision the IP WAN to allow for the worst-case scenario.
- Gatekeeper networks can scale to hundreds of sites, and the design is limited only by the WAN topology.

SIP devices provide resolution of E.164 numbers and SIP uniform resource identifiers (URIs) to enable endpoints to place calls to each other. Cisco Unified CallManager supports the use of E.164 numbers only.

The following best practices apply to the use of SIP proxies:

- Provide adequate redundancy for the SIP proxies.
- Ensure that the SIP proxies have the capacity for the call rate and number of calls required in the network.
- Planning for call admission control is outside the scope of this document.

Call Processing Agents for the Distributed Call Processing Model

Your choice of call processing agent will vary, based on many factors. The main factors, for the purpose of design, are the size of the site and the functionality required.

For a distributed call processing deployment, each site has its own call processing agent. The design of each site varies with the call processing agent, the functionality required, and the fault tolerance required. For example, in a site with 500 phones, a Cisco Unified CallManager cluster containing two servers can provide one-to-one redundancy, with the backup server being used as a publisher and TFTP (Trivial File Transfer Protocol) server.

The requirement for IP-based applications also greatly affects the choice of call processing agent because only Cisco Unified CallManager provides the required support for many Cisco IP applications.

[Table 2-1](#) lists recommended call processing agents.

Table 2-1 Recommended Call Processing Agents

Call Processing Agent	Recommended Size	Comments
Cisco Unified CME	Up to 240 phones	• For small remote sites. • Capacity depends on Cisco IOS platform; see platform-specific support documentation for more details.
Cisco Unified CallManager	50 to 30,000 phones	• Small to large sites, depending on the size of the Cisco Unified CallManager cluster. • Supports centralized or distributed call processing.
Legacy PBX with VoIP gateway	Depends on PBX	• Number of IP WAN calls and functionality depend on the PBX-to-VoIP gateway protocol and the gateway platform.

Design Considerations for Section 508 Conformance

Regardless of which deployment model you choose, you should consider designing your IP telephony network to make the telephony features more accessible to users with disabilities, in conformance with Section 255 of the Telecommunications Act and U.S. Section 508.

Observe the following basic design guidelines when configuring your IP telephony network to conform to Section 508:

- Enable Quality of Service (QoS) on the network.
- Configure only the G.711 codec for phones that will be connected to a terminal teletype (TTY) device or a Telephone Device for the Deaf (TDD). Although low bit-rate codecs such as G.729 are acceptable for audio transmissions, they do not work well for TTY/TDD devices if they have an error rate higher than one percent Total Character Error Rate (TCER).
- Configure TTY/TDD devices for G.711 across the WAN, if necessary.
- Enable (turn ON) Echo Cancellation for optimal performance.
- Voice Activity Detection (VAD) does not appear to have an effect on the quality of the TTY/TDD connection, so it may be disabled or enabled.
- Connect the TTY/TDD to the IP telephony network in either of the following ways:
 - Direct connection (Recommended method)

Plug a TTY/TDD with an RJ-11 analog line option directly into a Cisco FXS port. Any FXS port will work, such as the one on the Cisco VG224, Cisco VG248, Catalyst 6000, Cisco ATA 188 module, or any other Cisco voice gateway with an FXS port. We recommend this method of connection.
 - Acoustic coupling

Place the IP phone handset into a coupling device on the TTY/TDD. Acoustic coupling is less reliable than an RJ-11 connection because the coupling device is generally more susceptible to transmission errors caused by ambient room noise and other factors.
- If stutter dial tone is required, use an analog phone in conjunction with an FXS port on the Cisco VG224 or Cisco ATA 188.