Call Admission Control Design

This section describes how to apply the call admission control mechanisms to the various Unified CM deployment models and to the following IP WAN topologies:

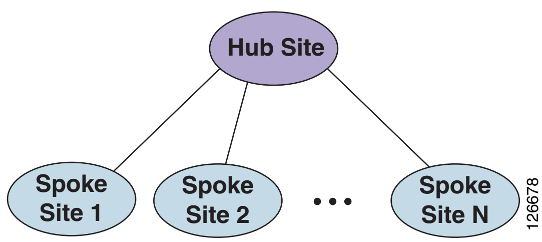
- Simple Hub-and-Spoke Topologies, page 9-37
- Two-Tier Hub-and-Spoke Topologies, page 9-41
- Simple MPLS Topologies, page 9-45
- Generic Topologies, page 9-52

For each topology, these sections present different sets of design considerations based on the Unified CM deployment model adopted.

Simple Hub-and-Spoke Topologies

Figure 9-21 depicts a simple hub-and-spoke topology, also known as a star topology. In this type of network topology, all sites (called *spoke sites*) are connected via a single IP WAN link to a central site (called the *hub site*). There are no direct links between the spoke sites, and every communication between them must transit through the hub site.

Figure 9-21 A Simple Hub-and-Spoke Topology



The design considerations in this section apply to simple hub-and-spoke topologies that use traditional Layer 2 IP WAN technologies such as:

- Frame Relay
- ATM
- Frame Relay/ATM Service Interworking
- Leased Lines

For IP WAN deployments based on the MPLS technology, refer to the section on Simple MPLS Topologies, page 9-45.

The remainder of this section contains design best practices for simple hub-and-spoke topologies according to the Unified CM deployment model adopted:

- Centralized Unified CM Deployments, page 9-38

  One or more Unified CM clusters are located at the hub site, but only phones and gateways are located at the spoke sites.

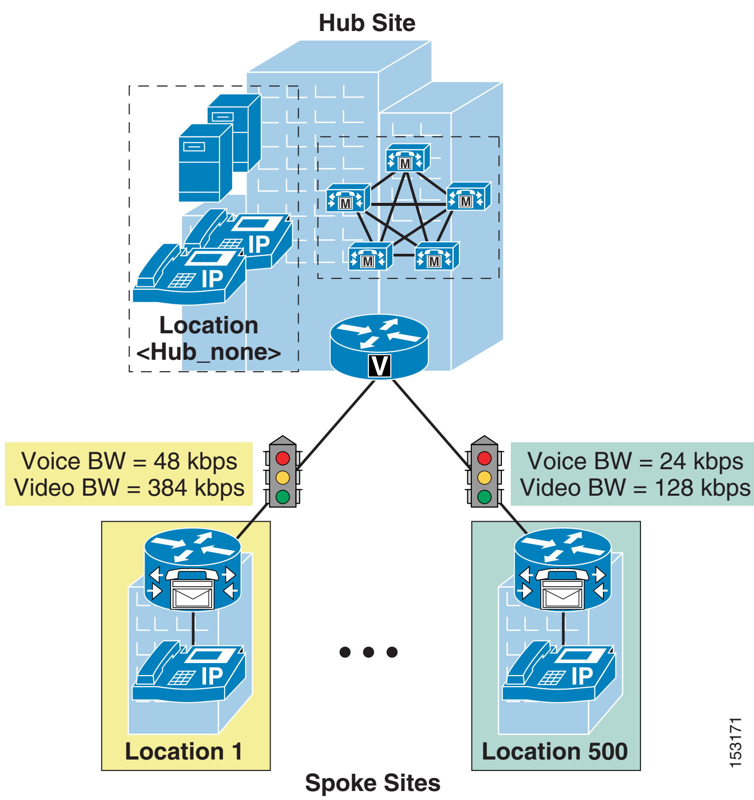
- Distributed Unified CM Deployments, page 9-39

  A Unified CM cluster or Cisco Unified Communications Manager Express (Unified CME) is located at each site.

Centralized Unified CM Deployments

In multisite WAN deployments with centralized call processing in a simple hub-and-spoke topology, use Unified CM static *locations* for implementing call admission control. Figure 9-22 shows an example of how to apply this mechanism to such a topology.

Figure 9-22 Call Admission Control for Simple Hub-and-Spoke Topologies Using Static Locations



Follow these guidelines when using static locations for call admission control:

- Configure a separate location in Unified CM for each spoke site.
- Configure the appropriate bandwidth limits for voice and video calls for each site according to the types of codecs used at that site. (See Table 9-2 for recommended bandwidth settings.)
- Assign all devices at each spoke site to the appropriate location.
- Leave devices at the hub site in the Hub_None location.
- If you move a device to another location, change its location configuration as well.
- Unified CM supports up to 2000 locations.
- If you require automatic rerouting over the PSTN when the WAN bandwidth is not sufficient, configure the automated alternate routing (AAR) feature on Unified CM. (See Automated Alternate Routing, page 10-86.)
- If multiple Unified CM clusters are located at the same hub site, leave the intercluster trunk devices in the Hub_None location. You may use a gatekeeper for dial plan resolution. However, gatekeeper call admission control is not necessary in this case because all IP WAN links are controlled by the locations algorithm.

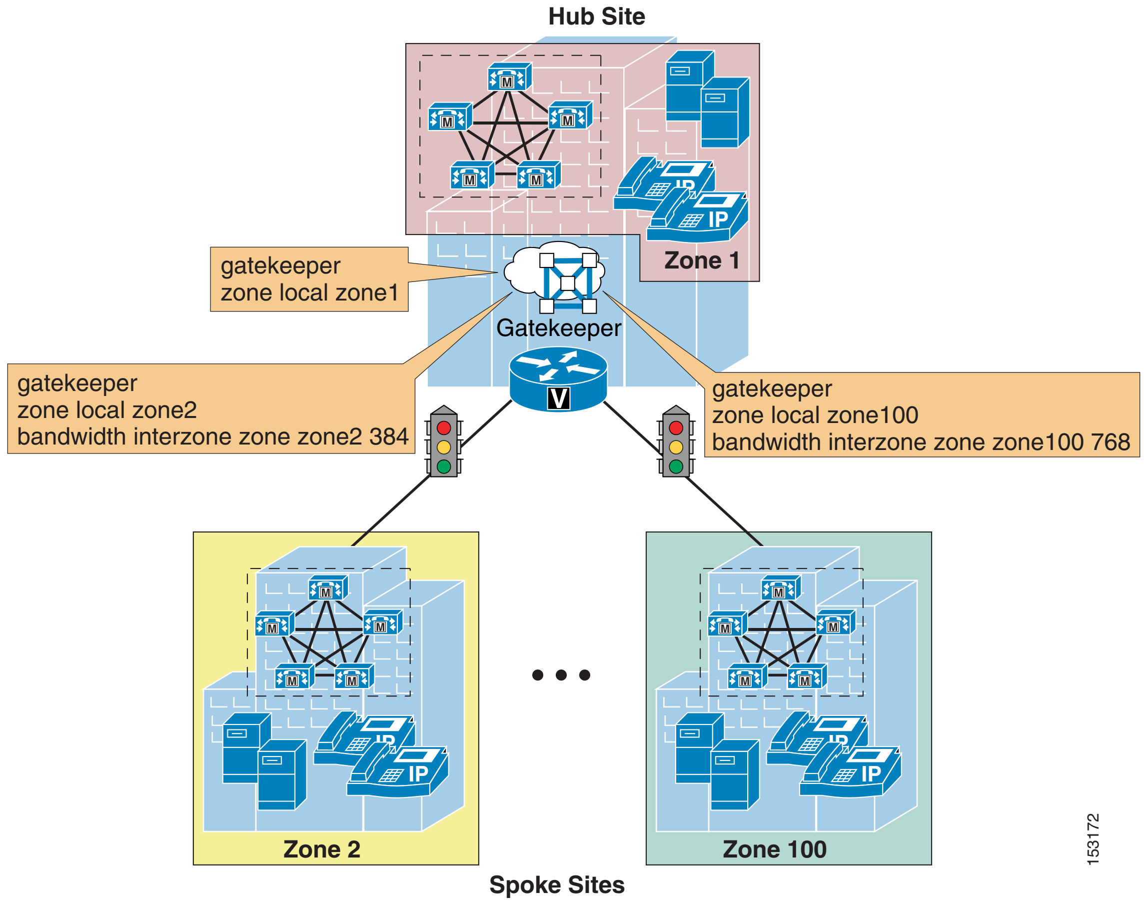
Note If one or more sites have dual connectivity to the IP WAN and you want to take full advantage of the bandwidth available on both links, Cisco recommends that you deploy topology-aware call admission control, as described in the section on Generic Topologies, page 9-52. See also Limitations of Topology-Unaware Call Admission Control, page 9-5, for more information.

Distributed Unified CM Deployments

For distributed call processing deployments on a simple hub-and-spoke topology, you can implement call admission control with a Cisco IOS gatekeeper. In this design, the call processing agent (which could be a Unified CM cluster, Cisco Unified Communications Manager Express (Unified CME), or an H.323 gateway) registers with the Cisco IOS gatekeeper and queries it each time the agent wants to place an IP WAN call. The Cisco IOS gatekeeper associates each call processing agent with a zone that has specific bandwidth limitations. Thus, the Cisco IOS gatekeeper can limit the maximum amount of bandwidth consumed by IP WAN voice calls into or out of a zone.

Figure 9-23 illustrates call admission control with a gatekeeper. In brief, when the call processing agent wants to place an IP WAN call, it first requests permission from the gatekeeper. If the gatekeeper grants permission, the call processing agent places the call across the IP WAN. If the gatekeeper denies the request, the call processing agent can try a secondary path (the PSTN, for example) or can simply fail the call.

Figure 9-23 Call Admission Control for Hub-and-Spoke Topologies Using a Gatekeeper



Follow these guidelines when deploying call admission control with a gatekeeper:

- In Unified CM, configure an H.225 gatekeeper-controlled trunk if you have a mixed environment with Cisco Unified Communications Manager Express (Unified CME) and H.323 gateways.
- In Unified CM, configure an intercluster gatekeeper-controlled trunk if you have an environment exclusively based on Unified CM clusters.
- Ensure that the zone configured in Unified CM matches the correct gatekeeper zone for the site.
- Each Unified CM subscriber listed in the device pool's Unified CM Redundancy Group registers a gatekeeper-controlled trunk with the gatekeeper. (Maximum of three.)
- Calls are load-balanced across the registered trunks in the Unified CM cluster.
- Unified CM supports multiple gatekeepers and trunks.
- You can place the trunk in a route group and route list construct to provide automatic PSTN failover. (See Dial Plan, page 10-1, for more details.)
- Configure a separate zone in the gatekeeper for each site supporting Unified CMs, Unified CME, or H.323 gateways.
- Use the **bandwidth interzone** command on the gatekeeper to control bandwidth between Unified CM clusters, Unified CME servers, and H.323 devices registered directly with the gatekeeper. (See Table 9-4 for bandwidth settings by codec type.)

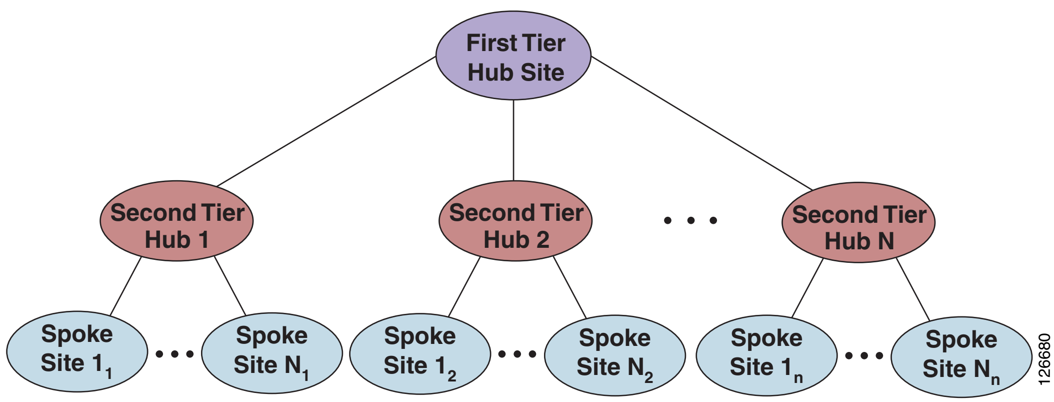
- A single Cisco IOS gatekeeper can support up to 100 zones or sites.
- You can provide gatekeeper redundancy by using gatekeeper clustering (alternate gatekeeper) or Cisco Hot Standby Router Protocol (HSRP). Use HSRP only if gatekeeper clustering is not available in your software feature set.

Note If one or more sites have dual connectivity to the IP WAN and you want to take full advantage of the bandwidth available on both links, Cisco recommends that you deploy topology-aware call admission control, as described in the section on Generic Topologies, page 9-52. See also Limitations of Topology-Unaware Call Admission Control, page 9-5, for more information.

Two-Tier Hub-and-Spoke Topologies

Figure 9-24 depicts a two-tier hub-and-spoke topology. This type of network topology consists of sites at three hierarchical levels: the first-tier hub site, the second-tier hub sites, and the spoke sites. A group of spoke sites are connected to a single second-tier hub site, and each second-tier hub site is in turn connected to the single first-tier hub site. As in the simple hub-and-spoke topology, there are no direct links between the spoke sites, and every communications between them must transit through the second-tier hub site. Similarly, there are no direct links between the second-tier hub sites, and all communications between them must transit through the first-tier hub site.

Figure 9-24 A Two-Tier Hub-and-Spoke Topology



The design considerations in this section apply to two-tier hub-and-spoke topologies that use traditional Layer 2 IP WAN technologies such as:

- Frame Relay
- ATM
- Frame Relay/ATM Service Interworking
- Leased Lines

For IP WAN deployments based on the MPLS technology, refer to the section on Simple MPLS Topologies, page 9-45.

The remainder of this section contains design best practices for two-tier hub-and-spoke topologies according to the Unified CM deployment model adopted:

- Centralized Unified CM Deployments, page 9-42

  One or more Unified CM clusters are located at the first-tier hub site, but only phones and gateways are located at the second-tier hub sites and the spoke sites.

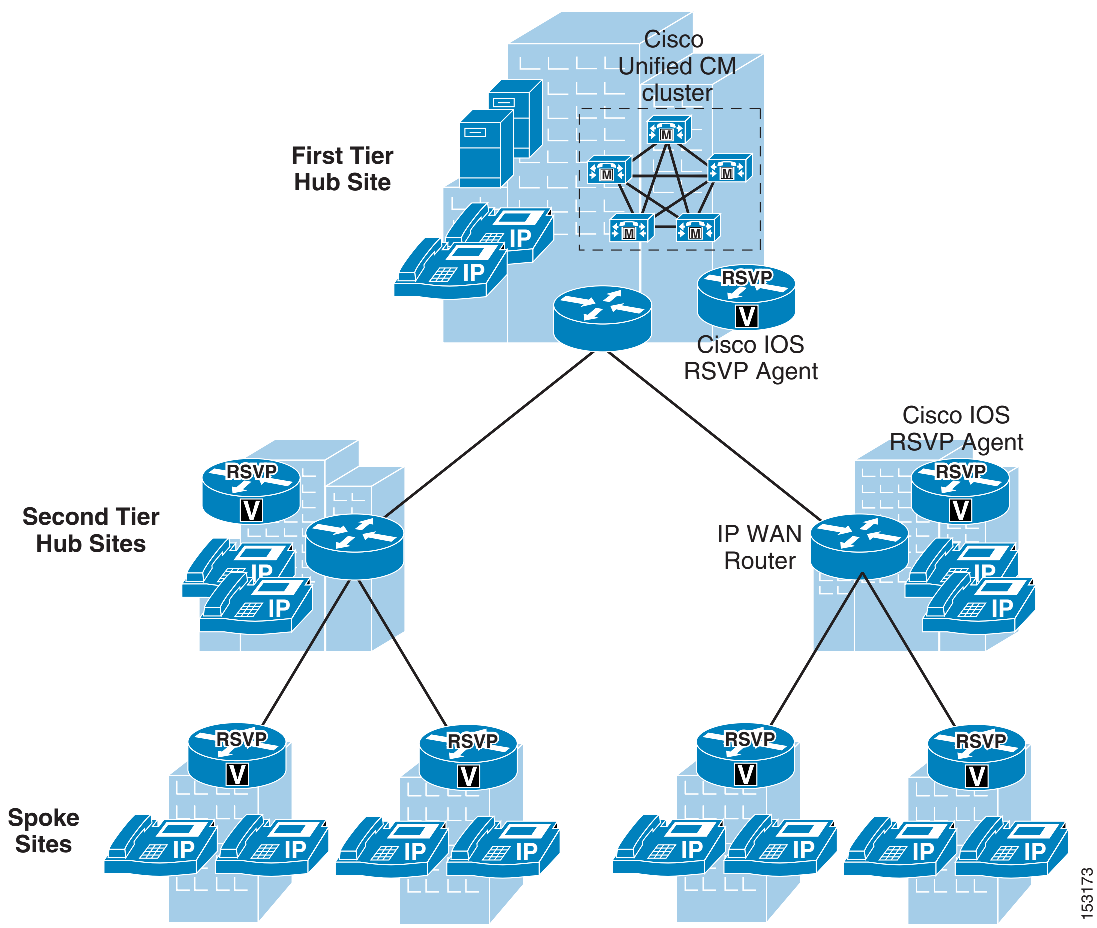
- Distributed Unified CM Deployments, page 9-44

  Unified CM clusters are located at the first-tier hub site and at the second-tier hub sites, while only endpoints and gateways are located at the spoke sites.

Centralized Unified CM Deployments

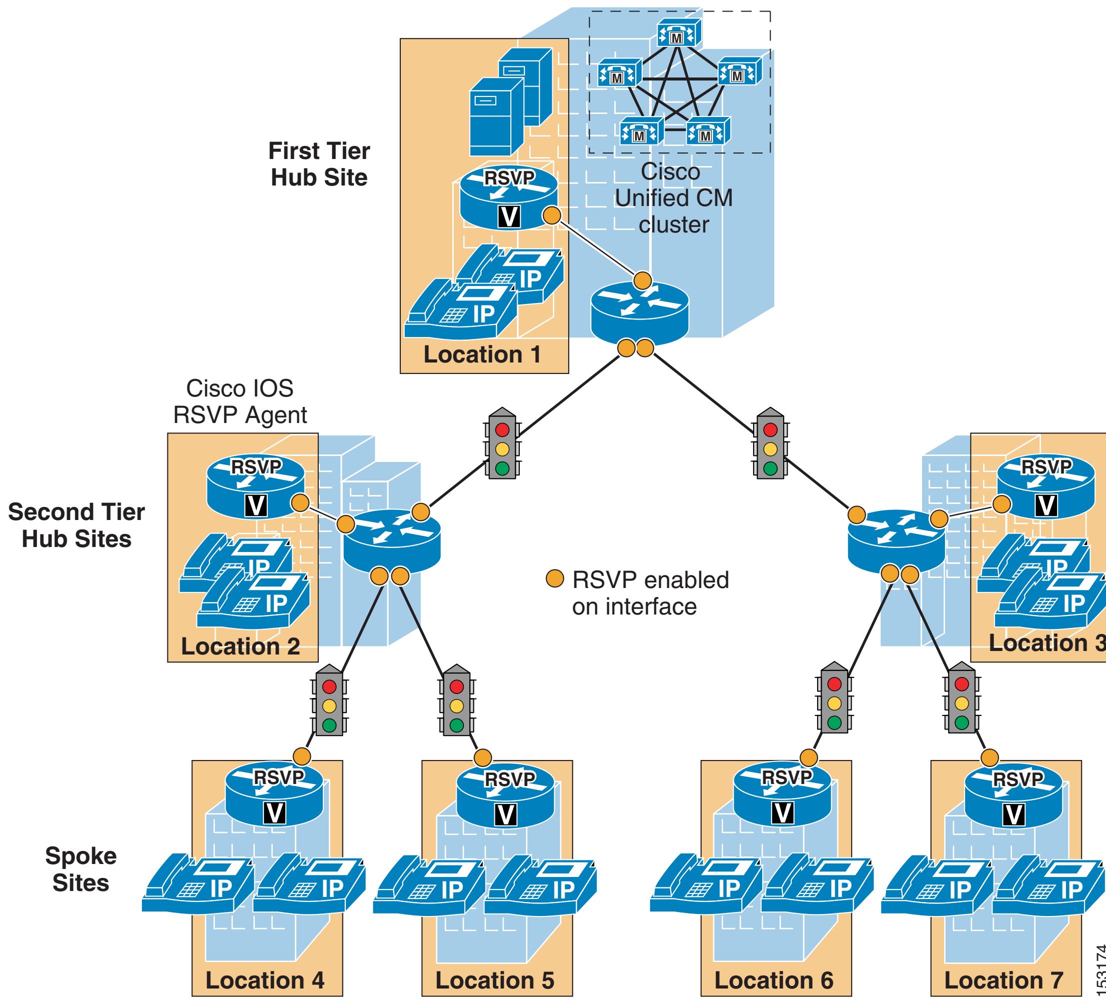
Figure 9-25 depicts a single centralized Unified CM cluster deployed in a two-tier hub-and-spoke IP WAN topology. In this scenario, the Unified CM cluster is located at the first-tier hub site, while all second-tier hub and spoke sites contain only endpoints and gateways.

Figure 9-25 Two-Tier Hub-and-Spoke Topology with Centralized Unified CM



This scenario requires that you deploy topology-aware call admission control, which for a single Unified CM cluster means using RSVP-enabled locations. Figure 9-26 shows how this mechanism can be deployed.

Figure 9-26 Call Admission Control for Two-Tier Hub-and-Spoke Topologies Using Locations with RSVP



The following guidelines apply to these deployments:

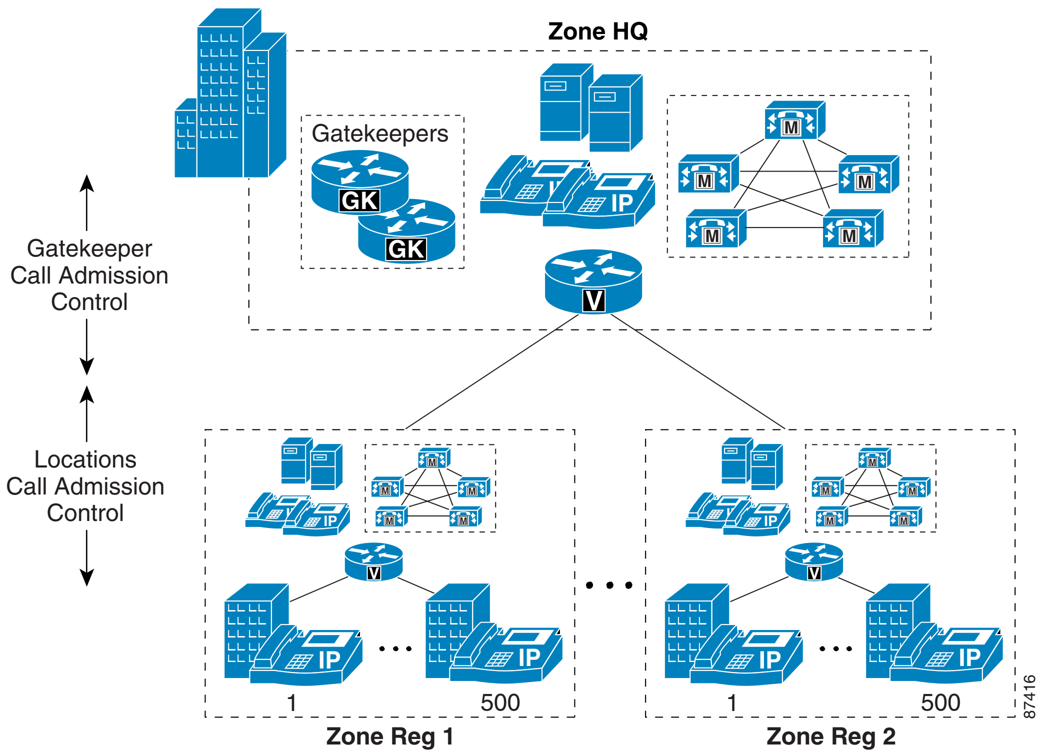
- Enable the Cisco IOS RSVP Agent feature on a Cisco IOS router at each site. At smaller sites, this router may coincide with the IP WAN router and PSTN gateway, while at larger sites they may be different platforms.
- In Unified CM, define a location for each site, and leave all bandwidth values as **Unlimited**.
- Assign all devices located at each site to the appropriate location (this includes endpoints, gateways, conferencing resources, and the Cisco RSVP Agent itself).
- Ensure that each Cisco RSVP Agent belongs to a media resource group (MRG) contained in the media resource group list (MRGL) of all devices at that site.
- In the Unified CM service parameters, set the **Default inter-location RSVP Policy** to **Mandatory** or **Mandatory (video desired)** and set the **Mandatory RSVP mid-call error handle option** to **Call fails following retry counter exceeded**.

- Enable RSVP on every WAN router interface in the network where congestion might occur, and configure the RSVP bandwidth based on the provisioning of the priority queue.
- If the Cisco RSVP Agent is not co-resident with the IP WAN router, enable RSVP on the LAN interfaces connecting the agent to the WAN router (as illustrated in Figure 9-26).

Distributed Unified CM Deployments

To provide call admission control in deployments that use a two-tier hub-and-spoke topology, with Unified CMs at the first-tier and second-tier hub sites, you can combine the static locations and gatekeeper zone mechanisms as illustrated in Figure 9-27.

Figure 9-27 Combining the Locations and Gatekeeper Mechanisms for Call Admission Control



Follow these recommendations when combining gatekeeper zones with static locations for call admission control:

- Use call admission control based on static locations for sites with no local Unified CM (that is, the spoke sites).
- Use gatekeeper-based call admission control between Unified CM clusters (that is, between the first-tier hub site and the second-tier hub sites).
- For each site without a local Unified CM, configure a location for that site in the Unified CM cluster supporting the site.
- Configure the appropriate bandwidth limits for voice and video calls at each site according to the type of codec used at that site. (See Table 9-2 and Table 9-4 for bandwidth settings.)
- Assign each device configured in Unified CM to a location. If you move a device to another location, change its location configuration as well.

- Unified CM supports up to 2000 locations.
- Each Unified CM cluster registers a gatekeeper-controlled trunk with the gatekeeper.
- On the gatekeeper, configure a zone for each Unified CM cluster, and use the **bandwidth interzone** command to control the number of calls to and from each cluster.

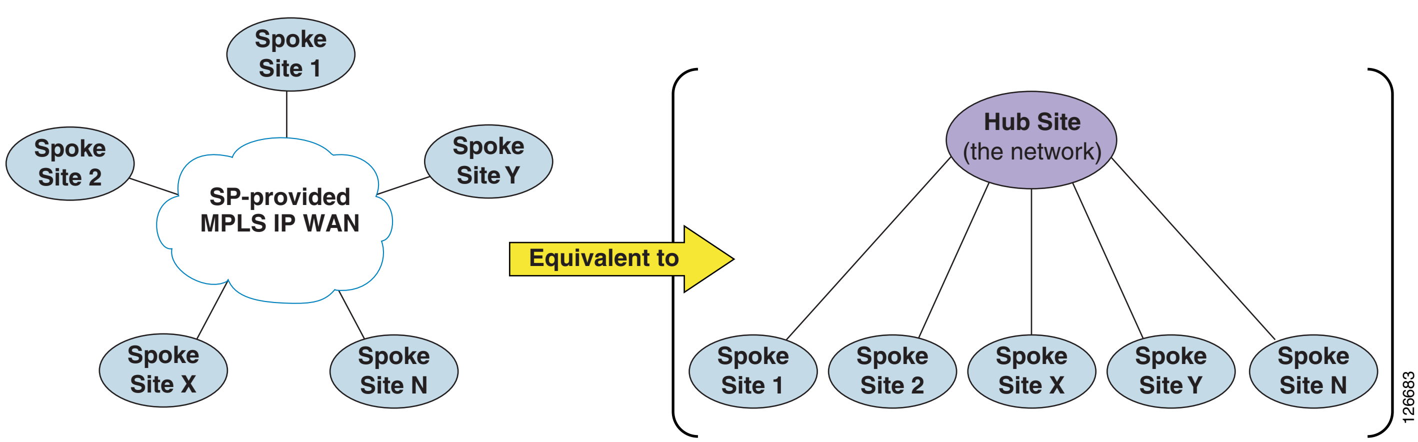
Note If one or more sites have dual connectivity to the IP WAN and you want to take full advantage of the bandwidth available on both links, Cisco recommends that you deploy topology-aware call admission control, as described in the section on Generic Topologies, page 9-52. See also Limitations of Topology-Unaware Call Admission Control, page 9-5, for more information.

Simple MPLS Topologies

Figure 9-28 shows an IP WAN (from a service provider) based on the Multiprotocol Label Switching (MPLS) technology. The main design difference between traditional Layer 2 WAN services offered by service providers and services based on MPLS is that, with MPLS, the IP WAN topology does not conform to a hub-and-spoke but instead provides "full-mesh" connectivity between all sites.

This topology difference means that, from an IP routing perspective on the enterprise side of the network, each site is one IP hop away from all other sites. Thus, there is no need to transit through a hub site to reach any other site. In fact, there is no concept of a "hub site." All sites are considered equal, and the only difference between them is the amount of bandwidth that they are allowed to use across the IP WAN.

Figure 9-28 MPLS IP WAN from a Service Provider, and Its Topology Equivalent



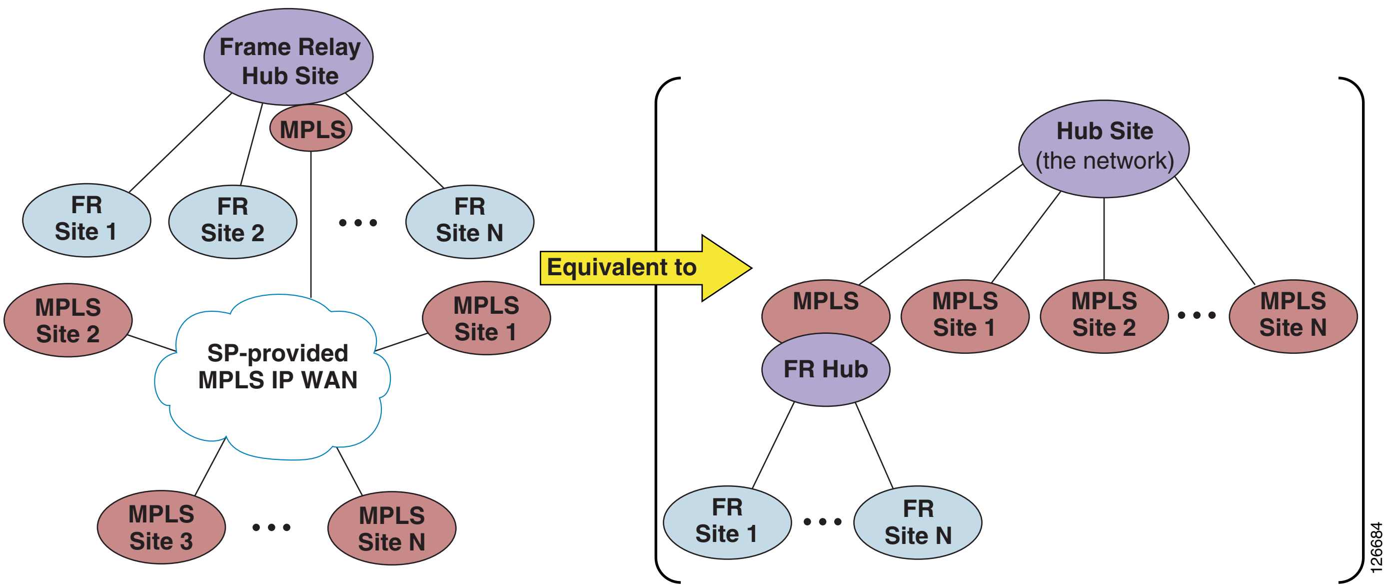
Based on these considerations, it is easy to see that, from a call admission control perspective, a service-provider IP WAN service based on MPLS is in reality equivalent to a hub-and-spoke topology without a hub site. (See Figure 9-28.) In fact, the network itself could be considered as the hub site, while all the enterprise sites (including the headquarters, or central site) are equivalent to spoke sites. These differences have implications on how to perform call admission control, which are described in the remainder of this section.

An exception to the above considerations that is worth mentioning here is represented by multisite deployments where an MPLS-based WAN co-exists with an IP WAN based on a traditional Layer 2 technology, such as Frame Relay or ATM. Such a scenario could occur, for example, in the case of network migration phases, company mergers, or various other situations.

As shown in Figure 9-29, integrating a hub-and-spoke IP WAN based on a traditional Layer 2 technology (such as Frame Relay) with an MPLS-based IP WAN results in a network topology that is neither a simple hub-and-spoke nor a full-mesh, but rather is equivalent to a two-tier hub-and-spoke.

In this case the first-tier hub site is represented by the MPLS network, the second-tier hub sites are represented by the MPLS-based sites as well as the MPLS-enabled Frame Relay hub site, and the spoke sites are represented by the Frame Relay spoke sites. Therefore, for design considerations on such deployments, refer to the section on Two-Tier Hub-and-Spoke Topologies, page 9-41.

Figure 9-29 Co-existence of MPLS Sites and Frame Relay Sites, and the Topology Equivalent



The remainder of this section contains design best practices for MPLS-based topologies according to the Unified CM deployment model adopted:

- Centralized Unified CM Deployments, page 9-47

  One or more Unified CM clusters are located at only one site, while only endpoints and gateways are located at all other sites.

- Distributed Unified CM Deployments, page 9-50

  Unified CM clusters are located at multiple sites, while endpoints and gateways are located at all other sites.

Note This section focuses on enterprise deployments where the MPLS WAN service is provided by a service provider. In cases where the MPLS network is deployed by the enterprise itself, call admission control can be performed effectively if one of the following two conditions is satisfied: (1) routing in the MPLS network is configured so that it is equivalent to a hub-and-spoke, or (2) bandwidth in the core of the MPLS network is heavily over-provisioned so that congestion can occur only at the edge.

> **Note** If one or more sites have dual connectivity to the IP WAN and you want to take full advantage of the bandwidth available on both links, Cisco recommends that you deploy topology-aware call admission control, as described in the section on Generic Topologies, page 9-52. Particular care must be taken to guarantee symmetric routing in the presence of load-balanced links. See also Limitations of Topology-Unaware Call Admission Control, page 9-5, and Special Considerations for MPLS Networks, page 9-11, for more information, and contact your local Cisco account team for further assistance.

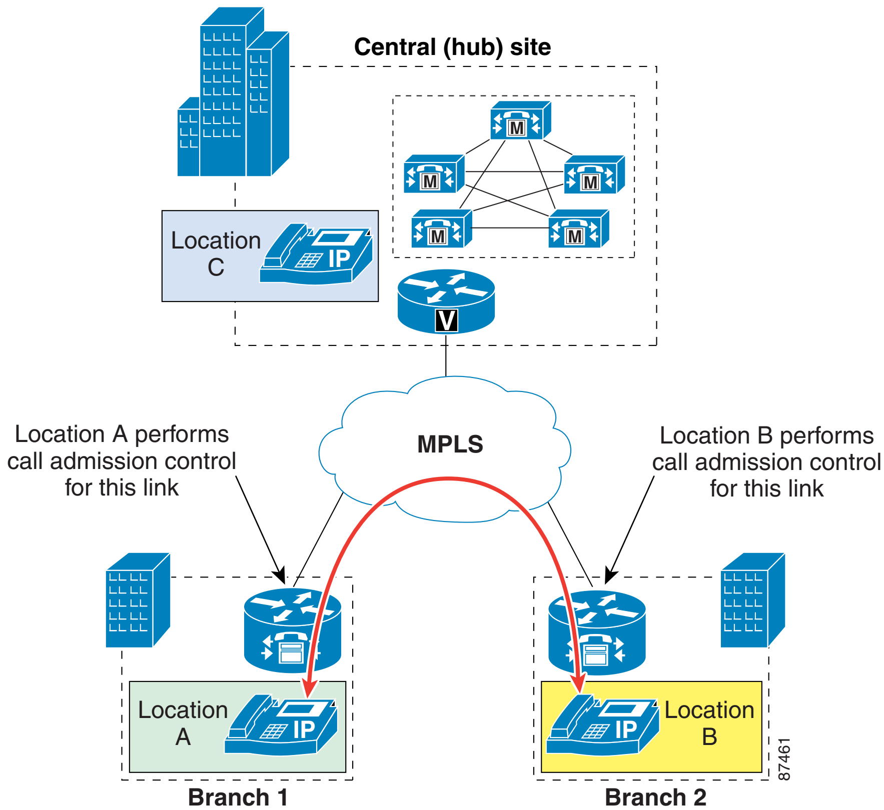
Centralized Unified CM Deployments

In multisite WAN deployments with centralized call processing in an MPLS topology, use Unified CM static *locations* for implementing call admission control.

In a hub-and-spoke WAN topology (for example, Frame Relay or ATM), each link to and from a branch site terminates at the central site. For example, in a Frame Relay network, all permanent virtual circuits (PVCs) from the branch routers are aggregated at the central site's head-end router. In such a scenario, there is no need to apply call admission control to devices at the central site because the bandwidth accounting occurs at the branch ends of the WAN links. Therefore, within the Unified CM locations configuration, devices at the central site are left in the Hub_None location, while devices at each branch are placed in their appropriate location to ensure proper call admission control.

With an MPLS WAN network, all branches are deemed to be adjacent at Layer 3, thus they do not have to rely on the central site for connectivity. Figure 9-30 illustrates a spoke-to-spoke call between two branch sites in this type of deployment.

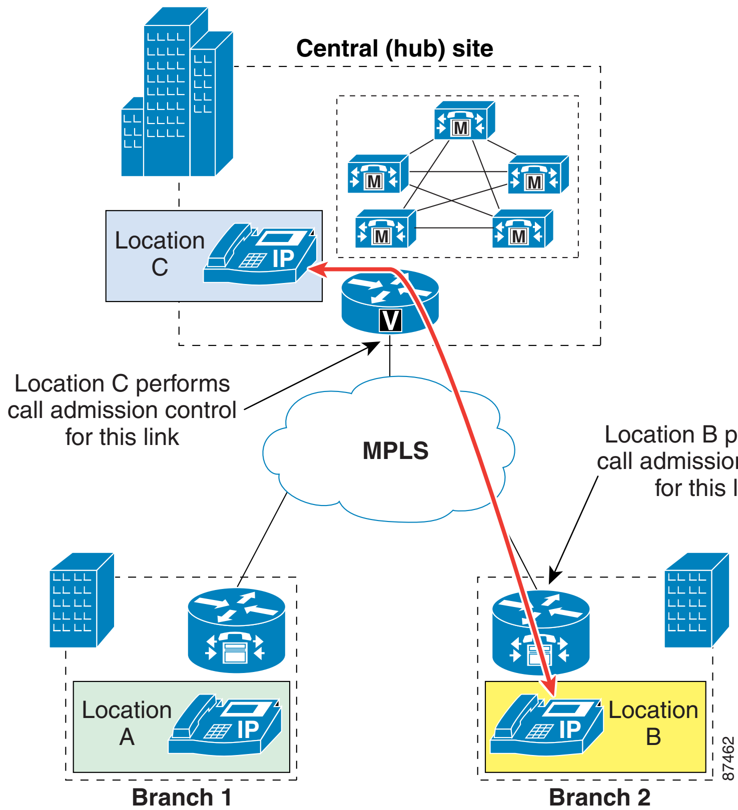
Figure 9-30 Spoke-to-Spoke Calls in an MPLS Deployment



Also, in an MPLS WAN, the link connecting the central site to the WAN does not aggregate every branch's WAN link. By placing all the central site devices in their own call admission control location (that is, not in the Hub_None location), this configuration requires that call admission control be performed on the central site link independently of the branch links. (See Figure 9-31.)

Note Some devices such as trunks do not terminate media and are normally left in the Hub_None location. However, to avoid errors in call admission control when an MTP is required on a trunk, the trunk must be assigned to a location other than Hub_None and any MTP in the trunk's MRGL must be physically located at the site associated with that location. This configuration is required because an MTP cannot be assigned a location directly, so it inherits the location of the device that selected it.

Figure 9-31 Calls to and from the Hub in an MPLS Deployment

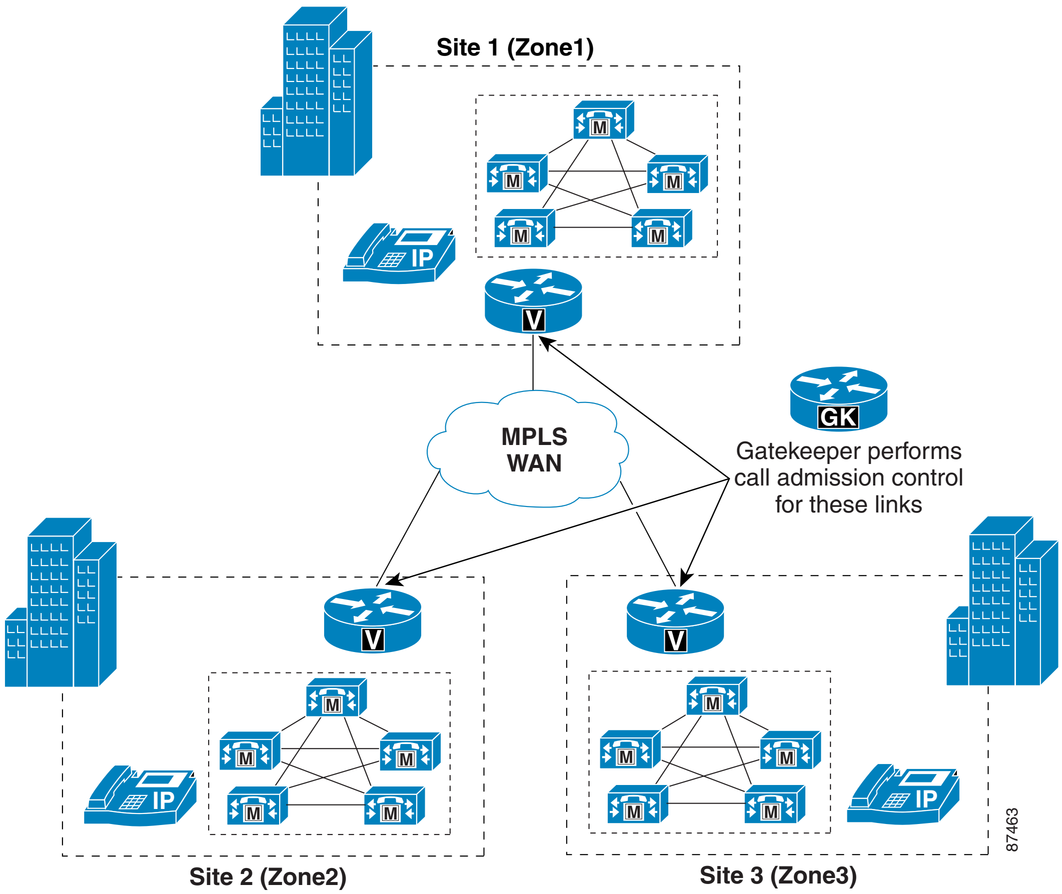


When all the available bandwidth for a particular site has been utilized, you can provide automatic failover to the PSTN by using the automated alternate routing (AAR) feature within Unified CM. (For more information on AAR, see Automated Alternate Routing, page 10-86.)

Distributed Unified CM Deployments

In multisite deployments where a Unified CM cluster is present at more than one site without any branch locations and the sites are linked through an MPLS WAN, a gatekeeper can provide dial-plan resolution as well as call admission control between the sites, with each site being placed in a different gatekeeper zone. This is the same mechanism adopted for hub-and-spoke topologies based on Layer 2 WAN technologies. (See Figure 9-32.)

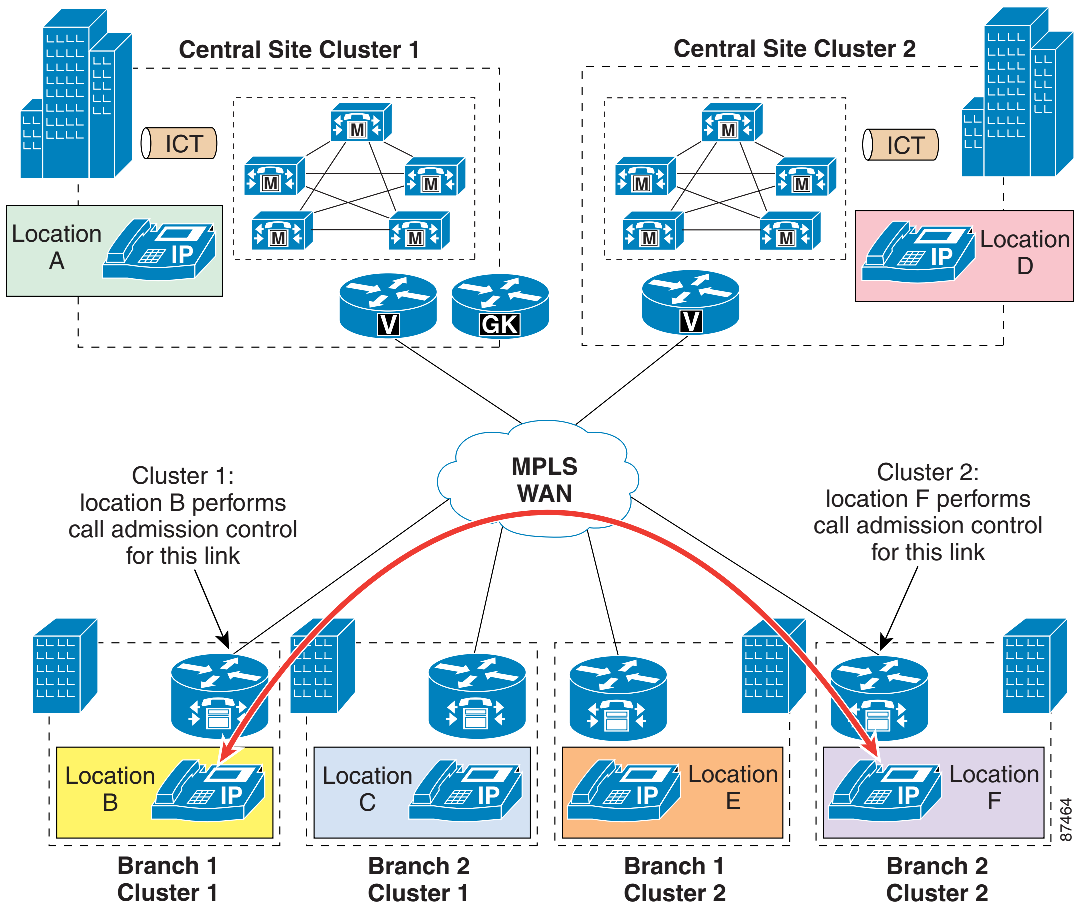
Figure 9-32 Gatekeeper Call Admission Control in a Distributed Deployment with MPLS



In deployments where branch sites are required, a gatekeeper can be used for dial-plan resolution between clusters, but a gatekeeper is not recommended for call admission control.

When calls occur between branches belonging to different clusters, the audio path is established between the two branches directly, with no media transiting through each cluster's central site. Therefore, call admission control is required only on the WAN links at the two branches. (See Figure 9-33.)

Figure 9-33 Multiple Clusters Connected by Intercluster Trunks (ICTs)



As in the centralized Unified CM deployments, devices that terminate media at each site (including the central sites for each cluster) must be placed in an appropriately configured location.

Note that the intercluster trunks are purely signaling devices, and there is no media transiting through them. Therefore, all intercluster trunks must be left in location Hub_None. The exception is when the trunk requires an MTP, in which case the trunk and MTP should both be in the location of the site in which they reside.

When all the available bandwidth for a particular site has been used, you can provide automatic failover to the PSTN by using a combination of the following two methods:

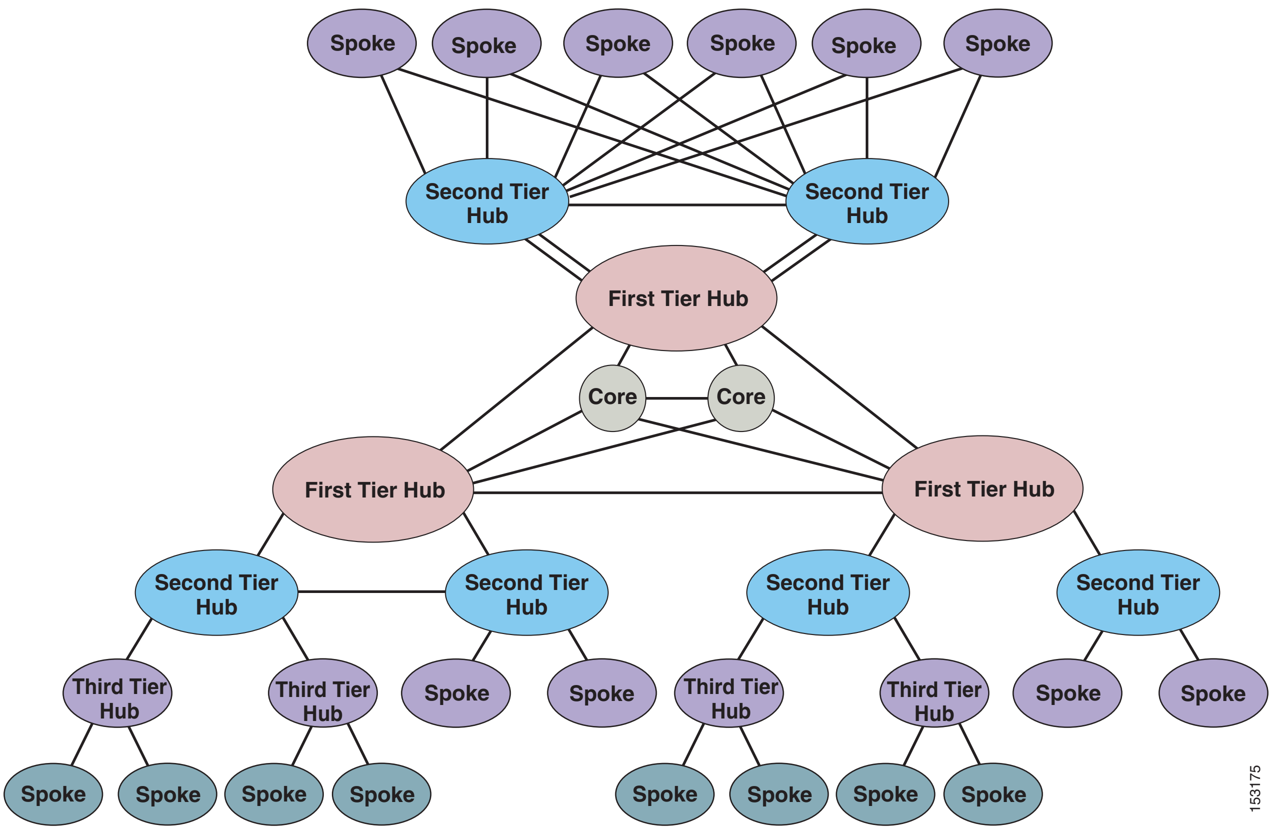
- The route list and route group construct for calls across multiple Unified CM clusters
- The automated alternate routing (AAR) feature for calls within a Unified CM cluster (For more information on AAR, see Automated Alternate Routing, page 10-86.)

Generic Topologies

In the context of this chapter, a generic topology is a network topology that cannot be reduced to a simple or two-tier hub-and-spoke or to a simple MPLS-based network.

As Figure 9-34 illustrates, a generic topology can present full-mesh features, hub-and-spoke features, partial-mesh features, or possibly all of them combined in a single network. It may also present dual connections between sites, as well as multiple paths from one site to another.

Figure 9-34 A Generic Topology



The complex nature of these networks requires the adoption of topology-aware call admission control mechanisms based on RSVP. In particular, these mechanisms can properly control bandwidth in presence of any of the following topology aspects:

- Remote sites dual-homed to different hub sites
- Multiple IP WAN links between any two sites, either in a primary/backup configuration or in an active/active load-balanced configuration
- Redundant hubs or data centers with a dedicated connection
- Fully-meshed core networks
- Multiple equal-cost IP paths between any two sites
- Multi-tiered architectures

The remainder of this section contains design best practices for generic network topologies according to the Unified CM deployment model adopted:

- Centralized Unified CM Deployments, page 9-53

  One or more Unified CM clusters are located at a given site, but only endpoints and gateways are located at all other sites.

- Distributed Unified CM Deployments, page 9-56

  Unified CM clusters are located at multiple sites, and endpoints and gateways are located at all other sites.

Centralized Unified CM Deployments

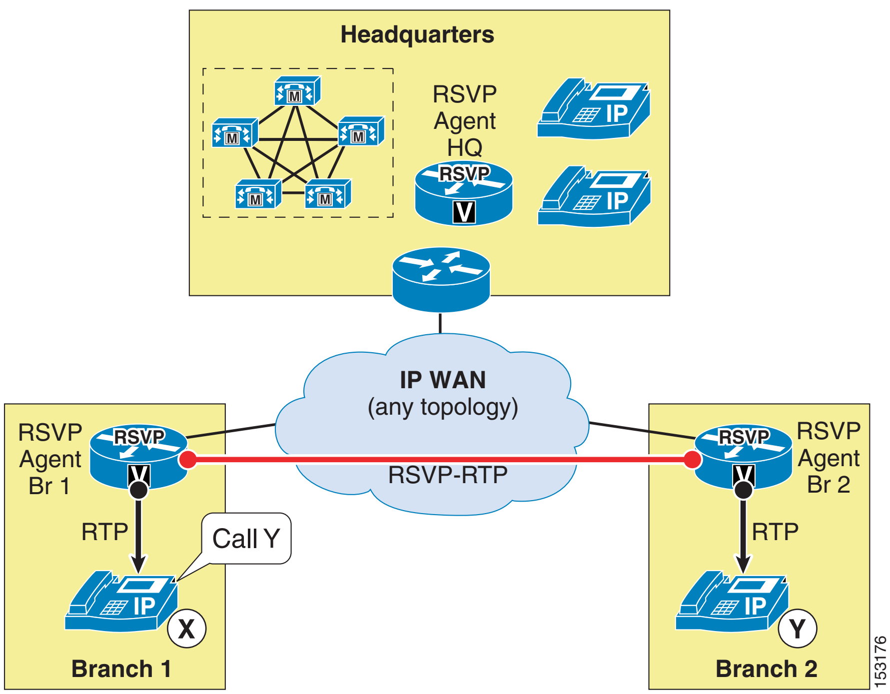
Centralized Unified CM deployments using a generic topology can be categorized into two sub-types:

- Single Unified CM Cluster, page 9-53
- Co-Located Unified CM Clusters, page 9-54

Single Unified CM Cluster

The recommendations in this section apply to a single Unified CM cluster deployed in a generic network topology, as illustrated in Figure 9-35.

Figure 9-35 A Single Unified CM Cluster in a Generic Topology



The following guidelines apply to this type of deployment:

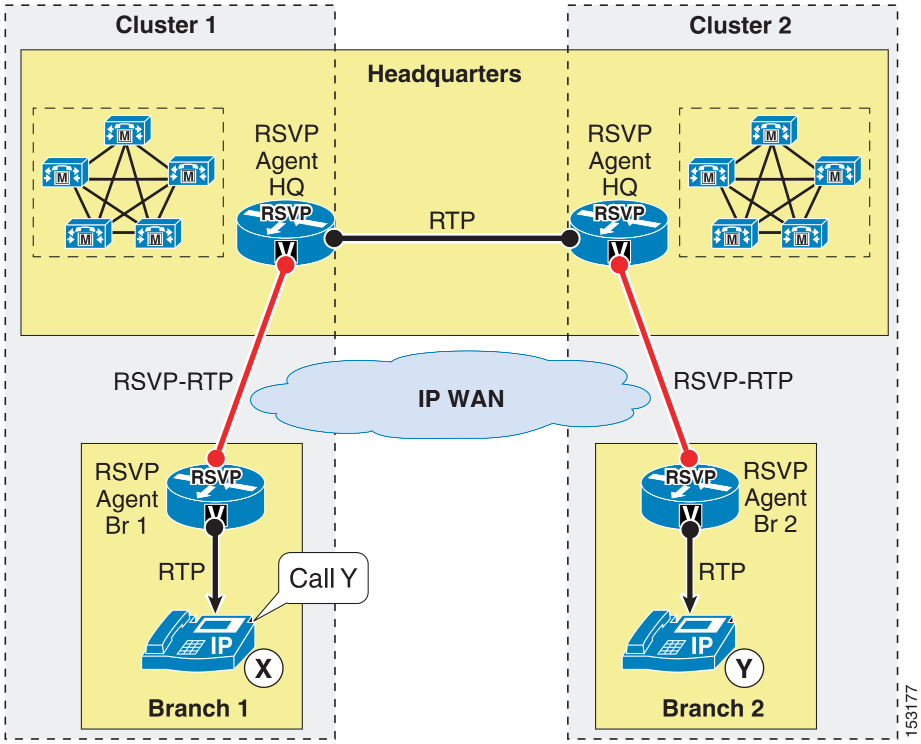
- Enable the Cisco IOS RSVP Agent feature on a Cisco IOS router at each site, including the central site where Unified CM resides. At smaller sites, this router may coincide with the IP WAN router and PSTN gateway, while at larger sites they may be different platforms.
- In Unified CM, define a location for each site, and leave all bandwidth values as **Unlimited**.
- Assign all devices located at each site to the appropriate location (this includes endpoints, gateways, conferencing resources, and the Cisco RSVP Agents themselves).
- Ensure that each Cisco RSVP Agent belongs to a media resource group (MRG) contained in the media resource group list (MRGL) of all devices at that site.
- In the Unified CM service parameters, set the **Default inter-location RSVP Policy** to **Mandatory** or **Mandatory (video desired)** and set the **Mandatory RSVP mid-call error handle option** to **Call fails following retry counter exceeded**.
- Enable RSVP on every WAN router interface in the network where congestion might occur, and configure the RSVP bandwidth based on the provisioning of the priority queue. (See RSVP Design Best Practices, page 3-58.)
- If you need to provision bandwidth separately for voice and video calls, also configure an RSVP application ID on the same WAN router interfaces.
- If the Cisco RSVP Agent is not co-resident with the IP WAN router, enable RSVP on the LAN interfaces connecting the agent to the WAN router.

Co-Located Unified CM Clusters

The recommendations in this section apply to deployments where multiple Unified CM clusters are located on the same LAN or MAN. However, the same considerations may also be valid if the sites where the Unified CM clusters reside are connected via a high-speed link, provided that no congestion occurs in the priority queue of the link and that bandwidth for voice and video can therefore be considered unlimited.

Figure 9-36 illustrates a deployment with two Unified CM clusters located at a given site (HQ) and a number of remote sites with endpoints and gateways, which are controlled either by Cluster 1 (for example, Branch 1) or Cluster 2 (for example, Branch 2).

Figure 9-36 Co-located Unified CM Clusters in a Generic Topology



In addition to the guidelines listed in Single Unified CM Cluster, page 9-53, observe the following best practices for this type of deployment:

- For each cluster, define an intercluster trunk to enable communications to the other clusters. A gatekeeper may be used for dial plan resolution, but it is not needed for call admission control.
- Assign the intercluster trunk to the same location used for all devices located at the central site (HQ in the example of Figure 9-36).
- Ensure that the intercluster trunk is assigned to a device pool that specifies an MRGL that in turn points to an MRG containing the Cisco RSVP Agent located at the central site (Cisco RSVP Agent HQ1 in the case of Cluster 1 in Figure 9-36).
- Use the AAR feature to provide automatic PSTN failover in case of call admission control failure within a cluster.
- Use the route list and route group constructs to provide automatic PSTN failover in case of call admission control failure across clusters.
- Both media and signaling traffic are hair-pinned via the central site for calls between two branch sites belonging to different clusters (as shown in Figure 9-36, where the call between phone X in Branch 1 and phone Y in Branch 2 is hair-pinned via the HQ site).

Distributed Unified CM Deployments

In order to provide call admission control for distributed Unified CM deployments in a generic network topology, two approaches are possible, depending on the number of Unified CM clusters involved:

- Remote Cisco RSVP Agent Approach, page 9-57

  This solution applies to deployments where three or fewer Unified CM clusters are located at different sites connected via an IP WAN with limited bandwidth.

- Cisco Unified Border Element Approach, page 9-60

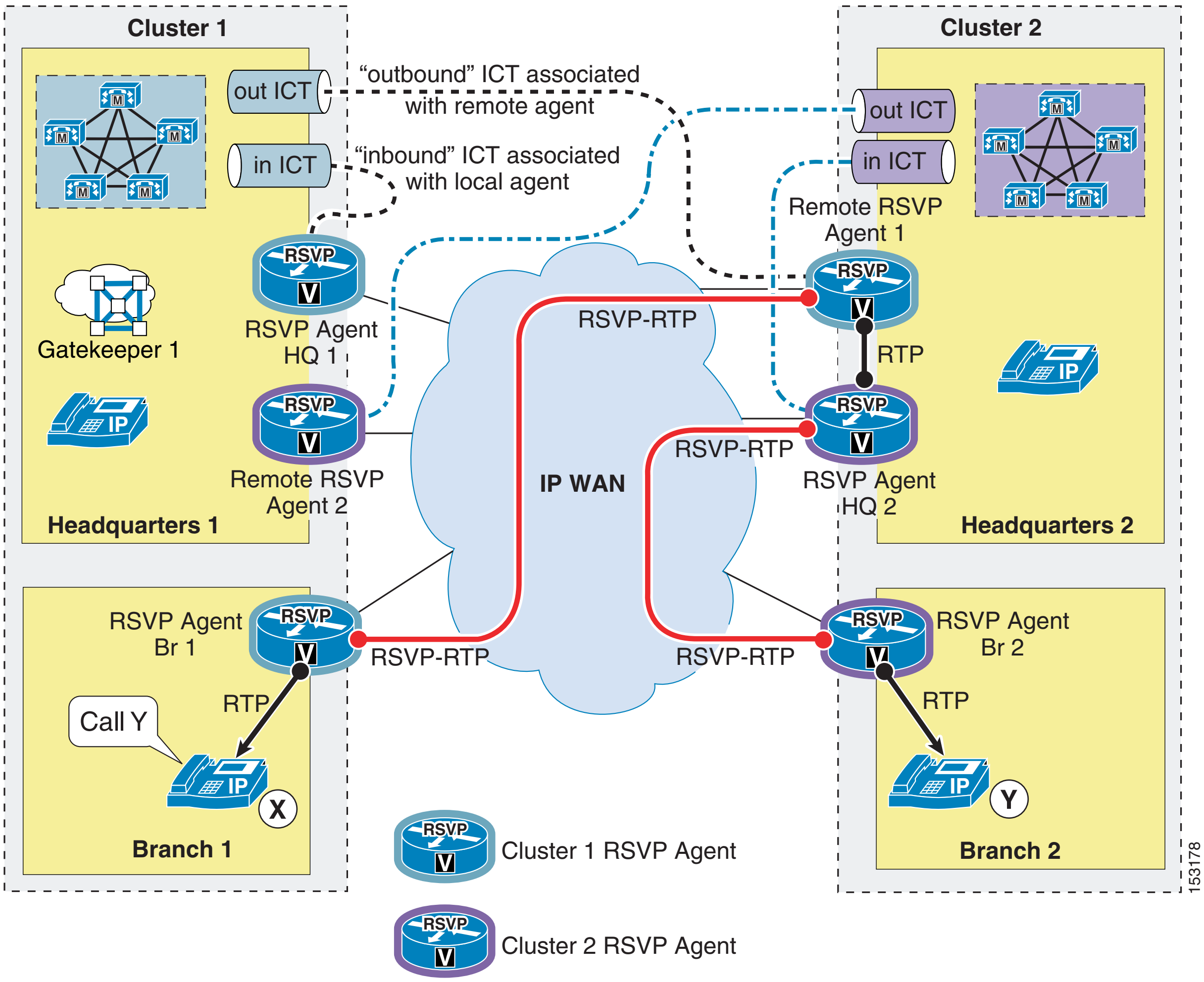
  This solution applies to deployments where any number of Unified CM clusters are located at different sites connected via an IP WAN with limited bandwidth.

Note In deployments where the Unified CM clusters are located at sites connected via a high-speed IP WAN link, it is possible to treat this scenario similar to that described in the section on Co-Located Unified CM Clusters, page 9-54, provided that congestion does not occur in the priority queue of the IP WAN link.

Remote Cisco RSVP Agent Approach

To provide call admission control in generic topologies where three or fewer Unified CM clusters are located at different sites, you can extend the concept of RSVP-enabled locations to cover intercluster calls by defining "remote" Cisco RSVP Agents, as shown in Figure 9-37.

Figure 9-37 Remote Cisco RSVP Agent Approach for Distributed Clusters in a Generic Topology



Note For simplicity, the descriptions in this section are based on an example with two Unified CM clusters, as shown in Figure 9-37. For deployments with three Unified CM clusters, some additional remarks are provided at the end of the section.

In addition to the guidelines listed in the section on Single Unified CM Cluster, page 9-53, observe the following best practices for these deployments:

- For each cluster, define two intercluster trunks (ICTs) to enable communications to the other cluster: one "outbound" intercluster trunk and one "inbound" intercluster trunk.
- Configure a Cisco IOS gatekeeper for dial plan resolution (not for call admission control), defining one zone per Unified CM cluster. For example:

```
gatekeeper
  zone local cluster1 customer.com 10.10.10.10
  zone local cluster2 customer.com
```

- For each cluster, register the inbound trunk with the gatekeeper in the normal zone for that cluster. (For example, the inbound trunk of Cluster 1 registers in zone cluster1, while the inbound trunk of Cluster 2 registers with zone cluster2.)
- For each cluster, register the outbound trunk with the gatekeeper in specially created zones. For example:

```
gatekeeper
  zone local cluster1 customer.com 10.10.10.10
  zone local cluster2 customer.com
  zone local cluster1-to-cluster2 customer.com
  zone local cluster2-to-cluster1 customer.com
```

- Set up your Unified CM dial plan so that outbound calls to the other cluster use the outbound intercluster trunk. (For example, in Cluster 1, have a 2XXX route pattern that points to the outbound trunk via a route list and route group construct.)
- Set up your gatekeeper dial plan so that calls destined to a given cluster are routed to its inbound trunk. For example:

```
gatekeeper
  zone local cluster1 customer.com 10.10.10.10
  zone local cluster2 customer.com
  zone local cluster1-to-cluster2 customer.com
  zone local cluster2-to-cluster1 customer.com
  zone prefix cluster1 1...
  zone prefix cluster2 2...
```

- Assign the inbound trunk to the same location as all devices located at that site (for example, location HQ1 for Cluster 1's inbound trunk and location HQ2 for Cluster 2's inbound trunk).
- Assign the outbound trunk to a newly created location (for example, location remote-to-HQ2 for Cluster 1's outbound trunk toward Cluster 2, and location remote-to-HQ1 for Cluster 2's outbound trunk toward Cluster 1).
- At each of the two sites where Unified CM resides, have an instance of Cisco RSVP Agent registered with the local cluster (for example, Cisco RSVP Agent HQ1 at site HQ1 and Cisco RSVP Agent HQ2 at site HQ2).
- Assign these local Cisco RSVP Agents to the same location as all devices located at that site (for example, location HQ1 for Cisco RSVP Agent HQ1 and location HQ2 for Cisco RSVP Agent HQ2).
- For each cluster, assign the local Cisco RSVP Agent to an MRG contained in the MRGL of all devices located at the central site, including the MRGL used by the inbound trunk via its device pool.
- At each of the two sites where Unified CM clusters reside, add an instance of Cisco RSVP Agent registered with the other Unified CM cluster. (For example, Remote Cisco RSVP Agent 1 is registered with Cluster 1 and placed at site HQ2, where Cluster 2 resides, while Remote Cisco RSVP Agent 2 is registered with Cluster 2 and is placed at site HQ1, where Cluster 1 resides).

- Assign these remote Cisco RSVP Agents to the locations created for the outbound trunks (for example, location remote-to-HQ2 within Cluster 1 for Remote Cisco RSVP Agent 1 and location remote-to-HQ1 within Cluster 2 for Remote Cisco RSVP Agent 2).
- For each cluster, assign the remote Cisco RSVP Agent to an MRG contained in the MRGL used by the outbound trunk (via its device pool).

Note While logically separate, the remote Cisco RSVP Agent instances may reside on the same router platform as the local Cisco RSVP Agent registered to the other cluster. For the example in Figure 9-37, this means that Remote Cisco RSVP Agent 1 and Cisco RSVP Agent HQ2 may actually be located on the same router platform, and the same applies to Remote Cisco RSVP Agent 2 and Cisco RSVP Agent HQ1.

In Figure 9-37, when phone X at Branch 1 places a call to phone Y at Branch 2, Unified CM Cluster 1 will route the call to the gatekeeper via the outbound trunk. Because phone X is assigned to location Branch 1 and the outbound trunk is associated to the location remote-to-HQ2, Unified CM Cluster 1 will initiate an RSVP reservation across the IP WAN between Cisco RSVP Agent Br 1 and Remote Cisco RSVP Agent 1. (The latter is located at HQ2, together with Cluster 2.)

The gatekeeper will then route the call to the inbound trunk of Cluster 2, based on its zone prefix configuration.

Unified CM Cluster 2 then receives a call from the inbound trunk (associated with location HQ1) and phone Y (associated with location Branch 2), so it will initiate another RSVP reservation across the IP WAN between Cisco RSVP Agent HQ2 and Cisco RSVP Agent Br 2.

The call is established across five call legs (four call legs if the Remote RSVP Agent 1 and the RSVP Agent HQ2 are co-resident), two of which traverse the IP WAN and are RSVP-enabled.

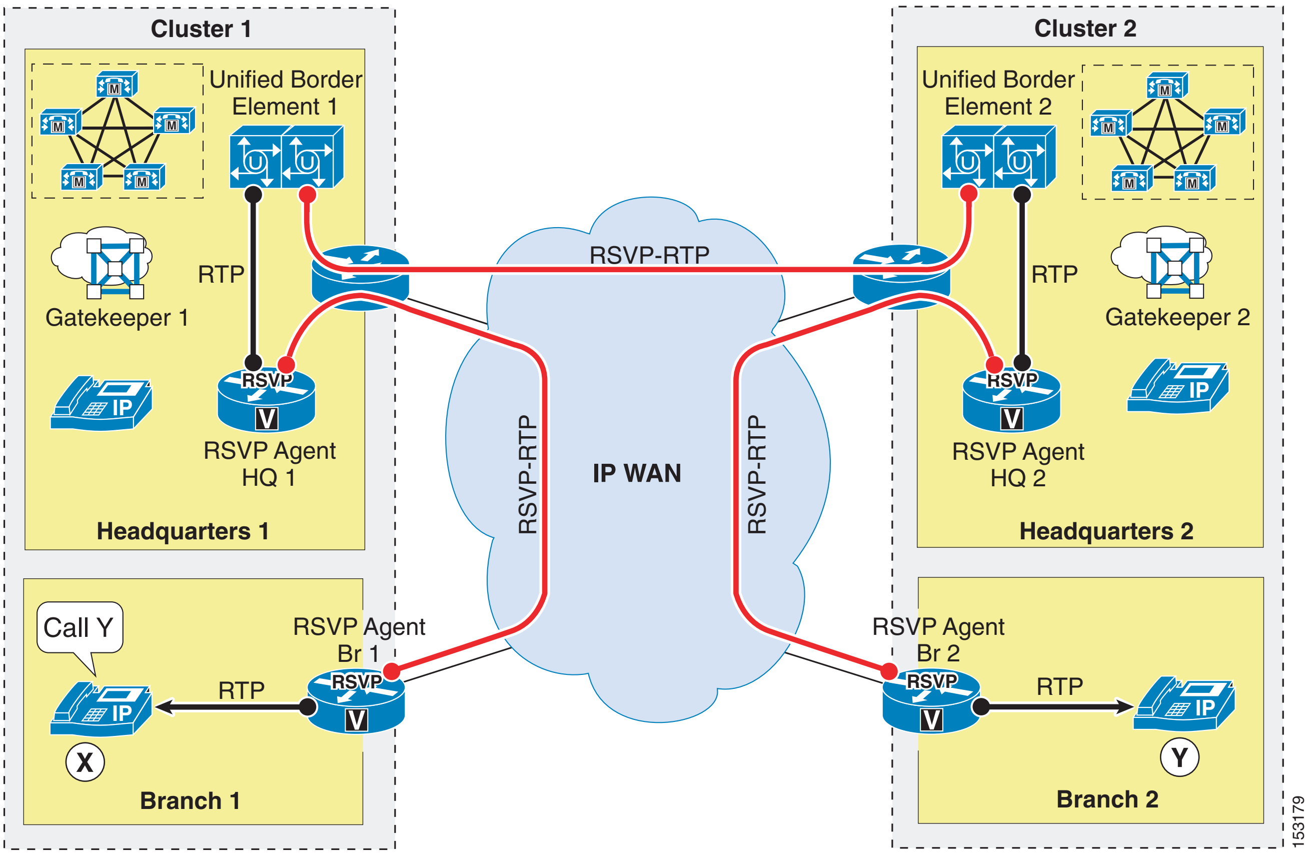
Note In presence of deployments involving three Unified CM clusters, the same considerations apply, except that instead of having a single outbound trunk per cluster, you need two of them, one for each of the other two clusters that will be called. Similarly, each cluster needs two remote Cisco RSVP Agents, each located with one of the other two clusters.

Cisco Unified Border Element Approach

The complexity of the remote Cisco RSVP Agent approach described in the previous section quickly increases with the number of Unified CM clusters, and it is therefore limited to a maximum of three clusters.

To provide call admission control in generic topologies where more than three Unified CM clusters are located at different sites, you can combine RSVP-enabled locations for calls within a cluster with RSVP-enabled Cisco Unified Border Elements for calls between clusters, as shown in Figure 9-38.

Figure 9-38 Cisco Unified Border Element Approach for Distributed Clusters in a Generic Topology



In addition to the guidelines listed in the section on Single Unified CM Cluster, page 9-53, observe the following best practices for these deployments:

- For each cluster, define a gatekeeper-controlled intercluster trunk to enable communications to the other clusters. (Gatekeeper zones are used for dial plan resolution but are not needed for call admission control in this scenario.)
- Assign the intercluster trunk to the same location used for all devices located at the central site for that cluster.
- Ensure that the intercluster trunk is assigned to a device pool that specifies an MRGL that in turn points to an MRG containing the Cisco RSVP Agent located at the central site (for example, Cisco RSVP Agent HQ1 for Cluster 1 in Figure 9-38).
- For each cluster, place a Cisco Unified Border Element at the central site for that cluster and enable it to use RSVP for VoIP calls across the IP WAN.

- For each cluster, configure a gatekeeper as a via-zone gatekeeper so that it invokes the local Cisco Unified Border Element for all calls in or out of the respective zone. (Note that the gatekeeper can be co-resident with the Cisco Unified Border Element.)
- Use the AAR feature to provide automatic PSTN failover in case of call admission control failure within a cluster.
- Use the route list and route group constructs to provide automatic PSTN failover in case of call admission control failure across clusters.
- For calls between two branch sites belonging to different clusters, both media and signaling traffic are hair-pinned via the central sites of the respective clusters (as shown in Figure 9-38, where the call between phone X in Branch 1 and phone Y in Branch 2 is hair-pinned via the HQ1 and HQ2 sites).

Note While logically separate, the Cisco RSVP Agents, gatekeeper, and Cisco Unified Border Element functions may reside on the same router platform. For example, in the scenario shown in Figure 9-38, Cisco Unified Border Element 1, Gatekeeper 1, and Cisco RSVP Agent HQ1 may resident on the same router platform, as may Cisco Unified Border Element 2, Gatekeeper 2, and Cisco RSVP Agent HQ2.

In Figure 9-38, when phone X at Branch 1 places a call to phone Y at Branch 2, Unified CM Cluster 1 will route the call to Gatekeeper 1 via the intercluster trunk. Because phone X is assigned to location Branch 1 and the intercluster trunk is associated to the location HQ1, Unified CM Cluster 1 will initiate an RSVP reservation across the IP WAN between Cisco RSVP Agent Br 1 and Cisco RSVP Agent HQ1.

Gatekeeper 1 then routes the call to Cisco Unified Border Element 1 based on the via-zone configuration, and Cisco Unified Border Element 1 establishes an RSVP reservation with Cisco Unified Border Element 2 across the IP WAN. Cisco Unified Border Element 2 in turn contacts Unified CM Cluster 2 via Gatekeeper 2.

Unified CM Cluster 2 then receives a call from the intercluster trunk associated with location HQ2 and directed to phone Y (associated with location Branch 2), so it will initiate another RSVP reservation across the IP WAN between Cisco RSVP Agent HQ2 and Cisco RSVP Agent Br 2.

The call is established across seven call legs, three of which traverse the IP WAN and are RSVP-enabled.

Design Recommendations for Call Admission Control

This section briefly summarizes the best practices for providing call admission control in various Cisco Unified Communications Manager (Unified CM) deployments.

The following recommendations apply to deployments with a single Unified CM cluster:

- For simple hub-and-spoke topologies with no dual links, use Unified CM static locations. Leave the hub site devices in the Hub_None location.
- For Multiprotocol Label Switching (MPLS) topologies with no dual links, use Unified CM static locations, with devices at every site (including the central site) assigned to a location.
- For any other topology, use Unified CM RSVP-enabled locations. Cisco recommends the **Mandatory** or **Mandatory (video desired)** policy as the default RSVP policy between sites. The Cisco RSVP Agent feature may reside on the IP WAN router in smaller sites or run on standalone platforms in larger sites.

The following recommendations apply to deployments with multiple Unified CM clusters:

- For simple hub-and-spoke topologies with no dual links, use Cisco IOS gatekeeper zones between sites where Unified CM clusters reside.
- For two-tier hub-and-spoke topologies with no dual links where Unified CM clusters are located at the first and second level hub sites, use Cisco IOS gatekeeper zones for the links between first- and second-level hub sites and use Unified CM static locations for the links between second-level hub sites and spoke sites.
- For MPLS topologies with no dual links, use Unified CM static locations, with every site in a location and with no gatekeeper zones. Leave intercluster trunks in the Hub_None location unless an MTP is required. You may use a gatekeeper for intercluster call routing, but it is not needed for call admission control.
- For any other topology and three or fewer clusters, use RSVP-enabled locations and the "remote agent" approach.
- For any other topology and more than three clusters, use RSVP-enabled locations within each cluster and gatekeepers with RSVP-enabled Cisco Unified Border Elements across clusters.