

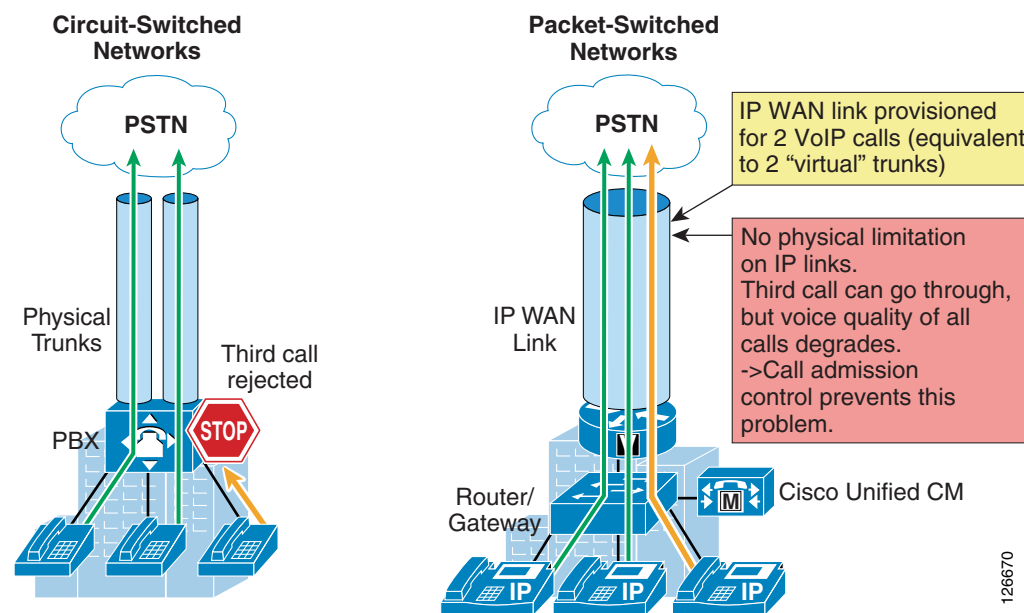


Call Admission Control

Revised: April 2, 2013; OL-29367-04

The call admission control function is an important component of any IP telephony system, especially those that involve multiple sites connected through an IP WAN. In order to better understand what call admission control does and why it is needed, consider the example in [Figure 13-1](#).

Figure 13-1 Why Call Admission Control is Needed



As shown on the left side of [Figure 13-1](#), traditional TDM-based PBXs operate within circuit-switched networks, where a circuit is established each time a call is set up. As a consequence, when a legacy PBX is connected to the PSTN or to another PBX, a certain number of physical trunks must be provisioned. When calls have to be set up to the PSTN or to another PBX, the PBX selects a trunk from those that are available. If no trunks are available, the call is rejected by the PBX and the caller hears a network-busy signal.

Now consider the IP telephony system shown on the right side of [Figure 13-1](#). Because it is based on a packet-switched network (the IP network), no circuits are established to set up an IP telephony call. Instead, the IP packets containing the voice samples are simply routed across the IP network together

with other types of data packets. Quality of Service (QoS) is used to differentiate the voice packets from the data packets, but bandwidth resources, especially on IP WAN links, are not infinite. Therefore, network administrators dedicate a certain amount of "priority" bandwidth to voice traffic on each IP WAN link. However, once the provisioned bandwidth has been fully utilized, the IP telephony system must reject subsequent calls to avoid oversubscription of the priority queue on the IP WAN link, which would cause quality degradation for all voice calls. This function is known as call admission control, and it is essential to guarantee good voice quality in a multisite deployment involving an IP WAN.

To preserve a satisfactory end-user experience, the call admission control function should always be performed during the call setup phase so that, if there are no network resources available, a message can be presented to the end-user or the call can be rerouted across a different network (such as the PSTN).

This chapter discusses the following main topics:

- **Call Admission Control Architecture, page 13-3**

This section describes the call admission control mechanism available through Cisco Unified Communications Manager called Enhanced Location Call Admission Control. For information regarding Cisco IOS gatekeeper, RSVP, and RSVP SIP Preconditions, refer to the *Call Admission Control* chapter of the *Cisco Unified Communications System 9.0 SRND*, available at

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/cac.html

- **Design Considerations for Call Admission Control, page 13-37**

This section shows how to apply Enhanced Location Call Admission Control based on the IP WAN topology.

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:	Revision Date
Video bandwidth utilization	Video Bandwidth Utilization and Admission Control, page 13-30	April 2, 2013
Removal of information on Cisco IOS Gatekeeper	For this information, refer to the <i>Call Admission Control</i> chapter of the <i>Cisco Unified Communications System 9.0 SRND</i> , available at http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/cac.html	April 2, 2013
Removal of information on Resource Reservation Protocol (RSVP)	For this information, refer to the <i>Call Admission Control</i> chapter of the <i>Cisco Unified Communications System 9.0 SRND</i> , available at http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/cac.html	April 2, 2013
Enhanced Location CAC	Unified CM Enhanced Location Call Admission Control, page 13-4	June 28, 2012
RSVP Agent support for RTCP, BFCP and FECC	No longer in this document	June 28, 2012
Migration to RSVP Agent call admission control	No longer in this document	June 28, 2012

Table 13-1 *New or Changed Information Since the Previous Release of This Document (continued)*

New or Revised Topic	Described in:	Revision Date
Example WAN topologies for call admission control	Design Considerations for Call Admission Control, page 13-37	June 28, 2012
Other updates for Cisco Unified Communications System Release 9.0	Various sections throughout this chapter	June 28, 2012

Call Admission Control Architecture

There are several mechanisms that perform the call admission control function in a Cisco IP Communications system. This section provides design and configuration guidelines for Enhanced Location Call Admission Control based on Cisco Unified CM 9.0 or later releases. For all other call admission control mechanisms, refer to the following information:

- Topology-unaware mechanisms
 - [Cisco IOS Gatekeeper Zones, page 13-3](#)
- Topology-aware mechanisms
 - [Unified Communications Architectures Using Resource Reservation Protocol \(RSVP\), page 13-3](#)

Cisco IOS Gatekeeper Zones

Gatekeeper is a legacy technology used for dial-plan resolution and admission control in H.323 networks. With most networks today supporting SIP or mixed H.323 and SIP trunking, gatekeeper is not an effective technology for admission control supporting these environments. For networks supporting H.323 networking or for internetworking call processing agents over H.323 and not mixed SIP networks, refer to the Cisco IOS Gatekeeper information in the *Call Admission Control* chapter of the *Cisco Unified Communications System 9.0 SRND*, available at

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/cac.html

Unified Communications Architectures Using Resource Reservation Protocol (RSVP)

Resource Reservation Protocol (RSVP) is a powerful technology useful in achieving endpoint QoS trust and admission control in Unified Communications voice and video networks. RSVP-enabled networks are possible although not recommended in Unified Communications and Collaboration deployments where TelePresence is pervasively deployed. The reason is that TelePresence requires a QoS overlay network and does not support RSVP call admission control. This means that deployments where RSVP call admission control for desktop video as well as telepresence-to-desktop video interoperability are desired, require a more complex design to ensure that telepresence-to-telepresence calls are provisioned without RSVP while all other calls are provisioned for RSVP. Due to this requirement, Enhanced Location Call Admission Control is preferred for pervasively deployed telepresence and mixed Unified Communications video and telepresence deployments.

For further information on RSVP deployments, refer to the Resource Reservation Protocol (RSVP) information in the *Call Admission Control* chapter of the *Cisco Unified Communications System 9.0 SRND*, available at

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/9x/cac.html

Unified CM Enhanced Location Call Admission Control

Cisco Unified CM 9.x provides Enhanced Location Call Admission Control (CAC) to support complex WAN topologies as well as distributed deployments of Unified CM for call admission control where multiple clusters manage devices in the same physical sites using the same WAN uplinks. The Enhanced Location CAC feature also supports immersive video, allowing the administrator to control call admissions for immersive video calls such as TelePresence separately from other video calls.

To support more complex WAN topologies Unified CM has implemented a location-based network modeling functionality. This provides Unified CM with the ability to support multi-hop WAN connections between calling and called parties. This network modeling functionality has also been incrementally enhanced to support multi-cluster distributed Unified CM deployments. This allows the administrator to effectively "share" locations between clusters by enabling the clusters to communicate with one another to reserve, release, and adjust allocated bandwidth for the same locations across clusters. In addition, an administrator has the ability to provision bandwidth separately for immersive video calls such as TelePresence by allocating a new field to the Location configuration called **immersive video bandwidth**.

There are also a number of tools to administer and troubleshoot Enhanced Location CAC. The CAC enhancements and design are discussed in detail in this chapter, but the troubleshooting and serviceability tools are discussed in separate product documentation.

Network Modeling with Locations, Links, and Weights

Enhanced Location CAC is a model-based static CAC mechanism. Enhanced Location CAC involves using the administration interface in Unified CM to configure Locations and Links to model the "Routed WAN Network" in an attempt to represent how the WAN network topology routes media between groups of endpoints for end-to-end audio, video, and immersive calls. Although Unified CM provides configuration and serviceability interfaces in order to model the network, it is still a "static" CAC mechanism that does not take into account network failures and network protocol rerouting such as RSVP CAC. Therefore, the model needs to be updated when the WAN network topology changes. Enhanced Location CAC is also call oriented, and bandwidth deductions are per-call not per-stream, so asymmetric media flows where the bit-rate is higher in one direction than in the other will always deduct for the highest bit rate. In addition, unidirectional media flows will be deducted as if they were bidirectional media flows.

The administrator builds the network model using locations and links. Enhanced Location CAC incorporates the following configuration components:

- **Locations** — A Location represents a LAN. It could contain endpoints or simply serve as a transit location between links for WAN network modeling.
- **Links** — Links interconnect locations and are used to define bandwidth available between locations. Links logically represent the WAN link and are configured in the Location user interface (UI).
- **Weights** — A weight provides the relative priority of a link in forming the effective path between any pair of locations. The effective path is the path used by Unified CM for the bandwidth calculations, and it has the least cumulative weight of all possible paths. Weights are used on links to provide a "cost" for the "effective path" and are pertinent only when there is more than one path between any two locations.
- **Path** — A path is a sequence of links and intermediate locations connecting a pair of locations. Unified CM calculates shortest paths (least cost) from each location to all other locations and builds the paths. Only one "effective path" is used between a pair of locations.
- **Effective Path** — The effective path is the path with the least cumulative weight.
- **Bandwidth Allocation** — The amount of bandwidth allocated in the model for each type of traffic: audio, video, and immersive video (TelePresence).
- **Location Bandwidth Manager (LBM)** — The active service in Unified CM that assembles a network model from configured location and link data in one or more clusters, determines the effective paths between pairs of locations, determines whether to admit calls between a pair of locations based on the availability of bandwidth for each type of call, and deducts (reserves) bandwidth for the duration of each call that is admitted.
- **Location Bandwidth Manager Hub** — A Location Bandwidth Manager (LBM) service that has been designated to participate directly in intercluster replication of fixed locations, links data, and dynamic bandwidth allocation data. LBMs assigned to an LBM hub group discover each other through their common connections and form a fully-meshed intercluster replication network. Other LBM services in a cluster with an LBM hub participate indirectly in intercluster replication through the LBM hubs in their cluster.

Locations and Links

Unified CM uses the concept of locations to represent a physical site and to create an association with media devices such as endpoints, voice messaging ports, trunks, gateways, and so forth, through direct configuration on the device itself, through a device pool, or even through device mobility.

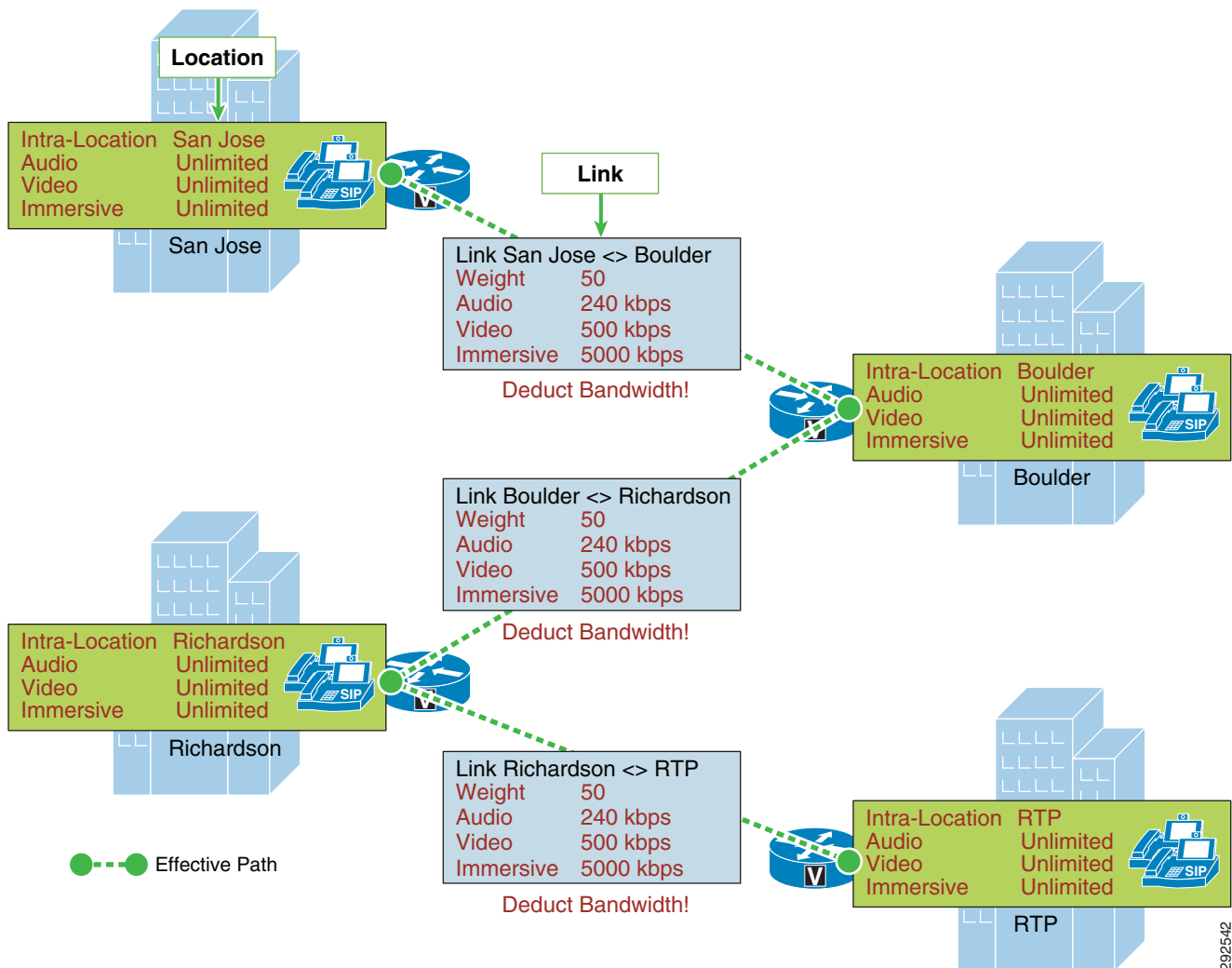
Unified CM 9.x also uses a new location configuration parameter called *links*. Links interconnect locations and are used to define bandwidth available between locations. Links logically represent the WAN links. This section describes locations and links and how they are used.

The location configuration itself consists of three main parts: links, intra-location bandwidth parameters, and RSVP locations settings. The RSVP locations settings are not considered here for Enhanced Location CAC because they apply only to RSVP implementations. In the configuration, the link bandwidth parameters are displayed first while the intra-location bandwidth parameters are hidden and displayed by selecting the **Show advanced** link.

The intra-location bandwidth parameters allow the administrator to configure bandwidth allocations for three call types: audio, video, and immersive. They limit the amount of traffic within, as well as to or from, any given location. When any device makes or receives a call, bandwidth is deducted from the applicable bandwidth allocation for that call type. This feature allows administrators to effectively limit the amount of bandwidth used on the LAN or transit location. In most networks today that consist of at least 100BASE-T or Gigabit LANs, there is little or no reason to limit bandwidth on those LANs. However, there are some deployments that can benefit from limiting high-bandwidth video calls. A

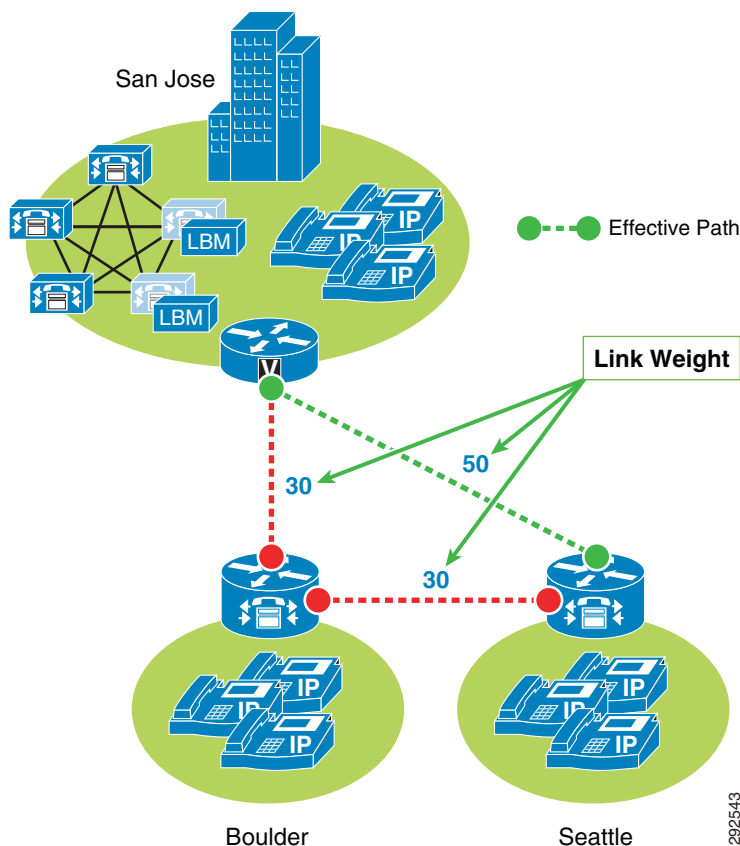
simple example might be an enterprise site with video deployed pervasively on the desktop and/or endpoints. If user calls are mostly all video-enabled, it is easy to see how a large number of 1 to 2 Mbps video calls might utilize a large percentage of available bandwidth on a LAN, and an administrator might consider limiting the number of video calls to a smaller percentage of that available LAN bandwidth. Keep in mind that this utilization might occur only during the busy hour of business or during a specific time of the year when specific traffic levels spike, and the bandwidth limit would be reached only during that time when it would be needed to ensure that the LAN is not over-subscribed with video traffic. It is also noteworthy to mention that video devices can be enabled to **Retry Video Call as Audio** if a video call to that device fails for any reason. This is configured on the video endpoint configuration page in Unified CM and is applicable to video endpoints or trunks receiving calls. It should also be noted that for some video endpoints **Retry Video Call as Audio** is enabled by default and not configurable on the endpoint.

The link bandwidth parameters allow the administrator to characterize the provisioned bandwidth for audio, video, and immersive calls between "adjacent locations" (that is, locations that have a link configured between them). This feature offers the administrator the ability to create a string of location pairings in order to model a multi-hop WAN network. To illustrate this, consider a simple three-hop WAN topology connecting four physical sites, as shown in [Figure 13-2](#). In this topology we want to create links between San Jose and Boulder, between Boulder and Richardson, and between Richardson and RTP. Note that when we create a link from San Jose to Boulder, for example, the inverse link (Boulder to San Jose) also exists. Therefore, the administrator needs to create the link pairing only once from either location configuration page. In the example in [Figure 13-2](#), each of the three links has the same settings: a weight of 50, 240 kbps of audio bandwidth, 500 kbps of video bandwidth, and 5000 kbps (or 5 MB) of immersive bandwidth.

Figure 13-2 Simple Link Example with Three WAN Hops

When a call is made between San Jose and RTP, Unified CM calculates the bandwidth of the requested call, which is determined by the region pair between the two devices (see [Locations, Links, and Region Settings, page 13-10](#)) and verifies the effective path between the two locations. That is to say, Unified CM verifies the locations and links that make up the path between the two locations and accordingly deducts bandwidth from each link and (if applicable) from each location in the path. The intra-location bandwidth also is deducted along the path if any of the locations has configured a bandwidth value other than unlimited.

Weight is configurable on the link only and provides the ability to force a specific path choice when multiple paths between two locations are available. When multiple paths are configured, only one will be selected based on the cumulative weight, and this path is referred to as the *effective path*. This weight is static and the effective path does not change dynamically. [Figure 13-3](#) illustrates weight configured on links between three locations: San Jose, Boulder, and Seattle.

Figure 13-3 Cumulative Path Weights

San Jose to Seattle has two paths, one direct link between the locations and another path through the Boulder location (link San Jose/Boulder and link Boulder/Seattle). The weight configured on the direct link between San Jose and Seattle is 50 and is less than the cumulative weight of links San Jose/Boulder and Boulder/Seattle which is 60 (30+30). Thus, the direct link is chosen as the effective path because the cumulative link weight is 50.

When you configure a device in Unified CM, the device can be assigned to a location. A location can be configured with links to other locations in order to build a topology. The locations configured in Unified CM are virtual locations and not real, physical locations. As mentioned, Unified CM has no knowledge of the actual physical topology of the network. Therefore, any changes to the physical network must be made manually in Unified CM to map the real underlying network topology with the Unified CM locations model. If a device is moved from one physical location to another, the system administrator must either perform a manual update on its location configuration or else implement the device mobility feature so that Unified CM can correctly calculate bandwidth allocations for calls to and from that device. Each device is in location **Hub_None** by default. Location **Hub_None** is an example location that typically serves as a hub linking two or more locations, and it is configured by default with unlimited intra-location bandwidth allocations for audio, video, and immersive bandwidth.

Unified CM allows you to define separate voice, video, and immersive video bandwidth pools for each location and link between locations. Typically the locations intra-location bandwidth configuration is left as a default of **Unlimited** while the link between locations is set to a finite number of kilobits per second (kbps) to match the capacity of a WAN links between physical sites. If the location's intra-location audio, video, and immersive bandwidths are configured as **Unlimited**, there will be unlimited bandwidth available for all calls (audio, video, and immersive) within that location and

transiting that location. On the other hand, if the bandwidth values are set to a finite number of kilobits per second (kbps), Unified CM will track all calls within the location and all calls that use the location as a transit location (a location that is in the calculation path but is not the originating or terminating location in the path).

For video calls, the video location bandwidth takes into account both the audio and the video portions of the video call. Therefore, for a video call, no bandwidth is deducted from the audio bandwidth pool. The same applies to immersive video calls.

The devices that can specify membership in a location include:

- IP phones
- CTI ports
- H.323 clients
- CTI route points
- Conference bridges
- Music on hold (MoH) servers
- Gateways
- Trunks
- Media termination point (via device pool)
- Trusted relay point (via device pool)
- Annunciator (via device pool)

The Enhanced Location Call Admission Control mechanism also takes into account the mid-call changes in call type. For example, if an inter-site video call is established, Unified CM will subtract the appropriate amount of video bandwidth from the respective locations and links in the path. If this video call changes to an audio-only call as the result of a transfer to a device that is not capable of video, Unified CM will return the allocated bandwidth to the video pool and allocate the appropriate amount of bandwidth from the audio pool along the same path. Calls that change from audio to video will cause the opposite change of bandwidth allocation.

Table 13-2 lists the amount of bandwidth requested by the static locations algorithm for various call speeds. For an audio call, Unified CM counts the media bit rates plus the IP and UDP overhead. For example, a G.711 audio call consumes 80 kbps (64k bit rate + 16k IP/UDP headers) deducted from the location's and link's audio bandwidth allocation. For a video call, Unified CM counts only the media bit rates for both the audio and video streams. For example, for a video call at a bit rate of 384 kbps, Unified CM will allocate 384 kbps from the video bandwidth allocation.

Table 13-2 **Amount of Bandwidth Requested by the Locations and Links Bandwidth Deduction Algorithm**

Call Speed	Static Location and Link Bandwidth Value
G.711 audio call (64 kbps)	80 kbps
G.729 audio call (8 kbps)	24 kbps
128 kbps video call	128 kbps
384 kbps video call	384 kbps
512 kbps video call	512 kbps
768 kbps video call	768 kbps

For a complete list of codecs and location and link bandwidth values, refer to the bandwidth calculations information in the *Call Admission Control* section of the *Cisco Unified Communications Manager System Guide*, available at

http://www.cisco.com/en/US/products/sw/voicesw/ps556/prod_maintenance_guides_list.html

For example, assume that the link configuration for the location Branch 1 to Hub_None allocates 256 kbps of available audio bandwidth and 384 kbps of available video bandwidth. In this case the path from Branch 1 to Hub_None can support up to three G.711 audio calls (at 80 kbps per call) or ten G.729 audio calls (at 24 kbps per call), or any combination of both that does not exceed 256 kbps. The link between locations can also support different numbers of video calls depending on the video and audio codecs being used (for example, one video call requesting 384 kbps of bandwidth or three video calls with each requesting 128 kbps of bandwidth).

When a call is placed from one location to the other, Unified CM deducts the appropriate amount of bandwidth from the effective path of locations and links from one location to another. Using Figure 13-2 as an example, a G.729 call between San Jose and RTP locations causes Unified CM to deduct 24 kbps from the available bandwidth at the links between San Jose and Boulder, between Boulder and Richardson, and between Richardson and RTP. When the call has completed, Unified CM returns the bandwidth to those same links over the effective path. If there is not enough bandwidth at any one of the links over the path, the call is denied by Unified CM and the caller receives the network busy tone. If the calling device is an IP phone with a display, that device also displays the message "Not Enough Bandwidth."

When an inter-location call is denied by call admission control, Unified CM can automatically reroute the call to the destination through the PSTN connection by means of the Automated Alternate Routing (AAR) feature. For detailed information on the AAR feature, see [Automated Alternate Routing, page 14-73](#).

**Note**

AAR is invoked only when Enhanced Location Call Admission Control denies the call due to a lack of network bandwidth along the effective path. AAR is not invoked when the IP WAN is unavailable or other connectivity issues cause the called device to become unregistered with Unified CM. In such cases, the calls are redirected to the target specified in the Call Forward No Answer field of the called device.

Locations, Links, and Region Settings

Locations work in conjunction with regions to define the characteristics of a call over the effective path of locations and links. Regions define the type of compression or bit rate (8 kbps or G.729, 64 kbps or G.722/G.711, and so forth) that is used between devices, and location links define the amount of available bandwidth for the effective path between devices. You assign each device in the system to both a region (by means of a device pool) and a location (by means of a device pool or by direct configuration on the device itself).

You can configure locations in Unified CM to define:

- Physical sites (for example, a branch office) or transit sites (for example, an MPLS cloud) — A location represents a LAN. It could contain endpoints or simply serve as a transit location between links for WAN network modeling.
- Link bandwidth between adjacent locations — Links interconnect locations and are used to define bandwidth available between locations. Links logically represent the WAN link between physical sites.
 - Audio Bandwidth — The amount of bandwidth that is available in the WAN link for voice and fax calls being made from devices in the location to the configured adjacent location. This bandwidth value is used by Unified CM for Enhanced Location Call Admission Control.

- Video Bandwidth — The amount of video bandwidth that is available in the WAN link for video calls being made from devices in the location to the configured adjacent location. This bandwidth value is used by Unified CM for Enhanced Location Call Admission Control.
- Immersive Video Bandwidth — The amount of immersive bandwidth that is available in the WAN link for TelePresence calls being made from devices in the location to the configured adjacent location. This bandwidth value is used by Unified CM for Enhanced Location Call Admission Control.
- Intra-location bandwidth
 - Audio Bandwidth — The amount of bandwidth that is available in the WAN link for voice and fax calls being made from devices in the location to the configured adjacent location. This bandwidth value is used by Unified CM for Enhanced Location Call Admission Control.
 - Video Bandwidth — The amount of video bandwidth that is available in the WAN link for video calls being made from devices in the location to the configured adjacent location. This bandwidth value is used by Unified CM for Enhanced Location Call Admission Control.
 - Immersive Video Bandwidth — The amount of immersive bandwidth that is available in the WAN link for TelePresence calls being made from devices in the location to the configured adjacent location. This bandwidth value is used by Unified CM for Enhanced Location Call Admission Control.
- The settings for RSVP call admission control between locations — Possible settings are No Reservation, Optional, Optional (Video Desired), Mandatory, and Mandatory (Video Desired).

You can configure regions in Unified CM to define:

- The Max Audio Bit Rate used for intraregion calls
- The Max Audio Bit Rate used for interregion calls
- The Max Video Call Bit Rate (Includes Audio) used for intraregion and interregion calls. This also includes the maximum bit rate for immersive calls when applied to TelePresence endpoints.
- Audio codec preference lists

Unified CM Support for Locations and Regions

Cisco Unified Communications Manager supports 2,000 locations and 2,000 regions with Cisco MCS-7845 servers. To deploy up to 2,000 locations and regions, you must configure the following service parameters in the **Clusterwide Parameters > (System - Location and Region)** and **Clusterwide Parameters > (System - RSVP)** configuration menus:

- Default Intraregion Max Audio Bit Rate
- Default Interregion Max Audio Bit Rate
- Default Intraregion Max Video Call Bit Rate (Includes Audio)
- Default Interregion Max Video Call Bit Rate (Includes Audio)
- Audio codec preference lists

When adding regions, you should select **Use System Default** for the Max Audio Bit Rate and Max Video Call Bit Rate values. If you are using RSVP call admission control, you should also select **Use System Default** for the RSVP parameter.

Changing these values for individual regions and locations from the default has an impact on server initialization and publisher upgrade times. Hence, with a total of 2,000 regions and 2,000 locations, you can modify up to 200 of them to use non-default values. With a total of 1,000 or fewer regions and locations, you can modify up to 500 of them to use non-default values. Table 13-3 summarizes these limits.

Table 13-3 **Number of Allowed Non-Default Regions and Locations**

Number of non-default regions and locations	Maximum number of regions	Maximum number of locations
0 to 200	2,000	2,000
200 to 500	1,000	1,000

**Note**

The Max Audio Bit Rate is used by both voice calls and fax calls. If you plan to use G.729 as the interregion codec, use T.38 Fax Relay for fax calls. If you plan to use fax pass-through over the WAN, change the default Interregion Max Audio Bit Rate to 64 kbps (G.722 or G.711), or else add a region for fax machines to each location with a non-default bit rate of 64 kbps (G.722 or G.711), subject to the limits in Table 13-3.

**Note**

Irrespective of the MCS model you are using, your Cisco Partner or Cisco Systems Engineer should always use the Cisco Unified Communications Sizing Tool (<http://tools.cisco.com/cucst>) to validate all designs that incorporate a large number of remote sites, because there are many interdependent variables that can affect Unified CM cluster scalability (such as regions, locations, gateways, media resources, and so forth). Use the Sizing Tool to accurately determine the number of servers or clusters required to meet your design criteria.

Location Bandwidth Manager

The Location Bandwidth Manager (LBM) is a Unified CM Feature Service managed from the serviceability web pages and responsible for all of the Enhanced Location CAC bandwidth functions. The LBM can run on any Unified CM subscriber or as a standalone service on a dedicated Unified CM server in the cluster. A minimum of one instance of LBM must run in each cluster to enable Enhanced Location CAC in the cluster. The LBM performs the following functions:

- Locations and links path assembly
- Bandwidth calculations over the effective paths in the assembly
- Servicing bandwidth requests from the Cisco CallManager service (Unified CM call control)
- Replication of bandwidth information to other LBMs within the cluster and between clusters when intercluster Enhanced Location CAC is enabled
- Providing configured and dynamic information to serviceability
- Updating Location Real-Time Monitoring Tool (RTMT) counters
- Using Extensible Markup Language (XML) over TCP for communication to/from the Cisco CallManager service as well as between LBMs.

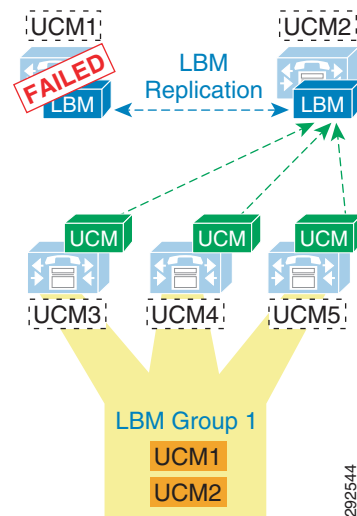
The LBM Service is enabled by default when upgrading to Cisco Unified CM 9.x from earlier releases. For new installations, the LBM service must be activated manually.

During initialization, the LBM reads local locations information from the database, such as: locations audio, video, and immersive bandwidth values; intra-location bandwidth data; and location-to-location link audio, video, and immersive bandwidth values and weight (inter-location bandwidth data). Using the link data, each LBM in a cluster creates a local assembly of the paths from one location to every other location. This is referred to as the *assembled topology*. In a cluster, each LBM accesses the same data and thus creates the same local copy of the assembled topology during initialization.

At runtime, the LBM applies reservations along the computed paths in the local assembled topology of locations and links, and it replicates the reservations to other LBMs in the cluster. If intercluster Enhanced Location CAC is configured and activated, the LBM replicates the assembled topology to other clusters (see [Intercluster Enhanced Location CAC](#), page 13-14, for more details).

By default the Cisco CallManager service communicates with the local LBM service; however, LBM groups can be used to manage this communication. LBM groups provide an active and standby LBM in order to create redundancy for Unified CM call control. [Figure 13-4](#) illustrates LBM redundancy.

Figure 13-4 Location Bandwidth Manager Redundancy



[Figure 13-4](#) shows five Unified CM servers: UCM1 and UCM2 are dedicated LBM servers (only LBM service enabled); UCM3, UCM4, and UCM5 are Unified CM subscribers (Cisco CallManager service enabled). An LBM Group has been configured with UCM1 as active and UCM2 as standby, and it is applied to subscribers UCM3, UCM4, and UCM5. This configuration allows for UCM3, UCM4, and UCM5 to query UCM1 for all bandwidth requests. If UCM1 fails for any reason, the subscribers will fail-over to the standby UCM2.

The order in which the Unified CM Cisco CallManager service uses the LBM is as follows:

- LBM Group designation
- Local LBM
- Service parameter **Call Treatment when no LBM available** (Default = **allow calls**)

Enhanced Location CAC Design and Deployment Recommendations and Considerations

- The Location Bandwidth Manager (LBM) is a Unified CM Feature Service. It is responsible for modeling the topology and servicing Unified CM bandwidth requests.
- LBMs within the cluster create a fully meshed communications network via XML over TCP for the replication of bandwidth change notifications between LBMs.
- Recommendations for LBM Group usage are as follows:
 - Manage how the Cisco CallManager service interacts with LBM (co-resident or dedicated).
 - Minimize LBM full-mesh bandwidth requirements in clustering over the WAN or dual data center deployments.
 - Deploy a minimum of two LBMs per call processing site for redundancy, either co-resident or dedicated.
- Current recommendation is to deploy the LBM service co-resident with a Unified CM subscriber running the Cisco CallManager call processing service.

LBM Group Recommendations

- Configure each Unified CM subscriber to have a local LBM running and active.
- A minimum of two LBMs in a redundant LBM group configuration should be active at each call processing site, such as in clustering over the WAN designs.

Intercluster Enhanced Location CAC

Intercluster Enhanced Location CAC extends the concept of network modeling across multiple clusters. In intercluster Enhanced Location CAC, each cluster manages its locally configured topology of locations and links and then propagates this local topology to other remote clusters that are part of the LBM intercluster replication network. Upon receiving a remote cluster's topology, the LBM assembles this into its own local topology and creates a global topology. Through this process the global topology is then identical across all clusters, providing each cluster a global view of enterprise network topology for end-to-end CAC. [Figure 13-5](#) illustrates the concept of a global topology with a simplistic hub-and-spoke network topology as an example.

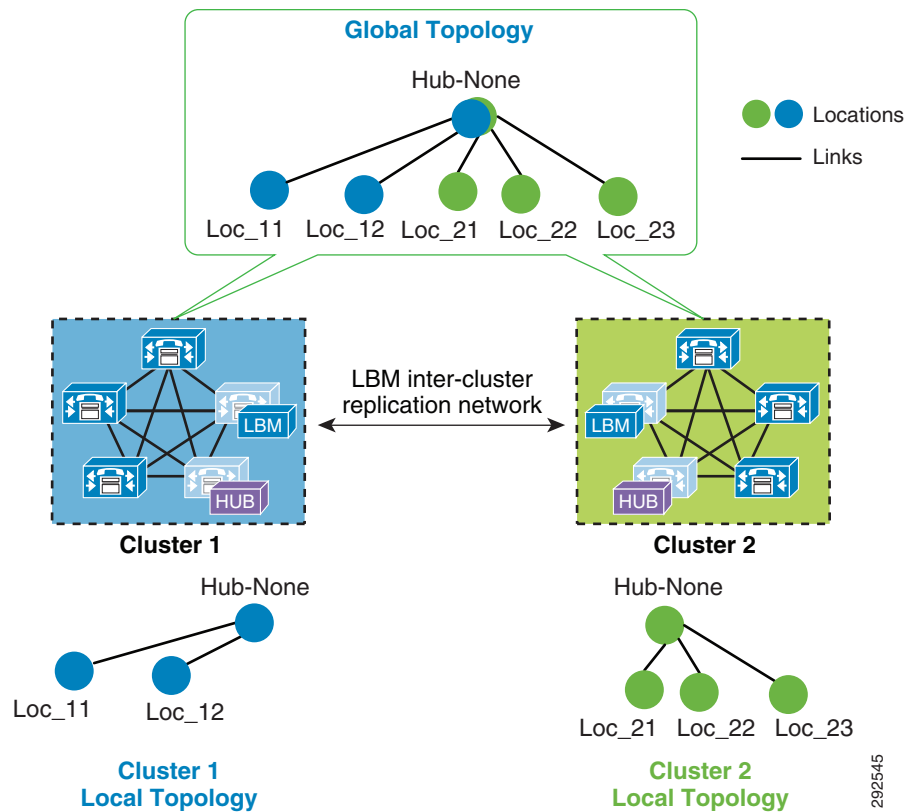
Figure 13-5 Example of a Global Topology for a Simple Hub-and-Spoke Network

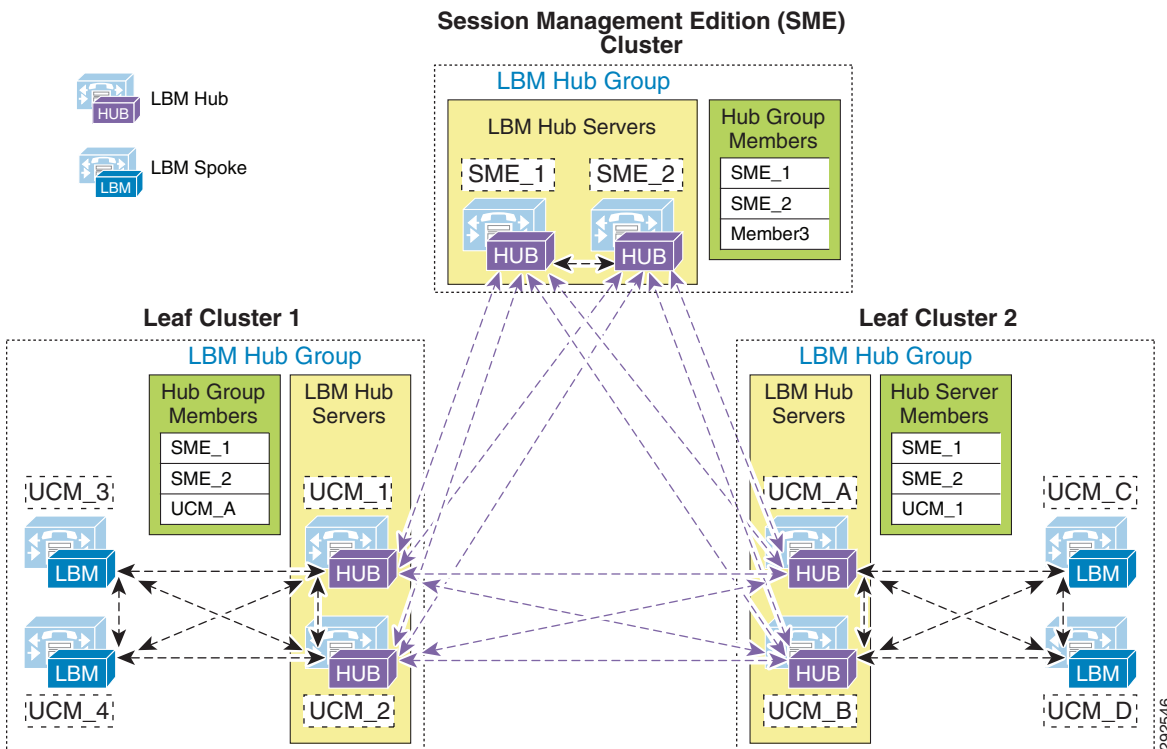
Figure 13-5 shows two clusters, Cluster 1 and Cluster 2, each with a locally configured hub-and-spoke network topology. Cluster 1 has configured Hub_None with links to Loc_11 and Loc_12, while Cluster 2 has configured Hub_None with links to Loc_21, Loc_22, and Loc_23. Upon enabling intercluster Enhanced Location CAC, Cluster 1 sends its local topology to Cluster 2, as does Cluster 2 to Cluster 1. After each cluster obtains a copy of the remote cluster's topology, each cluster overlays the remote cluster's topology over their own. The overlay is accomplished through common locations, which are locations that are configured with the same name. Because both Cluster 1 and Cluster 2 have the common location Hub_None with the same name, each cluster will overlay the other's network topology with Hub_None as a common location, thus creating a global topology where Hub_None is the hub and Loc_11, Loc_12, Loc_21, Loc_22 and Loc_23 are all spoke locations. This is an example of a simple network topology, but more complex topologies would be processed in the same way.

LBM Hub Replication Network

The intercluster LBM replication network is a network of designated LBMs that create a full-mesh with one another and replicate their local cluster's topology. In turn, each receives all remote clusters' topologies in order to create the global topology. The designated LBMs for the intercluster replication network are called LBM hubs, and the LBMs that replicate only within a cluster are called LBM spokes. The LBM hubs are designated in configuration through the LBM hub group. The LBM hub group has two main configuration areas called hub group members and hub group usage information. The hub group members are LBM hubs in remote clusters that are part of the LBM replication network. A maximum of three members can be configured. The members designated in the LBM hub group members serve as bootstrap servers for the entire intercluster replication network, providing each LBM

hub in each cluster with the connectivity details of other remote clusters with whom they are connected. The LBM Hub group usage information consists of the LBM hubs and spokes in the local cluster. Moving an LBM service into or out of the LBM Hub group determines the hub or spoke role. (See Cisco Unified Communications Manager product documentation for further information on the LBM hub group configuration.) Once the LBM hub group is configured on each cluster in the designated LBM, hubs will create the full mesh intercluster replication network. [Figure 13-6](#) illustrates an intercluster replication network configuration with LBM hub groups set up between three clusters (Leaf Cluster 1, Leaf Cluster 2 and a Session Management Edition (SME) Cluster) to form the intercluster replication network.

Figure 13-6 Example Intercluster Replication Network for Three Cluster



In [Figure 13-6](#), two LBM servers from each cluster have been designated as the LBM hubs for their cluster. These LBM hub servers form the intercluster LBM replication network. The LBM hub group members configured in each LBM hub group are designated as SME_1 and SME_2. These two LBM servers from the SME cluster serve as points of contact for the entire intercluster LBM replication network. This means that each LBM in each cluster connects to SME_1, replicates its local topology to SME_1, and gets the remote topology from SME_1. They also get the connectivity information for the other leaf clusters from SME_1, connect to the other remote clusters, and replicate their topologies. This creates the full-mesh replication network. If SME_1 is unavailable, the LBM hubs will connect to SME_2. If SME_2 is unavailable, Leaf Cluster 1 LBM hubs will connect to UCM_A and Leaf Cluster 2 LBM hubs will connect to UCM_1 as a backup measure in case the SME Cluster is unavailable. This is just an example configuration to illustrate the components of the intercluster LBM replication network.

The LBM has the following roles with respect to the LBM intercluster replication network:

- LBM Hub group members
 - Remote hub servers responsible for interconnecting all LBM hubs in the replication network
 - Can be any hub in the network
 - Can indicate up to 3 per hub group
- LBM Hub Servers (Local LBMs)
 - Communicate directly to other remote hub servers as part of the intercluster LBM replication network
- LBM Spoke Servers (Local LBMs)
 - Communicate directly to local LBM hubs in the cluster and indirectly to the remote LBM hubs through the local LBM hubs
- LBM Hub Replication Network — Bandwidth deduction and adjustment messages
 - If a cluster has multiple LBM hubs, the LBM hub with the lowest IPv4 (entire) address will function as the sender of messages to other remote clusters. Only one hub per cluster will forward messages to remote clusters. This limits the amount of replication traffic in the intercluster replication network.
 - The LBM hub that functions as the sender for messages in the cluster selects one LBM hub from each cluster and forwards messages to that LBM.
 - The LBM hubs that receive messages from remote clusters, in turn forward the received messages to the LBM spokes in their local cluster.
 - Forwarded messages have a unique random string associated with them that allows receivers to determine if a messages has already been received and thus drop messages that they have received twice to prevent any replication storm or looping.
 - Other LBM hubs in the cluster that receive the forwarded message will not forward on to LBM spokes because the message is not directly from a remote cluster. This avoids hubs sending duplicate messages from remote clusters.

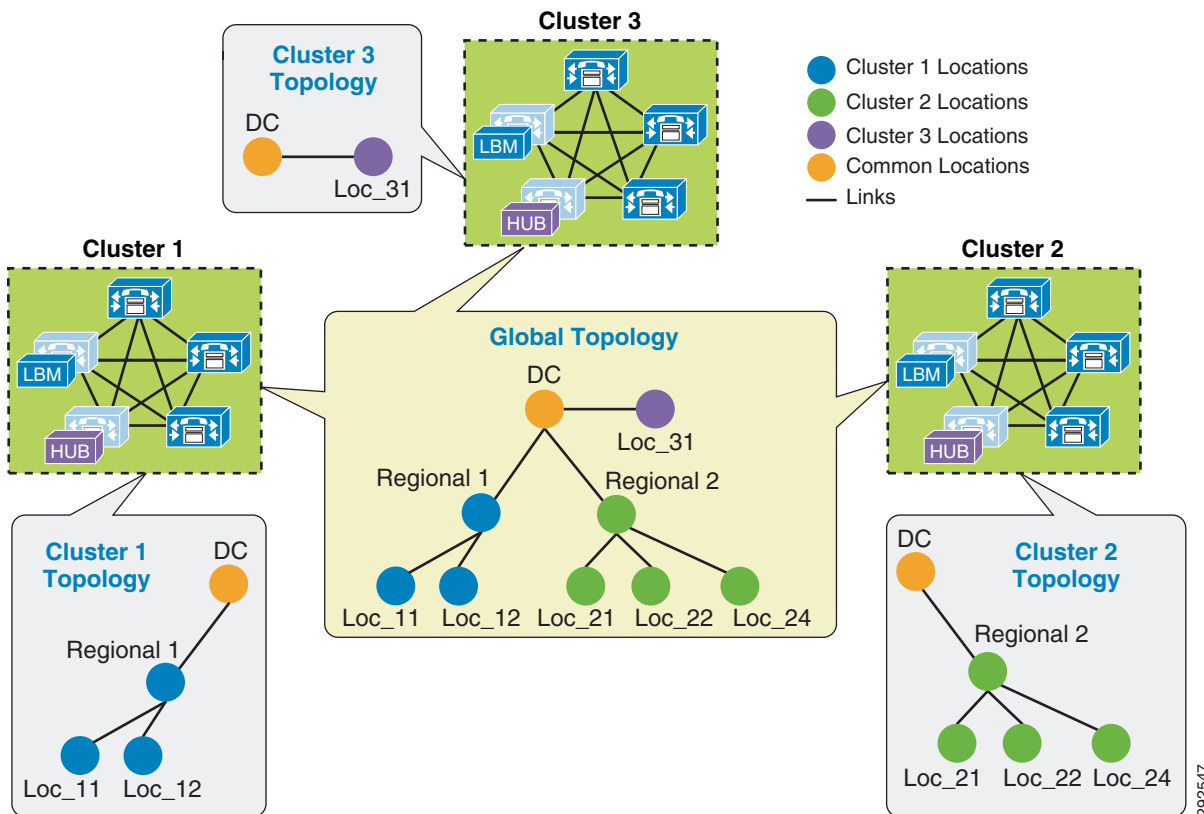
Common Locations (Shared Locations) and Links

As mentioned previously, common locations are locations that are named the same across clusters. Common locations play a key role in how the LBM creates the global topology and how it associates a single location across multiple clusters. A location with the same name between two or more clusters is considered the same location and is thus a shared location across those clusters. So if a location is meant to be shared between multiple clusters, it is required to have exactly the same name. After replication, the LBM will check for configuration discrepancies across locations and links. Any discrepancy in bandwidth value or weight between common locations and links can be seen in serviceability, and the LBM calculates the locations and link paths with the most restrictive values for bandwidth and the lowest value (least cost) for weight.

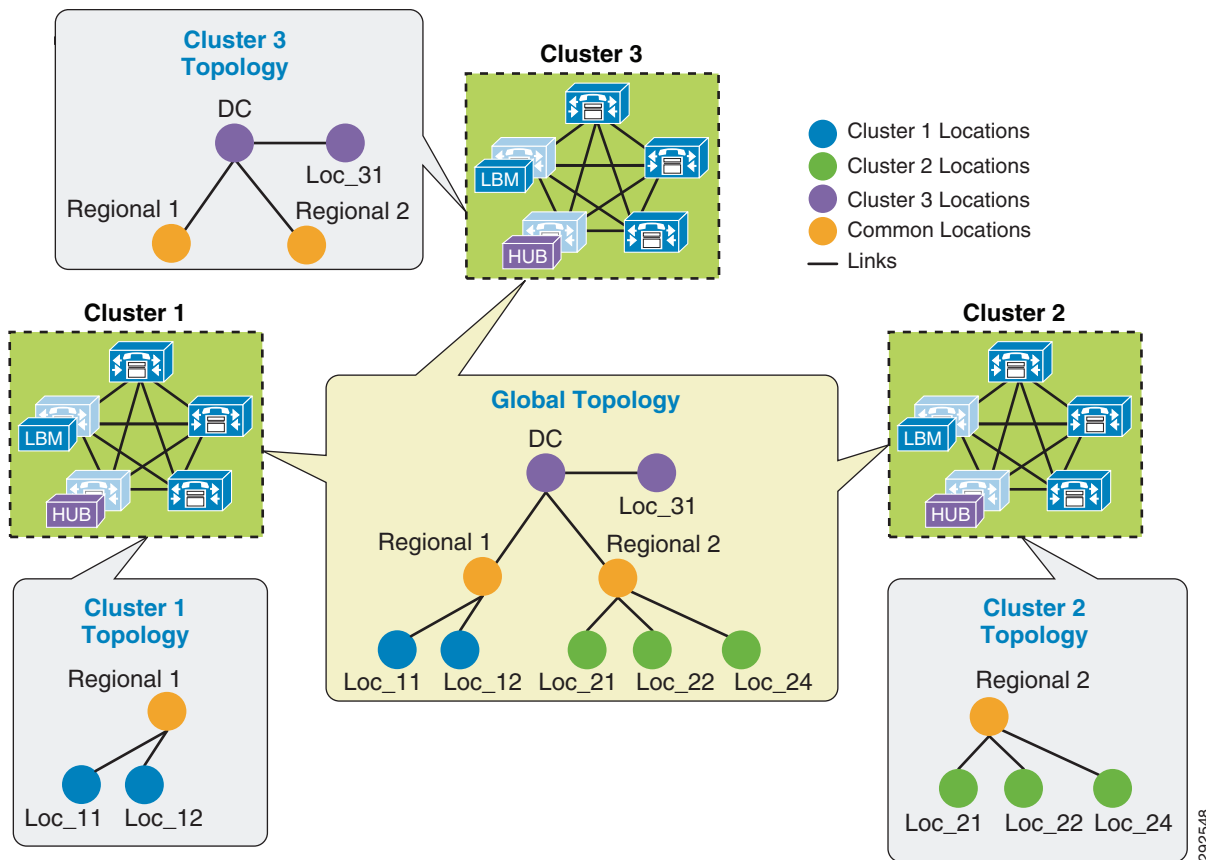
Common locations and links can be configured across clusters for a number of different reasons. You might have a number of clusters that manage devices in the same physical site and use the same WAN uplinks, and therefore the same location needs to be configured on each cluster in order to associate that location to the local devices on each cluster. You might also have clusters that manage their own topology, yet these topologies interconnect at specific locations and you will have to configure these locations as common locations across each cluster so that, when the global topology is being created, the

clusters have the common interconnecting locations and links on each cluster to link each remote topology together effectively. Figure 13-7 illustrates linking topologies together and shows the common topology that each cluster shares.

Figure 13-7 Using Common Locations and Links to Create a Global Topology



In Figure 13-7, Cluster 1 has devices in locations Regional 1, Loc_11, and Loc_12, but it requires configuring DC and a link from Regional 1 to DC in order to link to the rest of the global topology. Cluster 2 is similar, with devices in Regional 2 and Loc_21, Loc_22, and Loc_23, and it requires configuring DC and a link from DC to Regional 2 to map into the global topology. Cluster 3 has devices in Loc_31 only, and it requires configuring DC and a link to DC from Loc_31 to map into Cluster 1 and Cluster 2 topologies. Alternatively, Regional 1 and Regional 2 could be the common locations configured on all clusters instead of DC, as is illustrated in Figure 13-8.

Figure 13-8 Alternative Topology Using Different Common Locations

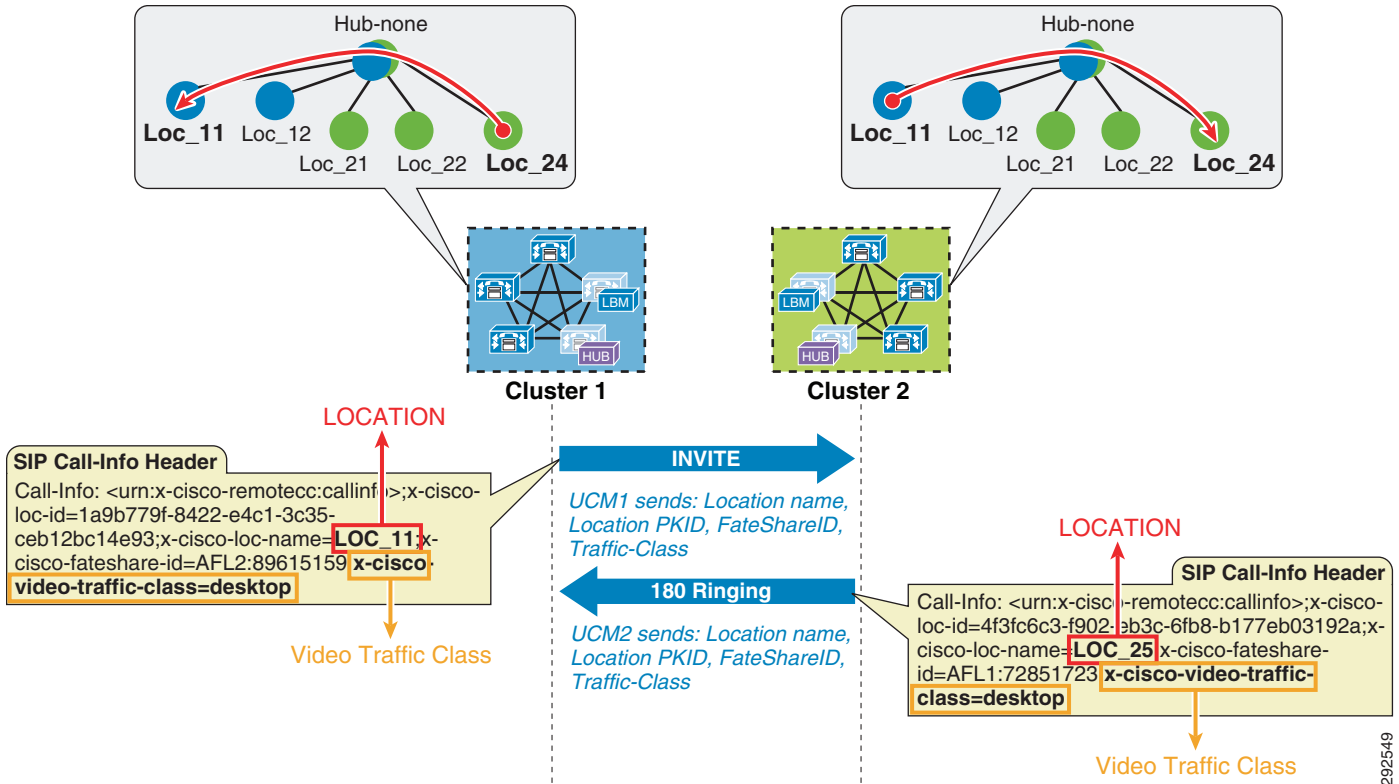
The key to topology mapping from cluster to cluster is to ensure that at least one cluster has a common location with another cluster so that the topologies interconnect accordingly.

Shadow Location

The *shadow location* is used to enable a SIP trunk to pass Enhanced Location CAC information such as location name and Video-Traffic-Class (discussed below), among other things, required for Enhanced Location CAC to function between clusters. In order to pass this location information across clusters, the SIP intercluster trunk (ICT) must be assigned to the "shadow" location. The shadow location cannot have a link to other locations, and therefore no bandwidth can be reserved between the shadow location and other locations. Any device other than a SIP ICT that is assigned to the shadow location will be treated as if it was associated to Hub_None. That is important to know because if a device other than a SIP ICT ends up in the shadow location, bandwidth deductions will be made from that device as if it were in Hub_None, and that could have varying effects depending on the location and links configuration.

When the SIP ICT is enabled for Enhanced Location CAC, it passes information in the SIP Call-Info header that allows the originating and terminating clusters to process the location bandwidth deductions end-to-end. Figure 13-9 illustrates an example of a call between two clusters and some details about the information passed. This is only to illustrate how location information is passed from cluster to cluster and how bandwidth deductions are made.

Figure 13-9 Location Information Passed Between Clusters over SIP



In Figure 13-9, Cluster 1 sends an invite to Cluster 2 and populates the call-info header with the calling parties location name and Video-Traffic-Class, among other pertinent information such as unique call-ID. When Cluster 2 receives the invite with the information, it looks up the terminating party and performs a CAC request on the path between the calling party's and called party's locations from the global topology that it has in memory from LBM replication. If it is successful, Cluster 2 will replicate the reservation and extend the call to the terminating device and return a 180 ringing with the location information of the called party back to Cluster 1. When Cluster 1 receives the 180 ringing, it gets the terminating device's location name and goes through the same bandwidth lookup process using the same unique call-ID that it calculates from the information passed in the call-info header. If it is successful, it too continues with the call flow. Because both clusters use the same information in the call-info header, they will deduct bandwidth for the same call using the same call-ID, thus avoiding any double bandwidth deductions.

Location and Link Management Cluster

In order to avoid configuration overhead, a Location and Link Management Cluster can be configured to manage all locations and links in the global topology. All other locations uniquely configure the locations that they require for location-to-device association and do not configure links or any bandwidth values other than unlimited. It should be noted that the Location and Link Management Cluster is a design concept and is simply any cluster that is configured with the entire global topology of locations and links, while all other clusters in the LBM replication network are configured only with locations with unlimited bandwidth values and no configured links. When intercluster Enhanced Location CAC is enabled and the LBM replication network is configured, all clusters replicate their view of the network. The designated Location and Link Management Cluster has the entire global topology with locations,

links, and bandwidth values; and once those values are replicated, all clusters use those values because they are the most restrictive. This design alleviates configuration overhead in deployments where a large number of common locations are required across multiple clusters.

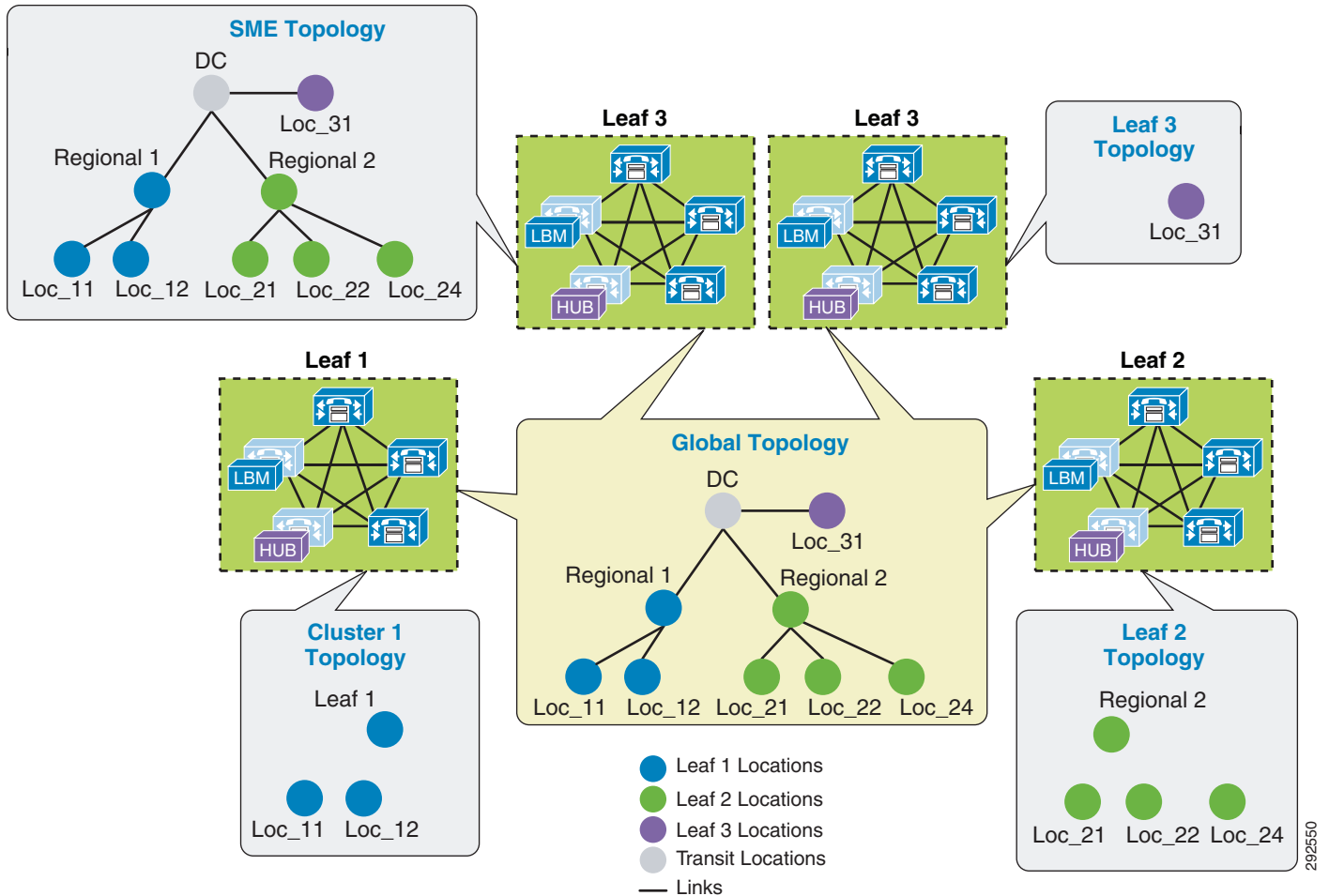
Recommendations

- Locations and link management cluster:
 - One cluster should be chosen as the management cluster (the cluster chosen to manage locations and links).
 - The management cluster should be configured with the following:
 - All locations within the enterprise will be configured in this cluster.
 - All bandwidth values and weights for all locations and links will be managed in this cluster.
- All other clusters in the enterprise:
 - All other clusters in the enterprise should configure *only* the locations required for association to devices but should *not* configure the links between locations. This link information will come from the management cluster when intercluster Enhanced Location CAC is enabled.
 - When intercluster Enhanced Location CAC is enabled, all of the locations and links will be replicated from the management cluster and LBM will use the lowest, most restrictive bandwidth and weight value.
- LBM will always use the lowest most restrictive bandwidth and lowest weight value after replication.

Benefits

- Manage enterprise CAC topology from a single cluster.
- Alleviates location and link configuration overhead when clusters share a large number of common locations.
- Alleviates configuration mistakes in locations and links across clusters.
- Other clusters in the enterprise require the configuration only of locations needed for location-to-device and endpoint association.
- Provides a single cluster for monitoring of the global locations topology.

Figure 13-10 illustrates Cisco Unified Communications Manager Session Management Edition (SME) as a Location and Link Management Cluster for three leaf clusters.

Figure 13-10 Example of SME as a Location and Link Management Cluster

In [Figure 13-10](#) there are three leaf clusters, each with devices in only a regional and remote locations. SME has the entire global topology configured with locations and links, and intercluster LBM replication is enabled between all four clusters. None of the clusters in this example share locations, although all of the locations are common locations because SME has configured the entire location and link topology. Note that Leaf 1, Leaf 2, and Leaf 3 configure only locations that they require to associate to devices and endpoints, while SME has the entire global topology configured. After intercluster replication, all clusters will have the global topology.

Intercluster Enhanced Location CAC Design and Deployment Recommendations and Considerations

- Oversubscription in bandwidth reservations can be incurred since reservations are made locally and replicated out to the rest of the LBM replication network. To avoid QoS impacts during oversubscription, observe the following guidelines:
 - Oversubscription is transient and will correct itself as the calls, using the oversubscribed locations and links in the path, clear and relinquish the bandwidth.
 - Bandwidth overhead should be provisioned in the QoS network policy to accommodate oversubscription. Cisco recommends over-provisioning by a minimum of one call of the highest bandwidth value in each QoS class (audio and video) for applicable locations and links. For audio-only implementations this may be a single call at 24 kbps or 80 kbps. For video implementations this may be the bit rate of a single video call.
 - Locations and links where CAC limits are often reached during the busy hour or in general, are prime candidates for over-provisioning of the QoS bandwidth capacity.
 - In cases of very high busy hour call completions (BHCC) and long delays between Unified CM clusters, Cisco recommends monitoring the locations and links to determine the amount of oversubscription during the busy hour, then ensure that the network is over-provisioned with an equal amount of bandwidth.
- A cluster requires the location to be configured locally for location-to-device association.
- Each cluster should be configured with the immediately neighboring locations so that each cluster's topology can inter-connect. This does not apply to Location and Link Management Cluster deployments.
- Links need to be configured to establish points of interconnect between remote topologies. This does not apply to Location and Link Management Cluster deployments.
- Discrepancies of bandwidth limits and weights on common locations and links are resolved by using the lowest bandwidth and weight values.
- Naming locations consistently across clusters is critical. Follow the practice, "Same location, same name; different location, different name."
- The Hub_None location should be renamed to be unique in each cluster or else it will be a common (shared) location by other clusters.
- Cluster-ID should be unique on each cluster for serviceability reports to be usable.
- All LBM hubs are fully meshed between clusters.
- An LBM hub is responsible for communicating to hubs in remote clusters.
- An LBM spoke does not directly communicate with other remote clusters. LBM spokes receive and send messages to remote clusters through the Local LBM Hub.
- LBM Hub Groups
 - Used to assign LBMs to the Hub role
 - Used to define three remote hub members that replicate hub contact information for all of the hubs in the LBM hub replication network
 - An LBM is a hub when it is assigned to an LBM hub group.
 - An LBM is a spoke when it is not assigned to an LBM hub group.
- If a cluster has no LBM hub, or if the LBM hub is not running, the cluster will be isolated and will not participate in the intercluster LBM replication network.

Performance Guidelines

- Maximum of 2,000 locally configured locations. This limit of 2,000 locations also applies to the Location and Link Management Cluster.
- Maximum of 8,000 total replicated locations with intercluster CAC

Enhanced Location CAC for TelePresence Immersive Video

Since TelePresence endpoints now provide a diverse range of collaborative experiences from the desktop to the conference room, Enhanced Location CAC includes support to provide CAC for TelePresence immersive video calls. This section discusses the features in Enhanced Location CAC that support TelePresence immersive video CAC.

Video Call Traffic Class

Video Call Traffic Class is an attribute that is assigned to all endpoints, and that can also be enabled on SIP trunks, to determine the video classification type of the endpoint or trunk. This enables Unified CM to classify various call flows as either immersive, desktop video, or both, and to deduct accordingly from the appropriate location and/or link bandwidth allocations of video bandwidth, immersive bandwidth, or both. For TelePresence endpoints there is a non-configurable Video Call Traffic Class of **immersive** assigned to the endpoint. A SIP trunk can be classified as desktop, immersive, or mixed video in order to deduct bandwidth reservations of a SIP trunk call. All other endpoints and trunks have a non-configurable Video Call Traffic Class of **desktop video**. More detail on endpoint and trunk classification is provided in the subsections below.

TelePresence immersive endpoints mark their media with a DSCP value of CS4 by default, and desktop video endpoints mark their media with AF41 by default, as per recommended QoS settings. For Cisco endpoints this is accomplished through the configurable Unified CM QoS service parameters **DSCP for Video calls** and **DSCP for TelePresence calls**. When a Cisco TelePresence endpoint registers with Unified CM, it downloads a configuration file and applies the QoS setting of **DSCP for TelePresence calls**. When a Unified Communications video-capable endpoint registers with Unified CM, it downloads a configuration file and applies the QoS setting of **DSCP for Video calls**. All third-party video endpoints require manual configuration of the endpoints themselves and are statically configured, meaning they do not change QoS marking depending on the call type; therefore, it is important to match the Enhanced Location CAC bandwidth allocation to the correct DSCP. Unified CM achieves this by deducting desktop video calls from the Video Bandwidth location and link allocation for devices that have a Video Call Traffic Class of **desktop**. End-to-end TelePresence immersive video calls are deducted from the Immersive Video Bandwidth location and link allocation for devices or trunks with the Video Call Traffic Class of **immersive**. This ensures that end-to-end desktop video and immersive video calls are marked correctly and counted correctly for call admission control. For calls between desktop devices and TelePresence immersive devices, bandwidth is deducted from both the Video Bandwidth and the Immersive Video Bandwidth location and link allocations.

Endpoint Classification

Cisco TelePresence endpoints have a fixed non-configurable Video Call Traffic Class of **immersive** and are identified by Unified CM as immersive. Telepresence endpoints are defined in Unified CM by the device type. When a device is added in Unified CM, any device with TelePresence in the name of the device type has a fixed Video Call Traffic Class of **immersive**. The generic single-screen and multi-screen room systems are also classified as immersive. Another way to check the capabilities of the

endpoints in the Unified CM is to go to the Cisco Unified Reporting Tool and select **Immersive Video Support for TelePresence Devices**. This will display all of the devices that are configured in Unified CM with the video call traffic class attribute of **immersive**.

All other endpoints have a fixed Video Call Traffic Class of **desktop**.

Bandwidth reservations are determined by the classification of endpoints in a video call, and they deduct bandwidth from the locations and links bandwidth pools as listed in [Table 13-4](#).

Table 13-4 Bandwidth Pool Usage per Endpoint Type

Endpoint A	Endpoint B	Locations and Links Pool Used
Immersive video	Immersive video	Immersive bandwidth
Immersive video	Desktop video	Immersive and video bandwidth
Desktop video	Desktop video	Video bandwidth
Audio-only call	Any	Audio bandwidth

SIP Trunks

A SIP trunk can also be classified as desktop, immersive, or mixed video in order to deduct bandwidth reservations of a SIP trunk call, and the classification is determined by the calling device type and Video Call Traffic Class of the SIP trunk. The SIP trunk can be configured through the SIP Profile trunk-specific information as:

- Immersive — High-definition immersive video
- Desktop — Standard desktop video
- Mixed — A mix of immersive and desktop video

A SIP trunk can be classified with any of these three classifications and is used primarily to classify Video or TelePresence Multipoint Control Units (MCUs), a video device at a fixed location, a Unified Communications system such as Unified CM prior to version 9.0, or a Cisco TelePresence System Video Communications Server (VCS).

Bandwidth reservations are determined by the classification of an endpoint and a SIP trunk in a video call, and they deduct bandwidth from the locations and links bandwidth pools as listed in [Table 13-5](#).

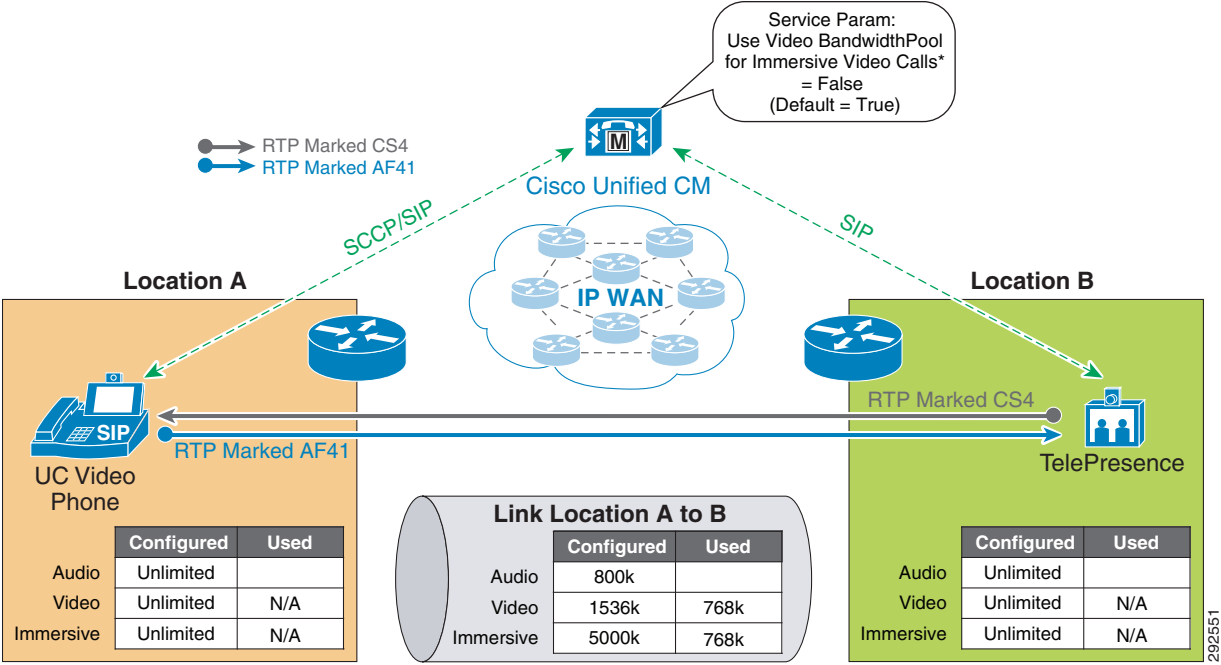
Table 13-5 Bandwidth Pool Usage per SIP Trunk and Endpoint Type

Endpoint	SIP Trunk	Locations and Links Pool Used
TelePresence endpoint	Immersive	Immersive bandwidth
TelePresence endpoint	Desktop	Immersive and video bandwidth
TelePresence endpoint	Mixed	Immersive and video bandwidth
Desktop endpoint	Immersive	Immersive and video bandwidth
Desktop endpoint	Desktop	Video bandwidth
Desktop endpoint	Mixed	Immersive and video bandwidth
Non-video endpoint	Any	Audio bandwidth

By default, all video calls from either immersive or desktop endpoints are deducted from the locations and links video bandwidth pool. To change this behavior, set the Unified CM service parameter **Use Video BandwidthPool for Immersive Video Calls** to **False**, and this will enable the immersive video bandwidth deductions. After this is enabled, immersive and desktop video calls will be deducted out of their respective pools.

As described earlier, a video call between a Unified Communications video endpoint (desktop Video Call Traffic Class) and a TelePresence endpoint (immersive Video Call Traffic Class) will mark their media asymmetrically and, when immersive video CAC is enabled, will deduct bandwidth from both video and immersive locations and links bandwidth pools. Figure 13-11 illustrates this.

Figure 13-11 Enhanced Location CAC Bandwidth Deductions and Media Marking for a Multi-Site Deployment



Examples of Various Call Flows and Location and Link Bandwidth Pool Deductions

The following call flows depict the expected behavior of locations and links bandwidth deductions when the Unified CM service parameter **Use Video BandwidthPool for Immersive Video Calls** is set to **False**.

Figure 13-12 illustrates an end-to-end TelePresence immersive video call between TP-A in Location L1 and TP-B in Location L2. End-to-end immersive video endpoint calls deduct bandwidth from the immersive bandwidth pool of the locations and the links along the effective path.

Figure 13-12 Call Flow for End-to-End TelePresence Immersive Video

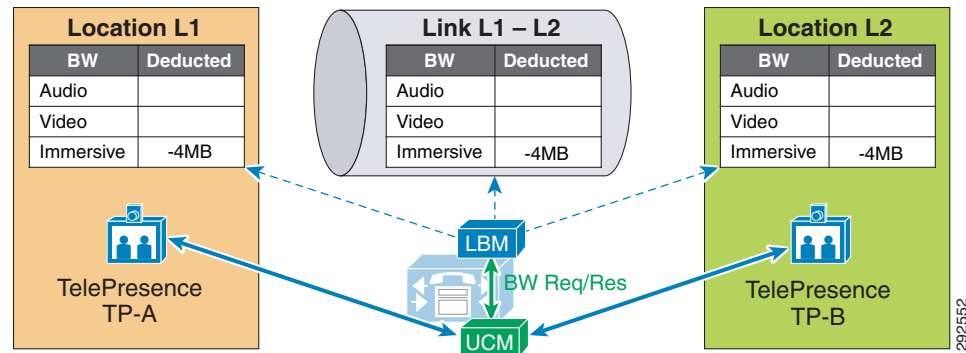


Figure 13-13 illustrates an end-to-end desktop video call between DP-A in Location L1 and DP-B in Location L2. End-to-end desktop video endpoint calls deduct bandwidth from the video bandwidth pool of the locations and the links along the effective path.

Figure 13-13 Call Flow for End-to-End Desktop Video

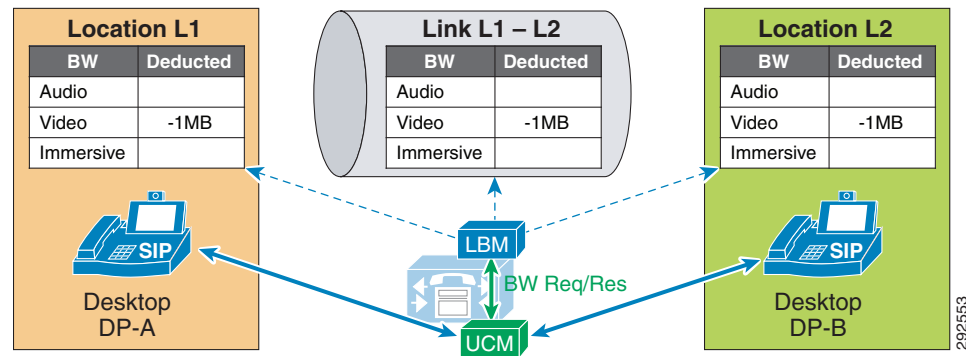
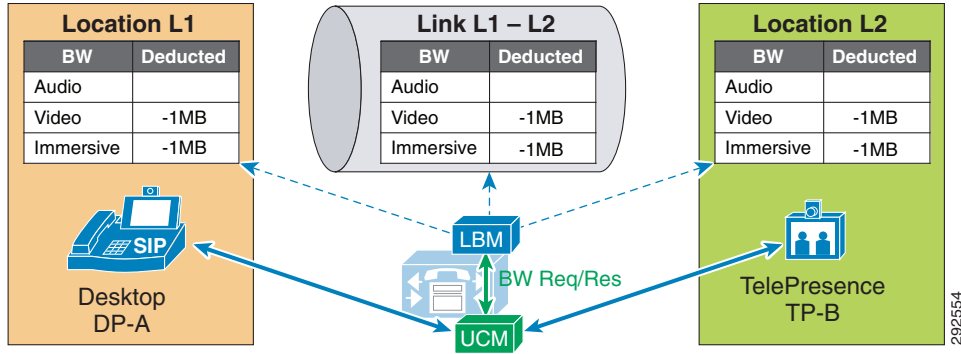


Figure 13-14 illustrates a video call between desktop video endpoint DP-A in Location L1 and TelePresence video endpoint TP-B in Location L2. Interoperating calls between desktop video endpoints and TelePresence video endpoints deduct bandwidth from both video and immersive locations and the links bandwidth pools along the effective path.

Figure 13-14 Call Flow for Desktop-to-TelePresence Video



In Figure 13-15, a desktop video endpoint and two TelePresence endpoints call a SIP trunk configured with a Video Traffic Class of **immersive** that points to a TelePresence MCU. Bandwidth is deducted along the effective path from the immersive locations and the links bandwidth pools for the calls that are end-to-end immersive and from both video and immersive locations and the links bandwidth pools for the call that is desktop-to-immersive.

Figure 13-15 Call Flow for a Video Conference with an MCU

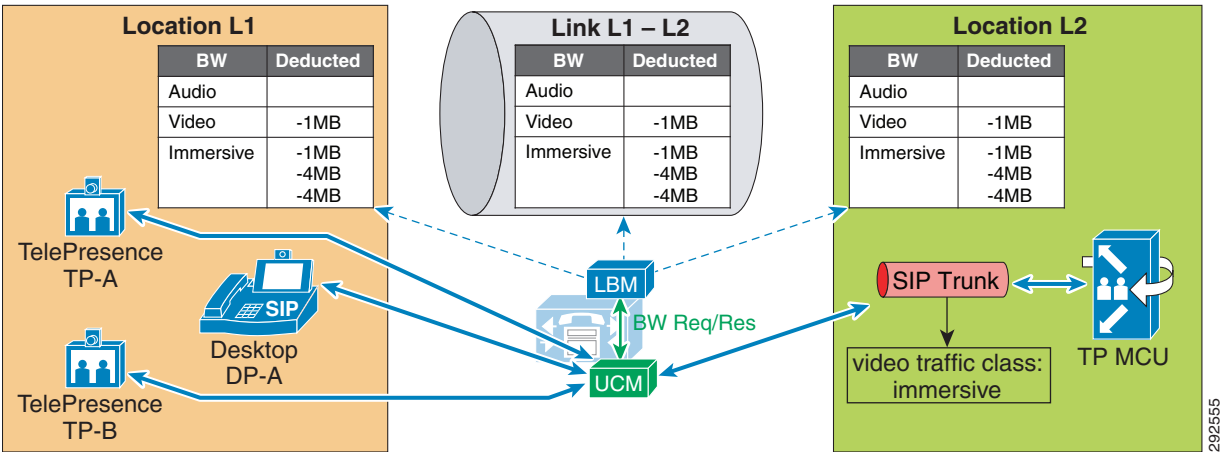


Figure 13-16 illustrates an end-to-end immersive video call across clusters, which deduces bandwidth from the immersive bandwidth pool of the locations and links along the effective path.

Figure 13-16 Call Flow for End-to-End TelePresence Immersive Video Across Clusters

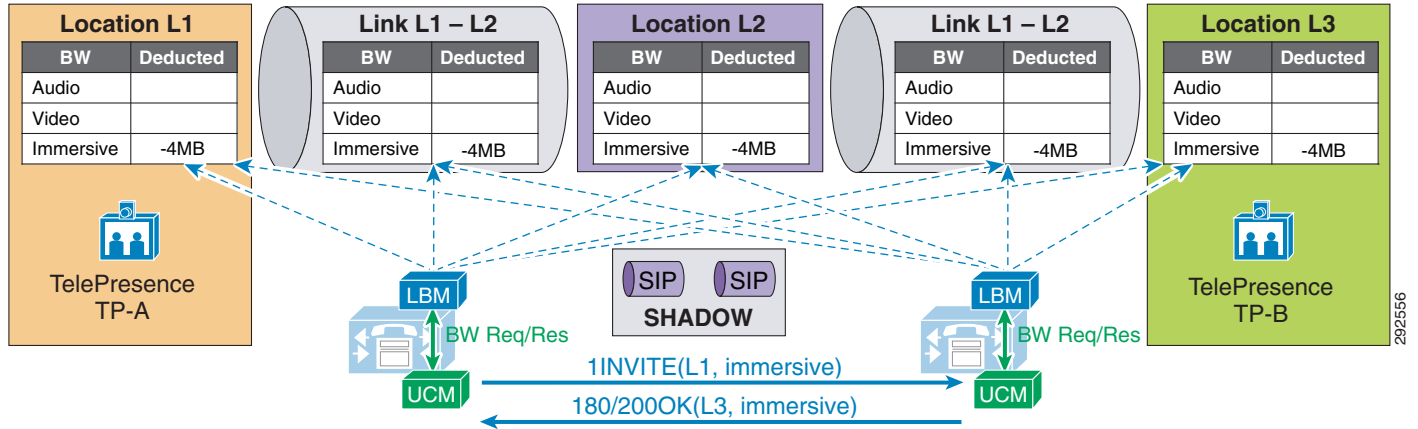


Figure 13-17 illustrates an end-to-end desktop video call across clusters, which deduces bandwidth from the video bandwidth pool of the locations and links along the effective path.

Figure 13-17 Call Flow for End-to-End Desktop Video Call Across Clusters

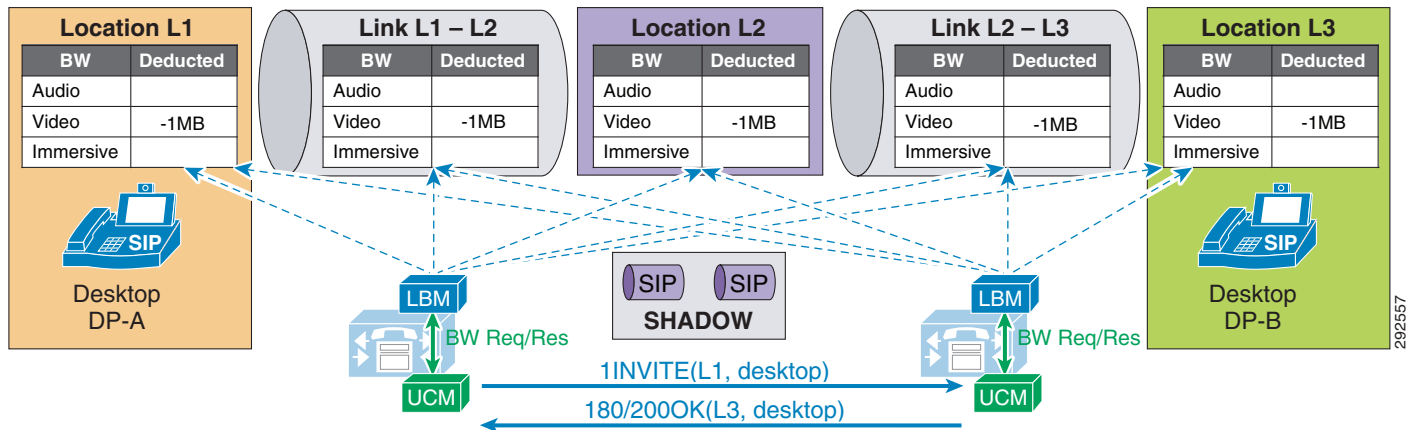
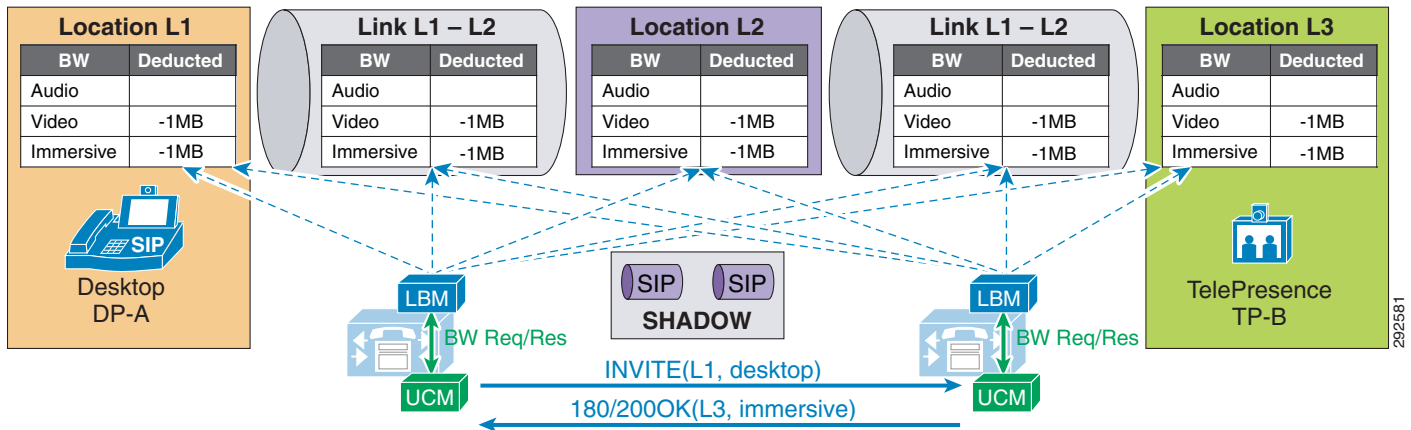


Figure 13-18 illustrates a desktop video endpoint calling a TelePresence endpoint across clusters. the call deducts bandwidth from both video and immersive bandwidth pools of the locations and links along the effective path.

Figure 13-18 Call Flow for Desktop-to-TelePresence Video Across Clusters



Video Bandwidth Utilization and Admission Control

When Unified CM negotiates an audio or video call, a number of separate streams are established between the endpoints involved in the call. For video calls with content sharing, this can result in as many as 8 (or possibly more) unidirectional streams. For an audio-only call typically the bare minimum is 2 streams, one in each direction. This section discusses bandwidth utilization on the network and how Unified CM accounts for this in admission control bandwidth accounting.

For the purpose of the discussion in this section, please note the following:

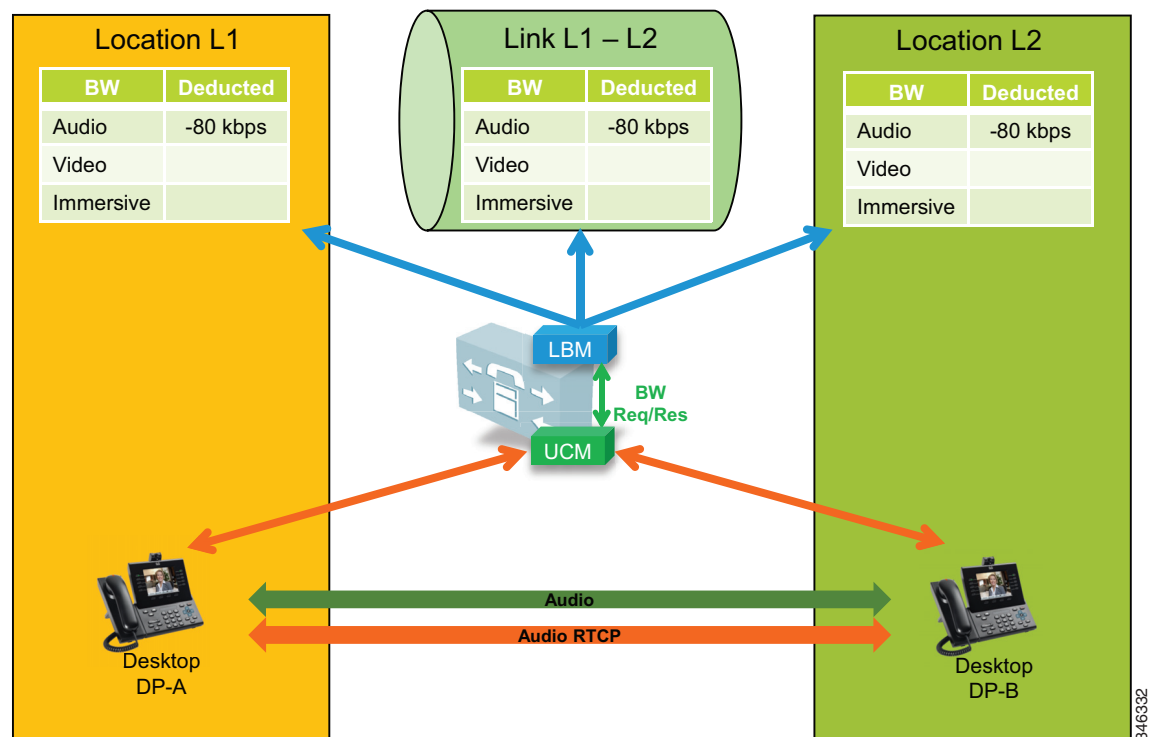
- The figures in this section use a bidirectional arrow (<-->) to represent two unidirectional streams.
- The following points summarize how Unified CM Enhanced Location CAC deducts bandwidth from the configured audio, video, and immersive allocations. For more information, see the section on [Locations and Links](#), page 13-5:
 - Audio (audio-only calls): RTP bit rate + IP and UDP header overhead
 - Video (video calls): RTP bit rate only
 - Immersive (video calls by Cisco TelePresence endpoints): RTP bit rate only
- Bandwidth deductions in Enhanced Location CAC:
 - Bandwidth deductions are made for bidirectional RTP streams and are assumed to be symmetrically routed (both streams routed over the same path). For example, a G.711 audio call of 80 kbps is 80 kbps in each direction over a full-duplex network; that is 80 kbps on the transmit pair of wires and 80 kbps on the receive pair of wires, equating to 80 kbps full-duplex. (See [Figure 13-19](#).) Note that traffic is not always routed symmetrically in the WAN. Check with your network administrator when necessary to ensure that admission control is correctly accounting for the media as it is routed in the network over the WAN.

- Real-Time Transport Control Protocol (RTCP) bandwidth overhead is not part of Unified CM bandwidth allocation and should be part of network provisioning. RTCP is quite common in most call flows and is commonly used for statistical information about the streams. It is also used to synchronize audio in video calls to ensure proper lip-sync. In some cases it can be enabled or disabled on the endpoint. RFC 3550 recommends that the fraction of the session bandwidth added for RTCP should be fixed at 5%. What this means is that it is common practice for the RTCP session to be up to 5% of the associated RTP session. So when calculating bandwidth consumption on the network, you should add the RTCP overhead for each RTP session. For example, if you have a G.711 audio call of 80 kbps with RTCP enabled, you will be using up to 84 kbps per session (4 kbps RTCP overhead) for both RTP and RTCP. This calculation is not part of Enhanced Location CAC deductions but should be part of network provisioning.

**Note**

There are, however, methods to re-mark this traffic to another Differentiated Services Code Point (DSCP). For example, RTCP uses odd-numbered UDP ports while RTP uses even-numbered UDP ports. Therefore, classification based on UDP port ranges is possible. Network-Based Application Recognition (NBAR) is another option as it allows for classification and re-marking based on the RTP header **Payload Type** field. For more information on NBAR, see <http://www.cisco.com>. However, if the endpoint marking is trusted in the network, then RTCP overhead should be provisioned in the network within the same QoS class as audio RTP (default marking is EF). It should also be noted that RTCP is marked by the endpoint with the same marking as RTP; by default this is EF (verify that RTCP is also marked as EF).

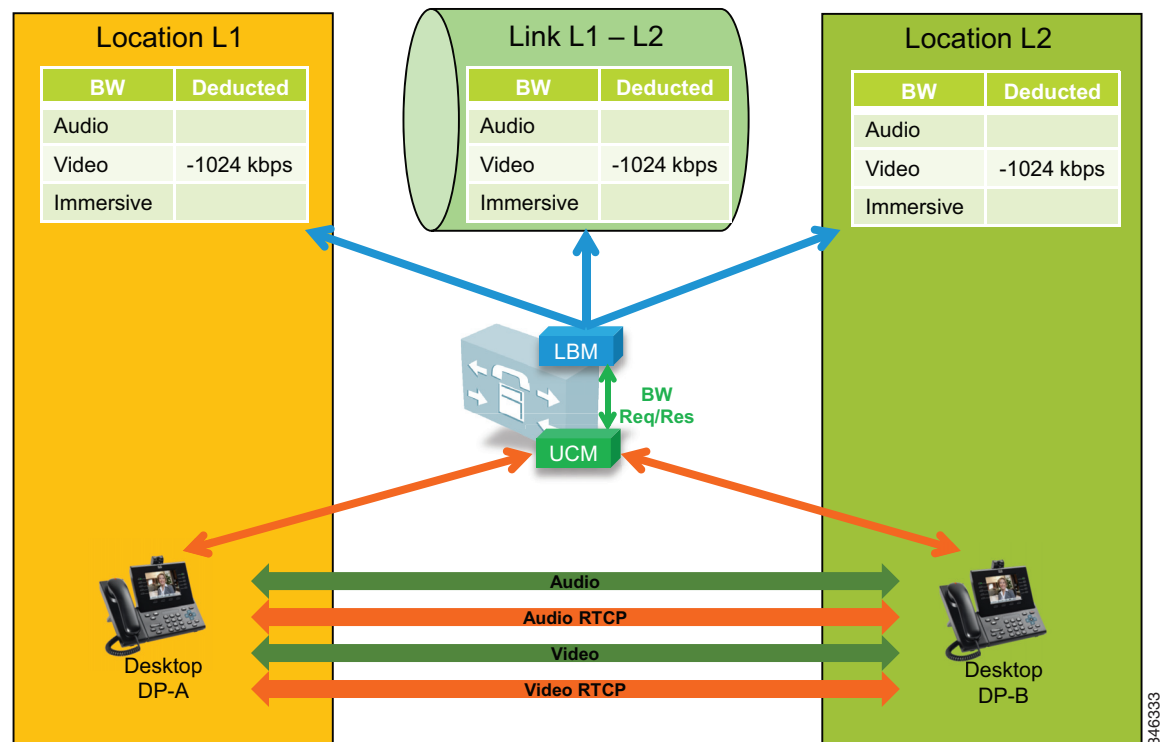
Figure 13-19 A Basic Audio-Only Call with RTCP Enabled



In [Figure 13-19](#) two desktop video phones have established an audio-only call. In this call flow four streams are negotiated: two audio streams illustrated by a single bidirectional arrow and two RTCP streams also illustrated by a bidirectional arrow. For this call, the Location Bandwidth Manager (LBM) deducts 80 kbps (bit rate + IP/UDP overhead) between location L1 and location L2 for a call established between desktop phones DP-A and DP-B. The actual bandwidth consumed at Layer 3 in the network with RTCP enabled would be between 80 kbps and 84 kbps, as discussed previously in this section.

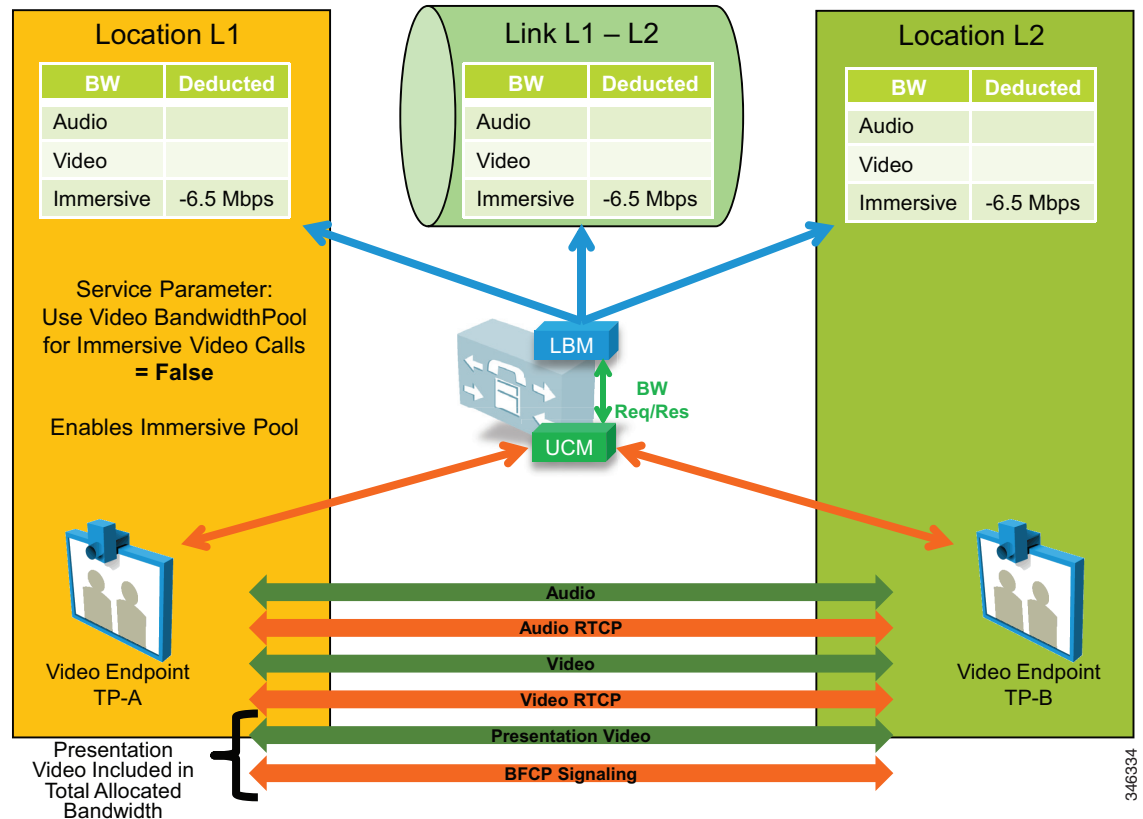
In [Figure 13-20](#) two desktop video phones have established a video call. In this call flow eight streams are negotiated: two audio streams, two audio-associated RTCP streams, two video streams, and two video-associated RTCP streams. Again for this illustration one bidirectional arrow is used to depict two unidirectional streams. This particular call is 1024 kbps, with 64 kbps of G.711 audio and 960 kbps of video (bit rate only for audio and video allocations of video calls). So in this case the LBM deducts 1024 kbps between locations L1 and L2 for a video call established between desktop phones DP-A and DP-B. RTCP is overhead that should be accounted for in provisioning, depending on how it is marked or re-marked by the network.

Figure 13-20 A Basic Video Call with RTCP Enabled



The example in [Figure 13-21](#) is of a video call with presentation sharing. This is a more complex call with regard to the number of associated streams and bandwidth allocation when compared to bandwidth used on the network, and therefore it must be provisioned in the network. [Figure 13-21](#) illustrates a video call with RTCP enabled and Binary Flow Control Protocol (BFCP) enabled for presentation sharing. All SIP-enabled telepresence multipurpose or personal endpoints such as the Cisco TelePresence System EX, MX, SX, C, and Profile Series function in the same manner.

Figure 13-21 *Video Call with RTCP and BFCP Enabled and Presentation Sharing*



When a video call is established between two video endpoints, audio and video streams are established and bandwidth is deducted for the negotiated rate. Unified CM uses regions to determine the maximum bit rate for the call. For example, with a Cisco TelePresence System EX90 at the highest detail of 1080p at 30 frames per second (fps), the negotiated rate between regions would have to be set at 6.5 Mbps. EX90s used in this scenario would average around 6.1 Mbps for this session. When the endpoints start presentation sharing during the session, BFCP is negotiated between the endpoints and a new video stream is enabled at either 5 fps or 30 fps, depending on endpoint configuration. When this occurs, the endpoints will throttle down their main video stream to include the presentation video so that the entire session does not use more than the allocated bandwidth of 6.5 Mbps. Thus, the average bandwidth consumption remains the same with or without presentation sharing.



Note

The presentation video stream is typically unidirectional in the direction of the person or persons viewing the presentation.

Telepresence immersive and office endpoints such as the Cisco TelePresence System 500, 1000, 3000, and TX9000 Series that negotiate a call between one another function a little differently in the sense that the video for presentation sharing is an additional bandwidth above and beyond what is allocated for the main video session, and thus it is not deducted from Enhanced Location CAC. [Figure 13-22](#) illustrates this.

Location L1

BW	Deducted
Audio	
Video	
Immersive	-4 Mbps

Service Parameter:
Use Video BandwidthPool
for Immersive Video Calls
= **False**

Enables Immersive Pool

Link L1 - L2

BW	Deducted
Audio	
Video	
Immersive	-4 Mbps

Location L2

BW	Deducted
Audio	
Video	
Immersive	-4 Mbps

LBM (Link Bandwidth Manager) and **UCM** (User Call Manager) manage the bandwidth.

Video Endpoint TP-A and **Video Endpoint TP-B** are connected to the LBM/UCM.

Bandwidth Flow:

- Blue arrows: L1 and L2 send bandwidth requests to LBM/UCM.
- Green arrows: LBM/UCM sends bandwidth to L1 and L2.
- Red arrows: L1 and L2 send bandwidth to Video Endpoints.
- Green arrows: Video Endpoints send Presentation Video to each other.
- Red arrows: Video Endpoints send BFCP Signaling to each other.

346335

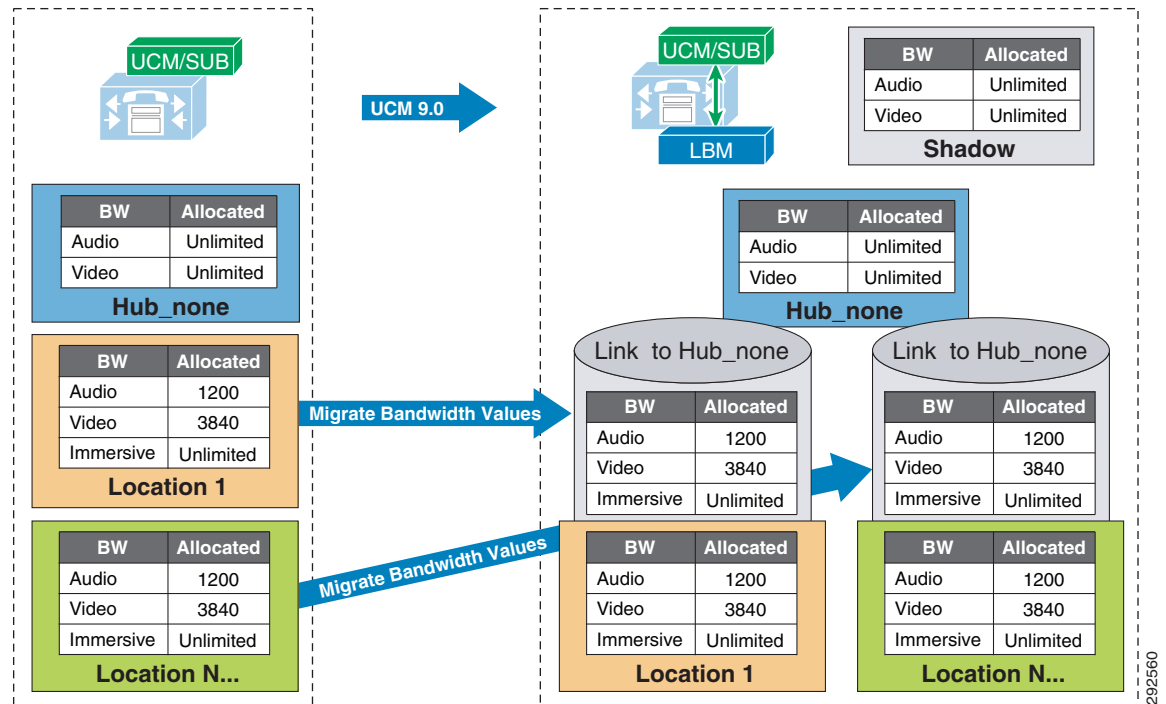

Note

346335

Upgrade and Migration from Location CAC to Enhanced Location CAC

Upgrading to Cisco Unified CM 9.x from a previous release will result in the migration of Location CAC to Enhanced Location CAC. The migration consists of taking all previously defined locations bandwidth limits of audio and video bandwidth and migrating them to a link between the user-defined location and Hub_None. This effectively recreates the hub-and-spoke model that previous versions of Unified CM Location CAC supported. Figure 13-23 illustrates the migration of bandwidth information.

Figure 13-23 Migration from Location CAC to Enhanced Location CAC After Unified CM Upgrade



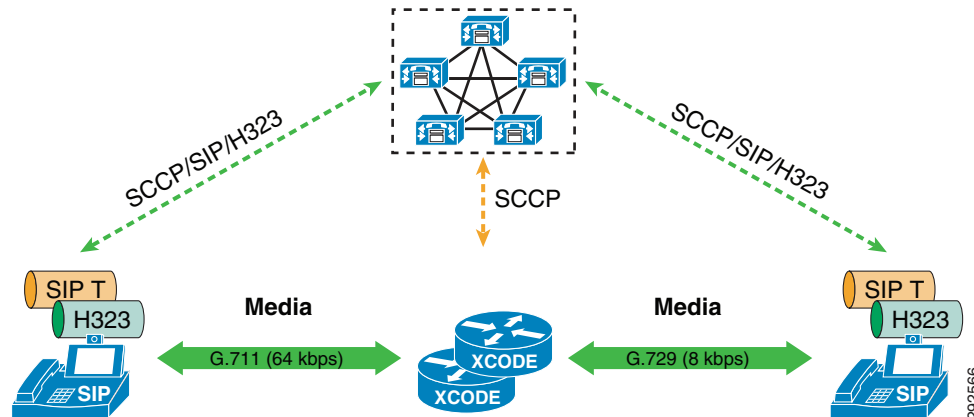
Settings after an upgrade to Cisco Unified CM 9.x:

- The LBM is activated on each Unified CM subscriber running the Cisco CallManager service.
- The Cisco CallManager service communicates directly with the local LBM.
- No LBM group or LBM hub group is created.
- All LBM services are fully meshed.
- Intercluster Enhanced Location CAC is not enabled.
- All intra-location bandwidth values are set to unlimited.
- Bandwidth values assigned to locations are migrated to a link connecting the user-defined location and Hub_None.
- Immersive bandwidth is set to unlimited.
- A shadow location is created.

- Phantom and shadow locations have no links.
- Enhanced Location CAC bandwidth adjustment for MTPs and transcoders:

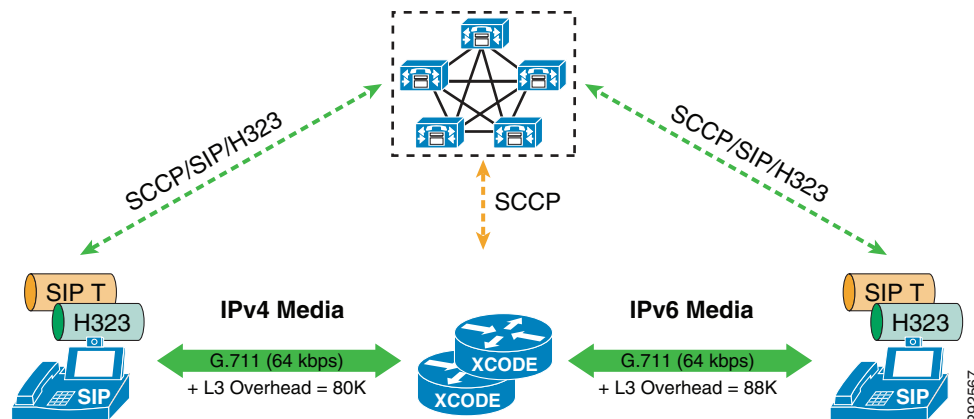
For transcoding insertion, the bit rate is different on each leg of the connection. [Figure 13-24](#) illustrates this.

Figure 13-24 Example of Different Bit Rate for Transcoding



For dual stack MTP insertion, the bit rate is different on each connection but the bandwidth is different due to IP header overhead. [Figure 13-25](#) illustrates the difference in bandwidth used for IPv4 and IPv6 networks with dual stack MTP insertion.

Figure 13-25 Bandwidth Differences for Dual Stack MTP Insertion



Enhanced Location CAC does not account for these differences in bandwidth between MTPs and transcoders. The service parameter **Locations Media Resource Audio Bit Rate Policy** determines whether the largest or smallest bandwidths should be used along the locations and links path. Lowest Bit Rate (default) or Highest Bit Rate can be used to manage these differences in bandwidth consumption.

Design Considerations for Call Admission Control

This section describes how to apply the call admission control mechanisms to various IP WAN topologies. With Unified CM Enhanced Location CAC network modeling support, Unified CM is no longer limited to supporting simple hub-and-spoke or MPLS topologies but, together with intercluster enhanced locations, can now support most any network topology in any Unified CM deployment model. Enhanced Location CAC is still a statically defined mechanism that does not query the network, and therefore the administrator still needs to provision Unified CM accordingly whenever network changes affect admission control. This is where a network-aware mechanism such as RSVP can fill that gap and provide support for dynamic changes in the network, such as when network failures occur and media streams take different paths in the network. This is often the case in designs with load-balanced dual or multi-homed WAN uplinks or unequally sized primary and backup WAN uplinks.

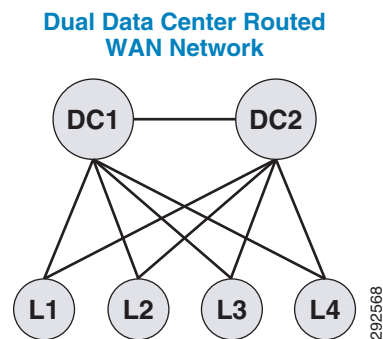
To learn how Enhanced Location CAC functions and how to design and deploy Enhanced location CAC, see the section on [Unified CM Enhanced Location Call Admission Control](#), page 13-4.

In this section explores a few typical topologies and explains how Enhanced Location CAC can be designed to manage them.

Dual Data Center Design

Figure 13-26 illustrates a simple dual data center WAN network design where each remote site has a single WAN uplink to each data center. The data centers are interconnected by a high-speed WAN connection that is over-provisioned for data traffic.

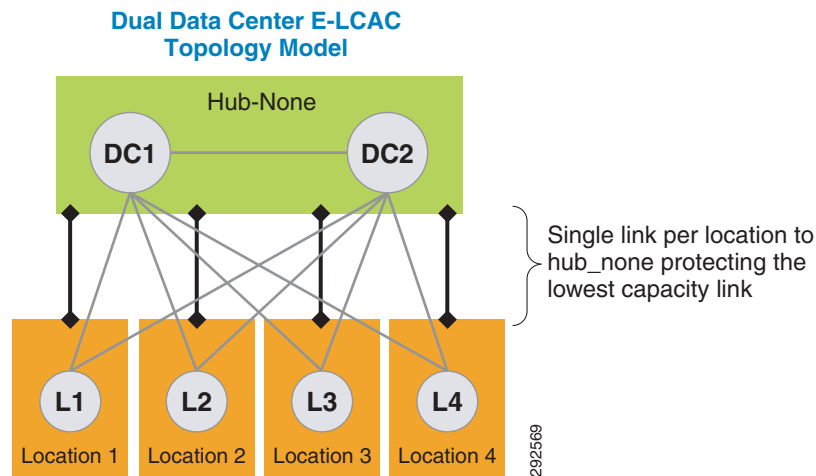
Figure 13-26 Dual Data Center WAN Network



Typically these WAN uplinks from the remote sites to the data centers are load-balanced or in a primary/backup configuration, and there are limited ways for a static CAC mechanism to handle these scenarios. Although you could configure this multi-path topology in Enhanced Location CAC, only one path would be calculated as the effective path and would remain statically so until the weight metric was

changed. A better way to support this type of network topology is to configure the two data centers as one data center or hub location in Enhanced Location CAC and configure a single link to each remote site location. Figure 13-27 illustrates an Enhanced Location (E-L) CAC locations and links overlay.

Figure 13-27 *Enhanced Location CAC Topology Model for Dual Data Centers*



Design Recommendations

The following design recommendations for dual data centers with remote dual or more links to remote locations apply to both load-balanced and primary/backup WAN designs:

- A single location (Hub_None) represents both data centers.
- A single link between the remote locations and Hub_None protects the remote site uplinks from over-subscription during normal conditions or failure of the highest bandwidth capacity links.
- The capacity of link bandwidth allocation between the remote site and Hub_None should be equal to the lowest bandwidth capacity for the applicable Unified Communications media for a single link. For example, if each WAN uplink can support 2 Mbps of audio traffic marked EF, then the link audio bandwidth value should be no more than 2 Mbps to support a failure condition or equal-cost path routing.

MPLS Clouds

When designing for Multiprotocol Label Switching (MPLS) any-to-any connectivity type clouds in the E-L CAC network model, a single location can serve as the MPLS cloud. This location will not have any devices associated to it, but all of the sites that have uplinks to this cloud will have links configured to the location. In this way the MPLS cloud serves as a transit location for interconnecting multiple variable-sized bandwidth WAN uplinks to other remote locations. The illustrations in this section depict a number of different MPLS networks and their equivalent locations and links model.

In Figure 13-28, Hub_None represents the MPLS cloud serving as a transit location interconnecting the campus location where servers, endpoints, and devices are located, with remote locations where only endpoints and devices are located. Each link to Hub_None from the remote location may be sized according to the WAN uplink bandwidth allocated for audio, video, and immersive media.

Figure 13-28 Single MPLS Cloud

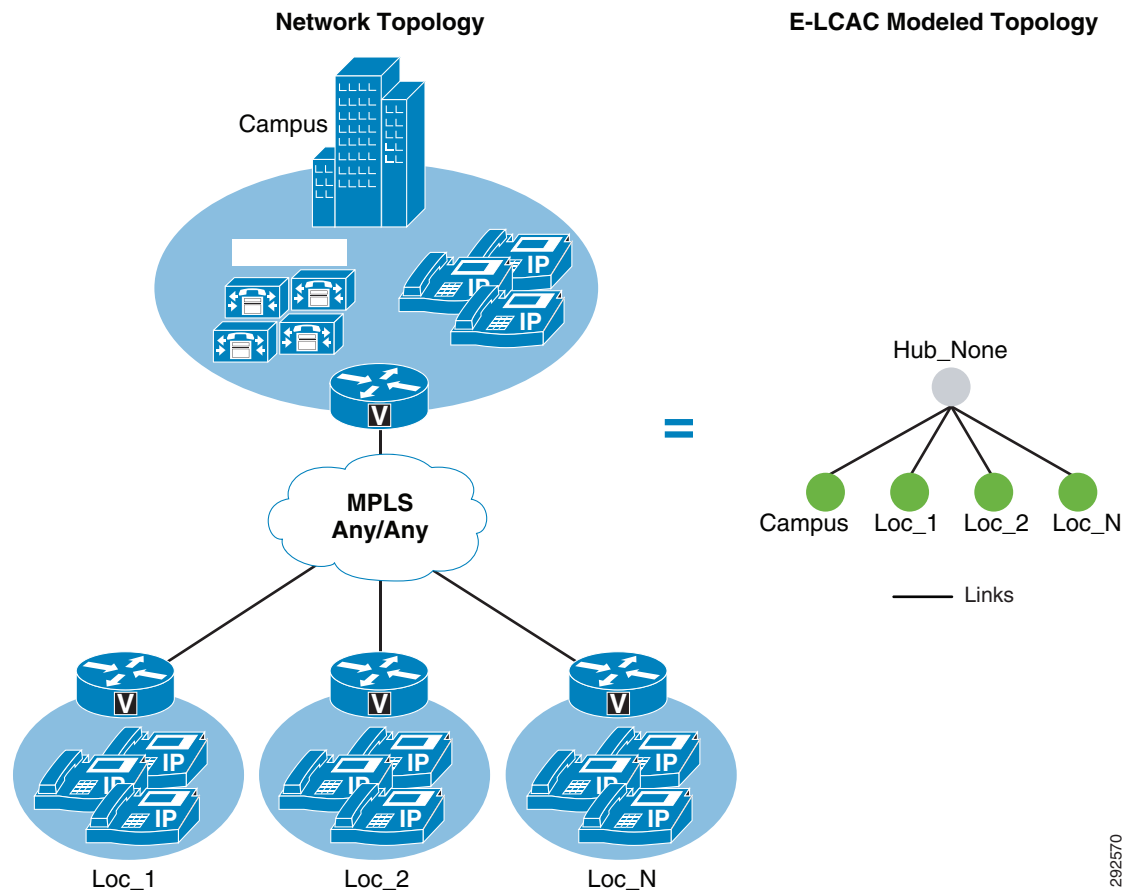


Figure 13-29 shows two MPLS clouds that serve as transit locations interconnecting the campus location where servers, endpoints, and devices are located, with remote locations where only endpoints and devices are located. The campus also connects to both clouds. Each link to the MPLS cloud from the remote location may be sized according to the WAN uplink bandwidth allocated for audio, video, and immersive media. This design is typical in enterprises that span continents, with a separate MPLS cloud from different providers in each geographical location.

292570

Figure 13-29 Separate MPLS Clouds

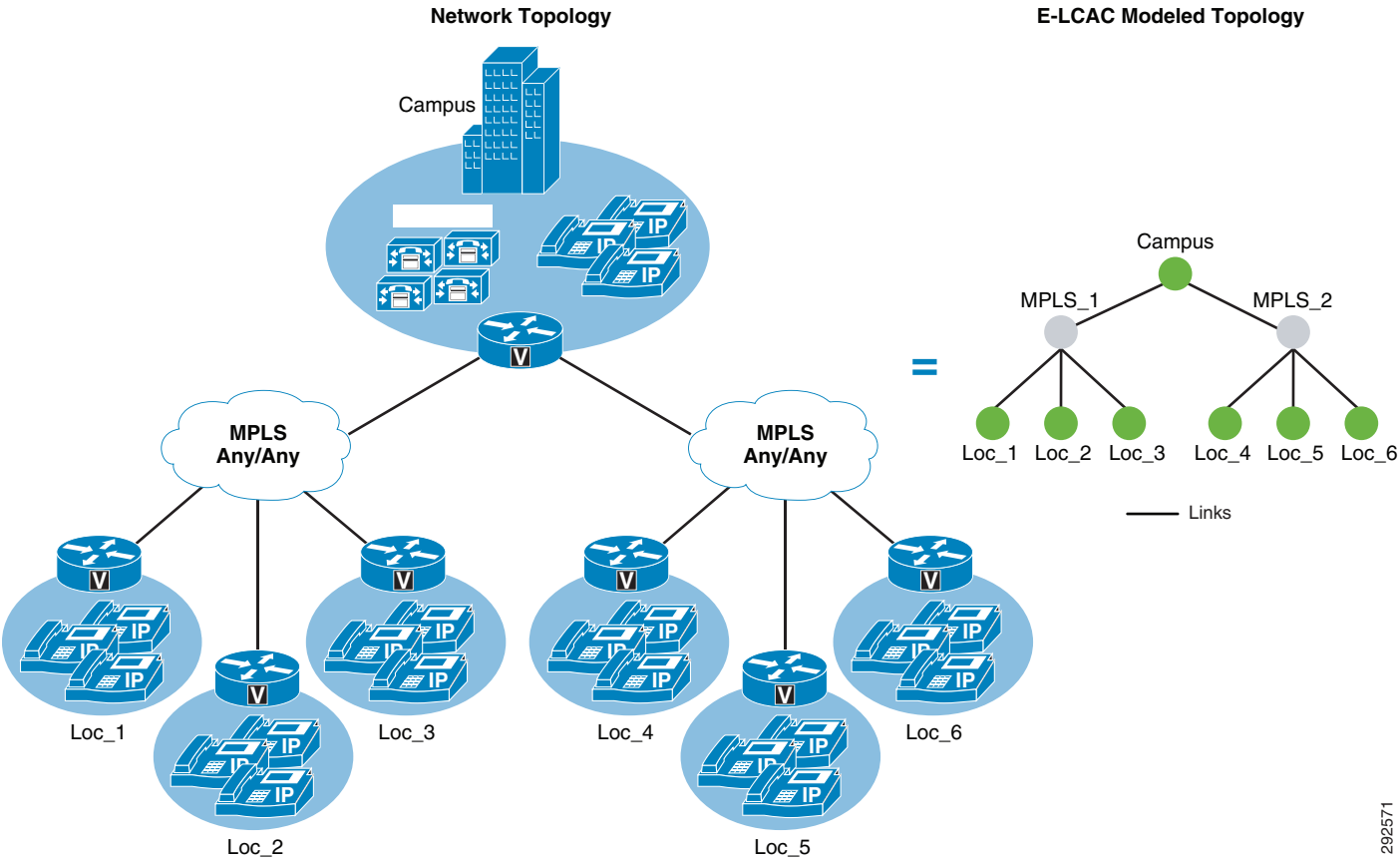
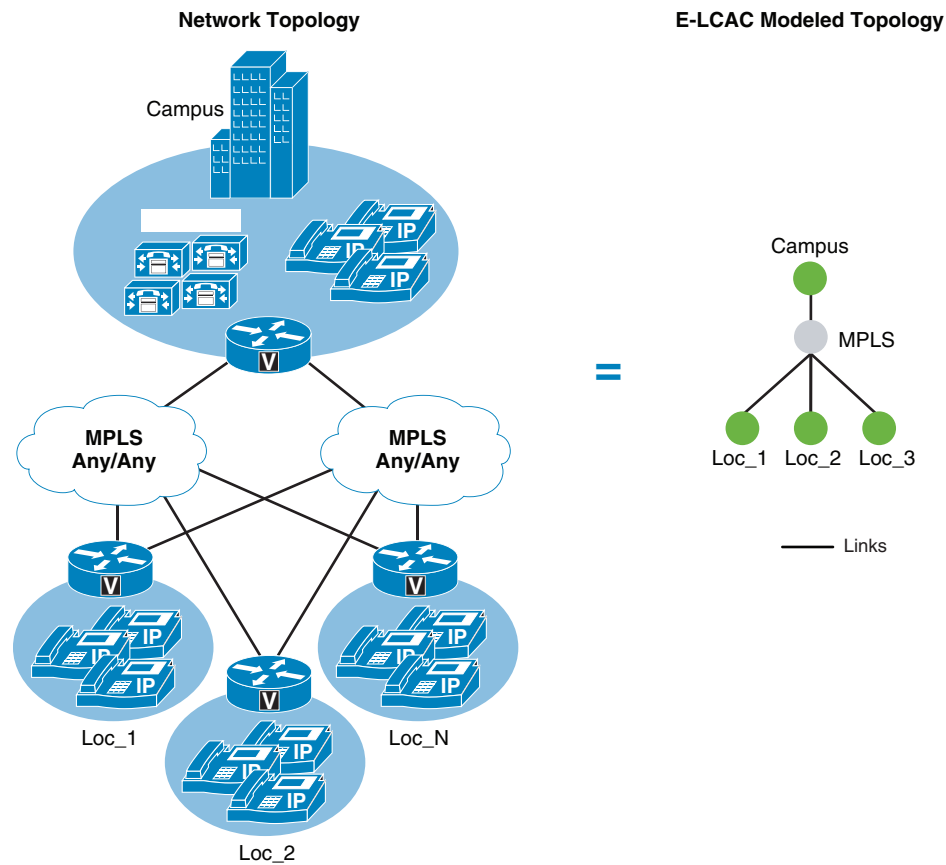


Figure 13-30 shows multiple MPLS clouds from different providers, where each site has one connection to each cloud and uses the MPLS clouds in either an equal-cost load-balanced manner or in a primary/backup scenario. In any case, this design is equivalent to the dual data center design where a single location represents both clouds and a single link represents the lowest capacity link of the two.

292571

Figure 13-30 Remote Sites Connected to Dual MPLS Clouds

292572

Design Recommendations

- The MPLS cloud should be configured as a location that does not contain any endpoints but is used as a hub to interconnect locations.
- The MPLS cloud serves as a transit location for interconnecting multiple variable-sized bandwidth WAN uplinks to other remote locations.
- Remote sites with connectivity to dual MPLS clouds should treat those connections as a single link and size to the lowest capacity of the links in order to avoid oversubscription during network failure conditions.

Call Admission Control Design Recommendations for TelePresence Video Interoperability Architectures

Video interoperability refers to the support for point-to-point video calls between Cisco TelePresence endpoints, Cisco Unified Communications video endpoints, and third-party video endpoints without requiring a Multipoint Control Unit (MCU). This section discusses the features in Enhanced Location CAC and the design considerations and recommendations applicable to Quality of Service (QoS) for interoperable video calls.

This section explains how to supplement the design to allow for video interoperation between Unified Communications and TelePresence endpoints. For information on Unified Communications endpoints, refer to the chapter on [Collaboration Endpoints](#), page 8-1; and for information on TelePresence endpoints, refer to the Cisco TelePresence Endpoints product documentation at <http://www.cisco.com/en/US/products/ps7060/index.html>.



Note

Third-party video endpoints should follow the same guidelines and recommendations as Cisco Unified Communications endpoints. Throughout this section, the term *UC endpoints* is used to refer to both third-party endpoints and Cisco Unified Communications Endpoints.

Starting with Cisco Unified CM 9.0, the Cisco TelePresence solution provides the ability to reserve network bandwidth and perform admission control for TelePresence calls separately from other video calls. It is important to be familiar with Enhanced Location Call Admission Control (CAC) prior to designing TelePresence video interoperability. This section addresses Enhanced Location CAC with regard to TelePresence video interoperability.

Additionally, TelePresence video interoperability in Unified CM enables Cisco Telepresence System (CTS) endpoints to communicate with non-CTS endpoints, provided that the installed CTS software supports such interoperability. For further information, refer to the document on *Interoperability Between CTS Endpoints and Other Cisco Endpoints or Devices*, available at

http://www.cisco.com/en/US/docs/telepresence/interop/endpoint_interop.html

Supported CAC Deployment Scenarios and Design Considerations

The design considerations for TelePresence video interoperability CAC are based on the following main deployment scenarios:

- **Mixed Single Cluster** — Mixed UC video endpoints and TelePresence endpoints registered to a single cluster
This is a single-cluster design where TelePresence and UC video endpoints are registered to the same cluster. The call processing deployment model can be any of the single-cluster designs such as clustering over the WAN, multi-site centralized call processing, or single campus designs.
- **Dedicated Multi-Cluster** — UC video endpoints and TelePresence endpoints on separate dedicated clusters

This is a multi-cluster design where the TelePresence endpoints are registered to a different cluster than the UC video endpoints. The call processing deployment model can be multi-site centralized or multi-site distributed cluster designs.

- **Mixed Multi-Cluster** — UC video endpoints and TelePresence endpoints mixed in multi-cluster distributed deployments

This is a multi-cluster design where TelePresence endpoints are spread across multiple clusters serviced by a single Cisco TelePresence System Manager (CTS-Manager). There are also deployment scenarios where some TelePresence endpoints can be co-located with the same Unified CM cluster as the UC endpoints (mixed single cluster model) while other TelePresence endpoints are registered to a dedicated cluster, and both clusters are serviced by a single CTS-Manager for TelePresence. The call processing deployment model can be multi-site, multi-cluster centralized, or multi-site/multi-cluster distributed designs. This is a hybrid model that combines aspects of the dedicated multi-cluster model with mixed single-cluster model where all endpoints are registered to the same cluster.

Cisco Unified Communications Manager Session Management Edition (SME) is also supported in the multi-cluster models. However, because Session Management Edition is a variation of the multi-site distributed call processing deployment model and is typically employed to interconnect large numbers of Unified Communications systems through a single front-end system, there are no specific guidelines for it with regard to TelePresence video interoperability.

Enhanced Location CAC Design Considerations and Recommendations

When designing Enhanced Location (E-L) CAC for TelePresence Video Interoperability, follow the design recommendations and considerations listed in this section.

Design Recommendations

The following design recommendations apply to TelePresence video interoperability solutions that employ Enhanced Location (E-L) CAC:

- E-L CAC for TelePresence video interoperability is supported in all three deployment models: mixed single cluster, dedicated multi-cluster, and mixed multi-cluster.
- If you are deploying Unified Communications video and TelePresence video interoperability where differentiation between desktop video and TelePresence video is a requirement, then ensure that the Unified CM service parameter **Use Video Bandwidth Pool for Immersive Video Calls** is set to **false**. This enables the immersive bandwidth pool for TelePresence calls.
- In E-L CAC TelePresence endpoints can be managed in the same location as Unified Communications video endpoints. If TelePresence calls are not to be tracked through E-L CAC, then set the immersive location and links bandwidth pool to **unlimited**. This will ensure that CAC will not be performed on TelePresence or SIP trunks classified as immersive. If TelePresence calls are to be tracked through E-L CAC, then set immersive location and links bandwidth pool to a value according to the bit rate and number of calls to be allowed over the locations and link paths.
- Intercluster SIP trunks should be associated with the shadow location.
- Cisco Unified CM uses two different cluster-wide QoS service parameter to differentiate between the Differentiated Services Code Point (DSCP) settings of UC video endpoints and TelePresence endpoints. TelePresence endpoints use the **DSCP for Telepresence calls** QoS parameter while the Cisco UC video endpoints use the **DSCP for video calls** QoS service parameter.
- For sites that deploy only UC endpoints and no TelePresence endpoints, ensure that the CS4 DSCP class is added to the AF41 QoS traffic class on inbound WAN QoS configurations to account for the inbound CS4 marked traffic, thus ensuring QoS treatment of CS4 marked media.

- For sites that deploy only UC TelePresence endpoints and no UC endpoints, ensure that the AF41 DSCP class is added to the CS4 QoS traffic class on inbound WAN QoS configurations to account for the inbound AF41 marked traffic, thus ensuring QoS treatment of AF41 marked media.

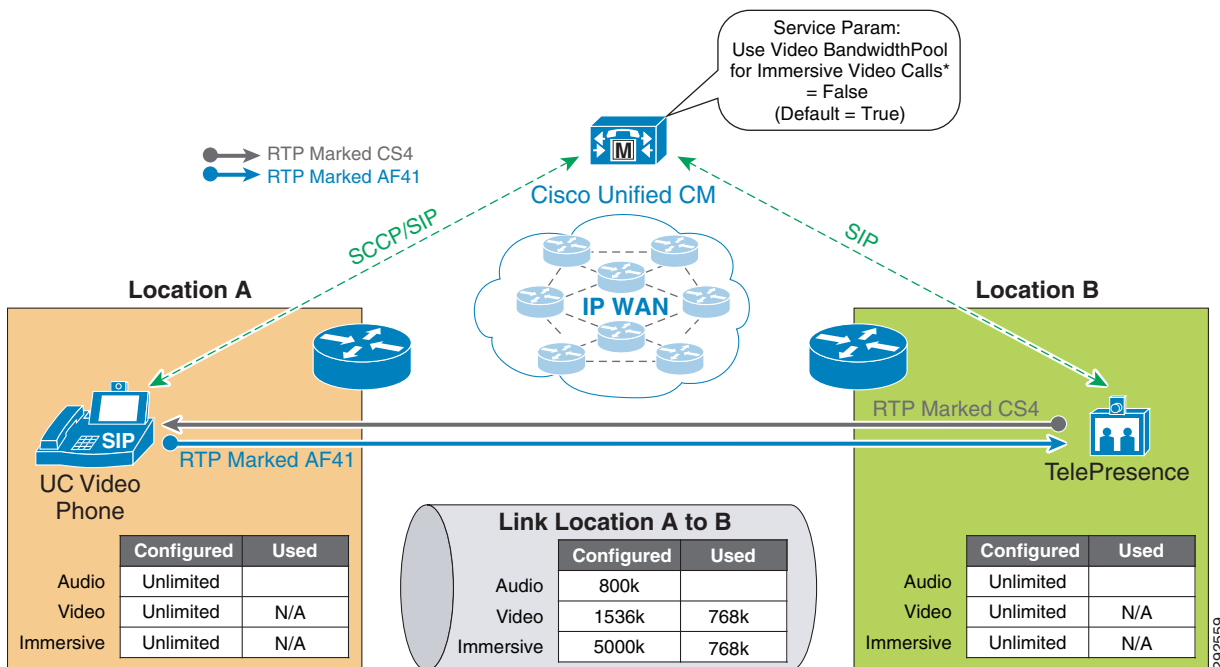
Design Considerations

When deploying Enhanced Location CAC for TelePresence video interoperable calls, consider the affects of DSCP marking for both QoS classes.

DSCP QoS Marking

The Differentiated Services Code Point (DSCP) QoS markings for TelePresence video interoperable calls are asymmetric, with AF41 used for the UC endpoints and CS4 for the TelePresence endpoints. AF41 and CS4 are default configurations in Unified CM, and changes to these defaults should align with the QoS configuration in the network infrastructure, as applicable. TelePresence endpoints mark video calls with a DSCP value of CS4, which is consistent with the default **DSCP for Telepresence calls** setting. UC endpoints mark calls with a DSCP value of AF41, which is consistent with the default **DSCP for Video calls** setting. [Figure 13-31](#) illustrates the media marking and bandwidth accounting.

Figure 13-31 Bandwidth Deductions and Media Marking in a Multi-Site Deployment with Enhanced Location CAC



Bandwidth Accounting for TelePresence Video Interoperability Calls

Enhanced Location CAC for TelePresence-to-UC video interoperable calls deducts bandwidth from both the video and immersive locations and links bandwidth pools, as illustrated in [Figure 13-31](#). This is by design to ensure that both types of QoS classified streams have the bandwidth required for media in both directions of the path between endpoints.

Enhanced Location CAC accounts for the bidirectional media of both AF41 and CS4 class traffic. In asymmetrically marked flows, however, the full allocated bit rate of the AF41 class is used in one direction but not the other. In the other direction, the full allocated bit rate is marked CS4. This does not

represent additional bandwidth consumption but simply a difference in marking and queuing in the network for each QoS class. This manner of bandwidth accounting is required to protect each flow in each direction.

If TelePresence video (CS4) has been provisioned in the network paths separately from Unified Communications video (AF41) and TelePresence is largely scheduled and in environments where the scheduling of calls is controlled and the utilization of TelePresence is deterministic, then immersive video bandwidth for locations and links can be set to **unlimited** to avoid the double bandwidth CAC calculations. This ensures that TelePresence-to-TelePresence calls always go through unimpeded and will not be subject to admission control, while desktop video and TelePresence-to-desktop video calls will be subject to admission control and accounted for in the video bandwidth allocation.

For more information on the call flows for Enhanced Location CAC and TelePresence interoperable calls, see the section on [Enhanced Location CAC for TelePresence Immersive Video](#), page 13-24.

Design Recommendations for Unified CM Session Management Edition Deployments with Enhanced Location CAC

Unified CM Session Management Edition (SME) is typically used for interconnecting multiple Unified CM clusters, third-party UC systems (IP- and TDM-based PBXs), PSTN connections, and centralized UC applications as well as for dial-plan and trunk aggregation. The following is a list of recommendations and design considerations to follow when deploying Unified CM SME with Enhanced Location (E-L) CAC. For more information on Unified CM SME, refer to the *Cisco Unified Communications Manager Session Management Edition Deployment Guide*, available at

http://www.cisco.com/en/US/products/ps10661/products_implementation_design_guides_list.html

Recommendations and Design Considerations

- All leaf clusters that support E-L CAC should be enabled for intercluster E-L CAC with SME.
- SME can be used as a centralized bootstrap hub for the E-L CAC intercluster hub replication network. See [LBM Hub Replication Network](#), page 13-15, for more information.
- All trunks to leaf clusters supporting E-L CAC should be SIP trunks placed in the shadow location to enable E-L CAC on the trunk between SME and the leaf clusters supporting E-L CAC.
- For TelePresence video interoperability, see the section on [Call Admission Control Design Recommendations for TelePresence Video Interoperability Architectures](#), page 13-42.
- Connectivity from SME to any trunk or device other than a Unified CM that supports E-L CAC (some examples are third-party PBXs, gateways, Unified CM clusters prior to release 9.0 that do not support E-L CAC, voice messaging ports or trunks to conference bridges, Cisco Video Communications Server, and so forth) should be configured in a location other than a phantom or shadow location. The reason for this is that both phantom and shadow locations are non-terminating locations; that is, they relay information about locations and are effectively placeholders for user-defined locations on other clusters. Phantom locations are legacy locations that allow for the transmission of location information in versions of Unified CM prior to 9.0, but they are not supported with Unified CM 9.x Enhanced Location CAC. Shadow locations are special locations that enable trunks between Unified CM clusters that support E-L CAC to accomplish it end-to-end.

- SME can be used as a locations and link management cluster. See Figure 13-32 as an example of this.
- SME can support a maximum of 2,000 locations configured locally.

Figure 13-32 Unified CM SME as a Location and Link Management Cluster

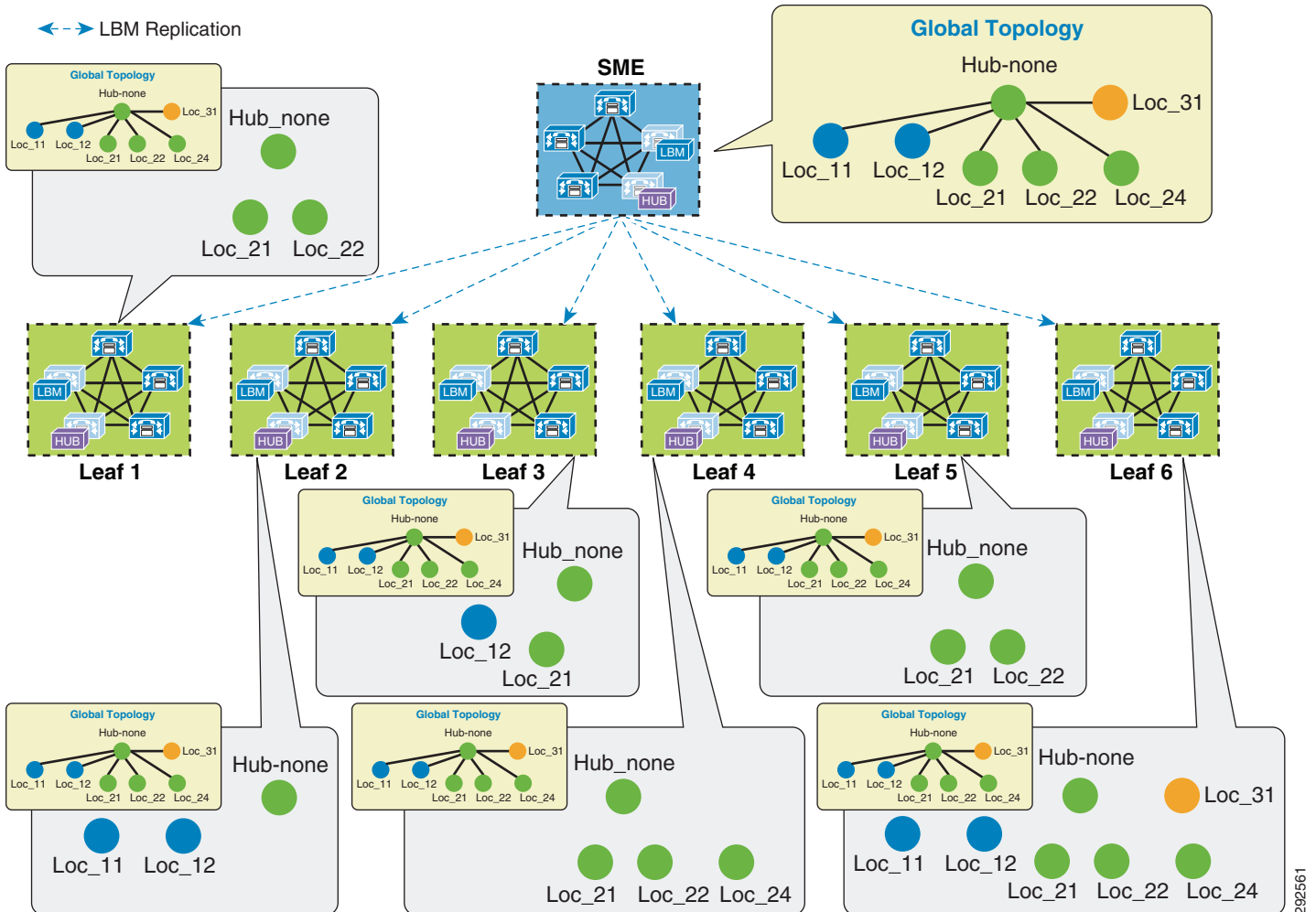


Figure 13-32 illustrates SME as a location and link management cluster. The entire location and link global topology is configured and managed in SME, and the leaf clusters configure locally only the locations that they require to associate with the end devices. When intercluster E-L CAC is enabled and locations and links are replicated, each leaf cluster will receive the global topology from SME and overlay this on their configured topology and use the global topology for call admission control. This simplifies configuration and location and link management across multiple clusters, and it diminishes the potential for misconfiguration across clusters. For more information and details on the design and deployment see the section on [Location and Link Management Cluster](#), page 13-20.

Figure 13-33 illustrates an SME design where intercluster E-L CAC has been enabled on one or more leaf clusters (right) and where one or more leaf clusters are running a version of Unified CM prior to 9.0 and are running traditional locations CAC (left). In this type of a deployment the locations managed by traditional locations CAC cannot be common or shared locations between E-L CAC-enabled clusters.

Leaf 1 has been configured in a traditional hub and spoke, where devices are managed at various remote sites. SME and the other leaf clusters that are enabled for intercluster E-L CAC share a global topology, as illustrated in the E-L CAC Modeled Topology. Leaf1_Hub is a user-defined location in SME assigned to the SIP or H.323 intercluster trunk that represents the hub of the Leaf 1 topology. This allows SME to deduct bandwidth for calls to and from Leaf 1 up to the Leaf1_Hub. In this way SME and Leaf 2 manage the E-L CAC locations and links while Leaf 1 manages its remote locations with traditional locations CAC.

Figure 13-33 SME Design with Enhanced Location CAC and Traditional Location CAC in Leaf Clusters

