Cisco Catalyst Center High Availability Guide, Release 2.3.7.x

First Published: 2023-12-20

Last Modified: 2024-04-08

Catalyst Center High Availability Guide, Release 2.3.7.x

This guide provides details of Catalyst Center's high availability (HA) implementation.

Note For a description of disaster recovery functionality in Catalyst Center, see the "Implement Disaster Recovery" chapter in the *Cisco Catalyst Center Administrator Guide*.

Catalyst Center High Availability Overview

Catalyst Center's HA framework is designed to reduce the amount of downtime that results from failures, and make your network more resilient. The HA framework achieves this by providing the near real-time synchronization of changes across your cluster nodes, giving your network a level of redundancy to deal with any issues that arise. The supported synchronization types include:

- Database changes, such as updates related to configuration, performance, and monitoring data.
- File changes, such as report configurations, configuration templates, TFTP root directory, administration settings, licensing files, and the key store.

This guide covers the requirements that need to be met to use HA, the deployment and administration best practices, and the failures, if any.

Note Catalyst Center provides HA support for both the Automation and the Assurance functionalities.

High Availability Requirements

To enable HA in your production environment, the following requirements must be met:

- Your cluster consists of three Catalyst Center appliances with the same number of cores (three second-generation 56-core appliances, for example).

Note To view a listing of the second and third-generation appliances and their corresponding Cisco part number, see the "Maglev Wizard Interface Configuration Order" topic in the *Installation Guide* for your appliance.

- Your secondary appliances must be running the same version of Catalyst Center as the primary appliance.
- Multinode cluster deployments require all of the member nodes to be in the same network and at the same site. The Catalyst Center appliance does not support the distribution of nodes across multiple networks or sites.
- Your cluster's round-trip time (RTT) is 10 milliseconds or less.

High Availability Functionality

Catalyst Center supports a three-node cluster configuration, which provides *both* software and hardware HA. A software failure occurs if a service on a node fails. Software HA involves the ability of the services on a node to be restarted. For example, if a service fails on one node in a three-node cluster, that service is either restarted on the same node or on one of the other two remaining nodes. A hardware failure occurs when the appliance malfunctions or fails. Hardware HA is enabled by the presence of multiple appliances in a cluster, multiple disk drives within each appliance's RAID configuration, and multiple power supplies. As a result, a failure by one of these components can be tolerated until the faulty component is restored or replaced.

Note Catalyst Center does not support a cluster with more than three nodes. For example, a multinode cluster with five or seven nodes is not currently supported.

Fault tolerance for a three-node cluster is designed to handle single-node failure. In other words, Catalyst Center tries to provide HA across specific services even if a single node fails. If two nodes fail, the quorum that is necessary to perform HA operations is lost and the cluster breaks.

Clustering and Database Replication

Catalyst Center provides a mechanism for distributed processing and database replication among multiple nodes. Clustering provides both sharing of resources and features, as well as enabling HA.

Security Replication

In a multinode environment, the security features of a single node are replicated to the other two nodes, including any X.509 certificates or trustpools. After you join the nodes to an existing cluster to form a three-node cluster, the Catalyst Center GUI user credentials are shared across the nodes. However, the CLI user credentials are not shared because they are separate for each node.

Software Upgrade

In a multinode cluster, you can trigger an upgrade of the whole cluster from the Catalyst Center GUI (the GUI represents the entire cluster and not just a single node). An upgrade triggered from the GUI automatically upgrades all the nodes in the cluster.

Note After you initiate a system upgrade (which updates Catalyst Center's core infrastructure), Catalyst Center goes into maintenance mode. In maintenance mode, Catalyst Center is unavailable until the upgrade process is completed. You should take this into account when scheduling a system upgrade. After the system upgrade is complete, you can verify its success in the GUI by choosing **System** > **Software Updates** > **Updates** and checking the installed version.

1. From the top-left corner, click the menu icon and choose **System** > **Software Updates** > **Updates**.
2. In the **System Update** area, confirm that the latest system package is installed.

High Availability Deployment

The topics in this section cover the best practices you should follow when deploying and administering an HA-enabled cluster in your production environment.

Deployment Recommendations

Catalyst Center supports three-node clusters. The odd number of nodes provides the quorum that is necessary to perform any operation in a distributed system. Instead of three separate nodes, Catalyst Center views them as one logical entity accessed through a virtual IP address.

When deploying HA, we recommend the following:

- When setting up a three-node cluster, do not configure the nodes to span a LAN across slow links, because this can make the cluster susceptible to network failures. It can also increase the amount of time needed for a failed service to recover. When configuring the cluster interface on a three-node cluster, ensure that all of the cluster nodes reside in the same subnet.
- Avoid overloading a single interface with management, data, and HA responsibilities, because this might negatively impact the HA operation. At a minimum, use the Cluster and Enterprise interfaces to keep cluster and enterprise traffic separate.
- In the appliance configuration wizards, Catalyst Center prepopulates the **Services Subnet** and **Cluster Services Subnet** fields with link-local (169.x.x.x) subnets. We recommend that you use the default subnets, but you can choose to specify different subnets. If you specify different subnets, they must conform to the IETF RFC 1918 and 6598 specifications for private networks.

  For details, see RFC 1918, Address Allocation for Private Internets, and RFC 6598, IANA-Reserved IPv4 Prefix for Shared Address Space.
- Enable HA during off-hours, because Catalyst Center enters maintenance mode and is unavailable until it finishes redistributing services.

Deploy a Cluster

To deploy Catalyst Center on a three-node cluster with HA enabled, complete the following procedure:

Procedure

Step 1 Configure Catalyst Center on the first node in your cluster.

See the topic that is specific to the configuration wizard you want to use and your appliance type in the *Installation Guide*:

- If you are configuring an appliance using the Maglev configuration wizard, see the "Configure the Primary Node Using the Maglev Wizard" topic.
- If you are configuring an appliance using the browser-based configuration wizard, see the "Configure the Primary Node Using the Advanced Install Configuration Wizard" topic.

Step 2 Configure Catalyst Center on the second node in your cluster.

See the topic that is specific to the configuration wizard you want to use and your appliance type in the *Installation Guide*:

- If you are configuring an appliance using the Maglev Configuration wizard, see the "Configure a Secondary Node Using the Maglev Wizard" topic.
- If you are configuring an appliance using the browser-based configuration wizard, see the "Configure a Secondary Node Using the Advanced Install Configuration Wizard" topic.

Step 3 Configure Catalyst Center on the third node in your cluster.

See the same secondary appliance configuration topic you viewed while completing the preceding step.

Step 4 Activate HA on your cluster:

a) From the top-left corner, click the menu icon and choose **System** > **Settings** > **System Configuration** > **High Availability**.

b) Click **Activate High Availability**.
After you click **Activate High Availability** in the GUI, Catalyst Center enters into maintenance mode. In this mode, Catalyst Center is unavailable until the process completes, which can take several hours. You should take this into account when scheduling an HA deployment.

Note
- Catalyst Center also goes into maintenance mode when you restore the database and perform a system upgrade (not a package upgrade).
- To enable external authentication with a AAA server in a three-node cluster environment, you must configure all the individual Catalyst Center node IP addresses and the virtual IP address for the three-node cluster on the AAA server.

Administer a Cluster

The topics in this section cover the administrative tasks you must complete when HA is enabled in your production environment.

Run maglev Commands

To make any changes to the IP address, static route, DNS server, or **maglev** user password that are currently configured for a Catalyst Center appliance, you'll need to run the `sudo maglev-config update` CLI command.

Typical Cluster Node Operations

The following are the operations that you should complete for the nodes in your cluster, including:

- Shutting down a cluster node before performing planned maintenance.
- Rebooting to restore a node that has been down or to save configuration changes.
- Preparing a node for Return Merchandise Authorization (RMA).
- Updating the Cisco IMC firmware installed on an appliance.

Note You cannot simultaneously reboot or shut down two nodes in an operational three-node cluster, because this breaks the cluster's quorum requirement.

Task	Action
From the CLI, shut down all of the nodes in a three-node cluster.	Run the **sudo shutdown -h now** command on all of the nodes at the same time. When powering nodes back on, be sure to power on all nodes at the same time through Cisco IMC.

Task	Action
Shut down or disconnect one node for maintenance (in situations where you are not just rebooting the node).	Run the following commands: 1. **maglev node drain** *node's-IP-address* 2. **maglev node drain_history** (to confirm that the node drained successfully) 3. **sudo shutdown -h now** (run on the node you are shutting down) After performing maintenance on the node, complete the following steps: 1. Log in to the Cisco IMC GUI as the Cisco IMC user. 2. From the hyperlinked menu, choose **Host Power > Power On** to power on the node. It should take 30–45 minutes for the node to come back up. 3. Run the **magctl node display** command and wait for the node's status to display as `Ready`. 4. Run the **maglev node allow** *node's-IP-address* command. 5. Run the **magctl workflow status** command and wait until its output indicates that the task you initiated in the previous step completed successfully before you proceed. 6. Run the **maglev service nodescale refresh** command, which puts the node in maintenance mode. **Note** Instead of running the command, you can also do the following: a. From the Catalyst Center GUI, click the menu icon and choose **System** > **Settings** > **System Configuration** > **High Availability**. b. Click **Activate High Availability**.
Reboot one or more nodes after making changes that may require a reboot.	Run the **sudo shutdown -r now** command on the relevant nodes.

Task	Action
Prepare a node for RMA.	1. Drain the node: **maglev node drain** *node-IP-address* To confirm that the node drained successfully, run the **maglev node drain_history** command. 2. Shut down the node: **sudo shutdown -h now** 3. Confirm that the node's status is listed as `NotReady,SchedulingDisabled`: **magctl node display** 4. Remove the node from the cluster: **maglev node remove** *node-IP-address* 5. Install the same Catalyst Center version that's already installed on the cluster's other two nodes. 6. Add the node back to the cluster by configuring it as a secondary node (see the *Installation Guide* for your second or third-generation appliance). 7. Enable service distribution, which puts the node in maintenance mode: **maglev service nodescale refresh** **Note** Instead of running the command, you can also do the following: a. From the Catalyst Center GUI, click the menu icon and choose **System** > **Settings** > **System Configuration** > **High Availability**. b. Click **Activate High Availability**.
Update an appliance's Cisco IMC firmware.	Do the following: 1. See the release notes for the Catalyst Center release that's installed on an appliance. In the release notes, the "Supported Firmware" section shows the Cisco IMC firmware version for your Catalyst Center release. 2. See the *Cisco Host Upgrade Utility User Guide* for instructions on updating the firmware. **Note** In a three-node cluster configuration, we recommend that you shut down all three nodes in the cluster before updating the Cisco IMC firmware. However, you can upgrade the cluster nodes individually if that's what you prefer. Follow the steps provided in this table to shut down one or all of the nodes for maintenance.

Replace a Failed Node

If a node fails, complete the following tasks in order to replace it:

1. Remove the failed node from your cluster.

   See Remove the Failed Node, on page 8.

2. Replace the failed node with another node.

   See Add a Replacement Node, on page 8.

Remove the Failed Node

If a node fails because of a hardware failure, you'll need to remove it from the cluster. For assistance with this task, contact the Cisco TAC.

Warning A two-node cluster (a transient configuration that's not supported for normal use) results when one of the following situations occurs:

- During the initial formation of a three-node cluster, only two of the cluster nodes are available.
- In an existing three-node cluster, one of the nodes has failed, or is currently down.

While a two-node cluster is active, you cannot remove either of its nodes.

Add a Replacement Node

After removing the failed node, you can add a replacement node to the cluster.

Complete the following tasks:

- Remove the failed node. For information, see Remove the Failed Node, on page 8.
- Allocate at least 30 minutes to perform this procedure.

Procedure

Step 1 On the replacement node, install the same software version that the other nodes in the cluster are running.

- If you are configuring an appliance using the Maglev Configuration wizard, use the wizard's **Join a Catalyst Center Cluster** option. See the "Configure a Secondary Node Using the Maglev Wizard" topic in the *Installation Guide* for your second or third-generation appliance.
- If you are configuring an appliance using the browser-based configuration wizard, use the wizard's **Join an existing cluster** option. See the "Configure a Secondary Node Using the Advanced Install Configuration Wizard" topic that's specific to your appliance in the *Installation Guide*.

Important In the **Maglev Cluster Details** screen (Maglev Configuration wizard) or the **Primary Cluster Details** screen (Advanced Install configuration wizard), enter the IP address that's configured for the Cluster port on either of the nodes that are still active.

Step 2 After the installation is complete, enter the following command:

magctl node display

The replacement node should show the Ready status.

Step 3 Redistribute services to the replacement node by activating HA on your cluster:

a. From the top-left corner, click the menu icon and choose **System** > **Settings** > **System Configuration** > **High Availability**.

b. Click **Activate High Availability**.

Step 4 Verify that the services have been redistributed:

magctl appstack status

The replacement node should show a `Running` status.

Minimize Failure and Outage Impact

In a typical three-node Catalyst Center cluster, each node is connected to a single cluster switch through the node's cluster port interface. Connectivity with the cluster switch requires two transceivers and a fiber optic cable, any of which can fail. The cluster switch itself can also fail (because of a loss of power or manual restart), which can result in an outage of your Catalyst Center cluster and loss of all controller functionality. To minimize the impact of a failure or outage on your cluster, do one or more of the following:

- Perform management operations such as software upgrades, configuration reloads, and power cycling during noncritical time periods, because these operations can result in a cluster outage.
- Connect your cluster nodes to a switch that supports the in-service software upgrade (ISSU) feature. This feature allows you to upgrade the system software while the system continues to forward traffic, using nonstop forwarding (NSF) with stateful switchover (SSO) to perform software upgrades with no system downtime.
- Connect your cluster nodes to a switch stack, which allows you to connect each cluster node to a different member of the switch stack joined using Cisco StackWise. Because the cluster is connected to multiple switches, the impact of one switch going down is mitigated.

High Availability Failure Scenarios

Nodes can fail because of issues in one or more of the following areas:

- Software
- Network access
- Hardware

When a failure occurs, Catalyst Center normally detects it within 5 minutes and resolves the failure on its own. Failures that persist for longer than 5 minutes might require user intervention.

The following table describes failure scenarios that your cluster might encounter, and how Catalyst Center responds to them. Pay attention to the table's first column, which indicates the scenarios that require action from you in order to restore the operation of your cluster.

Note For a cluster to operate, Catalyst Center's HA implementation requires at least two cluster nodes to be up at any given time.

Requires User Action	Failure Scenario	HA Behavior
Yes	Any node in the cluster goes down.	If you don't already have an Automation backup, we recommend that you perform one. See the "Backup and Restore" chapter in the *Cisco Catalyst Center Administrator Guide*.
No	A node fails, is unreachable, or experiences a service failure for less than 5 minutes.	• When using the virtual IP (VIP), the GUI and northbound interface (NBI) remain accessible. If using the failed node's IP instead, the GUI and NBI are unavailable until the node comes back up with all services running. • Services that were running on the failed node are not migrated to other nodes. After the node is restored: • Data on the restored node is synched with other cluster members. • Pending GUI and NBI calls that have not timed out are completed.

Requires User Action	Failure Scenario	HA Behavior
No	A node fails, is unreachable, or experiences a service failure for longer than 5 minutes.	• Catalyst Center displays a status message indicating that connectivity with a node has been lost. • The GUI remains usable on the remaining two nodes when using the VIP. • Services that were running on the failed node are migrated to other nodes. • The status of services running on the failed node may be set to either NodeLost or Unknown. • The NBI on the failed node is not accessible, while the NBI on the remaining two nodes remain operational. After the node is restored, the following actions take place: • Catalyst Center displays a status message indicating that cluster operation has resumed. • Pending GUI calls that have not timed out are completed. • Service requests that were pending on the failed node are completed on the node that the service was migrated to. • Data on the restored node is synched with other cluster members. • Services that were running on the failed node restart. • All the service requests that were pending on the failed node are stopped. • Assurance GUI selections operate as expected.
Yes	Two nodes fail or are unreachable.	The cluster is broken and the GUI is not accessible until connectivity is restored. • If the nodes recover, operations resume and the data shared by cluster members is synced. • If the nodes do not recover, contact the Cisco TAC for assistance.
Yes	A node fails and needs to be removed from a cluster.	Contact the Cisco TAC for assistance.
No	All the nodes lose connectivity with one another.	The GUI is not accessible until connectivity is restored. After connectivity is restored, operations resume and the data shared by cluster members is synced.

Requires User Action	Failure Scenario	HA Behavior
Yes	A backup is scheduled and a node goes down because of a hardware failure.	Contact the Cisco TAC for a replacement node, as well as assistance with joining the new node to the cluster.
Yes	A red banner in the GUI indicates that a node is down: "`Assurance services are currently down. Connectivity with host <IP-address> has been lost.`"	The banner indicates that the node is down. If the node comes back up, your Assurance functionality is restored. If the failure is related to a hardware failure, do the following: **1.** Remove the node that failed. Contact the Cisco TAC for assistance. **2.** Add a new node to replace the one that failed. See Add a Replacement Node, on page 8.
Yes	A red banner in the GUI indicates that a node is down, but eventually changes to yellow, with this message: "`This IP address is down.`"	The system is still usable. Investigate why the node is down, and bring it back up.
Yes	A failure occurs while upgrading a cluster.	Contact the Cisco TAC for assistance.
No	An appliance port fails.	• Cluster port: A banner appears, indicating the services that are currently unavailable. Service failover is completed within 10 minutes. The areas of the GUI that you can access depend on which services are restored. After the services that were unavailable are fully restored, the banner disappears. • Enterprise port: Catalyst Center might not be able to reach and manage your network. • Management port: Any upgrades and image downloads that are currently in progress fail and the northbound interface operations are affected. **Note** If a node's IP address is used for GUI or northbound interface operations and the node's management (on either the Enterprise or Management interface) goes down, the current user session is lost. You'll need to manually start a new session, pointing to one of the other node's IP or virtual IP address.

Requires User Action	Failure Scenario	HA Behavior
Yes	Appliance hardware fails.	Replace the hardware component (such as a fan, power supply, or disk drive) that failed. Because multiple instances of these components are found in an appliance, the failure of one component can be tolerated temporarily. As the RAID controller syncs a newly added disk drive with the other drives on the appliance, there might be a degradation in performance on the I/O system while this occurs.

Explanation of Pending State During a Failover

A pod that is in Pending state behaves as follows:

- Stateful set: The pod has some type of data storage. These pods are node bound using local persistent volume (LPV)—when the node is down, all the stateful sets on that node move to Pending state. Stateful examples are Mongodb, Elasticsearch, and Postgres.
- DaemonSet: By design, the pod is strictly node bound. DaemonSet examples are agent, broker-agent, and keepalived.
- Stateless/deployment:
  - While the pod doesn't have a data to store of its own, it uses a stateful set to store or retrieve data.
  - Deployment scale varies. Some deployments have 1x pod instance (such as spf-service-manager-service); some have 2x pod instances (such as apic-em-inventory-manager-service); and some have 3x pod instances (such as kong, platform-ui, collector-snmp).
  - The 1x stateless pods are free to move across nodes based on the current state of the cluster.
  - The 2x stateless pods have flexibility to move across nodes, but no two instances of stateless pods can run on the same node.
  - The 3x stateless pods have node antiaffinity, meaning no two instances can run on the same node.