Multinode Scalability and WAN Deployments

Multinode Scalability Feature

Multinode Scalability Requirements

IM and Presence Service supports multinode scalability:

- Six nodes per cluster
- 45,000 users per cluster with a maximum of 15,000 users per node in a full Unified Communication (UC) mode deployment
- 15,000 users per cluster in a presence redundancy group, and 45,000 users per cluster in a deployment with High Availability.
- Administrable customer-defined limit on the maximum contacts per user (default unlimited)
- The IM and Presence Service continues to support intercluster deployments with the multinode feature.

Scalability depends on the number of clusters in your deployment. For detailed VM configuration requirements and OVA templates, see *Virtualization for Unified CM IM and Presence* at the following url: http://docwiki.cisco.com/wiki/Virtualization_for_Unified_CM_IM_and_Presence

Scalability Options for Deployment

IM and Presence Service clusters can support up to six nodes. If you originally installed less than six nodes, then you can install additional nodes at any time. If you want to scale your IM and Presence Service deployment to support more users, you must consider the multinode deployment model you have configured. The following table describes the scalability options for each multinode deployment model.

Table 1: Multinode Scalability Options

Deployment Mode	Scalability Option	
	Add a New Node to an Existing Presence Redundancy Group	**Add a New Node to a New Presence Redundancy Group**
Balanced Non-Redundant High Availability Deployment	If you add a new node to an existing presence redundancy group, the new node can support the same number of users as the existing node; the presence redundancy group can now support twice the number of users. It also provides balanced High Availability for the users on the existing node and the new node in that presence redundancy group.	If you add a new node to a new presence redundancy group, you can support more users in your deployment. This does not provide balanced High Availability for the users in the presence redundancy group. To provide balanced High Availability, you must add a second node to the presence redundancy group.
Balanced Redundant High Availability Deployment	If you add a new node to an existing presence redundancy group, the new node can support the same users as the existing node. For example, if the existing node supports 5000 users, the new node supports the same 5000 users. It also provides balanced redundant High Availability for the users on the existing node and the new node in that presence redundancy group. **Note** You may have to reassign your users within the presence redundancy group, depending how many users were on the existing node.	If you add a new node to a new presence redundancy group, you can support more users in your deployment. This does not provide balanced High Availability for the users in the presence redundancy group. To provide balanced High Availability, you must add a second node to the presence redundancy group.
Active/Standby Redundant High Availability Deployment	If you add a new node to an existing presence redundancy group, you provide High Availability for the users in the existing node in the presence redundancy group. This provides a High Availability enhancement only; it does not increase the number of users you can support in your deployment.	If you add a new node in a new presence redundancy group, you can support more users in your deployment. This does not provide High Availability for the users in the presence redundancy group. To provide High Availability, you must add a second node to the presence redundancy group.

Cluster-Wide DNS SRV

For DNS configuration, you can define a cluster-wide IM and Presence Service address.

The SIP Publish Trunk on Cisco Unified Communications Manager uses the cluster-wide IM and Presence Service address to load-balance SIP PUBLISH messages from Cisco Unified Communications Manager to all nodes in the cluster. Notably this configuration ensures that the initial SIP PUBLISH messages are load-balanced across all nodes in the cluster. This configuration also provides a High Availability deployment as, in the event of a node failing, Cisco Unified Communications Manager will route the SIP PUBLISH messages to the remaining nodes.

The cluster-wide DNS configuration is not a required configuration. Cisco recommends this configuration as a method to load-balance the initial SIP PUBLISH messages across all nodes in the cluster. IM and Presence Service sends subsequent SIP PUBLISH messages for each device to the node where the user is homed on IM and Presence Service.

Even though IM and Presence Service supports multiple domains, you require only a single clusterwide DNS SRV record. You specify that DNS SRV record when you configure the Cisco Unified Communications Manager SIP trunk. Cisco recommends that you use the IM and Presence Service default domain as the destination address for that DNS SRV record.

Note You can specify any domain value as the destination address of the DNS SRV record; however, ensure that the SIP Proxy Service Parameter called SRV Cluster Name on IM and Presence Service matches the domain value you specify in the DNS SRV record. No users need to be assigned to the domain that is specified.

For more information, see topics related to configuring Cisco Unified Communications Manager for integration with IM and Presence Service and DNS SRV records.

Related Topics

Configure Cluster-Wide DNS SRV Name for SIP Publish Trunk

Local Failover

You can also deploy IM and Presence Service over WAN where one presence redundancy group is located in one geographic site, and a second presence redundancy group is located in another geographic site. The presence redundancy group can contain a single node, or a dual node for High Availability between the local nodes. This model provides no failover between geographic sites.

Presence Redundancy Group Failure Detection

The IM and Presence Service supports a failure detection mechanism for a presence redundancy group. Each node in the presence redundancy group monitors the status, or heartbeat, of the peer node. To configure the heartbeat connection and heartbeat intervals on IM and Presence Service, choose **Cisco Unified CM IM and Presence Administration** > **System** > **Service Parameters** > **Server Recovery Manager (service)**. In the section General Server Recovery Manager Parameters (Clusterwide), configure the following parameters:

- **Heart Beat Interval**: This parameter specifies how often in seconds the Server Recovery Manager sends a heartbeat message to the peer Server Recovery Manager in the same presence redundancy group. The heartbeat is used to determine network availability. The default value is 60 seconds.
- **Connect Timeout**: This parameter specifies how long in seconds the Server Recovery Manager waits to receive a response from a connection request to the peer Server Recovery Manager. The default value is 30 seconds.

Note Cisco recommend that you configure these parameters with the default values.

Method Event Routing

When you deploy IM and Presence Service over WAN we recommend that you configure TCP method event routing on IM and Presence Service. Choose **Cisco Unified CM IM and Presence Administration** > **Presence** > **Routing** > **Method/Event Routing** to configure method event routes.

External Database Recommendations

If you configure external database servers in your Clustering over WAN deployment, Cisco recommends that you co-locate the external database servers with the IM and Presence Service nodes that will use the external database servers.

You can connect the IM and Presence Service node to the external database server using either IPv4 or IPv6 Internet transport protocol.

For more information about external database servers and IM and Presence Service, see *Database Setup Guide for IM and Presence Service on Cisco Unified Communications Manager*.

Clustering Over WAN for Intracluster and Intercluster Deployments

IM and Presence Service supports Clustering over WAN for intracluster and intercluster deployments.

Intracluster Deployments Over WAN

IM and Presence Service supports intracluster deployments over WAN, using the bandwidth recommendations provided in this module. IM and Presence Service supports a single presence redundancy group geographically split over WAN, where one node in the presence redundancy group is in one geographic site and the second node in the presence redundancy group is in another geographic location.

This model can provide geographical redundancy and remote failover, for example failover to a backup IM and Presence Service node on a remote site. With this model, the IM and Presence Service node does not need to be co-located with the Cisco Unified Communications Manager database publisher node. The Cisco Jabber client can be either local or remote to the IM and Presence Service node.

This model also supports High Availability for the clients, where the clients fail over to the remote peer IM and Presence Service node if the services or hardware fails on the home IM and Presence Service node. When the failed node comes online again, the clients automatically reconnect to the home IM and Presence Service node.

When you deploy IM and Presence Service over WAN with remote failover, note the following restriction:

- This model only supports High Availability at the system level. Certain IM and Presence Service components may still have a single point of failure. These components are the Cisco Sync Agent, Cisco Intercluster Sync Agent, and Cisco Unified CM IM and Presence Administration interface.

IM and Presence Service also supports multiple presence redundancy groups in a Clustering over WAN deployment. For information about scale for a Clustering over WAN deployment, see the IM and Presence Service SRND.

For additional information, see the IM and Presence Service Solution Reference Network Design (SRND):

Multinode Configuration for Deployment Over WAN

When you configure the IM and Presence Service multinode feature for an intracluster deployment over WAN, configure the IM and Presence Service presence redundancy group, nodes and user assignment as described in the multinode section, but note the following recommendations:

- For optimum performance, Cisco recommends that you assign the majority of your users to the home IM and Presence Service node. This deployment model decreases the volume of messages sent to the remote IM and Presence Service node over WAN, however the failover time to the secondary node depends on the number of users failing over.
- If you wish to configure a High Availability deployment model over WAN, you can configure a presence redundancy group-wide DNS SRV address. In this case, IM and Presence Service sends the initial PUBLISH request message to the node specified by DNS SRV and the response message indicates the host node for the user. IM and Presence Service then sends all subsequent PUBLISH messages for that user to the host node. Before configuring this High Availability deployment model, you must consider if you have sufficient bandwidth for the potential volume of messages that may be sent over the WAN.

Related Topics

Intracluster Deployments Over WAN, on page 4
http://www.cisco.com/go/designzone

Intercluster Deployments

Intercluster Deployments Over WAN

IM and Presence Service supports intercluster deployments over WAN, using the bandwidth recommendations provided in this module.

Related Topics

WAN Bandwidth Requirements

Intercluster Peer Relationships

You can configure peer relationships that interconnect standalone IM and Presence Service clusters, known as intercluster peers. This intercluster peer functionality allows users in one IM and Presence Service cluster

to communicate and subscribe to the availability information of users in a remote IM and Presence Service cluster within the same domain. Keep in mind that if you delete an intercluster peer from one cluster, then you must also delete the corresponding peer in the remote cluster.

IM and Presence Service uses the AXL/SOAP interface to retrieve user information for the home cluster association. IM and Presence Service uses this user information to detect if a user is a local user (user on the home cluster), or a user on a remote IM and Presence Service cluster within the same domain.

IM and Presence Service uses the XMPP interface for the subscription and notification traffic. If IM and Presence Service detects a user to be on a remote cluster within the same domain, IM and Presence Service reroutes the messages to the remote cluster.

Caution Cisco highly recommends that you set up intercluster peers in a staggered manner, as the initial sync uses substantial bandwidth and CPU. Setting up multiple peers at the same time could result in excessive sync times.

Intercluster Router to Router Connections

By default, IM and Presence Service assigns all nodes in a cluster as intercluster router-to-router connectors. When IM and Presence Service establishes an intercluster peer connection between the clusters over the AXL interface, it synchronizes the information from all intercluster router-to-router connector nodes in the home and remote clusters.

You must restart the Cisco XCP Router service on all nodes in both local and remote clusters for IM and Presence Service to establish a connection between the intercluster router-to-router connector nodes. Each intercluster router-to-router connector in one cluster then either initiates or accepts an intercluster connection with router-to-router connectors in the other cluster.

Note In an intercluster deployment, when you add a new node to a cluster, you must restart the Cisco XCP router on all nodes in both the local and remote clusters.

Related Topics

Secure Intercluster Router to Router Connection, on page 7

Node Name Value for Intercluster Deployments

The node name defined for any IM and Presence Service node must be resolvable by every other IM and Presence Service node on every cluster. Therefore, each IM and Presence Service node name must be the FQDN of the node. If DNS is not deployed in your network, each node name must be an IP address.

Note Specifying the hostname as the node name is only supported if all nodes across all clusters share the same DNS domain.

Attention When using the Cisco Jabber client, certificate warning messages can be encountered if the IP address is configured as the IM and Presence Service node name. To prevent Cisco Jabber from generating certificate warning messages, the FQDN should be used as the node name. For instructions to set the IM and Presence Service node name value, see *Cisco Unified Communications Manager Administration Guide*.

Related Topics

IM and Presence Default Domain Value for Intercluster Deployments, on page 7

IM and Presence Default Domain Value for Intercluster Deployments

If you configure an intercluster deployment, note the following:

- The IM and Presence default domain value on the local cluster must match the IM and Presence default domain value on the remote cluster to ensure that intercluster functionality will work correctly.

See topics related to IM and Presence default domain configuration for detailed instructions.

Related Topics

IM and Presence Service Default Domain Configuration

Node Name Value for Intercluster Deployments, on page 6

IM Address Scheme for Intercluster Deployments

For intercluster deployments, all nodes in each of the clusters must use the same IM address scheme. If any node in a cluster is running a version of IM and Presence Service that is earlier than Release 10, all nodes must be set to use the *UserID@Default_Domain* IM address scheme for backward compatibility.

For more information, see topics related to IM address scheme configuration.

Related Topics

Configure IM Address Scheme

IM Address Using UserID@Default_Domain

IM Address Using Directory URI

Secure Intercluster Router to Router Connection

You can configure a secure XMPP connection between all router-to-router connectors in your IM and Presence Service deployment, incorporating both intracluster and intercluster router to router connections. Choose **Cisco Unified CM IM and Presence Administration** > **System** > **Security** > **Settings**, and check **Enable XMPP Router-to-Router Secure Mode**.

When you turn on the secure mode for XMPP router-to-router connections, IM and Presence Service enforces a secure SSL connection using XMPP trust certificates. For intercluster deployments, IM and Presence Service enforces a secure SSL connection between each router-to-router connector node in the local cluster, and each router connector node in the remote cluster.

Related Topics

Intercluster Router to Router Connections, on page 6