Configuring Cisco Unity Connection Cluster

- Introduction, on page 1
- Task List for Configuring a Unity Connection Cluster, on page 1
- Administering a Unity Connection Cluster, on page 2
- How a Unity Connection Cluster Works, on page 12
- Effects of Split Brain Condition in a Unity Connection Cluster, on page 13

Introduction

The Cisco Unity Connection cluster deployment provides high availability voice messaging through the two servers that run the same versions of Unity Connection. The first server in the cluster is the publisher server and the second server is the subscriber server.

Task List for Configuring a Unity Connection Cluster

Do the following tasks to create a Unity Connection cluster:

1. Gather Unity Connection cluster requirements. For more information, see *System Requirements for Cisco Unity Connection Release 15* at https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/connection/15/requirements/b_15cucsysreqs.html .
2. Install the publisher server. For more information, see the Installing the Publisher Server section.
3. Install the subscriber server. For more information, see the Installing the Subscriber Server section.
4. Configure the Cisco Unified Real-Time Monitoring Tool for both publisher and subscriber servers to send notifications for the following Unity Connection alerts:
    - AutoFailbackFailed
    - AutoFailbackSucceeded
    - AutoFailoverFailed
    - AutoFailoverSucceeded
    - NoConnectionToPeer
    - SbrFailed

For instructions on setting up alert notification for Unity Connection alerts, see the "Cisco Unified Real-Time Monitoring Tool" section of the Cisco Unified Real-Time Monitoring Tool Administration Guide for the required release, available at http://www.cisco.com/c/en/us/support/unified-communications/unity-connection/products-maintenance-guides-list.html.

5. *(Optional)* Do the following tasks to customize the cluster settings on the publisher server:

   Sign in to Cisco Unity Connection Administration.

   Expand **System Settings > Advanced and** select **Cluster Configuration**.

   On the Cluster Configuration page, change the server status and select Save. For more information on changing the server status in a cluster, see Help> This Page.

Administering a Unity Connection Cluster

You must check the Unity Connection cluster status to ensure that the cluster is correctly configured and working properly. It is also important to understand the different server status in a cluster and the effects of changing a server status in a cluster.

Checking the Cluster Status

You can check the Unity Connection cluster status either using web interface or Command Line Interface (CLI).

Steps to Check the Unity Connection Cluster Status from Web Interface

Procedure

Step 1 Sign in to Cisco Unity Connection Serviceability of either publisher or subscriber server.

Step 2 Expand Tools and select Cluster Management.

Step 3 On the Cluster Management page, check the server status. For more information about server status, see the Server Status and Its Functions in Unity Connection Cluster section.

Steps to Check Unity Connection Cluster Status from Command Line Interface (CLI)

Procedure

Step 1 You can run the show cuc cluster status CLI command on the publisher server or subscriber server to check the cluster status.

Step 2 For more information about server status and its related functions, see the Server Status and Its Functions in Unity Connection Cluster section.

Managing Messaging Ports in a Cluster

In a Unity Connection cluster, the servers share the same phone system integrations. Each server is responsible for handling a share of the incoming calls for the cluster (answering phone calls and taking messages).

Depending on the phone system integration, each voice messaging port is either assigned to a specific server or used by both servers. Table 4-1 describes the port assignments.

Table 1: Server Assignments and Usage of Voice Messaging Ports in a Unity Connection Cluster

Integration Type	Server Assignments and Usage of Voice Messaging Ports
Integration by Skinny Client Control Protocol (SCCP) with Cisco Unified Communications Manager or Cisco Unified Communications Manager Express	• The phone system is set up with twice the number of SCCP voicema that are needed to handle the voice messaging traffic. (For examp voicemail port devices are needed to handle all voice messaging t voicemail port devices must be set up on the phone system.) • In Cisco Unity Connection Administration, the voice messaging p configured so that half the number of the ports set up on the phon assigned to each server in the cluster. (For example, each server i has 16 voice messaging ports.) • On the phone system, a line group, hunt list, and hunt group are c enable the subscriber server answer most of the incoming calls fo • If one of the servers stops functioning (for example, when it is sh maintenance), the remaining server assumes responsibility of han incoming calls for the cluster. • When the server that stopped functioning is able to resume its nor and is activated, it resumes the responsibility of handling its share calls for the cluster.
Integration through a SIP Trunk with Cisco Unified Communications Manager or Cisco Unified Communications Manager Express	• In Cisco Unity Connection Administration, half the number of vo ports that are needed to handle voice messaging traffic are assigned in the cluster. (For example, if 16 voice messaging ports are need all voice messaging traffic for the cluster, each server in the cluste 8 voice messaging ports.) • On the phone system, a route group, route list, and route pattern a to distribute calls equally between both servers in the cluster. • If one of the servers stops functioning (for example, when it is sh maintenance), the remaining server assumes responsibility of han incoming calls for the cluster. • When the server that stopped functioning is able to resume its nor and is activated, it resumes responsibility of handling its share of i for the cluster.

Integration Type	Server Assignments and Usage of Voice Messaging Ports
Integration through PIMG/TIMG units	• The number of ports set up on the phone system is the same as the nu voice messaging ports on each server in the cluster so that the servers the voice messaging ports. (For example, if the phone system is set u voice messaging ports, each server in the cluster must have the same messaging ports.) • On the phone system, a hunt group is configured to distribute calls ec both servers in the cluster. • The PIMG/TIMG units are configured to balance the voice messagin between the servers. • If one of the servers stops functioning (for example, when it is shut d maintenance), the remaining server assumes responsibility of handlin incoming calls for the cluster. • When the server that stopped functioning is able to resume its norma and is activated, it resumes responsibility of handling its share of inco for the cluster.
Other integrations that use SIP	• In Cisco Unity Connection Administration, half the number of voice ports that are needed to handle voice messaging traffic are assigned to e in the cluster. (For example, if 16 voice messaging ports are needed t all voice messaging traffic for the cluster, each server in the cluster h messaging ports.) • On the phone system, a hunt group is configured to distribute calls ec both servers in the cluster. • If one of the servers stops functioning (for example, when it is shut d maintenance), the remaining server assumes responsibility of handlin incoming calls for the cluster. • When the server that stopped functioning is able to resume its normal it resumes responsibility of handling its share of incoming calls for th

Stopping All Ports from Taking New Calls

Follow the steps in this section to stop all the ports on a server from taking any new calls. Calls in progress continue until the callers hang up.

Tip Use the Port Monitor page in the Real-Time Monitoring Tool (RTMT) to determine whether any port is currently handling calls for the server. For more information, see the Step a

Stopping All Ports on a Unity Connection Server from Taking New Calls

Procedure

Step 1 Sign in to Cisco Unity Connection Serviceability.

Step 2 Expand the Tools menu, select **Cluster Management**.

Step 3 On the Cluster Management page, under Port Manager, in the Change Port Status column, select **Stop Taking Calls** for the server.

Restarting All Ports to Take Calls

Follow the steps in this section to restart all the ports on a Unity Connection server to allow them take calls again after they were stopped.

Procedure

Step 1 Sign in to Cisco Unity Connection Serviceability.

Step 2 Expand the Tools menu, select **Cluster Management**.

Step 3 On the Cluster Management page, under Port Manager, in the Change Port Status column, select **Take Calls** for the server.

Server Status and its Functions in a Unity Connection Cluster

Each server in the cluster has a status that appears on the Cluster Management page of Cisco Unity Connection Serviceability. The status indicates the functions that the server is currently performing in the cluster, as described in Table 4-2.

Table 2: Server Status in a Unity Connection Cluster

Server Status	Responsibilities of the Sever in a Unity Connection Cluster
Primary	• Publishes the database and message store both of which are replicated to the other server in t • Receives replicated data from the other server. • Displays and accepts changes to the administrative interfaces, such as Unity Connection Adm and Cisco Unified Operating System Administration. This data is replicated to the other ser cluster. • Answers phone calls and takes messages. • Sends message notifications and MWI requests. • Sends SMTP notifications and VPIM messages. • Synchronizes voice messages in Unity Connection and Exchange mailboxes if the unified m feature is configured. • Connects with the clients, such as email applications and the web tools available through Ci **Note** A server with Primary status cannot be deactivated.
Secondary	• Receives replicated data from the server with Primary status. Data includes the database and store. • Replicates data to the server with Primary status. • Displays and accepts changes to the administrative interfaces, such as Unity Connection Adm and Cisco Unified Operating System Administration. The data is replicated to the server wi status. • Answers phone calls and takes messages. • Connects with the clients, such as email applications and the web tools available through Ci **Note** Only a server with Secondary status can be deactivated.
Deactivated	• Receives replicated data from the server with Primary status. Data includes the database and store. • Does not display the administrative interfaces, such as Unity Connection Administration an Unified Operating System Administration. The data is replicated to the server with Primary • Does not answer phone calls or take messages. • Does not connect with the clients, such as email applications and the web tools available thr Cisco PCA.

Server Status	Responsibilities of the Sever in a Unity Connection Cluster
Not Functioning	• Does not receive replicated data from the server with Primary status. • Does not replicate data to the server with Primary status. • Does not display the administrative interfaces, such as Unity Connection Administration Unified Operating System Administration. • Does not answer phone calls or take messages. **Note** A server with Not Functioning status is usually shut down.
Starting	• Receives replicated database and message store from the server with Primary status. • Replicates data to the server with Primary status. • Does not answer phone calls or take messages. • Does not synchronize voice messages between Unity Connection and Exchange mailbo inbox). **Note** This status lasts only a few minutes, after which the server takes the applicable status.
Replicating Data	• Sends and receives data from the cluster. • Does not answer phone calls or take messages for sometime. • Does not connect with clients, such as email applications and the web tools available th Cisco PCA for sometime. **Note** This status lasts only a few minutes, after which the previous status resumes for the server.
Split Brain Recovery (*After detecting two servers with Primary status)*	• Updates the database and message store on the server that is determined to have Primary • Replicates data to the other server. • Does not answer phone calls or take messages for sometime. • Does not synchronize voice messages between Unity Connection and Exchange mailbo inbox is turned on for sometime. • Does not connect with clients, such as email applications and the web tools available th Cisco PCA for sometime. **Note** This status lasts only a few minutes, after which the previous status resumes for the server.

Changing Server Status in a Cluster and its Effects

The Unity Connection cluster status can be changed either automatically or manually.

You can manually change the status of servers in a cluster in the following ways:

1. A server with Secondary status can be manually changed to Primary status. See the Manually Changing the Server Status from Secondary to Primary section.
2. A server with Secondary status can be manually changed to Deactivated status. See the Manually Activating a Server with Deactivated Status.
3. A server with Deactivated status can be manually activated so that its status changes to Primary or Secondary, depending on the status of the other server. See the Manually Activating a Server with Deactivated Status section.

Manually Changing the Server Status from Secondary to Primary

Procedure

Step 1 Sign in to Cisco Unity Connection Serviceability.

Step 2 From the Tools menu, select **Cluster Management**.

Step 3 On the Cluster Management page, from the Server Manager menu, in the Change Server Status column of the server with Secondary status, select **Make Primary**.

Step 4 When prompted to confirm the change in server status, select **OK**.

The Server Status column displays the changed status when the change is complete.

Note
The server that originally had Primary status automatically changes to Secondary status.

Manually Changing from the Server Status from Secondary to Deactivated

Procedure

Step 1 Sign in to the Real-Time Monitoring Tool (RTMT).

Step 2 From the Cisco Unity Connection menu, select **Port Monitor**. The Port Monitor tool appears in the right pane.

Step 3 In the Node field, select the server with Secondary status.

Step 4 In the right pane, select **Start Polling**. Note whether any voice messaging ports are currently handling calls for the server.

Step 5 Sign in to Cisco Unity Connection Serviceability.

Step 6 From the Tools menu, select **Cluster Management**.

Step 7 If no voice messaging ports are currently handling calls for the server, skip to Step 8.

If there are voice messaging ports that are currently handling calls for the server, on the Cluster Management page, in the Change Port Status column, select **Stop Taking Calls** for the server and then wait until RTMT shows that all ports for the server are idle.

Step 8 On the Cluster Management page, from the Server Manager menu, in the Change Server Status column for the server with Secondary status, select **Deactivate**.

Deactivating a server terminates all the calls that the ports for the server are handling.

Step 9 When prompted to confirm the change in the server status, select **OK**.

The Server Status column displays the changed status when the change is complete.

Manually Activating a Server with Deactivated Status

Procedure

Step 1 Sign in to Cisco Unity Connection Serviceability.

Step 2 From the Tools menu, select **Cluster Management**.

Step 3 On the Cluster Management page, in the Server Manager menu, in the Change Server Status column for the server with Deactivated status, select **Activate**.

Step 4 When prompted to confirm the change in the server status, select **OK**.

The Server Status column displays the changed status when the change is complete.

Effect on Calls in Progress When Server Status Changes in a Unity Connection Cluster

When the status of a Unity Connection server changes, the effect on calls in progress depend upon the final status of the server that is handling a call and on the condition of the network. The following table describes the effects:

Table 3: Effect on Calls in Progress When Server Status Changes in a Unity Connection Cluster

Status Change	Effects
Primary to Secondary	When the status change is initiated manually, calls in progress are not affected. When the status change is automatic, effect on calls in progress depend on the critical service that stopped.
Secondary to Primary	When the status change is initiated manually, calls in progress are not affected. When the status change is automatic, effect on calls in progress depend upon the critical service that stopped.

Status Change	Effects
Secondary to Deactivated	Calls in progress are dropped. To prevent dropped calls, on the Cluster Management page in Cisco Unity Connection Serviceability, select Stop Taking Calls for the server and wait until all the calls get ended and deactivate the server.
Primary or Secondary to Replicating Data	Calls in progress are not affected.
Primary or Secondary to Split Brain Recovery	Calls in progress are not affected.

If network connections are lost, then calls in progress may be dropped depending upon the nature of the network problem.

Effect on Unity Connection Web Applications When the Server Status Changes

The functioning of the following web applications is not affected when the server status changes:

- Cisco Unity Connection Administration
- Cisco Unity Connection Serviceability
- Cisco Unity Connection web tools accessed through the Cisco PCA—the Messaging Assistant, Messaging Inbox, and Personal Call Transfer Rules web tools
- Cisco Web Inbox
- Representational state transfer (REST) API clients

Effect of Stopping a Critical Service on a Unity Connection Cluster

Critical services are necessary for the normal functioning of the Unity Connection system. The effects of stopping a critical service depend upon the server and its status described in the following table:

Table 4: Effects of Stopping a Critical Service on a Unity Connection Cluster

Server	Effects
Publisher	• When the server has Primary status, stopping a critical service in Cisco Unity Connection Serviceability causes the server status to change to Secondary and degrades the ability of the server to function normally. The status of the subscriber server changes to Primary if it does not have the Disabled or Not Functioning status. • When the server has Secondary status, stopping a critical service in Cisco Unity Connection Serviceability degrades the ability of the server to function normally. The status of the servers does not change.
Subscriber	When the server has Primary status, stopping a critical service in Cisco Unity Connection Serviceability degrades the ability of the server to function normally. The status of the servers does not change.

Shutting Down a Server in a Cluster

When a Unity Connection server has Primary or Secondary status, it is handles voice messaging traffic and cluster data replication. We do not recommend you to shutdown both the servers in a cluster at the same time to avoid abrupt termination of the calls and replication that are in progress.

Consider the following points when you want to shutdown a server in a Unity Connection cluster:

- Shutdown the server during non business hours when voice messaging traffic is low.
- Change the server status from Primary or Secondary to Deactivated before shutting down.

Procedure

Step 1 On the server that does not shut down, sign in to Cisco Unity Connection Serviceability.

Step 2 From the Tools menu, select **Cluster Management**.

Step 3 On the Cluster Management page, locate the server that you want to shut down.

Step 4 If the server that you want to shut down has Secondary status, skip to Step 5.

If the server that you want to shut down has Primary status, change the status:

a) In the Change Server Status column for the server with Secondary status, select **Make Primary**.
b) When prompted to confirm the change in the server status, select **OK**.
c) Confirm that the Server Status column indicates that the server has Primary status now and that the server you want to shut down has Secondary status.

Step 5 On the server with Secondary status (the one you want to shut down), change the status:

a) Sign in to the Real-Time Monitoring Tool (RTMT).
b) From the Cisco Unity Connection menu, select **Port Monitor**. The Port Monitor tool appears in the right pane.
c) In the Node field, select the server with Secondary status.
d) In the right pane, select **Start Polling**.
e) Note whether any voice messaging ports are currently handling calls for the server.
f) If no voice messaging ports are currently handling calls for the server, skip to Step5g..

   If there are voice messaging ports that are currently handling calls for the server, on the Cluster Management page, in the Change Port Status column, select **Stop Taking Calls** for the server and then wait until RTMT shows that all ports for the server are idle.

g) On the Cluster Management page, from the Server Manager menu, in the Change Server Status column for the server with Secondary status, select **Deactivate**.

   Caution
   Deactivating a server terminates all calls that the ports for the server are handling.

h) When prompted to confirm the change in the server status, select **OK**.
i) Confirm that the Server Status column indicates that the server now has Deactivated status.

Step 6 Shut down the server that you deactivated:

a) Sign in to Cisco Unity Connection Serviceability.
b) Expand Tools and select Cluster Management.

c) Make sure that the Server Status column shows Not Functioning status for the server that you shutdown.

Replacing Servers in a Cluster

Follow the steps in the given sections to replace publisher or subscriber server in a cluster:

- To replace the publisher server, see the Replacing a Publisher Server section.
- To replace the subscriber server, see the Replacing a Subscriber Server section.

How a Unity Connection Cluster Works

The Unity Connection cluster feature provides high availability voice messaging through two Unity Connection servers that are configured in a cluster.

The Unity Connection cluster behavior when both the servers are active:

- The cluster can be assigned a DNS name that is shared by the Unity Connection servers.
- Clients, such as email applications and the web tools available through the Cisco Personal Communications Assistant (PCA) can connect to either of the Unity Connection server.
- Phone systems can send calls to either of the Unity Connection server.
- Incoming phone traffic load is balanced between the Unity Connection servers by the phone system, PIMG/TIMG units, or other gateways that are required for the phone system integration.

Each server in a cluster is responsible for handling a share of the incoming calls for the cluster (answering phone calls and taking messages). The server with Primary status is responsible for the following functions:

- Homing and publishing the database and message store that are replicated to the other server.
- Sending message notifications and MWI requests (the Connection Notifier service is activated).
- Sending SMTP notifications and VPIM messages (the Connection Message Transfer Agent service is activated).
- Synchronizing voice messages between Unity Connection and Exchange mailboxes, if the unified messaging feature is configured (the Unity Connection Mailbox Sync service is activated).

When one of the servers stops functioning (for example, when it is shutdown for maintenance), the remaining server resumes the responsibility of handling all the incoming calls for the cluster. The database and message store are replicated to the other server when its functionality is restored.

When the server that stopped functioning is able to resume its normal functions and is activated, it resumes responsibility of handling its share of incoming calls for the cluster.

Note It is recommended to perform provisioning only on the Publisher server in Active-Active mode and on Subscriber (Acting Primary) in case of cluster failover. The password change and password setting modification for User PIN/Web application should be provisioned on Publisher server in Active-Active mode.

To monitor the server status, the Connection Server Role Manager service runs in Cisco Unity Connection Serviceability on both the servers. This service performs the following functions:

- Starts the applicable services on each server, depending on server status.
- Determines whether critical processes (such as voice message processing, database replication, voice message synchronization with Exchange, and message store replication) are functioning normally.
- Initiates changes to server status when the server with Primary status is not functioning or when critical services are not running.

Note the following limitations when the publisher server is not functioning:

- If the Unity Connection cluster is integrated with an LDAP directory, directory synchronization does not occur, although authentication continues to work when only the subscriber server is functioning. When the publisher server is resumes functioning, directory synchronization also resumes.
- If a digital or HTTPS network includes the Unity Connection cluster, directory updates do not occur, although messages continue to be sent to and from the cluster when only the subscriber server is functioning. When the publisher server is functioning again, directory updates resume.

The Connection Server Role Manager service sends a keep-alive events between the publisher and subscriber servers to confirm that the servers are functioning and connected. If one of the servers stops functioning or the connection between the servers is lost, the Connection Server Role Manager service waits for the keep-alive events and may require 30 to 60 seconds to detect that the other server is not available. While the Connection Server Role Manager service is waiting for the keep-alive events, users signing in to the server with Secondary status are not able to access their mailbox or send messages, because the Connection Server Role Manager service has not yet detected that the server with Primary status (which has the active message store) is unavailable. In this situation, callers who attempt to leave a message may hear dead air or may not hear the recording beep.

Note It is recommended to import and delete the LDAP users from the publisher node only.

Effects of Split Brain Condition in a Unity Connection Cluster

When both the servers in a Unity Connection cluster have Primary status at the same time (for example, when the servers have lost their connection with each other), both servers handle the incoming calls (answer phone calls and take messages), send message notifications, send MWI requests, accept changes to the administrative interfaces (such as Unity Connection Administration), and synchronize voice messages in Unity Connection and Exchange mailboxes if single inbox is turned on. However, the servers do not replicate the database and message store to each other and do not receive replicated data from each other.

When the connection between the servers is restored, the status of the servers temporarily changes to Split Brain Recovery while the data is replicated between the servers and MWI settings are coordinated. During the time when the server status is Split Brain Recovery, the Connection Message Transfer Agent service and the Connection Notifier service (in Cisco Unity Connection Serviceability) are stopped on both servers, so Unity Connection does not deliver any messages and does not send any message notifications. The Connection Mailbox Sync service is also stopped, so Unity Connection does not synchronize voice messages with Exchange (single inbox). The message stores are also briefly dismounted, so that Unity Connection tells users who are trying to retrieve their messages at this point that their mailboxes are temporarily unavailable.

When the recovery process is complete, the Connection Message Transfer Agent service and the Connection Notifier service are started on the publisher server. Delivery of the messages that arrived while during the recovery process may take additional time, depending on the number of messages to be delivered. The Connection Message Transfer Agent service and the Connection Notifier service are started on the subscriber server. Finally, the publisher server has Primary status and the subscriber server has Secondary status. At this point, the Connection Mailbox Sync service is started on the server with Primary status, so that Unity Connection can resume synchronizing voice messages with Exchange if single inbox is turned on.