Managing Nodes

Managing Nodes

Nodes are initially added to a storage cluster using the Create Cluster feature of the HX Data Platform Installer. Nodes are added to an existing storage cluster using the Expand Cluster feature of the HX Data Platform Installer. When nodes are added or removed from the storage cluster, the HX Data Platform adjusts the storage cluster status accordingly.

- Tasks for node maintenance with a failed nodes.

    - The ESXi or HX software needs to be reinstalled.

    - A node component needs to be replaced.

    - The node needs to be replaced.

    - The node needs to be removed.

- Tasks for node maintenance with a non-failed nodes.

    - Putting the node into maintenance mode.

    - Changing the ESX password.

Note Though there are subtle differences, the terms **server**, **host**, and **node** are used interchangeably throughout the HyperFlex documentation. Generally a server is a physical unit that runs software dedicated to a specific purpose. A node is a server within a larger group, typically a software cluster or a rack of servers. Cisco hardware documentation tends to use the term node. A host is a server that is running the virtualization and/or HyperFlex storage software, as it is 'host' to virtual machines. VMware documentation tends to use the term host.

Step 1 Monitor the nodes in the cluster.

HX storage cluster, node, and node component status is monitored and reported to HX Connect, HX Data Platform Plug-in, vCenter UI, and assorted logs as Operational status (online, offline) and Resiliency (healthy, warning) status values.

Note Functional state distinctions contribute to, but are separate from, the storage cluster operational and resiliency status reported in the HX Connect and HX Data Platform Plug-in views. For each Data Replication Factor (2 or 3), Cluster Access Policy (lenient or strict), and given number of nodes in the storage cluster, the storage cluster shifts between Read and Write, Read Only, or Shutdown state, depending on the number of failed nodes or failed disks in nodes.

Note A replication factor of three is highly recommended for all environments except HyperFlex Edge. A replication factor of two has a lower level of availability and resiliency. The risk of outage due to component or node failures should be mitigated by having active and regular backups.

Step 2 Analyze the node failure and determine the action to take.

This frequently requires monitoring the node state through HX Connect, HX Data Platform Plug-in, vCenter, or ESXi; checking the server beacons; and collecting and analyzing logs.

Step 3 Complete the identified tasks.

- Reinstall or upgrade software.

  For steps to reinstall ESXi or the HX Data Platform see *Cisco HyperFlex Systems Installation Guide for VMware ESXi*. For steps to upgrade software, see the *Cisco HyperFlex Systems Upgrade Guide*.

- Repair a component in the node.

  Node components, such as solid state drives (SSD), hard disk drives (HDD), power supply units (PSU), and network interface cards (NIC) components are not configurable through HX Connect or HX Data Platform Plug-in, but the HX Data Platform monitors them and adjusts the storage cluster status when any of these items are disrupted, added, removed, or replaced.

  The steps to add or remove disks, depends upon the type of disk. Field replaceable units (FRUs), such as PSUs and NICs are replaced following steps described in the server hardware guides.

- Replace a node in the cluster.

  Replacing a node in a storage cluster typically requires TAC assistance. Provided the requirements are met, nodes can be replaced without TAC assistance while the storage cluster is online (5+ node clusters only) or offline (4+ node clusters). To replace a node in a 3 node cluster always requires TAC assistance. For more information, see Removing a Node, on page 10.

- Remove a node from the cluster.

Note	Removing the node must not reduce the number of available nodes below the minimum 3 nodes, as this makes the storage cluster unhealthy. To remove a node in a 3 node cluster always requires TAC assistance.

You can remove a maximum of 2 nodes from an offline cluster. For more information, see Replacing a Node, on page 18.

Identify Node Maintenance Methods

When performing maintenance tasks on nodes, some of these tasks are performed while the storage cluster is offline, others can be performed while the cluster is online and only require that the node is in HX maintenance mode.

- **Online tasks -** require that the storage cluster is healthy before the task begins.

- **Offline tasks -** require that the storage cluster will be shutdown.

   If 2 or more nodes are down, then the storage cluster is automatically offline.

- **TAC assisted tasks -** typically require steps that are performed by the TAC representative.

Note	There are several considerations to keep in mind before replacing a node. For more information, see Replacing a Node, on page 18.

The following tables lists the methods available to perform the associated node maintenance task.

Repair Node Software

ESX and HX Data Platform software is installed on every node in the storage cluster. If it is determined after node failure analysis that either software item needs to be re-installed, see the *Cisco HyperFlex Systems Installation Guide for VMware ESXi*. For steps to upgrade software, see the *Cisco HyperFlex Systems Upgrade Guide*.

Repair Node Hardware

A reparable item on node fails. This includes FRUs and disks. Some node components require TAC assistance. Replacing a node's mother board, for example, requires TAC assistance.

No. Nodes in Cluster	No. Failed Nodes in Cluster	Method	Notes
3	1 or more	TAC assisted only node repair.	Node does not need to be removed to perform repair. Includes replacing disks on node.
4-8	1	Online or Offline node repair.	Node does not need to be removed to perform repair. Includes replacing disks on node.

Remove Node

A non-reparable item on node fails. Disks on the removed node are not reused in the storage cluster.

No. Nodes in Cluster	No. Failed Nodes in Cluster	Method	Notes
4	1	Offline node remove.	A 4 node cluster with 2 nodes down, requires TAC assistance.
5 or more	1	Online or Offline node remove.	
5 or more	2	Offline 2 node remove.	A 5 node cluster with 3 nodes down, requires TAC assistance.

Replace Node and Discard Storage

A non-reparable item on node fails. Disks on the removed node are not reused in the storage cluster.

No. Nodes in Cluster	No. Failed Nodes in Cluster	Method	Notes
3	1	TAC assisted only node replace.	TAC assisted node replacement required to return cluster to minimum 3 nodes. A 3 node cluster with 1 node down, requires TAC assistance.
4	1	Offline replace node. Not reusing the disks.	Use Expand cluster to add new nodes. All other nodes must be up and running. A 4 node cluster with 2 nodes down, requires TAC assistance.
5 or more	1	Online or offline replace node. Not reusing the disks.	Use Expand cluster to add new nodes. All other nodes must be up and running.
5 or more	2	Offline replace 1 or 2 nodes. Not reusing the disks.	Use Expand cluster to add new nodes. All other nodes must be up and running. Replacing up to 2 nodes is supported. Replacing 3 or more nodes requires TAC assistance.

Replace Node and Reuse Storage

A non-reparable item on node fails. Disks on the removed node are reused in the storage cluster.

No. Nodes in Cluster	No. Failed Nodes in Cluster	Method	Notes
3 or more	1 or more	TAC assisted only.	TAC assisted node replacement required to return cluster to minimum 3 nodes.
			Note Reusing disks requires assigning old node UUID to new node. Disks UUIDs to node UUID relationship is fixed and cannot be reassigned. This is a TAC assisted task.

Searching by DNS Address or Host Name

Sometimes for troubleshooting purposes it is useful to be able to search by the DNS server address or DNSserver host name. This is an optional task.

Step 1 Assign DNS search addresses

a) Login to the HX Data Platform Installer virtual machine. Use either `ssh` or the vSphere console interface.

b) Edit `resolv.conf.d` file.

```
# vi /etc/resolvconf/resolv.conf.d/base
```

c) Confirm the change.

```
# resolvconf -u
# cat /etc/resolv.conf
```

d) Confirm the DNS server can be queried from either the IP address or the host name.

```
# nslookup ip_address
# nslookup newhostname
```

Step 2 Assign a DNS host name.

a) Login to the HX Data Platform Installer virtual machine. Use either `ssh` or the vSphere console interface.

b) Open the hosts file for editing.

```
# vi /etc/hosts
```

c) Add the following line and save the file.

ip_address ubuntu *newhostname*

For each host *ip_address*, enter the host *newhostname*.

a) Add the *newhostname* to `hostname`.

```
# hostname newhostname
```

Changing ESXi Host Root Password

You can change the default ESXi password for the following scenarios:

- During creation of a standard and stretch cluster (supports only converged nodes)

- During expansion of a standard cluster (supports both converged or compute node expansion)

- During Edge cluster creation

Note In the above cases, the ESXi root password is secured as soon as installation is complete. In the event a subsequent password change is required, the procedure outlined below may be used after installation to manually change the root password.

As the ESXi comes up with the factory default password, you should change the password for security reasons. To change the default ESXi root password post-installation, do the following.

Note If you have forgotten the ESXi root password, for password recovery please contact Cisco TAC.

Step 1 Log in to the ESXi host service control using SSH.

Step 2 Acquire root privileges.

```
su -
```

Step 3 Enter the current root password.

Step 4 Change the root password.

```
passwd root
```

Step 5 Enter the new password, and press **Enter**. Enter the password a second time for confirmation.

Note If the password entered the second time does not match, you must start over.

Reinstalling Node Software

To re-install software on a node that is a member of an existing storage cluster, contact TAC. This task must be performed with TAC assistance.

Step 1 Reinstall ESX following the directions from TAC.

Ensure the server meets the required hardware and configuration listed in Host ESX Server Setting Requirements. HX configuration settings are applied during the HX Data Platform process.

Step 2   Reinstall HX Data Platform, following the directions from TAC.

The HX Data Platform must always be re-installed after ESX is re-installed.

Changing Node Identification Form in vCenter Cluster from IP to FQDN

This task describes how to change how vCenter identifies the nodes in the cluster, from IP address to Fully Qualified Domain Name (FQDN).

Step 1   Schedule a maintenance window to perform this task.

Step 2   Ensure the storage cluster is healthy.

Check the storage cluster status through either HX Connect, HX Data Platform Plug-in, or from the `stcli cluster info` command on the storage controller VM.

Step 3   Lookup the FQDN for each ESXi host in the storage cluster.

a) From the ESXi host command line.

```
# cat /etc/hosts
```

In this example, the FQDN is `sjs-hx-3-esxi-01.sjs.local`.

```
# Do not remove the following line, or various programs
# that require network functionality will fail.
127.0.0.1        localhost.localdomain localhost
::1              localhost.localdomain localhost
172.16.67.157    sjs-hx-3-esxi-01.sjs.local  sjs-hx-3-esxi-01
```

b) Repeat for each ESXi host in the storage cluster.

Step 4   Verify the FQDNs for each ESXi host are resolvable from vCenter, each other ESXi host, and the controller VMs.

a) From the vCenter command line.

```
# nslookup <fqdn_esx_host1>
# nslookup <fqdn_esx_host2>
# nslookup <fqdn_esx_host3>
...
```

b) Repeat for each ESXi host from an ESXi host.
c) Repeat for each ESXi host from each controller VM.

Step 5   If the FQDN name is not resolvable, then verify the DNS configuration on each ESXi host and each controller VM.

a) Check that the controller VMs have the correct IP address for the DNS server.

From a controller VM command line.

```
# stcli services dns show
10.192.0.31
```

a) Check the ESXi hosts have the same DNS configuration as the controller VMs.

From vCenter, select each ESXi host then **Configuration** > **DNS Servers**.

Step 6   Locate and note the Datacenter Name and the Cluster Name.

From vCenter client or web client, scroll through to see the Datacenter Name and Cluster Name. Write them down. They will be used in a later step.

Step 7   Delete the **cluster** from vCenter.

From vCenter, select **datacenter** > **cluster**. Right-click the **cluster** and select **Delete**.

Note   Do not delete the **datacenter**.

Step 8   Recreate the **cluster** in vCenter.
 a) From vCenter, right-click the **datacenter**. Select **New Cluster**.
 b) Enter the exact same name for the **Cluster Name** as the cluster you deleted. This is the name you wrote down from Step 6.

Step 9   Add ESXi hosts (nodes) to the **cluster** using the FQDN name. Perform these steps for all ESXi hosts.
 a) From vCenter, right-click the **datacenter** > **cluster**. Select **Add Host**.
 b) Select an ESXi host using their FQDN.
 c) Repeat for each ESXi host in the cluster.

Step 10   Reregister the cluster with vCenter.

```
#  stcli cluster reregister
--vcenter-datacenter <datacenter_name>
--vcenter-cluster <hx_cluster_name>
--vcenter-url <FQDN_name>
--vcenter-user <vCenter_username>
--vcenter-password <vCenter_Password>
```

The SSO URL is not required for HX version 1.8.1c or later. See Registering a Storage Cluster with a New vCenter Cluster for additional information on reregistering a cluster.

Replacing Node Components

Selected components on a node can be replaced. Some components can be replaced while the node is up and running. Replacing some components requires that the node be placed into a maintenance mode and shutdown. Refer to the hardware installation guide for your specific server for a complete list of field replaceable units (FRUs). Some components cannot be replaced or can only be replaced with TAC assistance. The following is a general list of components than can be replaced in a node.

Note   When disks are removed, the disk UUIDs continue to be listed, even when not physically present. To reuse disks on another node in the same cluster see TAC for assistance.

- Components that do not require the node be shutdown. These are hot-swappable.

  - HDD data drives. Front bays

    See Managing Disks for the storage cluster tasks and the hardware installation guides for the hardware focused tasks. Both sets of tasks are required to replace this component.

- SSD cache drive. Front bay 1

  See Managing Disks for the storage cluster tasks and the hardware installation guides for the hardware focused tasks. Both sets of tasks are required to replace this component.

- Fan Modules

  See the hardware installation guides to replace this component.

- Power Supplies

  See the hardware installation guides to replace this component.

- Components that do required the node be put into maintenance mode and shutdown.

  For all of the following components, see the hardware installation guides.

  - Housekeeping SSD

    Both the storage cluster tasks and hardware focused tasks are required to replace this component.

  - RTC Battery on motherboard

    Note The motherboard itself is not a replaceable component. You must purchase a battery from your local hardware store and replace it.

  - DIMMS

  - CPUs and Heatsinks

  - Internal SD Card

  - Internal USB Port

  - Modular HBA Riser (HX 220c servers)

  - Modular HBA Card

  - PCIe Riser Assembly

  - PCIe Card

  - Trusted Platform Module

  - mLOM Card

  - RAID Controller

  - Virtual Interface Card (VIC)

  - Graphic Processing Unit (GPU)

Removing a Node

Depending upon the node maintenance task, removing a node can be while the storage cluster is online or offline. Ensure you have completed the preparation steps before removing a node.

Note   It is highly recommended that you work with your account team when removing a converged node in a storage cluster.

Do not reuse the removed converged node or its disks in the original or another cluster.

The affecting context is based on the number of converged nodes. The number of compute nodes does not affect the replacing a node workflow.

Table 1: Removing a Node Workflows

Cluster Size	Nodes Removed	Workflow
3 node cluster	1 or more	Workflow requires TAC assistance.
4 node cluster	1	1. Cluster is healthy. 2. Affected node in Cisco HX Maintenance mode. 3. Shutdown the cluster (take cluster offline). Use the `stcli cluster shutdown` command. 4. Remove the node. Use the `stcli node remove` command. 5. Restart the cluster. Use the `stcli cluster start` command.
4 node cluster	2 or more	Workflow requires TAC assistance.
5 node cluster	1	1. Cluster is healthy. 2. Affected node in Cisco HX Maintenance mode. 3. Cluster remains online. 4. Remove the node. Use the `stcli node remove` command.

Cluster Size	Nodes Removed	Workflow

5 node cluster	2	1. Cluster is healthy.
		
		2. Affected node in Cisco HX Maintenance mode.
		
		3. Shutdown the cluster (take cluster offline).
		
		   Use the `stcli cluster shutdown` command.
		
		4. Remove the nodes.
		
		   Use the `stcli node remove` command.
		
		   Specify both nodes.
		
		5. Restart the cluster.
		
		   Use the `stcli cluster start` command.

5 node cluster	3 or more	Workflow requires TAC assistance.

Preparing to Remove a Node

Before you remove a node from a storage cluster, whether the cluster is online or offline, complete the following steps.

Note For all 3 node clusters, see TAC to assist with preparing, removing, and replacing a node.

Step 1 Ensure the cluster is healthy.

```
# stcli cluster info
```

Example response that indicates the storage cluster is online and heathy:

```
locale: English (United States)
state: online
upgradeState: ok
healthState: healthy
state: online
state: online
```

Step 2 Ensure that SSH is enabled in ESX on all the nodes in the storage cluster.

Step 3 Ensure that the Distributed Resource Scheduler (DRS) is enabled.

DRS migrates only powered-on VMs. If your network has powered-off VMs, you must manually migrate them to a node in the storage cluster that will not be removed.

Note If DRS is not available then manually move the Virtual Machines from the node.

Step 4 Rebalance the storage cluster.

This ensures that all datastores associated with the node will be removed.

The rebalance command is used to realign the distribution of stored data across changes in available storage and to restore storage cluster health. If you add or remove a node in the storage cluster, you can manually initiate a storage cluster rebalance using the `stcli rebalance` command.

Note　　Rebalancing might take some time depending on the disk capacity used on the failed node or disk.

a) Login to a controller VM in the storage cluster.

b) From the controller VM command line, run the command:

```
# stcli rebalance start --force
```

Step 5　Put the node to be removed into Cisco HX Maintenance mode. Choose a method: vSphere GUI or controller VM command line (CLI).

GUI

a) From vSphere web client, select **Home** > **Hosts and Clusters** > **Hosts** > *host*.

b) Right-click each host, scroll down the list, and select **Cisco HX Maintenance Mode** > **Enter HX Maintenance Mode**.

The vSphere Maintenance Mode option is at the top of the host right-click menu. Be sure to scroll to the bottom of the list to select Cisco HX Maintenance Mode.

CLI

a) On the ESX host, log in to a controller VM as a user with root privileges.

b) Identify the node.

```
# stcli node info

stNodes:
    ------------------------------------
    type: node
    id: 689324b2-b30c-c440-a08e-5b37c7e0eefe
    name: 192.168.92.144
    ------------------------------------
    type: node
    id: 9314ac70-77aa-4345-8f35-7854f71a0d0c
    name: 192.168.92.142
    ------------------------------------
    type: node
    id: 9e6ba2e3-4bb6-214c-8723-019fa483a308
    name: 192.168.92.141
    ------------------------------------
    type: node
    id: 575ace48-1513-0b4f-bfe1-e6abd5ff6895
    name: 192.168.92.143
    ------------------------------------
```

Under `stNodes` section the `id` is listed for each node in the cluster.

c) Move the ESX host into Maintenance mode.

```
# stcli node maintenanceMode (--id ID | --ip NAME) --mode enter
```

(see also `stcli node maintenanceMode --help`)

Step 6　Open a command shell and login to the storage controller VM. For example, using `ssh`.

```
# ssh root@controller_vm_ip
```

At the prompt, enter password, `Cisco123`.

What to do next

Proceed to Removing a Node. Choose the Online or Offline method per the condition of your storage cluster and the desired results listed in Managing Nodes.

Removing a Node from an Online Storage Cluster

Use the `stcli node remove` to clean up a deployment or remove a node from a storage cluster.

Removing a node from a storage cluster while the cluster remains online has slightly different requirements from removing a node while the cluster is offline.

Note It is highly recommended that you work with TAC when removing a converged node in a storage cluster.

The affecting context is based on the number of converged nodes. The number of compute nodes does not affect the replacing a node workflow.

Number of nodes in cluster	Method
3 node cluster	See TAC to remove and replace the node.
4 node cluster	Cluster must be offline. See Removing a Node from an Offline Storage Cluster, on page 15.
5 node cluster, removing 2 nodes	Cluster must be offline. See Removing a Node from an Offline Storage Cluster, on page 15.
5 node cluster, removing 1 node from a healthy cluster	Cluster can be online. Continue with the steps listed here.

Note Do not remove the controller VM or other HX Data Platform components before you complete the steps in this task.

Step 1 Complete the steps in Preparing for Maintenance Operations and Preparing to Remove a Node. This includes:
a) Ensure the cluster is healthy.

For 3 node clusters see TAC, as any node failure in a 3 node cluster means the cluster is not healthy.

b) Ensure DRS is enabled or manually move the VMs from the node.
c) Rebalance the storage cluster.
d) Put the node being removed into HX maintenance mode.
e) Login to the controller VM of a node that is not being removed.

Step 2 Rebalance the storage cluster.

    a) Run the rebalance command.

```
# stcli rebalance start -f
```

    b) Wait and confirm that rebalance has completed.

Step 3 Remove the desired node using the `stcli node remove` command.

stcli node remove [-h] {--id-1 ID1 | --ip-1 NAME1} [{--id-2 ID2 | --ip-2 NAME2}] [-f]

Syntax Description	Option	Required or Optional	Description
	--id-1 ID1	One of set required.	A unique ID number for the storage cluster node. The ID is listed in the `stcli cluster info` command under the `stNode` field `id`.
	--ip-1 NAME1	One of set required.	IP address of storage cluster node. The IP is listed in the `stcli cluster info` command under the `stNode` field `name`.
	--id-2 ID2	Optional.	A unique ID number for the storage cluster node. The ID is listed in the `stcli cluster info` command under the `stNode` field `id`.
	--ip-2 NAME2	Optional.	IP address of storage cluster node. The IP is listed in the `stcli cluster info` command under the `stNode` field `name`. The `--ip` option is currently not supported.
	-f, --force	Optional.	Forcibly remove storage cluster nodes.

Example:
```
stNodes for a 5 node cluster:
    --------------------------------------
    type: node
    id: 569c03dc-9af3-c646-8ac5-34b1f7e04b5c
    name: example1
    --------------------------------------
    type: node
    id: 0e0701a2-2452-8242-b6d4-bce8d29f8f17
    name: example2
    --------------------------------------
    type: node
    id: a2b43640-cf94-b042-a091-341358fdd3f4
    name: example3
--------------------------------------
    type: node
    id: c2d43691-fab5-30b2-a092-741368dee3c4
    name: example4
--------------------------------------
    type: node
    id: d2d43691-daf5-50c4-d096-941358fede374
    name: example5
```

The **stcli node remove** command to remove nodes from the 5 node cluster are:

    • To remove 1 node

- **stcli node remove –ip-1 example5** or

- **stcli node remove –id-1 d2d43691-daf5-50c4-d096-941358fede374**

- To remove 2 nodes at the same time:

    - **stcli node remove –ip-1 example5 –ip-2 example4** or

    - **stcli node remove –id-1 d2d43691-daf5-50c4-d096-941358fede374 –id-2 c2d43691-fab5-30b2-a092-741368dee3c4**

    This command unmounts all data stores, removes from the cluster ensemble, resets the EAM for this node, stops all services (stores, cluster management IP), and removes all firewall rules.

    This command does not: remove the node from the vCenter and it does not remove the installed HX Data Platform elements, such as the controller VM.

After the `stcli node remove` command completes successfully, the system rebalances the storage cluster until the storage cluster state is Healthy. Do not perform any failure tests during this time. The storage cluster remains healthy.

Because the node is no longer in the storage cluster, you do not need to exit HX maintenance mode.

Note If you want to reuse a removed node in another storage cluster, contact Technical Assistance Center (TAC). Additional steps are required to prepare the node for another storage cluster.

Step 4 Confirm the node is removed from the storage cluster.

a) Check the storage cluster information.

```
# stcli cluster info
```

b) Check the `ActiveNodes` entry in the response to verify the cluster has one less node.

Step 5 Confirm all the node-associated datastores are removed.

Note If any node-associated datastores are listed, then manually unmount and delete those datastores.

Step 6 Remove the host from the vCenter **Hosts and Cluster** view.

a) Log in to vSphere Web Client Navigator. Navigate to **Host** in the vSphere Inventory.
b) Right-click the host and select **Enter Maintenenace Mode**. Click **Yes**.
c) Right-click the host and select **All vCenter Actions** > **Remove from Inventory**. Click **Yes**.

Step 7 Decommission the host from UCS Manager.

a) Log in to UCS Manager. In the Navigation pane, click **Equipment**.
b) Expand **Equipment** > **Chassis**>*Chassis Number*>**Servers**.
c) Choose the HX server you want to decommission. In the work pane, click the **General** tab.
d) In the **Actions** area, click **Server Maintenance**. In the **Maintenance** dialog box, click **Decommission**. Click **OK**.

Removing a Node from an Offline Storage Cluster

Use the `stcli node remove` to clean up a deployment or remove a node from a storage cluster.

✎

Note It is highly recommended that you work with TAC when removing a converged node in a storage cluster.

The affecting context is based on the number of converged nodes. The number of compute nodes does not affect the replacing a node workflow.

Number of nodes in cluster	Method
3 node cluster	See TAC to remove and replace the node.
4 node cluster	Cluster must be offline.
5 node cluster, removing 2 nodes	Cluster must be offline.
5 node cluster, removing 1 node from a healthy cluster	Cluster can be online. See Removing a Node from an Online Storage Cluster, on page 13.

✎

Note Do not remove the controller VM or other HX Data Platform components before you complete the steps in this task.

You can remove a maximum of 2 nodes from an offline cluster.

Step 1 Complete the steps in Preparing for Maintenance Operations and Preparing to Remove a Node. This includes:

a) Ensure the cluster is healthy.

For 3 node clusters see TAC, as any node failure in a 3 node cluster means the cluster is not healthy.

b) Ensure DRS is enabled or manually move the VMs from the node.
c) Rebalance the storage cluster.
d) Put the node being removed into HX maintenance mode.
e) Login to the controller VM of a node that is not being removed.

Step 2 Prepare to shutdown, then shutdown the storage cluster.

This step is needed only for either of the following conditions:

• The cluster is less than 5 nodes.

• Removing 2 nodes from a 5 node cluster.

a) Gracefully shutdown all resident VMs on all the HX datastores.

Optionally, vMotion the VMs.

b) Gracefully shutdown all VMs on non-HX datastores on HX storage cluster nodes, and unmount.
c) From any controller VM command line, issue the `stcli cluster shutdown` command.

```
# stcli cluster shutdown
```

Step 3 Remove the desired node using the `stcli node remove` command.

For example, you can specify the node to be removed by either IP address or domain name.

```
# stcli node remove --ip-1 10.10.2.4 --ip-2 10.10.2.6
```

or

```
# stcli node remove --name-1 esx.SVHOST144A.complab --name-2 esx.SVHOST144B.complab.lab
```

Note   Enter the second IP address if you are removing a second node from a 5+ node storage cluster.

Response

```
Successfully removed node: EntityRef(type=3, id='', name='10.10.2.4' name='10.10.2.6')
```

This command unmounts all datastores, removes from the cluster ensemble, resets the EAM for this node, stops all services (stores, cluster management IP), and removes all firewall rules.

This command does not:

    • Remove the node from vCenter. The node remains in vCenter.

    • Remove the installed HX Data Platform elements, such as the controller VM.

After the `stcli node remove` command completes successfully, the system rebalances the storage cluster until the storage cluster state is Healthy. Do not perform any failure tests during this time. The storage cluster health remains Average.

Because the node is no longer in the storage cluster, you do not need to exit HX maintenance mode.

Note   If you want to reuse a removed node in another storage cluster, contact Technical Assistance Center (TAC). Additional steps are required to prepare the node for another storage cluster.

Step 4   Confirm the node is removed from the storage cluster.

    a) Check the storage cluster information.

```
# stcli cluster info
```

    b) Check the `ActiveNodes` entry in the response to verify the cluster has one less node.

Step 5   Confirm all the node-associated datastores are removed.

Note   If any node-associated datastores are listed, then manually unmount and delete those datastores.

Step 6   Restart the cluster.

```
# stcli cluster start
```

Removing a Compute Node

Step 1   Migrate all the VMs from a compute node that needs to be removed.

Step 2   Unmount the datastore from the compute node.

Step 3   Check if the cluster is in the healthy state, by running the following command:

```
stcli cluster info --summary
```

Step 4   Put ESXi host in the HX Maintenance mode.

Step 5 Remove the compute node using the `stcli node remove` command, from CMIP (use the Cisco HX connect IP address as it is the cluster IP address).

```
stcli node remove --ip-1
```

Where, IP is the IP address of the node to be removed.

Step 6 Remove any DVS from the EXSi host in vCenter, if there is a DVS.

Step 7 Remove the ESXi host from vCenter.

Step 8 Check if the cluster is in the healthy state, by running the following command:

```
stcli cluster info --summary
```

Step 9 Clear stale entries in the compute node by logging out of Cisco HX Connect and then logging into Cisco HX Connect.

Step 10 Disable and re-enable the High Availability (HA) and Distributed Resource Scheduler (DRS) services to reconfigure the services after node removal.

Deleting the Removed Node Data from a Disk and a Storage Controller VM

After removing a node from a storage cluster, you have to delete the removed node details from the disk and storage controller VM using the following procedure.

⚠

Warning Be aware that once the data is deleted you cannot recover it.

Step 1 Destroy the cluster by running the following command:

```
run destroycluster -sxy
```

Step 2 Remove the `stvboot.cfg` configuration file from the `/etc/` folder.

Step 3 Reboot the controller VM.

Note The reboot process takes few minutes.

Step 4 After rebooting the controller VM, run the following command:

```
# for d in $(/bin/lsblk -dpn -e 1,2,7,11 | awk '{ print $1 }');do  grep -qE "$d[0-9]"
/proc/mounts && continue; dd if=/dev/zero of=$d bs=1M oflag=direct & done;
```

The data deletion action takes few fours. On completion of the drive data deletion, you will get the message: `No space left on device`. Ignore this message.

Replacing a Node

Replacing a node uses Expand Cluster to add the replacement node after removing the failed node. Replacing a node can be performed while the HX storage cluster is online or offline, provided the requirements are met. Replacing a converged node in a storage cluster typically requires TAC assistance.

Note It is highly recommended that you work with TAC when replacing a node in a storage cluster.

Conditions that require TAC assistance to replace a converged node.

- **3 node cluster**

  All 3 node clusters require TAC assistance to replace a node. Replace the node during cluster maintenance.

- **4 node cluster**

  - Storage cluster is unhealthy.

  - Storage cluster will become unhealthy if a node is removed.

  - 2 or more nodes have failed.

  - Disks on the replaced node will be reused.

    When a node is added to a storage cluster, the HX Data Platform associates each disk UUID with the node UUID. This is a fixed relationship for the life of the storage cluster. Disks cannot be reassigned to nodes with different UUIDs. TAC will work with you to assign the old node's UUID to the new node to ensure the disk UUID to node UUID association.

  - Storage cluster to remain online while replacing a node.

- **5 node cluster**

  - Storage cluster is unhealthy.

  - Storage cluster will become unhealthy if a node is removed.

  - 3 or more nodes have failed.

  - Disks on the replaced node will be reused.

    When a node is added to a storage cluster, the HX Data Platform associates each disk UUID with the node UUID. This is a fixed relationship for the life of the storage cluster. Disks cannot be reassigned to nodes with different UUIDs. TAC will work with you to assign the old node's UUID to the new node to ensure the disk UUID to node UUID association.

  - Storage cluster to remain online while replacing 2 nodes.

  - Storage cluster to remain online and the cluster was initially 3 or 4 nodes.

    If your storage cluster was initially configured with either 3 or 4 nodes, adding nodes to make a total of 5 keeps your cluster as a 3+2 or 4+1 cluster. To keep the cluster online while replacing a node requires TAC assistance.

Workflows for Replacing a Node

The affecting context is based on the number of converged nodes. The number of compute nodes does not affect the replacing a node workflow.

Cluster Size	Nodes Replaced	Workflow
3 node cluster	1 or more	Workflow requires TAC assistance.

Cluster Size	Nodes Replaced	Workflow
4 node cluster	1	1. Cluster is healthy. 2. Affected node in Cisco HX Maintenance mode. 3. Shutdown the cluster (take cluster offline). Use the `stcli cluster shutdown` command. 4. Remove the node. Use the `stcli node remove` command. 5. Restart the cluster. Use the `stcli cluster start` command. 6. Wait until cluster is online and healthy. 7. Use **HX Installer** > **Expand Cluster** to add replacement node. **Note** Do not reuse the removed node or its disks in this or another cluster.
4 node cluster	2 or more	Workflow requires TAC assistance.
5 node cluster	1	1. Cluster is healthy. 2. Affected node in Cisco HX Maintenance mode. 3. Cluster remains online. 4. Remove the node. Use the `stcli node remove` command. 5. Restart the cluster. Use the `stcli cluster start` command. 6. Wait until cluster is online and healthy. 7. Use **HX Installer** > **Expand Cluster** to add replacement node. **Note** Do not reuse the removed node or its disks in this or another cluster.

Cluster Size	Nodes Replaced	Workflow
5 node cluster	2	1. Cluster is healthy. 2. Affected node in Cisco HX Maintenance mode. 3. Shutdown the cluster (take cluster offline). Use the `stcli cluster shutdown` command. 4. Remove the nodes. Use the `stcli node remove` command. Specify both nodes. 5. Restart the cluster. Use the `stcli cluster start` command. 6. Wait until cluster is online and healthy. 7. Use **HX Installer** > **Expand Cluster** to add replacement node **Note** Do not reuse the removed node or its disks in this or another cluster.
5 node cluster	3 or more	Workflow requires TAC assistance.

Replacing a node and discarding the failed node's disks.

Step 1   Remove the old node. Follow the steps in the appropriate topic:

     • *Removing a Node from an Online Storage Cluster*

     Use this method only if the HX cluster was initially configured with at least 5 nodes and currently still has at least 5 nodes.

     • *Removing a Node from an Offline Storage Cluster*

     Use this method for all other non-TAC assisted node removal.

Note   Even though you remove a node and its associated disks, the HX Data Platform remembers the disk UUIDs. When logs and reports are generated, messages indicate that the disks exist but cannot be found. Ignore these messages.

Step 2   Add the new node using the Expand option in the HX Data Platform Installer. See the *Cisco HyperFlex Systems Getting Started Guide*.

Replacing a Compute Node

If a compute node boot disk or blade is corrupted and the node needs to be replaced, perform the following steps:

1. Remove the compute node from the existing Hyper-V Hyperflex Cluster.

2. Reinstall OS and re-add the compute node into the cluster.

Note Compute nodes are supported in Hyperflex release 3.5.2 and later releases.

This section provides the procedure for replacing a compute node that needs to be replaced due to faulty boot disk or blade.

Step 1 Use Hyper-V failover cluster manager and remove the bad compute node from the failover cluster manager.

Step 2 Clean up the computer object of the compute node from the Active Directory.

Note There is no need to clean up DNS entry of the compute node.

Step 3 Navigate to any controller VM and run the `remcomputenode.py` script to clean up the stale entries associated with the compute node.

The remove compute node Python script can be executed by providing either the UUID or host name of the compute node as an argument.

The following sample shows how to run the script with UUID of the compute node:

```
python remcomputenode.py -u C2581942-55D2-8021-B1B1-A117F396D671
```

The following sample shows how to run the script with host name of the compute node:

```
python remcomputenode.py -n node-hv1.cloud.local
```

Note Ensure that the following .egg files are available in the controller VM:

- /usr/share/thrift-0.9.1.a-py2.7-linux-x86_64.egg

- /opt/springpath/storfs-mgmt-cli/stCli-1.0-py2.7.egg

Step 4 Replace the faulty MB, compute blade, or boot disk.

Step 5 Run the compute node expansion workflow from the Installer VM.

   a) Install Windows 2016.
   b) On the **HX Data Platform Installer** page, select the **I know what I'm doing...** check box.
   c) Select the expansion workflow and complete the procedure.