Cisco Unity Connection Clusters (Active/Active High Availability)

Cisco Unity Connection Clusters (active/active high availability) and disaster recovery are two key customer requirements for preserving voice messaging services in the event of a system outage or disaster. This chapter discusses the Connection cluster feature in Cisco Unity Connection Release 7.x. For information on disaster recovery, see the “Disaster Recovery” chapter.

Cisco Unity Connection Release 7.x is the first Connection release with the Connection cluster feature.

The Connection cluster feature is not supported for use with Cisco Unified Communications Manager Business Edition.

See the following sections:

- Cisco Unity Connection Cluster Overview, page 8-1
- Publisher Server, page 8-3
- Subscriber Server, page 8-3
- Requirements for Cisco Unity Connection Cisco Unity Connection Cluster, page 8-3
- Support for Installing the Cisco Unity Connection Servers in Separate Buildings or Sites, page 8-3
- Balancing the Load of Calls That the Cisco Unity Connection Servers Handle, page 8-4
- Load Balancing Clients in a Cisco Unity Connection Cluster, page 8-5
- Configuration for Dial-out Voice Messaging Ports, page 8-5
- For More Information, page 8-6

Cisco Unity Connection Cluster Overview

Cisco Unity Connection supports a Connection cluster configuration of two Connection servers to provide high availability and redundancy. The Connection servers in the Connection cluster are within one site and are connected by a LAN. The Connection servers handle calls, HTTP, and IMAP requests. If only one server in the Connection cluster is functioning, the remaining server preserves the system functionality by handling all calls, HTTP requests, and IMAP requests for the Connection cluster. Note that each server in the Connection cluster must have enough voice messaging ports to handle all calls for the Connection cluster.
The first server installed is the publisher server for the Connection cluster; the second server installed is the subscriber server. These terms are used to define the database relationship during installation. The separation of roles is consistent with the Cisco Unified Communications Manager cluster schema in which there is always one publisher server and multiple subscriber servers. (Note that Connection runs on the Cisco Unified CM platform). Unlike a Cisco Unified CM cluster, however, Connection supports only two Connection servers in the Connection cluster. For a network diagram of a Connection cluster integrated with Cisco Unified CM, see Figure 8-1.

Figure 8-1  Cisco Unity Connection Cluster Integrated with Cisco Unified Communications Manager

For systems that do not use web or email clients such as the Cisco Unity Inbox and IMAP clients, a Connection cluster server pair supports up to 10,000 users. In this configuration, both servers can support up to 144 voice messaging ports each for a cumulative total of 288 voice messaging ports when both servers are active. If only one server is active, the port capacity is lowered to a maximum of 144 ports.

For systems that use web or email clients such as the Cisco Unity Inbox and IMAP clients, a Connection cluster server pair supports up to 7,500 users. In this configuration, both servers can support up to 72 voice messaging ports each for a cumulative total of 144 voice messaging ports when both servers are active. If only one server is active, the port capacity is lowered to a maximum of 72 ports.

For more information on capacity planning for a Connection cluster, see the Cisco Unity Connection Supported Platform List at http://www.cisco.com/en/US/products/ps6509/products_data_sheets_list.html.

A Connection cluster server pair supports up to 7,500 IMAP Idle clients. If the IMAP clients that connect to the Connection server do not support IMAP Idle, each of these clients must be counted as 4 IMAP Idle clients. For example, deploying 4 non-IMAP Idle clients is the same as deploying 16 IMAP Idle clients. See the “IMAP Clients Used to Access Connection Voice Messages” section on page 3-6 for a discussion of IMAP Idle and non-IMAP Idle clients.
Publisher Server

The publisher server is required in a Connection cluster, and there can be only one publisher server in a Connection cluster server pair. The publisher server is the first server to be installed, and it provides the database and message store services to the subscriber server in the Connection cluster server pair.

For information on installing a Connection cluster server pair, see the “Overview of Mandatory Tasks for Installing a Cisco Unity Connection 7.x System” chapter of the Installation Guide for Cisco Unity Connection Release 7.x.

As a best practice, we recommend that you direct the majority of client traffic (for example, IMAP and the Cisco Personal Communications Assistant) and administration traffic (for example, Cisco Unity Connection Administration, the Bulk Administration Tool, and backup operations) to the publisher server in a Connection cluster server pair. However, we recommend that the majority of call traffic (for example, SCCP, SIP, or PIMG/TIMG) be directed to the subscriber server in a Connection cluster server pair rather than to the publisher server. Additional call traffic can be directed to the publisher server, if needed, but the call traffic should be directed to the subscriber server first.

Subscriber Server

When installing the subscriber server in a Connection cluster server pair, you provide the IP address or hostname of the publisher server. After the software is installed, the subscriber server subscribes to the publisher server to obtain a copy of the database and message store. There can be only one subscriber server in a Connection cluster server pair.

As a best practice, we recommend that you direct the majority of call traffic (for example, SCCP, SIP, or PIMG/TIMG) to the subscriber server in a Connection cluster server pair. Additional call traffic can be directed to the publisher server, if needed, but the call traffic should be directed to the subscriber server first. Most of the client traffic (for example, IMAP and the Cisco Personal Communications Assistant) and administration traffic (for example, Cisco Unity Connection Administration, the Bulk Administration Tool, and backup operations) should be directed to the publisher server in a Connection cluster server pair. Additional client and administration traffic can be directed to the subscriber server, if needed, but the client and administration traffic should be directed to the publisher server first.

Requirements for Cisco Unity Connection Cisco Unity Connection Cluster

For current Cisco Unity Connection cluster requirements, see the System Requirements for Cisco Unity Connection Release 7.x at http://www.cisco.com/en/US/docs/voice_ip_comm/connection/7x/requirements/7xcucsyseqs.html.

Support for Installing the Cisco Unity Connection Servers in Separate Buildings or Sites

Revised May 2009
Cisco Unity Connection supports the Connection cluster configuration in which the Connection servers are installed in separate buildings or sites. For requirements, see the System Requirements for Cisco Unity Connection Release 7.x at http://www.cisco.com/en/US/docs/voice_ip_comm/connection/7x/requirements/7xcucsreqs.html.

Balancing the Load of Calls That the Cisco Unity Connection Servers Handle

Revised July 9, 2009

Although it is possible to balance the load of calls that the Cisco Unity Connection servers handle in a Connection cluster, we recommend that most call traffic be directed to the subscriber server. This configuration follows the Cisco Unified Communications Manager cluster model of allowing call traffic only on subscriber servers.

Cisco Unified Communications Manager by Skinny Client Control Protocol (SCCP)

When integrating Connection with Cisco Unified CM by Skinny Client Control Protocol (SCCP), it is possible to balance the voice traffic that the Cisco Unity Connection server pair handles by using one of the following methods:

- *(Recommended)* In Cisco Unified Communications Manager Administration (on the Call Routing > Route/Hunt > Line Group page), use Top Down as the distribution algorithm for the line group that contains directory numbers of ports that will answer calls on both servers in the Connection cluster.

  In Connection Administration, all the ports that share the same device name prefix will be in one port group. (If there are ports that share a different device name prefix, they must be in a separate port group.) Beginning with the answering port that has the lowest number in its display name, assign half the answering ports to the subscriber server so that the subscriber server will answer most incoming calls. Assign the remaining answering ports to the publisher server. Then beginning with the dial-out port that has the lowest number in its display name, assign half the dial-out ports to the primary server so that the primary server will handle MWIs and notification calls. Assign the remaining dial-out ports to the subscriber server.

- In Cisco Unified Communications Manager Administration (on the Call Routing > Route/Hunt > Line Group page), use Longest Idle Time as the distribution algorithm for the line group that contains directory numbers of ports that will answer calls on both servers in the Connection cluster.

  In Connection Administration, all the ports will be in a single port group. The first half of the answering ports and dial-out ports will be assigned to the publisher server and the remaining ports will be assigned to the subscriber server in the Connection cluster.

Cisco Unified Communications Manager through a SIP Trunk

When integrating with Cisco Unified CM through a SIP trunk, it is possible to balance voice traffic that the Connection cluster server pair handles by using one of the following methods:

- *(Recommended)* Use a Route List in Cisco Unified CM.

  - Use DNS-SRV – RFC 2782.
  - Use a SIP gateway DNS-SRV.
When integrating with a TDM-based (circuit-switched) phone system through PIMG/TIMG units, it is possible to balance the load of voice traffic that the Connection cluster server pair handles by using one of the following methods:

- **(Recommended)** Turn on load balancing on the PIMG/TIMG units.
- Use load balancing on the TDM based PBX.

**Note** We recommend that you also turn on fault tolerance on the PIMG/TIMG units. This allows the PIMG/TIMG units to redirect calls to either server in the Connection cluster if one server is unavailable to take calls.

---

### Load Balancing Clients in a Cisco Unity Connection Cluster

Although it is possible to balance client and administration requests that the Cisco Unity Connection cluster server pair handles (for example, from the Cisco Personal Communications Assistant (PCA), IMAP, and Cisco Unity Connection Administration), we recommend that most client and administration traffic be directed to the publisher server.

In order to balance client requests, it is necessary to use DNS A-records. DNS A-records allow client DNS lookups to resolve to either server in a round-robin fashion.

**Note** If one server in a Connection cluster server pair stops functioning and failover occurs, clients such as the Cisco PCA and IMAP clients may need to authenticate again by logging in.

We do not recommend using DNS to load balance with multiple A-records because this method does not account for server unavailability (for example, if one of the servers in a Connection cluster server pair stops functioning). The DNS server cannot determine the availability of a server IP address that is listed in an A-record. It may be necessary for the clients to attempt DNS resolution multiple times before they connect to a functioning Connection server in a Connection cluster server pair.

---

### Configuration for Dial-out Voice Messaging Ports

Each Cisco Unity Connection server in a Connection cluster must have voice messaging ports designated for the following dial-out functions in case either server has an outage:

- Sending message waiting indicators (MWIs).
- Performing message notifications.
- Allowing telephone record and playback (TRAP) connections.

As a best practice, we recommend that you dedicate an adequate number of voice messaging ports for these dial-out functions. These dedicated dial-out ports should not receive incoming calls and should not be enabled for answering calls.
For More Information

Configuring Cisco Unity Connection Ports and Port Groups to Support a Cisco Unity Connection Cluster and the Various Phone System Integrations


Configuring Cisco Unity Connection Clients to Support a Cisco Unity Connection Cluster