Managing Disks

Managing Disks in the Cluster

Disks, SSDs or HDDs, might fail. If this occurs, you need to remove the failed disk and replace it. Follow the server hardware instructions for removing and replacing the disks in the host. The HX Data Platform identifies the SSD or HDD and incorporates it into the storage cluster.

To increase the datastore capacity of a storage cluster add the same size and type SSDs or HDDs to each converged node in the storage cluster. For hybrid servers, add hard disk drives (HDDs). For all flash servers, add SSDs.

Note When performing a hot-plug pull and replace on multiple drives from different vendors or of different types, pause for a few moments (30 seconds) between each action. Pull, pause for about 30 seconds and replace a drive, pause for 30 seconds. Then, pull, pause for 30 seconds and replace the next drive.

Sometimes, when a disk is removed it continues to be listed in cluster summary information. To refresh this, restart the HX cluster.

Note Removing a functional drive from one HX cluster and installing it into another HX cluster is not supported.

Disk Requirements

The disk requirements vary between converged nodes and compute-only nodes. To increase the available CPU and memory capacity, you can expand the existing cluster with compute-only nodes as needed. These compute-only nodes provide no increase to storage performance or storage capacity.

Alternatively, adding converged nodes increase storage performance and storage capacity alongside CPU and memory resources.

Servers with only Solid-State Disks (SSDs) are All-Flash servers. Servers with both SSDs and Hard Disk Drives (HDDs) are hybrid servers.

The following applies to all the disks in a HyperFlex cluster:

- All the disks in the storage cluster must have the same amount of storage capacity. All the nodes in the storage cluster must have the same number of disks.
- All **SSDs** must support TRIM and have TRIM enabled.
- All **HDDs** can be either SATA or SAS type. All SAS disks in the storage cluster must be in a pass-through mode.
- Disk partitions must be removed from SSDs and HDDs. Disks with partitions are ignored and not added to your HX storage cluster.
- Moving operational disks between servers within same cluster or moving them into expansion nodes within the same active cluster is not supported.
- Optionally, you can remove or backup existing data on disks. All existing data on a provided disk is overwritten.

Note New factory servers are shipped with appropriate disk partition settings. Do not remove disk partitions from new factory servers.

- Only the disks ordered directly from Cisco are supported.
- On servers with Self Encrypting Drives (SED), both the cache and persistent storage (capacity) drives must be SED capable. These servers support Data at Rest Encryption (DARE).
- In the event you see an error about unsupported drives or catalog upgrade, see the Catalog Update.
- To prevent data loss, ensure the data on the disk is not the last primary copy of the data.

  If needed, add disks to the servers on the cluster. Initiate or wait until a rebalance completes.
- To prevent data loss, ensure the data on the disk is not the last primary copy of the data. If needed, add disks to the servers on the cluster. Initiate or wait until a rebalance completes. After a successful rebalance the Cluster Flag Resiliency Status shows as Healthy.

In addition to the disks listed in the table below, all M5/M6 converged nodes have M.2 SATA SSD with ESXi installed.

Note Do not mix storage disks type or storage size on a server or across the storage cluster. Mixing storage disk types is not supported.

- When replacing cache or persistent disks, always use the same type and size as the original disk.
- Do not mix any of the persistent drives. Use all HDD or SSD and the same size drives in a server.
- Do not mix hybrid and All-Flash cache drive types. Use the hybrid cache device on hybrid servers and All-Flash cache devices on All-Flash servers.
- Do not mix encrypted and non-encrypted drive types. Use SED hybrid or SED All-Flash drives. On SED servers, both the cache and persistent drives must be SED type.
- All nodes must use same size and quantity of SSDs. Do not mix SSD types.

Please refer to the corresponding server model spec sheet for details of drives capacities and number of drives supported on the different servers.

For information on compatible PIDs when performing an expansion of existing cluster, please refer to the Cisco HyperFlex Drive Compatibility document.

Compute-Only Nodes

The following table lists the supported compute-only node configurations for compute-only functions. Storage on compute-only nodes is not included in the cache or capacity of storage clusters.

Note When adding compute nodes to your HyperFlex cluster, the compute-only service profile template automatically configures it for booting from an SD card. If you are using another form of boot media, update the local disk configuration policy. See the *Cisco UCS Manager Server Management Guide* for server-related policies.

Supported Compute-Only Node Servers	Supported Methods for Booting ESXi
• Cisco B200 M5/M6 • C240 M5/M6 • C220 M5/M6 • C480 M5 • B480 M5	Choose any method. **Important** Ensure that only one form of boot media is exposed to the server for ESXi installation. Post install, you may add in additional local or remote disks. USB boot is not supported for HX Compute-only nodes. • SD Cards in a mirrored configuration with ESXi installed. • Local drive HDD or SSD. • SAN boot. • M.2 SATA SSD Drive. **Note** HW RAID M.2 (UCS-M2-HWRAID and HX-M2-HWRAID) is a supported boot configuration starting with HX Data Platform version 4.5(1a) and later.

Replacing Self Encrypted Drives (SEDs)

Cisco HyperFlex Systems offers Data-At-Rest protection through Self-Encrypting Drives (SEDs) and Enterprise Key Management Support.

- Servers that are data at rest capable refer to servers with self encrypting drives.
- All servers in an encrypted HX Cluster must be data at rest capable.
- Encryption is configured on a HX Cluster , after the cluster is created, using HX Connect .
- Servers with self encrypting drives can be either solid state drive (SSD) or hybrid.

Important To ensure the encrypted data remains secure, the data on the drive must be **securely erased** prior to removing the SED.

Before you begin

Determine if the encryption is applied to the HX Cluster .

- **Encryption not configured—**No encryption related prerequisite steps are required to remove or replace the SED. See Replacing SSDs, on page 6 or Replacing or Adding Hard Disk Drives, on page 11 and the hardware guide for your server.
- **Encryption is configured—**Ensure the following:
  - If you are replacing the SED, obtain a Return to Manufacturer Authorization (RMA). Contact TAC.

- If you are using a local key for encryption, locate the key. You will be prompted to provide it.
- Complete the steps below before removing any SED.

Step 1 Ensure the HX Cluster is healthy.

Step 2 Log into HX Cluster .

Step 3 Select **System Information** > **Disks** page.

Step 4 Identify and verify the disk to remove.

a. Use the Turn On Locator LED button.

b. Physically view the disks on the server.

c. Use the Turn Off Locator LED button.

Step 5 Select the corresponding **Slot** row for the disk to be removed.

Step 6 Click **Secure erase**. This button is available only after a disk is selected.

Step 7 If you are using a local encryption key, enter the **Encryption Key** in the field and click **Secure erase**.

If you are using a remote encryption server, no action is needed.

Step 8 Confirm deleting the data on this disk, click **Yes, erase this disk**.

Warning **This deletes all your data from the disk.**

Step 9 Wait until the **Status** for the selected **Disk Slot** changes to **Ok To Remove**, then physically remove the disk as directed.

What to do next

Note Do not reuse a removed drive in a different server in this, or any other, HX Cluster . If you need to reuse the removed drive, contact TAC.

1. After securely erasing the data on the SED, proceed to the disk replacing tasks appropriate to the disk type: SSD or hybrid.

   Check the **Type** column for the disk type.

   - **Solid State** (SSDs)—See Replacing SSDs, on page 6 and the hardware guide for your server.
   - **Rotational** (hybrid drives)—See Replacing or Adding Hard Disk Drives, on page 11 and the hardware guide for your server.

2. Check the status of removed and replaced SEDs.

   When the SED is removed:

   - **Status**—Remains **Ok To Remove**.
   - **Encryption**—Changes from **Enabled** to **Unknown**.

When the SED is replaced, the new SED is automatically consumed by the HX Cluster . If encryption is not applied, the disk is listed the same as any other consumable disk. If encryption is applied, the security key is applied to the new disk.

- **Status—**Transitions from **Ignored** > **Claimed** > **Available**.
- **Encryption—**Transitions from **Disabled** > **Enabled** after the encryption key is applied.

Replacing SSDs

The procedures for replacing an SSD vary depending upon the type of SSD. Identify the failed SSD and perform the associated steps.

Note Mixing storage disks type or size on a server or across the storage cluster is not supported.

- Use all HDD, or all 3.8 TB SSD, or all 960 GB SSD
- Use the hybrid cache device on hybrid servers and all flash cache devices on all flash servers.
- When replacing cache or persistent disks, always use the same type and size as the original disk.

Step 1 Identify the failed SSD.

- For cache or persistent SSDs, perform a disk beacon check. See Setting a Beacon.

  Only cache and persistent SSDs respond to the beacon request. NVMe cache SSDs and housekeeping SSDs do not respond to beacon requests.
- For cache NVMe SSDs, perform a physical check. These drives are in Drive Bay 1 of the HX servers.
- For housekeeping SSDs on HXAF240c or HX240c servers, perform a physical check at the back of the server.
- For housekeeping SSDs on HXAF220c or HX220c servers, perform a physical check at Drive Bay 2 of the server.

Step 2 **If the failed SSD is a housekeeping SSD**, proceed based on the type of server.

- For HXAF240c M4 or HX240c M4 servers, contact Technical Assistance Center (TAC).
- For all other HX servers, see Replacing Housekeeping SSDs for Cisco HX Release 5.0(2a) and Earlier, on page 11.

Step 3 **If a failed SSD is a cache or persistent SSD**, proceed based on the type of disk.

- For NVMe SSDs, see Replacing NVMe SSDs, on page 7.
- For all other SSDs, follow the instructions for removing and replacing a failed SSD in the host, per the server hardware guide.

After the cache or persistent drive is replaced, the HX Data Platform identifies the SDD and updates the storage cluster.

When disks are added to a node, the disks are immediately available for HX consumption.

Step 4 To enable the Cisco UCS Manager to include new disks in the **UCS Manager** > **Equipment** > **Server** > **Inventory** > **Storage** tab, re-acknowledge the server node. This applies to cache and persistent disks.

Note Re-acknowledging a server is disruptive. Place the server into HXDP Maintenance Mode before doing so.

Step 5 If you replaced an SSD, and see a message *Disk successfully scheduled for repair*, it means that the disk is present, but is still not functioning properly. Check that the disk has been added correctly per the server hardware guide procedures.

Replacing NVMe SSDs

The procedures for replacing an SSD vary depending upon the type of SSD. This topic describes the steps for replacing NVMe cache SSDs.

Note Mixing storage disk types or size on a server or across the storage cluster is not supported.

When replacing NVMe disks, always use the same type and size as the original disk.

Before you begin

Ensure the following conditions are met when using NVMe SSDs in HX Cluster servers.

- NVMe SSDs are supported in HX240 and HX220 All-Flash and All-NVMe servers.
- Hot Swap NVME Drives in M5 and M6 Servers is supported in HX Release 4.5(1a) and later.
- Replacing NVMe SSDs with an HGST SN200 disk requires HX Data Platform Release 2.5(1a) or later.
- For All-Flash nodes, NVMe SSDs are only allowed in slot 1 of the server. Other server slots do not detect NVMe SSDs.
- For All-Flash nodes, NVMe SSDs are only used for cache.

Note You can not use NVMe SSDs as the capacity or housekeeping drive(s) in All-Flash nodes.

- For M5 servers: If you are replacing an NVMe cache drive with a non-NVMe drive (or vice versa, if you are replacing a non-NVMe cache drive with an NVMe drive), you must replace the cable with a different SAS cable (for example, UCSC-RNVME-240M5 = HXAF240c M5 Rear NVMe cable (1) or UCSC-RSAS-C240M5 = C240 Rear UCSC-RAID-M5 SAS cbl(1)). This is required to ensure that the drive is discovered properly.

Note For M6 servers: you cannot replace an NVMe cache drive with a non-NVMe cache drive, due to the placement of the slots which are in the front.

Step 1 Confirm the failed disk is an NVMe cache SSD.

Perform a physical check. NVMe cache SSDs and housekeeping SSDs do not respond to beacon requests.

If the failed SSD is not an NVMe SSD, see the Replacing SSD section of this guide.

Step 2 Put ESXi host into HXDP Maintenance Mode.

a) Log into HX Connect.
b) Select **System Information** > **Nodes** > ***node*** > **Enter HXDP Maintenance Mode**.

Step 3 Follow the instructions for removing and replacing a failed SSD in the host, per the server hardware guide.

Note When you remove an HGST NVMe disk, the controller VM will fail until you reinsert a disk of the same type into the same slot or reboot the host.

After the cache or persistent drive is replaced, the HX Data Platform identifies the SDD and updates the storage cluster.

When disks are added to a node, the disks are immediately available for HX consumption.

Step 4 Reboot the ESXi host. This enables ESXi to discover the NVMe SSD.

Step 5 Exit ESXi host from HXDP Maintenance Mode.

Step 6 To enable the Cisco UCS Manager to include new disks in the **UCS Manager > Equipment > Server > Inventory > Storage** tab, re-acknowledge the server node. This applies to cache and persistent disks.

Note Re-acknowledging a server is disruptive. Place the server into HXDP Maintenance Mode before doing so.

Step 7 If you replaced an SSD, and see a message *Disk successfully scheduled for repair*, it means that the disk is present, but is still not functioning properly. Check that the disk has been added correctly per the server hardware guide procedures.

Hot Swap NVME Drives in M5 and M6 Servers

Beginning with Cisco HyperFlex Release 4.5(1a), M5 and M6 servers with the VMD Enable bios option being active may hotswap NVME drives in new installs, and upgrades which have a combined HX+ UCS upgrade performed. VMD enabled is set in the bios which allows NVME drives to be hot swappable without requiring HXDP Maintenance Mode or a reboot of ESXi.

To verify that VMD is enabled:

Before you begin

Ensure the Replacing NVMe SSDs, on page 7 conditions are met when using NVMe SSDs in HX Cluster servers.

Step 1 In the Navigation pane, click **Servers**

Step 2 Go to **Policies** > **Root** > **BIOS Polices**

Step 3 Expand **root** > **Sub-Organizations** > ***your org***

Step 4 Select **hx-bios-af** (for M5) or **hx-bios-m6-af** (for m6)

Step 5 Click on **info**

Step 6 The BIOS Policy window apppears. Select **Advanced tab** > **LOM** > **PCIe Slots**

Step 7 Scroll down to see the **VMD Enable** settings and verify it is set to **Enabled**.

Replacing Housekeeping SSDs for Cisco HX Release 5.0(2b) and Later

Note This procedure applies to HXAF220c M4, HX220c M4,+ HXAF220c M5, HX220c M5, HXAF240c M5, HX240c M5, HXAF220c M6, HX220c M6, HXAF240c M6, and HX240c M6 servers only. To replace the housekeeping SSD on an HXAF240c M4 or HX240c M4 servers, contact Cisco TAC.

Identify the failed housekeeping SSD and perform the associated steps.

Step 1 Identify the failed housekeeping SSD.

Physically check the SSD drives, as housekeeping drives are not listed through a beacon check.

Step 2 Remove the SSD and replace with a new SSD of the same supported kind and size. Follow the steps in the server hardware guide.

The server hardware guide describes the physical steps required to replace the SSD.

Note Before performing the hardware steps, enter the node into HXDP Maintenance Mode. After performing the hardware steps, exit the node from HXDP Maintenance Mode.

Step 3 Using `SSH`, log into the storage controller VM of the cip node (any other working node) and run the following command to create **bootdev** partitions.

```
priv createBootdevPartitions --target 10.20.24.69
```

Sample response

```
hxshell:~$ priv createBootdevPartitions --target 10.20.24.69
Enter the root password:
create Bootdev Partitions initiated on 10.20.24.69
```

Note The target should be the storage controller VM IP of the affected node.

This command reboots the affected node.

Step 4 Wait for the storage controller VM to automatically reboot.

Step 5 When the storage controller VM completes its reboot, verify that partitions are created on the newly added SSD. Run the command.

```
# df -ah
```

Sample response

```
...........
/dev/sdb1 63G 324M 60G 1%
/var/stv /dev/sdb2 24G 173M 23G 1% /var/zookeeper
```

Step 6 Identify the HX Data Platform installer package version installed on the existing storage cluster.

```
# hxcli cluster version
```

The same version must be installed on all the storage cluster nodes. Run this command on the controller VM of any node in the storage cluster, but not the node with the new SSD.

Step 7 SFTP the HX Data Platform installer packages into the storage controller VM of the affected node using the admin account for **user name**/**password** and a file transfer application, such as **winscp**. This should upload the package to the `/tmp` directory. Untar the package after copying it to the `/tmp` directory.

```
# tar -zxvf storfs-packages-<version>.tgz
```

Step 8 Using SSH, log into the storage controller VM of the cip node (any other working) node and run the following command:

```
priv housekeeping-preinstall --target 10.20.24.69
```

Sample response:

```
hxshell:~$ priv housekeeping-preinstall --target 10.20.24.69
      Enter root password :
      Copied secure files
```

Note This step copies the secure files of `/etc/springpath/secure/*` folder from another working controller machine into the affected node.

Step 9 Run the following command on the storage controller VM of the cip node (any other working node) to install the HX Data Platform installer packages.

```
Priv housekeeping-inst-packages –target 10.20.24.69
```

Sample response:

```
hxshell:~$ priv housekeeping-inst-packages –target 10.20.24.69
         Enter root password :
  Installed packages successfully
```

The package installation takes about 10 to 15 minutes.

Step 10 Enter the following command on the storage controller VM of the cip node (any other working node) to perform the post-install tasks..

```
priv housekeeping-postinstall --target 10.20.24.69
```

Sample response:

```
hxshell:~$ priv housekeeping-postinstall --target 10.20.24.69
        Enter root password :
        Successfully done post install tasks
       Successfully installed SE core package on 10.20.24.69 (optional only when Software Encryption
 is enabled on the cluster
```

For post-installation tasks, take the following steps:

a) Install the SE core package (optional if SE is enabled on the cluster).
b) Reboot the CVM.

   This step reboots the affected node. Wait for the storage controller VM to automatically reboot.

Step 11 To confirm that the **cip-monitor** and **stofs** are in running status, run the `priv service cip-monitor status` and the `priv service storfs status` commands.

Example:

```
hxshell:~$ priv service cip-monitor status
       cip-monitor start/running, process 18251

       hxshell:~$ priv service storfs status
       storfs start/running, process 22057
```

Replacing Housekeeping SSDs for Cisco HX Release 5.0(2a) and Earlier

Note This procedure applies to HXAF220c M4, HX220c M4, HXAF220c M5, HX220c M5, HXAF240c M5, HX240c M5, servers only. To replace the housekeeping SSD on an HXAF240c M4 or HX240c M4 servers, contact Cisco TAC.

Identify the failed housekeeping SSD and perform the associated steps.

Step 1 Identify the failed housekeeping SSD.

Physically check the SSD drives, as housekeeping drives are not listed through a beacon check.

Step 2 Remove the SSD and replace with a new SSD of the same supported kind and size. Follow the steps in the server hardware guide.

The server hardware guide describes the physical steps required to replace the SSD.

Note Before performing the hardware steps, enter the node into HXDP Maintenance Mode. After performing the hardware steps, exit the node from HXDP Maintenance Mode.

Step 3 Contact TAC to complete the replacement process.

Replacing or Adding Hard Disk Drives

Note Mixing storage disks type or size on a server or across the storage cluster is not supported.

- Use all HDD, or all 3.8 TB SSD, or all 960 GB SSD
- Use the hybrid cache device on hybrid servers and all flash cache devices on all flash servers.
- When replacing cache or persistent disks, always use the same type and size as the original disk.

Step 1 Refer to the hardware guide for your server and follow the directions for adding or replacing disks.

Step 2 Add HDDs of the same size to each node in the storage cluster.

Step 3 Add the HDDs to each node within a reasonable amount of time.

The storage starts being consumed by storage cluster immediately.

The vCenter Event log displays messages reflecting the changes to the nodes.

Note When disks are added to a node, the disks are immediately available for HX consumption although they will not be seen in the UCSM server node inventory. This includes cache and persistent disks. To include the disks in the**Equipment** > **Manager** > **UCS** > **Equipment** > **Server** > **Inventory** > **Storage** tab, re-acknowledge the server node.

Note Re-acknowledging a server is disruptive. Place the server into HXDP Maintenance Mode before doing so.