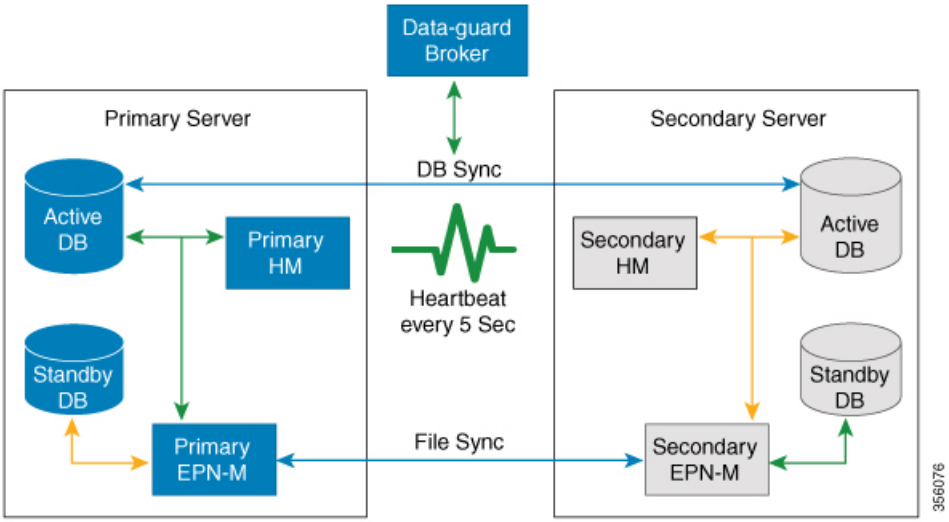
Configure and Manage High Availability

How High Availability Works

The Cisco EPN Manager high availability (HA) framework ensures continued system operation in case of failure. HA uses a pair of linked, synchronized Cisco EPN Manager servers to minimize or eliminate the impact of application or hardware failures that may take place on either server. Servers can fail due to issues in one or more of the following areas:

- Application processes—Server, TFTP, FTP, and other process failures. You can view the status of these processes using the CLI **ncs status** command.
- Database server—Database-related process failures (the database server runs as a service on Cisco EPN Manager).
- Network—Problems with network access or reachability.
- System—Problems with the server's physical hardware or operating system.
- Virtual machine (if HA is running in a VM environment)—Problems with the VM environment on which the primary and secondary servers are installed.

The following figure shows the main components and process flows for an HA setup.



An HA deployment consists of a primary and a secondary server with Health Monitor (HM) instances (running as an application process) on both servers. When the primary server fails (either automatically or because it is manually stopped), the secondary server takes over and manages the network while you restore access to the primary server. If the deployment is configured for automatic failover, the secondary server takes over the active role within two to three minutes after the failover. This HA is based on the *active/passive* or *cold standby* model of operation. Because it is not a clustered system, when the primary server fails, the sessions are not preserved in the secondary server.

When issues on the primary server are resolved and the server is in a running state, it remains in standby mode during which it begins syncing its data with the active secondary server. When the primary is available again, you can initiate a failback operation. When a failback is triggered, the primary server again takes over the active role. This role switching between the primary and secondary servers happens within two to three minutes.

Whenever the HA configuration determines that the primary server has changed, it synchronizes this change with the secondary server. These changes are of two types:

- File changes, which are synchronized using the HTTPS protocol. This includes items such as report configurations, configuration templates, TFTP-root directory, administration settings, licensing files, and the key store. File synchronization is done:
  - In batches, for files that are not updated frequently (such as license files). These files are synchronized once every 500 seconds.
  - Near real-time, for files that are updated frequently. These files are synchronized once every 11 seconds.
- Database changes, such as updates related to configuration, performance and monitoring data. Oracle Recovery Manager (RMAN) creates the initial standby database and Oracle Active Data Guard synchronizes the databases when there is any change.

The primary and secondary HA servers exchange the following messages to maintain synchronization between the two servers:

- Database Sync—Includes all the information necessary to ensure that the databases on the primary and secondary servers are running and synchronized.

- File Sync—Includes frequently updated configuration files. These are synchronized every 11 seconds, while other infrequently updated configuration files are synchronized every 500 seconds.

Note Configuration files that are updated manually on the primary are not synced to the secondary. When you update a configuration file manually on the primary, you must update the file on the secondary as well.

- Process Sync—Ensures that application- and database-related processes are running. These messages fall under the Heartbeat category.
- Health Monitor Sync—These messages check for the network, system, and health monitor failure conditions.

About Primary and Secondary Servers

In any EPN Manager high availability (HA) implementation, for a given instance of a primary server, there must be only one dedicated secondary server.

Typically, each HA server has its own IP address or host name. If you place the servers on the same subnet, they can share the same IP using virtual IP addressing, which simplifies device configuration.

To configure a HA deployment, you can use either a network interface eth0 or a NIC teaming interface. If you use NIC teaming interface for HA deployment, then you must designate it as "Northbound Interface". For more information, see NIC Teaming with HA, on page 12

Note Configuring virtual IP on the NIC teaming interface may work. However, this type of configuration is not officially certified.

Once HA is set up, you must avoid changing the IP addresses or host names of the HA servers. This breaks the HA setup.

For more information, see Reset the Server IP Address or Host Name, on page 39.

Note

- For HA configured servers, the EPNM title bar displays the type of server you are connected to, that is, whether you are connected to a primary server or secondary server.
- If the primary server is active and the secondary server is down for a period that is longer than the configured retention time, the HA configuration will be removed. By default, the configured retention time is 6 hours.

Planning HA Deployments

The HA feature supports the following deployment models:

- **Local**: Both of the HA servers are located on the same subnet (giving them Layer 2 proximity), usually in the same data center.
- **Campus**: Both HA servers are located in different subnets connected via LAN. Typically, they will be deployed on a single campus, but at different locations within the campus.
- **Remote**: Each HA server is located in a separate, remote subnet connected via WAN. Each server is in a different facility. The facilities are geographically dispersed across countries or continents.

The following sections explain the advantages and disadvantage of each model, and discusses underlying restrictions that affect all deployment models.

HA will function using any of the supported deployment models. The main restriction is on HA's performance and reliability, which depends on the bandwidth and latency criteria discussed in "Network Throughput Restrictions on HA". As long as you are able to successfully manage these parameters, it is a business decision (based on business parameters, such as cost, enterprise size, geography, compliance standards, and so on) as to which of the available deployment models you choose to implement.

Network Throughput Restrictions on HA

Cisco EPN Manager HA performance is always subject to the following limiting factors:

- The net bandwidth available to Cisco EPN Manager for handling all operations. These operations include (but are not restricted to) HA registration, database and file synchronization, and triggering failback.
- The net latency of the network across the links between the primary and secondary servers. Irrespective of the physical proximity of these two servers, high latency on these links can affect how the Cisco EPN Manager maintains sessions between the primary and secondary servers.
- The net throughput that can be delivered by the network that connects the primary and secondary servers. Net throughput varies with the net bandwidth and latency, and can be considered a function of these two factors.

These limits apply to at least some degree in every possible deployment model, although some models are more prone to problems than others. For example: Because of the high level of geographic dispersal, the Remote deployment model is more likely to have problems with both bandwidth and latency. But both the Local and Campus models, if not properly configured, are also highly susceptible to problems with throughput, as they can be saddled by low bandwidth and high latency on networks with high usage.

You rarely see throughput problems affecting a failback or failover, as the two HA servers are in more or less constant communication and the database changes are replicated quickly. Most failovers and failbacks take approximately two to three minutes.

The main exception to this rule is the delay for a full database copy operation. This kind of operation is triggered when the primary server has been down for more than the data retention period and you then bring it back up. The data retention period for the express, express-plus, and standard configurations server is six hours and for professional and Gen 2 appliance server it is 12 hours.

Cisco EPN Manager triggers a full database copy operation from the secondary to the primary. No failback is possible during this period, although the Health Monitor page displays any events encountered while the database copy is going on. When the copy is complete, the primary server goes to the "Primary Synching" state, and you can then trigger failback. Be sure not to restart the primary server or disconnect it from the network while the full database copy is in progress.

Variations in net throughput during a full database copy operation, irrespective of database size or other factors, can mean the difference between a database copy operation that completes successfully in under an hour and one that does not complete at all. Cisco has tested the impact of net throughput on HA deployment in configurations following the Remote model, using typical Cisco EPN Manager database sizes of 105–156 GB. Based on these tests, Cisco recommends for a typical database of 125 GB (generating a 10 GB backup file):

- For best results: With sub-millisecond latency, and net throughput of 977 Mbps or more, expect a complete database copy time of one hour or less.
- For good results: With latency of 70 milliseconds, and net throughput of 255 Mbps or more, expect a complete database copy time of two hours or less.
- For acceptable results: With latency of 220 milliseconds or less, and net throughput of 86 Mbps or more, expect a complete database copy time of 4.5 hours or less.

With latencies of 330 ms or higher, and throughput of 46 Mbps or less, you run the risk of the database copy not completing successfully.

Related Topics

Planning HA Deployments, on page 3
Using the Remote Model, on page 6

Using the Local Model

The main advantage of the Local deployment model is that it permits use of a virtual IP address as the single management address for the system. Users can use this virtual IP to connect to Cisco EPN Manager, and devices can use it as the destination for their SNMP trap and other notifications.

The only restriction on assigning a virtual IP address is to have that IP address in the same subnet as the IP address assignment for the primary and secondary servers. For example: If the primary and secondary servers have the following IP address assignments within the given subnet, the virtual IP address for both servers can be assigned as follows:

- Subnet mask: 255.255.255.224 (/27)
- Primary server IP address: 10.10.101.2
- Secondary server IP address: 10.10.101.3
- Virtual IP address: 10.10.101.[4-30] e.g., 10.10.101.4. Note that the virtual IP address can be any of a range of addresses that are valid and unused for the given subnet mask.

In addition to this main advantage, the Local model also has the following advantages:

- Usually provides the highest bandwidth and lowest latency.
- Simplified administration.
- Device configuration for forwarding syslogs and SNMP notifications is much easier.

The Local model has the following disadvantages:

- Being co-located in the same data center exposes them to site-wide failures, including power outages and natural disasters.

- Increased exposure to catastrophic site impacts will complicate business continuity planning and may increase disaster-recovery insurance costs.

Using the Campus Model

The Campus model assumes that the deploying organization is located at one or more geographical sites within a city, state or province, so that it has more than one location forming a "campus". This model has the following advantages:

- Usually provides bandwidth and latency comparable to the Local model, and better than the Remote model.
- Is simpler to administer than the Remote model.

The Campus model has the following disadvantages:

- More complicated to administer than the Local model.
- Does not permit use of a virtual IP address as the single management address for the system, so it requires more device configuration (see "What If I Cannot Use Virtual IP Addressing?" in Related Topics).
- May provide lower bandwidth and higher latency than the Local model. This can affect HA reliability and may require administrative intervention to remedy (see "Network Throughput Restrictions on HA" in Related Topics).
- While not located at the same site, it will still be exposed to city, state, or province-wide disasters. This may complicate business continuity planning and increase disaster-recovery costs.

Related Topics

Planning HA Deployments, on page 3
Network Throughput Restrictions on HA, on page 4
Using the Local Model, on page 5
Using the Remote Model, on page 6
What If I Cannot Use Virtual IP Addressing?, on page 10

Using the Remote Model

The Remote model assumes that the deploying organization has more than one site or campus, and that these locations communicate across geographical boundaries by WAN links. It has the following advantages:

- Least likely to be affected by natural disasters. This is usually the least complex and costly model with respect to business continuity and disaster recovery.
- May reduce business insurance costs.

The Remote model has the following disadvantages:

- More complicated to administer than the Local or Campus models.
- Does not permit use of a virtual IP address as the single management address for the system, so it requires more device configuration (see "What If I Cannot Use Virtual IP Addressing?" in Related Topics).

- Usually provides lower bandwidth and higher latency than the other two models. This can affect HA reliability and may require administrative intervention to remedy (see "Network Throughput Restrictions on HA" in Related Topics).

Related Topics

Planning HA Deployments, on page 3
Network Throughput Restrictions on HA, on page 4
Using the Local Model, on page 5
Using the Campus Model, on page 6
What If I Cannot Use Virtual IP Addressing?, on page 10

Automatic Versus Manual Failover

Configuring HA for automatic failover reduces the need for network administrators to manage HA. It also reduces the time taken to respond to the conditions that provoked the failover, since it brings up the secondary server automatically.

However, we recommend that the system be configured for Manual failover under most conditions. Following this recommendation ensures that the Cisco EPN Manager does not go into a state where it keeps failing over to the secondary server due to intermittent network outages. This scenario is most likely when deploying HA using the Remote model. This model is often especially susceptible to extreme variations in bandwidth and latency.

If the failover type is set to Automatic and the network connection goes down or the network link between the primary and secondary servers becomes unreachable, there is also a small possibility that the primary and secondary servers activate at the same time. We refer to this as the "split brain scenario".

To prevent this, the primary server always checks to see if the secondary server is Active. When the network connection or link is restored and the primary is able to reach the secondary again, the primary server checks the secondary server's state. If the secondary state is Active, then the primary server goes down on its own. Users can then trigger a normal, manual failback to the primary server.

This scenario *only* occurs when the primary HA server is configured for Automatic failover. Configuring the primary server for Manual failover eliminates the possibility of this scenario. This is another reason why we recommend Manual failover configuration.

Automatic failover is especially ill-advised for larger enterprises. If a particular HA deployment chooses to go with Automatic failover anyway, an administrator may be forced to choose between the data that was newly added to the primary or to the secondary. This means, essentially, that there is a possibility of data loss whenever a split-brain scenario occurs. To tackle this issue, see "How to Recover From Split-Brain Scenario" in Related Topics.

To ensure that HA is managed correctly, Cisco recommends that the Cisco EPN Manager administrators always confirm the overall health of the HA deployment before initiating failover or failback, including:

- The current state of the primary.
- The current state of the secondary.
- The current state of connectivity between the two servers.

Related Topics

Planning HA Deployments, on page 3
Network Throughput Restrictions on HA, on page 4

Trigger Failback, on page 22
How to Recover From Split-Brain Scenario, on page 32

Set Up High Availability

The Cisco Evolved Programmable Network Manager Installation Guide describes how to install the primary and secondary servers in your high availability deployment. As part of the installation, your administrator configures your HA deployment to use manual or automatic failover. You can check the current failover setting using the **ncs ha status** command or by checking the Health Monitor web page (see Use the Health Monitor Web Page, on page 18).

After the primary and secondary servers are installed, you must perform the HA configuration steps described in How to Configure HA Between the Primary and Secondary Servers, on page 10.

The following topics provide additional information about the HA deployment:

- Using Virtual IP Addressing With HA, on page 8
- What If I Cannot Use Virtual IP Addressing?, on page 10
- How to Configure HA Between the Primary and Secondary Servers, on page 10
- Configure an SSO Server in an HA Environment, on page 13

Using Virtual IP Addressing With HA

A virtual IP address represents the management IP address of the active HA server. During failover or failback, the virtual IP address automatically switches between the two HA servers. This provides two benefits:

- You do not need to know which server is active in order to connect to the Cisco EPN Manager web GUI. Using a virtual IP, your requests are automatically forwarded to the HA server that is active.
- You do not need to configure managed devices to forward notifications to both the primary server and the secondary server. Notifications only need to be forwarded to the virtual IP address.

Virtual IP addressing can be enabled when you configure the secondary server with the primary server. You will need to provide the virtual address (IPv4 is mandatory while IPv6 is optional) that you want both servers to share. See How to Configure HA Between the Primary and Secondary Servers, on page 10.

Using virtual IP addresses does not change the fact that active client-server sessions are terminated when a failover or failback occurs. Even though the virtual IP address will remain available, active client-server sessions (web GUI or NBI) are terminated as the new server begins servicing new requests. Web GUI users will have to log out and back in. For information on handling broken NBI sessions, see the Cisco Evolved Programmable Network Manager MTOSI API Guide for OSS Integration.

Note To use a virtual IP, the IP addresses of the primary and secondary servers must be on the same subnet.

Multiple Virtual IP Addressing with HA

With Cisco EPN Manager, you can configure up to three interfaces to have their own virtual IP address. In addition, a team (logical binding) of multiple interfaces can be configured with a virtual IP address. There are two ways to do this.

- (**recommended**) Configure all virtual IP addresses from CLI.

  In this case, do not select the **Enable Virtual IP** check box in Cisco EPN Manager UI. This field check box is automatically populated with the first virtual IP address that you configure from CLI.

- Configure the first virtual IP address from the Cisco EPN Manager UI and configure the remaining virtual IP addresses from CLI.

Note To avoid issues during HA registration, ensure that the first virtual IP you configure from CLI matches with what you have configured in UI. In case there is a mismatch, HA registration is blocked and an error message is displayed.

This process is a prerequisite for performing HA registration.

To enable multiple virtual IPs from CLI:

Step 1 Log into the server as the Cisco EPN Manager CLI admin user.

Step 2 Enter configuration mode.

```
configure terminal
```

Step 3 Choose the interface on which you would like to configure the virtual IP.

```
interface <name of interface>
```

Step 4 Enter the following command at the prompt.

```
virtual-ip
```

Step 5 Specify the IPv4 virtual IP address that is to be shared by the primary and secondary HA servers. Optionally, specify an IPv6 virtual IP address (An IPv4 address is mandatory while IPv6 address is optional).

- (mandatory) To configure an IPv4 address:

  ```
  ip-address IPv4 address
  ```

- (optional) To configure an IPv6 address:

  ```
  ipv6-address IPv6 address
  ```

Step 6 Exit the sub-menu.

```
exit
```

Step 7 Exit the interface configuration

```
exit
```

Step 8 Exit the configuration mode.

```
exit
```

Step 9 (Optional) Verify the configuration by running the following command on the interface

```
show running-config
```

The virtual IP addresses are enabled on the primary server once HA registration has completed successfully. The virtual IP addresses are copied to the secondary server during HA registration, but are enabled only in case of a failover.

Note

- The Cisco EPN Manager UI displays the virtual IP configured on the first interface – GigabitEthernet 0 (or Ethernet 0) only. Virtual IP addresses configured on remaining interfaces are not displayed in the web UI.
- To view all virtual IP addresses configured on an interface, run the `show running config` command in CLI.

What If I Cannot Use Virtual IP Addressing?

Depending on the deployment model you choose, not configuring a virtual IP address may result in the administrator having to perform additional steps in order to ensure that syslogs and SNMP notifications are forwarded to the secondary server in case of a failover. The usual method is to configure the devices to forward all syslogs and traps to both servers, usually via forwarding them to a given subnet or range of IP addresses that includes both the primary and secondary server.

This configuration work should be done at the same time HA is being set up: that is, after the secondary server is installed but before HA registration on the primary server. It must be completed before a failover so that the chance of losing data is eliminated or reduced. Not using a virtual IP address entails no change to the secondary server install procedure. The primary and secondary servers must be provisioned with their individual IP addresses, as normal.

Related Topics

Using Virtual IP Addressing With HA, on page 8
Planning HA Deployments, on page 3
Network Throughput Restrictions on HA, on page 4
Using the Campus Model, on page 6
Using the Remote Model, on page 6

How to Configure HA Between the Primary and Secondary Servers

To enable HA, you must configure HA on the primary server. The primary server does not need any configuration during the installation to participate in HA configuration. You must have the following information before configuring the primary server:

- The IP address or host name of the secondary HA server, which you have already installed and configured (installing the secondary server is described in the Cisco Evolved Programmable Network Manager Installation Guide).
- The authentication key that you set during installation of the secondary server.
- (Optional) One or more email addresses, to which notifications are to be sent.

- The Failover type (see Automatic Versus Manual Failover, on page 7).

If you plan to use virtual IP addressing, see Using Virtual IP Addressing With HA, on page 8).

If you plan to use NIC teaming interface, see NIC Teaming with HA, on page 12 for more information.

The following steps explain how to configure HA on a primary server. Follow the same steps while reconfiguring HA.

Before you begin

If you plan to use multiple virtual IP addresses, ensure that you configure them using CLI before this procedure. See Multiple Virtual IP Addressing with HA, on page 9 for more information.

Note If you plan to use only one virtual IP address, you can configure it from the Cisco EPN Manager UI during HA registration. There is no need to configure it through CLI.

Step 1 Log in to Cisco EPN Manager with a user ID and password that has administrator privileges.

Step 2 From the menu, select **Administration** > **Settings** > **High Availability**. Cisco EPN Manager displays the HA status page.

Step 3 Select **HA Configuration** and then complete the following fields:

a. **Secondary Server**: Enter the IP address or the host name of the secondary server.

   Note We recommended that you use a DNS server for resolving the host name to IP address. If you're using the "/etc/hosts" file instead of the DNS server, enter the secondary IP address instead of the host name.

b. **Authentication Key**: Enter the authentication key password that you set during the secondary server installation.

c. **Email Address** (Optional): Enter the address (or comma-separated list of addresses), to which notification about HA state changes should be mailed. If you have already configured email notifications using the Mail Server Configuration page, the email addresses you enter here will be appended to the list of addresses already configured for the mail server.

d. **Failover Type**: Select either **Manual** or **Automatic**. We recommend that you select **Manual**.

Step 4 (Ignore this step and go to Step 5 if you have already configured the virtual IP address using CLI) If you are using the virtual IP feature, then select the **Enable Virtual IP** check box and complete the additional fields as follows:

a. **IPV4 Virtual IP**: Enter the virtual IPv4 address that you want both HA servers to use.

b. **IPV6 Virtual IP**: (Optional) Enter the IPv6 address that you want both HA servers to use.

Note Virtual IP addressing does **not** work unless both the servers are on the same subnet.

Step 5 Click **Check Readiness** to ensure if the HA-related environmental parameters are ready for the configuration.

For more details, see Check Readiness for HA Registration/Configuration, on page 14.

Note The readiness check doesn't block the HA configuration. You can configure HA even if some of the tests do not pass.

Step 6 Click **Save** to save your changes. Cisco EPN Manager initiates the HA configuration process. When the configuration is successfully complete, **Configuration Mode** displays the value **HA Enabled**.

Note If an FTP or TFTP service is running on the primary server, you must restart the secondary server after the configuration is complete to ensure that failover does not fail.

Key points to note:

- The High availability feature does not manage virtual IP addresses added after HA registration. We recommended that you do not add virtual IP addresses after HA registration.
- HA registration failure deletes all configured virtual IP addresses. You must configure them again before HA registration.
- High availability fails if you remove the virtual IP addresses after high availability has been enabled.
- When a fiber is disconnected on a circuit, restore operation is triggered. During restoration if HA switch-over occurs between primary and secondary servers, **Discovery** state of the circuit *might* go to **Partial** on the switched over EPNM server. You can resolve this by manually syncing the devices or scheduling synchronization every night.
- To modify virtual IP addresses you have already configured:
  1. Remove the existing HA configuration.
  2. Configure the virtual IP addresses.
  3. Perform HA registration again.

NIC Teaming with HA

With Cisco EPN Manager, you must designate the NIC teaming interface as the “northbound interface” to be used for HA deployment. NIC teaming designation can be configured from CLI.

NIC teaming interface configuration and the designation as the “northbound interface” must be configured identically on primary and secondary servers as a prerequisite for HA deployment.

Note If the NIC teaming interface is configured with eth0 as a member, then the NIC teaming interface is automatically selected for NBI.

If the NIC teaming interface is configured without eth0 as a member, then the NIC teaming interface will be used only for SBI.

To designate NIC teaming interface as the “northbound interface” from CLI, follow these steps:

Step 1 Log in to Cisco EPN Manager with a user ID and password with CLI administrator privileges.

Step 2 Enter the following command at the prompt:

```
ncs ha northbound interface Team <0-2>
```

Step 3 Specify the NIC teaming interface number that is to be designated as “northbound interface” for HA deployment.

Step 4 Save the configuration:

```
write memory
```

Step 5 (Optional) Verify the configuration by running this command:

```
show running-config
```

Note The above procedure is certified only for NIC teaming interface.

Any other "northbound interface" configuration may work, but is not officially certified.

Configure an SSO Server in an HA Environment

Single Sign-On (SSO) authentication is used to authenticate and manage users in a multi-user, multi-repository environment. SSO is responsible for storing and retrieving the credentials that are used for logging into different systems. You can set up a Cisco EPN Manager as the SSO server for other instances of Cisco EPN Manager.

To configure an SSO server in the high-availability environment, choose one of the procedures listed in the Table 1: SSO Configuration in a HA Deployment. See these topics for more information:

- To configure the SSO server, see Add a RADIUS or TACACS+ Server to Cisco EPN Manager.
- To configure the HA servers, see the Cisco Evolved Programmable Network Manager Installation Guide.

Table 1: SSO Configuration in a HA Deployment

SSO Configuration	Setup SSO Server	Sever Failover Scenario	SSO Server Failure Scenario
SSO as a standalone server	**1.** Configure the standalone SSO server. **2.** Configure the primary and secondary HA servers.	When the primary server fails, the secondary server is activated. All machines that are connected to the primary server will be redirected to the secondary server.	When the SSO server fails, SSO functionality is disabled. Cisco EPN Manager will use local authentication.
SSO on the secondary Server	**1.** Configure one server to be the SSO server and the primary server (in other words, the primary server will also be the SSO server). **2.** Configure the secondary HA server.	When the primary server fails, the secondary server is activated. All machines that are connected to primary server will not be redirected to the secondary server (because SSO is configured on the primary server).	When the SSO (primary) server fails, the secondary server can be set as the failback option for SSO. This enables all instances to connect to the secondary server. If the secondary server is not set as the SSO server failback option, Cisco EPN Manager will use local authentication.

Check Readiness for HA Registration/Configuration

During the HA registration, other environmental parameters related to HA like system specification, network configuration and bandwidth between the servers determine the HA configuration.

An approximate of 15 checks are run in the system to ensure the HA configuration completion without any error or failure. The checklist name and the corresponding status with recommendations if any, will be displayed when you run the Check Readiness feature.

To check readiness for HA configuration, follow these steps:

Step 1 Log in to Cisco EPN Manager with a user ID and password that has administrator privileges.

Step 2 From the menu, select **Administration** > **Settings** > **High Availability**. Cisco EPN Manager displays the HA status page.

Step 3 Select **HA Configuration**.

Step 4 Provide the secondary server IP address in the **Secondary Server** field and secondary Authentication Key **Authentication Key** field.

Step 5 Click **Check Readiness**.

A pop-up window with the system specifications and other parameters are displayed. The screen shows the Checklist Item name, Status, Impact, and Recommendation details.

Below is the list of checklist test name and the description displayed for Check Readiness:

Table 2: Checklist name and description

Checklist Test Name	Test Description
SYSTEM - Check CPU Count	This validates the CPU count in primary and secondary server. The CPU count in primary server can be less than or equal to the secondary server.
DATABASE - LISTENER STATUS	This checks if the database listeners are up and running in both primary and secondary server. If there is a failure, the test restarts and reports the status. This checks if all the wcs instances exist under oracle "listener.ora" file. This is executed in both primary and secondary server.
DATABASE - CHECK MEMORY TARGET	This checks for "/dev/shm" database memory target size for HA setup.
DATABASE - CHECK LISTENER CONFIG CORRUPTION	This checks for all the database instances exist under database listener configuration. This is executed in both primary and secondary server.
SYSTEM - HEALTH MONITOR STATUS	This checks whether the health monitor process is running in both primary and secondary server.

SYSTEM - CHECK DISK IOPS	This validates the disk IOPS in both primary and secondary server. The minimum expected disk IOPS is 200 MBps.
NETWORK - CHECK FIREWALL FOR DATABASE PORT ACCESSIBILITY	This checks if the database port 1522 is open in the system firewall. If the port is disabled, the test grants permission for 1522 in the ip tables list.
NETWORK - CHECK NETWORK INTERFACE BANDWIDTH	This checks if the eth0 interface speed matches the recommended speed of 100 Mbps in both primary and secondary server. This test will not measure network bandwidth by transmitting data between primary and secondary server.
NETWORK - CHECK NETWORK BANDWIDTH SPEED	This checks if the network bandwidth speed matches the recommended speed of 100 Mbps in both primary and secondary server. This test measures the network bandwidth by transmitting data between primary and secondary server.
DATABASE - CHECK ONLINE STATUS	This checks if the database files status is online and accessible in both primary and secondary server.
DATABASE - CHECK TNS CONFIG CORRUPTION	This validates if the tnsping is successful in both primary and secondary server.
DATABASE - TNS REACHABILITY STATUS	This checks if all the wcs instances exist under oracle "listener.ora" file. This is executable in both primary and secondary server.
DATABASE - VALIDATE STANDBY DATABASE INSTANCE	This validates if the standby database instance (stbywcs) is available in both primary and secondary server.
SYSTEM - CHECK RAM SIZE	This checks if the disk size of primary server less than or equal to secondary server.
SYSTEM - CHECK SERVER PING REACHABILITY	This ensures that the primary server can run ping check with the remote (secondary) server.

Step 6 Once the check is completed for all the parameters, check their status, and click **Clear** to close the window.

Note The validation failback and failover events during Check Readiness will be sent to the Alarms and Events page; whereas, the registration failure event will not be present in the Alarms and Evens page.

How to Patch HA Servers

You can download and install UBF patches for your HA servers in one of the following ways, depending on your circumstances:

- Install the patch on HA servers that are not currently paired. Cisco recommends this method if you have not already set up HA for Cisco EPN Manager.
- Install the patch on existing paired HA servers using manual failover. This is the method Cisco recommends if you already have HA set up.
- Install the patch on existing paired HA servers using automatic failover.

For details on each method, see the Related Topics.

Related Topics

How to Patch New HA Servers, on page 16
How to Patch Paired HA Servers, on page 18

How to Patch New HA Servers

If you are setting up a new Cisco EPN Manager High Availability (HA) implementation and your new servers are not at the same patch level, follow the steps below to install patches on both servers and bring them to the same patch level.

Step 1 Download the patch and install it on the primary server:

a) Point your browser to the software patches listing for Cisco EPN Manager (see Software Download).
b) Click the **Download** button for the patch file you need to install (the file name ends with a UBF file extension), and save the file locally.
c) Log in to the primary server using an ID with administrator privileges and choose **Administration > Licenses and Software Updates > Software Update**.
d) Click the **Upload** link at the top of the page and browse to the location where you saved the patch file.
e) Select the UBF file and click **OK** to upload the file.
f) Use one of the following options to upload the UBF file.

1. Upload from local computer
   - Click the **Upload from local computer** radio button in the **Upload Update** window.
   - Click **Browse**, navigate to the file, and click **OK**. After the successful upload, the software will appear under the **Files** tab.
2. Copy from server's local disk
   - Click the **Copy from server's local disk** radio button in the **Upload Update** window.
   - Click **Select** , select the UBF file form the **Select file from local disk** pop-up and click **Select**. After the successful upload, the software will appear under the **Files** tab.

g) When the upload is complete: On the Software Upload page, verify that the Name, Published Date and Description of the patch file are correct.
h) Select the patch file and click **Install**.
i) Click **Yes** in the warning pop-up. When the installation is complete, the server will restart automatically. The restart typically takes 15 to 20 minutes.
j) After the installation is complete on the primary server, verify that the Status of Updates table on the Software Update page shows “Installed” for the patch.

Step 2 Install the same patch on the secondary server:

a) Access the secondary server’s Health Monitor (HM) web page by pointing your browser to the following URL:

   https://*ServerIP***:8082**

   where *ServerIP* is the IP address or host name of the secondary server.

b) You will be prompted for the secondary server authentication key. Enter it and click **Login**.
c) Click the HM web page’s **Software Update** link. You will be prompted for the authentication key a second time. Enter it and click **Login** again.
d) Click **Upload Update File** and browse to the location where you saved the patch file.
e) Select the UBF file and click **OK** to upload the file.
f) Click the **Upload** link at the top of the page.
g) Use one of the following options to upload the UBF file.

   1. Upload from local computer
      - Click the **Upload from local computer** radio button in the **Upload Update** window.
      - Click **Browse**, navigate to the file, and click **OK**. After the successful upload, the software will appear under the **Files** tab.
   2. Copy from server's local disk
      - Click the **Copy from server's local disk** radio button in the **Upload Update** window.
      - Click **Select** , select the UBF file form the **Select file from local disk** pop-up and click **Select**. After the successful upload, the software will appear under the **Files** tab.

h) When the upload is complete: On the Software Upload page, confirm that the Name, Published Date and Description of the patch file are correct.
i) Select the patch file and click **Install**.
j) Click **Yes** in the warning pop-up. When the installation is complete, the server will restart automatically. The restart typically takes 15 to 20 minutes.
k) After the installation is complete on the secondary server, verify that the Status of Updates table on the Software Update page shows “Installed” for the patch.

Step 3 Verify that the patch status is the same on both servers, as follows:

a) Log in to the primary server and access its Software Update page as you did in step 1, above. The **Status** column should show **Installed** for the installed patch.
b) Access the secondary server’s Health Monitor page as you did in step 2, above. The **Status** column should show **Installed** for the installed patch

Step 4 Register the servers.

For more information, see Software Download and Stop and Restart Cisco EPN Manager.

How to Patch Paired HA Servers

If your current Cisco EPN Manager implementation has High Availability servers that are not at the same patch level, or you have a new patch you must install on both your HA servers, follow the steps below.

Patching paired HA servers is not supported. You will receive a popup error message indicating that you cannot perform an update on Cisco EPN Manager servers while HA is configured. So, you must first disconnect the primary and secondary servers before attempting to apply the patch.

1. Follow the steps in “Remove HA Via the GUI” (see Related Topics) to disconnect the primary and secondary servers.
2. Follow the steps in “How to Patch New HA Servers” to apply the patch.
3. Follow the steps in “Set Up High Availability” to restore your HA configuration.

Related Topics

Set Up High Availability
Checking High Availability Status
Remove HA Via the GUI, on page 37
How to Patch New HA Servers, on page 16

Monitor HA Status and Events

These topics describe how to monitor the overall health of the HA environment:

Use the Health Monitor Web Page

The Health Monitor is one of the main components that manage the HA operations. Health Monitor instances run on both servers as an application process, with its own web page on each server. It performs the following functions:

- Synchronizes database and configuration data related to HA (this excludes databases that synchronize separately using Oracle Data Guard).

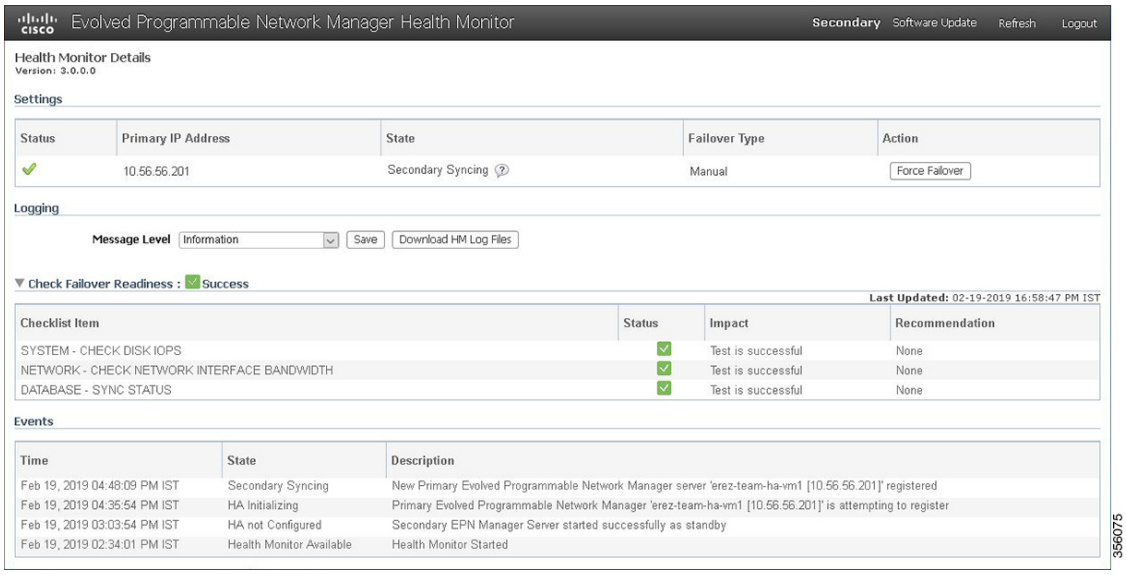
- Exchanges heartbeat messages between the primary and secondary servers every 5 seconds, to ensure communications are maintained between the servers. If the healthy server does not receive 3 consecutive heartbeats from the other redundant server, it waits for 10 seconds. The healthy server then attempts to open a web URL in the redundant server. If this attempt fails, the healthy server becomes the active server.
- Checks the available disk space on both servers at regular intervals and generates events when storage space runs low.
- Manages, controls, and monitors the overall health of the linked HA servers. If there is a failure on the primary server, the Health Monitor activates the secondary server.

After you have completed HA configuration successfully, you can access the Health Monitor web page from the primary or secondary server by entering the following URL on your browser:

`https://`*`ServerIP`***`:8082`**

where *ServerIP* is the primary or secondary server's IP address or host name.

The following example shows a Health Monitor web page for a secondary server in the **Secondary Syncing** state.



1	Settings—Displays the Health Monitor state and configuration detail in five separate sections.	2	Status—Indicates the current functional status of the HA setup (a green check mark indicates HA is enabled and working).
3	Events—Displays the current HA-related events in chronological order, with the most recent events at the top.	4	Primary/Secondary IP address—Displays the IP address of the paired servers. Because this Health Monitor instance is running on the secondary server, it shows the IP address of the primary server.
5	Download—Lets you download the Health Monitor log files.	6	State—Shows the current state of the server on which this Health Monitor instance is running (in this case, the secondary server).

7	Message Level—Indicates the current logging level, which you can change (Error, Informational, or Trace). You must click **Save** to change the logging level.	8	Title bar—Identifies the HA server whose Health Monitor web page you are viewing, along with the Refresh and Logout buttons. Note that the Software Update is only displayed for secondary servers.
9	Failover Type—Shows whether you have Manual or Automatic failover configured.	10	Action—Shows the actions you can perform, such as failover or failback. Only the available actions are displayed here.
11	Check Failover Readiness—Shows the outcome of the disk speed, network interface bandwidth and DB sync status checks after the HA configuration is enabled.		

Note The **Check Readiness** does not block failover to the secondary(either automatic or manual).

Check HA Status and Overall Health

You can use the Cisco EPN Manager web GUI or CLI to check HA status. Either of these approaches will list the state of the server. States are described in HA States and Transitions, on page 34.

To check the HA status from the web GUI, do one of the following:

- From the Cisco EPN Manager web GUI—Choose **Administration** > **Settings** > **High Availability**, then choose **HA Status**. The current HA status and the event states are displayed.
- From the Health Monitor. See Use the Health Monitor Web Page, on page 18.

To check HA status from the CLI, log into either server as a CLI admin user (see Establish an SSH Session With the Cisco EPN Manager Server). The **ncs ha status** command provides a HA-specific output similar to the below example:

```
ncs ha status
[Role] Secondary [Primary Server] cisco-ha1(192.0.2.133) [State] Secondary Active [Failover
 Type] Manual
```

Use the **ncs status** command to check the Health Monitor and other server processes. You will see an output similar to the following example:

```
ncs status
Health Monitor Server is running. ( [Role] Primary [State] Primary Active )
Database server is running
FTP Service is disabled
TFTP Service is disabled
NMS Server is running.
SAM Daemon is running ...
DA Daemon is running ...
```

View and Customize HA Events

HA-related alarms are listed in the Alarms and Events table. A list of these alarms is provided in Cisco Evolved Programmable Network Manager Supported Alarms. The following procedure explains how to view these alarms in the web GUI.

If desired, you can also:

- Adjust the severity for these alarms
- Configure notifications for these alarms

For more information, see Work With Server Internal SNMP Traps That Indicate System Problems.

To view HA-related alarms:

Step 1 Choose **Monitor** > **Monitoring Tools** > **Alarms and Events**, then click the **Alarms** tab.

Step 2 Choose **Quick Filter** from the **Show** drop-down list at the top right of the table.

Step 3 In the **Message** field, enter **High Availability**.

Use HA Error Logging

To save disk space and maximize performance, HA error logging is disabled by default. If you are having trouble with HA, complete the following procedure to enable error logging and examine the log files.

Step 1 Launch the Health Monitor on the server that is having trouble (see Use the Health Monitor Web Page, on page 18).

Step 2 In the **Logging** area, select the error-logging level from the **Message Level** drop-down list and then click **Save**.

Step 3 Download the log files you want to examine:

a. Click **Download**.

A .zip file is copied to your default download location.

b. Extract the log files and use any ASCII text editor to view them.

Trigger Failover

Failover activates the secondary server in response to a failure detected on the primary server.

The Health Monitor detects failure conditions using the heartbeat messages exchanged between the two HA servers. The heartbeat messages are sent every 5 seconds, and if the primary server is not responsive to three consecutive heartbeat messages from the secondary server, the Health Monitor deems the primary server to have failed. During the health check, the Health Monitor also checks the application process status and database health. If there is no proper response to these checks, these are also treated as having failed.

The HA system in the secondary server takes about 15 seconds to detect a process failure on the primary server. If the secondary server is unable to reach the primary server due to a network issue, it might take more time to discover the failure and initiate a failover. In addition, it may take additional time for the application processes on the secondary server to be fully operational.

As soon as the Health Monitor detects a failure, it sends an e-mail notification. The e-mail includes the failure status along with a link to the secondary server's Health Monitor web page. If HA is configured for automatic failover, the secondary server will activate automatically.

To perform a manual failover:

Before you begin

- Check the state of the primary and secondary servers.
- Validate the connectivity between the two servers.
- If you are not using virtual IP addresses, make sure all devices are configured to forward traps and syslogs to both servers.

Step 1 Access the secondary server's Health Monitor web page using the web link given in the email notification, or by entering the following URL on your browser:

`https://ServerIP:8082`

Step 2 Click **Failover**.

Trigger Failback

Failback is the process of re-activating the primary server once it is back online. It also transfers Active status from the secondary server to the primary server, and stops active network monitoring processes on the secondary server.

When a failback is triggered, the secondary server replicates its current database information and updated files to the primary server. The time it takes to complete the failback from the secondary server to the primary server will depend on the amount of data that needs to be replicated and the available network bandwidth.

After the data is replicated successfully, HA changes the state of the primary server to **Primary Active** and the state of the secondary server to **Secondary Syncing**.

During failback, the availability of the secondary server depends on whether the Cisco EPN Manager was re-installed on the primary server after the failover, as follows:

- If Cisco EPN Manager was re-installed on the primary server after the failover, a full database copy will be required and the secondary server will not be available during the failback process.
- If Cisco EPN Manager was not re-installed with primary server, the secondary server is available, except during the period when processes are started on the primary server and stopped on the secondary server. Both servers' Health Monitor web pages are accessible for monitoring the progress of the failback. Additionally, users can also connect to the secondary server to access all normal functionalities.

You must always trigger failback manually, as described in the procedure below. Note:

- Do not initiate configuration or provisioning activity while the failback is in progress.
- After a successful failback, the secondary server will go down and control will switch over to the primary server. During this process, Cisco EPN Manager will be inaccessible to the users for a few moments.

Before you begin

- Check the state of the primary and secondary servers.
- Validate the connectivity between the two servers.
- If you are not using virtual IP addresses, make sure all devices are configured to forward traps and syslogs to both servers.
- If you have re-installed Cisco EPN Manager on the primary server and you are using offline geo maps, you must re-install the geo maps resources on the primary server before triggering failback. See the Cisco Evolved Programmable Network Manager Installation Guide.

Step 1 Access the secondary server's Health Monitor web page using the link given in the e-mail notification, or by entering the following URL on your browser:

`https://ServerIP:8082`

Step 2 Click **Failback**.

Force Failover

A forced failover is the process of making the secondary server active while the primary server is still up. You will want to use this option when, for example, you want to test that your HA setup is fully functional.

Forced failover is available to you only when the primary is active, the secondary is in the “Secondary syncing” state, and all processes are running on both servers. Forced failover is disabled when the primary server is down. In this case, only the normal Failover is enabled.

Once the forced failover completes, the secondary server will be active and the primary will restart in standby automatically. You can return to an active primary server and standby secondary server by triggering a normal failback.

Step 1 Access the secondary server's Health Monitor web page using the steps in Use the Health Monitor Web Page.

Step 2 Trigger the forced failover by clicking the **Force Failover** button. The forced failover will complete in 2 to 3 minutes.

Respond to Other HA Events

All the HA related events are displayed on the HA Status page, the Health Monitor web pages, and under the Cisco EPN Manager Alarms and Events page. Most events require no response from you other than triggering failover and failback. A few events are more complex, as explained in the following topics:

HA Registration Fails

If HA registration fails, you will see the following HA state-change transitions for each server:

Primary HA State Transitions...	Secondary HA State Transitions...
From: HA Initializing	From: HA Initializing
To: HA Not Configured	To: HA Not Configured

To recover from failed HA registration, follow the steps below.

Step 1 Use ping and other tools to check the network connection between the two Cisco EPN Manager servers. Confirm that the secondary server is reachable from the primary, and vice versa.

Step 2 Check that the gateway, subnet mask, virtual IP address (if configured), server hostname, DNS, NTP settings are all correct.

Step 3 Check that the configured DNS and NTP servers are reachable from the primary and secondary servers, and that both are responding without latency or other network-specific issues.

Step 4 Check that all Cisco EPN Manager licenses are correctly configured.

Step 5 Once you have remedied any connectivity or setting issues, retry the steps in How to Configure HA Between the Primary and Secondary Servers, on page 10.

Network is Down (Automatic Failover)

If there is a loss of network connectivity between the two Cisco EPN Manager servers, you will see the following HA state-change transitions for each server, assuming that the Failover Type is set to “Automatic”:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Active	From: Secondary Syncing
To: Primary Lost Secondary	To: Secondary Lost Primary
To: Primary Lost Secondary	To: Secondary Failover
To: Primary Lost Secondary	To: Secondary Active

You get an email notification that the secondary is active.

Step 1 Check on and restore network connectivity between the two servers. Once network connectivity is restored and the primary server can detect that the secondary is active, all services on the primary will be restarted and made passive automatically. You will see the following state changes:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Lost Secondary	From: Secondary Active
To: Primary Failover	To: Secondary Active
To: Primary Syncing	To: Secondary Active

Step 2 Trigger a failback from the secondary to the primary. You will then see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Syncing	From: Secondary Active
To: Primary Failback	To: Secondary Failback
To: Primary Failback	To: Secondary Post Failback
To: Primary Active	To: Secondary Syncing

Network is Down (Manual Failover)

If there is a loss of network connectivity between the two Cisco EPN Manager servers, you will see the following HA state-change transitions for each server, assuming that the Failover Type is set to “Manual”:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Active	From: Secondary Syncing

Primary HA State Transitions...	Secondary HA State Transitions...
To: Primary Lost Secondary	To: Secondary Lost Primary

You will get email notifications that each server has lost the other.

Step 1 Check on and, if needed, restore the network connectivity between the two servers.

You will see the following state changes once network connectivity is restored.:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Lost Secondary	From: Secondary Lost Primary
To: Primary Active	To: Secondary Syncing

No administrator response is required.

Step 2 If network connection cannot be restored for any reason, use the HM web page for the secondary server to trigger a failover from the primary to the secondary server. You will see the following state changes:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Lost Secondary	From: Secondary Lost Primary
To: Primary Lost Secondary	To: Secondary Failover
To: Primary Failover	To: Secondary Active

You will get an email notification that the secondary server is now active.

Step 3 Check and restore network connectivity between the two servers. Once network connectivity is restored and the primary server detects that the secondary server is active, all services on the primary server will be restarted and made passive. You will see the following state changes:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Lost Secondary	From: Secondary Active
To: Primary Failover	To: Secondary Active
To: Primary Syncing	To: Secondary Active

Step 4 Trigger a failback from the secondary to the primary.

You will then see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Syncing	From: Secondary Active
To: Primary Failback	To: Secondary Failback
To: Primary Failback	To: Secondary Post Failback

Primary HA State Transitions...	Secondary HA State Transitions...
To: Primary Active	To: Secondary Syncing

Process Restart Fails (Automatic Failover)

The Cisco EPN Manager Health Monitor process is responsible for attempting to restart any Cisco EPN Manager server processes that have failed. Generally speaking, the current state of the primary and secondary servers should be “Primary Active” and “Secondary Syncing” at the time any such failures occur.

If HM cannot restart a critical process on the primary server, then the primary server is considered to have failed. If your currently configured Failover Type is “automatic”, you will see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Active	From: Secondary Syncing
To: Primary Uncertain	To: Secondary Lost Primary
To: Primary Failover	To: Secondary Failover
To: Primary Failover	To: Secondary Active

When this process is complete, you will get an email notification that the secondary server is now active.

Step 1 Restart the primary server and ensure that it is running. Once the primary is restarted, it will be in the state “Primary Syncing”. You will see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Failover	From: Secondary Active
To: Primary Preparing for Failback	To: Secondary Active
To: Primary Syncing	To: Secondary Active

Step 2 Trigger a failback from the secondary to the primary. You will then see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Syncing	From: Secondary Active
To: Primary Failback	To: Secondary Failback
To: Primary Failback	To: Secondary Post Failback
To: Primary Active	To: Secondary Syncing

Process Restart Fails (Manual Failover)

The Cisco EPN Manager Health Monitor process is responsible for attempting to restart any Cisco EPN Manager server processes that have failed. Generally speaking, the current state of the primary and secondary servers should be “Primary Active” and “Secondary Syncing” at the time any such failures occur. If HM cannot restart a critical process on the primary server, then the primary server is considered to have failed. You will receive an email notification of this failure. If your currently configured Failover Type is “Manual”, you will see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Active	From: Secondary Syncing
To: Primary Uncertain	To: Secondary Lost Primary

Step 1 Trigger on the secondary server a failover from the primary to the secondary. You will then see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Uncertain	From: Secondary Syncing
To: Primary Failover	To: Secondary Failover
To: Primary Failover	To: Secondary Active

Step 2 Restart the primary server and ensure that it is running. Once the primary server is restarted, the primary’s HA state will be “Primary Syncing”. You will see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Failover	From: Secondary Active
To: Primary Preparing for Failback	To: Secondary Active
To: Primary Syncing	To: Secondary Active

Step 3 Trigger a failback from the secondary to the primary. You will then see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Syncing	From: Secondary Active
To: Primary Failback	To: Secondary Failback
To: Primary Failback	To: Secondary Post Failback
To: Primary Active	To: Secondary Syncing

Primary Server Restarts During Synchronization (Manual Failover)

If the primary Cisco EPN Manager server is restarted while the secondary server is syncing, you will see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Active	From: Secondary Syncing
To: Primary Alone	To: Secondary Lost Primary
To: Primary Active	To: Secondary Syncing

The “Primary Alone” and “Primary Active” states occur immediately after the primary comes back online. No administrator response should be required.

Secondary Server Restarts During Synchronization

If the secondary Cisco EPN Manager server is restarted while syncing with the primary server, you will see the following state transitions:

Primary HA State Transitions...	Secondary HA State Transitions...
From: Primary Active	From: Secondary Syncing
To: Primary Lost Secondary	From: Secondary Lost Primary
To: Primary Active	To: Secondary Syncing

No administrator response should be required.

Both HA Servers Are Down

If both the primary and secondary servers are down at the same time, you can recover by bringing them back up in the correct order, as explained in the steps below.

Step 1 Restart the secondary server and the instance of Cisco EPN Manager running on it. If for some reason you cannot restart the secondary server, see “Both HA Servers Are Down and Secondary Will Not Restart” in Related Topics.

Step 2 When the Cisco EPN Manager is running on the secondary, access the secondary server’s Health Monitor web page. You will see the secondary server transition to the state “Secondary Lost Primary”.

Step 3 Restart the primary server and the instance of Cisco EPN Manager running on it. When the Cisco EPN Manager is running on the primary, the primary will automatically sync with the secondary. To verify this, access the primary server’s Health Monitor web page. You will see the two servers transition through the following series of HA states:

Primary HA State Transitions...	Secondary HA State Transitions...
To: Primary Lost Secondary	To: Secondary Lost Primary

Primary HA State Transitions...	Secondary HA State Transitions...
To: Primary Active	To: Secondary Syncing

Related Topics

Accessing the Health Monitor Web Page
Responding to Other HA Events

Both HA Servers Are Powered Down

If both the primary and secondary servers are powered down at the same time, you can recover by bringing them back up in the correct order, as explained in the steps below.

Step 1 Power on the secondary server and the Cisco EPN Manager instance running on it. The secondary HA restart will fail at this state because the primary server is not reachable. However, the secondary server's HM process will be running (with an error).

Step 2 When Cisco EPN Manager is running on the secondary server, access the secondary server's HM web page (see Use the Health Monitor Web Page, on page 18). You will see the secondary server transition to the **Secondary Lost Primary** state.

Step 3 Power on the primary server and the Cisco EPN Manager instance running on it.

Step 4 When Cisco EPN Manager is running on the primary server, the primary server will automatically begin syncing with the secondary server. To verify this, access the primary server's HM web page. You will see the two servers transition through the following series of HA states:

Primary HA State Transitions...	Secondary HA State Transitions...
To: Primary Lost Secondary	To: Secondary Lost Primary
To: Primary Active	To: Secondary Syncing

Step 5 Restart the secondary server and the Cisco EPN Manager instance running on it. This is required because not all processes will be running on the secondary server at this point.

If for some reason you cannot restart the secondary server, see Both HA Servers Are Down and Secondary Server Will Not Restart, on page 30.

Step 6 When Cisco EPN Manager finishes restarting on the secondary server, all processes should be running. Verify this by running the **ncs ha status** command.

Both HA Servers Are Down and Secondary Server Will Not Restart

If both HA servers are down at the same time and the secondary server will not restart, you will need to remove the HA configuration from the primary server in order to use it as a standalone server until you can replace the secondary server.

The following steps assume that you have already tried and failed to restart the secondary server.

Step 1 Attempt to restart the primary instance of Cisco EPN Manager. If the primary server is able to restart at all, the restart will abort with an error message indicating that you must remove the HA configuration.

Step 2 Open a CLI session with the primary server (see Establish an SSH Session With the Cisco EPN Manager Server).

Step 3 Enter the following command to remove the HA configuration on the primary server:

```
ncs ha remove
```

Note Once you remove the HA configuration, you need to reinstall the secondary server since the primary will no longer be able to register with the former secondary.

Step 4 Confirm that you want to remove the HA configuration.

You should now be able to restart the primary instance of Cisco EPN Manager without receiving an error message, and use it as a standalone server. When you are able to replace the secondary server, proceed as explained in How to Configure HA Between the Primary and Secondary Servers, on page 10.

How to Replace the Primary Server

Under normal circumstances, the state of your primary server will be **Primary Active** and your secondary server will be **Secondary Syncing**. If the primary server fails for any reason, a failover to the secondary will take place (automatically or manually).

You may find that restoring full HA access requires you to re-install the primary server using new hardware. If this happens, you can follow the steps below to bring up the new primary server without losing any data.

Before you begin

Make sure you have the password (authentication key) that was set when HA was configured on the secondary server. You will need it for this procedure.

Step 1 Ensure that the secondary server is in the **Secondary Active** state. If the primary server is configured for manual failover, you will need to trigger failover to the secondary server (see Trigger Failover, on page 21).

Step 2 Ensure that the old primary server you are replacing has been disconnected from the network.

Step 3 Ensure that the new primary server is ready for use. This will include connecting it to the network and configuring it similar to the old primary server (IP address, subnet mask, and so forth). You will need to enter the same authentication key that you entered when installing HA on the secondary server.

Step 4 Ensure that both the primary and secondary servers are at the same patch level and if you want to replace the primary server, then you must:

a) Ensure the primary and secondary server are in TOFU Mode by executing the following command in the secondary server CLI:

```
admin# ncs certvalidation certificate-check trust-on-first-use trustzone system
```

b) Login to Secondary server admin CLI.

c) Execute the following command in the secondary server CLI:

```
admin# ncs certvalidation tofu-certs deletecert host <primaryserver's-IP-address
appended with "_8082">
```

For example: `ncs certvalidation tofu-certs deletecert host 10.56.58.91_8082`

This is required to re-establish the communication between the Primary and Secondary servers.

Step 5 Update the IP table entries as listed below:

- On Primary - Add Secondary IP address and Virtual IP address (if configured) in iptables for 1522 port.
- On Secondary - Add Primary IP address and Virtual IP address (if configured) in iptables for 1522 port.

Example:

```
iptables -A INPUT -s IP address -p tcp --dport 1522 -j ACCEPT
iptables -A INPUT -s IP address -j ACCEPT
```

Step 6 Trigger a failback from the secondary server to the newly-installed primary server. During failback to the new primary HA server, a full database copy will be performed, so this operation will take time to complete depending on the available bandwidth and network latency. You will see the two servers transition through the following series of HA states:

Primary HA State Transitions...	Secondary HA State Transitions...
From: HA not configured	From: Secondary Active
To: Primary Failback	To: Secondary Failback
To: Primary Failback	To: Secondary Post Failback
To: Primary Active	To: Secondary Syncing

How to Recover From Split-Brain Scenario

In a split-brain scenario, both the primary and secondary servers become active at the same time, perhaps due to a network outage or a link that temporarily goes down. However, because the primary server constantly checks the secondary server, when the connection is reestablished, the primary server will go down due to the secondary server being active.

The possibility of data loss always exists on the rare occasions when a “split-brain scenario” occurs. In this case, you can choose to save the newly added data on the secondary and forget the data that was added on the primary, as explained in the following steps.

Step 1 Once the network is up, and the secondary server is up, the primary will restart itself automatically, using its standby database. The HA status of the primary server will be, first, “Primary Failover” transitioning to “Primary Syncing”. You can verify this by logging on to the primary server’s Health Monitor web page.

Step 2 Once the primary server’s status is “Primary Syncing, confirm that a user can log into the secondary server’s Cisco EPN Manager page using the web browser (for example, https://server-ip-address:443). Do not proceed until you have verified this.

Step 3 Once access to the secondary is verified, initiate a failback from the secondary server's Health Monitor web page (see Trigger Failback, on page 22). You can continue to perform monitoring activities on the secondary server until the switchover to the primary is completed.

Secondary Server Goes Down

In this scenario, the secondary server is acting as a standby server and it goes down.

To get the secondary server up and running again:

Step 1 Power on the secondary server.

Step 2 Start Cisco EPN Manager on the secondary server.

Step 3 On the primary server, verify that the primary server's HA status changes from "Primary Lost Secondary" to "Primary Active." Go to **Administration** > **Settings** > **High Availability** > **HA Configuration**.

Step 4 Log into the secondary server's Health Monitor page by entering the following URL in your browser: **https://*serverIP*:8082**.

Step 5 Verify that the secondary server's HA status changes from "Secondary Lost Primary" to "Secondary Syncing."
No further action is required once the above statuses are displayed. However, if the HA status does not change, the secondary server cannot be recovered automatically. In this case, continue with the following steps.

Step 6 Remove the HA configuration on the primary server. Go to **Administration** > **Settings** > **High Availability** > **HA Configuration** and click **Remove**.

Step 7 Register the secondary server with the primary server. See How to Configure HA Between the Primary and Secondary Servers, on page 10.
If HA registration is successful, no further action is required. However, if HA registration is unsuccessful, it indicates that the secondary server might have suffered hardware/software loss. In this case, continue with the following steps.

Step 8 Remove the HA configuration on the primary server.

Step 9 Reinstall the secondary server with the same release and patches (if any) as the primary server.

Step 10 Register the secondary server with the primary server. See How to Configure HA Between the Primary and Secondary Servers, on page 10.

How to Resolve Database Synchronization Issues

To resolve the database synchronization issue, when the primary server is in "Primary Active" state and the secondary server is in "Secondary Syncing" state, do the following:

Step 1 Remove HA, see Remove HA Via the CLI, on page 38 and Remove HA Via the GUI, on page 37.

Step 2 After both the primary and secondary servers reaches "HA not configured" state, perform the HA registration. See Set Up High Availability, on page 8

High Availability Reference Information

The following sections supply reference information on HA.

Related Topics

HA Configuration Modes, on page 34
HA States and Transitions, on page 34

HA Configuration Modes

HA configuration modes represent the overall status of the complete HA configuration (as opposed to HA states, which are specific to a server).

Mode	Description
HA Not Configured	HA is not configured on this server.
HA Initializing	HA configuration process between the primary and secondary servers has started.
HA Enabled	HA is enabled between the primary and secondary servers.
HA Alone	Server is running alone because one of the servers is down, out of sync, or unreachable.

HA States and Transitions

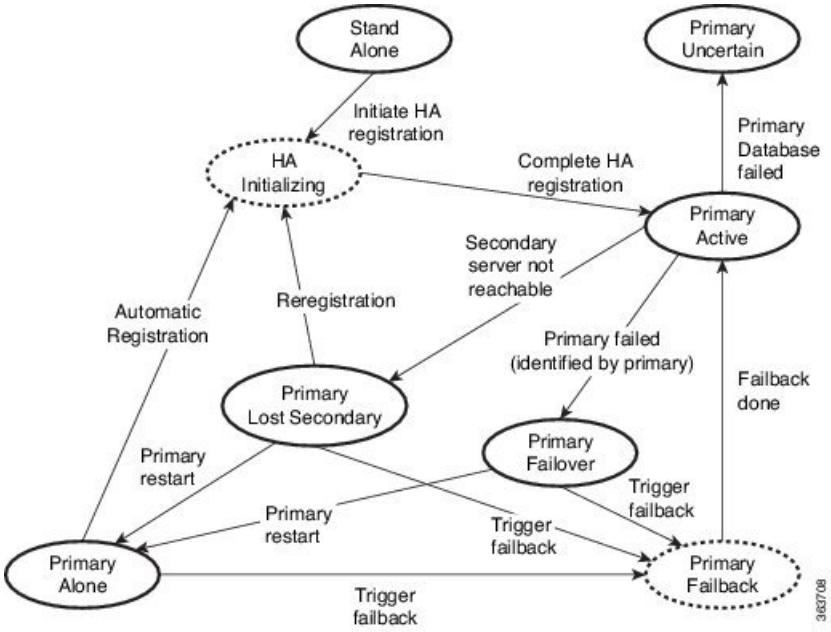
The following table lists the HA states, including those that require no response from you. You can view these states from the HA Status page (**Administration** > **Settings** > **High Availability** > **HA Status**) or from the Health Monitor. For a list of HA events and instructions for enabling, disabling, and adjusting them, see Customize Server Internal SNMP Traps and Forward the Traps.

State	Server	Description
Stand Alone	Both	HA is not configured on this server.
Primary Alone	Primary	Primary server has restarted after it lost the secondary server (only Health Monitor is running in this state).
HA Initializing	Both	HA configuration process between the primary and secondary server has started.
Primary Active	Primary	Primary server is now active and is synchronizing with the secondary server.

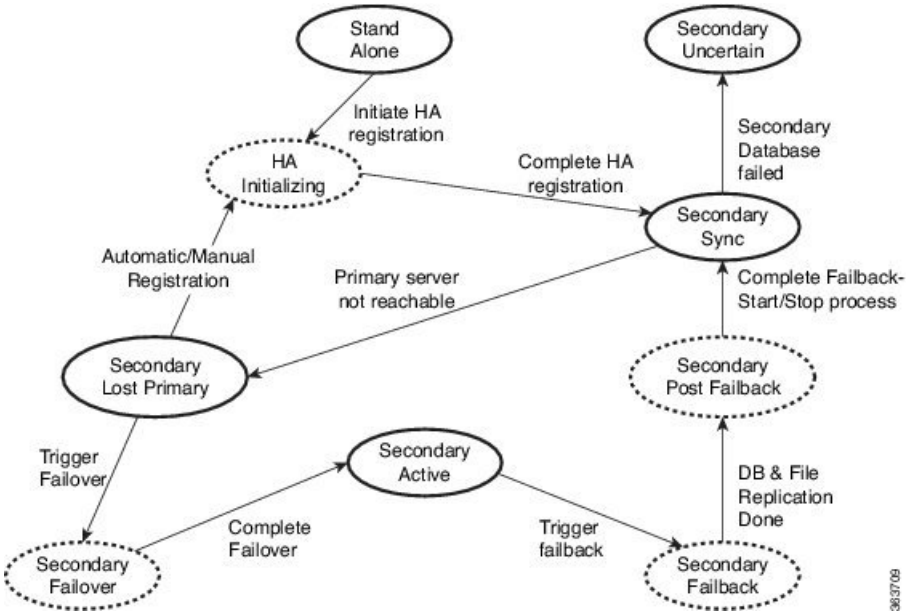
Primary Database Copy Failed	Primary	Restarted primary server detected a data gap, triggered a data copy from the active secondary server, and the database copy failed. When a primary server is restarted, it always checks to see if a data gap has occurred due to the primary server being down for 24 hours or more. This copy rarely fails but if it occurs, all attempts to failback to the primary are blocked until the database copy completes successfully. As soon as it does, the primary state is set to **Primary Syncing**.
Primary Failover	Primary	Primary server detected a failure.
Primary Failback	Primary	User-triggered failback is currently in progress.
Primary Lost Secondary	Primary	Primary server is unable to communicate with the secondary server.
Primary Preparing for Failback	Primary	Primary server has started up in standby mode after a failover (because the secondary server is still active). When the primary server is ready for failback, its state will be set to **Primary Syncing**.
Primary Syncing	Primary	Primary server is synchronizing the database and configuration files from the active secondary server. This occurs after a failover, when primary processes are brought up (and the secondary server is playing the active role).
Primary Uncertain	Primary	Primary server's application processes are not able to connect to its database.
Secondary Alone	Secondary	Primary server is not reachable from secondary server after a primary server restart.
Secondary Syncing	Secondary	Secondary server is synchronizing the database and configuration files from the primary server.
Secondary Active	Secondary	Failover from the primary server to the secondary server has completed successfully.
Secondary Lost Primary	Secondary	Secondary server is not able to connect to the primary server (occurs when the primary fails or network connectivity is lost). For automatic failover, the secondary server will automatically move to the **Secondary Active** state. For Manual failover, you must trigger the failover to make the secondary server active (see Trigger Failover, on page 21).
Secondary Failover	Secondary	Failover triggered and is in progress.
Secondary Failback	Secondary	Failback triggered and database and file replication is in progress.
Secondary Post Failback	Secondary	Failback triggered; associated process stops and restarts are in progress. Database and configuration files have been replicated from the secondary server to the primary server. The primary server status will change to **Primary Active**, and the secondary server HA status will change to **Secondary Syncing**.

Secondary Uncertain	Secondary	Secondary server's application processes cannot connect to the server's database.

The following figure illustrates the primary server HA state changes.



This figure illustrates the secondary server HA state changes.



High Availability CLI Command Reference

The following table lists the CLI commands available for HA management. Log in as admin to run these commands on the primary server (see Connect via CLI):

Table 3: High Availability Commands

Command	Description
ncs ha ?	Get help with high availability CLI commands
ncs ha authkey authkey	Update the authentication key for high availability
ncs ha remove	Remove the High Availability configuration
ncs ha status	Get the current status for High Availability

Related Topics

High Availability Reference Information, on page 33

Reset the HA Authentication Key

Users with administrator privileges can change the HA authentication key using the **ha authkey** command. You will need to ensure that the new authorization key meets the password standards.

Step 1 Log in to the primary server as a Cisco EPN Manager CLI admin user (see Establish an SSH Session With the Cisco EPN Manager Server).

Step 2 Enter the following at the command line:

```
ha authkey newAuthKey
```

Where *newAuthKey* is the new authorization key.

Remove HA Via the GUI

The simplest method for removing an existing HA implementation is via the GUI, as shown in the following steps. You can also remove the HA setup via the command line.

Note that, to use this method, you must ensure that the primary Cisco EPN Manager server is currently in the "Primary Active" state. If for any reason the secondary server is currently active, perform a failback and then try to remove the HA configuration after the failback is complete and the secondary's automatic restart has finished.

Step 1 Log in to the primary Cisco EPN Manager server with a user ID that has administrator privileges.

Step 2 Select **Administration > Settings > High Availability > HA Configuration**.

Step 3 Select **Remove**. Removing the HA configuration takes from 3 to 4 minutes.

Once the removal is complete, ensure that the HA configuration mode displayed on the page now reads “HA Not Configured”.

Remove HA Via the CLI

If for any reason you cannot access the Cisco EPN Manager GUI on the primary server, administrators can remove the HA setup via the command line, using the steps below.

To use this method, you must ensure that the primary Cisco EPN Manager server is currently in the “Primary Active” state. If for any reason the secondary server is currently active, perform a failback, and then try to remove the HA configuration after the failback is complete and the secondary’s automatic restart has finished.

Step 1 Connect to the primary server via CLI. Do not enter “configure terminal” mode.

Step 2 Enter the following at the command line:

admin# **ncs ha remove**. For more information, see Connect via CLI.

Related Topics

Remove HA Via the GUI, on page 37
Trigger Failback, on page 22
High Availability Reference Information, on page 33

Remove HA During Upgrade

To upgrade a Cisco EPN Manager implementation that uses HA, follow the steps below.

Step 1 Use the GUI to remove the HA settings from the primary server (see Remove HA Via the GUI, on page 37).

Step 2 Upgrade the primary server as needed.

Step 3 Reinstall the secondary server using the current image.

Note Upgrading the secondary server from the previous version or a beta version is not supported. The secondary server must always be a fresh installation.

Step 4 Once the upgrade is complete, perform the HA registration process again.

Note After upgrade, health monitor page displays the below health monitor event message:

Primary Authentication Key was changed by Admin

For more information, see Connect via CLI.

Related Topics

High Availability Reference Information, on page 33
How to Configure HA Between the Primary and Secondary Servers, on page 10

Remove HA During Restore

Cisco EPN Manager does not back up configuration settings related to high availability. If you are restoring an implementation that is using HA, you should only restore data to the primary server. The restored primary server will automatically replicate its data to the secondary server. If you try to run a restore on a secondary server, Cisco EPN Manager generates an error message.

Follow these steps when restoring an implementation that uses HA:

1. Use the GUI to remove the HA settings from the primary server. See Remove HA Via the GUI, on page 37.
2. Restore data on the primary server. See Restore Cisco EPN Manager Data.
3. When the restore process is complete, perform the HA configuration process again. See How to Configure HA Between the Primary and Secondary Servers, on page 10.

Reset the Server IP Address or Host Name

Avoid changing the IP address or hostname of the primary or secondary server, if possible. If you must change the IP address or hostname, remove the HA configuration from the primary server before making the change. When finished, re-register HA.

Resolve TOFU Failure at Any State

When the primary and secondary servers communicate, there is a possibility of a TOFU error as mentioned below.

You must correct the following error(s) before proceeding. 'A Trust-on-first-use (TOFU) based Certificate is configured for this connection. The current certificate on the remote host is different than what was used earlier.

To resolve this issue:

- Clear the existing certificate using the NCS CLI command on both the primary and secondary servers.

```
ncs certvalidation tofu-certs deletecert host <server-hostname>
```