



Deployment Models for Cisco UMG

First Published: February 28, 2008

If a business wants to deploy a Cisco UMG controlled messaging network with multiple sites, one of the key network design decisions that you must make is whether the Cisco UMGs will be centralized at one site or distributed on multiple sites.

This chapter discusses the advantages and disadvantages of the various types of Cisco UMG deployments and contains the following sections:

- [Centralized Cisco UMG Deployment, page 16](#)
- [Distributed Cisco UMG Deployment, page 17](#)
- [Hybrid Cisco UMG Deployment, page 19](#)
- [Best Practices for Cisco UMG Deployment, page 20](#)
- [Migrating from the Fully-Meshed VPIM Network to a Cisco UMG Network, page 20](#)

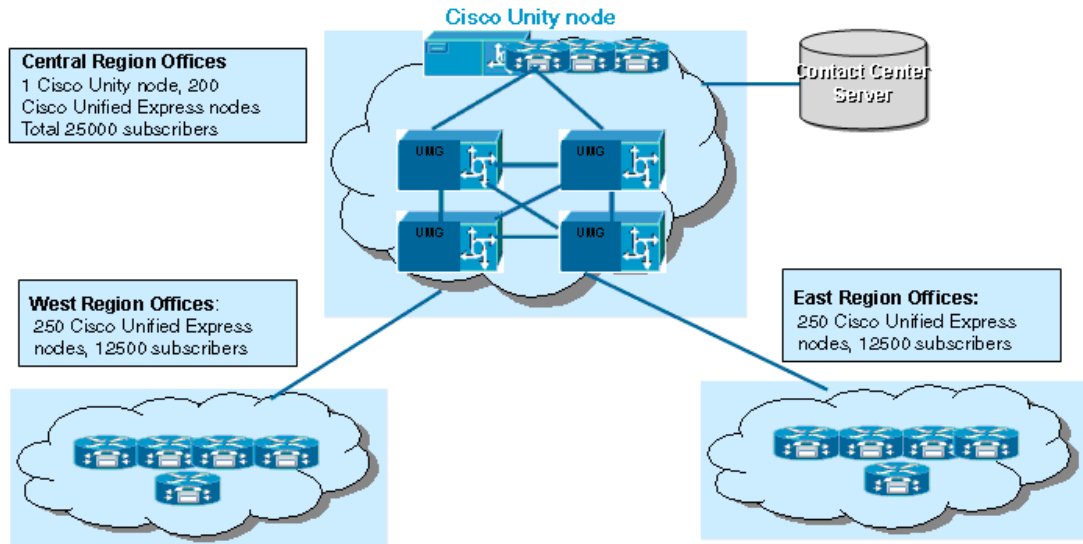


Note

The deployment scenarios in this chapter focus on medium to large enterprise businesses with at least 50 voicemail systems including Cisco Unity Express, Cisco Unity, and/or Avaya Interchange in the network.

Centralized Cisco UMG Deployment

Figure 1 Centralized Cisco Deployment Model



In the centralized Cisco UMG deployment model, all the Cisco UMGs are installed in one location (probably the company headquarters, labeled central region offices in Figure 1). Cisco Unity and Cisco Unity Express are used in the central region as voicemail systems. All branch offices in all regions (west and east in the figure) are equipped with Cisco Unity Express as voicemail systems. The WAN connections are setup between branches and the headquarters. The Contact Center Application is running in the central region.

The advantages of this model are:

- Easy installation on Cisco UMG hardware because all the Cisco UMGs are located in the same region.
- Directory exchange between Cisco UMGs does not consume the WAN bandwidth between remote branch offices (east and west regions) and the central region.
- Hardware cost is less when using Cisco UMG. A NME-UMG-EC can host up to 1000 nodes with one extra unit installed for failover scenario.

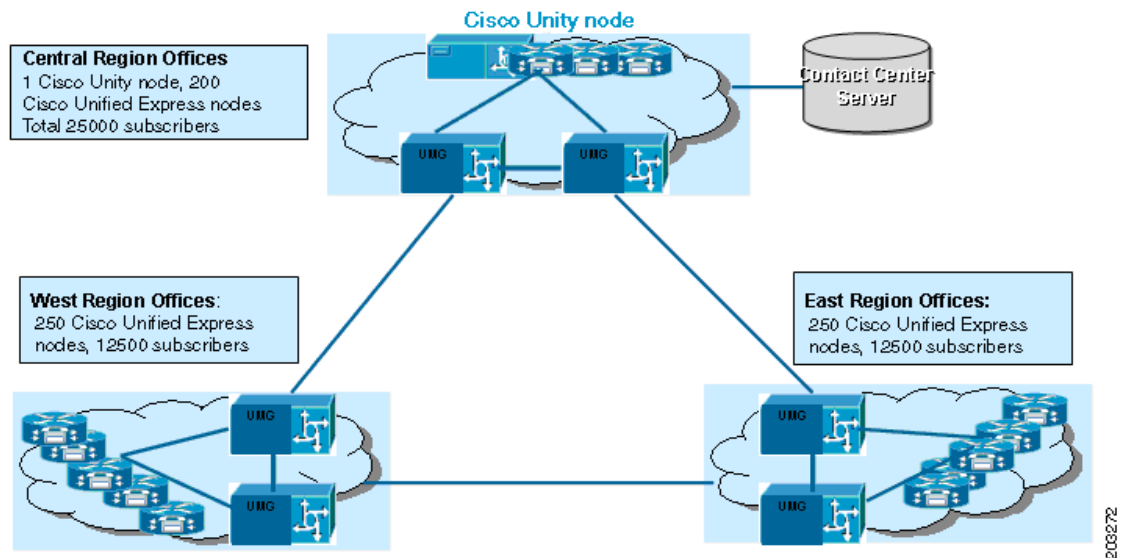
The disadvantages of this model are:

- No message survivability in the remote sites when the WAN link between remote branches and the central region is unreachable. In this messaging network, Cisco UMG is the proxy for any message sent between two nodes. If a WAN link is down, even in the same region, the subscriber cannot leave a message to any subscriber sitting on different Cisco Unity Express database.
- Increases the traffic load on the WAN link. For example, if a subscriber on CUE1 in the West Region is sending a message to a subscriber on CUE2 in the same region, CUE1 sends a VPIM message to the Cisco UMG across the WAN link, after the Cisco UMG searches its directory table to find a match, it delivers the message to the subscriber on CUE2 in the west region using the WAN link again. For every message within the same region, the same message must travel the WAN link twice and consume bandwidth.

- Extra WAN traffic load occurs between Cisco Unity Express nodes on the remote sites and the Cisco UMGs in the headquarters during directory exchanges and during Cisco Unity Express node remote lookups from any branch office to search remote subscribers.
- Limited network scalability especially when the remote office size grows. Although extra Cisco UMG units can be installed on the central site, the WAN link limitation between remote regions and the central site can impact how reliably messages are delivered.

Distributed Cisco UMG Deployment

Figure 2 Distributed Cisco Deployment Model



In the distributed Cisco UMG deployment model, every region has its own primary and secondary Cisco UMG. The Cisco Unity Express nodes on remote sites register with the local Cisco UMG instead of Cisco UMGs in the headquarters. The WAN link is used to connect remote sites to the central site and between remote branches.

The advantages of this model are:

- Full message survivability on the remote sites when the WAN link between the remote site and the central site is down. All the messages between different Cisco Unity Express mail systems within the same region do not have to be proxied by the Cisco UMG across the WAN link to the central site.
- Less network bandwidth is used on the WAN link. The local Cisco UMG can route and deliver the message within the region without utilizing the WAN link. Unlike the centralized model, every message between Cisco Unity Express nodes within a region must travel the WAN link twice. Because people work more closely within the same region, messages are delivered more frequently within the same region, saving more WAN bandwidth.
- Easily scaled on remote regions. On each remote site, a pair of Cisco UMGs (primary and secondary) can support up to 250 nodes with NME-UMG and up to 1000 nodes with NME-UMG-EC, without concern about WAN resource consumption during message delivery.

- No extra license cost compared to the centralized Cisco UMG model if the same number of nodes are deployed in the entire distributed Cisco UMG network. For example, as shown in [Figure 2](#), the west region and the east region have 250 nodes each, and the central site has 201 nodes. So total number of licenses needed is 750, which can be either installed on the central site or distributed to multiple regions. For example, you can put 250 node licenses on the east region, 250 node licenses on the west region, and 250 node licenses on the headquarter.
- No traffic load on the WAN connection during directory exchanges between Cisco Unity Express nodes in remote regions and their hosting Cisco UMGs. Also no traffic is loaded on the WAN connection during remote lookups from the remote site Cisco Unity Express endpoints.
- Increased flexibility on feature management, such as spoken-name. Instead of turning on the feature on the central site Cisco UMG that includes all the nodes in the entire message network. The Cisco UMGs in each region can control the features based on resource availability.

**Note**

The secondary Cisco UMG must purchase the same license as the primary Cisco UMG, in either the centralized or distributed model.

The disadvantages are:

- Extra Cisco UMG units must be purchased and installed on remote sites.
- During the directory exchange between Cisco UMGs, the traffic on the WAN connections may have spikes in volume. These bursts in traffic volume happen only when the Cisco UMG is inserted in to network for the first time and all endpoints register, or when out-of-sync directory information on the Cisco UMG is detected and is unrecoverable. After the system full directory exchange is complete, the subscriber information is stable. In most cases, Cisco UMGs are capable of handling the small updates needed for out-of-sync directory information. The burst traffic that floods into the WAN connections is not significant if the system administrator installs and configures the Cisco UMGs during off-peak hours and verifies that the WAN bandwidth meets the requirement before installation. For slower links, consider turning off the spoken-name confirmation feature to reduce bandwidth usage.

WAN Bandwidth Requirement for Distributed Model

The following three examples use the West Region in [Figure 2](#) as an example. This network information applies to all three examples:

- 250 Cisco Unity Express nodes in the West Region office
- 50 subscribers on each Cisco Unity Express node
- Total number of subscribers = $250 \times 50 = 12500$

No users have spoken-name turned on: Example 1

- SMTP message size during directory exchange: (assume a vCard size is about 180 bytes).
smtpPsize = $180 \times 12500 = 2250K$ bytes
- Add an extra 20%, total Bandwidth = $2250K \times (1+20\%) = 2.7M$ Bytes

20% of users have spoken-name turned on: Example 2

- The number of subscribers with spoken-name = $12500 \times 20\% = 2500$
- The number of subscribers without spoken-name = $12500 \times 80\% = 10000$

- Assume spoken-name is 2 second long at 32 Kbit/s exactly, SMTP message size for a vCard = 4K bytes x 2 = 8K bytes
- Bandwidth required for SMTP size = (8K x 2500) + (180 x 10000) = 18200K = 18.2M
- Add an extra 20%, total Bandwidth = 18.2M x (1+20%) = 22 MB
- The Cisco UMG provides the fragment mechanism when a directory exchange SMTP messaging is over 1 MB to avoid a huge package flood into the WAN link

All subscribers have spoken-name turned on: Example 3

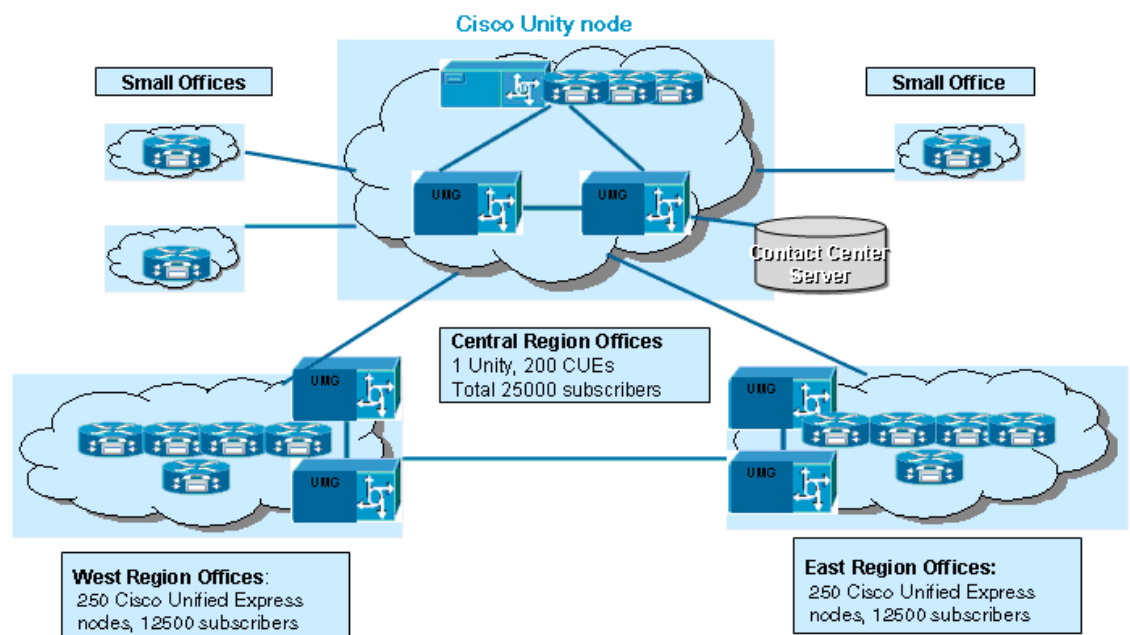
This scenario consumes a large amount of bandwidth. Carefully calculate the bandwidth requirement before deploying.

Additional Considerations

- The extra 20% in the previous estimates assume about 20% of the vCard traffic size is bigger than the average of 180 bytes.
- Fragmentation of SMTP messages to 1 MB units is implemented internally on the Cisco UMG. This is not user configurable.
- Directory-exchange traffic spikes most often during the first time sync-up between multiple regions. Thereafter, it really depends on how much the inter-region feature is used.
- Traffic for a Cisco UMG network is marked as best effort traffic. Therefore with QoS enabled network above traffic does not affect business critical or real time applications.

Hybrid Cisco UMG Deployment

Figure 3 Hybrid Cisco UMG Deployment Model



2003273

In this model, regions with a large number of subscribers on multiple Cisco Unity Express systems (east and west regions in [Figure 3](#)) are registered to the local Cisco UMGs, so that regions can take full advantage of the distributed Cisco UMG model. Remote office sites with a very small number of Cisco Unity Express installations (less than five), they can register with the Cisco UMGs in the central site to reduce hardware and license cost.

**Note**

With the hybrid Cisco UMG deployment model, you must carefully estimate the bandwidth between the remote regions and the central site to ensure the WAN link can handle the extra traffic during message delivery.

Best Practices for Cisco UMG Deployment

In summary, the following best practices apply to the Cisco UMG deployment model.

- If the business infrastructure is the distributed with multiple regions, we recommend the distributed Cisco UMG model. However, you should carefully calculate the WAN bandwidth requirement. In general, the distributed Cisco UMG model provides more geographic scalability and the local message survivability.
- If the message flow is heavy in the remote sites across Cisco Unity Express units, we recommend the distributed model to reduce the WAN link usage during message route and delivery.
- Deploy the centralized Cisco UMG model if:
 - A business is mostly located in one central office.
 - The remote sites have only a few Cisco Unity Express units installed
 - The customer is willing to sacrifice message network survivability when the WAN link is down

The customer must be aware that extra WAN bandwidth will be consumed even when messages are sent within the remote sites, if there are different mail systems. In general, we do not recommend this deployment.

- Deploy the hybrid Cisco UMG model for customers that have both large remote regional offices and some small offices in remote locations, if the overall topology of the business is centralized.

Migrating from the Fully-Meshed VPIM Network to a Cisco UMG Network

When migrating from the current messaging network to a Cisco UMG controlled messaging network, follow these recommendations:

- Upgrade Cisco Unity Express nodes to the version 3.1 to benefit from autoregistration. Earlier versions of Cisco Unity Express require manual provisioning on the Cisco UMGs.
- When registering the Cisco Unity Express nodes to the Cisco UMGs, Cisco Unity Express can keep all the existing VPIM network setup with remote network locations during the migration. The best practice is to:
 - a. Verify the directory tables on the Cisco UMG to ensure the correct subscriber information are saved on the Cisco UMG database.

