# Maintain

## StorageGRID 11.5

NetApp
April 11, 2024

This PDF was generated from https://docs.netapp.com/us-en/storagegrid-115/expand/guidelines-for-adding-object-capacity.html on April 11, 2024. Always check docs.netapp.com for the latest.

# Table of Contents

Maintain . . . . . 1
Expand your grid . . . . . 1
Maintain & recover . . . . . 54

# Maintain

## Expand your grid

Learn how to expand a StorageGRID system without interrupting system operations.

- Planning a StorageGRID expansion
- Preparing for an expansion
- Overview of expansion procedure
- Adding storage volumes to Storage Nodes
- Adding grid nodes to an existing site or adding a new site
- Configuring your expanded StorageGRID system
- Contacting technical support

### Planning a StorageGRID expansion

You can expand StorageGRID to increase storage capacity, to add metadata capacity, to add redundancy or new capabilities, or to add a new site. The number, type, and location of the nodes you need to add depends on the reason for the expansion.

- Adding storage capacity
- Adding metadata capacity
- Adding grid nodes to add capabilities to your system
- Adding a new site

#### Adding storage capacity

When existing Storage Nodes become full, you must increase the storage capacity of your StorageGRID system.

To increase storage capacity, you must first understand where data is stored currently and then add capacity in all required locations. For example, if you currently store copies of object data at several sites, you might need to increase the storage capacity of each site.

- Guidelines for adding object capacity
- Adding storage capacity for replicated objects
- Adding storage capacity for erasure-coded objects
- Considerations for rebalancing erasure-coded data

##### Guidelines for adding object capacity

You can expand the object storage capacity of your StorageGRID system by adding storage volumes to existing Storage Nodes or by adding new Storage Nodes to existing sites. You must add storage capacity in a way that meets the requirements of your information lifecycle management (ILM) policy.

### Guidelines for adding storage volumes

Before adding storage volumes to existing Storage Nodes, review the following guidelines and limitations:

- You must examine your current ILM rules to determine where and when to add storage volumes to increase the storage available for replicated or erasure-coded objects. See the instructions for managing objects with information lifecycle management.
- You cannot increase the metadata capacity of your system by adding storage volumes because object metadata is stored only on volume 0.
- Each software-based Storage Node can support a maximum of 16 storage volumes. If you need to add capacity beyond that, you must add new Storage Nodes.
- You can add one or two expansion shelves to each SG6060 appliance. Each expansion shelf adds 16 storage volumes. With both expansion shelves installed, the SG6060 can support a total of 48 storage volumes.
- You cannot add storage volumes to any other storage appliance.
- You cannot increase the size of an existing storage volume.
- You cannot add storage volumes to a Storage Node at the same time you are performing a system upgrade, recovery operation, or another expansion.

After you have decided to add storage volumes and have determined which Storage Nodes you must expand to satisfy your ILM policy, follow the instructions for your type of Storage Node:

- To add expansion shelves to an SG6060 storage appliance, see the instructions for SG6000 appliance installation and maintenance.

  SG6000 storage appliances

- For a software-based node, follow the instructions for adding storage volumes to Storage Nodes.

  Adding storage volumes to Storage Nodes

### Guidelines for adding Storage Nodes

Before adding Storage Nodes to existing sites, review the following guidelines and limitations:

- You must examine your current ILM rules to determine where and when to add Storage Nodes to increase the storage available for replicated or erasure-coded objects.
- You should not add more than 10 Storage Nodes in a single expansion procedure.
- You can add Storage Nodes to more than one site in a single expansion procedure.
- You can add Storage Nodes and other types of nodes in a single expansion procedure.
- Before starting the expansion procedure, you must confirm that all data-repair operations performed as part of a recovery are complete. See the steps for checking data repair jobs in the recovery and maintenance instructions.
- If you need to remove Storage Nodes before or after performing an expansion, you should not decommission more than 10 Storage Nodes in a single Decommission Node procedure.

### Guidelines for the ADC service on Storage Nodes

When configuring the expansion, you must choose whether to include the Administrative Domain Controller (ADC) service on each new Storage Node. The ADC service keeps track of the location and availability of grid

services.

- The StorageGRID system requires a quorum of ADC services to be available at each site and at all times.

  Learn about the ADC quorum in the recovery and maintenance instructions.

- At least three Storage Nodes at each site must include the ADC service.
- Adding the ADC service to every Storage Node is not recommended. Including too many ADC services can cause slowdowns due to the increased amount of communication between nodes.
- A single grid should not have more than 48 Storage Nodes with the ADC service. This is equivalent to 16 sites with three ADC services at each site.
- In general, when you select the **ADC Service** setting for a new node, you should select **Automatic**. Select **Yes** only if the new node will replace another Storage Node that includes the ADC service. Because you cannot decommission a Storage Node if too few ADC services would remain, this ensures that a new ADC service is available before the old service is removed.
- You cannot add the ADC service to a node after it is deployed.

**Related information**

Manage objects with ILM

SG6000 storage appliances

Adding storage volumes to Storage Nodes

Maintain & recover

Performing the expansion

#### Adding storage capacity for replicated objects

If the information lifecycle management (ILM) policy for your deployment includes a rule that creates replicated copies of objects, you must consider how much storage to add and where to add the new storage volumes or Storage Nodes.

For guidance on where to add additional storage, examine the ILM rules that create replicated copies. If ILM rules create two or more object copies, plan to add storage in each location where object copies are made. As a simple example, if you have a two-site grid and an ILM rule that creates one object copy at each site, you must add storage to each site to increase the overall object capacity of the grid.

For performance reasons, you should attempt to keep storage capacity and compute power balanced across sites. So, for this example, you should add the same number of Storage Nodes to each site or additional storage volumes at each site.

If you have a more complex ILM policy that includes rules that place objects in different locations based on criteria such as bucket name, or rules that change object locations over time, your analysis of where storage is required for the expansion will be similar, but more complex.

Charting how quickly overall storage capacity is being consumed can help you understand how much storage to add in the expansion, and when the additional storage space will be required. You can use the Grid Manager to monitor and chart storage capacity as described in the instructions for monitoring and troubleshooting StorageGRID.

When planning the timing of an expansion, remember to consider how long it might take to procure and install additional storage.

**Related information**

Manage objects with ILM

Monitor & troubleshoot

#### Adding storage capacity for erasure-coded objects

If your ILM policy includes a rule that makes erasure-coded copies, you must plan where to add new storage and when to add new storage. The amount of storage you add and the timing of the addition can affect the grid's usable storage capacity.

The first step in planning a storage expansion is to examine the rules in your ILM policy that create erasure-coded objects. Because StorageGRID creates *k+m* fragments for every erasure-coded object and stores each fragment on a different Storage Node, you must ensure that at least *k+m* Storage Nodes have space for new erasure-coded data after the expansion. If the erasure-coding profile provides site-loss protection, you must add storage to each site.

The number of nodes you need to add also depends on how full the existing nodes are when you perform the expansion.

**General recommendation for adding storage capacity for erasure-coded objects**

If you want to avoid detailed calculations, you can add two Storage Nodes per site when existing Storage Nodes reach 70% capacity.

This general recommendation provides reasonable results across a wide range of erasure-coding schemes for both single-site grids and for grids where erasure coding provides site-loss protection.

To better understand the factors that lead to this recommendation or to develop a more precise plan for your site, review the next section. For a custom recommendation optimized for your situation, contact your NetApp account representative.

**Calculating the number of expansion Storage Nodes to add for erasure-coded objects**

To optimize how you expand a deployment that stores erasure-coded objects, you must consider many factors:

- Erasure-coding scheme in use
- Characteristics of the storage pool used for erasure coding, including the number of nodes at each site and the amount of free space on each node
- Whether the grid was previously expanded (because the amount of free space per Storage Node might not be approximately the same on all nodes)
- Exact nature of the ILM policy, such as whether ILM rules make both replicated and erasure-coded objects

The following examples can help you understand the impact of the erasure-coding scheme, the number of nodes in the storage pool, and the amount of free space on each node.

Similar considerations affect the calculations for an ILM policy that stores both replicated and erasure-coded data and the calculations for a grid that has been previously expanded.

The examples in this section represent the best practices for adding storage capacity to a StorageGRID system. If you are unable to add the recommended number of nodes, you might need to run the EC rebalance procedure to allow additional erasure-coded objects to be stored.

Considerations for rebalancing erasure-coded data

**Example 1: Expanding a one-site grid that uses 2+1 erasure coding**

This example shows how to expand a simple grid that includes only three Storage Nodes.

This example uses only three Storage Nodes for simplicity. However, using only three Storage Nodes is not recommended: an actual production grid should use a minimum of *k+m* + 1 Storage Nodes for redundancy, which equals four Storage Nodes (2+1+1) for this example.

Assume the following:

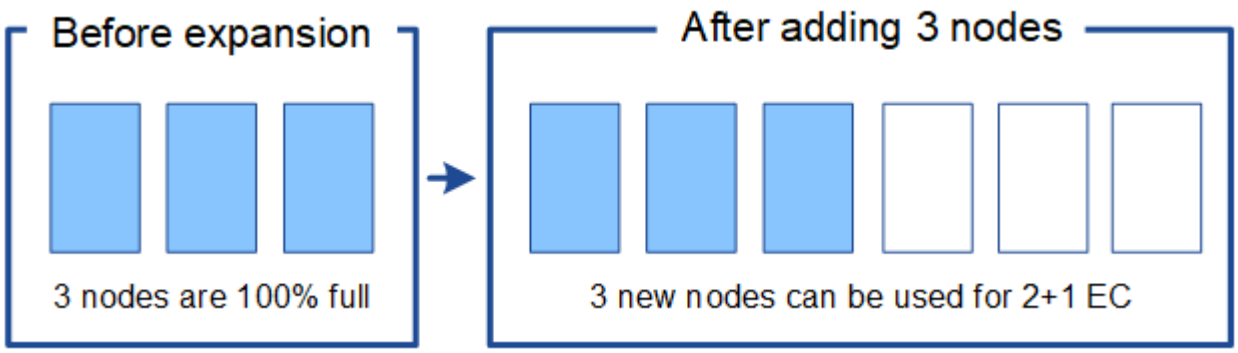
- All data is stored using the 2+1 erasure-coding scheme. With the 2+1 erasure coding scheme, every object is stored as three fragments, and each fragment is saved on a different Storage Node.
- You have one site with three Storage Nodes. Each Storage Node has a total capacity of 100 TB.
- You want to expand by adding new 100 TB Storage Nodes.
- You want to eventually balance erasure-coded data across the old and new nodes.

You have a number of options, based on how full the Storage Nodes are when you perform the expansion.

- **Add three 100 TB Storage Nodes when the existing nodes are 100% full**

  In this example, the existing nodes are 100% full. Because there is no free capacity, you must immediately add three nodes to continue 2+1 erasure coding.

  After the expansion is complete, when objects are erasure-coded, all fragments will be placed on the new nodes.



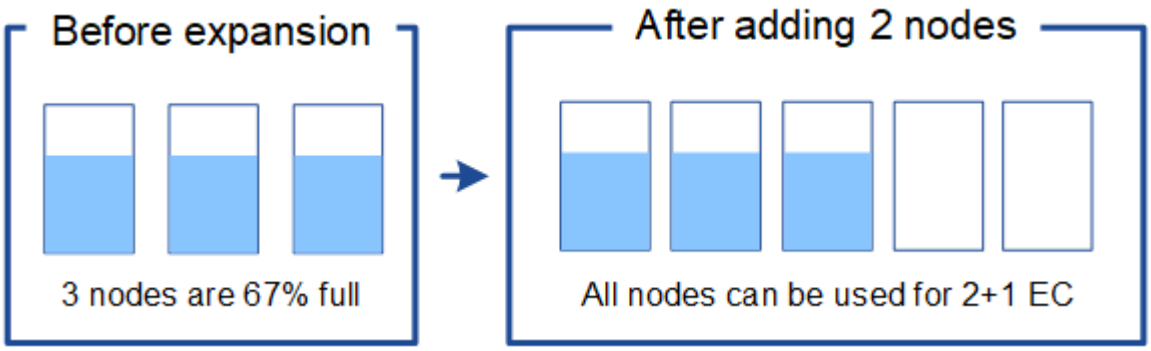

  This expansion adds *k+m* nodes. Adding four nodes is recommended for redundancy. If you add only *k+m* expansion Storage Nodes when existing nodes are 100% full, all new objects must be stored on the expansion nodes. If any of the new nodes become unavailable, even temporarily, StorageGRID cannot meet ILM requirements.

- **Add two 100 TB Storage Nodes, when the existing Storage Nodes are 67% full**

  In this example, the existing nodes are 67% full. Because there are 100 TB of free capacity on the existing

nodes (33 TB per node), you only need to add two nodes if you perform the expansion now.

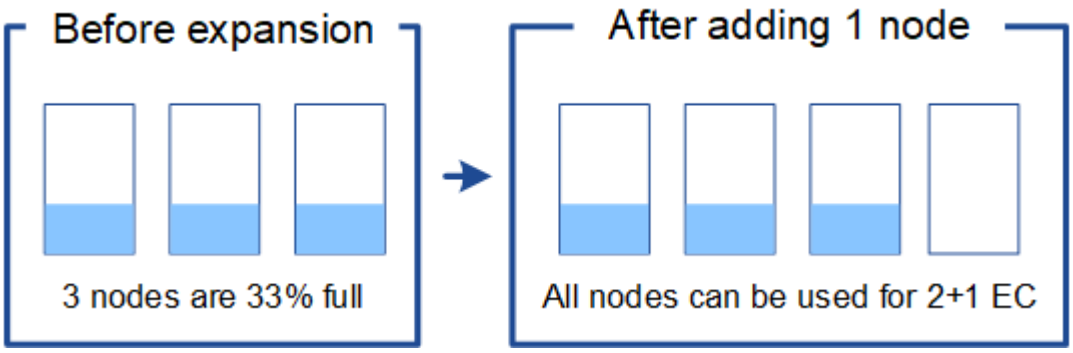
Adding 200 TB of additional capacity will allow you to continue 2+1 erasure coding and to eventually balance erasure-coded data across all nodes.




- **Add one 100 TB Storage Node when the existing Storage Nodes are 33% full**

  In this example, the existing nodes are 33% full. Because there are 200 TB of free capacity on the existing nodes (67 TB per node), you only need to add one node if you perform the expansion now.

  Adding 100 TB of additional capacity will allow you to continue 2+1 erasure coding and to eventually balance erasure-coded data across all nodes.




**Example 2: Expanding a three-site grid that uses 6+3 erasure coding**

This example shows how to develop an expansion plan for a multi-site grid that has an erasure-coding scheme with a larger number of fragments. Despite the differences between these examples, the recommended expansion plan is very similar.

Assume the following:

- All data is stored using the 6+3 erasure coding scheme. With the 6+3 erasure coding scheme, every object is stored as 9 fragments, and each fragment is saved to a different Storage Node.
- You have three sites, and each site has four Storage Nodes (12 nodes in total). Each node has a total capacity of 100 TB.
- You want to expand by adding new 100 TB Storage Nodes.
- You want to eventually balance erasure-coded data across the old and new nodes.

You have a number of options, based on how full the Storage Nodes are when you perform the expansion.

- **Add nine 100 TB Storage Nodes (three per site), when existing nodes are 100% full**

  In this example, the 12 existing nodes are 100% full. Because there is no free capacity, you must immediately add nine nodes (900 TB of additional capacity) to continue 6+3 erasure coding.

  After the expansion is complete, when objects are erasure-coded, all fragments will be placed on the new nodes.

  > This expansion adds *k+m* nodes. Adding 12 nodes (four per site) is recommended for redundancy. If you add only *k+m* expansion Storage Nodes when existing nodes are 100% full, all new objects must be stored on the expansion nodes. If any of the new nodes become unavailable, even temporarily, StorageGRID cannot meet ILM requirements.

- **Add six 100 TB Storage Nodes (two per site), when existing nodes are 75% full**

  In this example, the 12 existing nodes are 75% full. Because there are 300 TB of free capacity (25 TB per node), you only need to add six nodes if you perform the expansion now. You would add two nodes to each of the three sites.

  Adding 600 TB of storage capacity will allow you to continue 6+3 erasure coding and to eventually balance erasure-coded data across all nodes.

- **Add three 100 TB Storage Nodes (one per site), when existing nodes are 50% full**

  In this example, the 12 existing nodes are 50% full. Because there are 600 TB of free capacity (50 TB per node), you only need to add three nodes if you perform the expansion now. You would add one node to each of the three sites.

  Adding 300 TB of storage capacity will allow you to continue 6+3 erasure coding and to eventually balance erasure-coded data across all nodes.

**Related information**

Manage objects with ILM

Monitor & troubleshoot

Considerations for rebalancing erasure-coded data

#### Considerations for rebalancing erasure-coded data

If you are performing an expansion to add Storage Nodes and your ILM policy includes one or more ILM rules to erasure code data, you might need to perform the EC rebalance procedure after the expansion is complete.

For example, if you cannot add the recommended number of Storage Nodes in an expansion, you might need to run the EC rebalance procedure to allow additional erasure-coded objects to be stored.

**What is EC rebalancing?**

EC rebalancing is a StorageGRID procedure that might be required after a Storage Node expansion. The procedure is run as a command-line script from the primary Admin Node. When you run the EC rebalance procedure, StorageGRID redistributes erasure-coded fragments among the existing and the newly expanded Storage Nodes at a site.

When the EC rebalance procedure runs:

- It only moves erasure-coded object data. It does not move replicated object data.
- It redistributes the data within a site. It does not move data between sites.
- It redistributes data among all Storage Nodes at a site. It does not redistribute data within storage volumes.

When the EC rebalance procedure is complete:

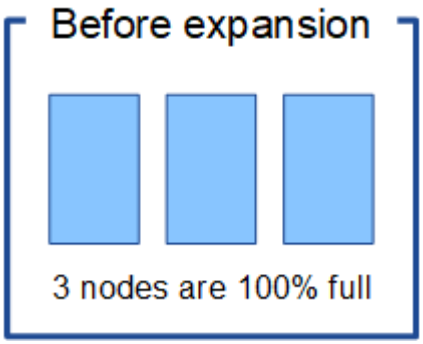
- Erasure-coded data is moved from Storage Nodes with less available space to Storage Nodes with more available space.
- Used (%) values might remain different between Storage Nodes because the EC rebalance procedure does not move replicated object copies.
- The data protection of erasure-coded objects will be unchanged.

When the EC rebalance procedure is running, the performance of ILM operations and S3 and Swift client operations are likely to be impacted. For this reason, you should only perform this procedure in limited cases.

**When not to perform an EC rebalance**

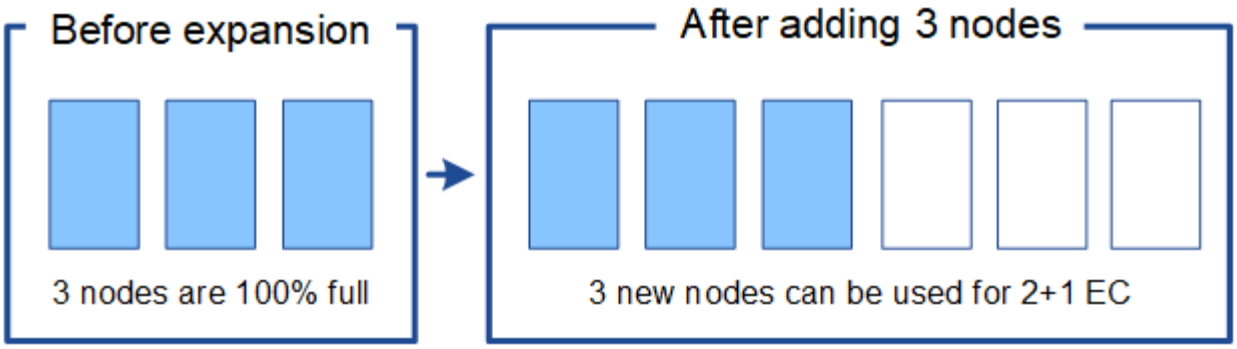
As an example of when you do not need to perform an EC rebalance, consider the following:

- StorageGRID is running at a single site, which contains three Storage Nodes.
- The ILM policy uses a 2+1 erasure-coding rule for all objects larger than 0.2 MB and a 2-copy replication rule for smaller objects.
- All Storage Nodes have become completely full, and the **Low Object Storage** alert has been triggered at the major severity level. The recommended action is to perform an expansion procedure to add Storage Nodes.



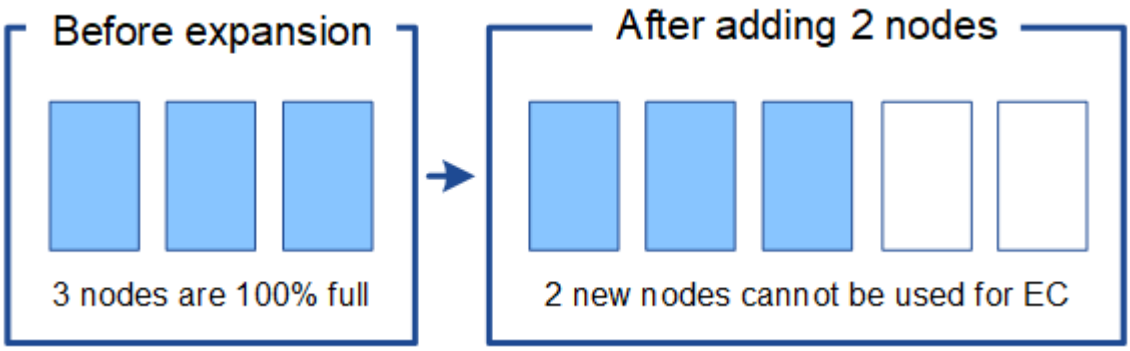

To expand the site in this example, it is recommended that you add three or more new Storage Nodes. StorageGRID requires three Storage Nodes for 2+1 erasure coding so that it can place the two data fragments and the one parity fragment on different nodes.

After you add the three Storage Nodes, the original Storage Nodes remain full, but objects can continue to be ingested into the 2+1 erasure coding scheme on the new nodes. Running the EC rebalance procedure is not recommended for this case: running the procedure will temporarily decrease performance, which might impact client operations.

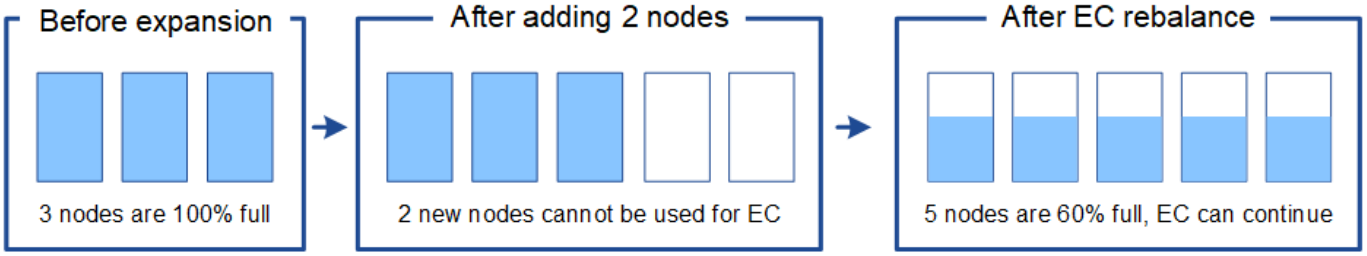



**When to perform an EC rebalance**

As an example of when you should perform the EC rebalance procedure, consider the same example, but assume that you can only add two Storage Nodes. Because 2+1 erasure coding requires at least three Storage Nodes, the new nodes cannot be used for erasure-coded data.




To resolve this issue and make use of the new Storage Nodes, you can run the EC rebalance procedure. When this procedure runs, StorageGRID redistributes erasure-coded data and parity fragments among all Storage Nodes at the site. In this example, when the EC rebalance procedure is complete, all five nodes are now only 60% full, and objects can continue to be ingested into the 2+1 erasure coding scheme on all Storage Nodes.




**Considerations for EC rebalancing**

In general, you should only run the EC rebalance procedure in limited cases. Specifically, you should perform EC rebalancing only if all of the following statements are true:

- You use erasure coding for your object data.
- The **Low Object Storage** alert has been triggered for one or more Storage Nodes at a site, indicating that the nodes are 80% or more full.
- You are unable to add the recommended number of new Storage Nodes for the erasure-coding scheme in use.

- Your S3 and Swift clients can tolerate lower performance for their write and read operations while the EC rebalance procedure is running.

### How the EC rebalance procedure interacts with other maintenance tasks

You cannot perform certain maintenance procedures at the same time you are running the EC rebalance procedure.

| Procedure | Allowed during EC rebalance procedure? |
|---|---|
| Additional EC rebalance procedures | No.<br><br>You can only run one EC rebalance procedure at a time. |
| Decommission procedure<br><br>EC data repair job | No.<br><br>• You are prevented from starting a decommission procedure or an EC data repair while the EC rebalance procedure is running.<br>• You are prevented from starting the EC rebalance procedure while a Storage Node decommission procedure or an EC data repair is running. |
| Expansion procedure | No.<br><br>If you need to add new Storage Nodes in an expansion, you should wait to run the EC rebalance procedure until after you have added all new nodes. If an EC rebalance procedure is in progress when you add new Storage Nodes, data will not be moved to those nodes. |
| Upgrade procedure | No.<br><br>If you need to upgrade StorageGRID software, you should perform the upgrade procedure before or after running the EC rebalance procedure. As required, you can terminate the EC rebalance procedure to perform a software upgrade. |
| Appliance node clone procedure | No.<br><br>If you need to clone an appliance Storage Node, you should wait to run the EC rebalance procedure until after you have added the new node. If an EC rebalance procedure is in progress when you add new Storage Nodes, data will not be moved to those nodes. |
| Hotfix procedure | Yes.<br><br>You can apply a StorageGRID hotfix while the EC rebalance procedure is running. |

| Procedure | Allowed during EC rebalance procedure? |
| --- | --- |
| Other maintenance procedures | No.<br><br>You must terminate the EC rebalance procedure before running other maintenance procedures. |

### How the EC rebalance procedure interacts with ILM

While the EC rebalance procedure is running, avoid making ILM changes that might change the location of existing erasure-coded objects. For example, do not start using an ILM rule that has a different Erasure Coding profile. If you need to make such ILM changes, you should abort the EC rebalance procedure.

#### Related information

Rebalancing erasure-coded data after adding Storage Nodes

## Adding metadata capacity

To ensure that adequate space is available for object metadata, you might need to perform an expansion procedure to add new Storage Nodes at each site.

StorageGRID reserves space for object metadata on volume 0 of each Storage Node. Three copies of all object metadata are maintained at each site, evenly distributed across all Storage Nodes.

You can use the Grid Manager to monitor the metadata capacity of Storage Nodes and to estimate how quickly metadata capacity is being consumed. In addition, the **Low metadata storage** alert is triggered for a Storage Node when the used metadata space reaches certain thresholds. See the instructions for monitoring and troubleshooting StorageGRID for details.

Note that a grid's object metadata capacity might be consumed faster than its object storage capacity, depending on how you use the grid. For example, if you typically ingest large numbers of small objects or add large quantities of user metadata or tags to objects, you might need to add Storage Nodes to increase metadata capacity even though sufficient object storage capacity remains.

#### Guidelines for increasing metadata capacity

Before adding Storage Nodes to increase metadata capacity, review the following guidelines and limitations:

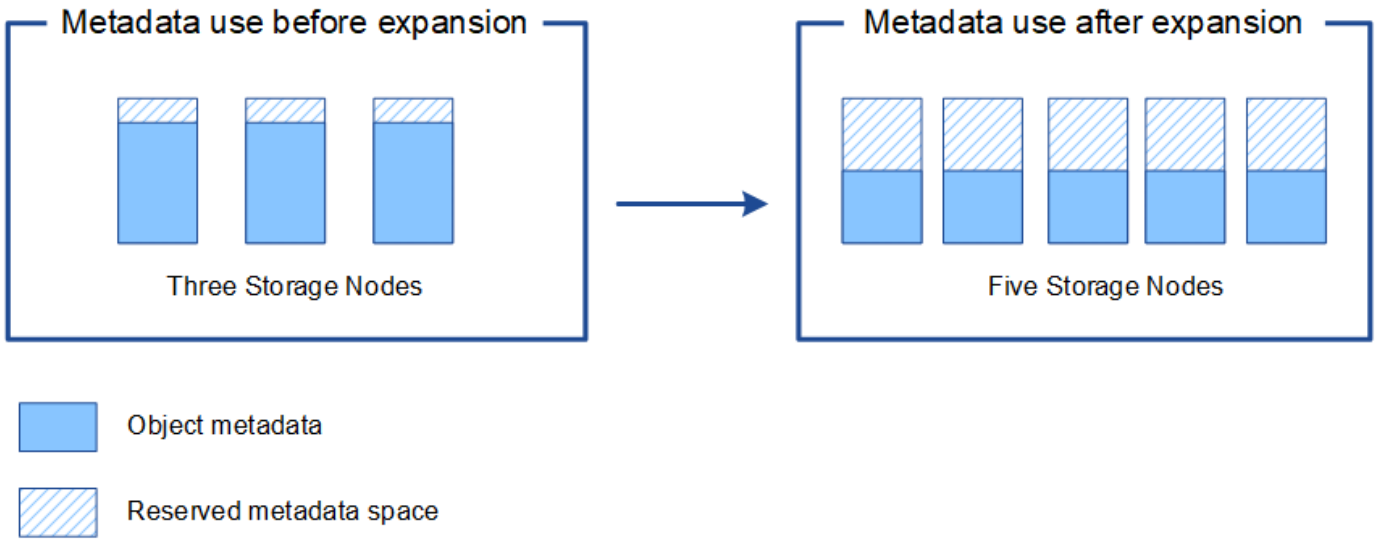
- Assuming sufficient object storage capacity is available, having more space available for object metadata increases the number of objects you can store in your StorageGRID system.
- You can increase a grid's metadata capacity by adding one or more Storage Nodes to each site.
- The actual space reserved for object metadata on any given Storage Node depends on the Metadata Reserved Space storage option (system-wide setting), the amount of RAM allocated to the node, and the size of the node's volume 0. See the instructions for administering StorageGRID for more information.
- You cannot increase metadata capacity by adding storage volumes to existing Storage Nodes, because metadata is stored only on volume 0.
- You cannot increase metadata capacity by adding a new site.
- StorageGRID keeps three copies of all object metadata at every site. For this reason, the metadata capacity for your system is limited by the metadata capacity of your smallest site.
- When adding metadata capacity, you should add the same number of Storage Nodes to each site.

#### How metadata is redistributed when you add Storage Nodes

When you add Storage Nodes in an expansion, StorageGRID redistributes the existing object metadata to the new nodes at each site, which increases the overall metadata capacity of the grid. No user action is required.

The following figure shows how StorageGRID redistributes object metadata when you add Storage Nodes in an expansion. The left side of the figure represents volume 0 of three Storage Nodes before an expansion. Metadata is consuming a relatively large portion of each node's available metadata space, and the **Low metadata storage** alert has been triggered.

The right side of the figure shows how the existing metadata is redistributed after two Storage Nodes are added to the site. The amount of metadata on each node has decreased, the **Low metadata storage** alert is no longer triggered, and the space available for metadata has increased.




**Related information**

Administer StorageGRID

Monitor & troubleshoot

#### Adding grid nodes to add capabilities to your system

You can add redundancy or additional capabilities to a StorageGRID system by adding new grid nodes to existing sites.

For example, you might choose to add additional Gateway Nodes to support the creation of High Availability groups of Gateway Nodes, or you might add an Admin Node at a remote site to permit monitoring using a local node.

You can add one or more of the following types of nodes to one or more existing sites in a single expansion operation:

- Non-primary Admin Nodes
- Storage Nodes
- Gateway Nodes
- Archive Nodes

When preparing to add grid nodes, be aware of the following limitations:

- The primary Admin Node is deployed during the initial installation. You cannot add a primary Admin Node during an expansion.
- You can add Storage Nodes and other types of nodes in the same expansion.
- When adding Storage Nodes, you must carefully plan the number and location of the new nodes.

  Adding storage capacity

- If you are adding Archive Nodes, note that each Archive Node only supports tape through Tivoli Storage Manager (TSM) middleware.
- If the **New Node Client Network Default** option is set to **Untrusted** on the Untrusted Client Networks page, client applications that connect to expansion nodes using the Client Network must connect using a load balancer endpoint port (**Configuration** > **Network Settings** > **Untrusted Client Network**). See the instructions for administering StorageGRID to change the setting for the new node and to configure load balancer endpoints.

**Related information**

Administer StorageGRID

### Adding a new site

You can expand your StorageGRID system by adding a new site.

#### Guidelines for adding a site

Before adding a site, review the following requirements and limitations:

- You can only add one site per expansion operation.
- You cannot add grid nodes to an existing site as part of the same expansion.
- All sites must include at least three Storage Nodes.
- Adding a new site does not automatically increase the number of objects you can store. The total object capacity of a grid depends on the amount of available storage, the ILM policy, and the metadata capacity at each site.
- When sizing a new site, you must ensure that it includes enough metadata capacity.

  StorageGRID keeps a copy of all object metadata at every site. When you add a new site, you must ensure that it includes enough metadata capacity for the existing object metadata and enough metadata capacity for growth.

  For information on monitoring object metadata capacity, see the instructions for monitoring and troubleshooting StorageGRID.

- You must consider the available network bandwidth between sites, and the level of network latency. Metadata updates are continually replicated between sites even if all objects are stored only at the site where they are ingested.
- Because your StorageGRID system remains operational during the expansion, you must review ILM rules before starting the expansion procedure. You must ensure that object copies are not stored to the new site until the expansion procedure is complete.

For example, before you begin the expansion, determine if any rules use the default storage pool (All Storage Nodes). If they do, you must create a new storage pool that contains the existing Storage Nodes and update your ILM rules to use the new storage pool. Otherwise, objects will be copied to the new site as soon as the first node at that site becomes active.

For more information about changing ILM when adding a new site, see the example for changing an ILM policy in the instructions for managing objects with information lifecycle management.

**Related information**

Manage objects with ILM

## Preparing for an expansion

You must prepare for the StorageGRID expansion by obtaining required materials and installing and configuring any new hardware and networks.

### Gathering required materials

Before you perform an expansion operation, you must gather the materials listed in the following table.

| Item | Notes |
|---|---|
| StorageGRID installation archive | If you are adding new grid nodes or a new site, you must download and extract the StorageGRID installation archive. You must use the same version that is currently running on the grid.<br><br>For details, see the instructions for downloading and extracting the StorageGRID installation files.<br><br>**Note:** You do not need to download files if you are adding new storage volumes to existing Storage Nodes or installing a new StorageGRID appliance. |
| Service laptop | The service laptop must meet the following requirements:<br><br>• Network port<br>• SSH client (for example, PuTTY)<br>• Supported browser |
| Provisioning passphrase | The passphrase is created and documented when the StorageGRID system is first installed. The provisioning passphrase is not in the `Passwords.txt` file. |
| StorageGRID documentation | • *Administering StorageGRID*<br>• *StorageGRID Release Notes*<br>• Installation instructions for your platform |

| Item | Notes |
| --- | --- |
| Current documentation for your platform | For supported versions, see the Interoperability Matrix. |

**Related information**

Administer StorageGRID

Release notes

Install VMware

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

NetApp Interoperability Matrix Tool

#### Web browser requirements

You must use a supported web browser.

| Web browser | Minimum supported version |
| --- | --- |
| Google Chrome | 87 |
| Microsoft Edge | 87 |
| Mozilla Firefox | 84 |

You should set the browser window to a recommended width.

| Browser width | Pixels |
| --- | --- |
| Minimum | 1024 |
| Optimum | 1280 |

### Downloading and extracting the StorageGRID installation files

Before you can add new grid nodes or a new site, you must download the appropriate StorageGRID installation archive and extract the files.

#### About this task

You must perform expansion operations using the version of StorageGRID that is currently running on the grid.

#### Steps

1. Go to the NetApp Downloads page for StorageGRID.

   NetApp Downloads: StorageGRID

2. Select the version of StorageGRID that is currently running on the grid.
3. Sign in with the username and password for your NetApp account.
4. Read the End User License Agreement, select the check box, and then select **Accept & Continue**.
5. In the **Install StorageGRID** column of the download page, select the `.tgz` or `.zip` file for your platform.

   The version shown in the installation archive file must match the version of the software that is currently installed.

   Use the `.zip` file if you are running Windows on the service laptop.

| Platform | Installation archive |
|---|---|
| VMware | `StorageGRID-Webscale-`*`version`*`-VMware-`*`uniqueID`*`.zip`<br><br>`StorageGRID-Webscale-`*`version`*`-VMware-`*`uniqueID`*`.tgz` |
| Red Hat Enterprise Linux or CentOS | `StorageGRID-Webscale-`*`version`*`-RPM-`*`uniqueID`*`.zip`<br><br>`StorageGRID-Webscale-`*`version`*`-RPM-`*`uniqueID`*`.tgz` |
| Ubuntu or Debian and Appliance | `StorageGRID-Webscale-`*`version`*`-DEB-`*`uniqueID`*`.zip`<br><br>`StorageGRID-Webscale-`*`version`*`-DEB-`*`uniqueID`*`.tgz` |
| OpenStack/other Hypervisor | To expand an existing deployment on OpenStack, you must deploy a virtual machine running one of the supported Linux distributions listed above and follow the appropriate instructions for Linux. |

6. Download and extract the archive file.
7. Follow the appropriate step for your platform to choose the files you need, based on your platform, planned grid topology, and how you will expand your StorageGRID system.

   The paths listed in the step for each platform are relative to the top-level directory installed by the archive file.

8. If you are expanding a VMware system, select the appropriate files.

| Path and file name | Description |
|---|---|
| `./vsphere/README` | A text file that describes all of the files contained in the StorageGRID download file. |
| `./vsphere/NLF000000.txt` | A free license that does not provide any support entitlement for the product. |

| Path and file name | Description |
| --- | --- |
| `./vsphere/NetApp-SG-version-SHA.vmdk` | The virtual machine disk file that is used as a template for creating grid node virtual machines. |
| `./vsphere/vsphere-primary-admin.ovf`<br><br>`./vsphere/vsphere-primary-admin.mf` | The Open Virtualization Format template file (`.ovf`) and manifest file (`.mf`) for deploying the primary Admin Node. |
| `./vsphere/vsphere-non-primary-admin.ovf`<br><br>`./vsphere/vsphere-non-primary-admin.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying non-primary Admin Nodes. |
| `./vsphere/vsphere-archive.ovf`<br><br>`./vsphere/vsphere-archive.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying Archive Nodes. |
| `./vsphere/vsphere-gateway.ovf`<br><br>`./vsphere/vsphere-gateway.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying Gateway Nodes. |
| `./vsphere/vsphere-storage.ovf`<br><br>`./vsphere/vsphere-storage.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying virtual machine-based Storage Nodes. |
| **Deployment scripting tool** | **Description** |
| `./vsphere/deploy-vsphere-ovftool.sh` | A Bash shell script used to automate the deployment of virtual grid nodes. |
| `./vsphere/deploy-vsphere-ovftool-sample.ini` | A sample configuration file for use with the `deploy-vsphere-ovftool.sh` script. |
| `./vsphere/configure-storagegrid.py` | A Python script used to automate the configuration of a StorageGRID system. |
| `./vsphere/configure-sga.py` | A Python script used to automate the configuration of StorageGRID appliances. |
| `./vsphere/storagegrid-ssoauth.py` | An example Python script that you can use to sign in to the Grid Management API when single sign-on is enabled. |
| `./vsphere/configure-storagegrid.sample.json` | A sample configuration file for use with the `configure-storagegrid.py` script. |
| `./vsphere/configure-storagegrid.blank.json` | A blank configuration file for use with the `configure-storagegrid.py` script. |

9. If you are expanding a Red Hat Enterprise Linux or CentOS system, select the appropriate files.

| Path and file name | Description |
|---|---|
| `./rpms/README` | A text file that describes all of the files contained in the StorageGRID download file. |
| `./rpms/NLF000000.txt` | A free license that does not provide any support entitlement for the product. |
| `./rpms/StorageGRID-Webscale-Images-version-SHA.rpm` | RPM package for installing the StorageGRID node images on your RHEL or CentOS hosts. |
| `./rpms/StorageGRID-Webscale-Service-version-SHA.rpm` | RPM package for installing the StorageGRID host service on your RHEL or CentOS hosts. |
| **Deployment scripting tool** | **Description** |
| `./rpms/configure-storagegrid.py` | A Python script used to automate the configuration of a StorageGRID system. |
| `./rpms/configure-sga.py` | A Python script used to automate the configuration of StorageGRID appliances. |
| `./rpms/configure-storagegrid.sample.json` | A sample configuration file for use with the `configure-storagegrid.py` script. |
| `./rpms/storagegrid-ssoauth.py` | An example Python script that you can use to sign in to the Grid Management API when single sign-on is enabled. |
| `./rpms/configure-storagegrid.blank.json` | A blank configuration file for use with the `configure-storagegrid.py` script. |
| `./rpms/extras/ansible` | Example Ansible role and playbook for configuring RHEL or CentOS hosts for StorageGRID container deployment. You can customize the role or playbook as necessary. |

10. If you are expanding an Ubuntu or Debian system, select the appropriate files.

| Path and file name | Description |
|---|---|
| `./debs/README` | A text file that describes all of the files contained in the StorageGRID download file. |
| `./debs/NLF000000.txt` | A non-production NetApp License File that you can use for testing and proof of concept deployments. |

| Path and file name | Description |
|---|---|
| `./debs/storagegrid-webscale-images-version-SHA.deb` | DEB package for installing the StorageGRID node images on Ubuntu or Debian hosts. |
| `./debs/storagegrid-webscale-images-version-SHA.deb.md5` | MD5 checksum for the file `/debs/storagegrid-webscale-images-version-SHA.deb`. |
| `./debs/storagegrid-webscale-service-version-SHA.deb` | DEB package for installing the StorageGRID host service on Ubuntu or Debian hosts. |
| **Deployment scripting tool** | **Description** |
| `./debs/configure-storagegrid.py` | A Python script used to automate the configuration of a StorageGRID system. |
| `./debs/configure-sga.py` | A Python script used to automate the configuration of StorageGRID appliances. |
| `./debs/storagegrid-ssoauth.py` | An example Python script that you can use to sign in to the Grid Management API when single sign-on is enabled. |
| `./debs/configure-storagegrid.sample.json` | A sample configuration file for use with the `configure-storagegrid.py` script. |
| `./debs/configure-storagegrid.blank.json` | A blank configuration file for use with the `configure-storagegrid.py` script. |
| `./debs/extras/ansible` | Example Ansible role and playbook for configuring Ubuntu or Debian hosts for StorageGRID container deployment. You can customize the role or playbook as necessary. |

11. If you are expanding a StorageGRID appliance-based system, select the appropriate files.

| Path and file name | Description |
|---|---|
| `./debs/storagegrid-webscale-images-version-SHA.deb` | DEB package for installing the StorageGRID node images on your appliances. |
| `./debs/storagegrid-webscale-images-version-SHA.deb.md5` | Checksum of the DEB installation package used by the StorageGRID Appliance Installer to validate that the package is intact after upload. |

For appliance installation, these files are only required if you need to avoid network traffic. The appliance can download the required files from the primary Admin Node.

### Verifying hardware and networking

Before you begin the expansion of your StorageGRID system, you must ensure that you have installed and configured the necessary hardware to support the new grid nodes or new site.

For information about supported versions, see the Interoperability Matrix.

You must also verify network connectivity between servers at the site, and confirm that the primary Admin Node can communicate with all expansion servers that are intended to host the StorageGRID system.

If you are performing an expansion activity that includes adding a new subnet, you must add the new Grid subnet before you start the expansion procedure.

Do not use network address translation (NAT) on the Grid Network between grid nodes or between StorageGRID sites. When you use private IPv4 addresses for the Grid Network, those addresses must be directly routable from every grid node at every site. As required, however, you can use NAT between external clients and grid nodes, such as to provide a public IP address for a Gateway Node. Using NAT to bridge a public network segment is supported only when you employ a tunneling application that is transparent to all nodes in the grid, meaning the grid nodes require no knowledge of public IP addresses.

**Related information**

NetApp Interoperability Matrix Tool

Updating subnets for the Grid Network

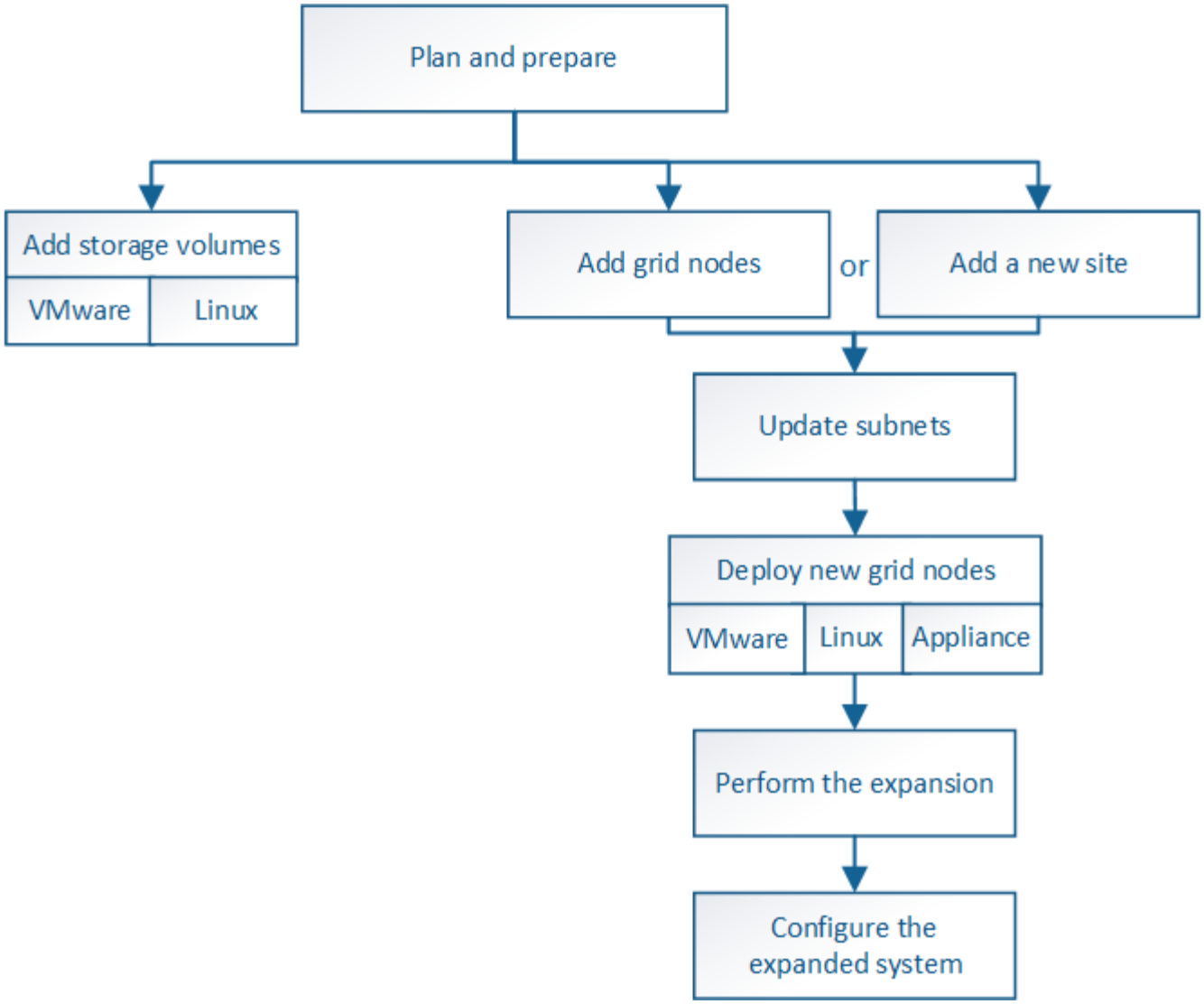
## Overview of expansion procedure

The basic steps for performing a StorageGRID expansion vary for the different types of expansion: adding storage volumes to a Storage Node, adding new nodes to an existing site, or adding a new site. In all cases, you can perform expansions without interrupting the operation of your current system.

The type of node that you are adding to the grid or the reason that you are adding nodes does not affect the basic expansion procedure. But as shown in the workflow diagram below, the steps for adding nodes vary slightly depending on whether you are adding StorageGRID appliances or hosts running VMware or Linux.

NetApp-provided virtual machine disk files and scripts for new installations or expansions of StorageGRID on OpenStack are no longer supported. To expand an existing deployment on OpenStack, refer to the steps for your Linux distribution.

“Linux” refers to a Red Hat® Enterprise Linux®, Ubuntu®, CentOS, or Debian® deployment. Use the NetApp Interoperability Matrix Tool to get a list of supported versions.




**Related information**

NetApp Interoperability Matrix Tool

Planning a StorageGRID expansion

Preparing for an expansion

Adding storage volumes to Storage Nodes

Adding grid nodes to an existing site or adding a new site

## Adding storage volumes to Storage Nodes

You can expand the storage capacity of Storage Nodes that have 16 or fewer storage volumes by adding additional storage volumes. You might need to add storage volumes to more than one Storage Node to satisfy ILM requirements for replicated or erasure-coded copies.

**What you'll need**

Before adding storage volumes, review the guidelines for adding storage capacity to ensure that you know where to add volumes to meet the requirements of your ILM policy.

Adding storage capacity

> These instructions apply to software-based Storage Nodes only. See the installation and maintenance instructions for the SG6060 appliance to learn how to add storage volumes to SG6060 by installing expansion shelves. Other appliance Storage Nodes cannot be expanded.
>
> SG6000 storage appliances

**About this task**

The underlying storage of a Storage Node is divided into a number of storage volumes. Storage volumes are block-based storage devices that are formatted by the StorageGRID system and mounted to store objects. Each Storage Node can support up to 16 storage volumes, which are called *object stores* in the Grid Manager.

> Object metadata is always stored in object store 0.

Each object store is mounted on a volume that corresponds to its ID. That is, the object store with an ID of 0000 corresponds to the `/var/local/rangedb/0` mount point.

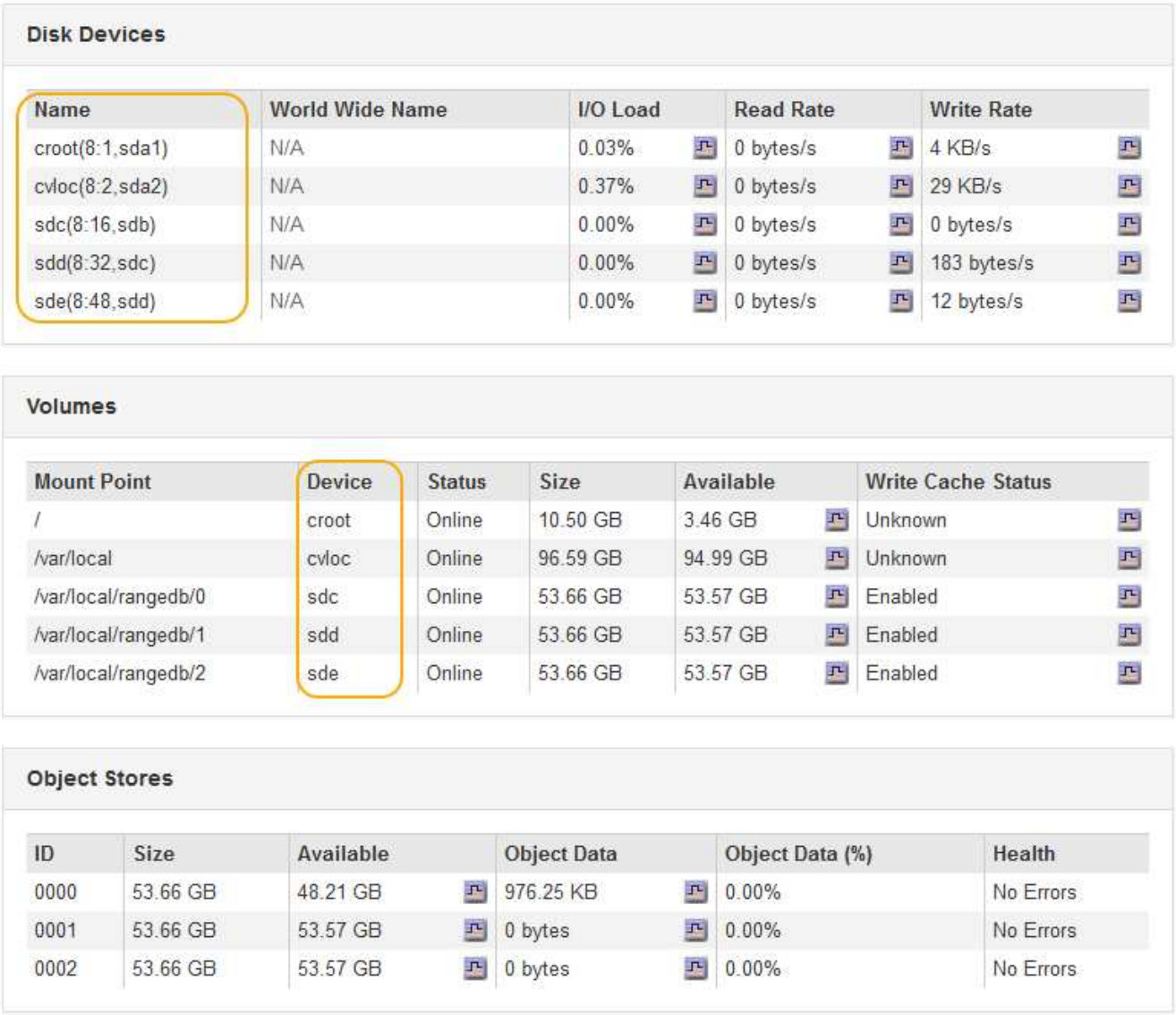
Before adding new storage volumes, use the Grid Manager to view the current object stores for each Storage Node as well as the corresponding mount points. You can use this information when adding storage volumes.

**Steps**

1. Select **Nodes** > ***site*** > ***Storage Node*** > **Storage**.
2. Scroll down to view the amounts of available storage for each volume and object store.

   For appliance Storage Nodes, the Worldwide Name for each disk matches the volume world-wide identifier (WWID) that appears when you view standard volume properties in SANtricity software (the management software connected to the appliance's storage controller).

   To help you interpret disk read and write statistics related to volume mount points, the first portion of the name shown in the **Name** column of the Disk Devices table (that is, *sdc*, *sdd*, *sde*, and so on) matches the value shown in the **Device** column of the Volumes table.



**Disk Devices**

| Name | World Wide Name | I/O Load | Read Rate | Write Rate |
|---|---|---|---|---|
| croot(8:1,sda1) | N/A | 0.03% | 0 bytes/s | 4 KB/s |
| cvloc(8:2,sda2) | N/A | 0.37% | 0 bytes/s | 29 KB/s |
| sdc(8:16,sdb) | N/A | 0.00% | 0 bytes/s | 0 bytes/s |
| sdd(8:32,sdc) | N/A | 0.00% | 0 bytes/s | 183 bytes/s |
| sde(8:48,sdd) | N/A | 0.00% | 0 bytes/s | 12 bytes/s |

**Volumes**

| Mount Point | Device | Status | Size | Available | Write Cache Status |
|---|---|---|---|---|---|
| / | croot | Online | 10.50 GB | 3.46 GB | Unknown |
| /var/local | cvloc | Online | 96.59 GB | 94.99 GB | Unknown |
| /var/local/rangedb/0 | sdc | Online | 53.66 GB | 53.57 GB | Enabled |
| /var/local/rangedb/1 | sdd | Online | 53.66 GB | 53.57 GB | Enabled |
| /var/local/rangedb/2 | sde | Online | 53.66 GB | 53.57 GB | Enabled |

**Object Stores**

| ID | Size | Available | Object Data | Object Data (%) | Health |
|---|---|---|---|---|---|
| 0000 | 53.66 GB | 48.21 GB | 976.25 KB | 0.00% | No Errors |
| 0001 | 53.66 GB | 53.57 GB | 0 bytes | 0.00% | No Errors |
| 0002 | 53.66 GB | 53.57 GB | 0 bytes | 0.00% | No Errors |

3. Follow the instructions for your platform to add new storage volumes to the Storage Node.
   - VMware: Adding storage volumes to a Storage Node
   - Linux: Adding direct-attached or SAN volumes to a Storage Node

### VMware: Adding storage volumes to a Storage Node

If a Storage Node includes fewer than 16 storage volumes, you can increase its capacity by using VMware vSphere to add volumes.

**What you'll need**

- You must have access to the instructions for installing StorageGRID for VMware deployments.
- You must have the `Passwords.txt` file.
- You must have specific access permissions.

> Do not attempt to add storage volumes to a Storage Node while a software upgrade, recovery procedure, or another expansion procedure is active.

**About this task**

The Storage Node is unavailable for a brief time when you add storage volumes. You should perform this procedure on one Storage Node at a time to avoid impacting client-facing grid services.

**Steps**

1. If necessary, install new storage hardware and create new VMware datastores.
2. Add one or more hard disks to the virtual machine for use as storage (object stores).

   a. Open VMware vSphere Client.

   b. Edit the virtual machine settings to add one or more additional hard disks.

   The hard disks are typically configured as Virtual Machine Disks (VMDKs). VMDKs are more commonly used and are easier to manage, while RDMs may provide better performance for workloads that use larger object sizes (for example, greater than 100 MB). For more information about adding hard disks to virtual machines, see the VMware vSphere documentation.

3. Restart the virtual machine by using the **Restart Guest OS** option in the VMware vSphere Client, or by entering the following command in an ssh session to the virtual machine:`sudo reboot`

   Do not use **Power Off** or **Reset** to restart the virtual machine.

4. Configure the new storage for use by the Storage Node:

   a. Log in to the grid node:

      i. Enter the following command: `ssh admin@grid_node_IP`

      ii. Enter the password listed in the `Passwords.txt` file.

      iii. Enter the following command to switch to root: `su -`

      iv. Enter the password listed in the `Passwords.txt` file. When you are logged in as root, the prompt changes from `$` to `#`.

   b. Configure the new storage volumes:

      ```
      sudo add_rangedbs.rb
      ```

      This script finds any new storage volumes and prompts you to format them.

   c. Enter **y** to accept the formatting.

   d. If any of the volumes have previously been formatted, decide if you want to reformat them.

      - Enter **y** to reformat.
      - Enter **n** to skip reformatting. The storage volumes are formatted.

   e. When asked, enter **y** to stop storage services.

      The storage services are stopped, and the `setup_rangedbs.sh` script runs automatically. After the volumes are ready for use as rangedbs, the services start again.

5. Check that the services start correctly:

   a. View a listing of the status of all services on the server:

      ```
      sudo storagegrid-status
      ```

The status is updated automatically.

b. Wait until all services are Running or Verified.

c. Exit the status screen:

```
Ctrl+C
```

6. Verify that the Storage Node is online:
   a. Sign in to the Grid Manager using a supported browser.
   b. Select **Support** > **Tools** > **Grid Topology**.
   c. Select ***site*** > ***Storage Node*** > **LDR** > **Storage**.
   d. Select the **Configuration** tab and then the **Main** tab.
   e. If the **Storage State - Desired** drop-down list is set to Read-only or Offline, select **Online**.
   f. Click **Apply Changes**.
7. To see the new object stores:
   a. Select **Nodes** > ***site*** > ***Storage Node*** > **Storage**.
   b. View the details in the **Object Stores** table.

**Result**

You can now use the expanded capacity of the Storage Nodes to save object data.

**Related information**

Install VMware

### Linux: Adding direct-attached or SAN volumes to a Storage Node

If a Storage Node includes fewer than 16 storage volumes, you can increase its capacity by adding new block storage devices, making them visible to the Linux hosts, and adding the new block device mappings to the StorageGRID configuration file used for the Storage Node.

**What you'll need**

- You must have access to the instructions for installing StorageGRID for your Linux platform.
- You must have the `Passwords.txt` file.
- You must have specific access permissions.

> Do not attempt to add storage volumes to a Storage Node while a software upgrade, recovery procedure, or another expansion procedure is active.

**About this task**

The Storage Node is unavailable for a brief time when you add storage volumes. You should perform this procedure on one Storage Node at a time to avoid impacting client-facing grid services.

**Steps**

1. Install the new storage hardware.

For more information, see the documentation provided by your hardware vendor.

2. Create new block storage volumes of the desired sizes.
   - Attach the new disk drives and update the RAID controller configuration as needed, or allocate the new SAN LUNs on the shared storage arrays and allow the Linux host to access them.
   - Use the same persistent naming scheme you used for the storage volumes on the existing Storage Node.
   - If you use the StorageGRID node migration feature, make the new volumes visible to other Linux hosts that are migration targets for this Storage Node. For more information, see the instructions for installing StorageGRID for your Linux platform.
3. Log into the Linux host supporting the Storage Node as root or with an account that has sudo permission.
4. Confirm that the new storage volumes are visible on the Linux host.

   You might have to rescan for devices.

5. Run the following command to temporarily disable the Storage Node:

   ```
   sudo storagegrid node stop <node-name>
   ```

6. Using a text editor such as vim or pico, edit the node configuration file for the Storage Node, which can be found at `/etc/storagegrid/nodes/<node-name>.conf`.
7. Locate the section of the node configuration file that contains the existing object storage block device mappings.

   In the example, `BLOCK_DEVICE_RANGEDB_00` to `BLOCK_DEVICE_RANGEDB_03` are the existing object storage block device mappings.

   ```
   NODE_TYPE = VM_Storage_Node
   ADMIN_IP = 10.1.0.2
   BLOCK_DEVICE_VAR_LOCAL = /dev/mapper/sgws-sn1-var-local
   BLOCK_DEVICE_RANGEDB_00 = /dev/mapper/sgws-sn1-rangedb-0
   BLOCK_DEVICE_RANGEDB_01 = /dev/mapper/sgws-sn1-rangedb-1
   BLOCK_DEVICE_RANGEDB_02 = /dev/mapper/sgws-sn1-rangedb-2
   BLOCK_DEVICE_RANGEDB_03 = /dev/mapper/sgws-sn1-rangedb-3
   GRID_NETWORK_TARGET = bond0.1001
   ADMIN_NETWORK_TARGET = bond0.1002
   CLIENT_NETWORK_TARGET = bond0.1003
   GRID_NETWORK_IP = 10.1.0.3
   GRID_NETWORK_MASK = 255.255.255.0
   GRID_NETWORK_GATEWAY = 10.1.0.1
   ```

8. Add new object storage block device mappings corresponding to the block storage volumes you added for this Storage Node.

   Make sure to start at the next `BLOCK_DEVICE_RANGEDB_nn`. Do not leave a gap.

   - Based on the example above, start at `BLOCK_DEVICE_RANGEDB_04`.

- In the example below, four new block storage volumes have been added to the node: `BLOCK_DEVICE_RANGEDB_04` to `BLOCK_DEVICE_RANGEDB_07`.

```
NODE_TYPE = VM_Storage_Node
ADMIN_IP = 10.1.0.2
BLOCK_DEVICE_VAR_LOCAL = /dev/mapper/sgws-sn1-var-local
BLOCK_DEVICE_RANGEDB_00 = /dev/mapper/sgws-sn1-rangedb-0
BLOCK_DEVICE_RANGEDB_01 = /dev/mapper/sgws-sn1-rangedb-1
BLOCK_DEVICE_RANGEDB_02 = /dev/mapper/sgws-sn1-rangedb-2
BLOCK_DEVICE_RANGEDB_03 = /dev/mapper/sgws-sn1-rangedb-3
<strong>BLOCK_DEVICE_RANGEDB_04 = /dev/mapper/sgws-sn1-rangedb-
4</strong>
<strong>BLOCK_DEVICE_RANGEDB_05 = /dev/mapper/sgws-sn1-rangedb-
5</strong>
<strong>BLOCK_DEVICE_RANGEDB_06 = /dev/mapper/sgws-sn1-rangedb-
6</strong>
<strong>BLOCK_DEVICE_RANGEDB_07 = /dev/mapper/sgws-sn1-rangedb-
7</strong>
GRID_NETWORK_TARGET = bond0.1001
ADMIN_NETWORK_TARGET = bond0.1002
CLIENT_NETWORK_TARGET = bond0.1003
GRID_NETWORK_IP = 10.1.0.3
GRID_NETWORK_MASK = 255.255.255.0
GRID_NETWORK_GATEWAY = 10.1.0.1
```

9. Run the following command to validate your changes to the node configuration file for the Storage Node:

   `sudo storagegrid node validate <node-name>`

   Address any errors or warnings before proceeding to the next step.

   If you observe an error similar to the following, it means that the node configuration file is attempting to map the block device used by `<node-name>` for `<PURPOSE>` to the given `<path-name>` in the Linux file system, but there is not a valid block device special file (or softlink to a block device special file) at that location.

   ```
   Checking configuration file for node <node-name>…
   ERROR: BLOCK_DEVICE_<PURPOSE> = <path-name>
   <path-name> is not a valid block device
   ```

   Verify that you entered the correct `<path-name>`.

10. Run the following command to restart the node with the new block device mappings in place:

    `sudo storagegrid node start <node-name>`

11. Log in to the Storage Node as admin using the password listed in the `Passwords.txt` file.
12. Check that the services start correctly:
    a. View a listing of the status of all services on the server:
    `sudo storagegrid-status`

    The status is updated automatically.

    b. Wait until all services are Running or Verified.
    c. Exit the status screen:

    `Ctrl+C`

13. Configure the new storage for use by the Storage Node:
    a. Configure the new storage volumes:

    `sudo add_rangedbs.rb`

    This script finds any new storage volumes and prompts you to format them.

    b. Enter **y** to format the storage volumes.
    c. If any of the volumes have previously been formatted, decide if you want to reformat them.
        - Enter **y** to reformat.
        - Enter **n** to skip reformatting. The storage volumes are formatted.
    d. When asked, enter **y** to stop storage services.

    The storage services are stopped, and the `setup_rangedbs.sh` script runs automatically. After the volumes are ready for use as rangedbs, the services start again.

14. Check that the services start correctly:
    a. View a listing of the status of all services on the server:

    `sudo storagegrid-status`

    The status is updated automatically.

    b. Wait until all services are Running or Verified.
    c. Exit the status screen:

    `Ctrl+C`

15. Verify that the Storage Node is online:
    a. Sign in to the Grid Manager using a supported browser.
    b. Select **Support** > **Tools** > **Grid Topology**.
    c. Select ***site*** > ***Storage Node*** > **LDR** > **Storage**.
    d. Select the **Configuration** tab and then the **Main** tab.
    e. If the **Storage State - Desired** drop-down list is set to Read-only or Offline, select **Online**.
    f. Click **Apply Changes**.

16. To see the new object stores:
    a. Select **Nodes** > ***site*** > ***Storage Node*** > **Storage**.
    b. View the details in the **Object Stores** table.

**Result**

You can now use the expanded capacity of the Storage Nodes to save object data.

**Related information**

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

## Adding grid nodes to an existing site or adding a new site

You can follow this procedure to add grid nodes to existing sites or to add a new site, but you cannot perform both types of expansion at the same time.

**What you'll need**

- You must have root or maintenance permissions. For details, see information about controlling system access with administration user accounts and groups.
- All existing nodes in the grid must be up and running across all sites.
- Any previous expansion, upgrade, decommissioning, or recovery procedures must be complete.

> You are prevented from starting an expansion while another expansion, upgrade, recovery, or active decommission procedure is in progress. However, if necessary, you can pause a decommission procedure to start an expansion.

**Steps**

1. Updating subnets for the Grid Network
2. Deploying new grid nodes
3. Performing the expansion

### Updating subnets for the Grid Network

When you add grid nodes or a new site in an expansion, you might need to update or add subnets to the Grid Network.

StorageGRID maintains a list of the network subnets used to communicate between grid nodes on the Grid Network (eth0). These entries include the subnets used for the Grid Network by each site in your StorageGRID system as well as any subnets used for NTP, DNS, LDAP, or other external servers accessed through the Grid Network gateway.

**What you'll need**

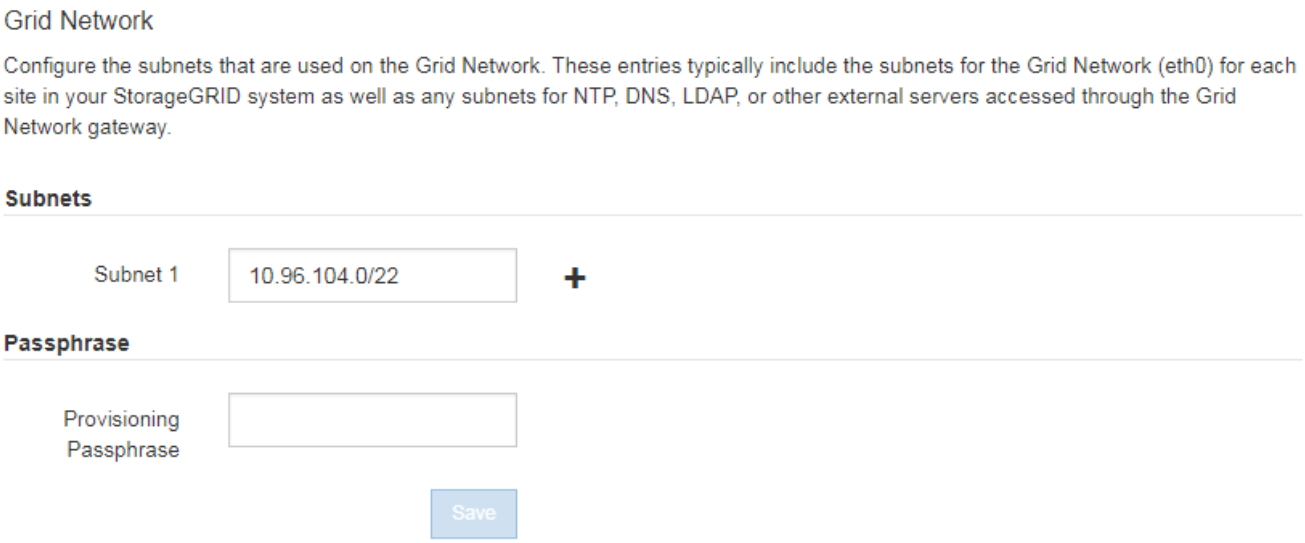
- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have the network addresses, in CIDR notation, of the subnets you want to configure.

### About this task

If you are performing an expansion activity that includes adding a new subnet, you must add the new Grid subnet before you start the expansion procedure.

### Steps

1. Select **Maintenance** > **Network** > **Grid Network**.




2. In the Subnets list, click the plus sign to add a new subnet in CIDR notation.

   For example, enter 10.96.104.0/22.

3. Enter the provisioning passphrase, and click **Save**.

   The subnets you have specified are configured automatically for your StorageGRID system.

### Deploying new grid nodes

The steps for deploying new grid nodes in an expansion are the same as the steps used when the grid was first installed. You must deploy all new grid nodes before you can perform the expansion.

When you expand the grid, the nodes you add do not have to match the existing node types. You can add VMware nodes, Linux container-based nodes, or appliance nodes.

#### VMware: Deploying grid nodes

You must deploy a virtual machine in VMware vSphere for each VMware node you want to add in the expansion.

### Steps

1. Deploy the new grid node as a virtual machine and connect it to one or more StorageGRID networks.

   When you deploy the node, you can optionally remap node ports or increase CPU or memory settings.

   Deploying a StorageGRID node as a virtual machine

2. After you have deployed all new VMware nodes, return to these instructions to perform the expansion procedure.

   Performing the expansion

#### Linux: Deploying grid nodes

You can deploy grid nodes on new Linux hosts or on existing Linux hosts. If you need additional Linux hosts to support the CPU, RAM, and storage requirements of the StorageGRID nodes you want to add to your grid, you prepare them in the same way you prepared the hosts when you first installed them. Then, you deploy the expansion nodes in the same way you deployed grid nodes during installation.

**What you'll need**

- You have the instructions for installing StorageGRID for your version of Linux, and you have reviewed the hardware and storage requirements.
- If you plan to deploy new grid nodes on existing hosts, you have confirmed the existing hosts have enough CPU, RAM, and storage capacity for the additional nodes.
- You have a plan to minimize failure domains. For example, you should not deploy all Gateway Nodes on a single physical host.

  > In a production deployment, do not run more than one Storage Node on a single physical or virtual host. Using a dedicated host for each Storage Node provides an isolated failure domain.

- If the StorageGRID node uses storage assigned from a NetApp AFF system, confirm that the volume does not have a FabricPool tiering policy enabled. Disabling FabricPool tiering for volumes used with StorageGRID nodes simplifies troubleshooting and storage operations.

  > Never use FabricPool to tier any data related to StorageGRID back to StorageGRID itself. Tiering StorageGRID data back to StorageGRID increases troubleshooting and operational complexity.

**Steps**

1. If you are adding new hosts, access the installation instructions for deploying StorageGRID nodes.
2. To deploy the new hosts, follow the instructions for preparing the hosts.
3. To create node configuration files and to validate the StorageGRID configuration, follow the instructions for deploying grid nodes.
4. If you are adding nodes to a new Linux host, start the StorageGRID host service.
5. If you are adding nodes to an existing Linux host, start the new nodes using the storagegrid host service CLI:`sudo storagegrid node start [<node name\>]`

**After you finish**

After deploying all new grid nodes, you can perform the expansion.

**Related information**

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

## Performing the expansion

### Appliances: Deploying Storage, Gateway, or non-primary Admin Nodes

To install the StorageGRID software on an appliance node, you use the StorageGRID Appliance Installer, which is included on the appliance. In an expansion, each storage appliance functions as a single Storage Node, and each services appliance functions as a single Gateway Node or non-primary Admin Node. Any appliance can connect to the Grid Network, the Admin Network, and the Client Network.

**What you'll need**

- The appliance has been installed in a rack or cabinet, connected to your networks, and powered on.
- You have used the StorageGRID Appliance Installer to complete all of the "configuring the hardware" steps in the appliance installation and maintenance instructions.

  Configuring appliance hardware includes the required steps for configuring StorageGRID connections (network links and IP addresses) as well the optional steps for enabling node encryption, changing the RAID mode, and remapping network ports.

- All Grid Network subnets listed on the IP Configuration page of the StorageGRID Appliance Installer have been defined in the Grid Network Subnet List on the primary Admin Node.
- The StorageGRID Appliance Installer version on the replacement appliance matches the software version of your StorageGRID system. (If the versions do not match, you must upgrade the StorageGRID Appliance Installer firmware.)

  For instructions, see the appliance installation and maintenance instructions.

  - SG100 & SG1000 services appliances
  - SG5600 storage appliances
  - SG5700 storage appliances
  - SG6000 storage appliances
- You have a service laptop with a supported web browser.
- You know one of the IP addresses assigned to the appliance's compute controller. You can use the IP address for any attached StorageGRID network.

**About this task**

The process of installing StorageGRID on an appliance node has the following phases:

- You specify or confirm the IP address of the primary Admin Node and the name of the appliance node.
- You start the installation and wait as volumes are configured and the software is installed.

  Partway through appliance installation tasks, the installation pauses. To resume the installation, you sign into the Grid Manager, approve all grid nodes, and complete the StorageGRID installation process.

> If you need to deploy multiple appliance nodes at one time, you can automate the installation process by using the `configure-sga.py` Appliance Installation Script.

**Steps**

1. Open a browser, and enter one of the IP addresses for the appliance's compute controller.

`https://Controller_IP:8443`

The StorageGRID Appliance Installer Home page appears.

2. In the **Primary Admin Node** connection section, determine whether you need to specify the IP address for the primary Admin Node.

   If you have previously installed other nodes in this data center, the StorageGRID Appliance Installer can discover this IP address automatically, assuming the primary Admin Node, or at least one other grid node with ADMIN_IP configured, is present on the same subnet.

3. If this IP address is not shown or you need to change it, specify the address:

| Option | Description |
|---|---|
| Manual IP entry | a. Unselect the **Enable Admin Node discovery** check box.<br>b. Enter the IP address manually.<br>c. Click **Save**.<br>d. Wait for the connection state for the new IP address to become ready. |
| Automatic discovery of all connected primary Admin Nodes | a. Select the **Enable Admin Node discovery** check box.<br>b. Wait for the list of discovered IP addresses to be displayed.<br>c. Select the primary Admin Node for the grid where this appliance Storage Node will be deployed.<br>d. Click **Save**.<br>e. Wait for the connection state for the new IP address to become ready. |

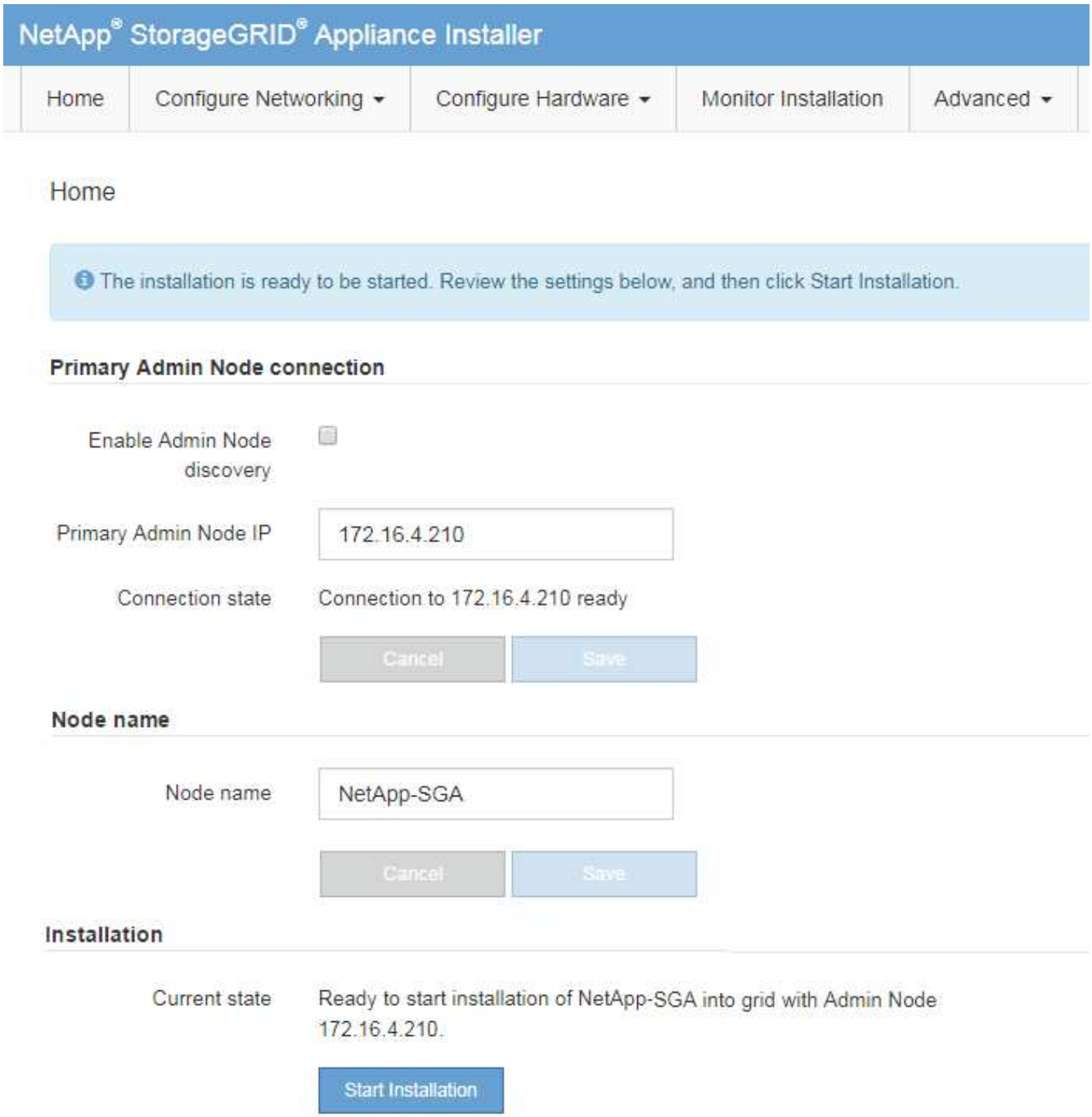
4. In the **Node name** field, enter the name you want to use for this appliance node, and click **Save**.

   The node name is assigned to this appliance node in the StorageGRID system. It is shown on the Nodes page (Overview tab) in the Grid Manager. If required, you can change the name when you approve the node.

5. In the **Installation** section, confirm that the current state is “Ready to start installation of *node name* into grid with primary Admin Node *admin_ip*” and that the **Start Installation** button is enabled.

   If the **Start Installation** button is not enabled, you might need to change the network configuration or port settings. For instructions, see the installation and maintenance instructions for your appliance.

6. From the StorageGRID Appliance Installer home page, click **Start Installation**.



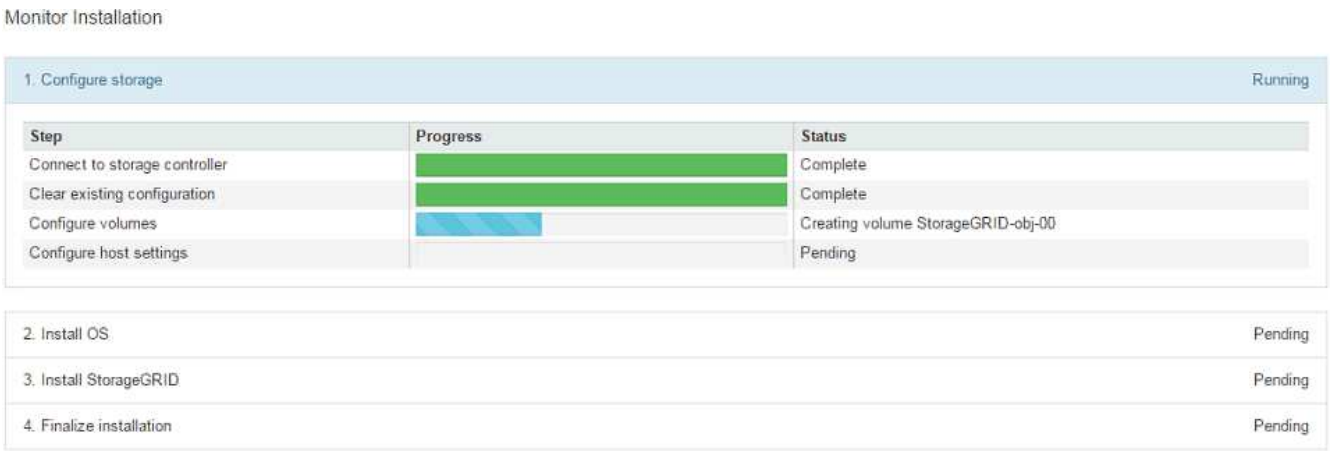

The Current state changes to “Installation is in progress,” and the Monitor Installation page is displayed.

7. If your expansion includes multiple appliance nodes, repeat the previous steps for each appliance.

> If you need to deploy multiple appliance Storage Nodes at one time, you can automate the installation process by using the configure-sga.py appliance installation script.

8. If you need to manually access the Monitor Installation page, click **Monitor Installation** from the menu bar.

   The Monitor Installation page shows the installation progress.




The blue status bar indicates which task is currently in progress. Green status bars indicate tasks that have completed successfully.

> The installer ensures that tasks completed in a previous install are not re-run. If you are re-running an installation, any tasks that do not need to be re-run are shown with a green status bar and a status of “Skipped.”

9. Review the progress of first two installation stages.

   **1. Configure appliance**

   During this stage, one of the following processes occurs:

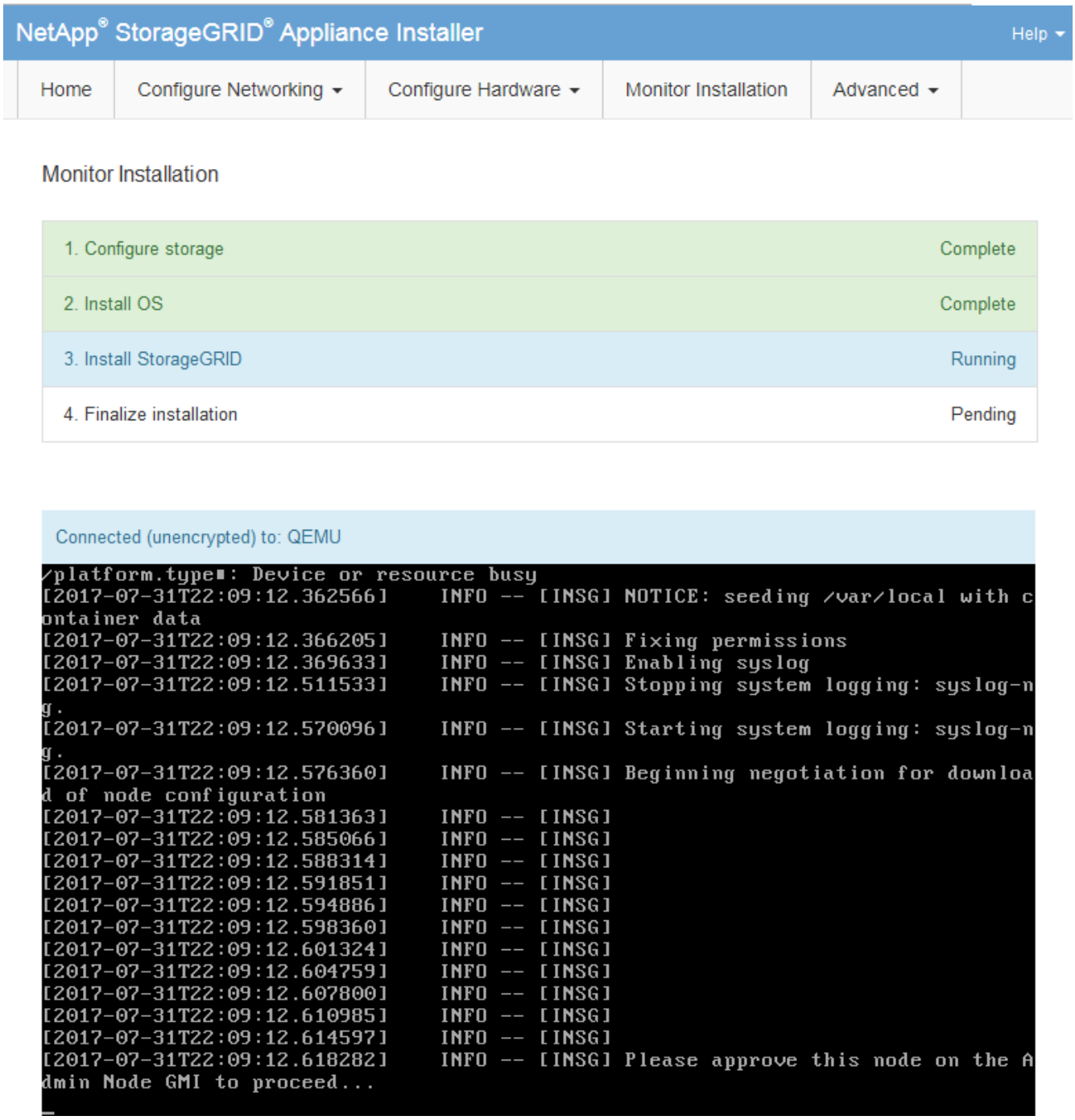
   - For a storage appliance, the installer connects to the storage controller, clears any existing configuration, communicates with SANtricity software to configure volumes, and configures host settings.
   - For a services appliance, the installer clears any existing configuration from the drives in the compute controller, and configures host settings.

   **2. Install OS**

   During this stage, the installer copies the base operating system image for StorageGRID to the appliance.

10. Continue monitoring the installation progress until a message appears in the console window, prompting you to use the Grid Manager to approve the node.

> Wait until all nodes you added in this expansion are ready for approval before going to the Grid Manager to approve the nodes.




**Related information**

SG5700 storage appliances

SG5600 storage appliances

SG6000 storage appliances

SG100 & SG1000 services appliances

**Performing the expansion**

When you perform the expansion, the new grid nodes are added to your existing

StorageGRID deployment.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have deployed all of the grid nodes that are being added in this expansion.
- If you are adding Storage Nodes, you must have confirmed that all data-repair operations performed as part of a recovery are complete. See the steps for checking data repair jobs in the recovery and maintenance instructions.
- If you are adding a new site, you must review and update ILM rules before starting the expansion procedure to ensure that object copies are not stored to the new site until after the expansion is complete. For example, if a rule uses the default storage pool (All Storage Nodes), you must create a new storage pool that contains only the existing Storage Nodes and update the ILM rule to use the new storage pool. Otherwise, objects will be copied to the new site as soon as the first node at that site becomes active. See the instructions for managing objects with information lifecycle management.

**About this task**

Performing the expansion includes these phases:

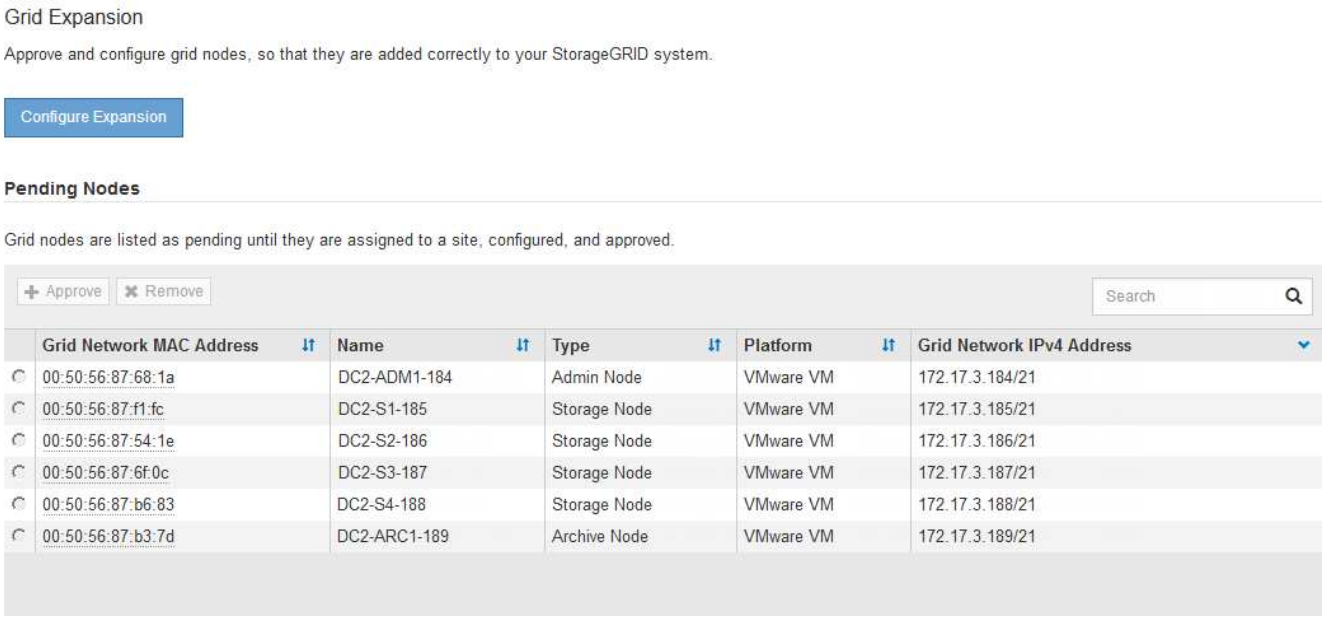
1. You configure the expansion by specifying whether you are adding new grid nodes or a new site and approving the grid nodes you want to add.
2. You start the expansion.
3. While the expansion process is running, you download a new Recovery Package file.
4. You monitor the status of the grid configuration tasks, which run automatically. The set of tasks depends on what types of grid nodes are being added and on whether a new site is being added.

Some tasks might take a significant amount of time to run on a large grid. For example, streaming Cassandra to a new Storage Node might take only a few minutes if the Cassandra database is relatively empty. However, if the Cassandra database includes a large amount of object metadata, this stage might take several hours or longer. You can look at the "streamed" percentage shown during the "Starting Cassandra and streaming data" stage to determine how complete the Cassandra streaming operation is.

**Steps**

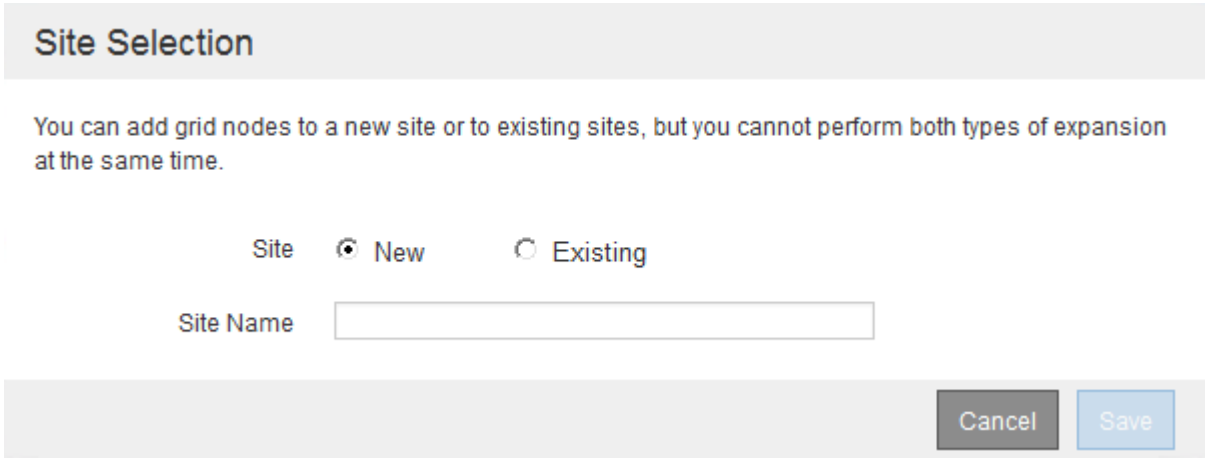
1. Select **Maintenance** > **Maintenance Tasks** > **Expansion**.

   The Grid Expansion page appears. The Pending Nodes section lists all nodes that are ready to be added.



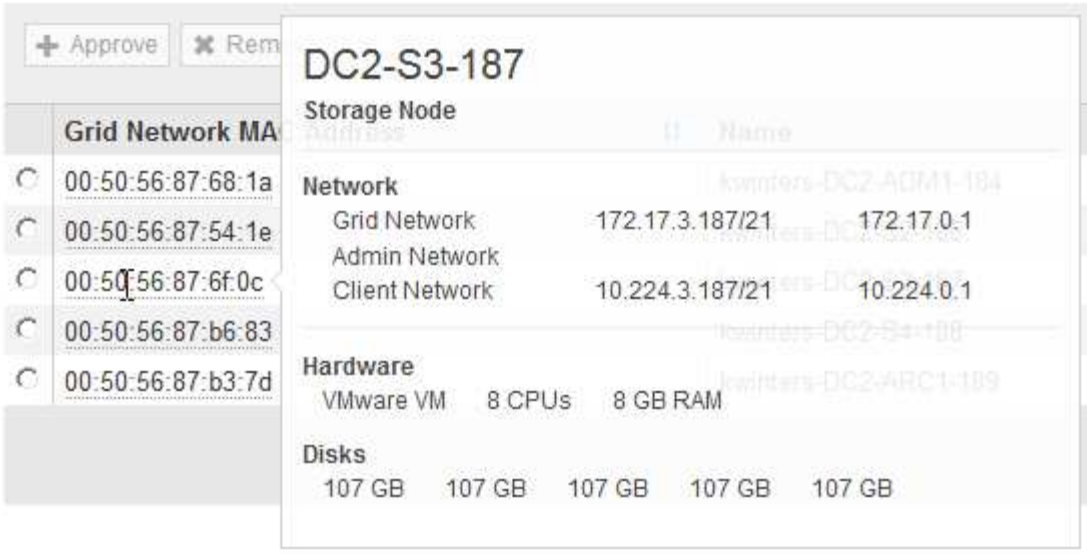

2. Click **Configure Expansion**.

   The Site Selection dialog box appears.




3. Select the type of expansion you are starting:
   - If you are adding a new site, select **New**, and enter the name of the new site.
   - If you are adding grid nodes to an existing site, select **Existing**.
4. Click **Save**.
5. Review the **Pending Nodes** list, and confirm that it shows all of the grid nodes you deployed.

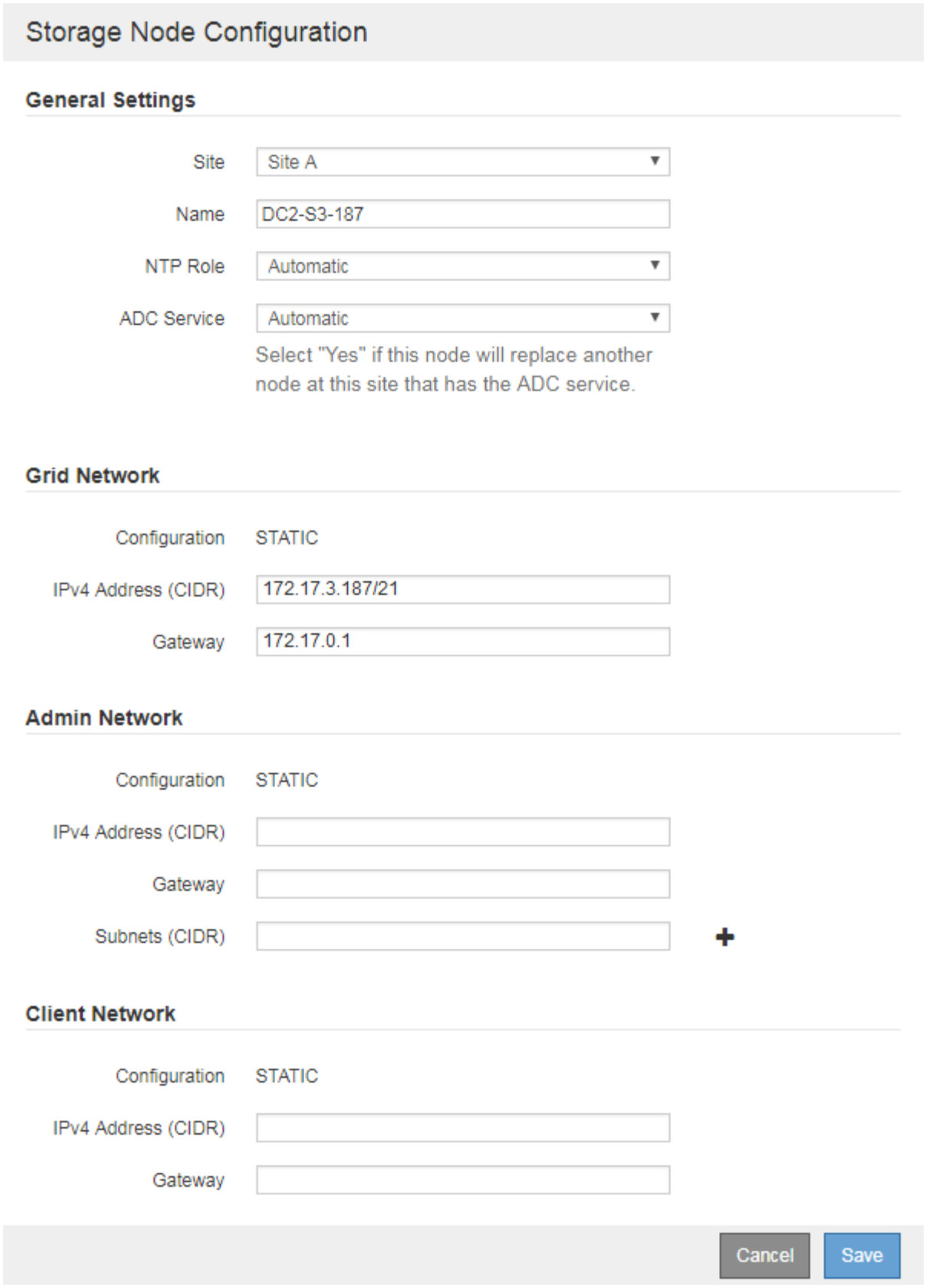
   As required, you can hover your cursor over a node's **Grid Network MAC Address** to see details about that node.




If a grid node is missing, confirm that it was deployed successfully.

6. From the list of pending nodes, approve the grid nodes for this expansion.
   a. Select the radio button next to the first pending grid node you want to approve.
   b. Click **Approve**.

      The grid node configuration form appears.




c. As required, modify the general settings:

- **Site**: The name of the site the grid node will be associated with. If you are adding multiple nodes, be sure to select the correct site for each node. If you are adding a new site, all nodes are added to the new site.

- **Name**: The hostname that will be assigned to the node, and the name that will be displayed in the Grid Manager.
- **NTP Role**: The Network Time Protocol (NTP) role of the grid node. The options are **Automatic**, **Primary**, and **Client**. Selecting **Automatic** assigns the Primary role to Admin Nodes, Storage Nodes with ADC services, Gateway Nodes, and any grid nodes that have non-static IP addresses. All other grid nodes are assigned the Client role.

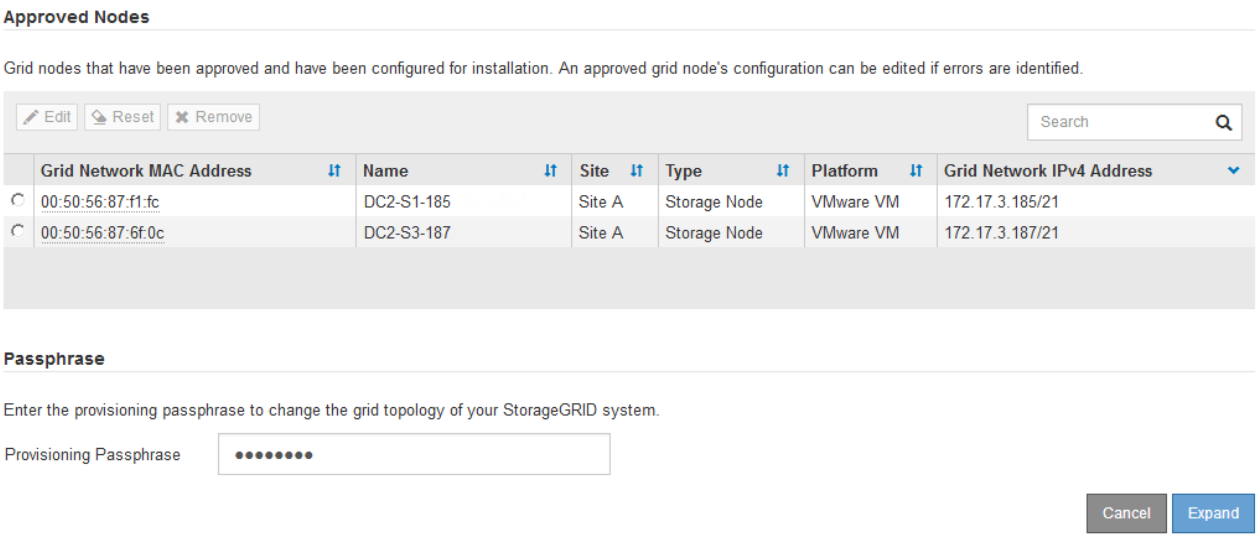
  > Assign the Primary NTP role to at least two nodes at each site. This provides redundant system access to external timing sources.

- **ADC Service** (Storage Nodes only): Whether this Storage Node will run the Administrative Domain Controller (ADC) service. The ADC service keeps track of the location and availability of grid services. At least three Storage Nodes at each site must include the ADC service. You cannot add the ADC service to a node after it is deployed.
  - If you are adding this node to replace a Storage Node, select **Yes** if the node you are replacing includes the ADC service. Because you cannot decommission a Storage Node if too few ADC services would remain, this ensures that a new ADC service is available before the old service is removed.
  - Otherwise, select **Automatic** to let the system determine whether this node requires the ADC service. Learn about the ADC quorum in the recovery and maintenance instructions.

d. As required, modify the settings for the Grid Network, Admin Network, and Client Network.
   - **IPv4 Address (CIDR)**: The CIDR network address for the network interface. For example: 172.16.10.100/24
   - **Gateway**: The default gateway of the grid node. For example: 172.16.10.1
   - **Subnets (CIDR)**: One or more subnetworks for the Admin Network.

e. Click **Save**.

   The approved grid node moves to the Approved Nodes list.




- To modify the properties of an approved grid node, select its radio button, and click **Edit**.
- To move an approved grid node back to the Pending Nodes list, select its radio button, and click **Reset**.

- To permanently remove an approved grid node, power the node off. Then, select its radio button, and click **Remove**.

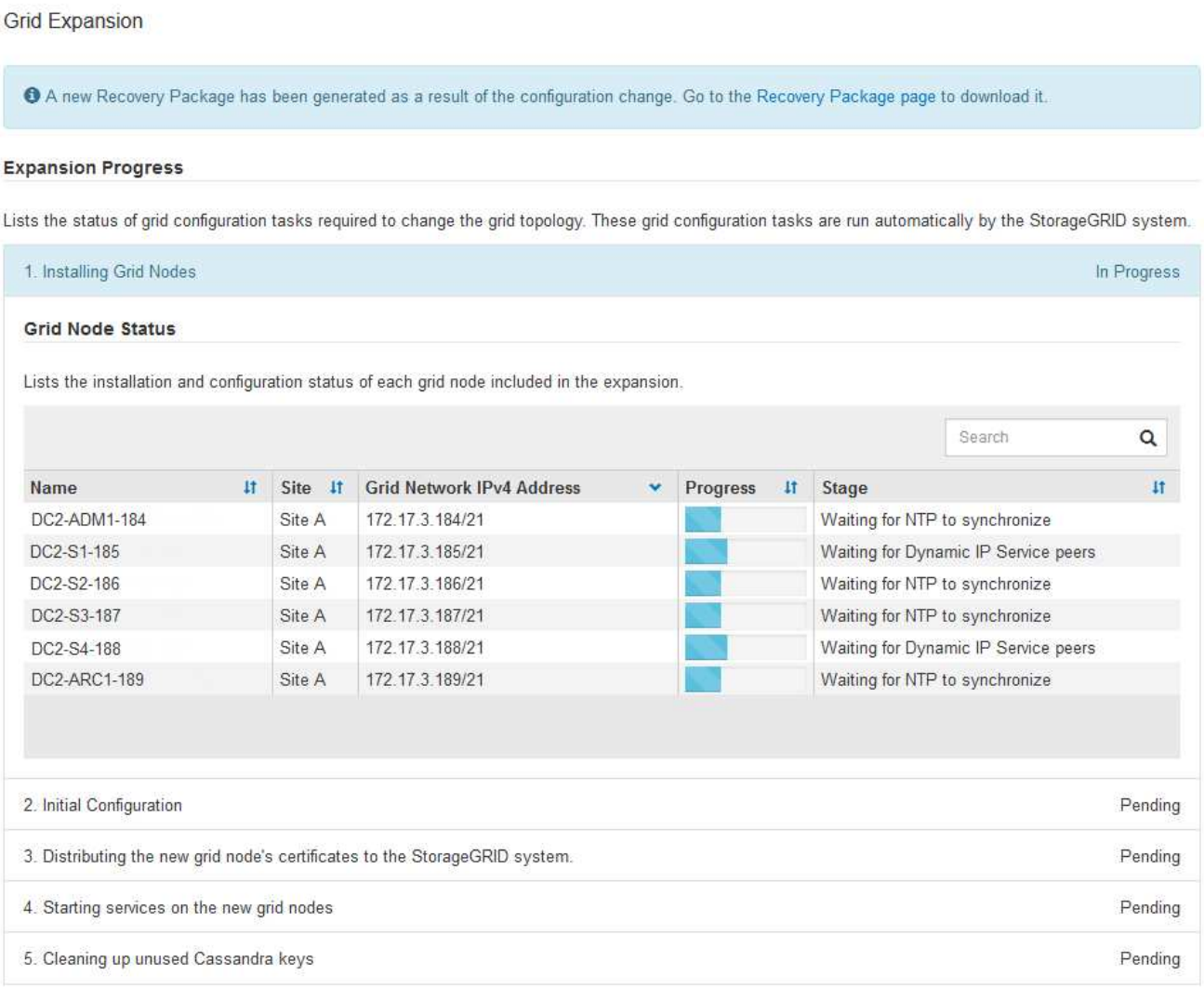
f. Repeat these steps for each pending grid node you want to approve.

> If possible, you should approve all pending grid notes and perform a single expansion. More time will be required if you perform multiple small expansions.

7. When you have approved all grid nodes, enter the **Provisioning Passphrase**, and click **Expand**.

   After a few minutes, this page updates to display the status of the expansion procedure. When tasks that affect individual grid node are in progress, the Grid Node Status section lists the current status for each grid node.

> During this process, for appliances the StorageGRID Appliance Installer shows installation moving from Stage 3 to Stage 4, Finalize Installation. When Stage 4 completes, the controller is rebooted.




> A site expansion includes an additional task to configure Cassandra for the new site.

8. As soon as the **Download Recovery Package** link appears, download the Recovery Package file.

You must download an updated copy of the Recovery Package file as soon as possible after making grid topology changes to the StorageGRID system. The Recovery Package file allows you to restore the system if a failure occurs.

a. Click the download link.

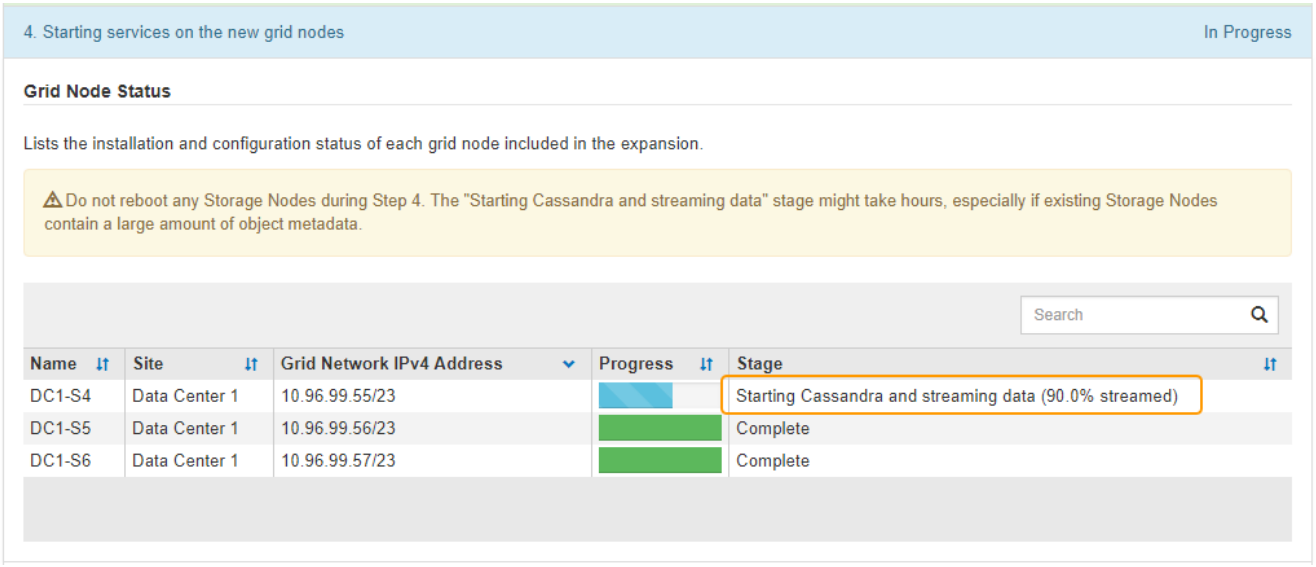
b. Enter the provisioning passphrase, and click **Start Download**.

c. When the download completes, open the `.zip` file and confirm it includes a `gpt-backup` directory and a `_SAID.zip` file. Then, extract the `_SAID.zip` file, go to the `/GID*_REV*` directory, and confirm you can open the `passwords.txt` file.

d. Copy the downloaded Recovery Package file (.zip) to two safe, secure, and separate locations.

> The Recovery Package file must be secured because it contains encryption keys and passwords that can be used to obtain data from the StorageGRID system.

9. If you are adding one or more Storage Nodes, monitor the progress of the “Starting Cassandra and streaming data” stage by reviewing the percentage shown in the status message.




This percentage estimates how complete the Cassandra streaming operation is, based on the total amount of Cassandra data available and the amount that has already been written to the new node.

> Do not reboot any Storage Nodes during Step 4 (Starting services on the new grid nodes). The “Starting Cassandra and streaming data” stage might take hours to complete for each new Storage Node, especially if existing Storage Nodes contain a large amount of object metadata.

10. Continue monitoring the expansion until all tasks are complete and the **Configure Expansion** button reappears.

### After you finish

Depending on which types of grid nodes you added, you must perform additional integration and configuration steps.

### Related information

Manage objects with ILM

# Configuring your expanded StorageGRID system

After completing an expansion, you must perform additional integration and configuration steps.

### About this task

You must complete the configuration tasks listed below for the grid nodes you are adding in your expansion. Some tasks might be optional, depending on the options selected when installing and administering your system, and how you want to configure the grid nodes added during the expansion.

### Steps

1. If you added a Storage Node, complete the following configuration tasks.

| Storage Node configuration tasks | For information |
|---|---|
| Review the storage pools used in your ILM rules to ensure the new storage will be used.<br>• If you added a site, create a storage pool for the site and update ILM rules to use the new storage pool.<br>• If you added a Storage Node to an existing site, confirm that the new node uses the correct storage grade.<br>**Note:** By default, a new Storage Node is assigned to the All Storage Nodes storage grade and added to storage pools that use that grade for the site. If you want a new node to use a custom storage grade, you must assign it to the custom grade manually (**ILM** > **Storage Grades**). | Manage objects with ILM |
| Verify that the Storage Node is ingesting objects. | Verifying that the Storage Node is active |
| Rebalance erasure-coded data (only if you were unable to add the recommended number of Storage Nodes). | Rebalancing erasure-coded data after adding Storage Nodes |

2. If you added a Gateway Node, complete the following configuration tasks.

| Gateway Node configuration tasks | For information |
|---|---|
| If High Availability Groups are used for client connections, add the Gateway Nodes to an HA group. Select **Configuration** > **Network Settings** > **High Availability Groups** to review the list of existing HA groups and to add the new nodes. | Administer StorageGRID |

3. If you added an Admin Node, complete the following configuration tasks.

| Admin Node configuration tasks | For information |
| --- | --- |
| If single sign-on is enabled for your StorageGRID system, you must create a relying party trust in Active Directory Federation Services (AD FS) for the new Admin Node. You cannot sign in to the node until you create this relying party trust. | Configuring single sign-on |
| If you plan to use the Load Balancer service on Admin Nodes, you might need to add the Admin Nodes to High Availability groups. Select **Configuration** > **Network Settings** > **High Availability Groups** to review the list of existing HA groups and to add the new nodes. | Administer StorageGRID |
| Optionally, copy the Admin Node database from the primary Admin Node to the expansion Admin Node if you want to keep the attribute and audit information consistent on each Admin Node. | Copying the Admin Node database |
| Optionally, copy the Prometheus database from the primary Admin Node to the expansion Admin Node if you want to keep the historical metrics consistent on each Admin Node. | Copying Prometheus metrics |
| Optionally, copy the existing audit logs from the primary Admin Node to the expansion Admin Node if you want to keep the historical log information consistent on each Admin Node. | Copying audit logs |
| Optionally, configure access to the system for auditing purposes through an NFS or a CIFS file share.<br><br>**Note:** Audit export through CIFS/Samba has been deprecated and will be removed in a future StorageGRID release. | Administer StorageGRID |
| Optionally, change the preferred sender for notifications. You can make the expansion Admin Node the preferred sender. Otherwise, an existing Admin Node configured as the preferred sender continues to send notifications, including AutoSupport messages, SNMP notifications, alert emails, and alarm emails (legacy system). | Administer StorageGRID |

4. If you added an Archive Node, complete the following configuration tasks.

| Archive Node configuration tasks | For information |
| --- | --- |
| Configure the Archive Node's connection to the targeted external archival storage system. When you complete the expansion, Archive Nodes are in an alarm state until you configure connection information through the **ARC** > **Target** component. | Administer StorageGRID |
| Update the ILM policy to archive object data through the new Archive Node. | Manage objects with ILM |

| Archive Node configuration tasks | For information |
| --- | --- |
| Configure custom alarms for the attributes that are used to monitor the speed and efficiency of object data retrieval from Archive Nodes. | Administer StorageGRID |

5. To check if expansion nodes were added with an untrusted Client Network or to change whether a node's Client Network is untrusted or trusted, go to **Configuration** > **Network Settings** > **Untrusted Client Network**.

   If the Client Network on the expansion node is untrusted, then connections to the node on the Client Network must be made using a load balancer endpoint. See the instructions for administering StorageGRID for more information.

6. Configure the Domain Name System (DNS).

   If you have been specifying DNS settings separately for each grid node, you must add custom per-node DNS settings for the new nodes. See information about modifying the DNS configuration for a single grid node in the recovery and maintenance instructions.

   The best practice is for the grid-wide DNS server list to contain some DNS servers that are accessible locally from each site. If you just added a new site, add new DNS servers for the site to the grid-wide DNS configuration.

   > Provide two to six IPv4 addresses for DNS servers. You should select DNS servers that each site can access locally in the event of network islanding. This is to ensure an islanded site continues to have access to the DNS service. After configuring the grid-wide DNS server list, you can further customize the DNS server list for each node. For details, see the information about modifying the DNS configuration in the recovery and maintenance instructions.

7. If you added a new site, confirm that Network Time Protocol (NTP) servers are accessible from that site.

   > Make sure that at least two nodes at each site can access at least four external NTP sources. If only one node at a site can reach the NTP sources, timing issues will occur if that node goes down. In addition, designating two nodes per site as primary NTP sources ensures accurate timing if a site is isolated from the rest of the grid.

   For more information, see the recovery and maintenance instructions.

**Related information**

Manage objects with ILM

Verifying that the Storage Node is active

Copying the Admin Node database

Copying Prometheus metrics

Copying audit logs

Upgrade software

Maintain & recover

### Verifying that the Storage Node is active

After an expansion operation that adds new Storage Nodes completes, the StorageGRID system should automatically start using the new Storage Nodes. You must use the StorageGRID system to verify that the new Storage Node is active.

#### Steps

1. Sign in to the Grid Manager using a supported browser.
2. Select **Nodes** > ***Expansion Storage Node*** > **Storage**.
3. Hover your cursor over the **Storage Used - Object Data** graph to view the value for **Used**, which is the amount of the Total usable space that has been used for object data.
4. Verify that the value of **Used** is increasing as you move your cursor to the right on the graph.

### Copying the Admin Node database

When adding Admin Nodes through an expansion procedure, you can optionally copy the database from the primary Admin Node to the new Admin Node. Copying the database allows you to retain historical information about attributes, alerts, and alerts.

#### What you'll need

- You must have completed the required expansion steps to add an Admin Node.
- You must have the `Passwords.txt` file.
- You must have the provisioning passphrase.

#### About this task

The StorageGRID software activation process creates an empty database for the NMS service on the expansion Admin Node. When the NMS service starts on the expansion Admin Node, it records information for servers and services that are currently part of the system or added later. This Admin Node database includes the following information:

- Alert history
- Alarm history
- Historical attribute data, which is used in the charts and text reports available from the **Support** > **Tools** > **Grid Topology** page

To ensure that the Admin Node database is consistent between nodes, you can copy the database from the primary Admin Node to the expansion Admin Node.

> Copying the database from the primary Admin Node (the*source Admin Node*) to an expansion Admin Node can take up to several hours to complete. During this period, the Grid Manager is inaccessible.

Use these steps to stop the MI service and the Management API service on both the primary Admin Node and the expansion Admin Node before copying the database.

#### Steps

1. Complete the following steps on the primary Admin Node:
   a. Log in to the Admin Node:

i. Enter the following command: `ssh admin@grid_node_IP`

ii. Enter the password listed in the `Passwords.txt` file.

iii. Enter the following command to switch to root: `su -`

iv. Enter the password listed in the `Passwords.txt` file.

b. Run the following command: `recover-access-points`

c. Enter the provisioning passphrase.

d. Stop the MI service: `service mi stop`

e. Stop the Management Application Program Interface (mgmt-api) service: `service mgmt-api stop`

2. Complete the following steps on the expansion Admin Node:

a. Log in to the expansion Admin Node:

i. Enter the following command: `ssh admin@grid_node_IP`

ii. Enter the password listed in the `Passwords.txt` file.

iii. Enter the following command to switch to root: `su -`

iv. Enter the password listed in the `Passwords.txt` file.

b. Stop the MI service: `service mi stop`

c. Stop the mgmt-api service: `service mgmt-api stop`

d. Add the SSH private key to the SSH agent. Enter:`ssh-add`

e. Enter the SSH Access Password listed in the `Passwords.txt` file.

f. Copy the database from the source Admin Node to the expansion Admin Node: `/usr/local/mi/bin/mi-clone-db.sh Source_Admin_Node_IP`

g. When prompted, confirm that you want to overwrite the MI database on the expansion Admin Node.

The database and its historical data are copied to the expansion Admin Node. When the copy operation is done, the script starts the expansion Admin Node.

h. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:`ssh-add -D`

3. Restart the services on the primary Admin Node: `service servermanager start`

### Copying Prometheus metrics

After adding a new Admin Node, you can optionally copy the historical metrics maintained by Prometheus from the primary Admin Node to the new Admin Node. Copying the metrics ensures that historical metrics are consistent between Admin Nodes.

#### What you'll need

- The new Admin Node must be installed and running.
- You must have the `Passwords.txt` file.
- You must have the provisioning passphrase.

#### About this task

When you add an Admin Node, the software installation process creates a new Prometheus database. You can keep the historical metrics consistent between nodes by copying the Prometheus database from the primary Admin Node (the *source Admin Node*) to the new Admin Node.

> Copying the Prometheus database might take an hour or more. Some Grid Manager features will be unavailable while services are stopped on the source Admin Node.

**Steps**

1. Log in to the source Admin Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.
2. From the source Admin Node, stop the Prometheus service: `service prometheus stop`
3. Complete the following steps on the new Admin Node:
   a. Log in to the new Admin Node:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.
   b. Stop the Prometheus service: `service prometheus stop`
   c. Add the SSH private key to the SSH agent. Enter:`ssh-add`
   d. Enter the SSH Access Password listed in the `Passwords.txt` file.
   e. Copy the Prometheus database from the source Admin Node to the new Admin Node: `/usr/local/prometheus/bin/prometheus-clone-db.sh Source_Admin_Node_IP`
   f. When prompted, press **Enter** to confirm that you want to destroy the new Prometheus database on the new Admin Node.

      The original Prometheus database and its historical data are copied to the new Admin Node. When the copy operation is done, the script starts the new Admin Node. The following status appears:

      ```
      Database cloned, starting services
      ```

   g. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:

      ```
      ssh-add -D
      ```

4. Restart the Prometheus service on the source Admin Node.

   ```
   service prometheus start
   ```

### Copying audit logs

When you add a new Admin Node through an expansion procedure, its AMS service only logs events and actions that occur after it joins the system. You can copy audit logs from a previously installed Admin Node to the new expansion Admin Node so that it is in sync with the rest of the StorageGRID system.

#### What you'll need

- You must have completed the required expansion steps to add an Admin Node.
- You must have the `Passwords.txt` file.

#### About this task

To make the historical audit messages from other Admin Nodes available on the expansion Admin Node, you must copy the audit log files manually from the primary Admin Node, or another existing Admin Node, to the expansion Admin Node.

#### Steps

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@_primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.
2. Stop the AMS service to prevent it from creating a new file: `service ams stop`
3. Rename the `audit.log` file to ensure that it does not overwrite the file on the expansion Admin Node you are copying it to:

   ```
   cd /var/local/audit/export
   ls -l
   mv audit.log new_name.txt
   ```

4. Copy all audit log files to the expansion Admin Node:

   ```
   scp -p * IP_address:/var/local/audit/export
   ```

5. If prompted for the passphrase for `/root/.ssh/id_rsa`, enter the SSH Access Password for the Primary Admin Node listed in the `Passwords.txt` file.
6. Restore the original `audit.log` file:

   ```
   mv new_name.txt audit.log
   ```

7. Start the AMS service:

   ```
   service ams start
   ```

8. Log out from the server:

```
exit
```

9. Log in to the expansion Admin Node:
   a. Enter the following command: `ssh admin@expansion_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

10. Update the user and group settings for the audit log files:

```
cd /var/local/audit/export
chown ams-user:bycast *
```

11. Log out from the server:

```
exit
```

### Rebalancing erasure-coded data after adding Storage Nodes

In some cases, you might need to rebalance erasure-coded data after you add new Storage Nodes.

#### What you'll need

- You must have completed the expansion steps to add the new Storage Nodes.
- You must have reviewed the considerations for rebalancing erasure-coded data.

  Considerations for rebalancing erasure-coded data

  > Only perform this procedure if the **Low Object Storage** alert has been triggered for one or more Storage Nodes at a site and you were unable to add the recommended number of new Storage Nodes.

- You must have the `Passwords.txt` file.

#### About this task

When the EC rebalance procedure is running, the performance of ILM operations and S3 and Swift client operations are likely to be impacted. For this reason, you should only perform this procedure in limited cases.

> The EC rebalance procedure temporarily reserves a large amount of storage. Storage alerts might be triggered, but will resolve when the rebalance is complete. If there is not enough storage for the reservation, the EC rebalance procedure will fail. Storage reservations are released when the EC rebalance procedure completes, whether the procedure failed or succeeded.

S3 and Swift API operations to upload objects (or object parts) might fail during the EC rebalancing procedure if they require more than 24 hours to complete. Long-duration PUT operations will fail if the applicable ILM rule uses Strict or Balanced placement on ingest. The following error will be reported:

```
500 Internal Server Error
```

**Steps**

1. Review the current object storage details for the site you plan to rebalance.
   a. Select **Nodes**.
   b. Select the first Storage Node at the site.
   c. Select the **Storage** tab.
   d. Hover your cursor over the Storage Used - Object Data chart to see the current amount of replicated data and erasure-coded data on the Storage Node.
   e. Repeat these steps to view the other Storage Nodes at the site.
2. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.
3. Enter the following command:

   ```
   rebalance-data start --site "site-name"
   ```

   For "*site-name*", specify the first site where you added new Storage Node or nodes. Enclose `site-name` in quotes.

   The EC rebalance procedure starts, and a job ID is returned.
4. Copy the job ID.
5. Monitor the status of the EC rebalance procedure.
   - To view the status of a single EC rebalance procedure:

     ```
     rebalance-data status --job-id job-id
     ```

     For *job-id*, specify the ID that was returned when you started the procedure.
   - To view the status of the current EC rebalance procedure and any previously completed procedures:

     ```
     rebalance-data status
     ```

To get help on the rebalance-data command:

```
rebalance-data --help
```

6. Perform additional steps, based on the status returned:
   - If the status is `In progress`, the EC rebalance operation is still running. You should periodically monitor the procedure until it completes.
   - If the status is `Failure`, perform the failure steps.
   - If the status is `Success`, perform the success step.
7. If the EC rebalance procedure is generating too much load (for example, ingest operations are affected), pause the procedure.

   ```
   rebalance-data pause --job-id job-id
   ```

8. If you need to terminate the EC rebalance procedure (for example, so you can perform a StorageGRID software upgrade), enter the following:

   ```
   rebalance-data abort --job-id job-id
   ```

   When you terminate an EC rebalance procedure, any data fragments that have already been moved remain in the new location. Data is not moved back to the original location.

9. If the status of the EC rebalance procedure is `Failure`, follow these steps:
   a. Confirm that all Storage Nodes at the site are connected to the grid.
   b. Check for and resolve any alerts that might be affecting these Storage Nodes.

      For information about specific alerts, see the monitoring and troubleshooting instructions.

   c. Restart the EC rebalance procedure:
      `rebalance-data start --job-id job-id`
   d. If the status of the EC rebalance procedure is still `Failure`, contact technical support.
10. If the status of the EC rebalance procedure is `Success`, optionally review object storage to see the updated details for the site.

    Erasure-coded data should now be more balanced among the Storage Nodes at the site.

    Replicated object data is not moved by the EC rebalance procedure.

11. If you are using erasure coding at more than one site, run this procedure for all other affected sites.

**Related information**

Considerations for rebalancing erasure-coded data

Monitor & troubleshoot

### Contacting technical support

If you encounter errors during the grid expansion process that you are unable to resolve, or if a grid task fails, contact technical support.

#### About this task

When you contact technical support, you must provide the required log files to assist in troubleshooting the errors you are encountering.

#### Steps

1. Connect to the expansion node that has experienced failures:

   a. Enter the following command:`ssh -p 8022 admin@grid_node_IP`

   > Port 8022 is the SSH port of the base OS, while port 22 is the SSH port of the Docker container running StorageGRID.

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   Once you are logged in as root, the prompt changes from `$` to `#`.

2. Depending on the stage the installation reached, retrieve any of the following logs that are available on the grid node:

| Platform | Logs |
|---|---|
| VMware | • `/var/log/daemon.log`<br>• `/var/log/storagegrid/daemon.log`<br>• `/var/log/storagegrid/nodes/<node-name>.log` |
| Linux | • `/var/log/storagegrid/daemon.log`<br>• `/etc/storagegrid/nodes/<node-name>.conf` (for each failed node)<br>• `/var/log/storagegrid/nodes/<node-name>.log` (for each failed node; might not exist) |

# Maintain & recover

Learn how to apply a hotfix; recover a failed grid node; decommission grid nodes and sites; and recover objects in the case of a system failure.

- Introduction to StorageGRID recovery and maintenance
- StorageGRID hotfix procedure

- Grid node recovery procedures
- How site recovery is performed by technical support
- Decommission procedure
- Network maintenance procedures
- Host-level and middleware procedures
- Grid node procedures
- Appliance node cloning

## Introduction to StorageGRID recovery and maintenance

The recovery and maintenance procedures for StorageGRID include applying a software hotfix, recovering grid nodes, recovering a failed site, decommissioning grid nodes or an entire site, performing network maintenance, performing host-level and middleware maintenance procedures, and performing grid node procedures.

All recovery and maintenance activities require a broad understanding of the StorageGRID system. You should review your StorageGRID system's topology to ensure that you understand the grid configuration.

You must follow all instructions exactly and heed all warnings.

Maintenance procedures not described are not supported or require a services engagement.

For hardware procedures, see the installation and maintenance instructions for your StorageGRID appliance.

"Linux" refers to a Red Hat® Enterprise Linux®, Ubuntu®, CentOS, or Debian® deployment. Use the NetApp Interoperability Matrix Tool to get a list of supported versions.

**Related information**

Grid primer

Network guidelines

Administer StorageGRID

SG100 & SG1000 services appliances

SG6000 storage appliances

SG5700 storage appliances

SG5600 storage appliances

NetApp Interoperability Matrix Tool

### Web browser requirements

You must use a supported web browser.

| Web browser | Minimum supported version |
|---|---|
| Google Chrome | 87 |
| Microsoft Edge | 87 |
| Mozilla Firefox | 84 |

You should set the browser window to a recommended width.

| Browser width | Pixels |
|---|---|
| Minimum | 1024 |
| Optimum | 1280 |

### Downloading the Recovery Package

The Recovery Package file allows you to restore the StorageGRID system if a failure occurs.

#### What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have the provisioning passphrase.
- You must have specific access permissions.

Download the current Recovery Package file before making grid topology changes to the StorageGRID system or before upgrading software. Then, download a new copy of the Recovery Package after making grid topology changes or after upgrading software.

#### Steps

1. Select **Maintenance** > **System** > **Recovery Package**.
2. Enter the provisioning passphrase, and select **Start Download**.

   The download starts immediately.

3. When the download completes:
   a. Open the `.zip` file.
   b. Confirm it includes a gpt-backup directory and an inner `.zip` file.
   c. Extract the inner `.zip` file.
   d. Confirm you can open the `Passwords.txt` file.
4. Copy the downloaded Recovery Package file (`.zip`) to two safe, secure, and separate locations.

> The Recovery Package file must be secured because it contains encryption keys and passwords that can be used to obtain data from the StorageGRID system.

**Related information**

Administer StorageGRID

## StorageGRID hotfix procedure

You might need to apply a hotfix to your StorageGRID system if issues with the software are detected and resolved between feature releases.

StorageGRID hotfixes contain software changes that are made available outside of a feature or patch release. The same changes are included in a future release. In addition, each hotfix release contains a roll-up of all previous hotfixes within the feature or patch release.

- Considerations for applying a hotfix
- How your system is affected when you apply a hotfix
- Obtaining the required materials for a hotfix
- Downloading the hotfix file
- Checking the system's condition before applying a hotfix
- Applying the hotfix

### Considerations for applying a hotfix

When you apply a hotfix, a cumulative series of software updates is applied to the nodes in your StorageGRID system.

You cannot apply a StorageGRID hotfix when another maintenance procedure is running. For example, you cannot apply a hotfix while a decommission, expansion, or recovery procedure is running.

> If a node or site decommission procedure is paused, you can safely apply a hotfix. In addition, you might be able to apply a hotfix during the final stages of a StorageGRID upgrade procedure. See the instructions for upgrading StorageGRID software for details.

After you upload the hotfix in the Grid Manager, the hotfix is applied automatically to the primary Admin Node. Then, you can approve the application of the hotfix to the rest of the nodes in your StorageGRID system.

If a hotfix fails to be applied to one or more nodes, the reason for the failure appears in the Details column of the hotfix progress table. You must resolve whatever issues caused the failures and then retry the entire process. Nodes with a previously successful application of the hotfix will be skipped in subsequent applications. You can safely retry the hotfix process as many times as required until all nodes have been updated. The hotfix must be successfully installed on all grid nodes in order for the application to be complete.

While grid nodes are updated with the new hotfix version, the actual changes in a hotfix might only affect specific services on specific types of nodes. For example, a hotfix might only affect the LDR service on Storage Nodes.

#### How hotfixes are applied for recovery and expansion

After a hotfix has been applied to your grid, the primary Admin Node automatically installs the same hotfix version to any nodes restored by recovery operations or added in an expansion.

However, if you need to recover the primary Admin Node, you must manually install the correct StorageGRID release and then apply the hotfix. The final StorageGRID version of the primary Admin Node must match the

version of the other nodes in the grid.

The following example illustrates how to apply a hotfix when recovering the primary Admin Node:

1. Assume the grid is running a StorageGRID 11.*A.B* version with the latest hotfix. The “grid version” is 11.*A.B.y*.
2. The primary Admin Node fails.
3. You redeploy the primary Admin Node using StorageGRID 11.*A.B*, and perform the recovery procedure.

   As required to match the grid version, you can use a minor release when deploying the node; you do not need to deploy the major release first.

4. You then apply hotfix 11.*A.B.y* to the primary Admin Node.

**Related information**

Configuring the replacement primary Admin Node

### How your system is affected when you apply a hotfix

You must understand how your StorageGRID system will be affected when you apply a hotfix.

#### Client applications might experience short-term disruptions

The StorageGRID system can ingest and retrieve data from client applications throughout the hotfix process; however, client connections to individual Gateway Nodes or Storage Nodes might be disrupted temporarily if the hotfix needs to restart services on those nodes. Connectivity will be restored after the hotfix process completes and services resume on the individual nodes.

You might need to schedule downtime to apply a hotfix if loss of connectivity for a short period is not acceptable. You can use selective approval to schedule when certain nodes are updated.

You can use multiple gateways and high availability (HA) groups to provide automatic failover during the hotfix process. To configure high availability groups, see the instructions for administering StorageGRID.

#### Alerts and SNMP notifications might be triggered

Alerts and SNMP notifications might be triggered when services are restarted and when the StorageGRID system is operating as a mixed-version environment (some grid nodes running an earlier version, while others have been upgraded to a later version). In general, these alerts and notifications will clear when the hotfix completes.

#### Configuration changes are restricted

When applying a hotfix to StorageGRID:

- Do not make any grid configuration changes (for example, specifying Grid Network subnets or approving pending grid nodes) until the hotfix has been applied to all nodes.
- Do not update the ILM configuration until the hotfix has been applied to all nodes.

### Obtaining the required materials for a hotfix

Before applying a hotfix, you must obtain all required materials.

| Item | Notes |
|---|---|
| StorageGRID hotfix file | You must download the StorageGRID hotfix file. |
| • Network port<br>• Supported web browser<br>• SSH client (for example, PuTTY) | See “Web browser requirements.” |
| Recovery Package (`.zip`) file | Before applying a hotfix, download the most recent Recovery Package file in case any problems occur during the hotfix.Then, after the hotfix has been applied, download a new copy of the Recovery Package file and save it in a safe location. The updated Recovery Package file allows you to restore the system if a failure occurs. |
| `Passwords.txt file` | Optional and used only if you are applying a hotfix manually using the SSH client. The `Passwords.txt` file is included in the SAID package, which is part of the Recovery Package `.zip` file. |
| Provisioning passphrase | The passphrase is created and documented when the StorageGRID system is first installed. The provisioning passphrase is not listed in the `Passwords.txt` file. |
| Related documentation | `readme.txt` file for the hotfix. This file is included on the hotfix download page. Be sure to review the `readme` file carefully before applying the hotfix. |

**Related information**

Downloading the hotfix file

Downloading the Recovery Package

### Downloading the hotfix file

You must download the hotfix file before you can apply the hotfix.

**Steps**

1. Go to the NetApp Downloads page for StorageGRID.

   NetApp Downloads: StorageGRID

2. Select the down arrow under **Available Software** to see a list of hotfixes that are available to download.

   Hotfix file versions have the form: 11.4.*x.y*.

3. Review the changes that are included in the update.

   If you have just recovered the primary Admin Node and you need to apply a hotfix, select the same hotfix version that is installed on the other grid nodes.

   a. Select the hotfix version you want to download, and select **Go**.
   b. Sign in using the username and password for your NetApp account.
   c. Read and accept the End User License Agreement.

      The download page for the version you selected appears.

   d. Download the hotfix `readme.txt` file to view a summary of the changes included in the hotfix.

4. Select the download button for the hotfix, and save the file.

   Do not change the name of this file.

   If you are using a macOS device, the hotfix file might be automatically saved as a `.txt` file. If it is, you must rename the file without the `.txt` extension.

5. Select a location for the download, and select **Save**.

**Related information**

Configuring the replacement primary Admin Node

### Checking the system's condition before applying a hotfix

You must verify the system is ready to accommodate the hotfix.

1. Sign in to the Grid Manager using a supported browser.
2. If possible, ensure that the system is running normally and that all grid nodes are connected to the grid.

   Connected nodes have green check marks ✔ on the Nodes page.

3. Check for and resolve any current alerts, if possible.

   For information about specific alerts, see the instructions for monitoring and troubleshooting StorageGRID.

4. Ensure no other maintenance procedures are in progress, such as an upgrade, recovery, expansion, or decommission procedure.

   You should wait for any active maintenance procedures to complete before applying a hotfix.

   You cannot apply a StorageGRID hotfix when another maintenance procedure is running. For example, you cannot apply a hotfix while a decommission, expansion, or recovery procedure is running.

   If a node or site decommission procedure is paused, you can safely apply a hotfix. In addition, you might be able to apply a hotfix during the final stages of a StorageGRID upgrade procedure. See the instructions for upgrading StorageGRID software for details.

**Related information**

### Applying the hotfix

The hotfix is first applied automatically to the primary Admin Node. Then, you must approve the application of the hotfix to other grid nodes until all nodes are running the same software version. You can customize the approval sequence by selecting to approve individual grid nodes, groups of grid nodes, or all grid nodes.

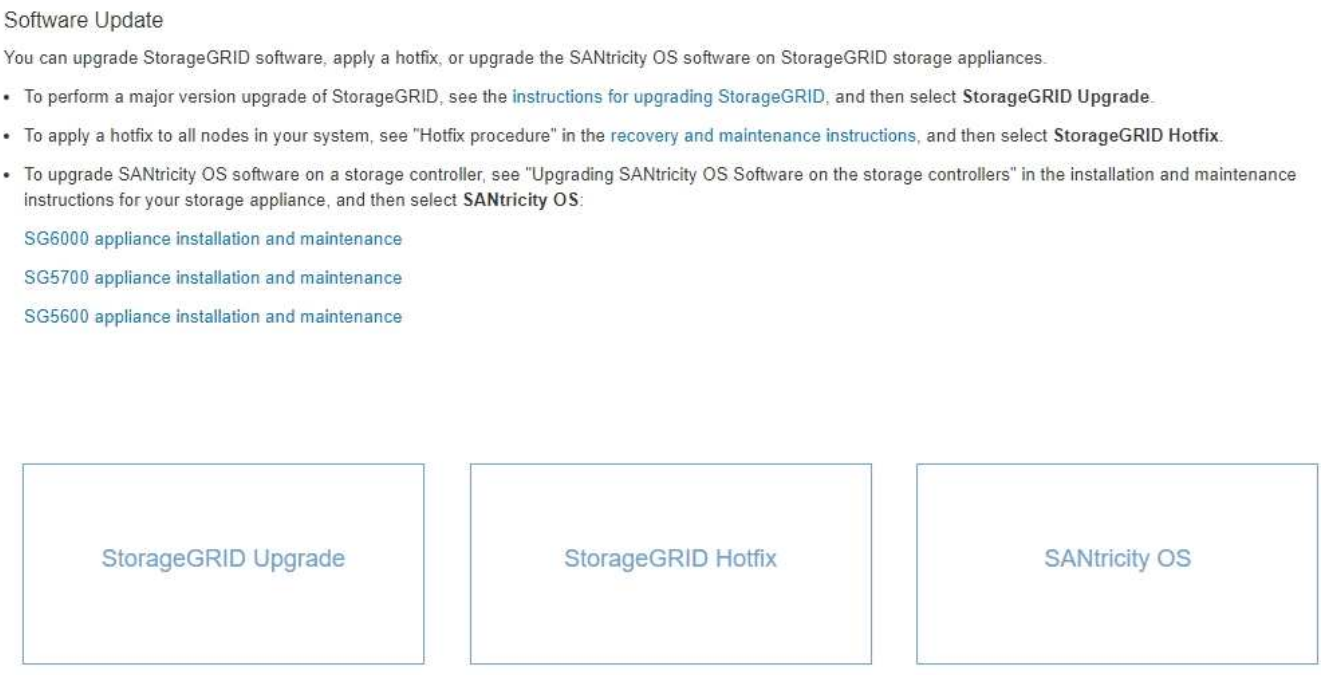
**What you'll need**

- You have reviewed all of the considerations and completed all of the steps in "Hotfix planning and preparation."
- You must have the provisioning passphrase.
- You must have Root Access or the Maintenance permission.
- You can delay applying a hotfix to a node, but the hotfix process is not complete until you apply the hotfix to all nodes.
- You cannot perform a StorageGRID software upgrade or a SANtricity OS upgrade until you have completed the hotfix process.

**Steps**

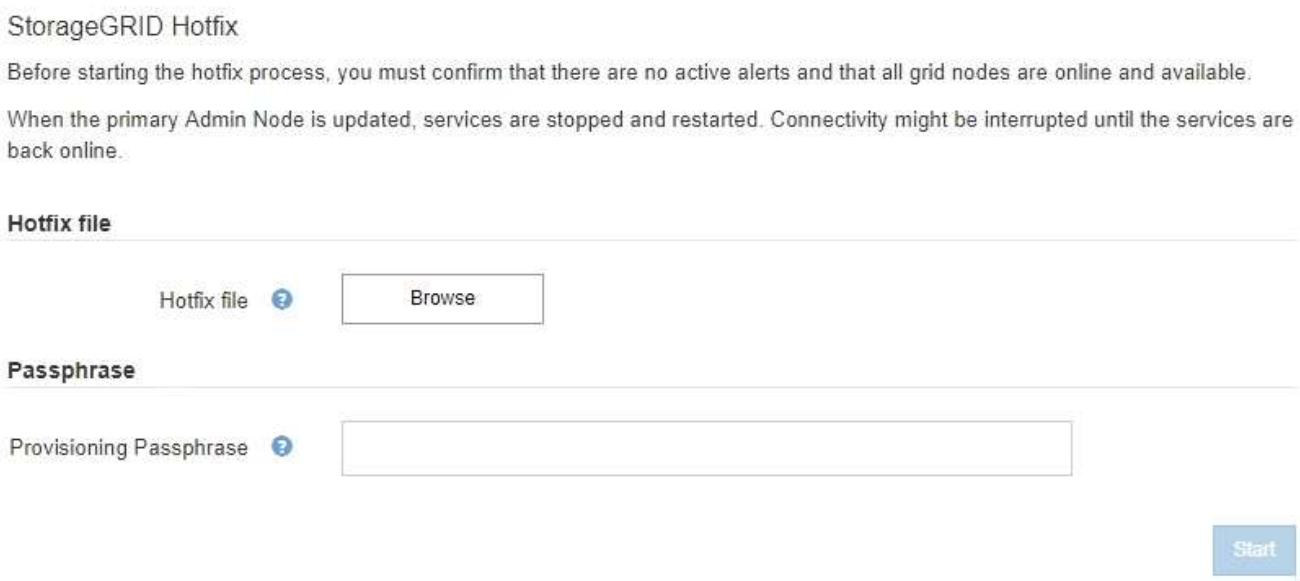
1. Sign in to the Grid Manager using a supported browser.
2. Select **Maintenance** > **System** > **Software Update**.

   The Software Update page appears.




3. Select **StorageGRID Hotfix**.

   The StorageGRID Hotfix page appears.



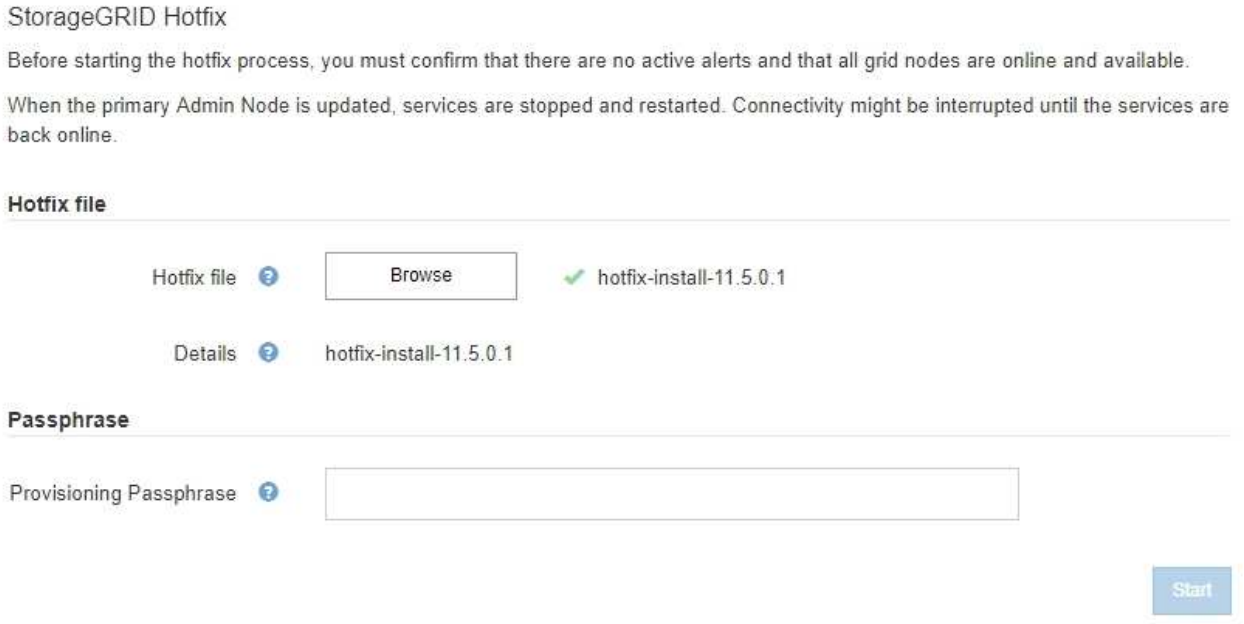

4. Select the hotfix file you downloaded from the NetApp support site.
   a. Select **Browse**.
   b. Locate and select the file.

      `hotfix-install-version`

   c. Select **Open**.

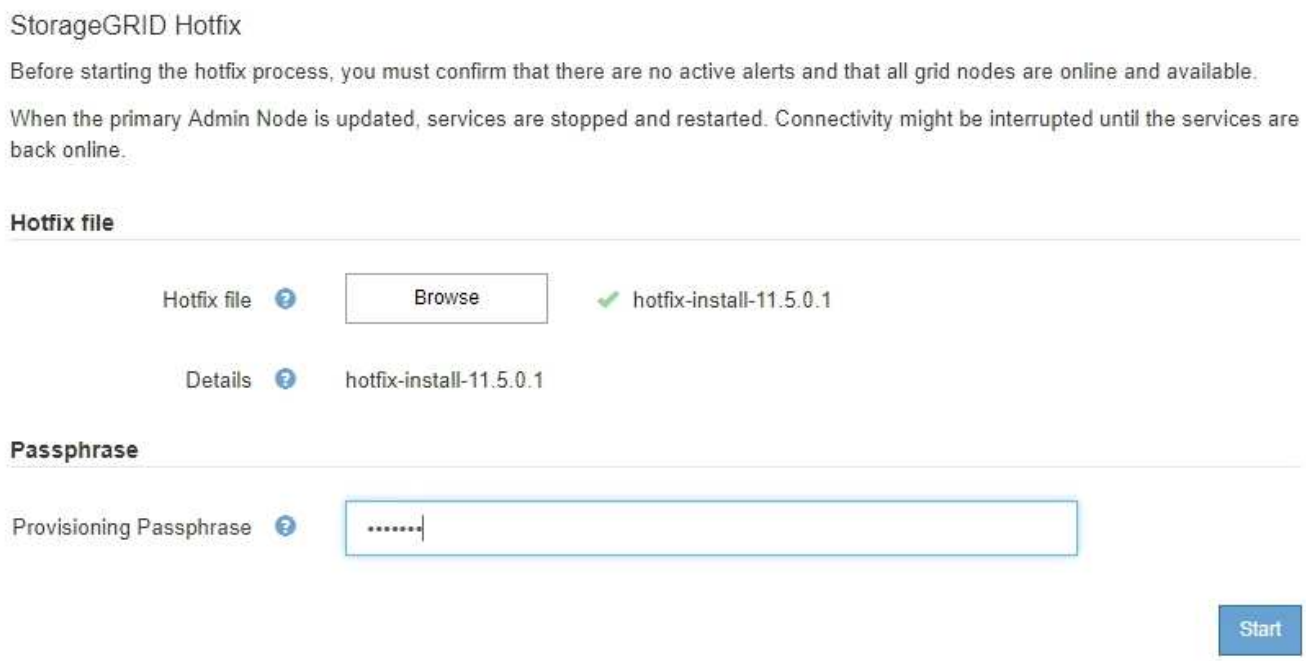
      The file is uploaded. When the upload is finished, the file name is shown in the Details field.

      Do not change the file name since it is part of the verification process.



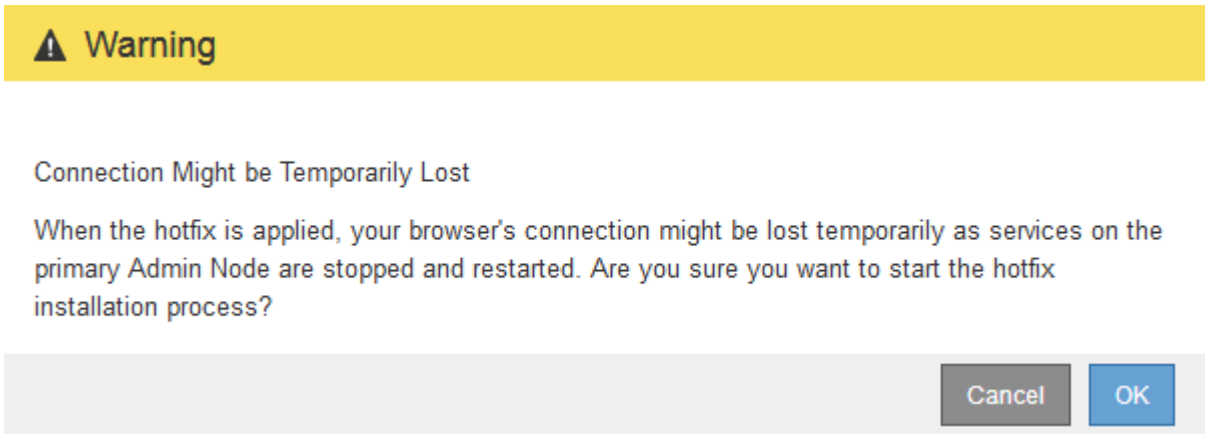

5. Enter the provisioning passphrase in the text box.

   The **Start** button becomes enabled.




6. Select **Start**.

   A warning appears stating that your browser's connection might be lost temporarily as services on the primary Admin Node are restarted.




7. Select **OK** to start applying the hotfix to the primary Admin Node.

   When the hotfix starts:

   a. The hotfix validations are run.

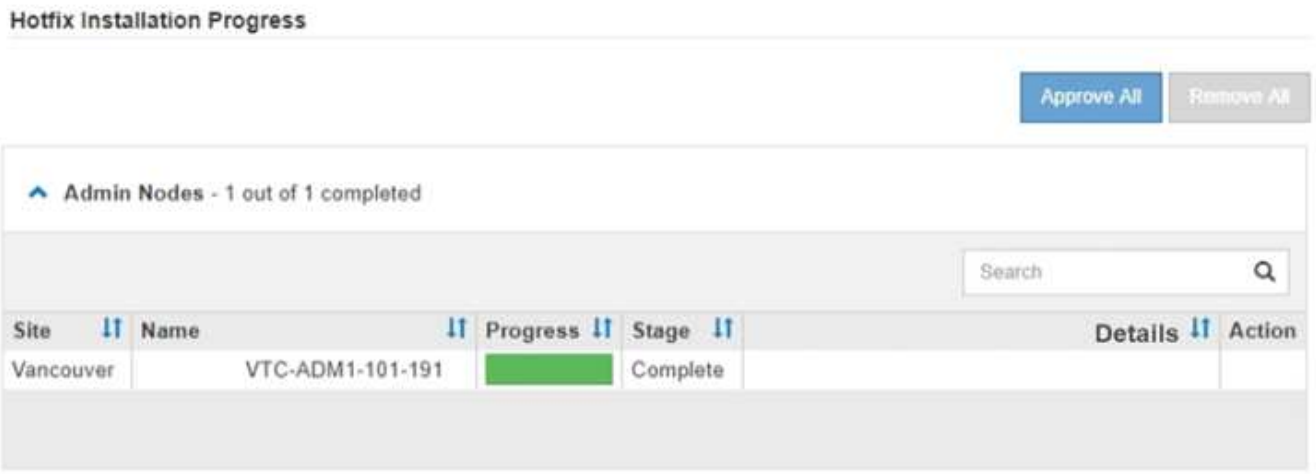
      If any errors are reported, resolve them, re-upload the hotfix file, and select **Start** again.

   b. The hotfix installation progress table appears. This table shows all nodes in your grid and the current stage of the hotfix installation for each node. The nodes in the table are grouped by type:

      - Admin Nodes
      - Gateway Nodes
      - Storage Nodes

- Archive Nodes

> The progress bar reaches completion, and then the primary Admin Node is shown first with stage “Complete.”




8. Optionally, sort the lists of nodes in each grouping in ascending or descending order by **Site**, **Name**, **Progress**, **Stage**, or **Details**. Or, enter a term in the **Search** box to search for specific nodes.
9. Approve the grid nodes that are ready to be updated. Approved nodes of the same type are upgraded one at a time.

> Do not approve the hotfix for a node unless you are sure the node is ready to be updated.When the hotfix is applied to a grid node, some services on that node might be restarted. These operations might cause service interruptions for clients that are communicating with the node.

   - Select one or more **Approve** buttons to add one or more individual nodes to the hotfix queue.
   - Select the **Approve All** button within each grouping to add all nodes of the same type to the hotfix queue. If you have entered search criteria in the **Search** box, the **Approve All** button applies to all the nodes selected by the search criteria.

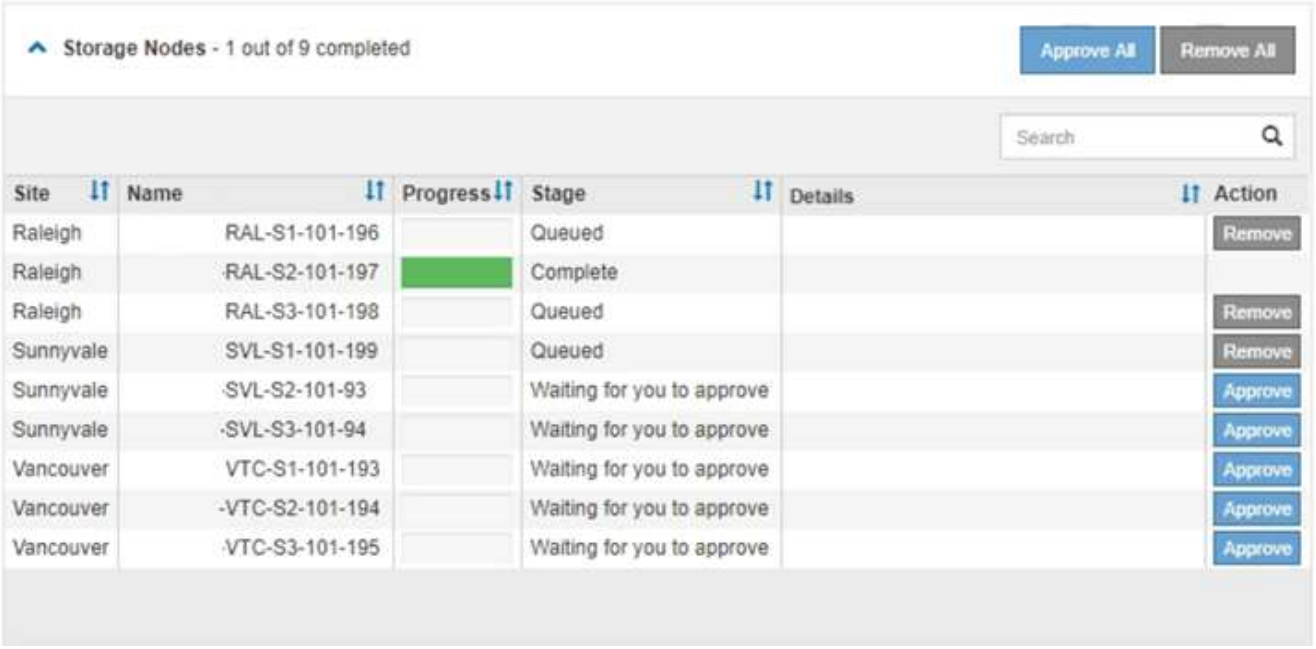
> The **Approve All** button at the top of the page approves all nodes listed on the page, while the **Approve All** button at the top of a table grouping only approves all nodes in that group. If the order in which nodes are upgraded is important, approve nodes or groups of nodes one at a time and wait until the upgrade is complete on each node before approving the next node(s).

   - Select the top-level **Approve All** button at the top of the page to add all nodes in the grid to the hotfix queue.

> You must complete the StorageGRID hotfix before you can start a different software update. If you are unable to complete the hotfix, contact technical support.

10. If you need to remove a node or all nodes from the hotfix queue, select **Remove** or **Remove All**.

    As shown in the example, when the Stage progresses beyond “Queued,” the **Remove** button is hidden and you can no longer remove the node from the hotfix process.




11. Wait while the hotfix is applied to each approved grid node.

    When the hotfix has been successfully installed on all nodes, the Hotfix Installation Progress table closes. A green banner shows the date and time the hotfix was completed.

12. If the hotfix could not be applied to any nodes, review the error for each node, resolve the issue, and repeat these steps.

    The procedure is not complete until the hotfix is successfully applied to all nodes. You can safely retry the hotfix process as many times as required until it is complete.

**Related information**

Hotfix planning and preparation

Administer StorageGRID

Monitor & troubleshoot

## Grid node recovery procedures

If a grid node fails, you can recover it by replacing the failed physical or virtual server, reinstalling StorageGRID software, and restoring recoverable data.

Grid nodes can fail if a hardware, virtualization, operating system, or software fault renders the node inoperable or unreliable. There are many kinds of failure that can trigger the need to recover a grid node.

The steps to recover a grid node vary, depending on the platform where the grid node is hosted and on the type of grid node. Each type of grid node has a specific recovery procedure, which you must follow exactly.

Generally, you try to preserve data from the failed grid node where possible, repair or replace the failed node, use the Grid Manager to configure the replacement node, and restore the node's data.

If an entire StorageGRID site has failed, contact technical support. Technical support will work with you to develop and execute a site recovery plan that maximizes the amount of data that is recovered, and meets your business objectives.

**Related information**

How site recovery is performed by technical support

### Warnings and considerations for grid node recovery

If a grid node fails, you must recover it as soon as possible. You must review all warnings and considerations for node recovery before you begin.

StorageGRID is a distributed system composed of multiple nodes working with each other. Do not use disk snapshots to restore grid nodes. Instead, refer to the recovery and maintenance procedures for each type of node.

Some of the reasons for recovering a failed grid node as soon as possible include the following:

- A failed grid node can reduce the redundancy of system and object data, leaving you vulnerable to the risk of permanent data loss if another node fails.
- A failed grid node can impact the efficiency of day-to-day operations.
- A failed grid node can reduce your ability to monitor system operations.
- A failed grid node can cause a 500 internal server error if strict ILM rules are in place.
- If a grid node is not recovered promptly, recovery times might increase. For example, queues might develop that need to be cleared before recovery is complete.

Always follow the recovery procedure for the specific type of grid node you are recovering. Recovery procedures vary for primary or non-primary Admin Nodes, Gateway Nodes, Archive Nodes, appliance nodes, and Storage Nodes.

#### Preconditions for recovering grid nodes

All of the following conditions are assumed when recovering grid nodes:

- The failed physical or virtual hardware has been replaced and configured.
- The StorageGRID Appliance Installer version on the replacement appliance matches the software version of your StorageGRID system, as described in hardware installation and maintenance for verifying and upgrading the StorageGRID Appliance Installer version.
  - SG100 & SG1000 services appliances
  - SG5600 storage appliances
  - SG5700 storage appliances
  - SG6000 storage appliances
- If you are recovering a grid node other than the primary Admin Node, there is connectivity between the grid node being recovered and the primary Admin Node.

#### Order of node recovery if a server hosting more than one grid node fails

If a server that is hosting more than one grid node fails, you can recover the nodes in any order. However, if

the failed server is hosting the primary Admin Node, you must recover that node first. Recovering the primary Admin Node first prevents other node recoveries from halting as they wait to contact the primary Admin Node.

#### IP addresses for recovered nodes

Do not attempt to recover a node using an IP address that is currently assigned to any other node. When you deploy the new node, use the failed node's current IP address or an unused IP address.

### Gathering required materials for grid node recovery

Before performing maintenance procedures, you must ensure you have the necessary materials to recover a failed grid node.

| Item | Notes |
|---|---|
| StorageGRID installation archive | If you need to recover a grid node, you need the StorageGRID installation archive for your platform.<br><br>**Note:** You do not need to download files if you are recovering failed storage volumes on a Storage Node. |
| Recovery Package `.zip` file | Obtain a copy of the most recent Recovery Package `.zip` file: `sgws-recovery-package-`*`id-revision`*`.zip`<br><br>The contents of the `.zip` file are updated each time the system is modified. You are directed to store the most recent version of the Recovery Package in a secure location after making such changes. Use the most recent copy to recover from grid failures.<br><br>If the primary Admin Node is operating normally, you can download the Recovery Package from the Grid Manager. Select **Maintenance** > **System** > **Recovery Package**.<br><br>If you cannot access the Grid Manager, you can find encrypted copies of the Recovery Package on some Storage Nodes that contain the ADC service. On each Storage Node, examine this location for the Recovery Package: `/var/local/install/sgws-recovery-package-`*`grid-id-revision`*`.zip.gpg` Use the Recovery Package with the highest revision number. |
| `Passwords.txt` file | Contains the passwords required to access grid nodes on the command line. Included in the Recovery Package. |
| Provisioning passphrase | The passphrase is created and documented when the StorageGRID system is first installed. The provisioning passphrase is not in the `Passwords.txt` file. |

| Item | Notes |
| --- | --- |
| Current documentation for your platform | For the current supported versions of your platform, see the Interoperability Matrix Tool.<br><br>NetApp Interoperability Matrix Tool<br><br>Go to the platform vendor's website for documentation. |

**Related information**

Downloading and extracting the StorageGRID installation files

Web browser requirements

#### Downloading and extracting the StorageGRID installation files

## Before you can recover StorageGRID grid nodes, you must download the software and extract the files.

You must use the version of StorageGRID that is currently running on the grid.

**Steps**

1. Determine which version of the software is currently installed. From the Grid Manager, go to **Help** > **About**.
2. Go to the NetApp Downloads page for StorageGRID.

   NetApp Downloads: StorageGRID

3. Select the version of StorageGRID that is currently running on the grid.

   StorageGRID software versions have this format: 11.x.y.

4. Sign in with the username and password for your NetApp account.
5. Read the End User License Agreement, select the check box, and then select **Accept & Continue**.
6. In the **Install StorageGRID** column of the download page, select the `.tgz` or `.zip` file for your platform.

   The version shown in the installation archive file must match the version of the software that is currently installed.

   Use the `.zip` file if you are running Windows.

   | Platform | Installation archive |
   | --- | --- |
   | VMware | `StorageGRID-Webscale-`*`version`*`-VMware-`*`uniqueID`*`.zip`<br><br>`StorageGRID-Webscale-`*`version`*`-VMware-`*`uniqueID`*`.tgz` |
   | Red Hat Enterprise Linux or CentOS | `StorageGRID-Webscale-`*`version`*`-RPM-`*`uniqueID`*`.zip`<br><br>`StorageGRID-Webscale-`*`version`*`-RPM-`*`uniqueID`*`.tgz` |

| Platform | Installation archive |
| --- | --- |
| Ubuntu or Debian<br><br>or Appliances | `StorageGRID-Webscale-version-DEB-uniqueID.zip`<br><br>`StorageGRID-Webscale-version-DEB-uniqueID.tgz` |
| OpenStack or other hypervisor | NetApp-provided virtual machine disk files and scripts for OpenStack are no longer supported for recovery operations. If you need to recover a node running in an OpenStack deployment, download the files for your Linux operating system. Then, follow the procedure for replacing a Linux node. |

7. Download and extract the archive file.
8. Follow the appropriate step for your platform to choose the files you need, based on your platform and which grid nodes you need to recover.

   The paths listed in the step for each platform are relative to the top-level directory installed by the archive file.

9. If you are recovering a VMware system, select the appropriate files.

| Path and file name | Description |
| --- | --- |
| `./vsphere/README` | A text file that describes all of the files contained in the StorageGRID download file. |
| `./vsphere/NLF000000.txt` | A free license that does not provide any support entitlement for the product. |
| `./vsphere/NetApp-SG-version-SHA.vmdk` | The virtual machine disk file that is used as a template for creating grid node virtual machines. |
| `./vsphere/vsphere-primary-admin.ovf`<br><br>`./vsphere/vsphere-primary-admin.mf` | The Open Virtualization Format template file (`.ovf`) and manifest file (`.mf`) for deploying the primary Admin Node. |
| `./vsphere/vsphere-non-primary-admin.ovf`<br><br>`./vsphere/vsphere-non-primary-admin.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying non-primary Admin Nodes. |
| `/vsphere/vsphere-archive.ovf`<br><br>`./vsphere/vsphere-archive.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying Archive Nodes. |
| `./vsphere/vsphere-gateway.ovf`<br><br>`./vsphere/vsphere-gateway.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying Gateway Nodes. |
| `./vsphere/vsphere-storage.ovf`<br><br>`./vsphere/vsphere-storage.mf` | The template file (`.ovf`) and manifest file (`.mf`) for deploying virtual machine-based Storage Nodes. |

| Path and file name | Description |
| --- | --- |
| **Deployment scripting tool** | **Description** |
| `./vsphere/deploy-vsphere-ovftool.sh` | A Bash shell script used to automate the deployment of virtual grid nodes. |
| `./vsphere/deploy-vsphere-ovftool-sample.ini` | A sample configuration file for use with the `deploy-vsphere-ovftool.sh` script. |
| `./vsphere/configure-storagegrid.py` | A Python script used to automate the configuration of a StorageGRID system. |
| `./vsphere/configure-sga.py` | A Python script used to automate the configuration of StorageGRID appliances. |
| `./vsphere/storagegrid-ssoauth.py` | An example Python script that you can use to sign in to the Grid Management API when single sign-on is enabled. |
| `./vsphere/configure-storagegrid.sample.json` | A sample configuration file for use with the `configure-storagegrid.py` script. |
| `./vsphere/configure-storagegrid.blank.json` | A blank configuration file for use with the `configure-storagegrid.py` script. |

10. If you are recovering a Red Hat Enterprise Linux or CentOS system, select the appropriate files.

| Path and file name | Description |
| --- | --- |
| `./rpms/README` | A text file that describes all of the files contained in the StorageGRID download file. |
| `./rpms/NLF000000.txt` | A free license that does not provide any support entitlement for the product. |
| `./rpms/StorageGRID-Webscale-Images-`*`version`*`-SHA.rpm` | RPM package for installing the StorageGRID node images on your RHEL or CentOS hosts. |
| `./rpms/StorageGRID-Webscale-Service-`*`version`*`-SHA.rpm` | RPM package for installing the StorageGRID host service on your RHEL or CentOS hosts. |
| **Deployment scripting tool** | **Description** |
| `./rpms/configure-storagegrid.py` | A Python script used to automate the configuration of a StorageGRID system. |
| `./rpms/configure-sga.py` | A Python script used to automate the configuration of StorageGRID appliances. |

| Path and file name | Description |
| --- | --- |
| `./rpms/configure-storagegrid.sample.json` | A sample configuration file for use with the `configure-storagegrid.py` script. |
| `./rpms/storagegrid-ssoauth.py` | An example Python script that you can use to sign in to the Grid Management API when single sign-on is enabled. |
| `./rpms/configure-storagegrid.blank.json` | A blank configuration file for use with the `configure-storagegrid.py` script. |
| `./rpms/extras/ansible` | Example Ansible role and playbook for configuring RHEL or CentOS hosts for StorageGRID container deployment. You can customize the role or playbook as necessary. |

11. If you are recovering an Ubuntu or Debian system, select the appropriate files.

| Path and file name | Description |
| --- | --- |
| `./debs/README` | A text file that describes all of the files contained in the StorageGRID download file. |
| `./debs/NLF000000.txt` | A non-production NetApp License File that you can use for testing and proof of concept deployments. |
| `./debs/storagegrid-webscale-images-`*`version`*`-SHA.deb` | DEB package for installing the StorageGRID node images on Ubuntu or Debian hosts. |
| `./debs/storagegrid-webscale-images-`*`version`*`-SHA.deb.md5` | MD5 checksum for the file `/debs/storagegrid-webscale-images-version-SHA.deb` |
| `./debs/storagegrid-webscale-service-`*`version`*`-SHA.deb` | DEB package for installing the StorageGRID host service on Ubuntu or Debian hosts. |
| **Deployment scripting tool** | **Description** |
| `./debs/configure-storagegrid.py` | A Python script used to automate the configuration of a StorageGRID system. |
| `./debs/configure-sga.py` | A Python script used to automate the configuration of StorageGRID appliances. |
| `./debs/storagegrid-ssoauth.py` | An example Python script that you can use to sign in to the Grid Management API when single sign-on is enabled. |

| Path and file name | Description |
| --- | --- |
| `./debs/configure-storagegrid.sample.json` | A sample configuration file for use with the `configure-storagegrid.py` script. |
| `./debs/configure-storagegrid.blank.json` | A blank configuration file for use with the `configure-storagegrid.py` script. |
| `./debs/extras/ansible` | Example Ansible role and playbook for configuring Ubuntu or Debian hosts for StorageGRID container deployment. You can customize the role or playbook as necessary. |

12. If you are recovering a StorageGRID appliance-based system, select the appropriate files.

| Path and file name | Description |
| --- | --- |
| `./debs/storagegrid-webscale-images-`*`version`*`-SHA.deb` | DEB package for installing the StorageGRID node images on your appliances. |
| `./debs/storagegrid-webscale-images-`*`version`*`-SHA.deb.md5` | Checksum of the DEB installation package used by the StorageGRID Appliance Installer to validate that the package is intact after upload. |

**Note:** For appliance installation, these files are only required if you need to avoid network traffic. The appliance can download the required files from the primary Admin Node.

**Related information**

Install VMware

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

### Selecting a node recovery procedure

You must select the correct recovery procedure for the type of node that has failed.

| Grid node | Recovery procedure |
| --- | --- |
| More than one Storage Node | Contact technical support. If more than one Storage Node has failed, technical support must assist with recovery to prevent database inconsistencies that could lead to data loss. A site recovery procedure might be required.<br><br>How site recovery is performed by technical support |

| Grid node | Recovery procedure |
| --- | --- |
| A single Storage Node | The Storage Node recovery procedure depends on the type and duration of the failure.<br><br>Recovering from Storage Node failures |
| Admin Node | The Admin Node procedure depends on whether you need to recover the primary Admin Node or a non-primary Admin Node.<br><br>Recovering from Admin Node failures |
| Gateway Node | Recovering from Gateway Node failures. |
| Archive Node | Recovering from Archive Node failures. |

If a server that is hosting more than one grid node fails, you can recover the nodes in any order. However, if the failed server is hosting the primary Admin Node, you must recover that node first. Recovering the primary Admin Node first prevents other node recoveries from halting as they wait to contact the primary Admin Node.

#### Recovering from Storage Node failures

## The procedure for recovering a failed Storage Node depends on the type of failure and the type of Storage Node that has failed.

Use this table to select the recovery procedure for a failed Storage Node.

| Issue | Action | Notes |
| --- | --- | --- |
| • More than one Storage Node has failed.<br>• A second Storage Node has failed less than 15 days after a Storage Node failure or recovery.<br><br>This includes the case where a Storage Node fails while recovery of another Storage Node is still in progress. | You must contact technical support. | If all failed Storage Nodes are at the same site, it might be necessary to perform a site recovery procedure.<br><br>Technical support will assess your situation and develop a recovery plan.<br><br>How site recovery is performed by technical support<br><br>Recovering more than one Storage Node (or more than one Storage Node within 15 days) might affect the integrity of the Cassandra database, which can cause data loss.<br><br>Technical support can determine when it is safe to begin recovery of a second Storage Node.<br><br>**Note**: If more than one Storage Node that contains the ADC service fails at a site, you lose any pending platform service requests for that site. |
| A Storage Node has been offline for more than 15 days. | Recovering a Storage Node that has been down more than 15 days | This procedure is required to ensure Cassandra database integrity. |
| An appliance Storage Node has failed. | Recovering a StorageGRID appliance Storage Node | The recovery procedure for appliance Storage Nodes is the same for all failures. |
| One or more storage volumes have failed, but the system drive is intact | Recovering from storage volume failure where the system drive is intact | This procedure is used for software-based Storage Nodes. |
| The system drive has failed. | Recovering from system drive failure | The node replacement procedure depends on the deployment platform and on whether any storage volumes have also failed. |

Some StorageGRID recovery procedures use Reaper to handle Cassandra repairs. Repairs occur automatically as soon as the related or required services have started. You might notice script output that mentions “reaper” or “Cassandra repair.” If you see an error message indicating the repair has failed, run the command indicated in the error message.

#### Recovering a Storage Node that has been down more than 15 days

# If a single Storage Node has been offline and not connected to other Storage Nodes for more than 15 days, you must rebuild Cassandra on the node.

**What you'll need**

- You have checked that a Storage Node decommissioning is not in progress, or you have paused the node decommission procedure. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Decommission**.)
- You have checked that an expansion is not in progress. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Expansion**.)

**About this task**

Storage Nodes have a Cassandra database that includes object metadata. If a Storage Node has not been able to communicate with other Storage Nodes for more than 15 days, StorageGRID assumes that node's Cassandra database is stale. The Storage Node cannot rejoin the grid until Cassandra has been rebuilt using information from other Storage Nodes.

Use this procedure to rebuild Cassandra only if a single Storage Node is down. Contact technical support if additional Storage Nodes are offline or if Cassandra has been rebuilt on another Storage Node within the last 15 days; for example, Cassandra might have been rebuilt as part of the procedures to recover failed storage volumes or to recover a failed Storage Node.

If more than one Storage Node has failed (or is offline), contact technical support. Do not perform the following recovery procedure. Data loss could occur.

If this is the second Storage Node failure in less than 15 days after a Storage Node failure or recovery, contact technical support. Do not perform the following recovery procedure. Data loss could occur.

If more than one Storage Node at a site has failed, a site recovery procedure might be required. Contact technical support.

How site recovery is performed by technical support

**Steps**

1. If necessary, power on the Storage Node that needs to be recovered.
2. Log in to the grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.+

If you are unable to log in to the grid node, the system disk might not be intact. Go to the procedure for recovering from system drive failure. Recovering from system drive failure

3. Perform the following checks on the Storage Node:
   a. Issue this command: `nodetool status`

      The output should be `Connection refused`

   b. In the Grid Manager, select **Support** > **Tools** > **Grid Topology**.
   c. Select *site* > **Storage Node** > **SSM** > **Services**. Verify that the Cassandra service displays `Not Running`.
   d. Select **Storage Node** > **SSM** > **Resources**. Verify that there is no error status in the Volumes section.
   e. Issue this command: `grep -i Cassandra /var/local/log/servermanager.log`

      You should see the following message in the output:

      ```
      Cassandra not started because it has been offline for more than 15
      day grace period - rebuild Cassandra
      ```

4. Issue this command, and monitor the script output: `check-cassandra-rebuild`
   - If storage services are running, you will be prompted to stop them. Enter: **y**
   - Review the warnings in the script. If none of them apply, confirm that you want to rebuild Cassandra. Enter: **y**

   > Some StorageGRID recovery procedures use Reaper to handle Cassandra repairs. Repairs occur automatically as soon as the related or required services have started. You might notice script output that mentions “reaper” or “Cassandra repair.” If you see an error message indicating the repair has failed, run the command indicated in the error message.

5. After the rebuild completes, perform the following checks:
   a. In the Grid Manager, select **Support** > **Tools** > **Grid Topology**.
   b. Select *site* > **recovered Storage Node** > **SSM** > **Services**.
   c. Confirm that all services are running.
   d. Select **DDS** > **Data Store**.
   e. Confirm that the **Data Store Status** is “Up” and the **Data Store State** is “Normal.”

**Related information**

Recovering from system drive failure

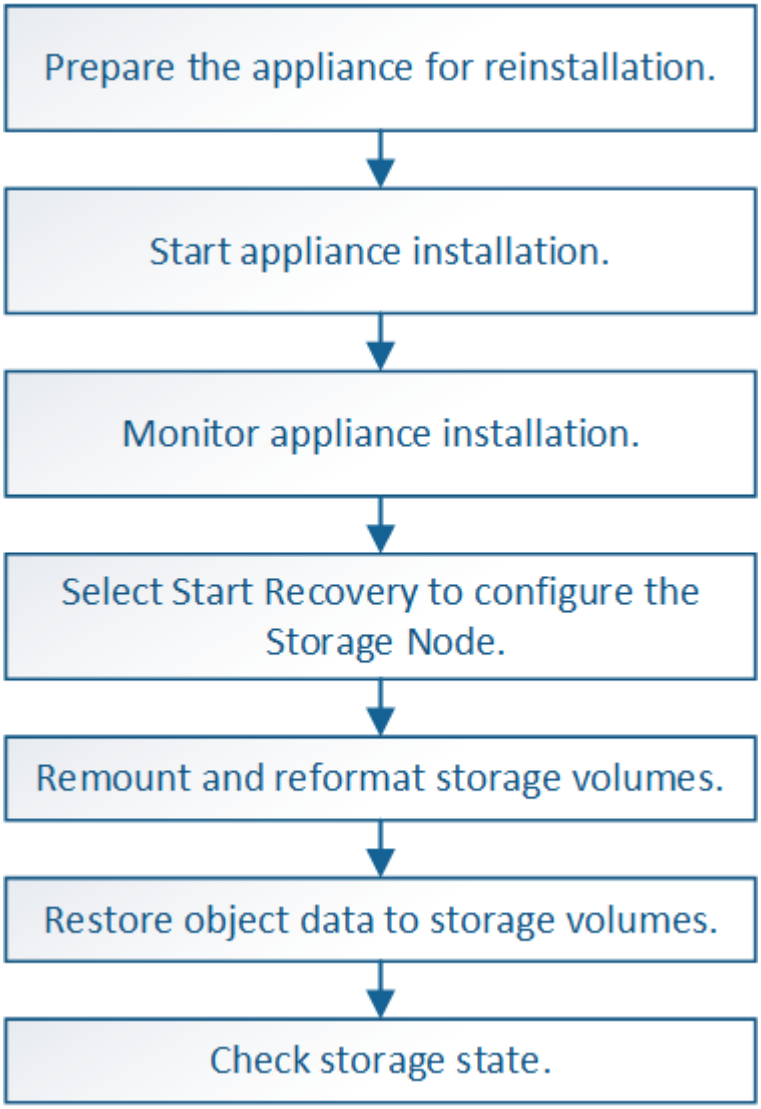
#### Recovering a StorageGRID appliance Storage Node

## The procedure for recovering a failed StorageGRID appliance Storage Node is the same whether you are recovering from the loss of the system drive or from the loss of storage volumes only.

**About this task**

You must prepare the appliance and reinstall software, configure the node to rejoin the grid, reformat storage,

and restore object data.




If more than one Storage Node has failed (or is offline), contact technical support. Do not perform the following recovery procedure. Data loss could occur.

If this is the second Storage Node failure in less than 15 days after a Storage Node failure or recovery, contact technical support. Rebuilding Cassandra on two or more Storage Nodes within 15 days can result in data loss.

If more than one Storage Node at a site has failed, a site recovery procedure might be required. Contact technical support.

How site recovery is performed by technical support

If ILM rules are configured to store only one replicated copy and the copy exists on a storage volume that has failed, you will not be able to recover the object.

If you encounter a Services: Status - Cassandra (SVST) alarm during recovery, see the monitoring and troubleshooting instructions to recover from the alarm by rebuilding Cassandra. After Cassandra is rebuilt, alarms should clear. If alarms do not clear, contact technical support.

For hardware maintenance procedures, such as instructions for replacing a controller or reinstalling SANtricity OS, see the installation and maintenance instructions for your storage appliance.

**Related information**

Monitor & troubleshoot

SG6000 storage appliances

SG5700 storage appliances

SG5600 storage appliances

**Steps**

- Preparing an appliance Storage Node for reinstallation
- Starting StorageGRID appliance installation
- Monitoring StorageGRID appliance installation
- Selecting Start Recovery to configure an appliance Storage Node
- Remounting and reformatting appliance storage volumes (“Manual Steps”)
- Restoring object data to a storage volume for an appliance
- Checking the storage state after recovering an appliance Storage Node

### Preparing an appliance Storage Node for reinstallation

When recovering an appliance Storage Node, you must first prepare the appliance for reinstallation of StorageGRID software.

1. Log in to the failed Storage Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
2. Prepare the appliance Storage Node for the installation of StorageGRID software. `sgareinstall`
3. When prompted to continue, enter: `y`

   The appliance reboots, and your SSH session ends. It usually takes about 5 minutes for the StorageGRID Appliance Installer to become available, although in some cases you might need to wait up to 30 minutes.

   The StorageGRID appliance Storage Node is reset, and data on the Storage Node is no longer accessible. IP addresses configured during the original installation process should remain intact; however, it is

recommended that you confirm this when the procedure completes.

After executing the `sgareinstall` command, all StorageGRID-provisioned accounts, passwords, and SSH keys are removed, and new host keys are generated.

### Starting StorageGRID appliance installation

To install StorageGRID on an appliance Storage Node, you use the StorageGRID Appliance Installer, which is included on the appliance.

#### What you'll need

- The appliance has been installed in a rack, connected to your networks, and powered on.
- Network links and IP addresses have been configured for the appliance using the StorageGRID Appliance Installer.
- You know the IP address of the primary Admin Node for the StorageGRID grid.
- All Grid Network subnets listed on the IP Configuration page of the StorageGRID Appliance Installer have been defined in the Grid Network Subnet List on the primary Admin Node.
- You have completed these prerequisite tasks by following the installation and maintenance instructions for your storage appliance:
  - SG5600 storage appliances
  - SG5700 storage appliances
  - SG6000 storage appliances
- You are using a supported web browser.
- You know one of the IP addresses assigned to the compute controller in the appliance. You can use the IP address for the Admin Network (management port 1 on the controller), the Grid Network, or the Client Network.

#### About this task

To install StorageGRID on an appliance Storage Node:

- You specify or confirm the IP address of the primary Admin Node and the name of the node.
- You start the installation and wait as volumes are configured and the software is installed.
- Partway through the process, the installation pauses. To resume the installation, you must sign into the Grid Manager and configure the pending Storage Node as a replacement for the failed node.
- After you have configured the node, the appliance installation process completes, and the appliance is rebooted.

#### Steps

1. Open a browser and enter one of the IP addresses for the compute controller in the appliance.

   `https://Controller_IP:8443`

   The StorageGRID Appliance Installer Home page appears.

2. In the Primary Admin Node connection section, determine whether you need to specify the IP address for the primary Admin Node.

   The StorageGRID Appliance Installer can discover this IP address automatically, assuming the primary

Admin Node, or at least one other grid node with ADMIN_IP configured, is present on the same subnet.

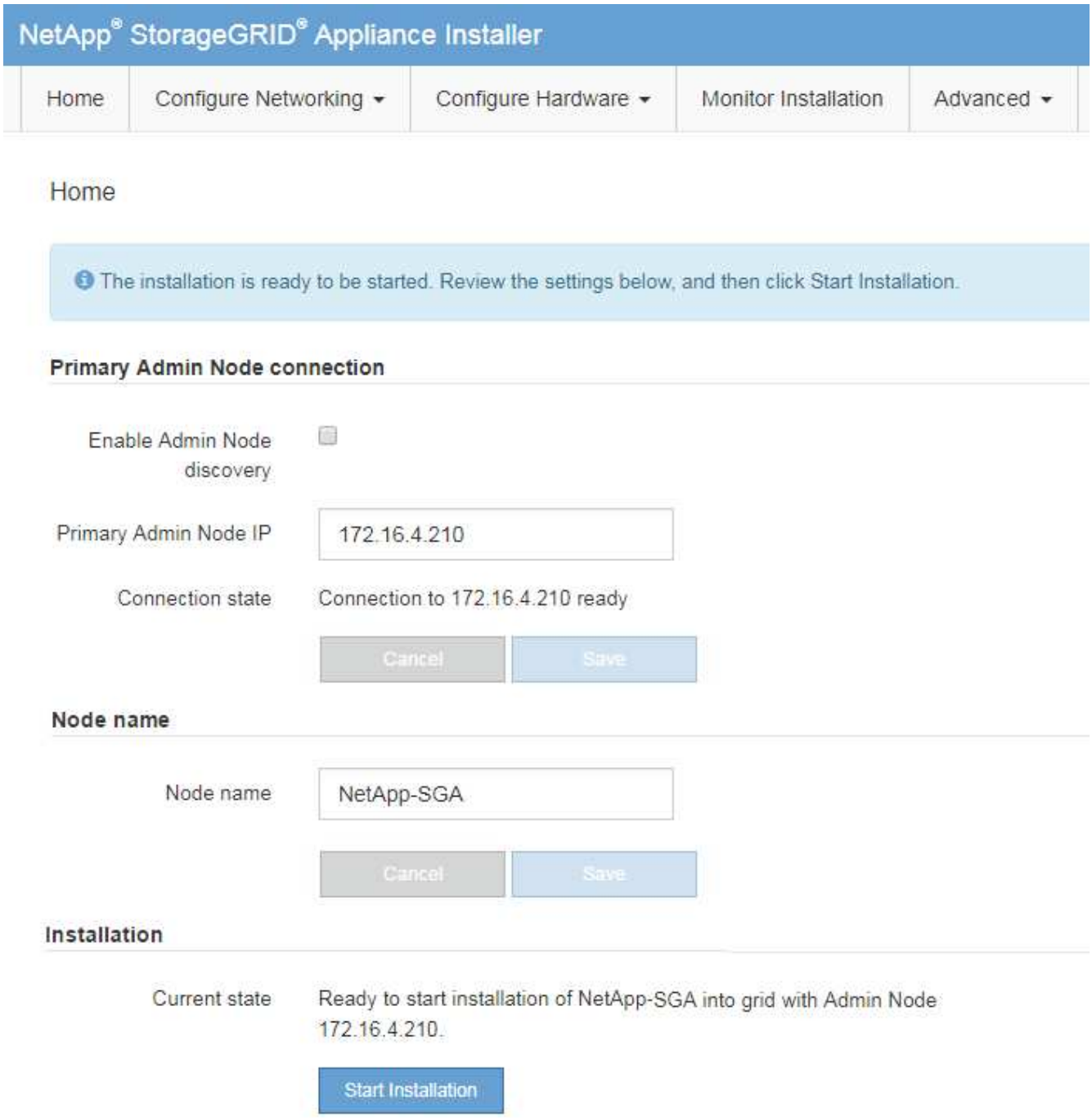
3. If this IP address is not shown or you need to change it, specify the address:

| Option | Steps |
| --- | --- |
| Manual IP entry | a. Unselect the **Enable Admin Node discovery** check box.<br>b. Enter the IP address manually.<br>c. Click **Save**.<br>d. Wait while the connection state for the new IP address becomes “ready.” |
| Automatic discovery of all connected primary Admin Nodes | a. Select the **Enable Admin Node discovery** check box.<br>b. From the list of discovered IP addresses, select the primary Admin Node for the grid where this appliance Storage Node will be deployed.<br>c. Click **Save**.<br>d. Wait while the connection state for the new IP address becomes “ready.” |

4. In the **Node Name** field, enter the same name that was used for the node you are recovering, and click **Save**.
5. In the Installation section, confirm that the current state is “Ready to start installation of node name into grid with Primary Admin Node admin_ip” and that the **Start Installation** button is enabled.

   If the **Start Installation** button is not enabled, you might need to change the network configuration or port settings. For instructions, see the installation and maintenance instructions for your appliance.

6. From the StorageGRID Appliance Installer home page, click **Start Installation**.




The Current state changes to “Installation is in progress,” and the Monitor Installation page is displayed.

If you need to access the Monitor Installation page manually, click **Monitor Installation** from the menu bar.

**Related information**

SG100 & SG1000 services appliances

SG6000 storage appliances

SG5700 storage appliances

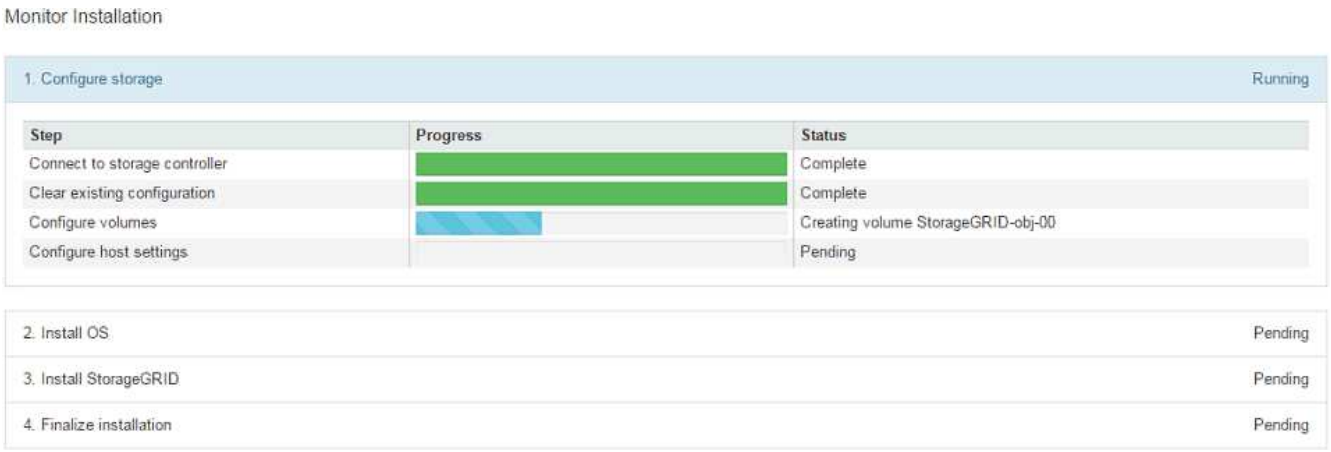
SG5600 storage appliances

### Monitoring StorageGRID appliance installation

The StorageGRID Appliance Installer provides status until installation is complete. When the software installation is complete, the appliance is rebooted.

1. To monitor the installation progress, click **Monitor Installation** from the menu bar.

   The Monitor Installation page shows the installation progress.




   The blue status bar indicates which task is currently in progress. Green status bars indicate tasks that have completed successfully.

   > The installer ensures that tasks completed in a previous install are not re-run. If you are re-running an installation, any tasks that do not need to be re-run are shown with a green status bar and a status of “Skipped.”

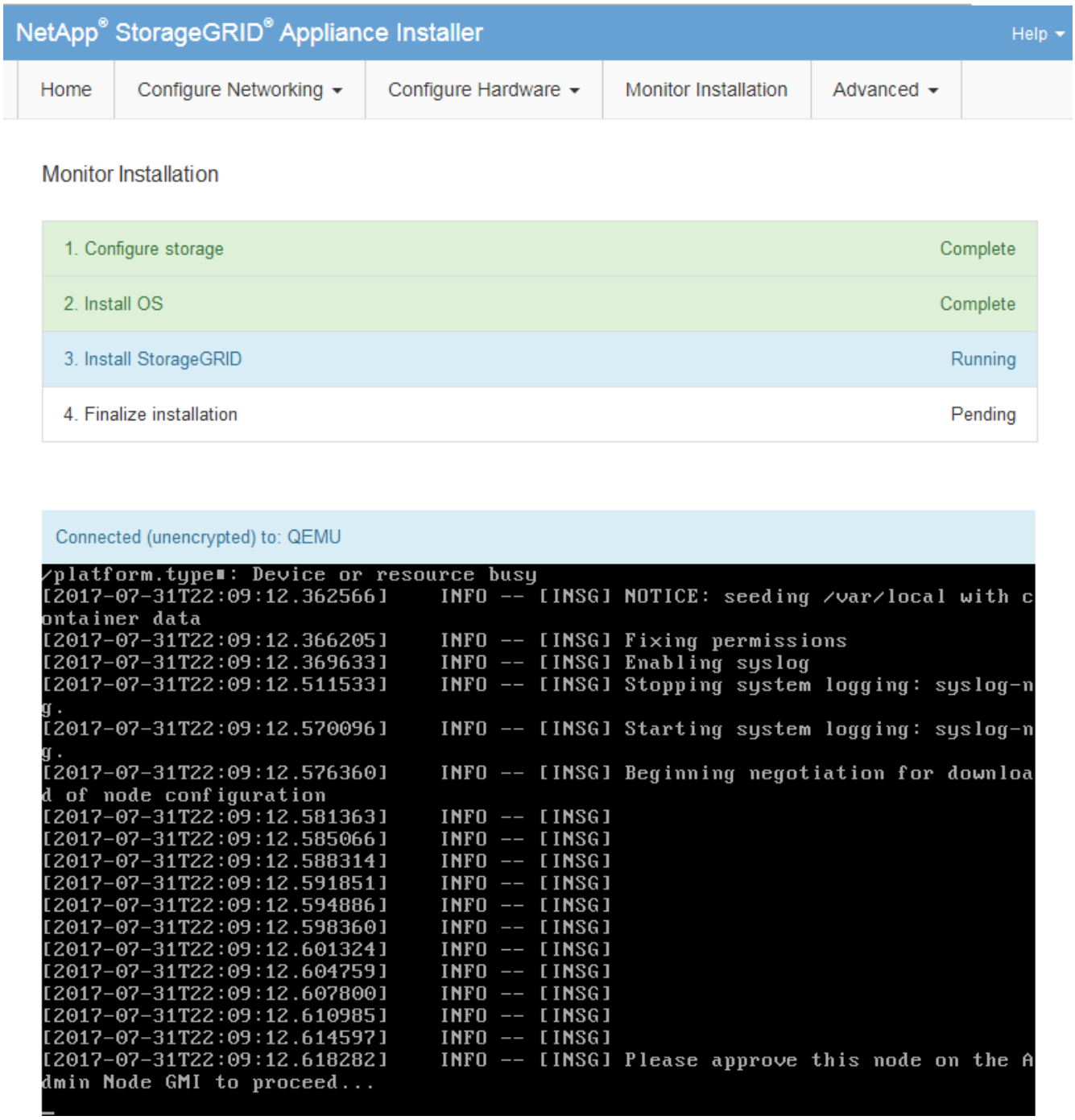
2. Review the progress of first two installation stages.
   - **1. Configure storage**

     During this stage, the installer connects to the storage controller, clears any existing configuration, communicates with SANtricity software to configure volumes, and configures host settings.

   - **2. Install OS**

     During this stage, the installer copies the base operating system image for StorageGRID to the appliance.

3. Continue monitoring the installation progress until the **Install StorageGRID** stage pauses and a message appears on the embedded console prompting you to approve this node on the Admin Node using the Grid Manager.




4. Go to the procedure to configure the appliance Storage Node.

### Selecting Start Recovery to configure an appliance Storage Node

You must select Start Recovery in the Grid Manager to configure an appliance Storage Node as a replacement for the failed node.

#### What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.

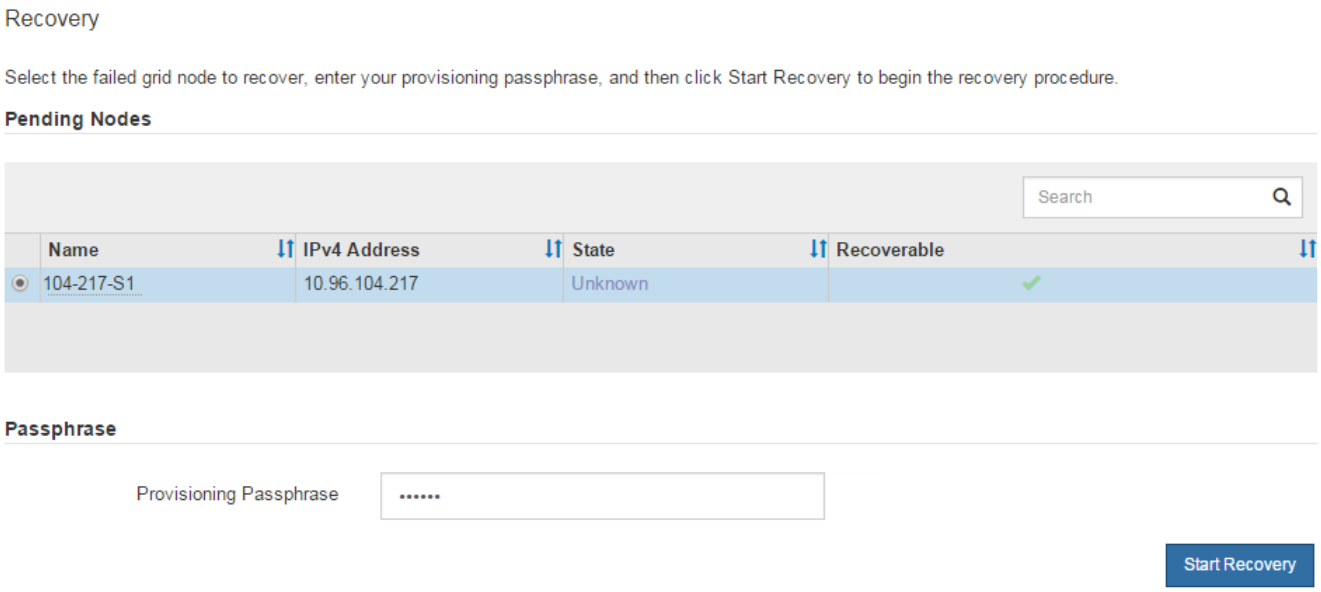
- You must have deployed a recovery appliance Storage Node.
- You must know the start date of any repair jobs for erasure-coded data.
- You must have verified that the Storage Node has not been rebuilt within the last 15 days.

**Steps**

1. From the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Recovery**.
2. Select the grid node you want to recover in the Pending Nodes list.

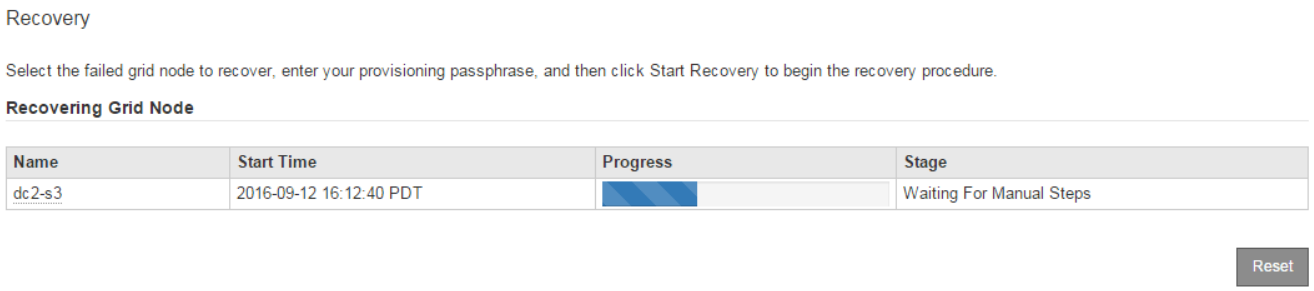
   Nodes appear in the list after they fail, but you cannot select a node until it has been reinstalled and is ready for recovery.

3. Enter the **Provisioning Passphrase**.
4. Click **Start Recovery**.

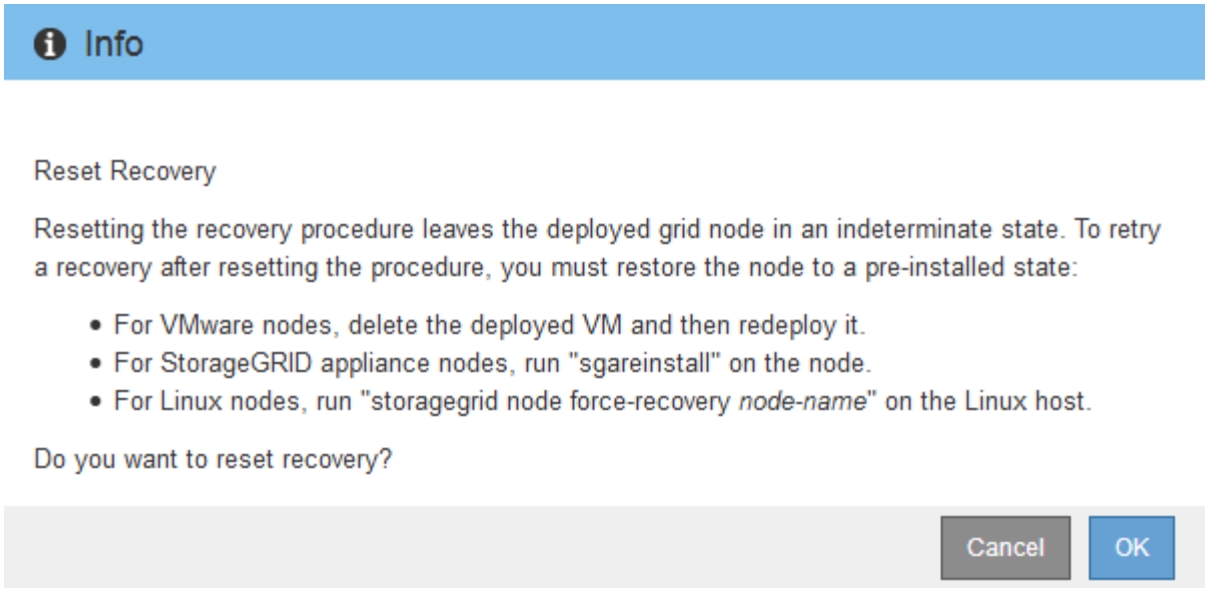



5. Monitor the progress of the recovery in the Recovering Grid Node table.

   When the grid node reaches the “Waiting for Manual Steps” stage, go to the next topic and perform the manual steps to remount and reformat appliance storage volumes.




> At any point during the recovery, you can click **Reset** to start a new recovery. An Info dialog box appears, indicating that the node will be left in an indeterminate state if you reset the procedure.




If you want to retry the recovery after resetting the procedure, you must restore the appliance node to a pre-installed state by running `sgareinstall` on the node.

### Remounting and reformatting appliance storage volumes (“Manual Steps”)

You must manually run two scripts to remount preserved storage volumes and reformat any failed storage volumes. The first script remounts volumes that are properly formatted as StorageGRID storage volumes. The second script reformats any unmounted volumes, rebuilds the Cassandra database, if needed, and starts services.

#### What you’ll need

- You have already replaced the hardware for any failed storage volumes that you know require replacement.

  Running the `sn-remount-volumes` script might help you identify additional failed storage volumes.

- You have checked that a Storage Node decommissioning is not in progress, or you have paused the node decommission procedure. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Decommission**.)
- You have checked that an expansion is not in progress. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Expansion**.)

> Contact technical support if more than one Storage Node is offline or if a Storage Node in this grid has been rebuilt in the last 15 days. Do not run the `sn-recovery-postinstall.sh` script. Rebuilding Cassandra on two or more Storage Nodes within 15 days of each other might result in data loss.

#### About this task

To complete this procedure, you perform these high-level tasks:

- Log in to the recovered Storage Node.
- Run the `sn-remount-volumes` script to remount properly formatted storage volumes. When this script runs, it does the following:

- Mounts and unmounts each storage volume to replay the XFS journal.
- Performs an XFS file consistency check.
- If the file system is consistent, determines if the storage volume is a properly formatted StorageGRID storage volume.
- If the storage volume is properly formatted, remounts the storage volume. Any existing data on the volume remains intact.

- Review the script output and resolve any issues.
- Run the `sn-recovery-postinstall.sh` script. When this script runs, it does the following.

> Do not reboot a Storage Node during recovery before running `sn-recovery-postinstall.sh` (step 4) to reformat the failed storage volumes and restore object metadata. Rebooting the Storage Node before `sn-recovery-postinstall.sh` completes causes errors for services that attempt to start and causes StorageGRID appliance nodes to exit maintenance mode.

  - Reformats any storage volumes that the `sn-remount-volumes` script could not mount or that were found to be improperly formatted.

    > If a storage volume is reformatted, any data on that volume is lost. You must perform an additional procedure to restore object data from other locations in the grid, assuming that ILM rules were configured to store more than one object copy.

  - Rebuilds the Cassandra database on the node, if needed.
  - Starts the services on the Storage Node.

**Steps**

1. Log in to the recovered Storage Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Run the first script to remount any properly formatted storage volumes.

   > If all storage volumes are new and need to be formatted, or if all storage volumes have failed, you can skip this step and run the second script to reformat all unmounted storage volumes.

   a. Run the script: `sn-remount-volumes`

      This script might take hours to run on storage volumes that contain data.

   b. As the script runs, review the output and answer any prompts.

As required, you can use the `tail -f` command to monitor the contents of the script's log file (`/var/local/log/sn-remount-volumes.log`) . The log file contains more detailed information than the command line output.

```
root@SG:~ # sn-remount-volumes
The configured LDR noid is 12632740

====== Device /dev/sdb ======
Mount and unmount device /dev/sdb and checking file system
consistency:
The device is consistent.
Check rangedb structure on device /dev/sdb:
Mount device /dev/sdb to /tmp/sdb-654321 with rangedb mount options
This device has all rangedb directories.
Found LDR node id 12632740, volume number 0 in the volID file
Attempting to remount /dev/sdb
Device /dev/sdb remounted successfully

====== Device /dev/sdc ======
Mount and unmount device /dev/sdc and checking file system
consistency:
Error: File system consistency check retry failed on device /dev/sdc.
You can see the diagnosis information in the /var/local/log/sn-
remount-volumes.log.

This volume could be new or damaged. If you run sn-recovery-
postinstall.sh, this volume and any data on this volume will be
deleted. If you only had two copies of object data, you will
temporarily have only a single copy.
StorageGRID Webscale will attempt to restore data redundancy by
making additional replicated copies or EC fragments, according to the
rules in the active ILM policy.

Do not continue to the next step if you believe that the data
remaining on this volume cannot be rebuilt from elsewhere in the grid
(for example, if your ILM policy uses a rule that makes only one copy
or if volumes have failed on multiple nodes). Instead, contact
support to determine how to recover your data.

====== Device /dev/sdd ======
Mount and unmount device /dev/sdd and checking file system
consistency:
Failed to mount device /dev/sdd
This device could be an uninitialized disk or has corrupted
superblock.
File system check might take a long time. Do you want to continue? (y
```

```
or n) [y/N]? y

Error: File system consistency check retry failed on device /dev/sdd.
You can see the diagnosis information in the /var/local/log/sn-
remount-volumes.log.

This volume could be new or damaged. If you run sn-recovery-
postinstall.sh, this volume and any data on this volume will be
deleted. If you only had two copies of object data, you will
temporarily have only a single copy.
StorageGRID Webscale will attempt to restore data redundancy by
making additional replicated copies or EC fragments, according to the
rules in the active ILM policy.

Do not continue to the next step if you believe that the data
remaining on this volume cannot be rebuilt from elsewhere in the grid
(for example, if your ILM policy uses a rule that makes only one copy
or if volumes have failed on multiple nodes). Instead, contact
support to determine how to recover your data.

====== Device /dev/sde ======
Mount and unmount device /dev/sde and checking file system
consistency:
The device is consistent.
Check rangedb structure on device /dev/sde:
Mount device /dev/sde to /tmp/sde-654321 with rangedb mount options
This device has all rangedb directories.
Found LDR node id 12000078, volume number 9 in the volID file
Error: This volume does not belong to this node. Fix the attached
volume and re-run this script.
```

In the example output, one storage volume was remounted successfully and three storage volumes had errors.

- `/dev/sdb` passed the XFS file system consistency check and had a valid volume structure, so it was remounted successfully. Data on devices that are remounted by the script is preserved.
- `/dev/sdc` failed the XFS file system consistency check because the storage volume was new or corrupt.
- `/dev/sdd` could not be mounted because the disk was uninitialized or the disk's superblock was corrupted. When the script cannot mount a storage volume, it asks if you want to run the file system consistency check.
  - If the storage volume is attached to a new disk, answer **N** to the prompt. You do not need check the file system on a new disk.
  - If the storage volume is attached to an existing disk, answer **Y** to the prompt. You can use the results of the file system check to determine the source of the corruption. The results are saved in the `/var/local/log/sn-remount-volumes.log` log file.

- `/dev/sde` passed the XFS file system consistency check and had a valid volume structure; however, the LDR node ID in the `volID` file did not match the ID for this Storage Node (the `configured LDR noid` displayed at the top). This message indicates that this volume belongs to another Storage Node.

3. Review the script output and resolve any issues.

> If a storage volume failed the XFS file system consistency check or could not be mounted, carefully review the error messages in the output. You must understand the implications of running the `sn-recovery-postinstall.sh` script on these volumes.

a. Check to make sure that the results include an entry for all of the volumes you expected. If any volumes are not listed, rerun the script.

b. Review the messages for all mounted devices. Make sure there are no errors indicating that a storage volume does not belong to this Storage Node.

In the example, the output for /dev/sde includes the following error message:

```
Error: This volume does not belong to this node. Fix the attached
volume and re-run this script.
```

> If a storage volume is reported as belonging to another Storage Node, contact technical support. If you run the `sn-recovery-postinstall.sh` script, the storage volume will be reformatted, which might cause data loss.

c. If any storage devices could not be mounted, make a note of the device name, and repair or replace the device.

> You must repair or replace any storage devices that could not be mounted.

You will use the device name to look up the volume ID, which is required input when you run the `repair-data` script to restore object data to the volume (the next procedure).

d. After repairing or replacing all unmountable devices, run the `sn-remount-volumes` script again to confirm that all storage volumes that can be remounted have been remounted.

> If a storage volume cannot be mounted or is improperly formatted, and you continue to the next step, the volume and any data on the volume will be deleted. If you had two copies of object data, you will have only a single copy until you complete the next procedure (restoring object data).

> Do not run the `sn-recovery-postinstall.sh` script if you believe that the data remaining on a failed storage volume cannot be rebuilt from elsewhere in the grid (for example, if your ILM policy uses a rule that makes only one copy or if volumes have failed on multiple nodes). Instead, contact technical support to determine how to recover your data.

4. Run the `sn-recovery-postinstall.sh` script: `sn-recovery-postinstall.sh`

This script reformats any storage volumes that could not be mounted or that were found to be improperly formatted; rebuilds the Cassandra database on the node, if needed; and starts the services on the Storage Node.

Be aware of the following:

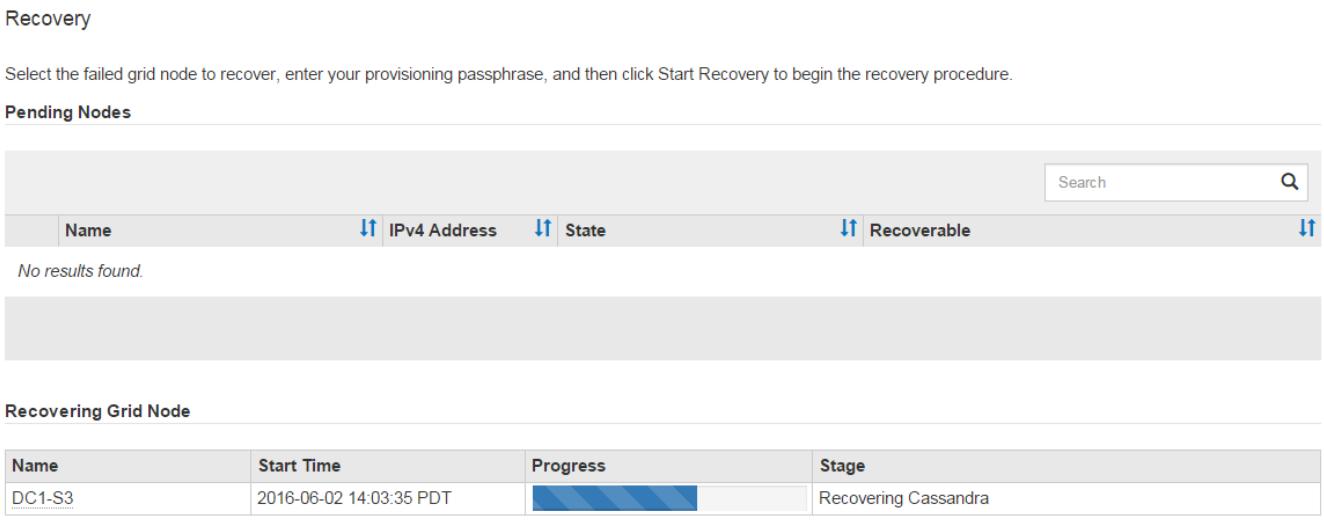
- The script might take hours to run.
- In general, you should leave the SSH session alone while the script is running.
- Do not press **Ctrl+C** while the SSH session is active.
- The script will run in the background if a network disruption occurs and terminates the SSH session, but you can view the progress from the Recovery page.
- If the Storage Node uses the RSM service, the script might appear to stall for 5 minutes as node services are restarted. This 5-minute delay is expected whenever the RSM service boots for the first time.

> The RSM service is present on Storage Nodes that include the ADC service.

> Some StorageGRID recovery procedures use Reaper to handle Cassandra repairs. Repairs occur automatically as soon as the related or required services have started. You might notice script output that mentions “reaper” or “Cassandra repair.” If you see an error message indicating the repair has failed, run the command indicated in the error message.

5. As the `sn-recovery-postinstall.sh` script runs, monitor the Recovery page in the Grid Manager.

   The Progress bar and the Stage column on the Recovery page provide a high-level status of the `sn-recovery-postinstall.sh` script.




6. Return to the Monitor Install page of the StorageGRID Appliance Installer by entering `http://Controller_IP:8080`, using the IP address of the compute controller.

   The Monitor Install page shows the installation progress while the script is running.

After the `sn-recovery-postinstall.sh` script has started services on the node, you can restore object data to any storage volumes that were formatted by the script, as described in the next procedure.

**Related information**

Reviewing warnings for Storage Node system drive recovery

Restoring object data to a storage volume for an appliance

### Restoring object data to a storage volume for an appliance

After recovering storage volumes for the appliance Storage Node, you can restore the object data that was lost when the Storage Node failed.

**What you'll need**

- You must have confirmed that the recovered Storage Node has a Connection State of **Connected*** ✔ **on the *Nodes** > **Overview** tab in the Grid Manager.

**About this task**

Object data can be restored from other Storage Nodes, an Archive Node, or a Cloud Storage Pool, assuming that the grid's ILM rules were configured such that object copies are available.

> If an ILM rule was configured to store only one replicated copy and that copy existed on a storage volume that failed, you will not be able to recover the object.

> If the only remaining copy of an object is in a Cloud Storage Pool, StorageGRID must issue multiple requests to the Cloud Storage Pool endpoint to restore object data. Before performing this procedure, contact technical support for help in estimating the recovery time frame and the associated costs.

> If the only remaining copy of an object is on an Archive Node, object data is retrieved from the Archive Node. Due to the latency associated with retrievals from external archival storage systems, restoring object data to a Storage Node from an Archive Node takes longer than restoring copies from other Storage Nodes.

To restore object data, you run the `repair-data` script. This script begins the process of restoring object data and works with ILM scanning to ensure that ILM rules are met. You use different options with the `repair-data` script, based on whether you are restoring replicated data or erasure coded data, as follows:

- **Replicated data**: Two commands are available for restoring replicated data, based on whether you need to repair the entire node or only certain volumes on the node:

```
repair-data start-replicated-node-repair
```

```
repair-data start-replicated-volume-repair
```

- **Erasure coded (EC) data**: Two commands are available for restoring erasure coded data, based on whether you need to repair the entire node or only certain volumes on the node:

```
repair-data start-ec-node-repair
```

```
repair-data start-ec-volume-repair
```

Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available. You can track repairs of erasure coded data with this command:

```
repair-data show-ec-repair-status
```

> The EC repair job temporarily reserves a large amount of storage. Storage alerts might be triggered, but will resolve when the repair is complete. If there is not enough storage for the reservation, the EC repair job will fail. Storage reservations are released when the EC repair job completes, whether the job failed or succeeded.

For more information on using the `repair-data` script, enter `repair-data --help` from the command line of the primary Admin Node.

**Steps**

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
2. Use the `/etc/hosts` file to find the hostname of the Storage Node for the restored storage volumes. To see a list of all nodes in the grid, enter the following: `cat /etc/hosts`
3. If all storage volumes have failed, repair the entire node. (If only some volumes have failed, go to the next step.)

   > You cannot run `repair-data` operations for more than one node at the same time. To recover multiple nodes, contact technical support.

   - If your grid includes replicated data, use the `repair-data start-replicated-node-repair` command with the `--nodes` option to repair the entire Storage Node.

     This command repairs the replicated data on a Storage Node named SG-DC-SN3:

     ```
     repair-data start-replicated-node-repair --nodes SG-DC-SN3
     ```

     > As object data is restored, the **Objects Lost** alert is triggered if the StorageGRID system cannot locate replicated object data. Alerts might be triggered on Storage Nodes throughout the system. You should determine the cause of the loss and if recovery is possible. See the instructions for monitoring and troubleshooting StorageGRID.

- If your grid contains erasure coded data, use the `repair-data start-ec-node-repair` command with the `--nodes` option to repair the entire Storage Node.

  This command repairs the erasure coded data on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-node-repair --nodes SG-DC-SN3
  ```

  The operation returns a unique `repair ID` that identifies this `repair_data` operation. Use this `repair ID` to track the progress and result of the `repair_data` operation. No other feedback is returned as the recovery process completes.

  > Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available.

- If your grid has both replicated and erasure coded data, run both commands.

4. If only some of the volumes have failed, repair the affected volumes.

   Enter the volume IDs in hexadecimal. For example, `0000` is the first volume and `000F` is the sixteenth volume. You can specify one volume, a range of volumes, or multiple volumes that are not in a sequence.

   All the volumes must be on the same Storage Node. If you need to restore volumes for more than one Storage Node, contact technical support.

   - If your grid contains replicated data, use the `start-replicated-volume-repair` command with the `--nodes` option to identify the node. Then add either the `--volumes` or `--volume-range` option, as shown in the following examples.

     **Single volume**: This command restores replicated data to volume `0002` on a Storage Node named SG-DC-SN3:

     ```
     repair-data start-replicated-volume-repair --nodes SG-DC-SN3
     --volumes 0002
     ```

     **Range of volumes**: This command restores replicated data to all volumes in the range `0003` to `0009` on a Storage Node named SG-DC-SN3:

     ```
     repair-data start-replicated-volume-repair --nodes SG-DC-SN3 --volume
     -range 0003-0009
     ```

     **Multiple volumes not in a sequence**: This command restores replicated data to volumes `0001`, `0005`, and `0008` on a Storage Node named SG-DC-SN3:

     ```
     repair-data start-replicated-volume-repair --nodes SG-DC-SN3
     --volumes 0001,0005,0008
     ```

> As object data is restored, the **Objects Lost** alert is triggered if the StorageGRID system cannot locate replicated object data. Alerts might be triggered on Storage Nodes throughout the system. You should determine the cause of the loss and if recovery is possible. See the instructions for monitoring and troubleshooting StorageGRID.

- If your grid contains erasure coded data, use the `start-ec-volume-repair` command with the `--nodes` option to identify the node. Then add either the `--volumes` or `--volume-range` option, as shown in the following examples.

  **Single volume**: This command restores erasure coded data to volume `0007` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volumes 0007
  ```

  **Range of volumes**: This command restores erasure coded data to all volumes in the range `0004` to `0006` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volume-range
  0004-0006
  ```

  **Multiple volumes not in a sequence**: This command restores erasure coded data to volumes `000A`, `000C`, and `000E` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volumes
  000A,000C,000E
  ```

  The `repair-data` operation returns a unique `repair ID` that identifies this `repair_data` operation. Use this `repair ID` to track the progress and result of the `repair_data` operation. No other feedback is returned as the recovery process completes.

  > Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available.

- If your grid has both replicated and erasure coded data, run both commands.

5. Monitor the repair of replicated data.
   a. Select **Nodes** > **Storage Node being repaired** > **ILM**.
   b. Use the attributes in the Evaluation section to determine if repairs are complete.

      When repairs are complete, the Awaiting - All attribute indicates 0 objects.

   c. To monitor the repair in more detail, select **Support** > **Tools** > **Grid Topology**.
   d. Select **grid** > **Storage Node being repaired** > **LDR** > **Data Store**.
   e. Use a combination of the following attributes to determine, as well as possible, if replicated repairs are complete.

Cassandra inconsistencies might be present, and failed repairs are not tracked.

- **Repairs Attempted (XRPA)**: Use this attribute to track the progress of replicated repairs. This attribute increases each time a Storage Node tries to repair a high-risk object. When this attribute does not increase for a period longer than the current scan period (provided by the **Scan Period — Estimated** attribute), it means that ILM scanning found no high-risk objects that need to be repaired on any nodes.

  High-risk objects are objects that are at risk of being completely lost. This does not include objects that do not satisfy their ILM configuration.

- **Scan Period — Estimated (XSCM)**: Use this attribute to estimate when a policy change will be applied to previously ingested objects. If the **Repairs Attempted** attribute does not increase for a period longer than the current scan period, it is probable that replicated repairs are done. Note that the scan period can change. The **Scan Period — Estimated (XSCM)** attribute applies to the entire grid and is the maximum of all node scan periods. You can query the **Scan Period — Estimated** attribute history for the grid to determine an appropriate time frame.

6. Monitor the repair of erasure coded data, and retry any requests that might have failed.
   a. Determine the status of erasure coded data repairs:
      - Use this command to see the status of a specific `repair-data` operation:

```
repair-data show-ec-repair-status --repair-id repair ID
```

      - Use this command to list all repairs:

```
repair-data show-ec-repair-status
```

      The output lists information, including `repair ID`, for all previously and currently running repairs.

```
root@DC1-ADM1:~ # repair-data show-ec-repair-status

Repair ID Scope Start Time  End Time  State  Est Bytes
Affected/Repaired Retry Repair
================================================================
==================
 949283 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:27:06.9 Success
17359 17359 No
 949292 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:37:06.9 Failure
17359 0     Yes
 949294 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:47:06.9 Failure
17359 0     Yes
 949299 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:57:06.9 Failure
17359 0     Yes
```

b. If the output shows that the repair operation failed, use the `--repair-id` option to retry the repair.

   This command retries a failed node repair, using the repair ID `83930030303133434`:

   ```
   repair-data start-ec-node-repair --repair-id 83930030303133434
   ```

   This command retries a failed volume repair, using the repair ID `83930030303133434`:

   ```
   repair-data start-ec-volume-repair --repair-id 83930030303133434
   ```

**Related information**

Monitor & troubleshoot

### Checking the storage state after recovering an appliance Storage Node

After recovering an appliance Storage Node, you must verify that the desired state of the appliance Storage Node is set to online and ensure that the state will be online by default whenever the Storage Node server is restarted.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- The Storage Node has been recovered, and data recovery is complete.

**Steps**

1. Select **Support** > **Tools** > **Grid Topology**.
2. Check the values of **Recovered Storage Node** > **LDR** > **Storage** > **Storage State — Desired** and **Storage State — Current**.

   The value of both attributes should be Online.

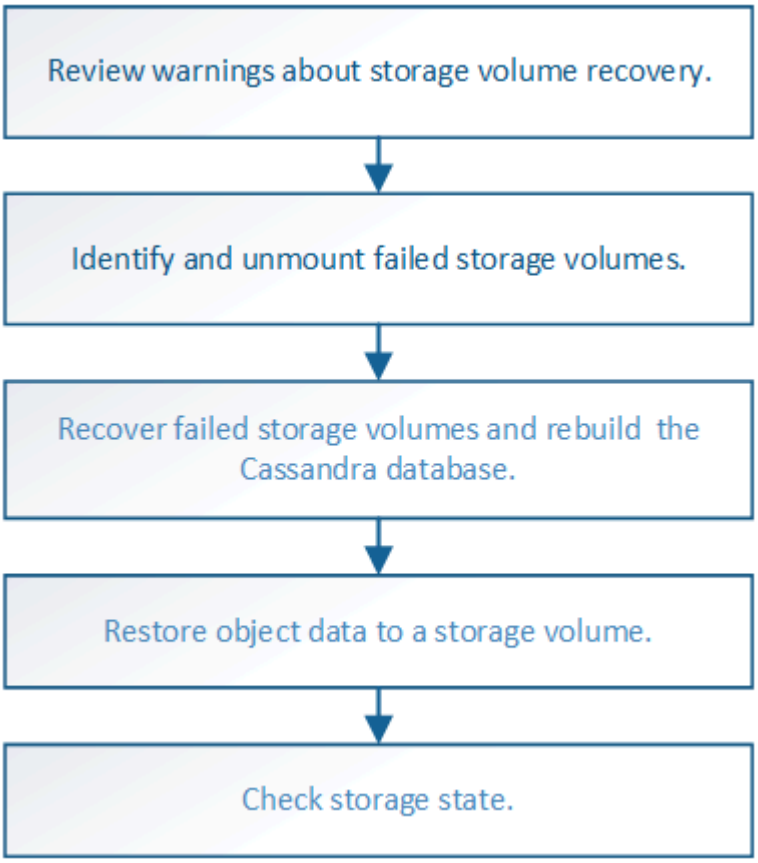
3. If the Storage State — Desired is set to Read-only, complete the following steps:
   a. Click the **Configuration** tab.
   b. From the **Storage State — Desired** drop-down list, select **Online**.
   c. Click **Apply Changes**.
   d. Click the **Overview** tab and confirm that the values of **Storage State — Desired** and **Storage State — Current** are updated to Online.

#### Recovering from storage volume failure where the system drive is intact

You must complete a series of tasks to recover a software-based Storage Node where one or more storage volumes on the Storage Node have failed, but the system drive is intact. If only storage volumes have failed, the Storage Node is still available to the StorageGRID system.

**About this task**

This recovery procedure applies to software-based Storage Nodes only. If storage volumes have failed on an appliance Storage Node, use the procedure for “Recovering a StorageGRID appliance Storage Node.”




**Related information**

Recovering a StorageGRID appliance Storage Node

**Steps**

- Reviewing warnings about storage volume recovery
- Identifying and unmounting failed storage volumes
- Recovering failed storage volumes and rebuilding the Cassandra database
- Restoring object data to a storage volume where the system drive is intact
- Checking the storage state after recovering storage volumes

### Reviewing warnings about storage volume recovery

Before recovering failed storage volumes for a Storage Node, you must review the following warnings.

The storage volumes (or rangedbs) in a Storage Node are identified by a hexadecimal number, which is known as the volume ID. For example, 0000 is the first volume and 000F is the sixteenth volume. The first object store (volume 0) on each Storage Node uses up to 4 TB of space for object metadata and Cassandra database operations; any remaining space on that volume is used for object data. All other storage volumes are used exclusively for object data.

If volume 0 fails and needs to be recovered, the Cassandra database might be rebuilt as part of the volume recovery procedure. Cassandra might also be rebuilt in the following circumstances:

- A Storage Node is brought back online after having been offline for more than 15 days.
- The system drive and one or more storage volumes fails and is recovered.

When Cassandra is rebuilt, the system uses information from other Storage Nodes. If too many Storage Nodes are offline, some Cassandra data might not be available. If Cassandra has been rebuilt recently, Cassandra data might not yet be consistent across the grid. Data loss can occur if Cassandra is rebuilt when too many Storage Nodes are offline or if two or more Storage Nodes are rebuilt within 15 days of each other.

If more than one Storage Node has failed (or is offline), contact technical support. Do not perform the following recovery procedure. Data loss could occur.

If this is the second Storage Node failure in less than 15 days after a Storage Node failure or recovery, contact technical support. Rebuilding Cassandra on two or more Storage Nodes within 15 days can result in data loss.

If more than one Storage Node at a site has failed, a site recovery procedure might be required. Contact technical support.

How site recovery is performed by technical support

If ILM rules are configured to store only one replicated copy and the copy exists on a storage volume that has failed, you will not be able to recover the object.

If you encounter a Services: Status - Cassandra (SVST) alarm during recovery, see the monitoring and troubleshooting instructions to recover from the alarm by rebuilding Cassandra. After Cassandra is rebuilt, alarms should clear. If alarms do not clear, contact technical support.

**Related information**

Monitor & troubleshoot

Warnings and considerations for grid node recovery

### Identifying and unmounting failed storage volumes

When recovering a Storage Node with failed storage volumes, you must identify and unmount the failed volumes. You must verify that only the failed storage volumes are reformatted as part of the recovery procedure.

**What you'll need**

You must be signed in to the Grid Manager using a supported browser.

**About this task**

You should recover failed storage volumes as soon as possible.

The first step of the recovery process is to detect volumes that have become detached, need to be unmounted, or have I/O errors. If failed volumes are still attached but have a randomly corrupted file system, the system might not detect any corruption in unused or unallocated parts of the disk.

You must finish this procedure before performing manual steps to recover the volumes, such as adding or re-attaching the disks, stopping the node, starting the node, or rebooting. Otherwise, when you run the `reformat_storage_block_devices.rb` script, you might encounter a file system error that causes the script to hang or fail.

Repair the hardware and properly attach the disks before running the `reboot` command.

Identify failed storage volumes carefully. You will use this information to verify which volumes must be reformatted. Once a volume has been reformatted, data on the volume cannot be recovered.

To correctly recover failed storage volumes, you need to know both the device names of the failed storage volumes and their volume IDs.

At installation, each storage device is assigned a file system universal unique identifier (UUID) and is mounted to a rangedb directory on the Storage Node using that assigned file system UUID. The file system UUID and the rangedb directory are listed in the `/etc/fstab` file. The device name, rangedb directory, and the size of the mounted volume are displayed in the Grid Manager.

In the following example, device `/dev/sdc` has a volume size of 4 TB, is mounted to `/var/local/rangedb/0`, using the device name `/dev/disk/by-uuid/822b0547-3b2b-472e-ad5e-e1cf1809faba` in the `/etc/fstab` file:

**Steps**

1. Complete the following steps to record the failed storage volumes and their device names:
   a. Select **Support** > **Tools** > **Grid Topology**.
   b. Select **site** > **failed Storage Node** > **LDR** > **Storage** > **Overview** > **Main**, and look for object stores with alarms.

**Object Stores**

| ID | Total | Available | Stored Data | Stored (%) | Health |
|---|---|---|---|---|---|
| 0000 | 96.6 GB | 96.6 GB | 823 KB | 0.001 % | Error |
| 0001 | 107 GB | 107 GB | 0 B | 0 % | No Errors |
| 0002 | 107 GB | 107 GB | 0 B | 0 % | No Errors |

c. Select **site** > **failed Storage Node** > **SSM** > **Resources** > **Overview** > **Main**. Determine the mount point and volume size of each failed storage volume identified in the previous step.

Object stores are numbered in hex notation. For example, 0000 is the first volume and 000F is the sixteenth volume. In the example, the object store with an ID of 0000 corresponds to `/var/local/rangedb/0` with device name sdc and a size of 107 GB.

**Volumes**

| Mount Point | Device | Status | Size | Space Available | Total Entries | Entries Available | Write Cache |
|---|---|---|---|---|---|---|---|
| / | croot | Online | 10.4 GB | 4.17 GB | 655,360 | 554,806 | Unknown |
| /var/local | cvloc | Online | 96.6 GB | 96.1 GB | 94,369,792 | 94,369,423 | Unknown |
| /var/local/rangedb/0 | sdc | Online | 107 GB | 107 GB | 104,857,600 | 104,856,202 | Enabled |
| /var/local/rangedb/1 | sdd | Online | 107 GB | 107 GB | 104,857,600 | 104,856,536 | Enabled |
| /var/local/rangedb/2 | sde | Online | 107 GB | 107 GB | 104,857,600 | 104,856,536 | Enabled |

2. Log in to the failed Storage Node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

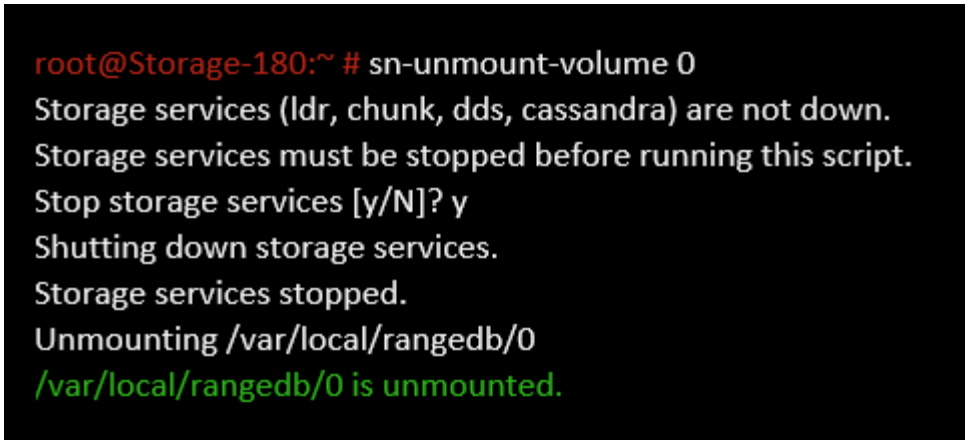
3. Run the following script to stop the storage services and unmount a failed storage volume:

   ```
   sn-unmount-volume object_store_ID
   ```

   The `object_store_ID` is the ID of the failed storage volume. For example, specify `0` in the command for an object store with ID 0000.

4. If prompted, press **y** to stop the storage services on the Storage Node.

   > If the storage services are already stopped, you are not prompted. The Cassandra service is stopped only for volume 0.




In a few seconds, the storage services are stopped and the volume is unmounted. Messages appear indicating each step of the process. The final message indicates that the volume is unmounted.

#### Recovering failed storage volumes and rebuilding the Cassandra database

You must run a script that reformats and remounts storage on failed storage volumes, and rebuilds the Cassandra database on the Storage Node if the system determines that it is necessary.

- You must have the `Passwords.txt` file.
- The system drives on the server must be intact.
- The cause of the failure must have been identified and, if necessary, replacement storage hardware must already have been acquired.
- The total size of the replacement storage must be the same as the original.
- You have checked that a Storage Node decommissioning is not in progress, or you have paused the node decommission procedure. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Decommission**.)
- You have checked that an expansion is not in progress. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Expansion**.)
- You have reviewed the warnings about storage volume recovery.

  Reviewing warnings about storage volume recovery

1. As needed, replace failed physical or virtual storage associated with the failed storage volumes that you identified and unmounted earlier.

   After you replace the storage, make sure you rescan or reboot to make sure that it is recognized by the operating system, but do not remount the volumes. The storage is remounted and added to `/etc/fstab` in a later step.

2. Log in to the failed Storage Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

1. Use a text editor (vi or vim) to delete failed volumes from the `/etc/fstab` file and then save the file.

   > Commenting out a failed volume in the `/etc/fstab` file is insufficient. The volume must be deleted from `fstab` as the recovery process verifies that all lines in the `fstab` file match the mounted file systems.

2. Reformat any failed storage volumes and rebuild the Cassandra database if it is necessary. Enter: `reformat_storage_block_devices.rb`
   - If storage services are running, you will be prompted to stop them. Enter: **y**
   - You will be prompted to rebuild the Cassandra database if it is necessary.
     - Review the warnings. If none of them apply, rebuild the Cassandra database. Enter: **y**
     - If more than one Storage Node is offline or if another Storage Node has been rebuilt in the last 15 days. Enter: **n**

       The script will exit without rebuilding Cassandra. Contact technical support.

   - For each rangedb drive on the Storage Node, when you are asked: `Reformat the rangedb drive <name> (device <major number>:<minor number>)? [y/n]?`, enter one of the following responses:
     - **y** to reformat a drive that had errors. This reformats the storage volume and adds the reformatted storage volume to the `/etc/fstab` file.
     - **n** if the drive contains no errors, and you do not want to reformat it.

       > Selecting **n** exits the script. Either mount the drive (if you think the data on the drive should be retained and the drive was unmounted in error) or remove the drive. Then, run the `reformat_storage_block_devices.rb` command again.

       > Some StorageGRID recovery procedures use Reaper to handle Cassandra repairs. Repairs occur automatically as soon as the related or required services have started. You might notice script output that mentions “reaper” or “Cassandra repair.” If you see an error message indicating the repair has failed, run the command indicated in the error message.

       In the following example output, the drive `/dev/sdf` must be reformatted, and Cassandra did not need to be rebuilt:

```
root@DC1-S1:~ # reformat_storage_block_devices.rb
Storage services must be stopped before running this script.
Stop storage services [y/N]? **y**
Shutting down storage services.
Storage services stopped.
Formatting devices that are not in use...
Skipping in use device /dev/sdc
Skipping in use device /dev/sdd
Skipping in use device /dev/sde
Reformat the rangedb drive /dev/sdf (device 8:64)? [Y/n]? **y**
Successfully formatted /dev/sdf with UUID c817f87f-f989-4a21-
8f03-b6f42180063f
Skipping in use device /dev/sdg
All devices processed
Running: /usr/local/ldr/setup_rangedb.sh 12075630
Cassandra does not need rebuilding.
Starting services.

Reformatting done.  Now do manual steps to
restore copies of data.
```

**Related information**

Reviewing warnings about storage volume recovery

### Restoring object data to a storage volume where the system drive is intact

After recovering a storage volume on a Storage Node where the system drive is intact, you can restore the object data that was lost when the storage volume failed.

**What you'll need**

- You must have confirmed that the recovered Storage Node has a Connection State of **Connected*** ✔ **on the *Nodes** > **Overview** tab in the Grid Manager.

**About this task**

Object data can be restored from other Storage Nodes, an Archive Node, or a Cloud Storage Pool, assuming that the grid's ILM rules were configured such that object copies are available.

> ⓘ If an ILM rule was configured to store only one replicated copy and that copy existed on a storage volume that failed, you will not be able to recover the object.

> ⓘ If the only remaining copy of an object is in a Cloud Storage Pool, StorageGRID must issue multiple requests to the Cloud Storage Pool endpoint to restore object data. Before performing this procedure, contact technical support for help in estimating the recovery time frame and the associated costs.

If the only remaining copy of an object is on an Archive Node, object data is retrieved from the Archive Node. Due to the latency associated with retrievals from external archival storage systems, restoring object data to a Storage Node from an Archive Node takes longer than restoring copies from other Storage Nodes.

To restore object data, you run the `repair-data` script. This script begins the process of restoring object data and works with ILM scanning to ensure that ILM rules are met. You use different options with the `repair-data` script, based on whether you are restoring replicated data or erasure coded data, as follows:

- **Replicated data**: Two commands are available for restoring replicated data, based on whether you need to repair the entire node or only certain volumes on the node:

```
repair-data start-replicated-node-repair
```

```
repair-data start-replicated-volume-repair
```

- **Erasure coded (EC) data**: Two commands are available for restoring erasure coded data, based on whether you need to repair the entire node or only certain volumes on the node:

```
repair-data start-ec-node-repair
```

```
repair-data start-ec-volume-repair
```

  Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available. You can track repairs of erasure coded data with this command:

```
repair-data show-ec-repair-status
```

The EC repair job temporarily reserves a large amount of storage. Storage alerts might be triggered, but will resolve when the repair is complete. If there is not enough storage for the reservation, the EC repair job will fail. Storage reservations are released when the EC repair job completes, whether the job failed or succeeded.

For more information on using the `repair-data` script, enter `repair-data --help` from the command line of the primary Admin Node.

**Steps**

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

2. Use the `/etc/hosts` file to find the hostname of the Storage Node for the restored storage volumes. To see a list of all nodes in the grid, enter the following: `cat /etc/hosts`
3. If all storage volumes have failed, repair the entire node. (If only some volumes have failed, go to the next step.)

> You cannot run `repair-data` operations for more than one node at the same time. To recover multiple nodes, contact technical support.

- If your grid includes replicated data, use the `repair-data start-replicated-node-repair` command with the `--nodes` option to repair the entire Storage Node.

  This command repairs the replicated data on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-node-repair --nodes SG-DC-SN3
  ```

  > As object data is restored, the **Objects Lost** alert is triggered if the StorageGRID system cannot locate replicated object data. Alerts might be triggered on Storage Nodes throughout the system. You should determine the cause of the loss and if recovery is possible. See the instructions for monitoring and troubleshooting StorageGRID.

- If your grid contains erasure coded data, use the `repair-data start-ec-node-repair` command with the `--nodes` option to repair the entire Storage Node.

  This command repairs the erasure coded data on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-node-repair --nodes SG-DC-SN3
  ```

  The operation returns a unique `repair ID` that identifies this `repair_data` operation. Use this `repair ID` to track the progress and result of the `repair_data` operation. No other feedback is returned as the recovery process completes.

  > Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available.

- If your grid has both replicated and erasure coded data, run both commands.

4. If only some of the volumes have failed, repair the affected volumes.

Enter the volume IDs in hexadecimal. For example, `0000` is the first volume and `000F` is the sixteenth volume. You can specify one volume, a range of volumes, or multiple volumes that are not in a sequence.

All the volumes must be on the same Storage Node. If you need to restore volumes for more than one Storage Node, contact technical support.

- If your grid contains replicated data, use the `start-replicated-volume-repair` command with the `--nodes` option to identify the node. Then add either the `--volumes` or `--volume-range` option, as shown in the following examples.

  **Single volume**: This command restores replicated data to volume `0002` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-volume-repair --nodes SG-DC-SN3
  --volumes 0002
  ```

  **Range of volumes**: This command restores replicated data to all volumes in the range `0003` to `0009` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-volume-repair --nodes SG-DC-SN3 --volume
  -range 0003-0009
  ```

  **Multiple volumes not in a sequence**: This command restores replicated data to volumes `0001`, `0005`, and `0008` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-volume-repair --nodes SG-DC-SN3
  --volumes 0001,0005,0008
  ```

  > As object data is restored, the **Objects Lost** alert is triggered if the StorageGRID system cannot locate replicated object data. Alerts might be triggered on Storage Nodes throughout the system. You should determine the cause of the loss and if recovery is possible. See the instructions for monitoring and troubleshooting StorageGRID.

- If your grid contains erasure coded data, use the `start-ec-volume-repair` command with the `--nodes` option to identify the node. Then add either the `--volumes` or `--volume-range` option, as shown in the following examples.

  **Single volume**: This command restores erasure coded data to volume `0007` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volumes 0007
  ```

  **Range of volumes**: This command restores erasure coded data to all volumes in the range `0004` to `0006` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volume-range
  0004-0006
  ```

  **Multiple volumes not in a sequence**: This command restores erasure coded data to volumes `000A`,

000C, and 000E on a Storage Node named SG-DC-SN3:

```
repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volumes
000A,000C,000E
```

The repair-data operation returns a unique repair ID that identifies this repair_data operation. Use this repair ID to track the progress and result of the repair_data operation. No other feedback is returned as the recovery process completes.

> Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available.

- If your grid has both replicated and erasure coded data, run both commands.

5. Monitor the repair of replicated data.
   a. Select **Nodes** > **Storage Node being repaired** > **ILM**.
   b. Use the attributes in the Evaluation section to determine if repairs are complete.

      When repairs are complete, the Awaiting - All attribute indicates 0 objects.

   c. To monitor the repair in more detail, select **Support** > **Tools** > **Grid Topology**.
   d. Select **grid** > **Storage Node being repaired** > **LDR** > **Data Store**.
   e. Use a combination of the following attributes to determine, as well as possible, if replicated repairs are complete.

      > Cassandra inconsistencies might be present, and failed repairs are not tracked.

      - **Repairs Attempted (XRPA)**: Use this attribute to track the progress of replicated repairs. This attribute increases each time a Storage Node tries to repair a high-risk object. When this attribute does not increase for a period longer than the current scan period (provided by the **Scan Period — Estimated** attribute), it means that ILM scanning found no high-risk objects that need to be repaired on any nodes.

        > High-risk objects are objects that are at risk of being completely lost. This does not include objects that do not satisfy their ILM configuration.

      - **Scan Period — Estimated (XSCM)**: Use this attribute to estimate when a policy change will be applied to previously ingested objects. If the **Repairs Attempted** attribute does not increase for a period longer than the current scan period, it is probable that replicated repairs are done. Note that the scan period can change. The **Scan Period — Estimated (XSCM)** attribute applies to the entire grid and is the maximum of all node scan periods. You can query the **Scan Period — Estimated** attribute history for the grid to determine an appropriate time frame.

6. Monitor the repair of erasure coded data, and retry any requests that might have failed.
   a. Determine the status of erasure coded data repairs:
      - Use this command to see the status of a specific repair-data operation:

```
repair-data show-ec-repair-status --repair-id repair ID
```

- Use this command to list all repairs:

```
repair-data show-ec-repair-status
```

The output lists information, including `repair ID`, for all previously and currently running repairs.

```
root@DC1-ADM1:~ # repair-data show-ec-repair-status

 Repair ID Scope  Start Time  End Time  State  Est Bytes
Affected/Repaired Retry Repair
================================================================
======================
 949283 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:27:06.9 Success
17359 17359 No
 949292 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:37:06.9 Failure
17359 0     Yes
 949294 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:47:06.9 Failure
17359 0     Yes
 949299 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:57:06.9 Failure
17359 0     Yes
```

b. If the output shows that the repair operation failed, use the `--repair-id` option to retry the repair.

This command retries a failed node repair, using the repair ID 83930030303133434:

```
repair-data start-ec-node-repair --repair-id 83930030303133434
```

This command retries a failed volume repair, using the repair ID 83930030303133434:

```
repair-data start-ec-volume-repair --repair-id 83930030303133434
```

**Related information**

Administer StorageGRID

Monitor & troubleshoot

### Checking the storage state after recovering storage volumes

After recovering storage volumes, you must verify that the desired state of the Storage Node is set to online and ensure that the state will be online by default whenever the

Storage Node server is restarted.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- The Storage Node has been recovered, and data recovery is complete.

**Steps**

1. Select **Support** > **Tools** > **Grid Topology**.
2. Check the values of **Recovered Storage Node** > **LDR** > **Storage** > **Storage State — Desired** and **Storage State — Current**.

   The value of both attributes should be Online.

3. If the Storage State — Desired is set to Read-only, complete the following steps:
   a. Click the **Configuration** tab.
   b. From the **Storage State — Desired** drop-down list, select **Online**.
   c. Click **Apply Changes**.
   d. Click the **Overview** tab and confirm that the values of **Storage State — Desired** and **Storage State — Current** are updated to Online.

#### Recovering from system drive failure

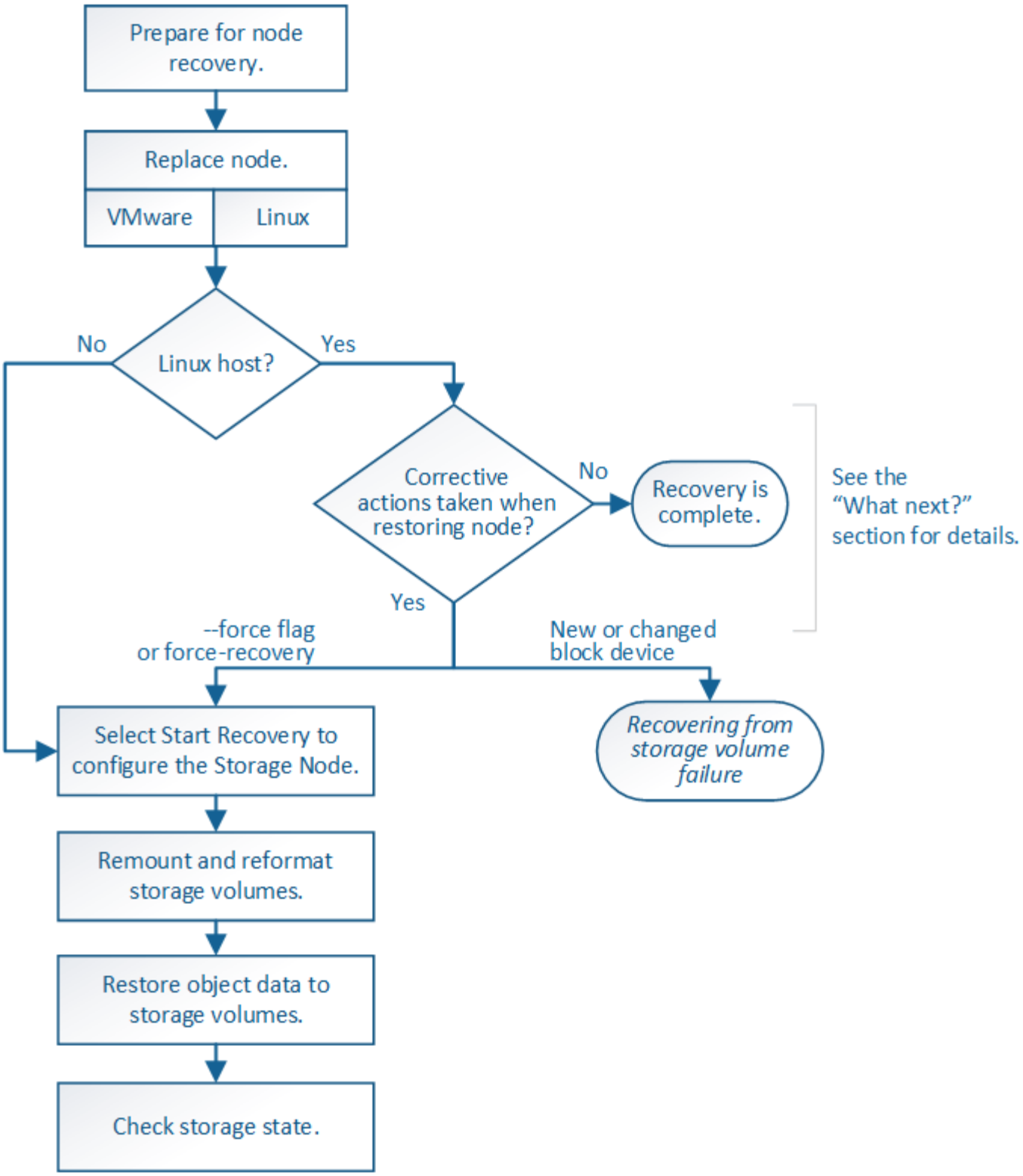
## If the system drive on a software-based Storage Node has failed, the Storage Node is not available to the StorageGRID system. You must complete a specific set of tasks to recover from a system drive failure.

**About this task**

Use this procedure to recover from a system drive failure on a software-based Storage Node. This procedure includes the steps to follow if any storage volumes also failed or cannot be remounted.

> This procedure applies to software-based Storage Nodes only. You must follow a different procedure to recover an appliance Storage Node.

Recovering a StorageGRID appliance Storage Node




**Steps**

- Reviewing warnings for Storage Node system drive recovery
- Replacing the Storage Node
- Selecting Start Recovery to configure a Storage Node
- Remounting and reformatting storage volumes ("Manual Steps")
- Restoring object data to a storage volume, if required

- Checking the storage state after recovering a Storage Node system drive

### Reviewing warnings for Storage Node system drive recovery

Before recovering a failed system drive of a Storage Node, you must review the following warnings.

Storage Nodes have a Cassandra database that includes object metadata. The Cassandra database might be rebuilt in the following circumstances:

- A Storage Node is brought back online after having been offline for more than 15 days.
- A storage volume has failed and been recovered.
- The system drive and one or more storage volumes fails and is recovered.

When Cassandra is rebuilt, the system uses information from other Storage Nodes. If too many Storage Nodes are offline, some Cassandra data might not be available. If Cassandra has been rebuilt recently, Cassandra data might not yet be consistent across the grid. Data loss can occur if Cassandra is rebuilt when too many Storage Nodes are offline or if two or more Storage Nodes are rebuilt within 15 days of each other.

If more than one Storage Node has failed (or is offline), contact technical support. Do not perform the following recovery procedure. Data loss could occur.

If this is the second Storage Node failure in less than 15 days after a Storage Node failure or recovery, contact technical support. Rebuilding Cassandra on two or more Storage Nodes within 15 days can result in data loss.

If more than one Storage Node at a site has failed, a site recovery procedure might be required. Contact technical support.

How site recovery is performed by technical support

If this Storage Node is in read-only maintenance mode to allow for the retrieval of objects by another Storage Node with failed storage volumes, recover volumes on the Storage Node with failed storage volumes before recovering this failed Storage Node. See the instructions for recovering from loss of storage volumes where the system drive is intact.

If ILM rules are configured to store only one replicated copy and the copy exists on a storage volume that has failed, you will not be able to recover the object.

If you encounter a Services: Status - Cassandra (SVST) alarm during recovery, see the monitoring and troubleshooting instructions to recover from the alarm by rebuilding Cassandra. After Cassandra is rebuilt, alarms should clear. If alarms do not clear, contact technical support.

**Related information**

Monitor & troubleshoot

Warnings and considerations for grid node recovery

Recovering from storage volume failure where the system drive is intact

### Replacing the Storage Node

## If the system drive has failed, you must first replace the Storage Node.

You must select the node replacement procedure for your platform. The steps to replace a node are the same for all types of grid nodes.

> This procedure applies to software-based Storage Nodes only. You must follow a different procedure to recover an appliance Storage Node.

Recovering a StorageGRID appliance Storage Node

**Linux:** If you are not sure if your system drive has failed, follow the instructions to replace the node to determine which recovery steps are required.

| Platform | Procedure |
| --- | --- |
| VMware | Replacing a VMware node |
| Linux | Replacing a Linux node |
| OpenStack | NetApp-provided virtual machine disk files and scripts for OpenStack are no longer supported for recovery operations. If you need to recover a node running in an OpenStack deployment, download the files for your Linux operating system. Then, follow the procedure for replacing a Linux node. |

### Selecting Start Recovery to configure a Storage Node

## After replacing a Storage Node, you must select Start Recovery in the Grid Manager to configure the new node as a replacement for the failed node.

#### What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have deployed and configured the replacement node.
- You must know the start date of any repair jobs for erasure-coded data.
- You must have verified that the Storage Node has not been rebuilt within the last 15 days.

#### About this task

If the Storage Node is installed as a container on a Linux host, you must perform this step only if one of these is true:

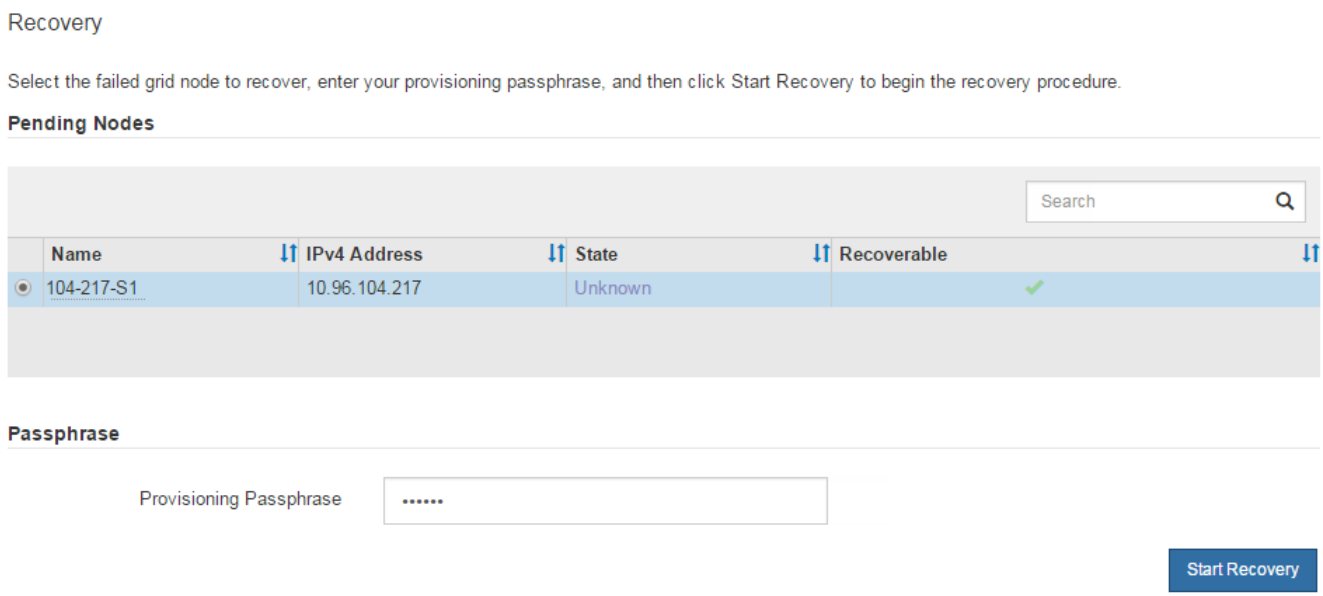
- You had to use the `--force` flag to import the node, or you issued `storagegrid node force-recovery` *`node-name`*
- You had to do a full node reinstall, or you needed to restore /var/local.

#### Steps

1. From the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Recovery**.
2. Select the grid node you want to recover in the Pending Nodes list.

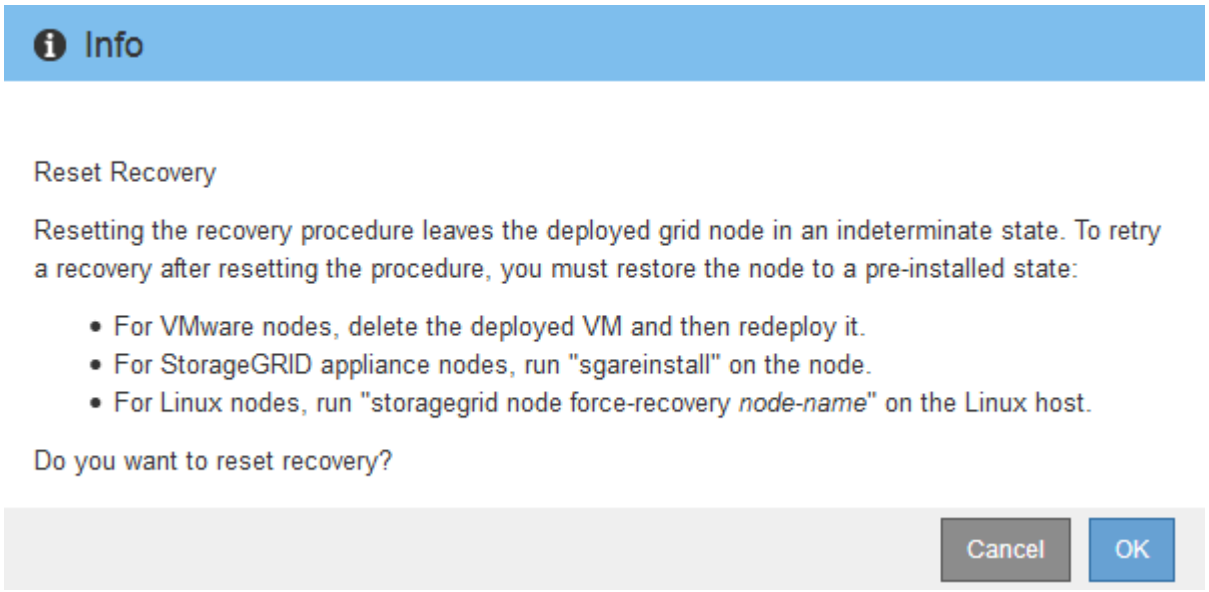
   Nodes appear in the list after they fail, but you cannot select a node until it has been reinstalled and is ready for recovery.

3. Enter the **Provisioning Passphrase**.
4. Click **Start Recovery**.




5. Monitor the progress of the recovery in the Recovering Grid Node table.

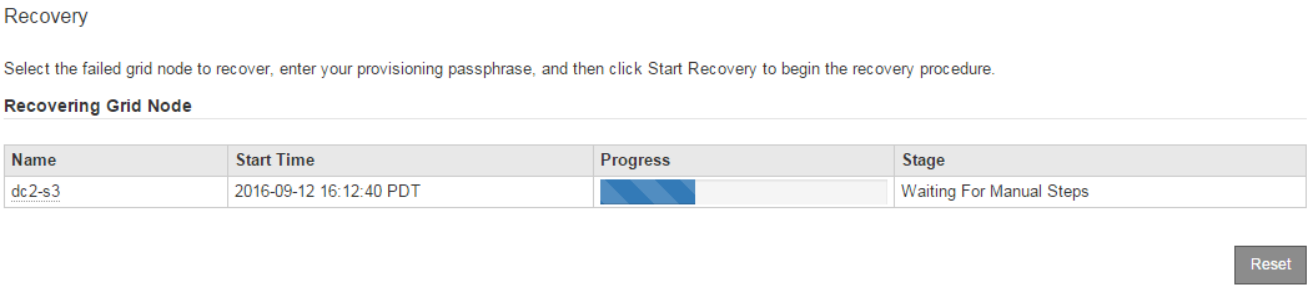
> While the recovery procedure is running, you can click **Reset** to start a new recovery. An Info dialog box appears, indicating that the node will be left in an indeterminate state if you reset the procedure.




If you want to retry the recovery after resetting the procedure, you must restore the node to a pre-installed state, as follows:

- **VMware**: Delete the deployed virtual grid node. Then, when you are ready to restart the recovery, redeploy the node.
- **Linux**: Restart the node by running this command on the Linux host: `storagegrid node force-recovery node-name`

6. When the Storage Node reaches the stage “Waiting for Manual Steps” stage, go to the next task in the recovery procedure to remount and reformat storage volumes.




**Related information**

Preparing an appliance for reinstallation (platform replacement only)

### Remounting and reformatting storage volumes (“Manual Steps”)

You must manually run two scripts to remount preserved storage volumes and to reformat any failed storage volumes. The first script remounts volumes that are properly formatted as StorageGRID storage volumes. The second script reformats any unmounted volumes, rebuilds Cassandra, if needed, and starts services.

**What you’ll need**

- You have already replaced the hardware for any failed storage volumes that you know require replacement.

  Running the `sn-remount-volumes` script might help you identify additional failed storage volumes.

- You have checked that a Storage Node decommissioning is not in progress, or you have paused the node decommission procedure. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Decommission**.)
- You have checked that an expansion is not in progress. (In the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Expansion**.)
- You have reviewed the warnings for Storage Node system drive recovery.

  Reviewing warnings for Storage Node system drive recovery

> Contact technical support if more than one Storage Node is offline or if a Storage Node in this grid has been rebuilt in the last 15 days. Do not run the `sn-recovery-postinstall.sh` script. Rebuilding Cassandra on two or more Storage Nodes within 15 days of each other might result in data loss.

**About this task**

To complete this procedure, you perform these high-level tasks:

- Log in to the recovered Storage Node.
- Run the `sn-remount-volumes` script to remount properly formatted storage volumes. When this script runs, it does the following:
  - Mounts and unmounts each storage volume to replay the XFS journal.
  - Performs an XFS file consistency check.
  - If the file system is consistent, determines if the storage volume is a properly formatted StorageGRID storage volume.
  - If the storage volume is properly formatted, remounts the storage volume. Any existing data on the volume remains intact.
- Review the script output and resolve any issues.
- Run the `sn-recovery-postinstall.sh` script. When this script runs, it does the following.

> Do not reboot a Storage Node during recovery before running `sn-recovery-postinstall.sh` (see the step for post-install script) to reformat the failed storage volumes and restore object metadata. Rebooting the Storage Node before `sn-recovery-postinstall.sh` completes causes errors for services that attempt to start and causes StorageGRID appliance nodes to exit maintenance mode.

  - Reformats any storage volumes that the `sn-remount-volumes` script could not mount or that were found to be improperly formatted.

> If a storage volume is reformatted, any data on that volume is lost. You must perform an additional procedure to restore object data from other locations in the grid, assuming that ILM rules were configured to store more than one object copy.

  - Rebuilds the Cassandra database on the node, if needed.
  - Starts the services on the Storage Node.

**Steps**

1. Log in to the recovered Storage Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Run the first script to remount any properly formatted storage volumes.

> If all storage volumes are new and need to be formatted, or if all storage volumes have failed, you can skip this step and run the second script to reformat all unmounted storage volumes.

   a. Run the script: `sn-remount-volumes`

      This script might take hours to run on storage volumes that contain data.

b. As the script runs, review the output and answer any prompts.

> As required, you can use the `tail -f` command to monitor the contents of the script's log file (`/var/local/log/sn-remount-volumes.log`) . The log file contains more detailed information than the command line output.

```
root@SG:~ # sn-remount-volumes
The configured LDR noid is 12632740

====== Device /dev/sdb ======
Mount and unmount device /dev/sdb and checking file system
consistency:
The device is consistent.
Check rangedb structure on device /dev/sdb:
Mount device /dev/sdb to /tmp/sdb-654321 with rangedb mount options
This device has all rangedb directories.
Found LDR node id 12632740, volume number 0 in the volID file
Attempting to remount /dev/sdb
Device /dev/sdb remounted successfully

====== Device /dev/sdc ======
Mount and unmount device /dev/sdc and checking file system
consistency:
Error: File system consistency check retry failed on device /dev/sdc.
You can see the diagnosis information in the /var/local/log/sn-
remount-volumes.log.

This volume could be new or damaged. If you run sn-recovery-
postinstall.sh,
this volume and any data on this volume will be deleted. If you only
had two
copies of object data, you will temporarily have only a single copy.
StorageGRID Webscale will attempt to restore data redundancy by
making
additional replicated copies or EC fragments, according to the rules
in
the active ILM policy.

Do not continue to the next step if you believe that the data
remaining on
this volume cannot be rebuilt from elsewhere in the grid (for
example, if
your ILM policy uses a rule that makes only one copy or if volumes
have
failed on multiple nodes). Instead, contact support to determine how
to
```

```
recover your data.

====== Device /dev/sdd ======
Mount and unmount device /dev/sdd and checking file system
consistency:
Failed to mount device /dev/sdd
This device could be an uninitialized disk or has corrupted
superblock.
File system check might take a long time. Do you want to continue? (y
or n) [y/N]? y

Error: File system consistency check retry failed on device /dev/sdd.
You can see the diagnosis information in the /var/local/log/sn-
remount-volumes.log.

This volume could be new or damaged. If you run sn-recovery-
postinstall.sh,
this volume and any data on this volume will be deleted. If you only
had two
copies of object data, you will temporarily have only a single copy.
StorageGRID Webscale will attempt to restore data redundancy by
making
additional replicated copies or EC fragments, according to the rules
in
the active ILM policy.

Do not continue to the next step if you believe that the data
remaining on
this volume cannot be rebuilt from elsewhere in the grid (for
example, if
your ILM policy uses a rule that makes only one copy or if volumes
have
failed on multiple nodes). Instead, contact support to determine how
to
recover your data.

====== Device /dev/sde ======
Mount and unmount device /dev/sde and checking file system
consistency:
The device is consistent.
Check rangedb structure on device /dev/sde:
Mount device /dev/sde to /tmp/sde-654321 with rangedb mount options
This device has all rangedb directories.
Found LDR node id 12000078, volume number 9 in the volID file
Error: This volume does not belong to this node. Fix the attached
volume and re-run this script.
```

In the example output, one storage volume was remounted successfully and three storage volumes had errors.

- `/dev/sdb` passed the XFS file system consistency check and had a valid volume structure, so it was remounted successfully. Data on devices that are remounted by the script is preserved.
- `/dev/sdc` failed the XFS file system consistency check because the storage volume was new or corrupt.
- `/dev/sdd` could not be mounted because the disk was uninitialized or the disk's superblock was corrupted. When the script cannot mount a storage volume, it asks if you want to run the file system consistency check.
  - If the storage volume is attached to a new disk, answer **N** to the prompt. You do not need check the file system on a new disk.
  - If the storage volume is attached to an existing disk, answer **Y** to the prompt. You can use the results of the file system check to determine the source of the corruption. The results are saved in the `/var/local/log/sn-remount-volumes.log` log file.
- `/dev/sde` passed the XFS file system consistency check and had a valid volume structure; however, the LDR node ID in the volID file did not match the ID for this Storage Node (the `configured LDR noid` displayed at the top). This message indicates that this volume belongs to another Storage Node.

3. Review the script output and resolve any issues.

> If a storage volume failed the XFS file system consistency check or could not be mounted, carefully review the error messages in the output. You must understand the implications of running the `sn-recovery-postinstall.sh` script on these volumes.

a. Check to make sure that the results include an entry for all of the volumes you expected. If any volumes are not listed, rerun the script.

b. Review the messages for all mounted devices. Make sure there are no errors indicating that a storage volume does not belong to this Storage Node.

In the example, the output for `/dev/sde` includes the following error message:

```
Error: This volume does not belong to this node. Fix the attached
volume and re-run this script.
```

> If a storage volume is reported as belonging to another Storage Node, contact technical support. If you run the `sn-recovery-postinstall.sh` script, the storage volume will be reformatted, which might cause data loss.

c. If any storage devices could not be mounted, make a note of the device name, and repair or replace the device.

> You must repair or replace any storage devices that could not be mounted.

You will use the device name to look up the volume ID, which is required input when you run the `repair-data` script to restore object data to the volume (the next procedure).

d. After repairing or replacing all unmountable devices, run the `sn-remount-volumes` script again to confirm that all storage volumes that can be remounted have been remounted.

> If a storage volume cannot be mounted or is improperly formatted, and you continue to the next step, the volume and any data on the volume will be deleted. If you had two copies of object data, you will have only a single copy until you complete the next procedure (restoring object data).

> Do not run the `sn-recovery-postinstall.sh` script if you believe that the data remaining on a failed storage volume cannot be rebuilt from elsewhere in the grid (for example, if your ILM policy uses a rule that makes only one copy or if volumes have failed on multiple nodes). Instead, contact technical support to determine how to recover your data.

4. Run the `sn-recovery-postinstall.sh` script: `sn-recovery-postinstall.sh`

   This script reformats any storage volumes that could not be mounted or that were found to be improperly formatted; rebuilds the Cassandra database on the node, if needed; and starts the services on the Storage Node.

   Be aware of the following:

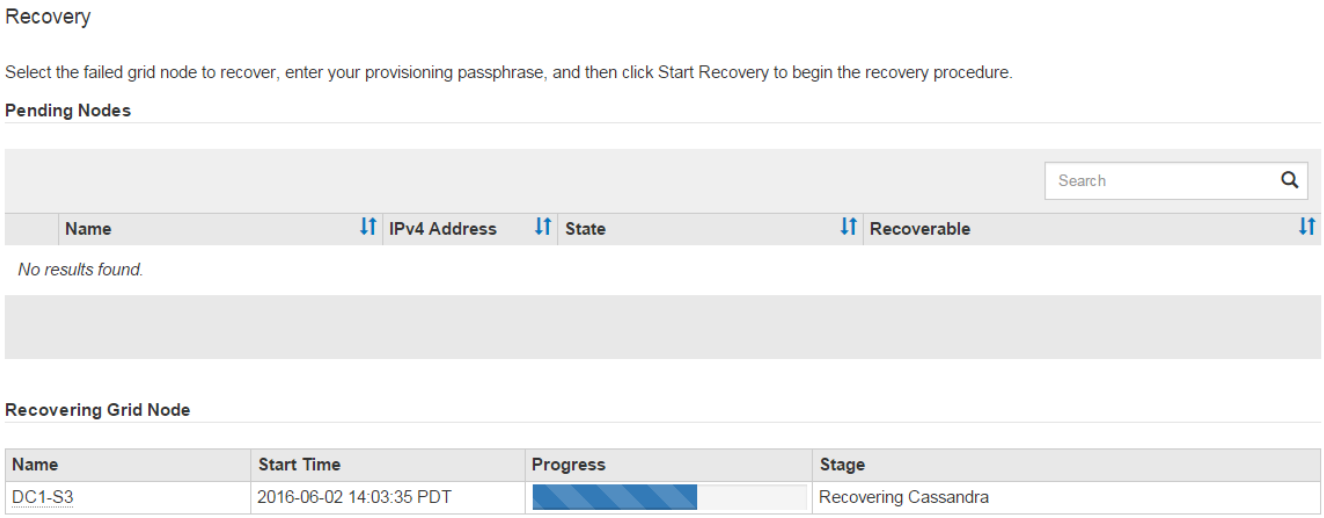
   - The script might take hours to run.
   - In general, you should leave the SSH session alone while the script is running.
   - Do not press **Ctrl+C** while the SSH session is active.
   - The script will run in the background if a network disruption occurs and terminates the SSH session, but you can view the progress from the Recovery page.
   - If the Storage Node uses the RSM service, the script might appear to stall for 5 minutes as node services are restarted. This 5-minute delay is expected whenever the RSM service boots for the first time.

     > The RSM service is present on Storage Nodes that include the ADC service.

   > Some StorageGRID recovery procedures use Reaper to handle Cassandra repairs. Repairs occur automatically as soon as the related or required services have started. You might notice script output that mentions “reaper” or “Cassandra repair.” If you see an error message indicating the repair has failed, run the command indicated in the error message.

5. As the `sn-recovery-postinstall.sh` script runs, monitor the Recovery page in the Grid Manager.

   The Progress bar and the Stage column on the Recovery page provide a high-level status of the `sn-recovery-postinstall.sh` script.




After the `sn-recovery-postinstall.sh` script has started services on the node, you can restore object data to any storage volumes that were formatted by the script, as described in that procedure.

**Related information**

Reviewing warnings for Storage Node system drive recovery

Restoring object data to a storage volume, if required

### Restoring object data to a storage volume, if required

If the `sn-recovery-postinstall.sh` script is needed to reformat one or more failed storage volumes, you must restore object data to the reformatted storage volume from other Storage Nodes and Archive Nodes. These steps are not required unless one or more storage volumes were reformatted.

**What you'll need**

- You must have confirmed that the recovered Storage Node has a Connection State of **Connected*** ✔ **on the *Nodes** > **Overview** tab in the Grid Manager.

**About this task**

Object data can be restored from other Storage Nodes, an Archive Node, or a Cloud Storage Pool, assuming that the grid's ILM rules were configured such that object copies are available.

ⓘ If an ILM rule was configured to store only one replicated copy and that copy existed on a storage volume that failed, you will not be able to recover the object.

ⓘ If the only remaining copy of an object is in a Cloud Storage Pool, StorageGRID must issue multiple requests to the Cloud Storage Pool endpoint to restore object data. Before performing this procedure, contact technical support for help in estimating the recovery time frame and the associated costs.

If the only remaining copy of an object is on an Archive Node, object data is retrieved from the Archive Node. Due to the latency associated with retrievals from external archival storage systems, restoring object data to a Storage Node from an Archive Node takes longer than restoring copies from other Storage Nodes.

To restore object data, you run the `repair-data` script. This script begins the process of restoring object data and works with ILM scanning to ensure that ILM rules are met. You use different options with the `repair-data` script, based on whether you are restoring replicated data or erasure coded data, as follows:

- **Replicated data**: Two commands are available for restoring replicated data, based on whether you need to repair the entire node or only certain volumes on the node:

```
repair-data start-replicated-node-repair
```

```
repair-data start-replicated-volume-repair
```

- **Erasure coded (EC) data**: Two commands are available for restoring erasure coded data, based on whether you need to repair the entire node or only certain volumes on the node:

```
repair-data start-ec-node-repair
```

```
repair-data start-ec-volume-repair
```

  Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available. You can track repairs of erasure coded data with this command:

```
repair-data show-ec-repair-status
```

The EC repair job temporarily reserves a large amount of storage. Storage alerts might be triggered, but will resolve when the repair is complete. If there is not enough storage for the reservation, the EC repair job will fail. Storage reservations are released when the EC repair job completes, whether the job failed or succeeded.

For more information on using the `repair-data` script, enter `repair-data --help` from the command line of the primary Admin Node.

**Steps**

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

2. Use the `/etc/hosts` file to find the hostname of the Storage Node for the restored storage volumes. To see a list of all nodes in the grid, enter the following: `cat /etc/hosts`
3. If all storage volumes have failed, repair the entire node. (If only some volumes have failed, go to the next step.)

> You cannot run `repair-data` operations for more than one node at the same time. To recover multiple nodes, contact technical support.

- If your grid includes replicated data, use the `repair-data start-replicated-node-repair` command with the `--nodes` option to repair the entire Storage Node.

  This command repairs the replicated data on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-node-repair --nodes SG-DC-SN3
  ```

  > As object data is restored, the **Objects Lost** alert is triggered if the StorageGRID system cannot locate replicated object data. Alerts might be triggered on Storage Nodes throughout the system. You should determine the cause of the loss and if recovery is possible. See the instructions for monitoring and troubleshooting StorageGRID.

- If your grid contains erasure coded data, use the `repair-data start-ec-node-repair` command with the `--nodes` option to repair the entire Storage Node.

  This command repairs the erasure coded data on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-node-repair --nodes SG-DC-SN3
  ```

  The operation returns a unique `repair ID` that identifies this `repair_data` operation. Use this `repair ID` to track the progress and result of the `repair_data` operation. No other feedback is returned as the recovery process completes.

  > Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available.

- If your grid has both replicated and erasure coded data, run both commands.

4. If only some of the volumes have failed, repair the affected volumes.

Enter the volume IDs in hexadecimal. For example, `0000` is the first volume and `000F` is the sixteenth volume. You can specify one volume, a range of volumes, or multiple volumes that are not in a sequence.

All the volumes must be on the same Storage Node. If you need to restore volumes for more than one Storage Node, contact technical support.

- If your grid contains replicated data, use the `start-replicated-volume-repair` command with the `--nodes` option to identify the node. Then add either the `--volumes` or `--volume-range` option, as shown in the following examples.

  **Single volume**: This command restores replicated data to volume `0002` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-volume-repair --nodes SG-DC-SN3
  --volumes 0002
  ```

  **Range of volumes**: This command restores replicated data to all volumes in the range `0003` to `0009` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-volume-repair --nodes SG-DC-SN3 --volume
  -range 0003-0009
  ```

  **Multiple volumes not in a sequence**: This command restores replicated data to volumes `0001`, `0005`, and `0008` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-replicated-volume-repair --nodes SG-DC-SN3
  --volumes 0001,0005,0008
  ```

  > As object data is restored, the **Objects Lost** alert is triggered if the StorageGRID system cannot locate replicated object data. Alerts might be triggered on Storage Nodes throughout the system. You should determine the cause of the loss and if recovery is possible. See the instructions for monitoring and troubleshooting StorageGRID.

- If your grid contains erasure coded data, use the `start-ec-volume-repair` command with the `--nodes` option to identify the node. Then add either the `--volumes` or `--volume-range` option, as shown in the following examples.

  **Single volume**: This command restores erasure coded data to volume `0007` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volumes 0007
  ```

  **Range of volumes**: This command restores erasure coded data to all volumes in the range `0004` to `0006` on a Storage Node named SG-DC-SN3:

  ```
  repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volume-range
  0004-0006
  ```

  **Multiple volumes not in a sequence**: This command restores erasure coded data to volumes `000A`,

000C, and 000E on a Storage Node named SG-DC-SN3:

```
repair-data start-ec-volume-repair --nodes SG-DC-SN3 --volumes
000A,000C,000E
```

The repair-data operation returns a unique repair ID that identifies this repair_data operation. Use this repair ID to track the progress and result of the repair_data operation. No other feedback is returned as the recovery process completes.

> Repairs of erasure coded data can begin while some Storage Nodes are offline. Repair will complete after all nodes are available.

- If your grid has both replicated and erasure coded data, run both commands.

5. Monitor the repair of replicated data.
   a. Select **Nodes** > **Storage Node being repaired** > **ILM**.
   b. Use the attributes in the Evaluation section to determine if repairs are complete.

      When repairs are complete, the Awaiting - All attribute indicates 0 objects.

   c. To monitor the repair in more detail, select **Support** > **Tools** > **Grid Topology**.
   d. Select **grid** > **Storage Node being repaired** > **LDR** > **Data Store**.
   e. Use a combination of the following attributes to determine, as well as possible, if replicated repairs are complete.

      > Cassandra inconsistencies might be present, and failed repairs are not tracked.

      - **Repairs Attempted (XRPA)**: Use this attribute to track the progress of replicated repairs. This attribute increases each time a Storage Node tries to repair a high-risk object. When this attribute does not increase for a period longer than the current scan period (provided by the **Scan Period — Estimated** attribute), it means that ILM scanning found no high-risk objects that need to be repaired on any nodes.

        > High-risk objects are objects that are at risk of being completely lost. This does not include objects that do not satisfy their ILM configuration.

      - **Scan Period — Estimated (XSCM)**: Use this attribute to estimate when a policy change will be applied to previously ingested objects. If the **Repairs Attempted** attribute does not increase for a period longer than the current scan period, it is probable that replicated repairs are done. Note that the scan period can change. The **Scan Period — Estimated (XSCM)** attribute applies to the entire grid and is the maximum of all node scan periods. You can query the **Scan Period — Estimated** attribute history for the grid to determine an appropriate time frame.

6. Monitor the repair of erasure coded data, and retry any requests that might have failed.
   a. Determine the status of erasure coded data repairs:
      - Use this command to see the status of a specific repair-data operation:

```
repair-data show-ec-repair-status --repair-id repair ID
```

- Use this command to list all repairs:

```
repair-data show-ec-repair-status
```

The output lists information, including `repair ID`, for all previously and currently running repairs.

```
root@DC1-ADM1:~ # repair-data show-ec-repair-status

Repair ID Scope Start Time End Time State Est Bytes
Affected/Repaired Retry Repair
================================================================================
================
949283 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:27:06.9 Success
17359 17359 No
949292 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:37:06.9 Failure
17359 0 Yes
949294 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:47:06.9 Failure
17359 0 Yes
949299 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:57:06.9 Failure
17359 0 Yes
```

b. If the output shows that the repair operation failed, use the `--repair-id` option to retry the repair.

This command retries a failed node repair, using the repair ID 83930030303133434:

```
repair-data start-ec-node-repair --repair-id 83930030303133434
```

This command retries a failed volume repair, using the repair ID 83930030303133434:

```
repair-data start-ec-volume-repair --repair-id 83930030303133434
```

**Related information**

Administer StorageGRID

Monitor & troubleshoot

#### Checking the storage state after recovering a Storage Node system drive

After recovering the system drive for a Storage Node, you must verify that the desired state of the Storage Node is set to online and ensure that the state will be online by

default whenever the Storage Node server is restarted.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- The Storage Node has been recovered, and data recovery is complete.

**Steps**

1. Select **Support** > **Tools** > **Grid Topology**.
2. Check the values of **Recovered Storage Node** > **LDR** > **Storage** > **Storage State — Desired** and **Storage State — Current**.

   The value of both attributes should be Online.

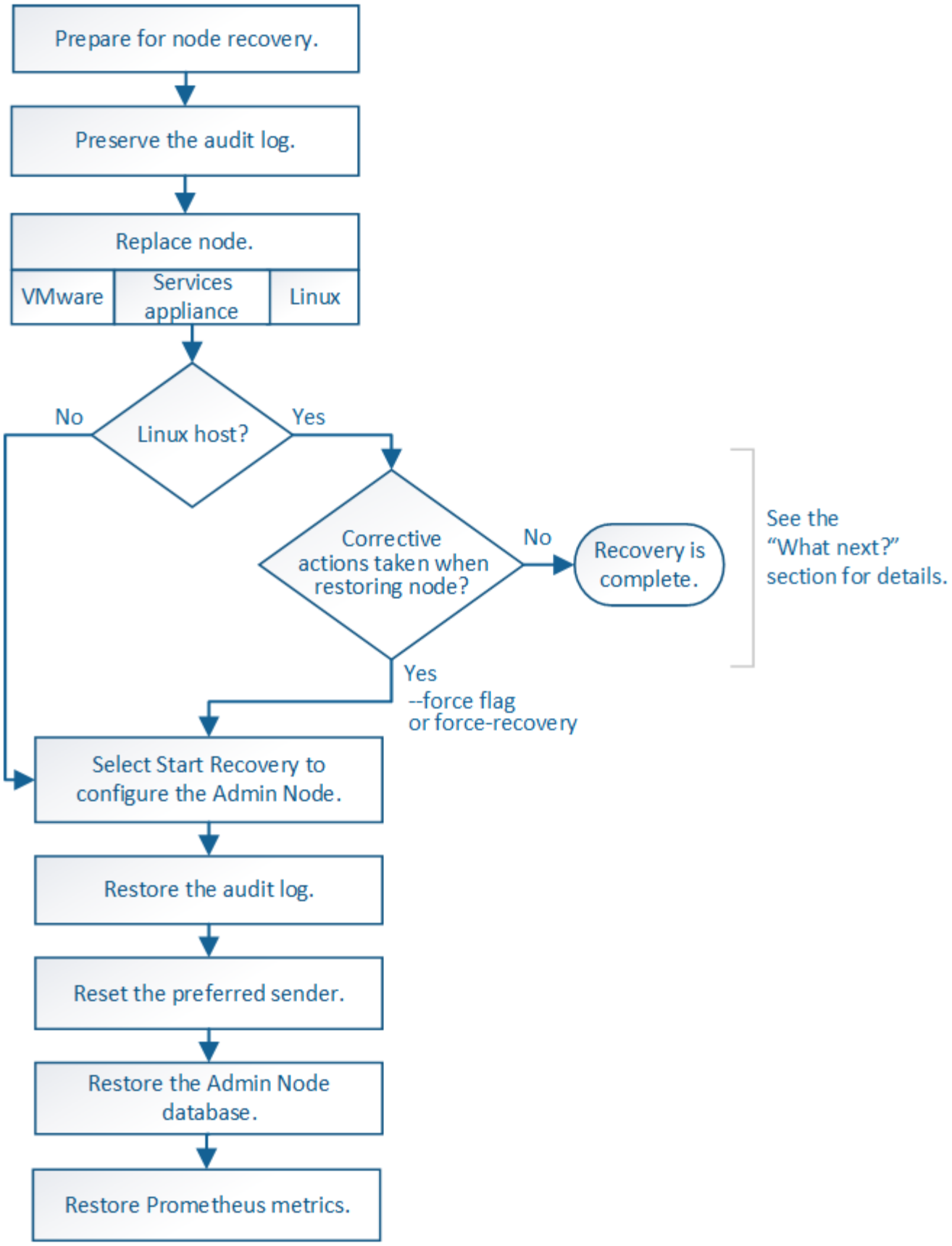
3. If the Storage State — Desired is set to Read-only, complete the following steps:
   a. Click the **Configuration** tab.
   b. From the **Storage State — Desired** drop-down list, select **Online**.
   c. Click **Apply Changes**.
   d. Click the **Overview** tab and confirm that the values of **Storage State — Desired** and **Storage State — Current** are updated to Online.

### Recovering from Admin Node failures

The recovery process for an Admin Node depends on whether it is the primary Admin Node or a non-primary Admin Node.

**About this task**

The high-level steps for recovering a primary or non-primary Admin Node are the same, although the details of the steps differ.




Always follow the correct recovery procedure for the Admin Node you are recovering. The procedures look the same at a high level, but differ in the details.

**Related information**

SG100 & SG1000 services appliances

**Choices**

- Recovering from primary Admin Node failures
- Recovering from non-primary Admin Node failures

#### Recovering from primary Admin Node failures

You must complete a specific set of tasks to recover from a primary Admin Node failure. The primary Admin Node hosts the Configuration Management Node (CMN) service for the grid.

**About this task**

A failed primary Admin Node should be replaced promptly. The Configuration Management Node (CMN) service on the primary Admin Node is responsible for issuing blocks of object identifiers for the grid. These identifiers are assigned to objects as they are ingested. New objects cannot be ingested unless there are identifiers available. Object ingest can continue while the CMN is unavailable because approximately one month's supply of identifiers is cached in the grid. However, after cached identifiers are exhausted, no new objects can be added.

> ⓘ You must repair or replace a failed primary Admin Node within approximately a month or the grid might lose its ability to ingest new objects. The exact time period depends on your rate of object ingest: if you need a more accurate assessment of the time frame for your grid, contact technical support.

**Steps**

- Copying audit logs from the failed primary Admin Node
- Replacing the primary Admin Node
- Configuring the replacement primary Admin Node
- Restoring the audit log on the recovered primary Admin Node
- Resetting the preferred sender on the recovered primary Admin Node
- Restoring the Admin Node database when recovering a primary Admin Node
- Restoring Prometheus metrics when recovering a primary Admin Node

#### Copying audit logs from the failed primary Admin Node

If you are able to copy audit logs from the failed primary Admin Node, you should preserve them to maintain the grid's record of system activity and usage. You can restore the preserved audit logs to the recovered primary Admin Node after it is up and running.

This procedure copies the audit log files from the failed Admin Node to a temporary location on a separate grid node. These preserved audit logs can then be copied to the replacement Admin Node. Audit logs are not automatically copied to the new Admin Node.

Depending on the type of failure, you might not be able to copy audit logs from a failed Admin Node. If the deployment has only one Admin Node, the recovered Admin Node starts recording events to the audit log in a new empty file and previously recorded data is lost. If the deployment includes more than one Admin Node, you can recover the audit logs from another Admin Node.

If the audit logs are not accessible on the failed Admin Node now, you might be able to access them later, for example, after host recovery.

1. Log in to the failed Admin Node if possible. Otherwise, log in to the primary Admin Node or another Admin Node, if available.
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
2. Stop the AMS service to prevent it from creating a new log file:`service ams stop`
3. Rename the audit.log file so that it does not overwrite the existing file when you copy it to the recovered Admin Node.

   Rename audit.log to a unique numbered file name such as yyyy-mm-dd.txt.1. For example, you can rename the audit.log file to 2015-10-25.txt.1`cd /var/local/audit/exportls -l``mv audit.log 2015-10-25.txt.1`
4. Restart the AMS service: `service ams start`
5. Create the directory to copy all audit log files to a temporary location on a separate grid node: `ssh admin@grid_node_IP mkdir -p /var/local/tmp/saved-audit-logs`

   When prompted, enter the password for admin.
6. Copy all audit log files: `scp -p * admin@grid_node_IP:/var/local/tmp/saved-audit-logs`

   When prompted, enter the password for admin.
7. Log out as root: `exit`

**Replacing the primary Admin Node**

## To recover a primary Admin Node, you must first replace the physical or virtual hardware.

You can replace a failed primary Admin Node with a primary Admin Node running on the same platform, or you can replace a primary Admin Node running on VMware or a Linux host with a primary Admin Node hosted on a services appliance.

Use the procedure that matches the replacement platform you select for the node. After you complete the node replacement procedure (which is suitable for all node types), that procedure will direct you to the next step for primary Admin Node recovery.

| Replacement platform | Procedure |
| --- | --- |
| VMware | Replacing a VMware node |

| Replacement platform | Procedure |
| --- | --- |
| Linux | Replacing a Linux node |
| SG100 and SG1000 services appliances | Replacing a services appliance |
| OpenStack | NetApp-provided virtual machine disk files and scripts for OpenStack are no longer supported for recovery operations. If you need to recover a node running in an OpenStack deployment, download the files for your Linux operating system. Then, follow the procedure for replacing a Linux node. |

#### Configuring the replacement primary Admin Node

## The replacement node must be configured as the primary Admin Node for your StorageGRID system.

#### What you'll need

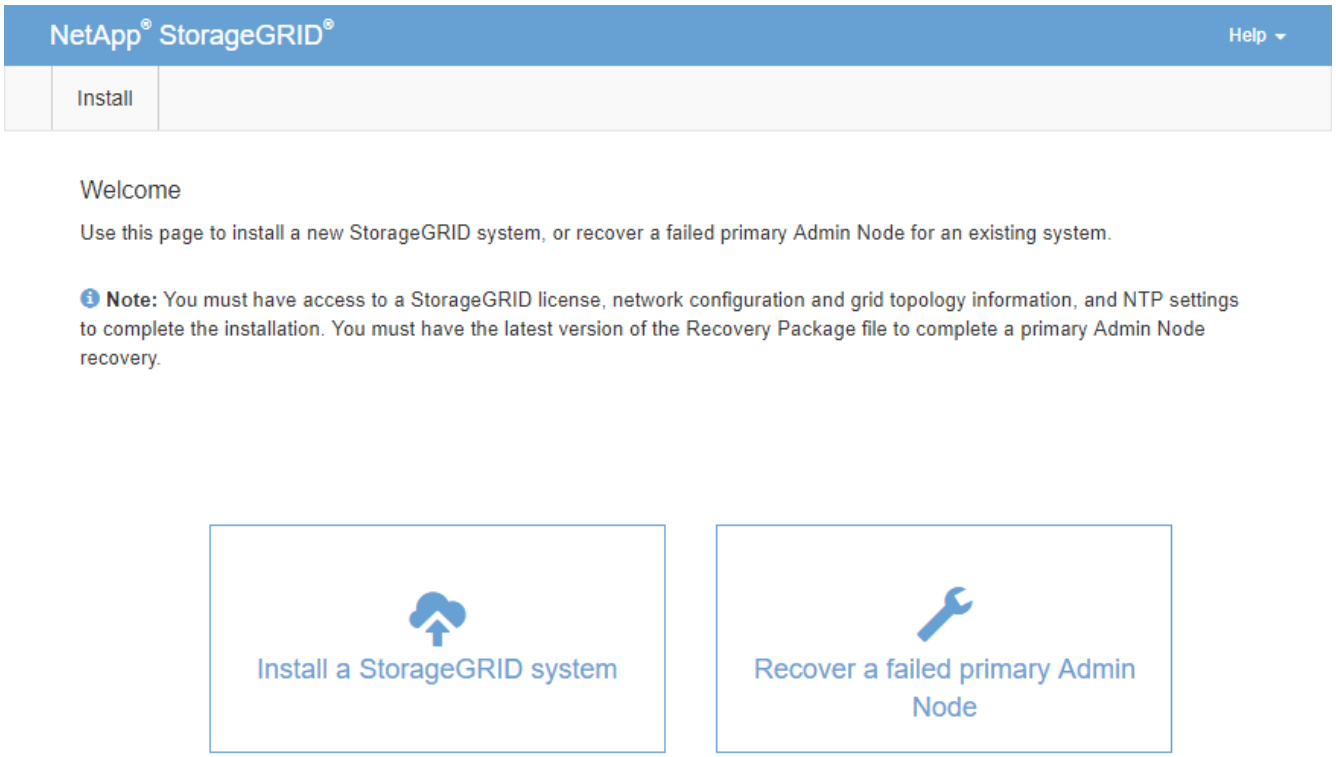
- For primary Admin Nodes hosted on virtual machines, the virtual machine must be deployed, powered on, and initialized.
- For primary Admin Nodes hosted on a services appliance, you have replaced the appliance and have installed software. See the installation guide for your appliance.

  SG100 & SG1000 services appliances

- You must have the latest backup of the Recovery Package file (`sgws-recovery-package-id-revision.zip`).
- You must have the provisioning passphrase.

#### Steps

1. Open your web browser and navigate to `https://primary_admin_node_ip`.




2. Click **Recover a failed primary Admin Node**.
3. Upload the most recent backup of the Recovery Package:
   a. Click **Browse**.
   b. Locate the most recent Recovery Package file for your StorageGRID system, and click **Open**.
4. Enter the provisioning passphrase.
5. Click **Start Recovery**.

   The recovery process begins. The Grid Manager might become unavailable for a few minutes as the required services start. When the recovery is complete, the sign in page is displayed.

6. If single sign-on (SSO) is enabled for your StorageGRID system and the relying party trust for the Admin Node you recovered was configured to use the default Management Interface Server Certificate, update (or delete and recreate) the node's relying party trust in Active Directory Federation Services (AD FS). Use the new default server certificate that was generated during the Admin Node recovery process.

   To configure a relying party trust, see the instructions for administering StorageGRID. To access the default server certificate, log in to the command shell of the Admin Node. Go to the `/var/local/mgmt-api` directory, and select the `server.crt` file.

7. Determine if you need to apply a hotfix.
   a. Sign in to the Grid Manager using a supported browser.
   b. Select **Nodes**.
   c. From the list on the left, select the primary Admin Node.
   d. On the Overview tab, note the version displayed in the **Software Version** field.
   e. Select any other grid node.

f. On the Overview tab, note the version displayed in the **Software Version** field.

- If the versions displayed in the **Software Version** fields are the same, you do not need to apply a hotfix.
- If the versions displayed in the **Software Version** fields are different, you must apply a hotfix to update the recovered primary Admin Node to the same version.

**Related information**

Administer StorageGRID

StorageGRID hotfix procedure

### Restoring the audit log on the recovered primary Admin Node

If you were able to preserve the audit log from the failed primary Admin Node, you can copy it to the primary Admin Node you are recovering.

- The recovered Admin Node must be installed and running.
- You must have copied the audit logs to another location after the original Admin Node failed.

If an Admin Node fails, audit logs saved to that Admin Node are potentially lost. It might be possible to preserve data from loss by copying audit logs from the failed Admin Node and then restoring these audit logs to the recovered Admin Node. Depending on the failure, it might not be possible to copy audit logs from the failed Admin Node. In that case, if the deployment has more than one Admin Node, you can recover audit logs from another Admin Node as audit logs are replicated to all Admin Nodes.

If there is only one Admin Node and the audit log cannot be copied from the failed node, the recovered Admin Node starts recording events to the audit log as if the installation is new.

You must recover an Admin Node as soon as possible to restore logging functionality.

1. Log in to the recovered Admin Node:

   a. Enter the following command: `ssh admin@recovery_Admin_Node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   After you are logged in as root, the prompt changes from `$` to `#`.

2. Check which audit files have been preserved: `cd /var/local/audit/export`

3. Copy the preserved audit log files to the recovered Admin Node: `scp admin@grid_node_IP:/var/local/tmp/saved-audit-logs/YYYY* .`

   When prompted, enter the password for admin.

4. For security, delete the audit logs from the failed grid node after verifying that they have been copied successfully to the recovered Admin Node.

5. Update the user and group settings of the audit log files on the recovered Admin Node: `chown ams-user:bycast *`

6. Log out as root: `exit`

You must also restore any pre-existing client access to the audit share. For more information, see the instructions for administering StorageGRID.

**Related information**

Administer StorageGRID

### Resetting the preferred sender on the recovered primary Admin Node

If the primary Admin Node you are recovering is currently set as the preferred sender of alert notifications, alarm notifications, and AutoSupport messages, you must reconfigure this setting.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.
- The recovered Admin Node must be installed and running.

**Steps**

1. Select **Configuration** > **System Settings** > **Display Options**.
2. Select the recovered Admin Node from the **Preferred Sender** drop-down list.
3. Click **Apply Changes**.

**Related information**

Administer StorageGRID

### Restoring the Admin Node database when recovering a primary Admin Node

If you want to retain the historical information about attributes, alarms, and alerts on a primary Admin Node that has failed, you can restore the Admin Node database. You can only restore this database if your StorageGRID system includes another Admin Node.

- The recovered Admin Node must be installed and running.
- The StorageGRID system must include at least two Admin Nodes.
- You must have the `Passwords.txt` file.
- You must have the provisioning passphrase.

If an Admin Node fails, the historical information stored in its Admin Node database is lost. This database includes the following information:

- Alert history
- Alarm history
- Historical attribute data, which is used in the charts and text reports available from the **Support** > **Tools** > **Grid Topology** page.

When you recover an Admin Node, the software installation process creates an empty Admin Node database on the recovered node. However, the new database only includes information for servers and services that are

currently part of the system or added later.

If you restored a primary Admin Node and your StorageGRID system has another Admin Node, you can restore the historical information by copying the Admin Node database from a non-primary Admin Node (the *source Admin Node*) to the recovered primary Admin Node. If your system has only a primary Admin Node, you cannot restore the Admin Node database.

> Copying the Admin Node database might take several hours. Some Grid Manager features will be unavailable while services are stopped on the source Admin Node.

1. Log in to the source Admin Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.
2. From the source Admin Node, stop the MI service: `service mi stop`
3. From the source Admin Node, stop the Management Application Program Interface (mgmt-api) service: `service mgmt-api stop`
4. Complete the following steps on the recovered Admin Node:
   a. Log in to the recovered Admin Node:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.
   b. Stop the MI service: `service mi stop`
   c. Stop the mgmt-api service: `service mgmt-api stop`
   d. Add the SSH private key to the SSH agent. Enter:`ssh-add`
   e. Enter the SSH Access Password listed in the `Passwords.txt` file.
   f. Copy the database from the source Admin Node to the recovered Admin Node: `/usr/local/mi/bin/mi-clone-db.sh Source_Admin_Node_IP`
   g. When prompted, confirm that you want to overwrite the MI database on the recovered Admin Node.

      The database and its historical data are copied to the recovered Admin Node. When the copy operation is done, the script starts the recovered Admin Node.

   h. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:`ssh-add -D`
5. Restart the services on the source Admin Node: `service servermanager start`

#### Restoring Prometheus metrics when recovering a primary Admin Node

Optionally, you can retain the historical metrics maintained by Prometheus on a primary

## Admin Node that has failed. The Prometheus metrics can only be restored if your StorageGRID system includes another Admin Node.

- The recovered Admin Node must be installed and running.
- The StorageGRID system must include at least two Admin Nodes.
- You must have the `Passwords.txt` file.
- You must have the provisioning passphrase.

If an Admin Node fails, the metrics maintained in the Prometheus database on the Admin Node are lost. When you recover the Admin Node, the software installation process creates a new Prometheus database. After the recovered Admin Node is started, it records metrics as if you had performed a new installation of the StorageGRID system.

If you restored a primary Admin Node and your StorageGRID system has another Admin Node, you can restore the historical metrics by copying the Prometheus database from a non-primary Admin Node (the *source Admin Node*) to the recovered primary Admin Node. If your system has only a primary Admin Node, you cannot restore the Prometheus database.

> Copying the Prometheus database might take an hour or more. Some Grid Manager features will be unavailable while services are stopped on the source Admin Node.

1. Log in to the source Admin Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.
2. From the source Admin Node, stop the Prometheus service: `service prometheus stop`
3. Complete the following steps on the recovered Admin Node:
   a. Log in to the recovered Admin Node:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.
   b. Stop the Prometheus service: `service prometheus stop`
   c. Add the SSH private key to the SSH agent. Enter:`ssh-add`
   d. Enter the SSH Access Password listed in the `Passwords.txt` file.
   e. Copy the Prometheus database from the source Admin Node to the recovered Admin Node: `/usr/local/prometheus/bin/prometheus-clone-db.sh Source_Admin_Node_IP`
   f. When prompted, press **Enter** to confirm that you want to destroy the new Prometheus database on the recovered Admin Node.

      The original Prometheus database and its historical data are copied to the recovered Admin Node. When the copy operation is done, the script starts the recovered Admin Node. The following status

appears:

Database cloned, starting services

g. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:`ssh-add -D`

4. Restart the Prometheus service on the source Admin Node.`service prometheus start`

#### Recovering from non-primary Admin Node failures

You must complete the following tasks to recover from a non-primary Admin Node failure. One Admin Node hosts the Configuration Management Node (CMN) service and is known as the primary Admin Node. Although you can have multiple Admin Nodes, each StorageGRID system includes only one primary Admin Node. All other Admin Nodes are non-primary Admin Nodes.

**Related information**

SG100 & SG1000 services appliances

**Steps**

- Copying audit logs from the failed non-primary Admin Node
- Replacing a non-primary Admin Node
- Selecting Start Recovery to configure a non-primary Admin Node
- Restoring the audit log on the recovered non-primary Admin Node
- Resetting the preferred sender on the recovered non-primary Admin Node
- Restoring the Admin Node database when recovering a non-primary Admin Node
- Restoring Prometheus metrics when recovering a non-primary Admin Node

#### Copying audit logs from the failed non-primary Admin Node

If you are able to copy audit logs from the failed Admin Node, you should preserve them to maintain the grid's record of system activity and usage. You can restore the preserved audit logs to the recovered non-primary Admin Node after it is up and running.

This procedure copies the audit log files from the failed Admin Node to a temporary location on a separate grid node. These preserved audit logs can then be copied to the replacement Admin Node. Audit logs are not automatically copied to the new Admin Node.

Depending on the type of failure, you might not be able to copy audit logs from a failed Admin Node. If the deployment has only one Admin Node, the recovered Admin Node starts recording events to the audit log in a new empty file and previously recorded data is lost. If the deployment includes more than one Admin Node, you can recover the audit logs from another Admin Node.

> If the audit logs are not accessible on the failed Admin Node now, you might be able to access them later, for example, after host recovery.

1. Log in to the failed Admin Node if possible. Otherwise, log in to the primary Admin Node or another Admin Node, if available.

a. Enter the following command: `ssh admin@grid_node_IP`

b. Enter the password listed in the `Passwords.txt` file.

c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

2. Stop the AMS service to prevent it from creating a new log file:`service ams stop`
3. Rename the audit.log file so that it does not overwrite the existing file when you copy it to the recovered Admin Node.

   Rename audit.log to a unique numbered file name such as yyyy-mm-dd.txt.1. For example, you can rename the audit.log file to 2015-10-25.txt.1`cd /var/local/audit/exportls -l``mv audit.log 2015-10-25.txt.1`
4. Restart the AMS service: `service ams start`
5. Create the directory to copy all audit log files to a temporary location on a separate grid node: `ssh admin@grid_node_IP mkdir -p /var/local/tmp/saved-audit-logs`

   When prompted, enter the password for admin.
6. Copy all audit log files: `scp -p * admin@grid_node_IP:/var/local/tmp/saved-audit-logs`

   When prompted, enter the password for admin.
7. Log out as root: `exit`

**Replacing a non-primary Admin Node**

## To recover a non-primary Admin Node, you first must replace the physical or virtual hardware.

You can replace a failed non-primary Admin Node with a non-primary Admin Node running on the same platform, or you can replace a non-primary Admin Node running on VMware or a Linux host with a non-primary Admin Node hosted on a services appliance.

Use the procedure that matches the replacement platform you select for the node. After you complete the node replacement procedure (which is suitable for all node types), that procedure will direct you to the next step for non-primary Admin Node recovery.

| Replacement platform | Procedure |
|---|---|
| VMware | Replacing a VMware node |
| Linux | Replacing a Linux node |
| SG100 and SG1000 services appliances | Replacing a services appliance |

| Replacement platform | Procedure |
| --- | --- |
| OpenStack | NetApp-provided virtual machine disk files and scripts for OpenStack are no longer supported for recovery operations. If you need to recover a node running in an OpenStack deployment, download the files for your Linux operating system. Then, follow the procedure for replacing a Linux node. |

### Selecting Start Recovery to configure a non-primary Admin Node

After replacing a non-primary Admin Node, you must select Start Recovery in the Grid Manager to configure the new node as a replacement for the failed node.

#### What you'll need

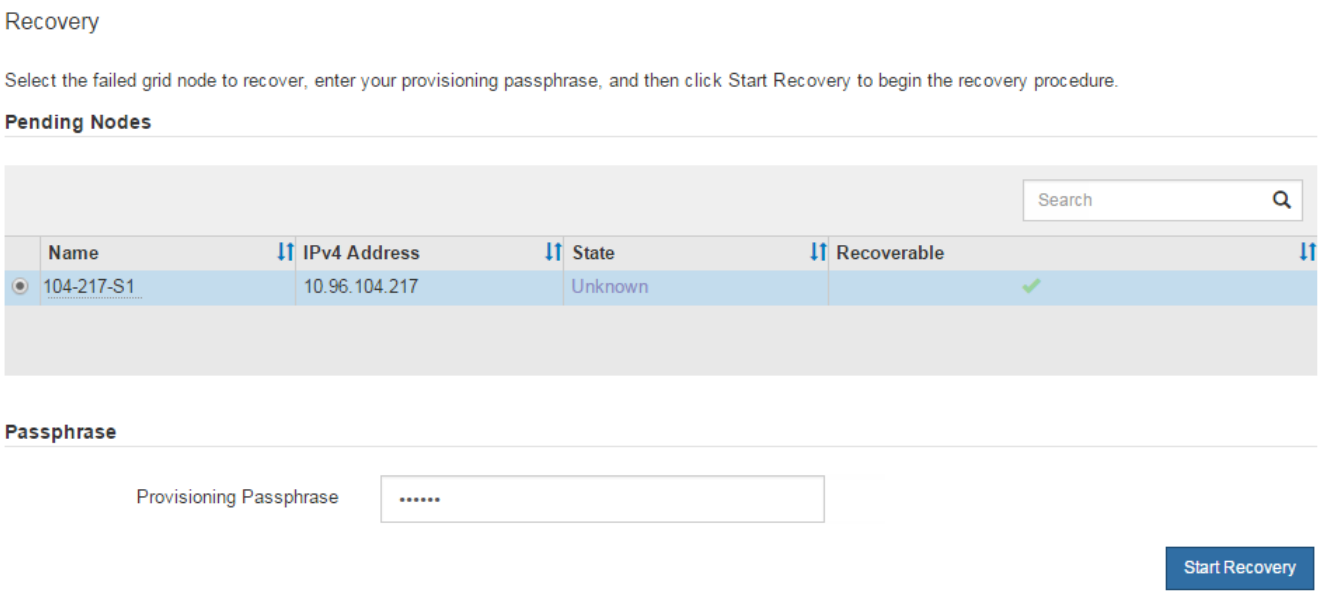
- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have deployed and configured the replacement node.

#### Steps

1. From the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Recovery**.
2. Select the grid node you want to recover in the Pending Nodes list.

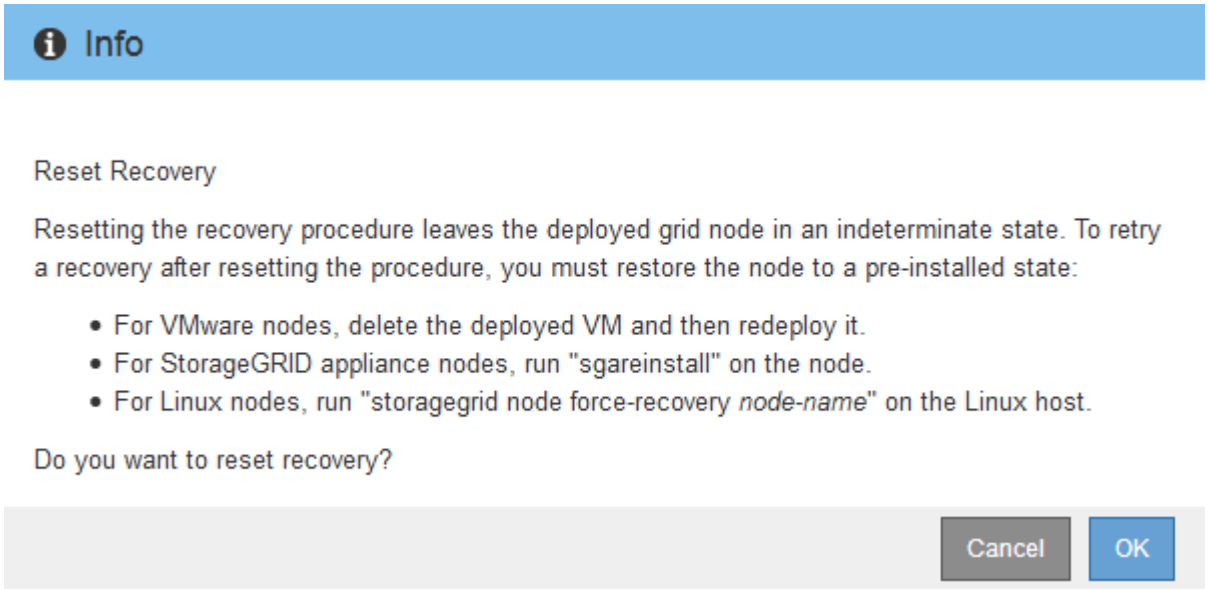
   Nodes appear in the list after they fail, but you cannot select a node until it has been reinstalled and is ready for recovery.

3. Enter the **Provisioning Passphrase**.
4. Click **Start Recovery**.




5. Monitor the progress of the recovery in the Recovering Grid Node table.

While the recovery procedure is running, you can click **Reset** to start a new recovery. An Info dialog box appears, indicating that the node will be left in an indeterminate state if you reset the procedure.




If you want to retry the recovery after resetting the procedure, you must restore the node to a pre-installed state, as follows:

- **VMware**: Delete the deployed virtual grid node. Then, when you are ready to restart the recovery, redeploy the node.
- **Linux**: Restart the node by running this command on the Linux host: `storagegrid node force-recovery node-name`
- **Appliance**: If you want to retry the recovery after resetting the procedure, you must restore the appliance node to a pre-installed state by running `sgareinstall` on the node.

6. If single sign-on (SSO) is enabled for your StorageGRID system and the relying party trust for the Admin Node you recovered was configured to use the default Management Interface Server Certificate, update (or delete and recreate) the node's relying party trust in Active Directory Federation Services (AD FS). Use the new default server certificate that was generated during the Admin Node recovery process.

To configure a relying party trust, see the instructions for administering StorageGRID. To access the default server certificate, log in to the command shell of the Admin Node. Go to the `/var/local/mgmt-api` directory, and select the `server.crt` file.

**Related information**

Administer StorageGRID

Preparing an appliance for reinstallation (platform replacement only)

### Restoring the audit log on the recovered non-primary Admin Node

If you were able to preserve the audit log from the failed non-primary Admin Node, so that historical audit log information is retained, you can copy it to the non-primary Admin Node you are recovering.

- The recovered Admin Node must be installed and running.
- You must have copied the audit logs to another location after the original Admin Node failed.

If an Admin Node fails, audit logs saved to that Admin Node are potentially lost. It might be possible to preserve data from loss by copying audit logs from the failed Admin Node and then restoring these audit logs to the recovered Admin Node. Depending on the failure, it might not be possible to copy audit logs from the failed Admin Node. In that case, if the deployment has more than one Admin Node, you can recover audit logs from another Admin Node as audit logs are replicated to all Admin Nodes.

If there is only one Admin Node and the audit log cannot be copied from the failed node, the recovered Admin Node starts recording events to the audit log as if the installation is new.

You must recover an Admin Node as soon as possible to restore logging functionality.

1. Log in to the recovered Admin Node:

   a. Enter the following command: + `ssh admin@recovery_Admin_Node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   After you are logged in as root, the prompt changes from `$` to `#`.

2. Check which audit files have been preserved:

   ```
   cd /var/local/audit/export
   ```

3. Copy the preserved audit log files to the recovered Admin Node:

   ```
   scp admin@grid_node_IP:/var/local/tmp/saved-audit-logs/YYYY*
   ```

   When prompted, enter the password for admin.

4. For security, delete the audit logs from the failed grid node after verifying that they have been copied successfully to the recovered Admin Node.
5. Update the user and group settings of the audit log files on the recovered Admin Node:

   ```
   chown ams-user:bycast *
   ```

6. Log out as root: `exit`

You must also restore any pre-existing client access to the audit share. For more information, see the instructions for administering StorageGRID.

**Related information**

Administer StorageGRID

**Resetting the preferred sender on the recovered non-primary Admin Node**

If the non-primary Admin Node you are recovering is currently set as the preferred sender of alert notifications, alarm notifications, and AutoSupport messages, you must

reconfigure this setting in the StorageGRID system.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.
- The recovered Admin Node must be installed and running.

**Steps**

1. Select **Configuration** > **System Settings** > **Display Options**.
2. Select the recovered Admin Node from the **Preferred Sender** drop-down list.
3. Click **Apply Changes**.

**Related information**

Administer StorageGRID

### Restoring the Admin Node database when recovering a non-primary Admin Node

If you want to retain the historical information about attributes, alarms, and alerts on a non-primary Admin Node that has failed, you can restore the Admin Node database from the primary Admin Node.

- The recovered Admin Node must be installed and running.
- The StorageGRID system must include at least two Admin Nodes.
- You must have the `Passwords.txt` file.
- You must have the provisioning passphrase.

If an Admin Node fails, the historical information stored in its Admin Node database is lost. This database includes the following information:

- Alert history
- Alarm history
- Historical attribute data, which is used in the charts and text reports available from the **Support** > **Tools** > **Grid Topology** page.

When you recover an Admin Node, the software installation process creates an empty Admin Node database on the recovered node. However, the new database only includes information for servers and services that are currently part of the system or added later.

If you restored a non-primary Admin Node, you can restore the historical information by copying the Admin Node database from the primary Admin Node (the *source Admin Node*) to the recovered node.

> Copying the Admin Node database might take several hours. Some Grid Manager features will be unavailable while services are stopped on the source node.

1. Log in to the source Admin Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.

c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

2. Run the following command from the source Admin Node. Then, enter the provisioning passphrase if prompted. `recover-access-points`

3. From the source Admin Node, stop the MI service: `service mi stop`

4. From the source Admin Node, stop the Management Application Program Interface (mgmt-api) service: `service mgmt-api stop`

5. Complete the following steps on the recovered Admin Node:

   a. Log in to the recovered Admin Node:

      i. Enter the following command: `ssh admin@grid_node_IP`

      ii. Enter the password listed in the `Passwords.txt` file.

      iii. Enter the following command to switch to root: `su -`

      iv. Enter the password listed in the `Passwords.txt` file.

   b. Stop the MI service: `service mi stop`

   c. Stop the mgmt-api service: `service mgmt-api stop`

   d. Add the SSH private key to the SSH agent. Enter:`ssh-add`

   e. Enter the SSH Access Password listed in the `Passwords.txt` file.

   f. Copy the database from the source Admin Node to the recovered Admin Node: `/usr/local/mi/bin/mi-clone-db.sh Source_Admin_Node_IP`

   g. When prompted, confirm that you want to overwrite the MI database on the recovered Admin Node.

      The database and its historical data are copied to the recovered Admin Node. When the copy operation is done, the script starts the recovered Admin Node.

   h. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:`ssh-add -D`

6. Restart the services on the source Admin Node: `service servermanager start`

**Restoring Prometheus metrics when recovering a non-primary Admin Node**

## Optionally, you can retain the historical metrics maintained by Prometheus on a non-primary Admin Node that has failed.

- The recovered Admin Node must be installed and running.
- The StorageGRID system must include at least two Admin Nodes.
- You must have the `Passwords.txt` file.
- You must have the provisioning passphrase.

If an Admin Node fails, the metrics maintained in the Prometheus database on the Admin Node are lost. When you recover the Admin Node, the software installation process creates a new Prometheus database. After the recovered Admin Node is started, it records metrics as if you had performed a new installation of the StorageGRID system.

If you restored a non-primary Admin Node, you can restore the historical metrics by copying the Prometheus database from the primary Admin Node (the *source Admin Node*) to the recovered Admin Node.

> Copying the Prometheus database might take an hour or more. Some Grid Manager features will be unavailable while services are stopped on the source Admin Node.

1. Log in to the source Admin Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.
2. From the source Admin Node, stop the Prometheus service: `service prometheus stop`
3. Complete the following steps on the recovered Admin Node:
   a. Log in to the recovered Admin Node:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.
   b. Stop the Prometheus service: `service prometheus stop`
   c. Add the SSH private key to the SSH agent. Enter:`ssh-add`
   d. Enter the SSH Access Password listed in the `Passwords.txt` file.
   e. Copy the Prometheus database from the source Admin Node to the recovered Admin Node: `/usr/local/prometheus/bin/prometheus-clone-db.sh Source_Admin_Node_IP`
   f. When prompted, press **Enter** to confirm that you want to destroy the new Prometheus database on the recovered Admin Node.

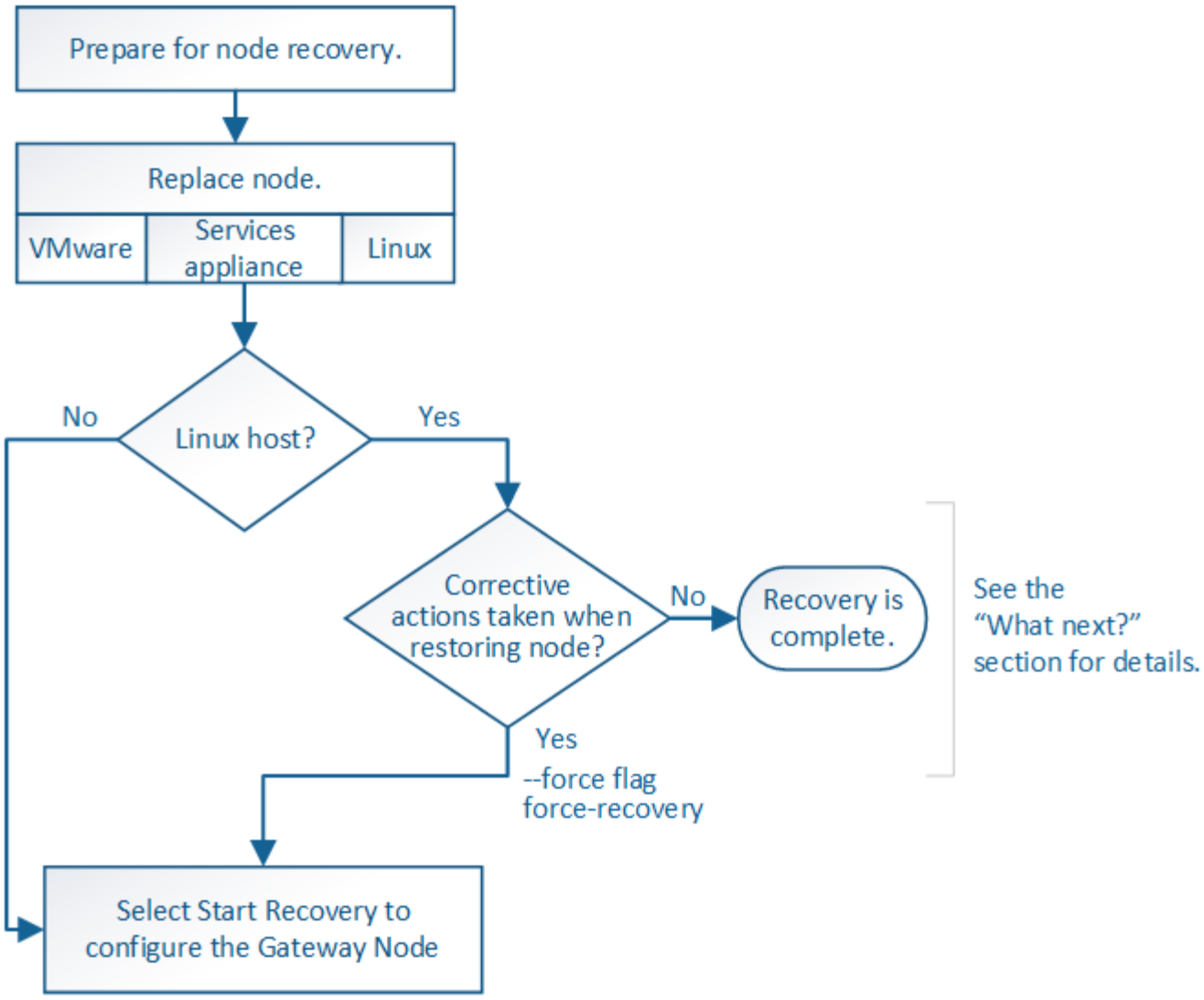
      The original Prometheus database and its historical data are copied to the recovered Admin Node. When the copy operation is done, the script starts the recovered Admin Node. The following status appears:

      Database cloned, starting services

   g. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:`ssh-add -D`
4. Restart the Prometheus service on the source Admin Node.`service prometheus start`

#### Recovering from Gateway Node failures

You must complete a sequence of tasks in exact order to recover from a Gateway Node failure.




**Related information**

SG100 & SG1000 services appliances

**Steps**

- Replacing a Gateway Node
- Selecting Start Recovery to configure a Gateway Node

#### Replacing a Gateway Node

You can replace a failed Gateway Node with a Gateway Node running on the same physical or virtual hardware, or you can replace a Gateway Node running on VMware or a Linux host with a Gateway Node hosted on a services appliance.

The node replacement procedure you must follow depends on which platform will be used by the replacement node. After you complete the node replacement procedure (which is suitable for all node types), that procedure will direct you to the next step for Gateway Node recovery.

| Replacement platform | Procedure |
| --- | --- |
| VMware | Replacing a VMware node |

| Replacement platform | Procedure |
| --- | --- |
| Linux | Replacing a Linux node |
| SG100 and SG1000 services appliances | Replacing a services appliance |
| OpenStack | NetApp-provided virtual machine disk files and scripts for OpenStack are no longer supported for recovery operations. If you need to recover a node running in an OpenStack deployment, download the files for your Linux operating system. Then, follow the procedure for replacing a Linux node. |

#### Selecting Start Recovery to configure a Gateway Node

After replacing a Gateway Node, you must select Start Recovery in the Grid Manager to configure the new node as a replacement for the failed node.

**What you'll need**

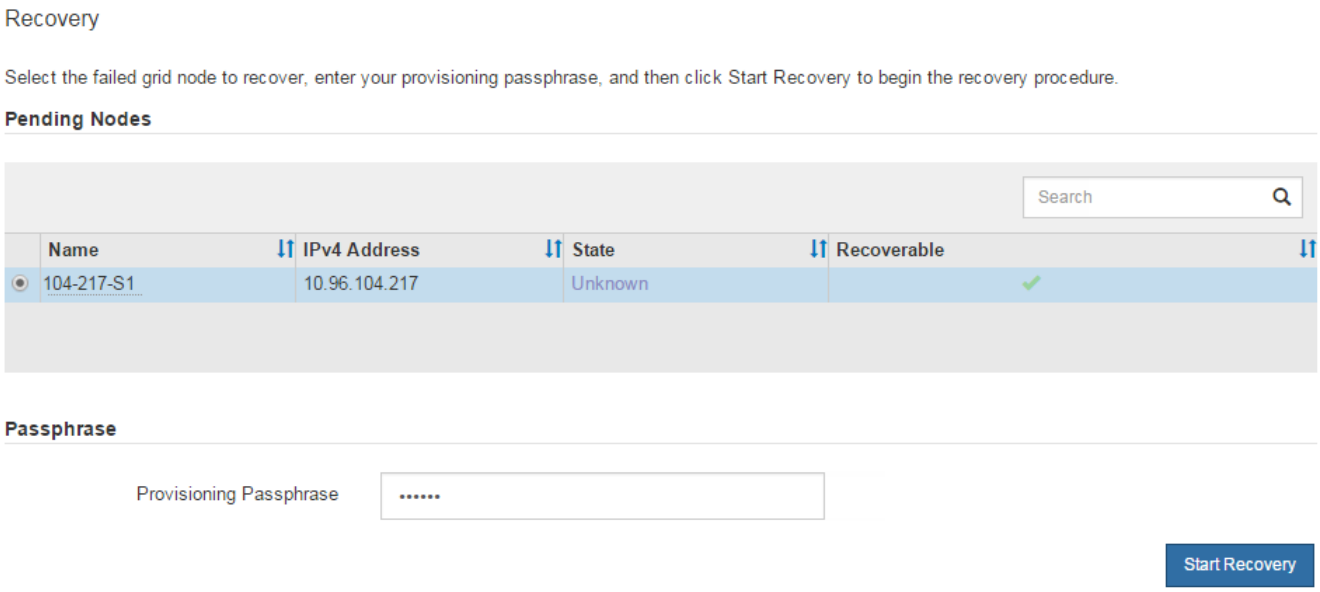
- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have deployed and configured the replacement node.

**Steps**

1. From the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Recovery**.
2. Select the grid node you want to recover in the Pending Nodes list.

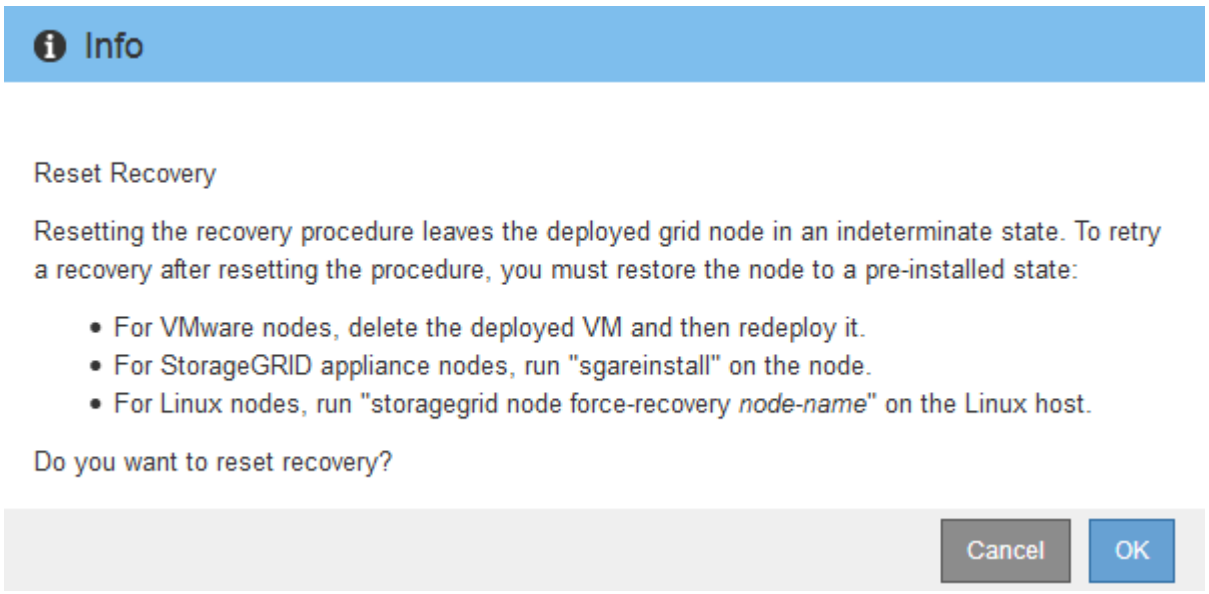
   Nodes appear in the list after they fail, but you cannot select a node until it has been reinstalled and is ready for recovery.

3. Enter the **Provisioning Passphrase**.
4. Click **Start Recovery**.




5. Monitor the progress of the recovery in the Recovering Grid Node table.

> While the recovery procedure is running, you can click **Reset** to start a new recovery. An Info dialog box appears, indicating that the node will be left in an indeterminate state if you reset the procedure.




If you want to retry the recovery after resetting the procedure, you must restore the node to a pre-installed state, as follows:

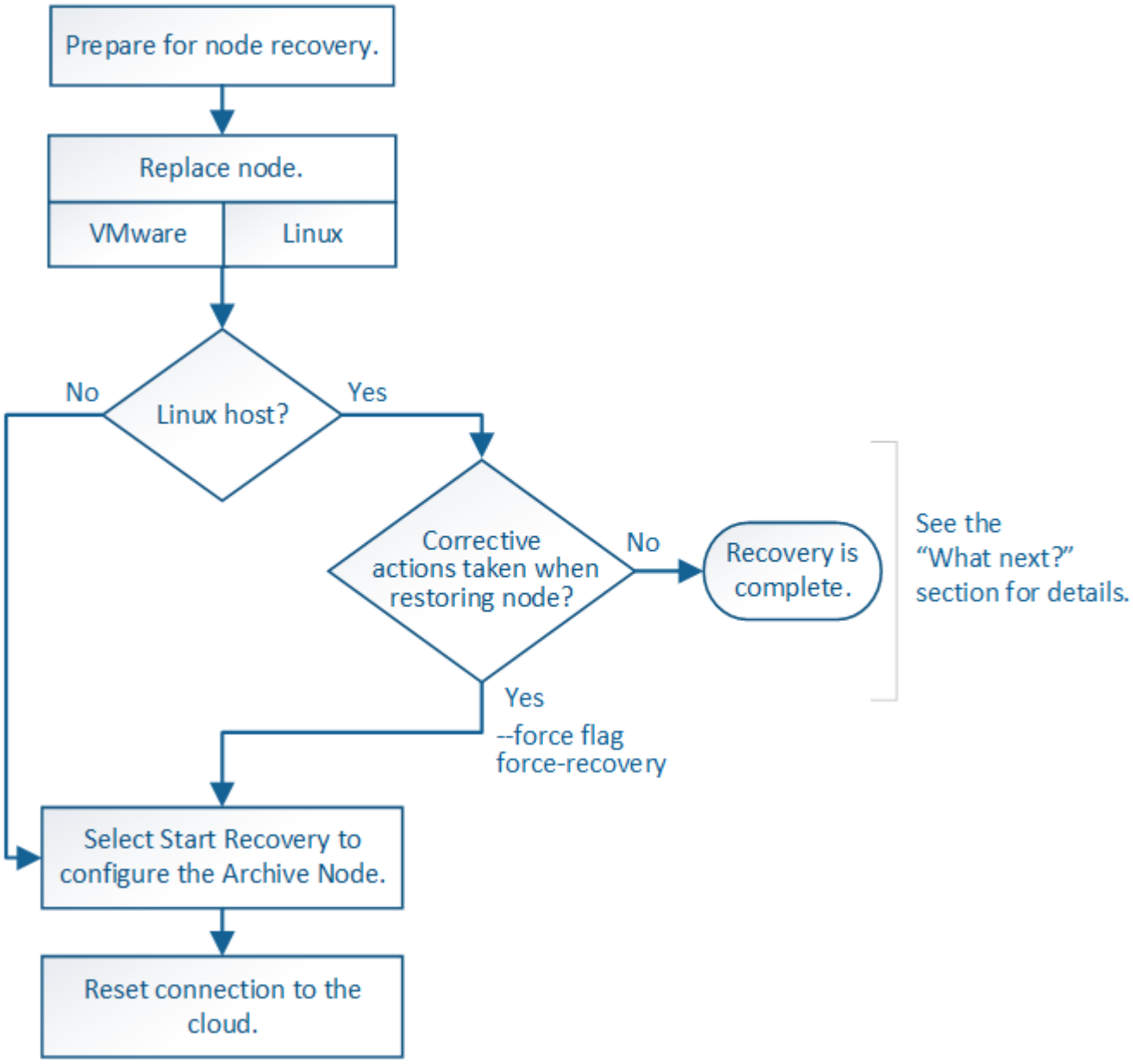
- **VMware**: Delete the deployed virtual grid node. Then, when you are ready to restart the recovery, redeploy the node.
- **Linux**: Restart the node by running this command on the Linux host: `storagegrid node force-recovery node-name`
- **Appliance**: If you want to retry the recovery after resetting the procedure, you must restore the appliance node to a pre-installed state by running `sgareinstall` on the node.

**Related information**

Preparing an appliance for reinstallation (platform replacement only)

### Recovering from Archive Node failures

You must complete a sequence of tasks in exact order to recover from an Archive Node failure.




**About this task**

Archive Node recovery is affected by the following issues:

- If the ILM policy is configured to replicate a single copy.

  In a StorageGRID system that is configured to make a single copy of objects, an Archive Node failure might result in an unrecoverable loss of data. If there is a failure, all such objects are lost; however, you must still perform recovery procedures to "clean up" your StorageGRID system and purge lost object information from the database.

- If an Archive Node failure occurs during Storage Node recovery.

  If the Archive Node fails while processing bulk retrievals as part of a Storage Node recovery, you must repeat the procedure to recover copies of object data to the Storage Node from the beginning to ensure that all object data retrieved from the Archive Node is restored to the Storage Node.

**Steps**

- Replacing an Archive Node
- Selecting Start Recovery to configure an Archive Node
- Resetting Archive Node connection to the cloud

#### Replacing an Archive Node

## To recover an Archive Node, you must first replace the node.

You must select the node replacement procedure for your platform. The steps to replace a node are the same for all types of grid nodes.

| Platform | Procedure |
|---|---|
| VMware | Replacing a VMware node |
| Linux | Replacing a Linux node |
| OpenStack | NetApp-provided virtual machine disk files and scripts for OpenStack are no longer supported for recovery operations. If you need to recover a node running in an OpenStack deployment, download the files for your Linux operating system. Then, follow the procedure for replacing a Linux node. |

#### Selecting Start Recovery to configure an Archive Node

## After replacing an Archive Node, you must select Start Recovery in the Grid Manager to configure the new node as a replacement for the failed node.

**What you'll need**

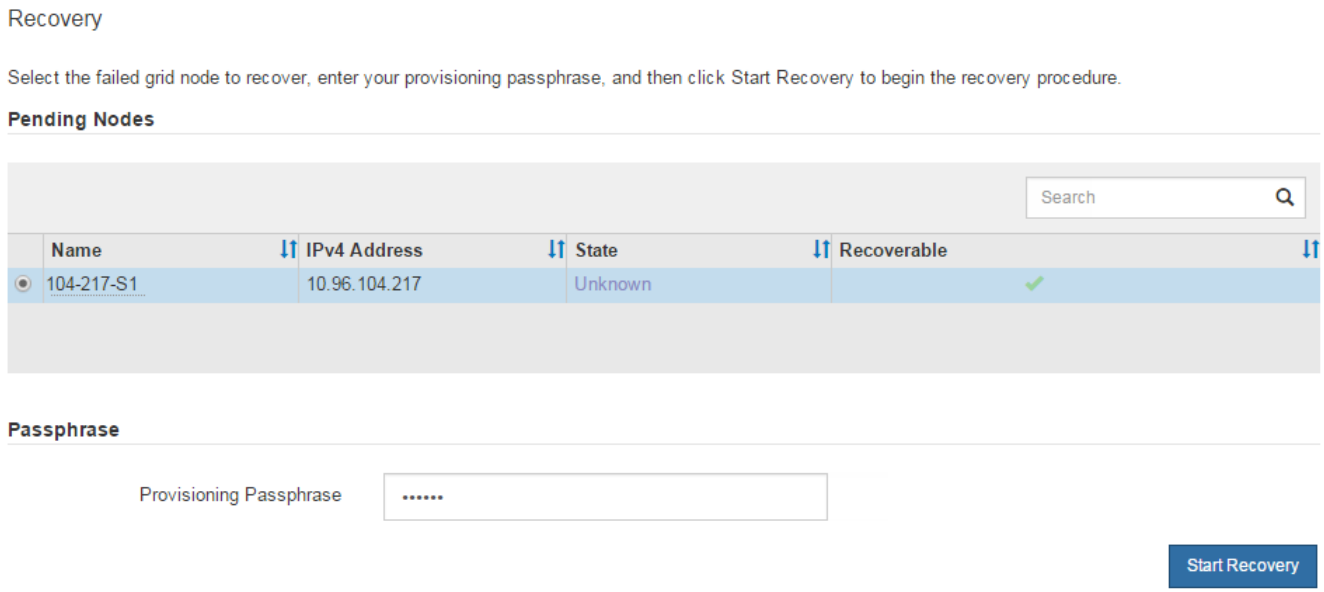
- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have deployed and configured the replacement node.

**Steps**

1. From the Grid Manager, select **Maintenance** > **Maintenance Tasks** > **Recovery**.
2. Select the grid node you want to recover in the Pending Nodes list.

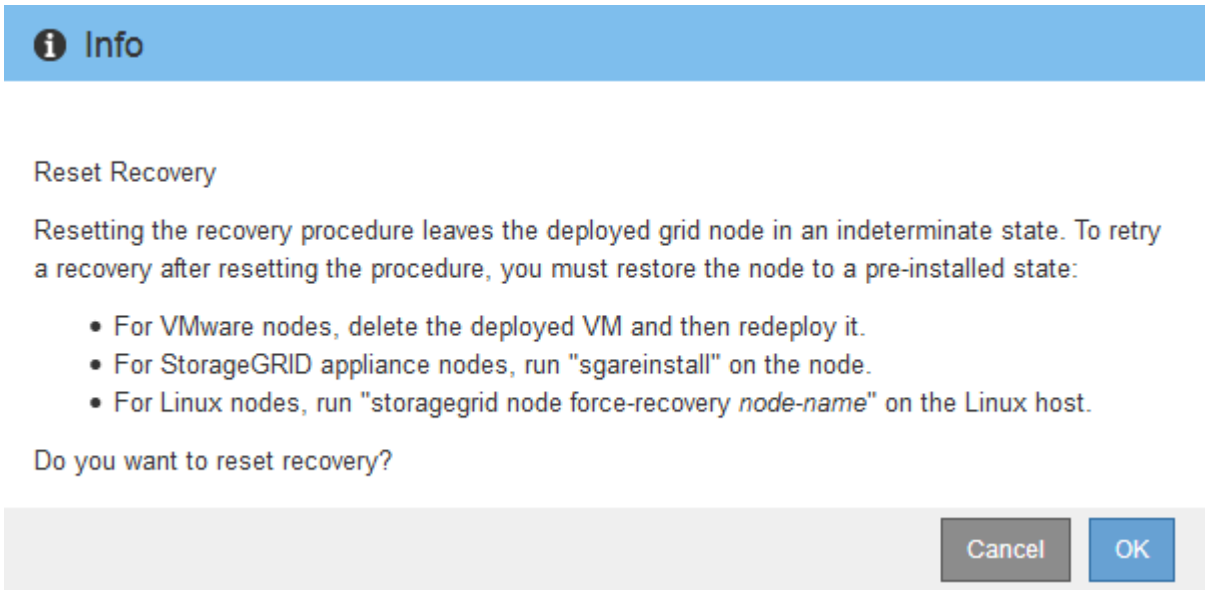
   Nodes appear in the list after they fail, but you cannot select a node until it has been reinstalled and is ready for recovery.

3. Enter the **Provisioning Passphrase**.
4. Click **Start Recovery**.




5. Monitor the progress of the recovery in the Recovering Grid Node table.

> While the recovery procedure is running, you can click **Reset** to start a new recovery. An Info dialog box appears, indicating that the node will be left in an indeterminate state if you reset the procedure.




If you want to retry the recovery after resetting the procedure, you must restore the node to a pre-installed state, as follows:

- **VMware**: Delete the deployed virtual grid node. Then, when you are ready to restart the recovery, redeploy the node.
- **Linux**: Restart the node by running this command on the Linux host: `storagegrid node force-recovery node-name`

#### Resetting Archive Node connection to the cloud

After you recover an Archive Node that targets the cloud through the S3 API, you need to modify configuration settings to reset connections. An Outbound Replication Status (ORSU) alarm is triggered if the Archive Node is unable to retrieve object data.

> If your Archive Node connects to external storage through TSM middleware, then the node resets itself automatically and you do not need to reconfigure.

**What you'll need**

You must be signed in to the Grid Manager using a supported browser.

**Steps**

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select **Archive Node** > **ARC** > **Target**.
3. Edit the **Access Key** field by entering an incorrect value and click **Apply Changes**.
4. Edit the **Access Key** field by entering the correct value and click **Apply Changes**.

### All grid node types: Replacing a VMware node

When you recover a failed StorageGRID node that was hosted on VMware, you must remove the failed node and deploy a recovery node.

**What you'll need**

You must have determined that the virtual machine cannot be restored, and must be replaced.

**About this task**

You use the VMware vSphere Web Client to first remove the virtual machine associated with the failed grid node. Then, you can deploy a new virtual machine.

This procedure is only one step in the grid node recovery process. The node removal and deployment procedure is the same for all VMware nodes, including Admin Nodes, Storage Nodes, Gateway Nodes, and Archive Nodes.

**Steps**

1. Log in to VMware vSphere Web Client.
2. Navigate to the failed grid node virtual machine.
3. Make a note of all of the information required to deploy the recovery node.
   a. Right-click the virtual machine, select the **Edit Settings** tab, and note the settings in use.
   b. Select the **vApp Options** tab to view and record the grid node network settings.
4. If the failed grid node is a Storage Node, determine if any of the virtual hard disks used for data storage are undamaged and preserve them for reattachment to the recovered grid node.
5. Power off the virtual machine.
6. Select **Actions** > **All vCenter Actions** > **Delete from Disk** to delete the virtual machine.
7. Deploy a new virtual machine to be the replacement node, and connect it to one or more StorageGRID networks.

When you deploy the node, you can optionally remap node ports or increase CPU or memory settings.

> After deploying the new node, you can add new virtual disks according to your storage requirements, reattach any virtual hard disks preserved from the previously removed failed grid node, or both.

For instructions:

Install VMware > Deploying a StorageGRID node as a virtual machine

8. Complete the node recovery procedure, based on the type of node you are recovering.

| Type of node | Go to |
|---|---|
| Primary Admin Node | Configuring the replacement primary Admin Node |
| Non-primary Admin Node | Selecting Start Recovery to configure a non-primary Admin Node |
| Gateway Node | Selecting Start Recovery to configure a Gateway Node |
| Storage Node | Selecting Start Recovery to configure a Storage Node |
| Archive Node | Selecting Start Recovery to configure an Archive Node |

**All grid node types: Replacing a Linux node**

If a failure requires that you deploy one or more new physical or virtual hosts or reinstall Linux on an existing host, you must deploy and configure the replacement host before you can recover the grid node. This procedure is one step of the grid node recovery process for all types of grid nodes.

“Linux” refers to a Red Hat® Enterprise Linux®, Ubuntu®, CentOS, or Debian® deployment. Use the NetApp Interoperability Matrix Tool to get a list of supported versions.

This procedure is only performed as one step in the process of recovering software-based Storage Nodes, primary or non-primary Admin Nodes, Gateway Nodes, or Archive Nodes. The steps are identical regardless of the type of grid node you are recovering.

If more than one grid node is hosted on a physical or virtual Linux host, you can recover the grid nodes in any order. However, recovering a primary Admin Node first, if present, prevents the recovery of other grid nodes from stalling as they try to contact the primary Admin Node to register for recovery.

1. Deploying new Linux hosts
2. Restoring grid nodes to the host
3. What’s next: Performing additional recovery steps, if required

**Related information**

NetApp Interoperability Matrix Tool

#### Deploying new Linux hosts

## With a few exceptions, you prepare the new hosts as you did during the initial installation process.

To deploy new or reinstalled physical or virtual Linux hosts, follow the procedure for preparing the hosts in the StorageGRID installation instructions for your Linux operating system.

This procedure includes steps to accomplish the following tasks:

1. Install Linux.
2. Configure the host network.
3. Configure host storage.
4. Install Docker.
5. Install the StorageGRID host service.

> Stop after you complete the “Install StorageGRID host service” task in the installation instructions. Do not start the “Deploying grid nodes” task.

As you perform these steps, note the following important guidelines:

- Be sure to use the same host interface names you used on the original host.
- If you use shared storage to support your StorageGRID nodes, or you have moved some or all of the disk drives or SSDs from the failed to the replacement nodes, you must reestablish the same storage mappings that were present on the original host. For example, if you used WWIDs and aliases in `/etc/multipath.conf` as recommended in the installation instructions, be sure to use the same alias/WWID pairs in `/etc/multipath.conf` on the replacement host.
- If the StorageGRID node uses storage assigned from a NetApp AFF system, confirm that the volume does not have a FabricPool tiering policy enabled. Disabling FabricPool tiering for volumes used with StorageGRID nodes simplifies troubleshooting and storage operations.

> Never use FabricPool to tier any data related to StorageGRID back to StorageGRID itself. Tiering StorageGRID data back to StorageGRID increases troubleshooting and operational complexity.

**Related information**

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

#### Restoring grid nodes to the host

## To restore a failed grid node to a new Linux host, you restore the node configuration file using the appropriate commands.

When doing a fresh install, you create a node configuration file for each grid node to be installed on a host. When restoring a grid node to a replacement host, you restore or replace the node configuration file for any failed grid nodes.

If any block storage volumes were preserved from the previous host, you might have to perform additional recovery procedures. The commands in this section help you determine which additional procedures are required.

**Steps**

- Restoring and validating grid nodes
- Starting the StorageGRID host service
- Recovering nodes that fail to start normally

### Restoring and validating grid nodes

You must restore the grid configuration files for any failed grid nodes, and then validate the grid configuration files and resolve any errors.

**About this task**

You can import any grid node that should be present on the host, as long as its `/var/local` volume was not lost as a result of the failure of the previous host. For example, the `/var/local` volume might still exist if you used shared storage for StorageGRID system data volumes, as described in the StorageGRID installation instructions for your Linux operating system. Importing the node restores its node configuration file to the host.

If it is not possible to import missing nodes, you must recreate their grid configuration files.

You must then validate the grid configuration file, and resolve any networking or storage issues that might occur before going on to restart StorageGRID. When you re-create the configuration file for a node, you must use the same name for the replacement node that was used for the node you are recovering.

See the installation instructions for more information on the location of the `/var/local` volume for a node.

**Steps**

1. At the command line of the recovered host, list all currently configured StorageGRID grid nodes:`sudo storagegrid node list`

   If no grid nodes are configured, there will be no output. If some grid nodes are configured, expect output in the following format:

   ```
   Name               Metadata-Volume
   ============================================================
   dc1-adm1           /dev/mapper/sgws-adm1-var-local
   dc1-gw1            /dev/mapper/sgws-gw1-var-local
   dc1-sn1            /dev/mapper/sgws-sn1-var-local
   dc1-arc1           /dev/mapper/sgws-arc1-var-local
   ```

   If some or all of the grid nodes that should be configured on the host are not listed, you need to restore the missing grid nodes.

2. To import grid nodes that have a `/var/local` volume:

   a. Run the following command for each node you want to import:`sudo storagegrid node import node-var-local-volume-path`

The `storagegrid node import` command succeeds only if the target node was shut down cleanly on the host on which it last ran. If that is not the case, you will observe an error similar to the following:

```
This node (node-name) appears to be owned by another host (UUID host-uuid).

Use the --force flag if you are sure import is safe.
```

b. If you see the error about the node being owned by another host, run the command again with the `--force` flag to complete the import:`sudo storagegrid --force node import node-var-local-volume-path`

> Any nodes imported with the `--force` flag will require additional recovery steps before they can rejoin the grid, as described in “Performing additional recovery steps, if required.”

3. For grid nodes that do not have a `/var/local` volume, recreate the node’s configuration file to restore it to the host.

   Follow the guidelines in “Creating node configuration files” in the installation instructions.

> When you re-create the configuration file for a node, you must use the same name for the replacement node that was used for the node you are recovering. For Linux deployments, ensure that the configuration file name contains the node name. You should use the same network interfaces, block device mappings, and IP addresses when possible. This practice minimizes the amount of data that needs to be copied to the node during recovery, which could make the recovery significantly faster (in some cases, minutes rather than weeks).

> If you use any new block devices (devices that the StorageGRID node did not use previously) as values for any of the configuration variables that start with `BLOCK_DEVICE_` when you are recreating the configuration file for a node, be sure to follow all of the guidelines in “Fixing missing block device errors.”

4. Run the following command on the recovered host to list all StorageGRID nodes.

   ```
   sudo storagegrid node list
   ```

5. Validate the node configuration file for each grid node whose name was shown in the storagegrid node list output:

   ```
   sudo storagegrid node validate node-name
   ```

   You must address any errors or warnings before starting the StorageGRID host service. The following sections give more detail on errors that might have special significance during recovery.

**Related information**

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

Fixing missing network interface errors

Fixing missing block device errors

What's next: Performing additional recovery steps, if required

### Fixing missing network interface errors

If the host network is not configured correctly or a name is misspelled, an error occurs when StorageGRID checks the mapping specified in the `/etc/storagegrid/nodes/node-name.conf` file.

You might see an error or warning matching this pattern:

`Checking configuration file `/etc/storagegrid/nodes/node-name.conf` for node node-name...`ERROR: node-name: GRID_NETWORK_TARGET = host-interface-name`node-name: Interface 'host-interface-name' does not exist`

The error could be reported for the Grid Network, the Admin Network, or the Client Network. This error means that the `/etc/storagegrid/nodes/node-name.conf` file maps the indicated StorageGRID network to the host interface named `host-interface-name`, but there is no interface with that name on the current host.

If you receive this error, verify that you completed the steps in "Deploying new Linux hosts." Use the same names for all host interfaces as were used on the original host.

If you are unable to name the host interfaces to match the node configuration file, you can edit the node configuration file and change the value of the GRID_NETWORK_TARGET, the ADMIN_NETWORK_TARGET, or the CLIENT_NETWORK_TARGET to match an existing host interface.

Make sure the host interface provides access to the appropriate physical network port or VLAN, and that the interface does not directly reference a bond or bridge device. You must either configure a VLAN (or other virtual interface) on top of the bond device on the host, or use a bridge and virtual Ethernet (veth) pair.

**Related information**

Deploying new Linux hosts

### Fixing missing block device errors

The system checks that each recovered node maps to a valid block device special file or a valid softlink to a block device special file. If StorageGRID finds invalid mapping in the `/etc/storagegrid/nodes/node-name.conf` file, a missing block device error displays.

If you observe an error matching this pattern:

`Checking configuration file /etc/storagegrid/nodes/node-name.conf for node node-name... ERROR: node-name: BLOCK_DEVICE_PURPOSE = path-name`node-name: path-name does not exist`

It means that `/etc/storagegrid/nodes/node-name.conf` maps the block device used by *node-name* for PURPOSE to the given path-name in the Linux file system, but there is not a valid block device special file, or softlink to a block device special file, at that location.

Verify that you completed the steps in "Deploying new Linux hosts." Use the same persistent device names for all block devices as were used on the original host.

If you are unable to restore or recreate the missing block device special file, you can allocate a new block

device of the appropriate size and storage category and edit the node configuration file to change the value of BLOCK_DEVICE_PURPOSE to point to the new block device special file.

Determine the appropriate size and storage category from the tables in the “Storage requirements” section of the installation instructions for your Linux operating system. Review the recommendations in “Configuring host storage” before proceeding with the block device replacement.

If you must provide a new block storage device for any of the configuration file variables starting with `BLOCK_DEVICE_` because the original block device was lost with the failed host, ensure the new block device is unformatted before attempting further recovery procedures. The new block device will be unformatted if you are using shared storage and have created a new volume. If you are unsure, run the following command against any new block storage device special files.

Run the following command only for new block storage devices. Do not run this command if you believe the block storage still contains valid data for the node being recovered, as any data on the device will be lost.

```
sudo dd if=/dev/zero of=/dev/mapper/my-block-device-name bs=1G count=1
```

**Related information**

Deploying new Linux hosts

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

### Starting the StorageGRID host service

To start your StorageGRID nodes, and ensure they restart after a host reboot, you must enable and start the StorageGRID host service.

1. Run the following commands on each host:

   ```
   sudo systemctl enable storagegrid
   sudo systemctl start storagegrid
   ```

2. Run the following command to ensure the deployment is proceeding:

   ```
   sudo storagegrid node status node-name
   ```

   For any node that returns a status of Not-Running or Stopped, run the following command:

   ```
   sudo storagegrid node start node-name
   ```

3. If you have previously enabled and started the StorageGRID host service (or if you are unsure if the service has been enabled and started), also run the following command:

```
sudo systemctl reload-or-restart storagegrid
```

### Recovering nodes that fail to start normally

If a StorageGRID node does not rejoin the grid normally and does not show up as recoverable, it may be corrupted. You can force the node into recovery mode.

To force the node into recovery mode:

```
sudo storagegrid node force-recovery node-name
```

Before issuing this command, confirm that the node's network configuration is correct; it may have failed to rejoin the grid due to incorrect network interface mappings or an incorrect Grid Network IP address or gateway.

After issuing the `storagegrid node force-recovery node-name` command, you must perform additional recovery steps for *node-name*.

### Related information

What's next: Performing additional recovery steps, if required

#### What's next: Performing additional recovery steps, if required

Depending on the specific actions you took to get the StorageGRID nodes running on the replacement host, you might need to perform additional recovery steps for each node.

Node recovery is complete if you did not need to take any corrective actions while you replaced the Linux host or restored the failed grid node to the new host.

### Corrective actions and next steps

During node replacement, you may have needed to take one of these corrective actions:

- You had to use the `--force` flag to import the node.
- For any `<PURPOSE>`, the value of the `BLOCK_DEVICE_<PURPOSE>` configuration file variable refers to a block device that does not contain the same data it did before the host failure.
- You issued `storagegrid node force-recovery node-name` for the node.
- You added a new block device.

If you took **any** of these corrective actions, you must perform additional recovery steps.

| Type of recovery | Next step |
|---|---|
| Primary Admin Node | Configuring the replacement primary Admin Node |
| Non-primary Admin Node | Selecting Start Recovery to configure a non-primary Admin Node |

| Type of recovery | Next step |
| --- | --- |
| Gateway Node | Selecting Start Recovery to configure a Gateway Node |
| Archive Node | Selecting Start Recovery to configure an Archive Node |
| Storage Node (software-based):<br>• If you had to use the `--force` flag to import the node, or you issued `storagegrid node force-recovery node-name`<br>• If you had to do a full node reinstall, or you needed to restore /var/local | Selecting Start Recovery to configure a Storage Node |
| Storage Node (software-based):<br>• If you added a new block device.<br>• If, for any `<PURPOSE>`, the value of the `BLOCK_DEVICE_<PURPOSE>` configuration file variable refers to a block device that does not contain the same data it did before the host failure. | Recovering from storage volume failure where the system drive is intact |

#### Replacing a failed node with a services appliance

You can use an SG100 or SG1000 services appliance to recover a failed Gateway Node, a failed non-primary Admin Node, or a failed primary Admin Node that was hosted on VMware, a Linux host, or a services appliance. This procedure is one step of the grid node recovery procedure.

#### What you'll need

- You must have determined that one of the following situations is true:
  - The virtual machine hosting the node cannot be restored.
  - The physical or virtual Linux host for the grid node has failed, and must be replaced.
  - The services appliance hosting the grid node must be replaced.
- You must make sure that the StorageGRID Appliance Installer version on the services appliance matches the software version of your StorageGRID system, as described in hardware installation and maintenance for verifying and upgrading the StorageGRID Appliance Installer version.

  SG100 & SG1000 services appliances

> Do not deploy both an SG100 and an SG1000 service appliance in the same site. Unpredictable performance might result.

#### About this task

You can use an SG100 or SG1000 services appliance to recover a failed grid node in the following cases:

- The failed node was hosted on VMware or Linux (platform change)
- The failed node was hosted on a services appliance (platform replacement)

**Steps**

- Installing a services appliance (platform change only)
- Preparing an appliance for reinstallation (platform replacement only)
- Starting software installation on a services appliance
- Monitoring services appliance installation

#### Installing a services appliance (platform change only)

## When you are recovering a failed grid node that was hosted on VMware or a Linux host and you are using an SG100 or SG1000 services appliance for the replacement node, you must first install the new appliance hardware using the same node name as the failed node.

You must have the following information about the failed node:

- **Node name**: You must install the services appliance using the same node name as the failed node.
- **IP addresses**: You can assign the services appliance the same IP addresses as the failed node, which is the preferred option, or you can select a new unused IP address on each network.

Perform this procedure only if you are recovering a failed node that was hosted on VMware or Linux and are replacing it with a node hosted on a services appliance.

1. Follow the instructions for installing a new SG100 or SG1000 services appliance.
2. When prompted for a node name, use the node name of the failed node.

**Related information**

SG100 & SG1000 services appliances

#### Preparing an appliance for reinstallation (platform replacement only)

## When recovering a grid node that was hosted on a services appliance, you must first prepare the appliance for reinstallation of StorageGRID software.

Perform this procedure only if you are replacing a failed node that was hosted on a services appliance. Do not follow these steps if the failed node was originally hosted on VMware or a Linux host.

1. Log in to the failed grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Prepare the appliance for the installation of StorageGRID software. Enter: `sgareinstall`
3. When prompted to continue, enter: `y`

   The appliance reboots, and your SSH session ends. It usually takes about 5 minutes for the StorageGRID Appliance Installer to become available, although in some cases you might need to wait up to 30 minutes.

   The services appliance is reset, and data on the grid node is no longer accessible. IP addresses configured during the original installation process should remain intact; however, it is recommended that you confirm this when the procedure completes.

   After executing the `sgareinstall` command, all StorageGRID-provisioned accounts, passwords, and SSH keys are removed, and new host keys are generated.

#### Starting software installation on a services appliance

## To install a Gateway Node or Admin Node on an SG100 or SG1000 services appliance, you use the StorageGRID Appliance Installer, which is included on the appliance.

**What you'll need**

- The appliance must be installed in a rack, connected to your networks, and powered on.
- Network links and IP addresses must be configured for the appliance using the StorageGRID Appliance Installer.
- If you are installing a Gateway Node or non-primary Admin Node, you know the IP address of the primary Admin Node for the StorageGRID grid.
- All Grid Network subnets listed on the IP Configuration page of the StorageGRID Appliance Installer must be defined in the Grid Network Subnet List on the primary Admin Node.

For instructions for completing these prerequisite tasks, see the installation and maintenance instructions for an SG100 or SG1000 services appliance.

- You must be using a supported web browser.
- You must know one of the IP addresses assigned to the appliance. You can use the IP address for the Admin Network, the Grid Network, or the Client Network.
- If you are installing a primary Admin Node, you have the Ubuntu or Debian install files for this version of StorageGRID available.

> ⓘ A recent version of StorageGRID software is preloaded onto the services appliance during manufacturing. If the preloaded version of software matches the version being used in your StorageGRID deployment, you do not need the installation files.

**About this task**

To install StorageGRID software on an SG100 or SG1000 services appliance:

- For a primary Admin Node, you specify the name of the node and then upload the appropriate software packages (if required).
- For a non-primary Admin Node or a Gateway Node, you specify or confirm the IP address of the primary Admin Node and the name of the node.
- You start the installation and wait as volumes are configured and the software is installed.

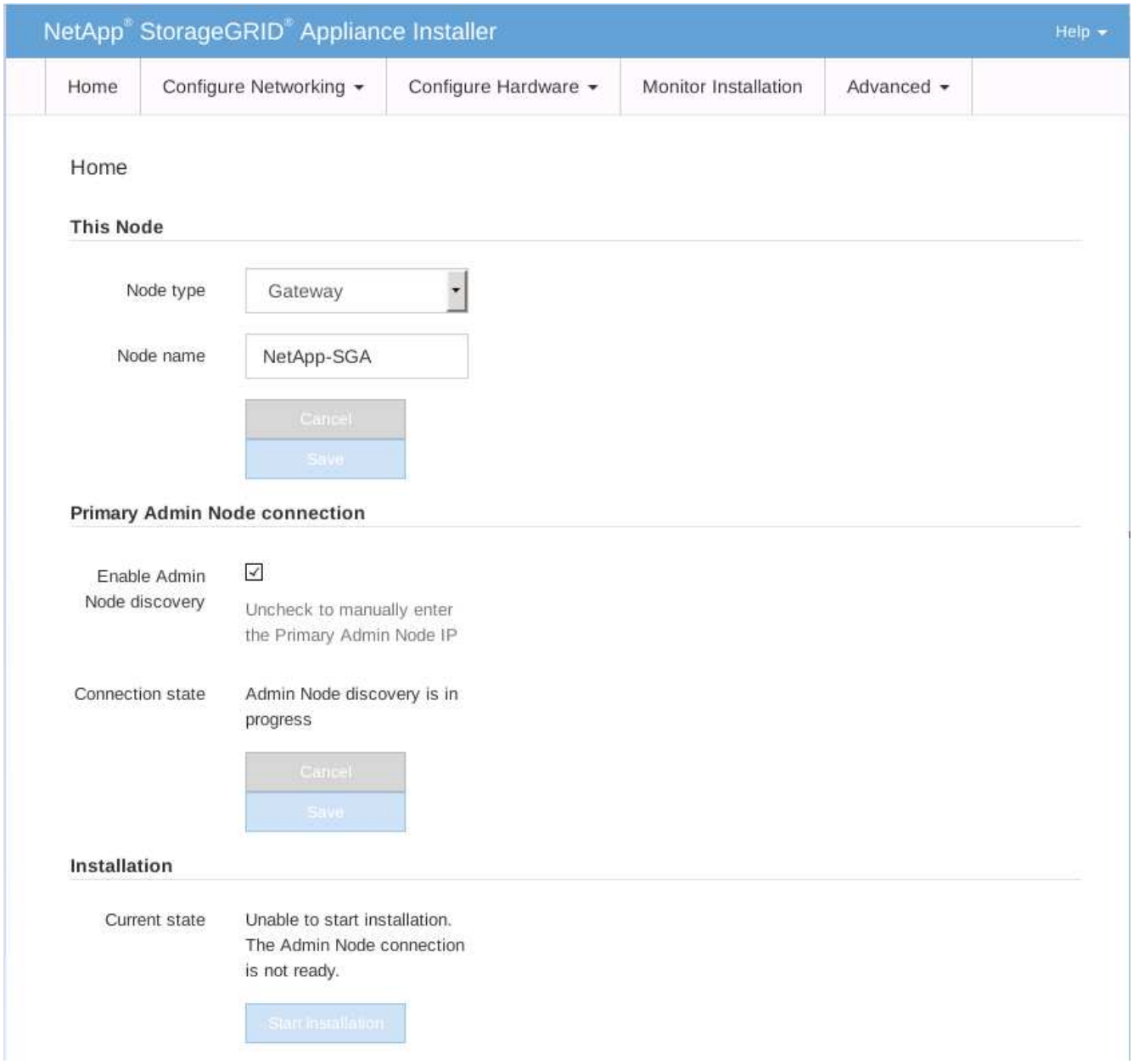
- Partway through the process, the installation pauses. To resume the installation, you must sign into the Grid Manager and configure the pending node as a replacement for the failed node.
- After you have configured the node, the appliance installation process completes, and the appliance is rebooted.

**Steps**

1. Open a browser and enter one of the IP addresses for the SG100 or SG1000 services appliance.

   `https://Controller_IP:8443`

   The StorageGRID Appliance Installer Home page appears.



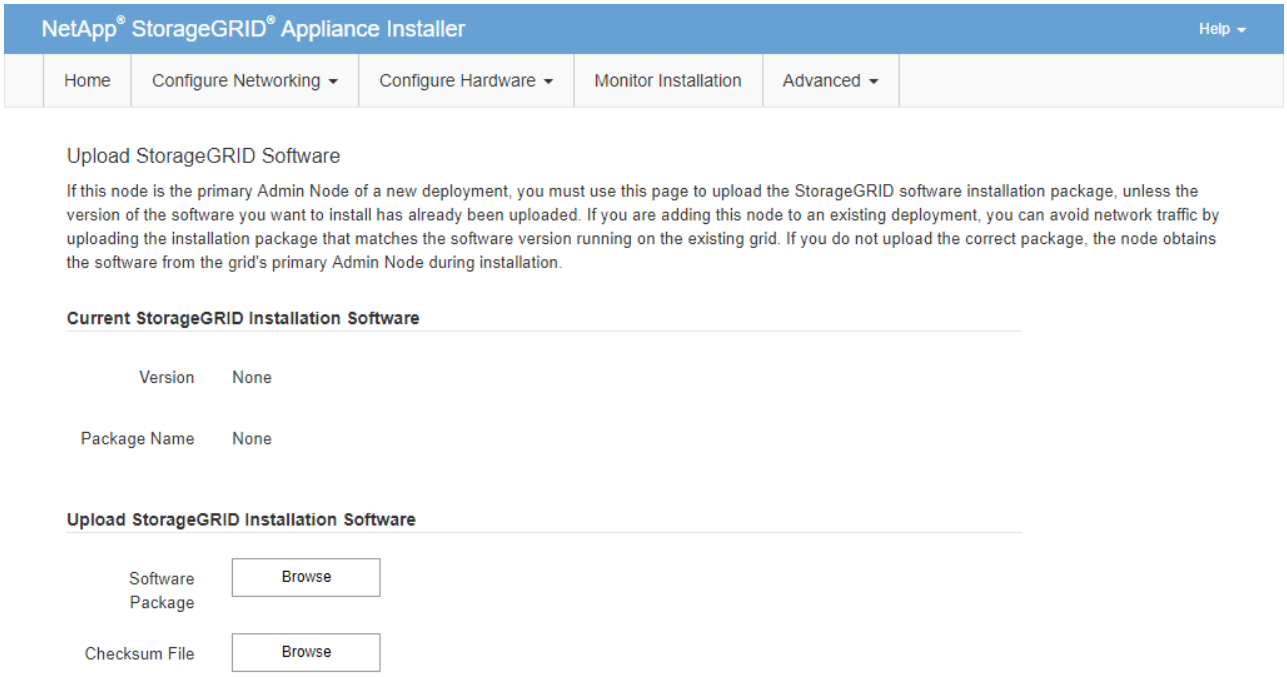

2. To install a Primary Admin Node:
   a. In the This Node section, for **Node Type**, select **Primary Admin**.
   b. In the **Node Name** field, enter the same name that was used for the node you are recovering, and click **Save**.

c. In the Installation section, check the software version listed under Current state

   If the version of software that is ready to install is correct, skip ahead to the Installation step.

d. If you need to upload a different version of software, under the **Advanced** menu, select **Upload StorageGRID Software**.

   TheUpload StorageGRID Software page appears.




e. Click **Browse** to upload the **Software Package** and **Checksum File** for StorageGRID software.

   The files are automatically uploaded after you select them.

f. Click **Home** to return to the StorageGRID Appliance Installer Home page.

3. To install a Gateway Node or non-Primary Admin Node:

   a. In the This Node section, for **Node Type**, select **Gateway** or **Non-Primary Admin**, depending on the type of node you are restoring.

   b. In the **Node Name** field, enter the same name that was used for the node you are recovering, and click **Save**.

   c. In the Primary Admin Node connection section, determine whether you need to specify the IP address for the primary Admin Node.

      The StorageGRID Appliance Installer can discover this IP address automatically, assuming the primary Admin Node, or at least one other grid node with ADMIN_IP configured, is present on the same subnet.

   d. If this IP address is not shown or you need to change it, specify the address:

| Option | Description |
| --- | --- |
| Manual IP entry | a. Unselect the **Enable Admin Node discovery** check box.<br>b. Enter the IP address manually.<br>c. Click **Save**.<br>d. Wait while the connection state for the new IP address becomes “ready.” |
| Automatic discovery of all connected primary Admin Nodes | a. Select the **Enable Admin Node discovery** check box.<br>b. From the list of discovered IP addresses, select the primary Admin Node for the grid where this services appliance will be deployed.<br>c. Click **Save**.<br>d. Wait while the connection state for the new IP address becomes “ready.” |

4. In the Installation section, confirm that the current state is Ready to start installation of node name and that the **Start Installation** button is enabled.

   If the **Start Installation** button is not enabled, you might need to change the network configuration or port settings. For instructions, see the installation and maintenance instructions for your appliance.

5. From the StorageGRID Appliance Installer home page, click **Start Installation**.

   The Current state changes to “Installation is in progress,” and the Monitor Installation page is displayed.

   If you need to access the Monitor Installation page manually, click **Monitor Installation** from the menu bar.

**Related information**

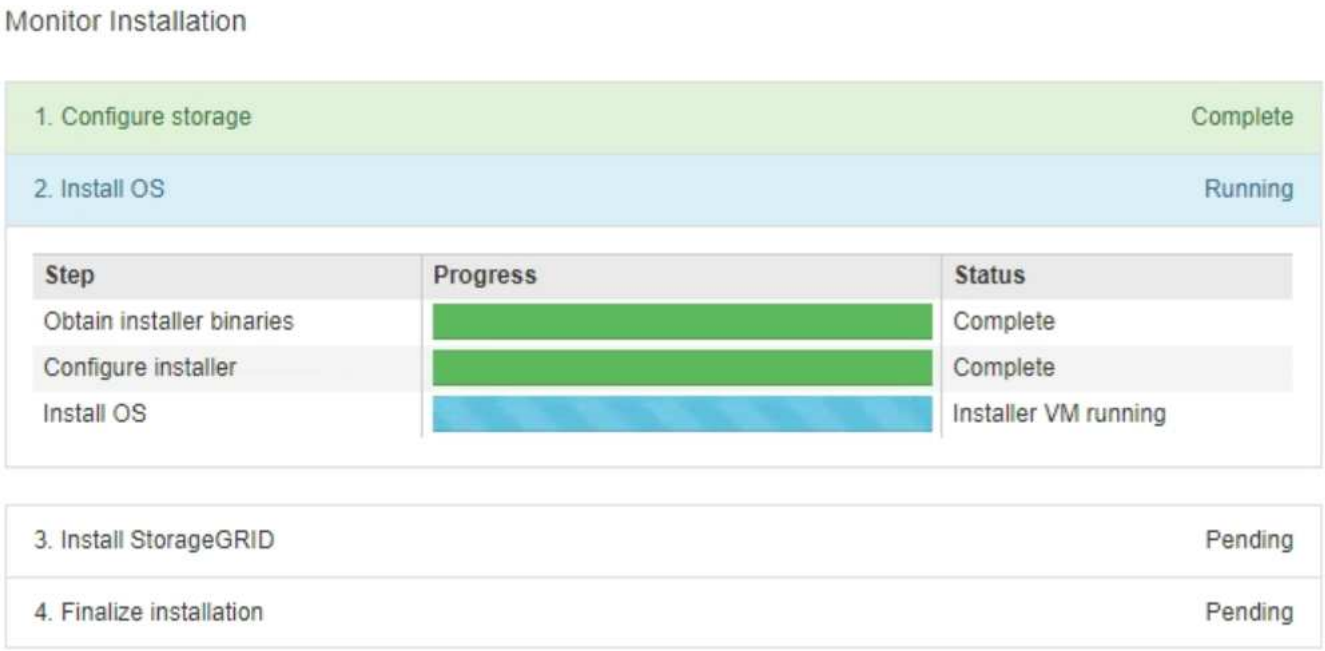
SG100 & SG1000 services appliances

#### Monitoring services appliance installation

The StorageGRID Appliance Installer provides status until installation is complete. When the software installation is complete, the appliance is rebooted.

1. To monitor the installation progress, click **Monitor Installation** from the menu bar.

   The Monitor Installation page shows the installation progress.




The blue status bar indicates which task is currently in progress. Green status bars indicate tasks that have completed successfully.

> ⓘ The installer ensures that tasks completed in a previous install are not re-run. If you are re-running an installation, any tasks that do not need to be re-run are shown with a green status bar and a status of “Skipped.”

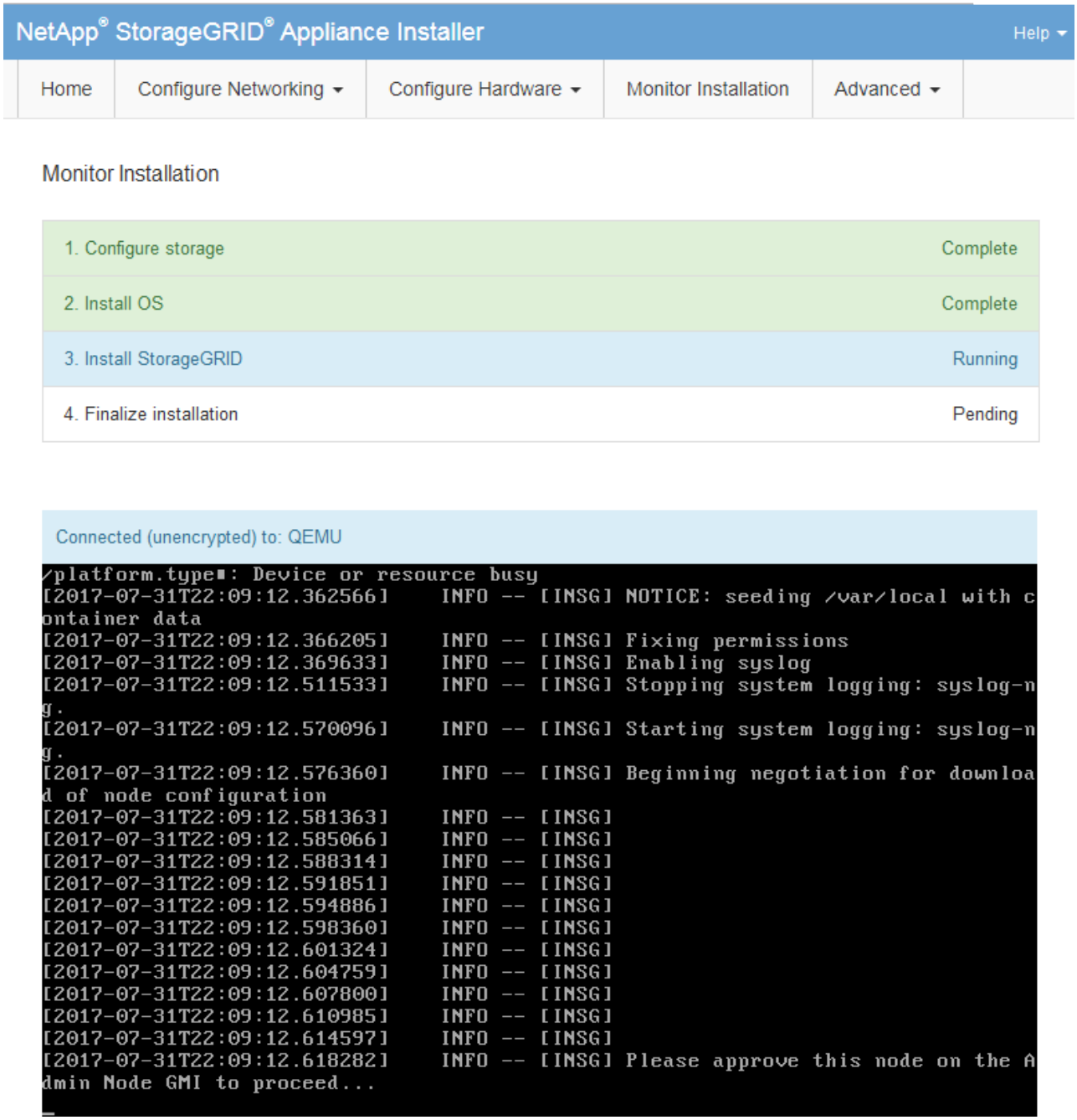
2. Review the progress of first two installation stages.
   - **1. Configure storage**

     During this stage, the installer clears any existing configuration from the drives, and configures host settings.

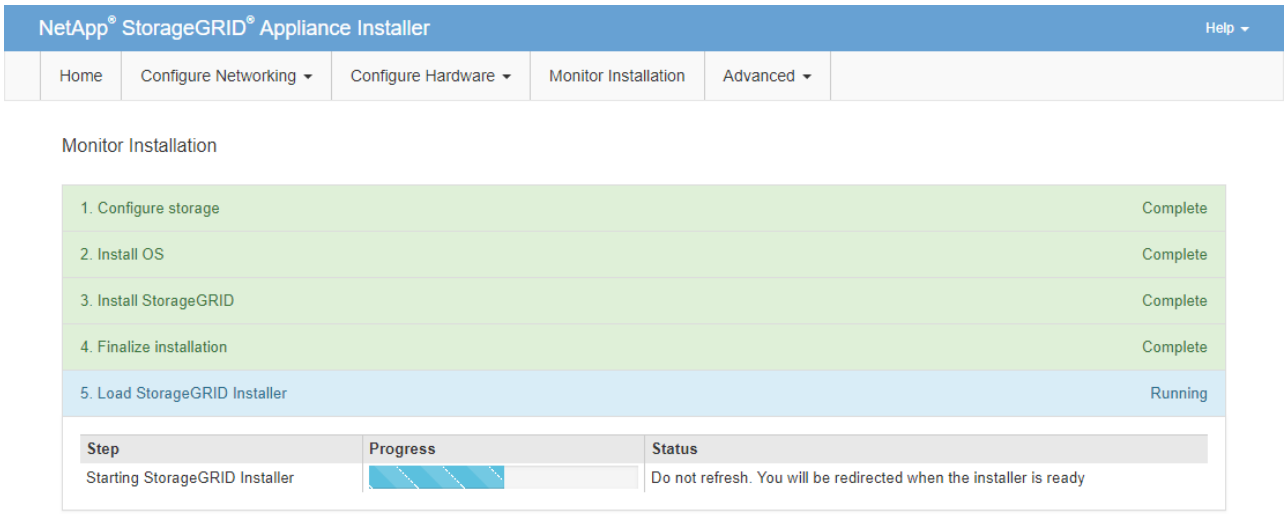
   - **2. Install OS**

     During this stage, the installer copies the base operating system image for StorageGRID from the primary Admin Node to the appliance or installs the base operating system from the installation package for the primary Admin Node.

3. Continue monitoring the installation progress until one of the following occurs:
   - For appliance Gateway Nodes or non-primary appliance Admin Nodes, the **Install StorageGRID** stage pauses and a message appears on the embedded console, prompting you to approve this node on the Admin Node using the Grid Manager.




- For appliance primary Admin Nodes, a fifth phase (Load StorageGRID Installer) appears. If the fifth phase is in progress for more than 10 minutes, refresh the page manually.




4. Go on to the next step of the recovery process for the type of appliance grid node that you are recovering.

| Type of recovery | Reference |
|---|---|
| Gateway Node | Selecting Start Recovery to configure a Gateway Node |
| Non-primary Admin Node | Selecting Start Recovery to configure a non-primary Admin Node |
| Primary Admin Node | Configuring the replacement primary Admin Node |

## How site recovery is performed by technical support

If an entire StorageGRID site fails or if multiple Storage Nodes fail, you must contact technical support. Technical support will assess your situation, develop a recovery plan, and then recover the failed nodes or site in a way that meets your business objectives, optimizes recovery time, and prevents unnecessary data loss.

Site recovery can only be performed by technical support.

StorageGRID systems are resilient to a wide variety of failures, and you can successfully perform many recovery and maintenance procedures yourself. However, it is difficult to create a simple, generalized site recovery procedure because the detailed steps depend on factors that are specific to your situation. For example:

- **Your business objectives**: After the complete loss of a StorageGRID site, you should evaluate how best to meet your business objectives. For example, do you want to rebuild the lost site in-place? Do you want to replace the lost StorageGRID site in a new location? Every customer's situation is different, and your recovery plan must be designed to address your priorities.
- **Exact nature of the failure**: Before beginning a site recovery, it is important to establish if any nodes at the failed site are intact or if any Storage Nodes contain recoverable objects. If you rebuild nodes or storage volumes that contain valid data, unnecessary data loss could occur.
- **Active ILM policy**: The number, type, and location of object copies in your grid is controlled by your active ILM policy. The specifics of your ILM policy can affect the amount of recoverable data, as well as the specific techniques required for recovery.

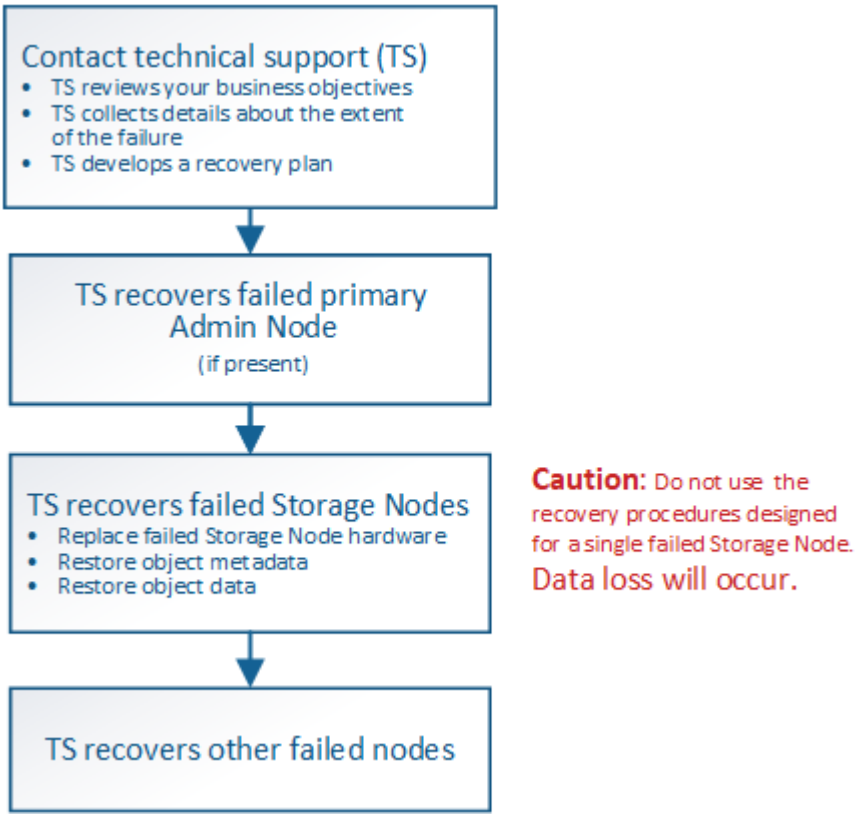
If a site contains the only copy of an object and the site is lost, the object is lost.

- **Bucket (or container) consistency**: The consistency level applied to a bucket (or container) affects whether StorageGRID fully replicates object metadata to all nodes and sites before telling a client that object ingest was successful. If your consistency level allows for eventual consistency, some object metadata might have been lost in the site failure. This can affect the amount of recoverable data and potentially the details of the recovery procedure.
- **History of recent changes**: The details of your recovery procedure can be affected by whether any maintenance procedures were in progress at the time of the failure or whether any recent changes were made to your ILM policy. Technical support must assess the recent history of your grid as well as its current situation before beginning a site recovery.

### Overview of site recovery

This is a general overview of the process that technical support uses to recover a failed site.

Site recovery can only be performed by technical support.




1. Contact technical support.

   Technical support does a detailed assessment of the failure and works with you to review your business objectives. Based on this information, technical support develops a recovery plan tailored for your situation.

2. Technical support recovers the primary Admin Node if it has failed.
3. Technical support recovers all Storage Nodes, following this outline:
   a. Replace Storage Node hardware or virtual machines as required.
   b. Restore object metadata to the failed site.

c. Restore object data to the recovered Storage Nodes.

> Data loss will occur if the recovery procedures for a single failed Storage Node are used.

> When an entire site has failed, specialized commands are required to successfully restore objects and object metadata.

4. Technical support recovers other failed nodes.

   After object metadata and data have been recovered, failed Gateway Nodes, non-primary Admin Nodes, or Archive Nodes can be recovered using standard procedures.

**Related information**

Site decommissioning

## Decommission procedure

You can perform a decommission procedure to permanently remove grid nodes or an entire site from the StorageGRID system.

To remove a grid node or a site, you perform one of the following decommission procedures:

- Perform a **node decommission** to remove one or more nodes, which can be at one or more sites. The nodes you remove can be online and connected to the StorageGRID system, or they can be offline and disconnected.
- Perform a **connected site decommission** to remove a site in which all nodes are connected to StorageGRID.
- Perform a **disconnected site decommission** to remove a site in which all nodes are disconnected from StorageGRID.

> Before performing a disconnected site decommission, you must contact your NetApp account representative. NetApp will review your requirements before enabling all steps in the Decommission Site wizard. You should not attempt a disconnected site decommission if you believe it might be possible to recover the site or to recover object data from the site.

If a site contains a mixture of connected (✔) and disconnected nodes (⬢ or ?), you must bring all offline nodes back online.

**Related information**

Grid node decommissioning

Site decommissioning

### Grid node decommissioning

You can use the node decommission procedure to remove one or more Storage Nodes, Gateway Nodes, or non-primary Admin Nodes at one or more sites. You cannot decommission the primary Admin Node or an Archive Node.

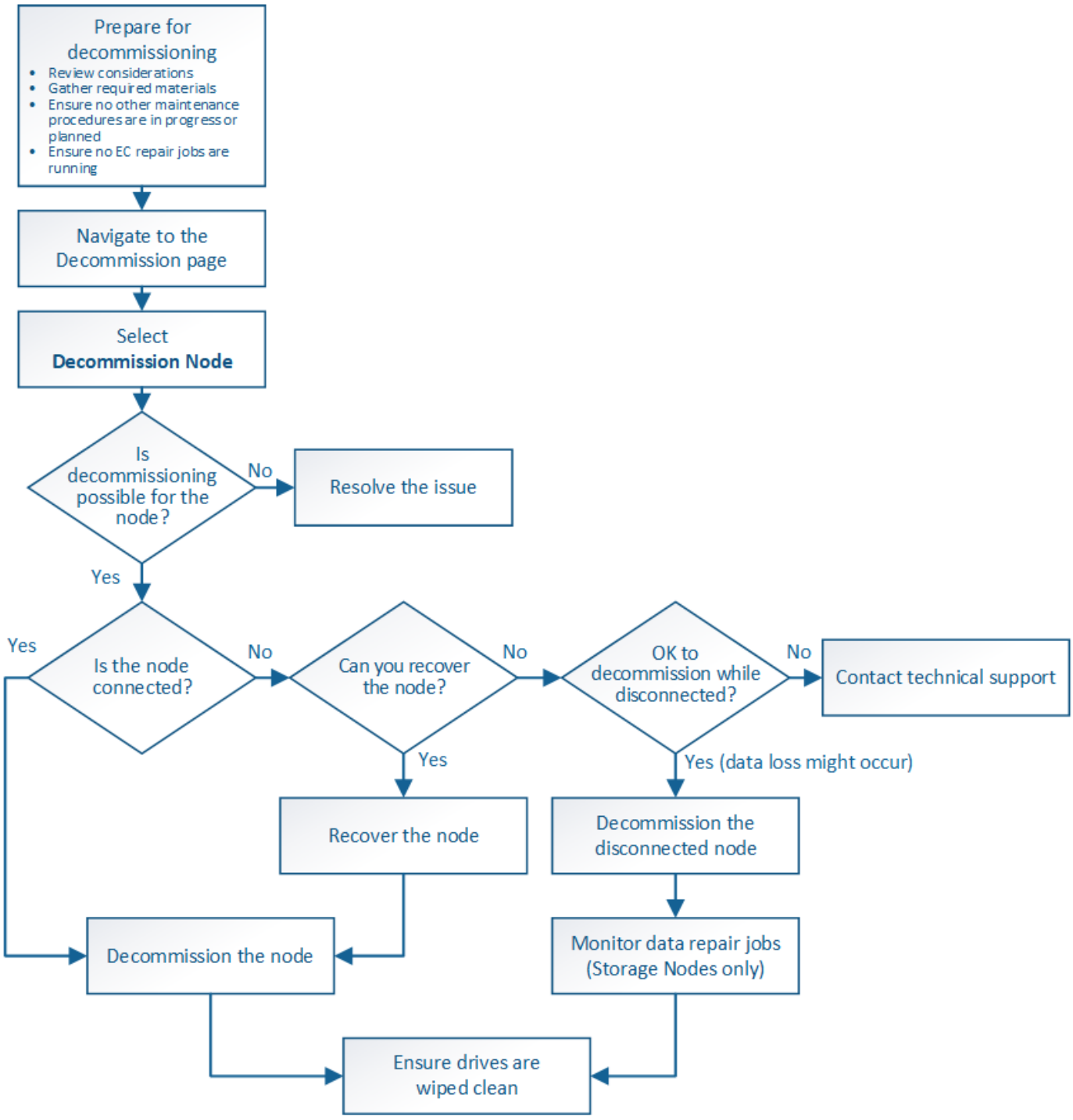
In general, you should decommission grid nodes only while they are connected to the StorageGRID system

and all nodes are in normal health (have green icons on the **Nodes** pages and on the **Decommission Nodes** page). However, if required, you can decommission a grid node that is disconnected. Before removing a disconnected node, make sure you understand the implications and restrictions of that process.

Use the node decommission procedure when any of the following are true:

- You have added a larger Storage Node to the system and you want to remove one or more smaller Storage Nodes, while at the same time preserving objects.
- You require less total storage.
- You no longer require a Gateway Node.
- You no longer require a non-primary Admin Node.
- Your grid includes a disconnected node that you cannot recover or bring back online.

The flowchart shows the high-level steps for decommissioning grid nodes.




**Steps**

- Preparing to decommission grid nodes
- Gathering required materials
- Accessing the Decommission Nodes page
- Decommissioning disconnected grid nodes
- Decommissioning connected grid nodes
- Pausing and resuming the decommission process for Storage Nodes
- Troubleshooting node decommissioning

#### Preparing to decommission grid nodes

You must review the considerations for removing grid nodes and confirm no repair jobs are active for erasure-coded data.

**Steps**

- Considerations for decommissioning Storage Nodes
- Checking data repair jobs

#### Considerations for decommissioning grid nodes

Before you start this procedure to decommission one or more nodes, you must understand the implications of removing each type of node. Upon the successful decommissioning of a node, its services will be disabled and the node will be automatically shut down.

You cannot decommission a node if doing so will leave the StorageGRID in an invalid state. The following rules are enforced:

- You cannot decommission the primary Admin Node.
- You cannot decommission Archive Nodes.
- You cannot decommission an Admin Node or a Gateway Node if one of its network interfaces is part of a high availability (HA) group.
- You cannot decommission a Storage Node if its removal would affect the ADC quorum.
- You cannot decommission a Storage Node if it is required for the active ILM policy.
- You should not decommission more than 10 Storage Nodes in a single Decommission Node procedure.
- You cannot decommission a connected node if your grid includes any disconnected nodes (nodes whose health is Unknown or Administratively Down). You must decommission or recover the disconnected nodes first.
- If your grid contains multiple disconnected nodes, the software requires you to decommission them at all the same time, which increases the potential for unexpected results.
- If a disconnected node cannot be removed (for example, a Storage Node that is required for the ADC quorum), no other disconnected node can be removed.
- If you want to replace an older appliance with a newer appliance, consider using the appliance node cloning procedure instead of decommissioning the old node and adding the new node in an expansion.

  Appliance node cloning

> Do not remove a grid node's virtual machine or other resources until instructed to do so in decommission procedures.

#### Considerations for decommissioning Admin Nodes or a Gateway Nodes

Review the following considerations before decommissioning an Admin Node or a Gateway Node.

- The decommission procedure requires exclusive access to some system resources, so you must confirm

that no other maintenance procedures are running.

- You cannot decommission the primary Admin Node.
- You cannot decommission an Admin Node or a Gateway Node if one of its network interfaces is part of a high availability (HA) group. You must first remove the network interfaces from the HA group. See the instructions for administering StorageGRID.
- As required, you can safely change the ILM policy while decommissioning a Gateway Node or an Admin Node.
- If you decommission an Admin Node and single sign-on (SSO) is enabled for your StorageGRID system, you must remember to remove the node's relying party trust from Active Directory Federation Services (AD FS).

**Related information**

Administer StorageGRID

### Considerations for decommissioning Storage Nodes

If you plan to decommission a Storage Node, you must understand how StorageGRID manages the object data and metadata on that node.

The following considerations and restrictions apply when decommissioning Storage Nodes:

- The system must, at all times, include enough Storage Nodes to satisfy operational requirements, including the ADC quorum and the active ILM policy. To satisfy this restriction, you might need to add a new Storage Node in an expansion operation before you can decommission an existing Storage Node.
- If the Storage Node is disconnected when you decommission it, the system must reconstruct the data using data from the connected Storage Nodes, which can result in data loss.
- When you remove a Storage Node, large volumes of object data must be transferred over the network. Although these transfers should not affect normal system operations, they can have an impact on the total amount of network bandwidth consumed by the StorageGRID system.
- Tasks associated with Storage Node decommissioning are given a lower priority than tasks associated with normal system operations. This means that decommissioning does not interfere with normal StorageGRID system operations, and does not need to be scheduled for a period of system inactivity. Because decommissioning is performed in the background, it is difficult to estimate how long the process will take to complete. In general, decommissioning finishes more quickly when the system is quiet, or if only one Storage Node is being removed at a time.
- It might take days or weeks to decommission a Storage Node. Plan this procedure accordingly. While the decommission process is designed to not impact system operations, it can limit other procedures. In general, you should perform any planned system upgrades or expansions before you remove grid nodes.
- Decommission procedures that involve Storage Nodes can be paused during certain stages to allow other maintenance procedures to run if needed, and resumed once they are complete.
- You cannot run data repair operations on any grid nodes when a decommission task is running.
- You should not make any changes to the ILM policy while a Storage Node is being decommissioned.
- When you remove a Storage Node, data on the node is migrated to other grid nodes; however, this data is not completely removed from the decommissioned grid node. To permanently and securely remove data, you must wipe the decommissioned grid node's drives after the decommission procedure is complete.
- When you decommission a Storage Node, the following alerts and alarms might be raised and you might receive related email and SNMP notifications:

- **Unable to communicate with node** alert. This alert is triggered when you decommission a Storage Node that includes the ADC service. The alert is resolved when the decommission operation completes.
- VSTU (Object Verification Status) alarm. This notice-level alarm indicates that the Storage Node is going into maintenance mode during the decommission process.
- CASA (Data Store Status) alarm. This major-level alarm indicates that the Cassandra database is going down because services have stopped.

**Related information**

Restoring object data to a storage volume, if required

Understanding the ADC quorum

Reviewing the ILM policy and storage configuration

Decommissioning disconnected Storage Nodes

Consolidating Storage Nodes

Decommissioning multiple Storage Nodes

### Understanding the ADC quorum

You might not be able to decommission certain Storage Nodes at a data center site if too few Administrative Domain Controller (ADC) services would remain after the decommissioning. This service, which is found on some Storage Nodes, maintains grid topology information and provides configuration services to the grid. The StorageGRID system requires a quorum of ADC services to be available at each site and at all times.

You cannot decommission a Storage Node if removing the node would cause the ADC quorum to no longer be met. To satisfy the ADC quorum during a decommissioning, a minimum of three Storage Nodes at each data center site must have the ADC service. If a data center site has more than three Storage Nodes with the ADC service, a simple majority of those nodes must remain available after the decommissioning ((0.5 * *`Storage Nodes with ADC`*) + 1).

For example, suppose a data center site currently includes six Storage Nodes with ADC services and you want to decommission three Storage Nodes. Because of the ADC quorum requirement, you must complete two decommission procedures, as follows:

- In the first decommission procedure, you must ensure that four Storage Nodes with ADC services remain available ((0.5 * 6) +1) . This means that you can only decommission two Storage Nodes initially.
- In the second decommission procedure, you can remove the third Storage Node because the ADC quorum now only requires three ADC services to remain available ((0.5 * 4) + 1).

If you need to decommission a Storage Node but are unable to because of the ADC quorum requirement, you must add a new Storage Node in an expansion and specify that it should have an ADC service. Then, you can decommission the existing Storage Node.

**Related information**

Expand your grid

### Reviewing the ILM policy and storage configuration

If you plan to decommission a Storage Node, you should review your StorageGRID system's ILM policy before starting the decommissioning process.

During decommissioning, all object data is migrated from the decommissioned Storage Node to other Storage Nodes.

> The ILM policy you have *during* the decommission will be the one used *after* the decommission. You must ensure this policy meets your data requirements both before you start the decommission and after the decommission is complete.

You should review the rules in the active ILM policy to ensure that the StorageGRID system will continue to have enough capacity of the correct type and in the correct locations to accommodate the decommissioning of a Storage Node.

Consider the following:

- Will it be possible for ILM evaluation services to copy object data such that ILM rules are satisfied?
- What happens if a site becomes temporarily unavailable while decommissioning is in progress? Can additional copies be made in an alternate location?
- How will the decommissioning process affect the final distribution of content? As described in "Consolidating Storage Nodes," you should add new Storage Nodes before decommissioning old ones. If you add a larger replacement Storage Node after decommissioning a smaller Storage Node, the old Storage Nodes could be close to capacity and the new Storage Node could have almost no content. Most write operations for new object data would then be directed at the new Storage Node, reducing the overall efficiency of system operations.
- Will the system, at all times, include enough Storage Nodes to satisfy the active ILM policy?

> An ILM policy that cannot be satisfied will lead to backlogs and alarms, and can halt operation of the StorageGRID system.

Verify that the proposed topology that will result from the decommissioning process satisfies the ILM policy by assessing the factors listed in the table.

| Area to assess | Notes |
|---|---|
| Available capacity | Will there be enough storage capacity to accommodate all of the object data stored in the StorageGRID system, including the permanent copies of object data currently stored on the Storage Node to be decommissioned?Will there be enough capacity to handle the anticipated growth in stored object data for a reasonable interval of time after decommissioning is complete? |
| Location of storage | If enough capacity remains in the StorageGRID system as a whole, is the capacity in the right locations to satisfy the StorageGRID system's business rules? |

| Area to assess | Notes |
| --- | --- |
| Storage type | Will there be enough storage of the appropriate type after decommissioning is complete? For example, ILM rules might dictate that content be moved from one type of storage to another as content ages. If so, you must ensure that enough storage of the appropriate type is available in the final configuration of the StorageGRID system. |

**Related information**

Consolidating Storage Nodes

Manage objects with ILM

Expand your grid

### Decommissioning disconnected Storage Nodes

You must understand what can happen if you decommission a Storage Node while it is disconnected (health is Unknown or Administratively Down).

When you decommission a Storage Node that is disconnected from the grid, StorageGRID uses data from other Storage Nodes to reconstruct the object data and metadata that was on the disconnected node. It does this by automatically starting data repair jobs at the end of the decommissioning process.

Before decommissioning a disconnected Storage Node, be aware of the following:

- You should never decommission a disconnected node unless you are sure it cannot be brought online or recovered.

  > Do not perform this procedure if you believe it might be possible to recover object data from the node. Instead, contact technical support to determine if node recovery is possible.

- If a disconnected Storage Node contains the only copy of an object, that object will be lost when you decommission the node. The data repair jobs can only reconstruct and recover objects if at least one replicated copy or enough erasure-coded fragments exist on Storage Nodes that are currently connected.
- When you decommission a disconnected Storage Node, the decommission procedure completes relatively quickly. However, the data repair jobs can take days or weeks to run and are not monitored by the decommission procedure. You must manually monitor these jobs and restart them as needed. See the instructions on monitoring data repair.

  Checking data repair jobs

- If you decommission more than one disconnected Storage Node at a time, data loss might occur. The system might not be able to reconstruct data if too few copies of object data, metadata, or erasure-coded fragments remain available.

  > If you have more than one disconnected Storage Node that you cannot recover, contact technical support to determine the best course of action.

### Consolidating Storage Nodes

## You can consolidate Storage Nodes to reduce the Storage Node count for a site or deployment while increasing storage capacity.

When you consolidate Storage Nodes, you expand the StorageGRID system to add new, larger capacity Storage Nodes and then decommission the old, smaller capacity Storage Nodes. During the decommission procedure, objects are migrated from the old Storage Nodes to the new Storage Nodes.

For example, you might add two new, larger capacity Storage Nodes to replace three older Storage Nodes. You would first use the expansion procedure to add the two new, larger Storage Nodes, and then use the decommission procedure to remove the three old, smaller capacity Storage Nodes.

By adding new capacity before removing existing Storage Nodes, you ensure a more balanced distribution of data across the StorageGRID system. You also reduce the possibility that an existing Storage Node might be pushed beyond the storage watermark level.

#### Related information

Expand your grid

### Decommissioning multiple Storage Nodes

## If you need to remove more than one Storage Node, you can decommission them either sequentially or in parallel.

- If you decommission Storage Nodes sequentially, you must wait for the first Storage Node to complete decommissioning before starting to decommission the next Storage Node.
- If you decommission Storage Nodes in parallel, the Storage Nodes simultaneously process decommission tasks for all Storage Nodes being decommissioned. This can result in a situation where all permanent copies of a file are marked as “read-only,” temporarily disabling deletion in grids where this functionality is enabled.

### Checking data repair jobs

## Before decommissioning a grid node, you must confirm that no data repair jobs are active. If any repairs have failed, you must restart them and allow them to complete before performing the decommission procedure.

If you need to decommission a disconnected Storage Node, you will also complete these steps after the decommission procedure completes in order to ensure the data repair job has completed successfully. You must ensure that any erasure-coded fragments that were on the removed node have been restored successfully.

These steps only apply to systems that have erasure-coded objects.

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@grid_node_IP`

      When you are logged in as root, the prompt changes from `$` to `#`.

   b. Enter the password listed in the `Passwords.txt` file.

c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

2. Check for running repairs: `repair-data show-ec-repair-status`

   - If you have never run a data repair job, the output is `No job found`. You do not need to restart any repair jobs.
   - If the data repair job was run previously or is running currently, the output lists information for the repair. Each repair has a unique repair ID. Go to the next step.

```
root@DC1-ADM1:~ # repair-data show-ec-repair-status

Repair ID Scope Start Time End Time State Est/Affected Bytes Repaired
Retry Repair
================================================================================
===========
949283 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:27:06.9 Success 17359
17359 No
949292 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:37:06.9 Failure 17359 0
Yes
949294 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:47:06.9 Failure 17359 0
Yes
949299 DC1-S-99-10(Volumes: 1,2) 2016-11-30T15:57:06.9 Failure 17359 0
Yes
```

3. If the State for all repairs is `Success`, you do not need to restart any repair jobs.
4. If the State for any repair is `Failure`, you must restart that repair.
   a. Obtain the repair ID for the failed repair from the output.
   b. Run the `repair-data start-ec-node-repair` command.

      Use the `--repair-id` option to specify the Repair ID. For example, if you want to retry a repair with repair ID 949292, run this command: `repair-data start-ec-node-repair --repair-id 949292`

   c. Continue to track the status of EC data repairs until the State for all repairs is `Success`.

#### Gathering required materials

Before performing a grid node decommission, you must obtain the following information.

| Item | Notes |
|---|---|
| Recovery Package `.zip` file | You must download the most recent Recovery Package `.zip` file (`sgws-recovery-package-`*`id-revision`*`.zip`). You can use the Recovery Package file to restore the system if a failure occurs. |

| Item | Notes |
| --- | --- |
| `Passwords.txt` file | This file contains the passwords required to access grid nodes on the command line and is included in the Recovery Package. |
| Provisioning passphrase | The passphrase is created and documented when the StorageGRID system is first installed. The provisioning passphrase is not in the `Passwords.txt` file. |
| Description of StorageGRID system's topology before decommissioning | If available, obtain any documentation that describes the system's current topology. |

**Related information**

Web browser requirements

Downloading the Recovery Package

#### Accessing the Decommission Nodes page

When you access the Decommission Nodes page in the Grid Manager, you can see at a glance which nodes can be decommissioned.

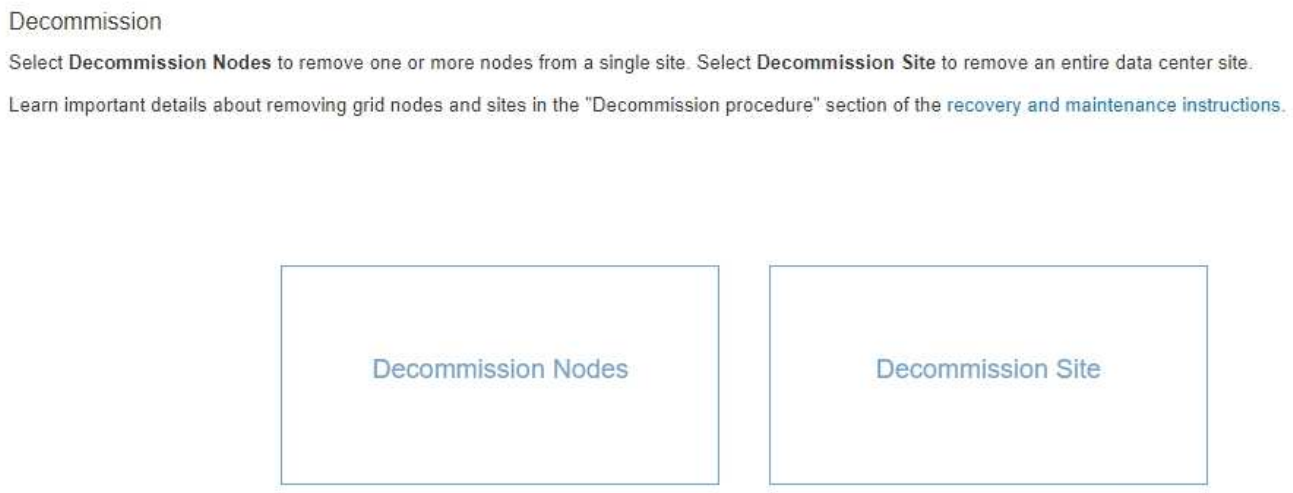
**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.

**Steps**

1. Select **Maintenance** > **Maintenance Tasks** > **Decommission**.

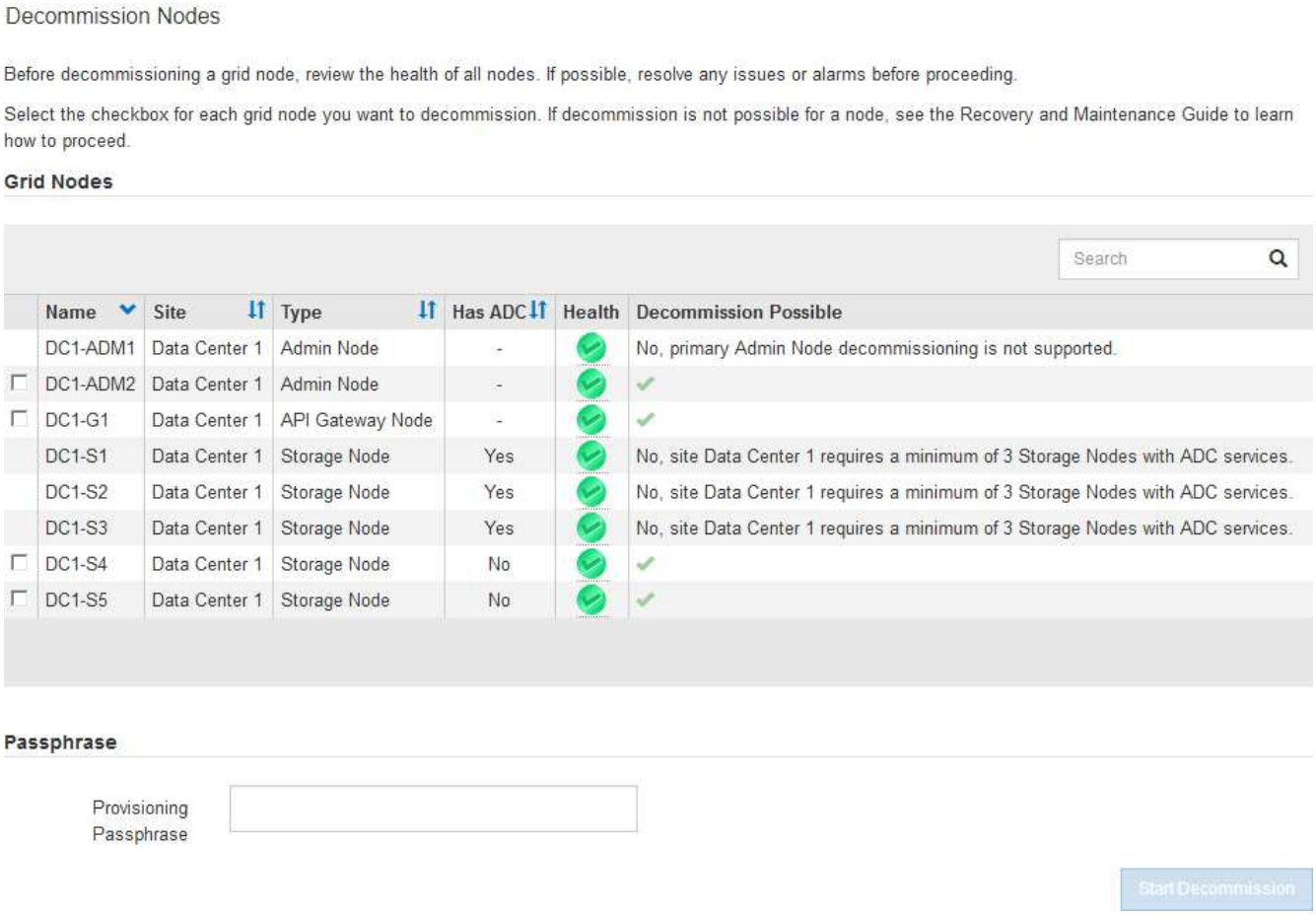
   The Decommission page appears.




2. Click the **Decommission Nodes** button.

   The Decommission Nodes page appears. From this page, you can:

- Determine which grid nodes can be decommissioned currently.
- See the health of all grid nodes
- Sort the list in ascending or descending order by **Name**, **Site**, **Type**, or **Has ADC**.
- Enter search terms to quickly find particular nodes. For example, this page shows all grid nodes in a single data center. The Decommission Possible column indicates that you can decommission the non-primary Admin Node, the Gateway Node, and two of the five Storage Nodes.




3. Review the **Decommission Possible** column for each node you want to decommission.

   If a grid node can be decommissioned, this column includes a green check mark, and the left-most column includes a check box. If a node cannot be decommissioned, this column describes the issue. If there is more than one reason a node cannot be decommissioned, the most critical reason is shown.

| Decommission Possible reason | Description | Steps to resolve |
|---|---|---|
| No, node type decommissioning is not supported. | You cannot decommission the primary Admin Node or an Archive Node. | None. |

| Decommission Possible reason | Description | Steps to resolve |
| --- | --- | --- |
| No, at least one grid node is disconnected.<br><br>**Note:** This message is shown for connected grid nodes only. | You cannot decommission a connected grid node if any grid node is disconnected.<br><br>The **Health** column includes one of these icons for grid nodes that are disconnected:<br><br>• (gray): Administratively Down<br>• (blue): Unknown | Go to the step that lists the decommission procedure choices. |
| No, one or more required nodes is currently disconnected and must be recovered.<br><br>**Note:** This message is shown for disconnected grid nodes only. | You cannot decommission a disconnected grid node if one or more required nodes is also disconnected (for example, a Storage Node that is required for the ADC quorum). | a. Review the Decommission Possible messages for all disconnected nodes.<br>b. Determine which nodes cannot be decommissioned because they are required.<br>◦ If the Health of a required node is Administratively Down, bring the node back online.<br>◦ If the health of a required node is Unknown, perform a node recovery procedure to recover the required node. |
| No, member of HA group(s): *x*. Before you can decommission this node, you must remove it from all HA groups. | You cannot decommission an Admin Node or a Gateway Node if a node interface belongs to a high availability (HA) group. | Edit the HA group to remove the node's interface or remove the entire HA group. See the instructions for administering StorageGRID. |
| No, site *x* requires a minimum of *n* Storage Nodes with ADC services. | **Storage Nodes only.** You cannot decommission a Storage Node if insufficient nodes would remain at the site to support ADC quorum requirements. | Perform an expansion. Add a new Storage Node to the site, and specify that it should have an ADC service. See information about the ADC quorum. |

| Decommission Possible reason | Description | Steps to resolve |
|---|---|---|
| No, one or more Erasure Coding profiles need at least *n* Storage Nodes. If the profile is not used in an ILM rule, you can deactivate it. | **Storage Nodes only.** You cannot decommission a Storage Node unless enough nodes would remain for the existing Erasure Coding profiles.<br><br>For example, if an Erasure Coding profile exists for 4+2 erasure coding, at least 6 Storage Nodes must remain. | For each affected Erasure Coding profile, perform one of the following steps, based on how the profile is being used:<br><br>• **Used in the active ILM policy**: Perform an expansion. Add enough new Storage Nodes to allow erasure coding to continue. See the instructions for expanding StorageGRID.<br>• **Used in an ILM rule but not in the active ILM policy**: Edit or delete the rule and then deactivate the Erasure Coding profile.<br>• **Not used in any ILM rule**: Deactivate the Erasure Coding profile.<br><br>**Note:** An error message appears if you attempt to deactivate an Erasure Coding profile and object data is still associated with the profile. You might need to wait several weeks before trying the deactivation process again.<br><br>Learn about deactivating an Erasure Coding profile in the instructions for managing objects with information lifecycle management. |

4. If decommissioning is possible for the node, determine which procedure you need to perform:

| If your grid includes… | Go to… |
|---|---|
| Any disconnected grid nodes | Decommissioning disconnected grid nodes |
| Only connected grid nodes | Decommissioning connected grid nodes |

**Related information**

Checking data repair jobs

Understanding the ADC quorum

Manage objects with ILM

Expand your grid

Administer StorageGRID

#### Decommissioning disconnected grid nodes

You might need to decommission a node that is not currently connected to the grid (one whose Health is Unknown or Administratively Down).

**What you'll need**

- You understand the requirements and considerations for decommissioning grid nodes.

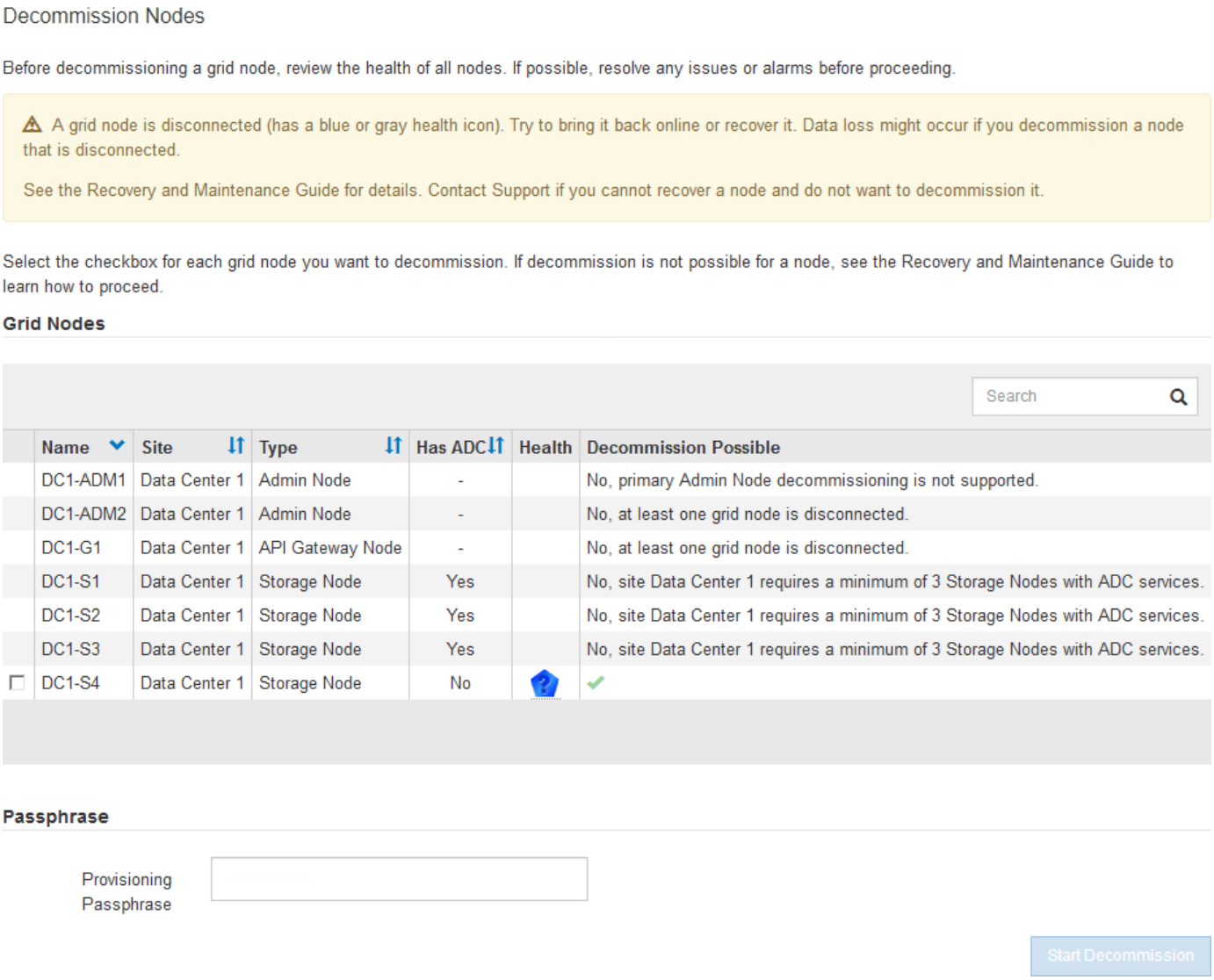
  Considerations for decommissioning grid nodes

- You have obtained all prerequisite items.
- You have ensured that no data repair jobs are active.

  Checking data repair jobs

- You have confirmed that Storage Node recovery is not in progress anywhere in the grid. If it is, you must wait until any Cassandra rebuild performed as part of the recovery is complete. You can then proceed with decommissioning.
- You have ensured that other maintenance procedures will not be run while the node decommission procedure is running, unless the node decommission procedure is paused.
- The **Decommission Possible** column for the disconnected node or nodes you want to decommission includes a green check mark.
- You must have the provisioning passphrase.

You can identify disconnected nodes by looking for Unknown (blue) or Administratively Down (gray) icons in the **Health** column. In the example, the Storage Node named DC1-S4 is disconnected; all of the other nodes are connected.




Before decommissioning any disconnected node, note the following:

- This procedure is primarily intended for removing a single disconnected node. If your grid contains multiple disconnected nodes, the software requires you to decommission them at all the same time, which increases the potential for unexpected results.

> Be very careful when decommissioning more than one disconnected grid node at a time, especially if you are selecting multiple disconnected Storage Nodes.

- If a disconnected node cannot be removed (for example, a Storage Node that is required for the ADC quorum), no other disconnected node can be removed.

Before decommissioning a disconnected **Storage Node**, note the following

- You should never decommission a disconnected Storage Node unless you are sure it cannot be brought online or recovered.

> If you believe that object data can still be recovered from the node, do not perform this procedure. Instead, contact technical support to determine if node recovery is possible.

- If you decommission more than one disconnected Storage Node, data loss might occur. The system might not be able to reconstruct data if not enough object copies, erasure-coded fragments, or object metadata remain available.

  > If you have more than one disconnected Storage Node that you cannot recover, contact technical support to determine the best course of action.

- When you decommission a disconnected Storage Node, StorageGRID starts data repair jobs at the end of the decommissioning process. These jobs attempt to reconstruct the object data and metadata that was stored on the disconnected node.
- When you decommission a disconnected Storage Node, the decommission procedure completes relatively quickly. However, the data repair jobs can take days or weeks to run and are not monitored by the decommission procedure. You must manually monitor these jobs and restart them as needed. See the instructions on monitoring data repair.

  Checking data repair jobs

- If you decommission a disconnected Storage Node that contains the only copy of an object, the object will be lost. The data repair jobs can only reconstruct and recover objects if at least one replicated copy or enough erasure-coded fragments exist on Storage Nodes that are currently connected.

Before decommissioning a disconnected **Admin Node** or **Gateway Node**, note the following:

- When you decommission a disconnected Admin Node, you will lose the audit logs from that node; however, these logs should also exist on the primary Admin Node.
- You can safely decommission a Gateway Node while it is disconnected.

**Steps**

1. Attempt to bring any disconnected grid nodes back online or to recover them.

   See the recovery procedures for instructions.

2. If you are unable to recover a disconnected grid node and you want to decommission it while it is disconnected, select the check box for that node.

   > If your grid contains multiple disconnected nodes, the software requires you to decommission them at all the same time, which increases the potential for unexpected results.

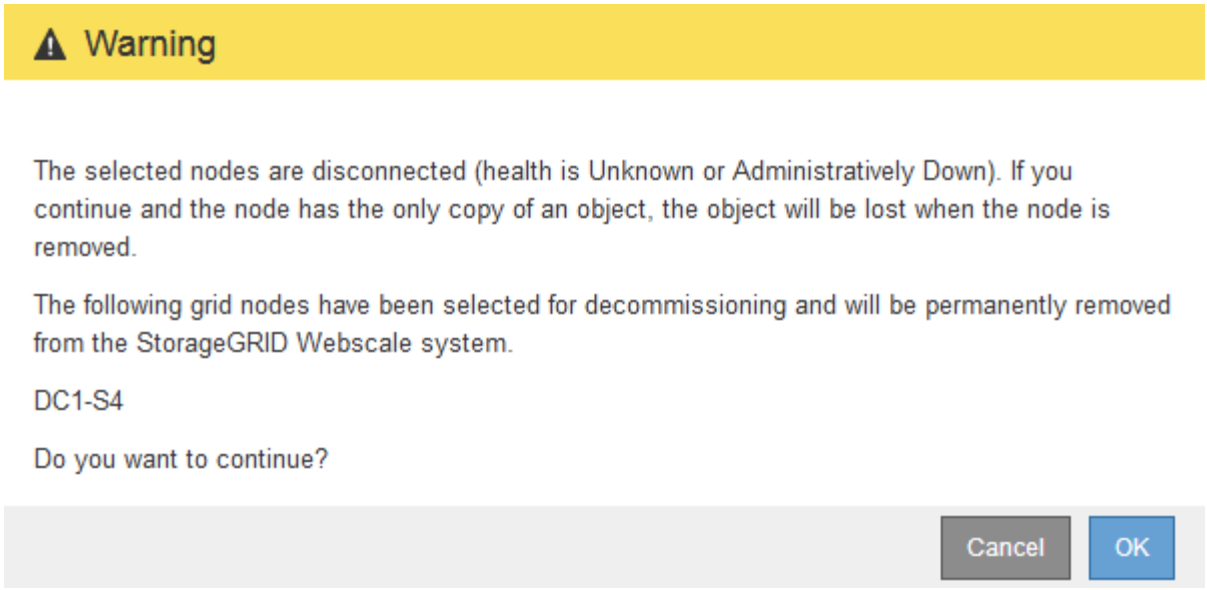
   > Be very careful when selecting to decommission more than one disconnected grid node at a time, especially if you are selecting multiple disconnected Storage Nodes. If you have more than one disconnected Storage Node that you cannot recover, contact technical support to determine the best course of action.

3. Enter the provisioning passphrase.

   The **Start Decommission** button is enabled.

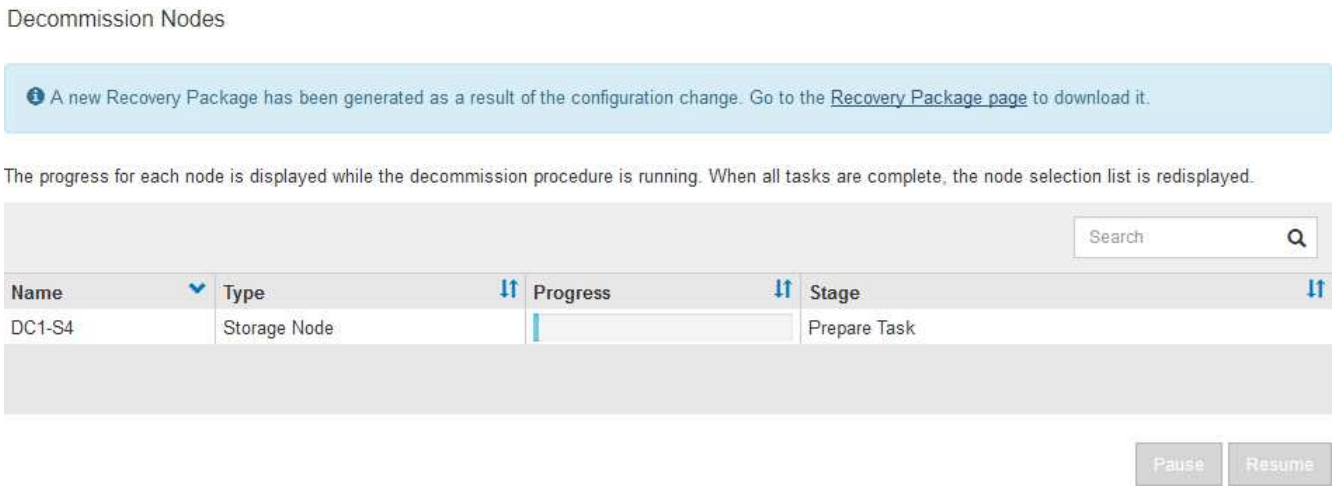
4. Click **Start Decommission**.

   A warning appears, indicating that you have selected a disconnected node and that object data will be lost if the node has the only copy of an object.




5. Review the list of nodes, and click **OK**.

   The decommission procedure starts, and the progress is displayed for each node. During the procedure, a new Recovery Package is generated containing the grid configuration change.




6. As soon as the new Recovery Package is available, click the link or select **Maintenance** > **System** > **Recovery Package** to access the Recovery Package page. Then, download the `.zip` file.

   See the instructions for downloading the Recovery Package.

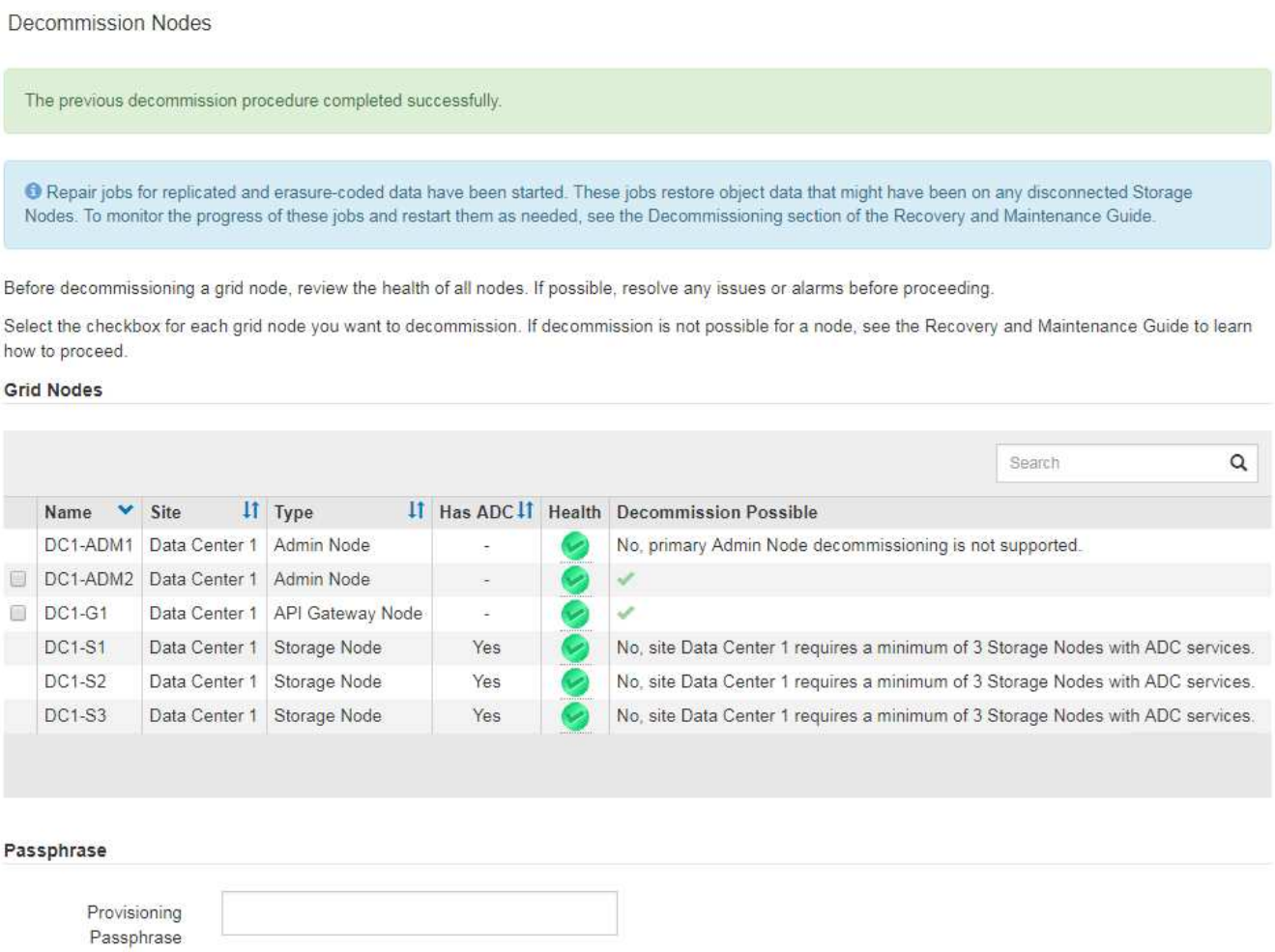
   > Download the Recovery Package as soon as possible to ensure you can recover your grid if something goes wrong during the decommission procedure.

   > The Recovery Package file must be secured because it contains encryption keys and passwords that can be used to obtain data from the StorageGRID system.

7. Periodically monitor the Decommission page to ensure that all selected nodes are decommissioned successfully.

   Storage Nodes can take days or weeks to decommission. When all tasks are complete, the node selection list is redisplayed with a success message. If you decommissioned a disconnected Storage Node, an

information message indicates that the repair jobs have been started.




8. After the nodes have shut down automatically as part of the decommission procedure, remove any remaining virtual machines or other resources that are associated with the decommissioned node.

   Do not perform this step until the nodes have shut down automatically.

9. If you are decommissioning a Storage Node, monitor the status of the data repair jobs that are automatically started during the decommissioning process.
   a. Select **Support** > **Tools** > **Grid Topology**.
   b. Select **StorageGRID deployment** at the top of the Grid Topology tree.
   c. On the Overview tab, locate the ILM Activity section.
   d. Use a combination of the following attributes to determine, as well as possible, if replicated repairs are complete.

      Cassandra inconsistencies might be present, and failed repairs are not tracked.

      - **Repairs Attempted (XRPA)**: Use this attribute to track the progress of replicated repairs. This attribute increases each time a Storage Node tries to repair a high-risk object. When this attribute does not increase for a period longer than the current scan period (provided by the **Scan Period — Estimated** attribute), it means that ILM scanning found no high-risk objects that need to be repaired on any nodes.

> ⓘ High-risk objects are objects that are at risk of being completely lost. This does not include objects that do not satisfy their ILM configuration.

- **Scan Period — Estimated (XSCM)**: Use this attribute to estimate when a policy change will be applied to previously ingested objects. If the **Repairs Attempted** attribute does not increase for a period longer than the current scan period, it is probable that replicated repairs are done. Note that the scan period can change. The **Scan Period — Estimated (XSCM)** attribute applies to the entire grid and is the maximum of all node scan periods. You can query the **Scan Period — Estimated** attribute history for the grid to determine an appropriate time frame.

e. Use the following commands to track or restart repairs:

   - Use the `repair-data show-ec-repair-status` command to track repairs of erasure coded data.
   - Use the `repair-data start-ec-node-repair` command with the `--repair-id` option to restart a failed repair. See the instructions for checking data repair jobs.

10. Continue to track the status of EC data repairs until all repair jobs have completed successfully.

    As soon as the disconnected nodes have been decommissioned and all data repair jobs have been completed, you can decommission any connected grid nodes as required.

Complete these steps after you complete the decommission procedure:

- Ensure that the drives of the decommissioned grid node are wiped clean. Use a commercially available data wiping tool or service to permanently and securely remove data from the drives.
- If you decommissioned an appliance node and the data on the appliance was protected using node encryption, use the StorageGRID Appliance Installer to clear the key management server configuration (Clear KMS). You must clear the KMS configuration if you want to add the appliance to another grid.

  SG100 & SG1000 services appliances

  SG5600 storage appliances

  SG5700 storage appliances

  SG6000 storage appliances

**Related information**

Grid node recovery procedures

Downloading the Recovery Package

Checking data repair jobs

#### Decommissioning connected grid nodes

You can decommission and permanently remove nodes that are connected to the grid.

**What you'll need**

- You understand the requirements and considerations for decommissioning grid nodes.

  Considerations for decommissioning grid nodes

- You have gathered all required materials.
- You have ensured that no data repair jobs are active.
- You have confirmed that Storage Node recovery is not in progress anywhere in the grid. If it is, you must wait until any Cassandra rebuild performed as part of the recovery is complete. You can then proceed with decommissioning.
- You have ensured that other maintenance procedures will not be run while the node decommission procedure is running, unless the node decommission procedure is paused.
- You have the provisioning passphrase.
- Grid nodes are connected.
- The **Decommission Possible** column for the node or nodes you want to decommission include a green checkmark.
- All grid nodes have Normal (green) health . If you see one of these icons in the **Health** column, you must try to resolve the issue:

| Icon | Color | Severity |
|---|---|---|
| | Yellow | Notice |
| | Light orange | Minor |
| | Dark orange | Major |
| | Red | Critical |

- If you previously decommissioned a disconnected Storage Node, the data repair jobs have all completed successfully. See the instructions for checking data repair jobs.

> Do not remove a grid node's virtual machine or other resources until instructed to do so in this procedure.

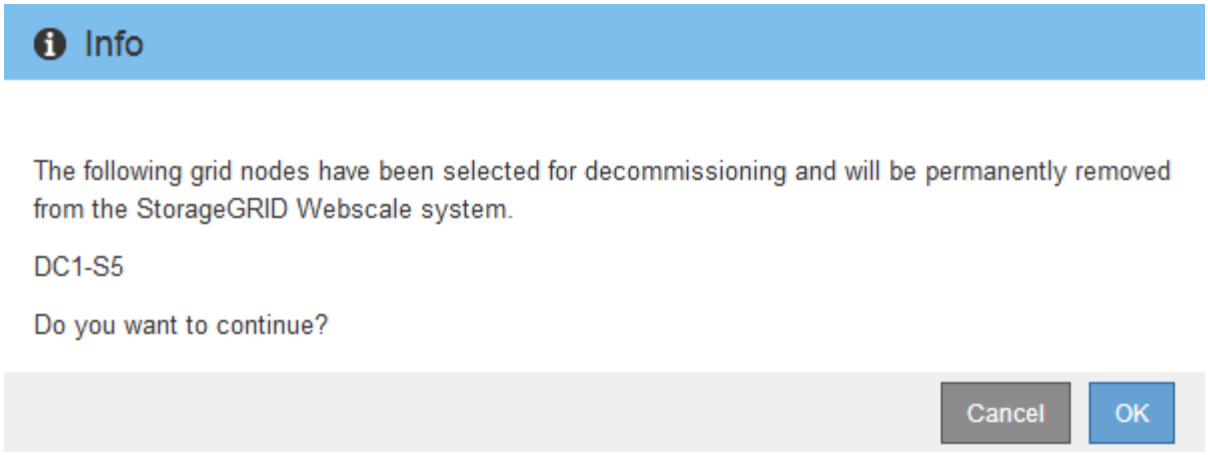
**Steps**

1. From the Decommission Nodes page, select the check box for each grid node you want to decommission.
2. Enter the provisioning passphrase.

   The **Start Decommission** button is enabled.

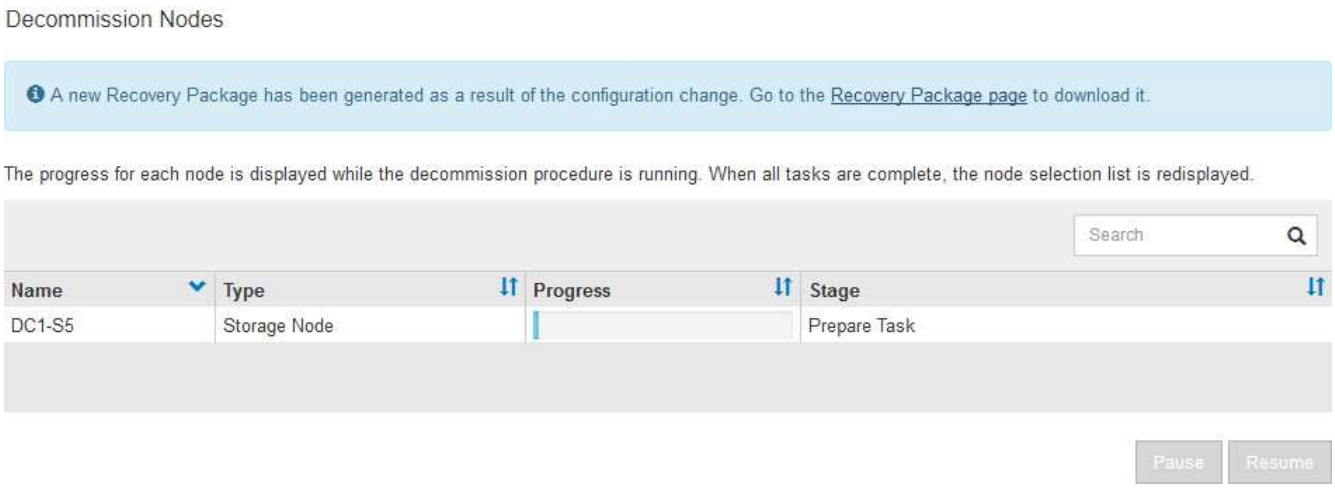
3. Click **Start Decommission**.

   A confirmation dialog box appears.




4. Review the list of selected nodes, and click **OK**.

   The node decommission procedure starts, and the progress is displayed for each node. During the procedure, a new Recovery Package is generated to show the grid configuration change.




> Do not take a Storage Node offline after the decommission procedure has started. Changing the state might result in some content not being copied to other locations.

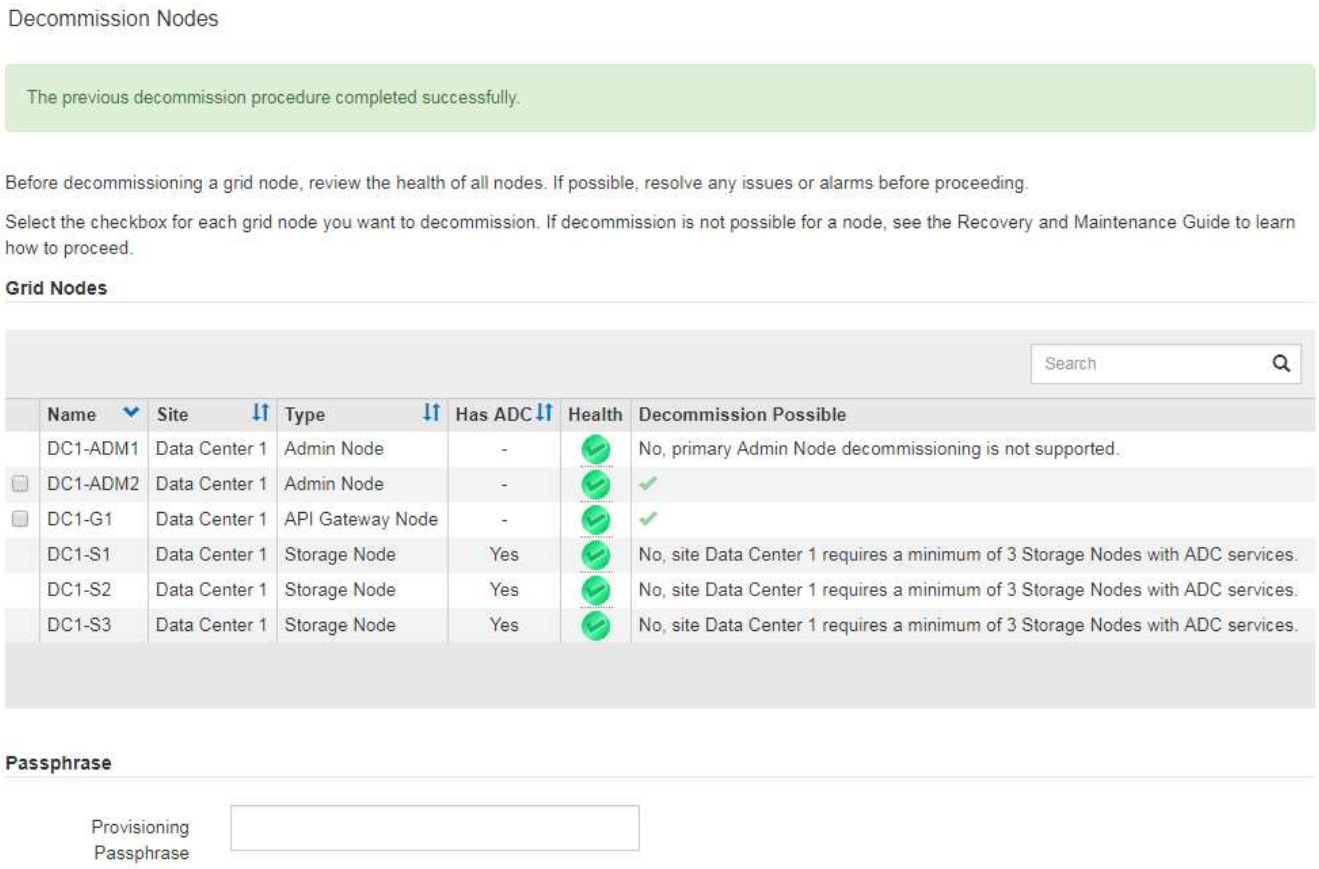
5. As soon as the new Recovery Package is available, click the link or select **Maintenance** > **System** > **Recovery Package** to access the Recovery Package page. Then, download the `.zip` file.

   See the instructions for downloading the Recovery Package.

> Download the Recovery Package as soon as possible to ensure you can recover your grid if something goes wrong during the decommission procedure.

6. Periodically monitor the Decommission Nodes page to ensure that all selected nodes are decommissioned successfully.

   Storage Nodes can take days or weeks to decommission. When all tasks are complete, the node selection list is redisplayed with a success message.




7. Follow the appropriate step for your platform. For example:
   - **Linux**: You might want to detach the volumes and delete the node configuration files you created during installation.
   - **VMware**: You might want to use the vCenter “Delete from Disk” option to delete the virtual machine. You might also need to delete any data disks that are independent of the virtual machine.
   - **StorageGRID appliance**: The appliance node automatically reverts to an undeployed state where you can access the StorageGRID Appliance Installer. You can power off the appliance or add it to another StorageGRID system.

Complete these steps after you complete the node decommission procedure:

- Ensure that the drives of the decommissioned grid node are wiped clean. Use a commercially available data wiping tool or service to permanently and securely remove data from the drives.
- If you decommissioned an appliance node and the data on the appliance was protected using node encryption, use the StorageGRID Appliance Installer to clear the key management server configuration (Clear KMS). You must clear the KMS configuration if you want to use the appliance in another grid.

  SG100 & SG1000 services appliances

  SG5600 storage appliances

  SG5700 storage appliances

  SG6000 storage appliances

**Related information**

Checking data repair jobs

Downloading the Recovery Package

Install Red Hat Enterprise Linux or CentOS

#### Pausing and resuming the decommission process for Storage Nodes

If necessary, you can pause the decommission procedure for a Storage Node during certain stages. You must pause decommissioning on a Storage Node before you can start a second maintenance procedure. After the other procedure is finished, you can resume decommissioning.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.

**Steps**

1. Select **Maintenance** > **Maintenance Tasks** > **Decommission**.

   The Decommission page appears.

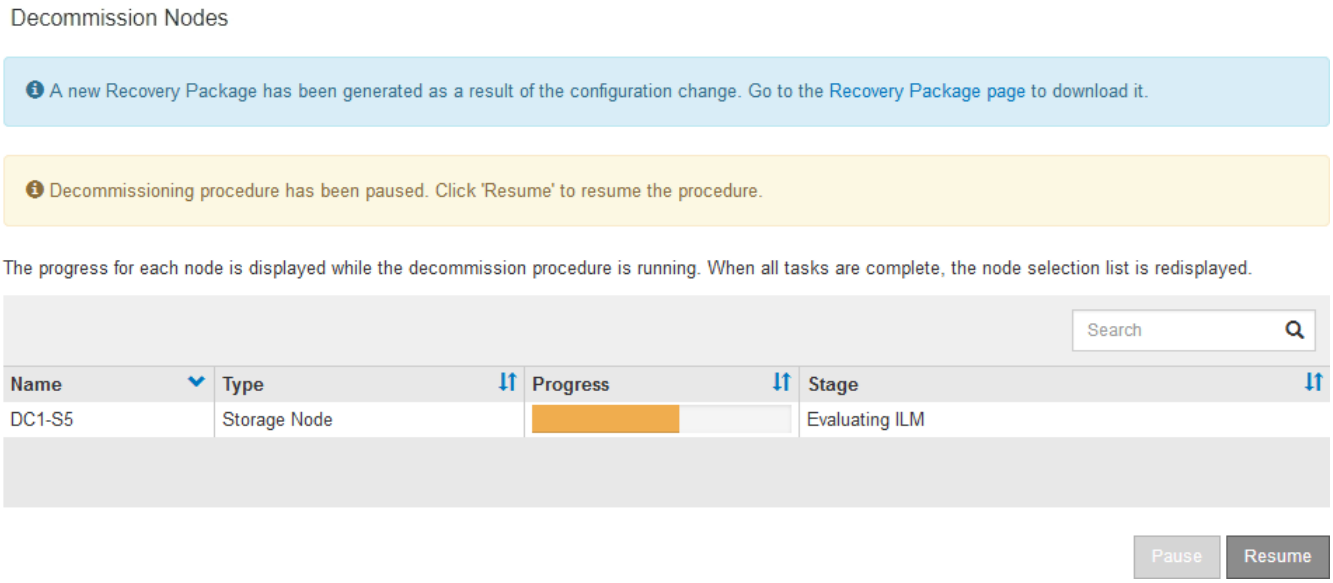
2. Click **Decommission Nodes**.

   The Decommission Nodes page appears. When the decommission procedure reaches either of the following stages, the **Pause** button is enabled.

   - Evaluating ILM
   - Decommissioning Erasure Coded data

3. Click **Pause** to suspend the procedure.

   The current stage is paused, and the **Resume** button is enabled.




4. After the other maintenance procedure is finished, click **Resume** to proceed with the decommission.

#### Troubleshooting node decommissioning

If the node decommission procedure stops because of an error, you can take specific steps to troubleshoot the problem.

**What you'll need**

You must be signed in to the Grid Manager using a supported browser.

**About this task**

If you shut down the grid node being decommissioned, the task stops until the grid node is restarted. The grid node must be online.

**Steps**

1. Select **Support** > **Tools** > **Grid Topology**.
2. In the Grid Topology tree, expand each Storage Node entry, and verify that the DDS and LDR services are both online.

   To perform Storage Node decommissioning, the StorageGRID system's DDS services (hosted by Storage Nodes) must be online. This is a requirement of the ILM re-evaluation.

3. To view the active grid tasks, select **primary Admin Node** > **CMN** > **Grid Tasks** > **Overview**.
4. Check the status of the decommissioning grid task.
   a. If the status of the decommissioning grid task indicates a problem with saving grid task bundles, select **primary Admin Node** > **CMN** > **Events** > **Overview**
   b. Check the number of Available Audit Relays.

      If the attribute Available Audit Relay is one or greater, the CMN service is connected to at least one ADC service. ADC services act as Audit Relays.

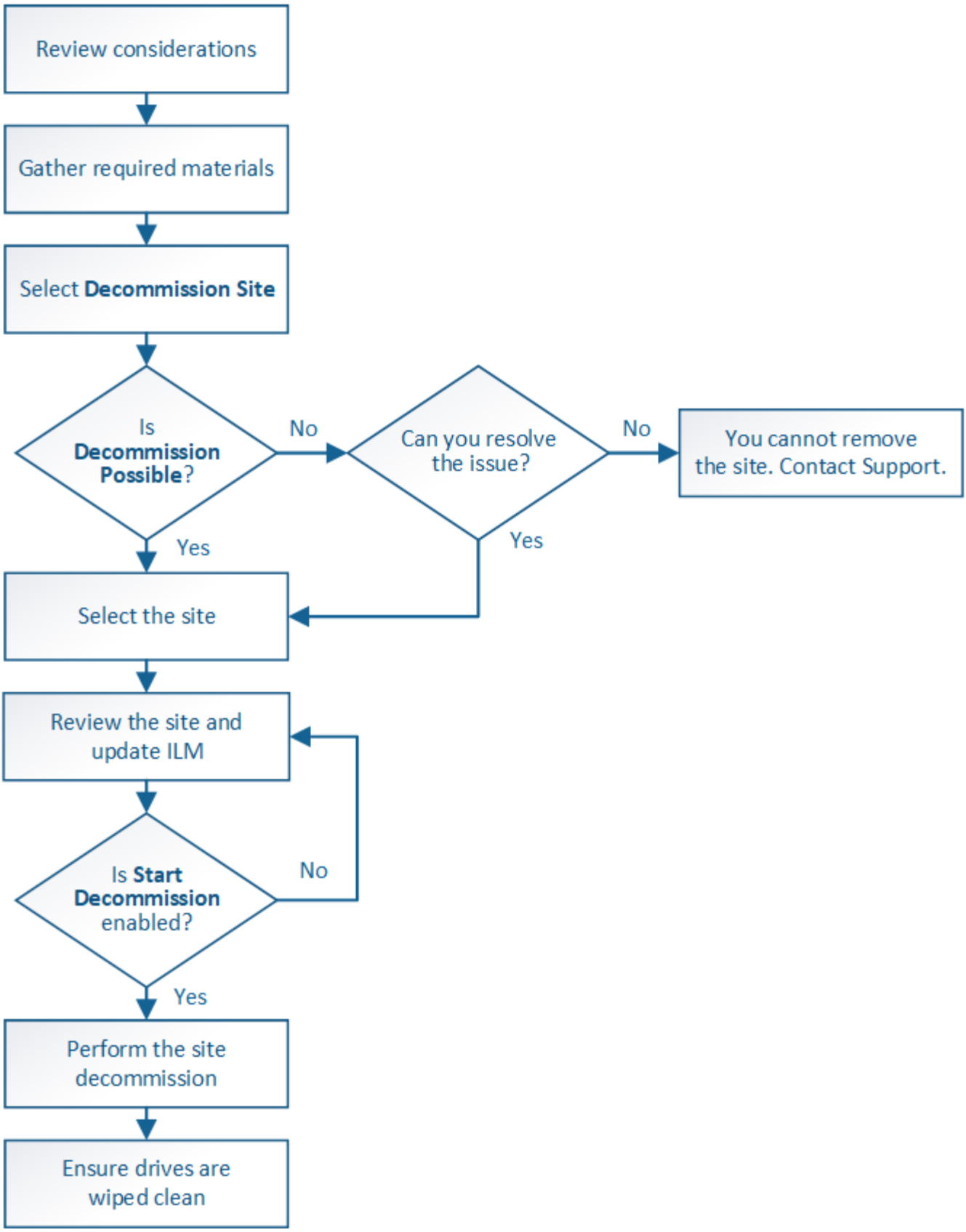
      The CMN service must be connected to at least one ADC service and a majority (50 percent plus one) of the StorageGRID system's ADC services must be available in order for a grid task to move from one stage of decommissioning to another and finish.

   c. If the CMN service is not connected to enough ADC services, ensure that Storage Nodes are online, and check network connectivity between the primary Admin Node and Storage Nodes.

#### Site decommissioning

You might need to remove a data center site from the StorageGRID system. To remove a site, you must decommission it.

The flowchart shows the high-level steps for decommissioning a site.




**Steps**

- Considerations for removing a site
- Gathering required materials

- Step 1: Select Site
- Step 2: View Details
- Step 3: Revise ILM Policy
- Step 4: Remove ILM References
- Step 5: Resolve Node Conflicts (and start decommission)
- Step 6: Monitor Decommission

**Considerations for removing a site**

## Before using the site decommission procedure to remove a site, you must review the considerations.

### What happens when you decommission a site

When you decommission a site, StorageGRID permanently removes all nodes at the site and the site itself from the StorageGRID system.

When the site decommission procedure is complete:

- You can no longer use StorageGRID to view or access the site or any of the nodes at the site.
- You can no longer use any storage pools or Erasure Coding profiles that referred to the site. When StorageGRID decommissions a site, it automatically removes these storage pools and deactivates these Erasure Coding profiles.

### Differences between connected site and disconnected site decommission procedures

You can use the site decommission procedure to remove a site in which all nodes are connected to StorageGRID (referred to as a connected site decommission) or to remove a site in which all nodes are disconnected from StorageGRID (referred to as a disconnected site decommission). Before you begin, you must understand the differences between these procedures.

If a site contains a mixture of connected (✔) and disconnected nodes ( or ), you must bring all offline nodes back online.

- A connected site decommission allows you to remove an operational site from the StorageGRID system. For example, you can perform a connected site decommission to remove a site that is functional but no longer needed.
- When StorageGRID removes a connected site, it uses ILM to manage the object data at the site. Before you can start a connected site decommission, you must remove the site from all ILM rules and activate a new ILM policy. The ILM processes to migrate object data and the internal processes to remove a site can occur at the same time, but the best practice is to allow the ILM steps to complete before you start the actual decommission procedure.
- A disconnected site decommission allows you to remove a failed site from the StorageGRID system. For example, you can perform a disconnected site decommission to remove a site that has been destroyed by a fire or flood.

  When StorageGRID removes a disconnected site, it considers all nodes to be unrecoverable and makes no attempt to preserve data. However, before you can start a disconnected site decommission, you must remove the site from all ILM rules and activate a new ILM policy.

Before performing a disconnected site decommission procedure, you must contact your NetApp account representative. NetApp will review your requirements before enabling all steps in the Decommission Site wizard. You should not attempt a disconnected site decommission if you believe it might be possible to recover the site or to recover object data from the site.

### General requirements for removing a connected or a disconnected site

Before removing a connected or disconnected site, you must be aware of the following requirements:

- You cannot decommission a site that includes the primary Admin Node.
- You cannot decommission a site that includes an Archive Node.
- You cannot decommission a site if any of the nodes have an interface that belongs to a high availability (HA) group. You must either edit the HA group to remove the node's interface or remove the entire HA group.
- You cannot decommission a site if it contains a mixture of connected (✔) and disconnected (? or ⬢) nodes.
- You cannot decommission a site if any node at any other site is disconnected (? or ⬢).
- You cannot start the site decommission procedure if an ec-node-repair operation is in progress. See the following topic to track repairs of erasure-coded data.

  Checking data repair jobs

- While the site decommission procedure is running:
  - You cannot create ILM rules that refer to the site being decommissioned. You also cannot edit an existing ILM rule to refer to the site.
  - You cannot perform other maintenance procedures, such as expansion or upgrade.

    If you need to perform another maintenance procedure during a connected site decommission, you can pause the procedure while the Storage Nodes are being removed. The **Pause** button is enabled during the "Decommissioning Replicated and Erasure Coded Data" stage.

  - If you need to recover any node after starting the site decommission procedure, you must contact support.
- You cannot decommission more than one site at a time.
- If the site includes one or more Admin Nodes and single sign-on (SSO) is enabled for your StorageGRID system, you must remove all relying party trusts for the site from Active Directory Federation Services (AD FS).

### Requirements for information lifecycle management (ILM)

As part of removing a site, you must update your ILM configuration. The Decommission Site wizard guides you through a number of prerequisite steps to ensure the following:

- The site is not referred to by the active ILM policy. If it is, you must create and activate a new ILM policy with new ILM rules.
- No proposed ILM policy exists. If you have a proposed policy, you must delete it.

- No ILM rules refer to the site, even if those rules are not used in the active or proposed policy. You must delete or edit all rules that refer to the site.

When StorageGRID decommissions the site, it will automatically deactivate any unused Erasure Coding profiles that refer to the site, and it will automatically delete any unused storage pools that refer to the site. The system-default All Storage Nodes storage pool is removed because it uses all sites.

> Before you can remove a site, you might be required to create new ILM rules and activate a new ILM policy. These instructions assume that you have a good understanding of how ILM works and that you are familiar with creating storage pools, Erasure Coding profiles, ILM rules, and simulating and activating an ILM policy. See the instructions for managing objects with information lifecycle management.

Manage objects with ILM

### Considerations for the object data at a connected site

If you are performing a connected site decommission, you must decide what to do with existing object data at the site when you create new ILM rules and a new ILM policy. You can do either or both of the following:

- Move object data from the selected site to one or more other sites in your grid.

  **Example for moving data**: Suppose you want to decommission a site in Raleigh because you added a new site in Sunnyvale. In this example, you want to move all object data from the old site to the new site. Before updating your ILM rules and ILM policy, you must review the capacity at both sites. You must ensure that the Sunnyvale site has enough capacity to accommodate the object data from the Raleigh site and that adequate capacity will remain in Sunnyvale for future growth.

  > To ensure that adequate capacity is available, you might need to add storage volumes or Storage Nodes to an existing site or add a new site before you perform this procedure. See the instructions for expanding a StorageGRID system.

- Delete object copies from the selected site.

  **Example for deleting data**: Suppose you currently use a 3-copy ILM rule to replicate object data across three sites. Before decommissioning a site, you can create an equivalent 2-copy ILM rule to store data at only two sites. When you activate a new ILM policy that uses the 2-copy rule, StorageGRID deletes the copies from the third site because they no longer satisfy ILM requirements. However, the object data will still be protected and the capacity of the two remaining sites will stay the same.

  > Never create a single-copy ILM rule to accommodate the removal of a site. An ILM rule that creates only one replicated copy for any time period puts data at risk of permanent loss. If only one replicated copy of an object exists, that object is lost if a Storage Node fails or has a significant error. You also temporarily lose access to the object during maintenance procedures such as upgrades.

### Additional requirements for a connected site decommission

Before StorageGRID can remove a connected site, you must ensure the following:

- All nodes in your StorageGRID system must have a Connection State of **Connected** (✔); however, the nodes can have active alerts.

You can complete Steps 1-4 of the Decommission Site wizard if one or more nodes are disconnected. However, you cannot complete Step 5 of the wizard, which starts the decommission process, unless all nodes are connected.

- If the site you plan to remove contains a Gateway Node or an Admin Node that is used for load balancing, you might need to perform an expansion procedure to add an equivalent new node at another site. Be sure clients can connect to the replacement node before starting the site decommission procedure.
- If the site you plan to remove contains any Gateway Node or Admin Nodes that are in an high availability (HA) group, you can complete Steps 1-4 of the Decommission Site wizard. However, you cannot complete Step 5 of the wizard, which starts the decommission process, until you remove these nodes from all HA groups. If existing clients connect to an HA group that includes nodes from the site, you must ensure they can continue to connect to StorageGRID after the site is removed.
- If clients connect directly to Storage Nodes at the site you are planning to remove, you must ensure that they can connect to Storage Nodes at other sites before starting the site decommission procedure.
- You must provide sufficient space on the remaining sites to accommodate any object data that will be moved because of changes to the active ILM policy. In some cases, you might need to expand your StorageGRID system by adding Storage Nodes, storage volumes, or new sites before you can complete a connected site decommission.
- You must allow adequate time for the decommission procedure to complete. StorageGRID ILM processes might take days, weeks, or even months to move or delete object data from the site before the site can be decommissioned.

Moving or deleting object data from a site might take days, weeks, or even months, depending on the amount of data at the site, the load on your system, network latencies, and the nature of the required ILM changes.

- Whenever possible, you should complete Steps 1-4 of the Decommission Site wizard as early as you can. The decommission procedure will complete more quickly and with fewer disruptions and performance impacts if you allow data to be moved from the site before starting the actual decommission procedure (by selecting **Start Decommission** in Step 5 of the wizard).

**Additional requirements for a disconnected site decommission**

Before StorageGRID can remove a disconnected site, you must ensure the following:

- You have contacted your NetApp account representative. NetApp will review your requirements before enabling all steps in the Decommission Site wizard.

You should not attempt a disconnected site decommission if you believe it might be possible to recover the site or to recover any object data from the site.

- All nodes at the site must have a Connection State of one of the following:
  - **Unknown** ( ): The node is not connected to the grid for an unknown reason. For example, the network connection between nodes has been lost or the power is down.
  - **Administratively Down** ( ): The node is not connected to the grid for an expected reason. For example, the node or services on the node have been gracefully shut down.
- All nodes at all other sites must have a Connection State of **Connected** ( ); however, these other nodes can have active alerts.
- You must understand that you will no longer be able to use StorageGRID to view or retrieve any object

data that was stored at the site. When StorageGRID performs this procedure, it makes no attempt to preserve any data from the disconnected site.

> If your ILM rules and policy were designed to protect against the loss of a single site, copies of your objects still exist on the remaining sites.

- You must understand that if the site contained the only copy of an object, the object is lost and cannot be retrieved.

### Considerations for consistency controls when you remove a site

The consistency level for an S3 bucket or Swift container determines whether StorageGRID fully replicates object metadata to all nodes and sites before telling a client that object ingest was successful. The consistency level makes a trade-off between the availability of the objects and the consistency of those objects across different Storage Nodes and sites.

When StorageGRID removes a site, it needs to ensure that no data is written to the site being removed. As a result, it temporarily overrides the consistency level for each bucket or container. After you start the site decommission process, StorageGRID temporarily uses strong-site consistency to prevent object metadata from being written to the site being removed.

As a result of this temporary override, be aware that any client write, update, and delete operations that occur during a site decommission can fail if multiple nodes become unavailable at the remaining sites.

#### Related information

How site recovery is performed by technical support

Manage objects with ILM

Expand your grid

#### Gathering required materials

Before you decommission a site, you must obtain the following materials.

| Item | Notes |
|---|---|
| Recovery Package `.zip` file | You must download the most recent Recovery Package `.zip` file (`sgws-recovery-package-id-revision.zip`). You can use the Recovery Package file to restore the system if a failure occurs. |
| `Passwords.txt` file | This file contains the passwords required to access grid nodes on the command line and is included in the Recovery Package. |
| Provisioning passphrase | The passphrase is created and documented when the StorageGRID system is first installed. The provisioning passphrase is not in the `Passwords.txt` file. |

| Item | Notes |
| --- | --- |
| Description of StorageGRID system's topology before decommissioning | If available, obtain any documentation that describes the system's current topology. |

**Related information**

Web browser requirements

Downloading the Recovery Package

#### Step 1: Select Site

## To determine if a site can be decommissioned, start by accessing the Decommission Site wizard.

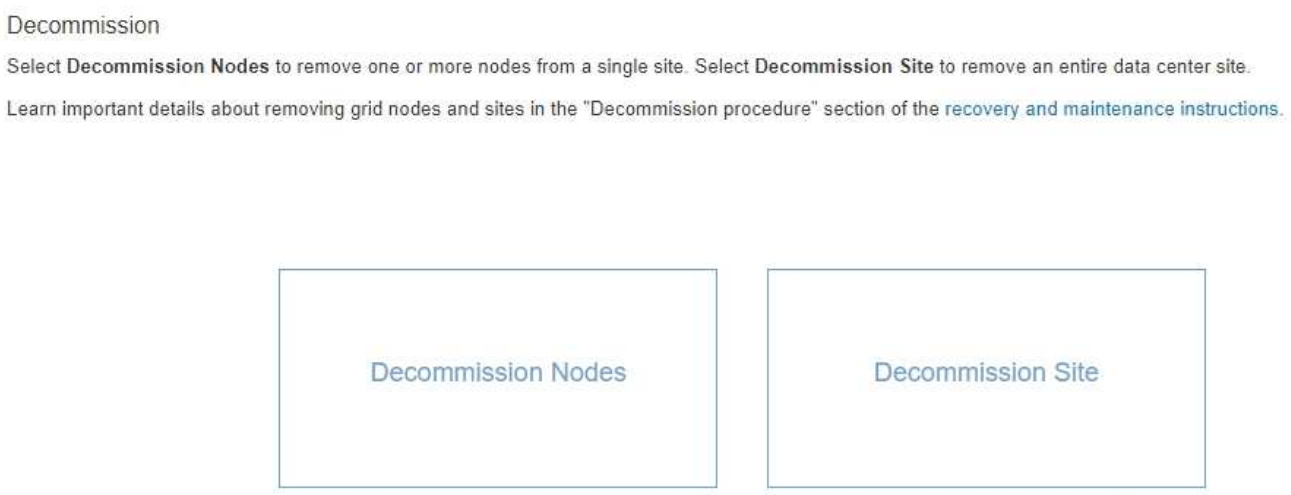
**What you'll need**

- You must have obtained all required materials.
- You must have reviewed the considerations for removing a site.
- You must be signed in to the Grid Manager using a supported browser.
- You must have the Root Access permission, or the Maintenance and ILM permissions.

**Steps**

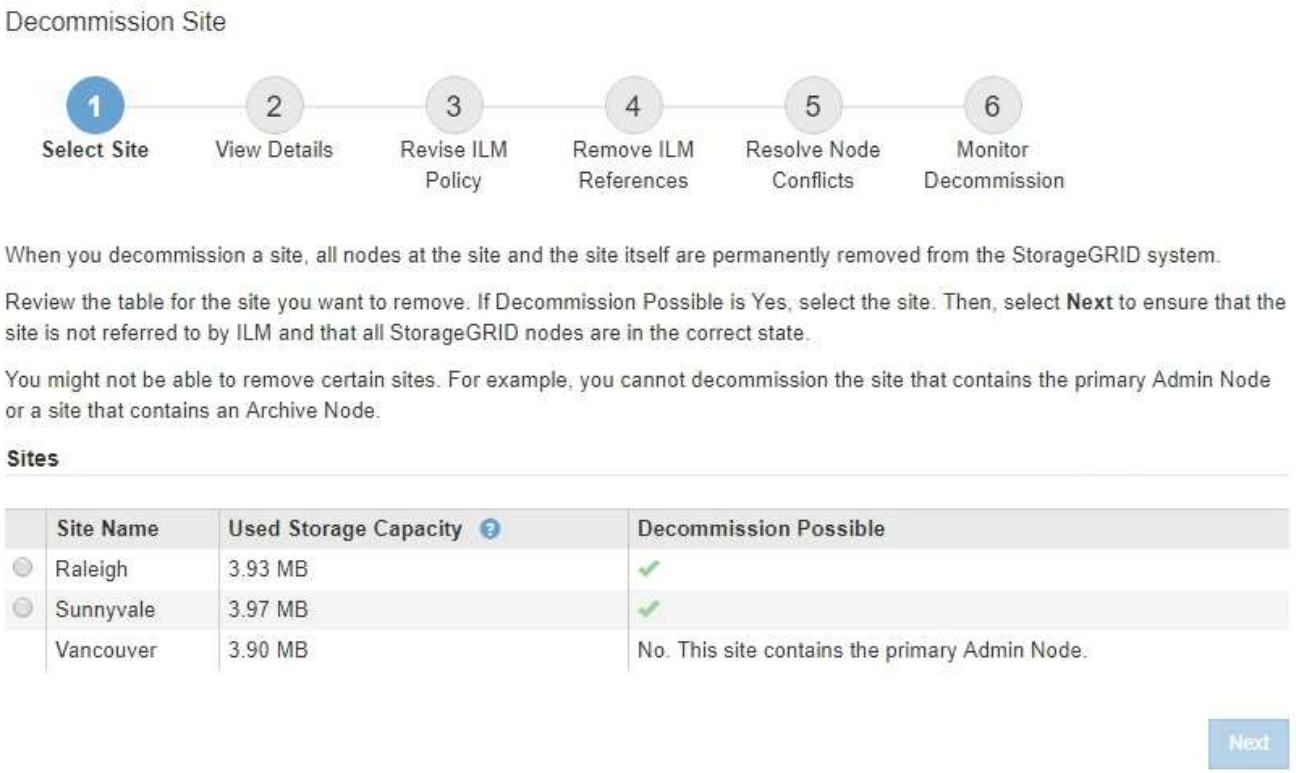
1. Select **Maintenance** > **Maintenance Tasks** > **Decommission**.

   The Decommission page appears.




2. Select the **Decommission Site** button.

   Step 1 (Select Site) of the Decommission Site wizard appears. This step includes an alphabetic list of the sites in your StorageGRID system.




3. View the values in the **Used Storage Capacity** column to determine how much storage is currently being used for object data at each site.

   The Used Storage Capacity is an estimate. If nodes are offline, the Used Storage Capacity is the last known value for the site.

   - For a connected site decommission, this value represents how much object data will need to be moved to other sites or deleted by ILM before you can safely decommission this site.
   - For a disconnected site decommission, this value represents how much of your system's data storage will become inaccessible when you decommission this site.

   > ⓘ If your ILM policy was designed to protect against the loss of a single site, copies of your object data should still exist on the remaining sites.

4. Review the reasons in the **Decommission Possible** column to determine which sites can be decommissioned currently.

   > ⓘ If there is more than one reason a site cannot be decommissioned, the most critical reason is shown.

| Decommission Possible reason | Description | Next step |
|---|---|---|
| Green checkmark (✔) | You can decommission this site. | Go to the next step. |
| No. This site contains the primary Admin Node. | You cannot decommission a site containing the primary Admin Node. | None. You cannot perform this procedure. |

| Decommission Possible reason | Description | Next step |
| --- | --- | --- |
| No. This site contains one or more Archive Nodes. | You cannot decommission a site containing an Archive Node. | None. You cannot perform this procedure. |
| No. All nodes at this site are disconnected. Contact your NetApp account representative. | You cannot perform a connected site decommission unless every node in the site is connected (✔). | If you want to perform a disconnected site decommission, you must contact your NetApp account representative, who will review your requirements and enable the rest of the Decommission Site wizard.<br><br>**IMPORTANT**: Never take online nodes offline so that you can remove a site. You will lose data. |

The example shows a StorageGRID system with three sites. The green checkmark (✔) for the Raleigh and Sunnyvale sites indicates that you can decommission those sites. However, you cannot decommission the Vancouver site because it contains the primary Admin Node.

5. If decommission is possible, select the radio button for the site.

   The **Next** button is enabled.

6. Select **Next**.

   Step 2 (View Details) appears.

#### Step 2: View Details

From Step 2 (View Details) of the Decommission Site wizard, you can review which nodes are included at the site, see how much space has been used on each Storage Node, and assess how much free space is available at the other sites in your grid.

**What you'll need**

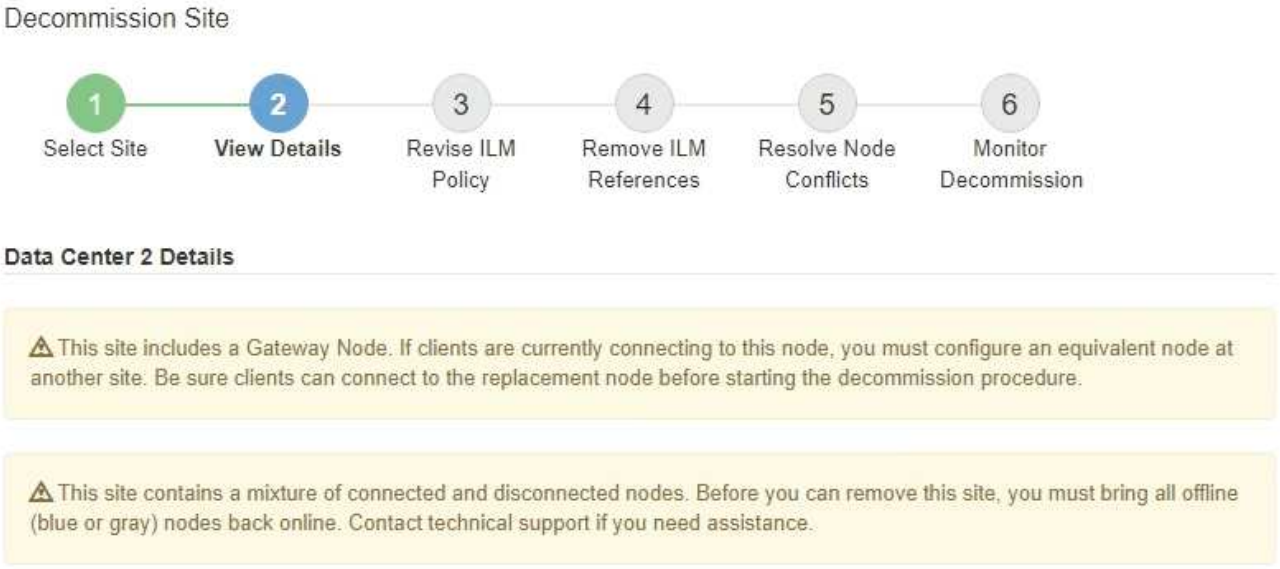
Before decommissioning a site, you must review how much object data exists at the site.

- If you are performing a connected site decommission, you must understand how much object data currently exists at the site before updating ILM. Based on site capacities and your data protection needs, you can create new ILM rules to move data to other sites or to delete object data from the site.
- Perform any required Storage Node expansions before starting the decommission procedure if possible.
- If you are performing a disconnected site decommission, you must understand how much object data will become permanently inaccessible when you remove the site.

> If you are performing a disconnected site decommission, ILM cannot move or delete object data. Any data that remains at the site will be lost. However, if your ILM policy was designed to protect against the loss of a single site, copies of your object data still exist on the remaining sites.

**Steps**

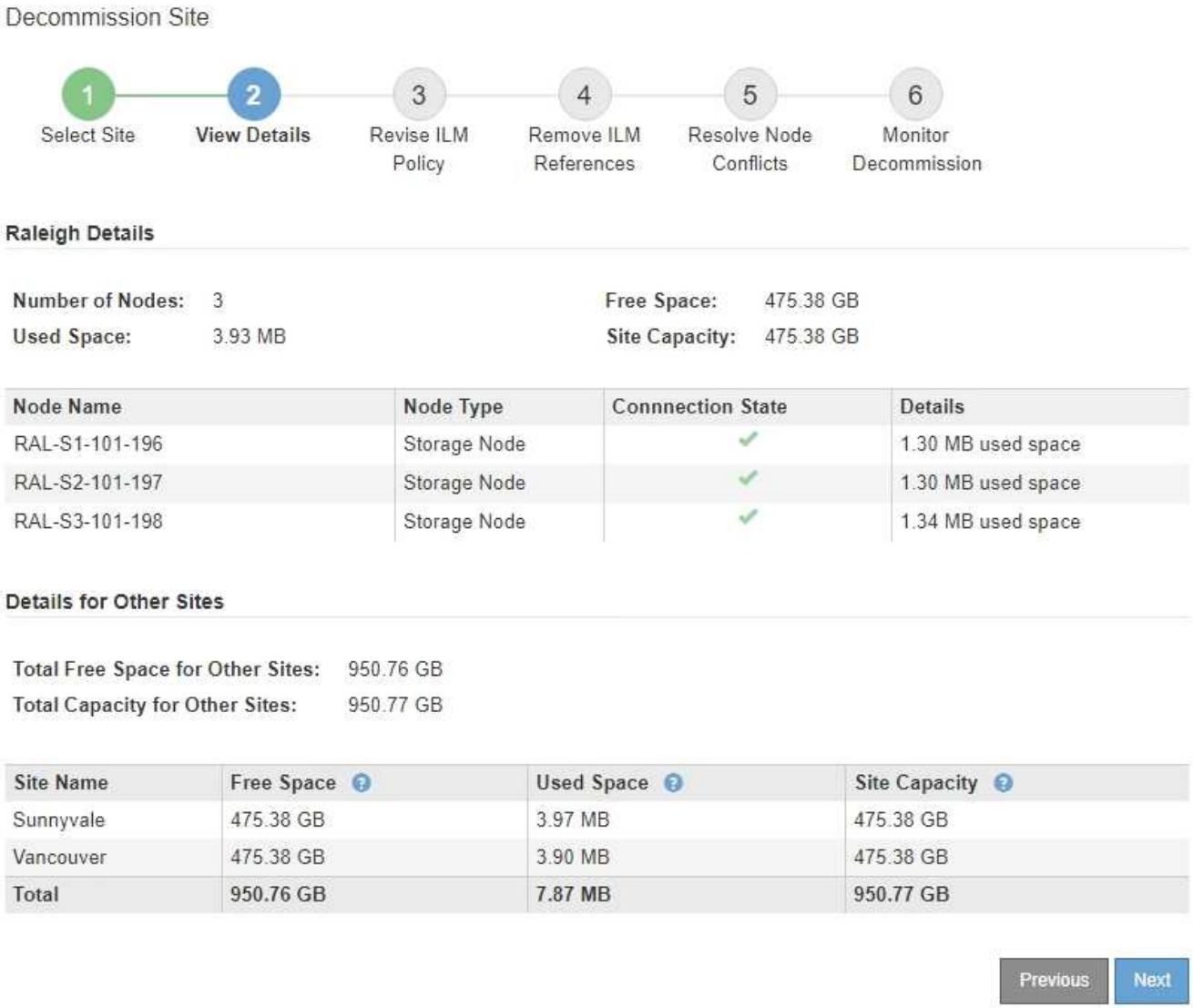
1. From Step 2 (View Details), review any warnings related to the site you selected to remove.




   A warning appears in these cases:

   - The site includes a Gateway Node. If S3 and Swift clients are currently connecting to this node, you must configure an equivalent node at another site. Be sure clients can connect to the replacement node before continuing with the decommission procedure.
   - The site contains a mixture of connected (✔) and disconnected nodes (⬢ or ?). Before you can remove this site, you must bring all offline nodes back online.

2. Review details about the site you selected to remove.



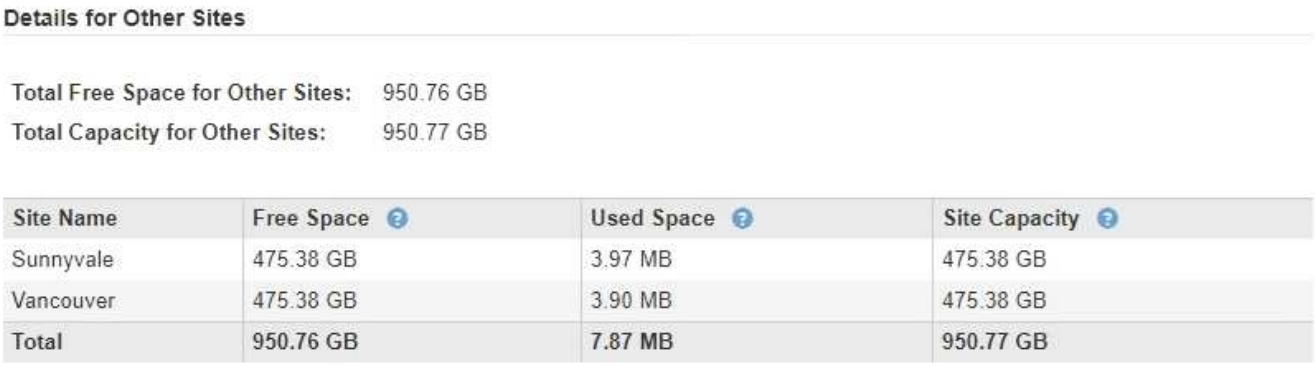

The following information is included for the selected site:

- Number of nodes
- The total used space, free space, and capacity of all Storage Nodes in the site.
  - For a connected site decommission, the **Used Space** value represents how much object data must be moved to other sites or deleted with ILM.
  - For a disconnected site decommission, the **Used Space** value indicates how much object data will become inaccessible when you remove the site.
- Node names, types, and connection states:
  - (Connected)
  - (Administratively Down)
  - (Unknown)
- Details about each node:
  - For each Storage Node, the amount of space that has been used for object data.
  - For Admin Nodes and Gateway Nodes, whether the node is currently used in a high availability (HA) group. You cannot decommission an Admin Node or a Gateway Node that is used in a HA

group. Before you start the decommission, you must edit HA groups to remove all nodes at the site. Or, you can remove the HA group if it only includes nodes from this site.

Administer StorageGRID

3. In the Details for Other Sites section of the page, assess how much space is available at the other sites in your grid.



**Details for Other Sites**

Total Free Space for Other Sites: 950.76 GB
Total Capacity for Other Sites: 950.77 GB

| Site Name | Free Space | Used Space | Site Capacity |
|---|---|---|---|
| Sunnyvale | 475.38 GB | 3.97 MB | 475.38 GB |
| Vancouver | 475.38 GB | 3.90 MB | 475.38 GB |
| Total | 950.76 GB | 7.87 MB | 950.77 GB |

If you are performing a connected site decommission and you plan to use ILM to move object data from the selected site (instead of just deleting it), you must ensure that the other sites have enough capacity to accommodate the moved data and that adequate capacity remains for future growth.

A warning appears if the **Used Space** for the site you want to remove is greater than the **Total Free Space for Other Sites**. To ensure that adequate storage capacity is available after the site is removed, you might need to perform an expansion before performing this procedure.

4. Select **Next**.

Step 3 (Revise ILM Policy) appears.

**Related information**

Manage objects with ILM

#### Step 3: Revise ILM Policy

From Step 3 (Revise ILM Policy) of the Decommission Site wizard, you can determine if the site is referred to by the active ILM policy.

**What you'll need**

You have a good understanding of how ILM works and you are familiar with creating storage pools, Erasure Coding profiles, ILM rules, and simulating and activating an ILM policy.

Manage objects with ILM

**About this task**

StorageGRID cannot decommission a site if that site is referred to by any ILM rule in the active ILM policy.

If your current ILM policy refers to the site you want to remove, you must activate a new ILM policy that meets certain requirements. Specifically, the new ILM policy:

- Cannot use a storage pool that refers to the site.
- Cannot use an Erasure Coding profile that refers to the site.
- Cannot use the default **All Storage Nodes** storage pool or the default **All Sites** site.
- Cannot use the stock **Make 2 Copies** rule.
- Must be designed to fully protect all object data.

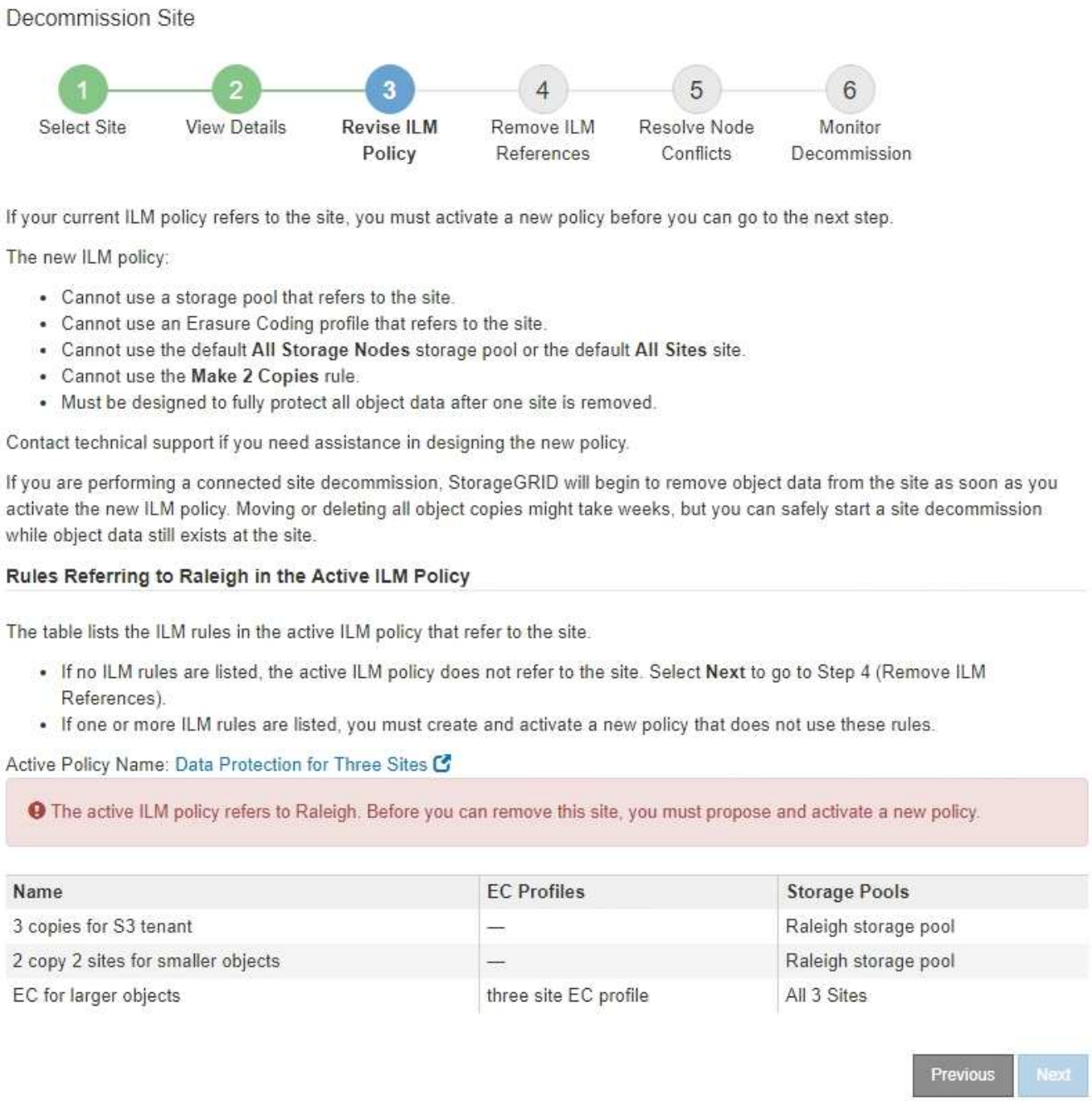
> Never create a single-copy ILM rule to accommodate the removal of a site. An ILM rule that creates only one replicated copy for any time period puts data at risk of permanent loss. If only one replicated copy of an object exists, that object is lost if a Storage Node fails or has a significant error. You also temporarily lose access to the object during maintenance procedures such as upgrades.

If you are performing a *connected site decommission*, you must consider how StorageGRID should manage the object data currently at the site you want to remove. Depending on your data protection requirements, the new rules can move existing object data to different sites or they can delete any extra object copies that are no longer needed.

Contact technical support if you need assistance designing the new policy.

**Steps**

1. From Step 3 (Revise ILM Policy), determine if any ILM rules in the active ILM policy refer to the site you selected to remove.




2. If no rules are listed, select **Next** to go to Step 4 (Remove ILM References)

   Step 4: Remove ILM References

3. If one or more ILM rules are listed in the table, select the link next to **Active Policy Name**.

   The ILM Policies page appears in a new browser tab. Use this tab to update ILM. The Decommission Site page will remain open on the other tab.

   a. If necessary, select **ILM** > **Storage Pools** to create one or more storage pools that do not refer to the site.

      For details, see the instructions for managing objects with information lifecycle management.

b. If you plan to use erasure coding, select **ILM** > **Erasure Coding** to create one or more Erasure Coding profiles.

   You must select storage pools that do not refer to the site.

   > Do not use the **All Storage Nodes** storage pool in the Erasure Coding profiles.

4. Select **ILM** > **Rules** and clone each of the rules listed in the table for Step 3 (Revise ILM Policy).

   > For details, see the instructions for managing objects with information lifecycle management.

   a. Use names that will make it easy to select these rules in a new policy.
   b. Update the placement instructions.

      Remove any storage pools or Erasure Coding profiles that refer to the site and replace them with new storage pools or Erasure Coding profiles.

      > Do not use the **All Storage Nodes** storage pool in the new rules.

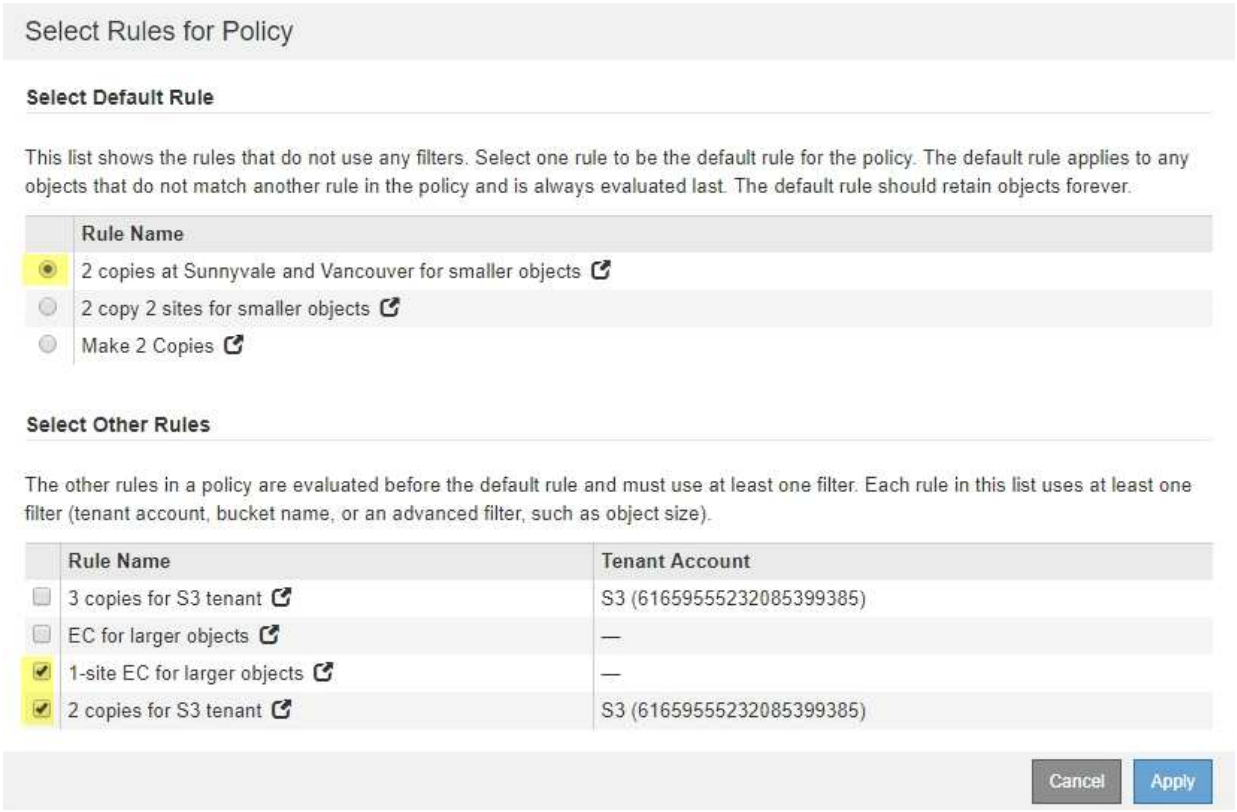
5. Select **ILM** > **Policies** and create a new policy that uses the new rules.

   > For details, see the instructions for managing objects with information lifecycle management.

   a. Select the active policy, and select **Clone**.
   b. Specify a policy name and a reason for change.
   c. Select rules for the cloned policy.
      - Unselect all rules listed for Step 3 (Revise ILM Policy) of the Decommission Site page.
      - Select a default rule that does not refer to the site.

        > Do not select the **Make 2 Copies** rule because that rule uses the **All Storage Nodes** storage pool, which is not allowed.

      - Select the other replacement rules you created. These rules should not refer to the site.




d. Select **Apply**.

e. Drag and drop the rows to reorder the rules in the policy.

You cannot move the default rule.

> You must confirm that the ILM rules are in the correct order. When the policy is activated, new and existing objects are evaluated by the rules in the order listed, starting at the top.

f. Save the proposed policy.

6. Ingest test objects, and simulate the proposed policy to ensure that the correct rules are applied.

> Errors in an ILM policy can cause unrecoverable data loss. Carefully review and simulate the policy before activating it to confirm that it will work as intended.

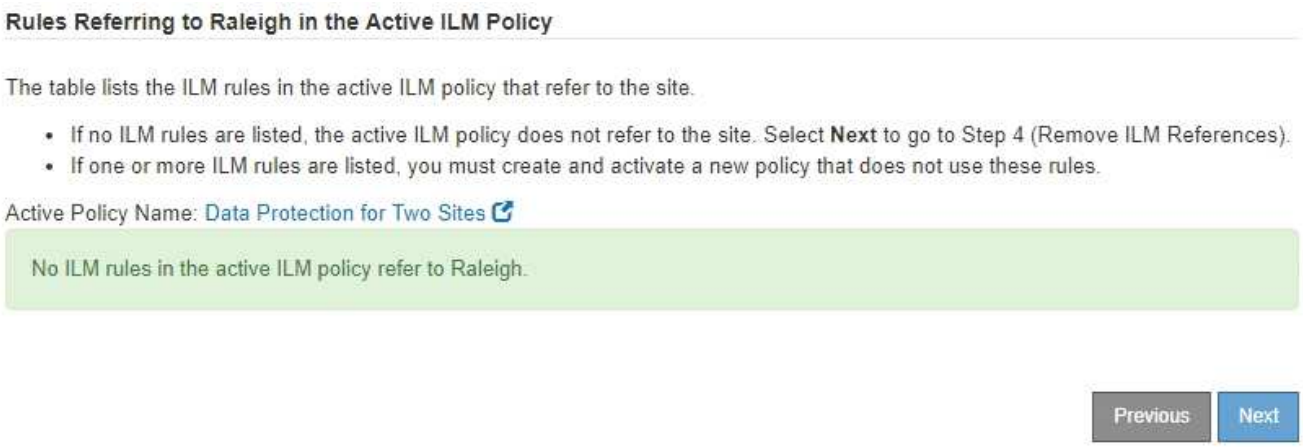
> When you activate a new ILM policy, StorageGRID uses it to manage all objects, including existing objects and newly ingested objects. Before activating a new ILM policy, review any changes to the placement of existing replicated and erasure-coded objects. Changing an existing object's location might result in temporary resource issues when the new placements are evaluated and implemented.

7. Activate the new policy.

If you are performing a connected site decommission, StorageGRID begins to remove object data from the selected site as soon as you activate the new ILM policy. Moving or deleting all object copies might take weeks. Although you can safely start a site decommission while object data still exists at the site, the decommission procedure will complete more quickly and with fewer disruptions and performance impacts if you allow data to be moved from the site before starting the actual decommission procedure (by selecting

**Start Decommission** in Step 5 of the wizard).

8. Return to **Step 3 (Revise ILM Policy)** to ensure that no ILM rules in the new active policy refer to the site and the **Next** button is enabled.




> If any rules are listed, you must create and activate a new ILM policy before you can continue.

9. If no rules are listed, select **Next**.

   Step 4 (Remove ILM References) appears.

#### Step 4: Remove ILM References

From Step 4 (Remove ILM References) of the Decommission Site wizard, you can remove the proposed policy if one exists and delete or edit any unused ILM rules that still refer to the site.

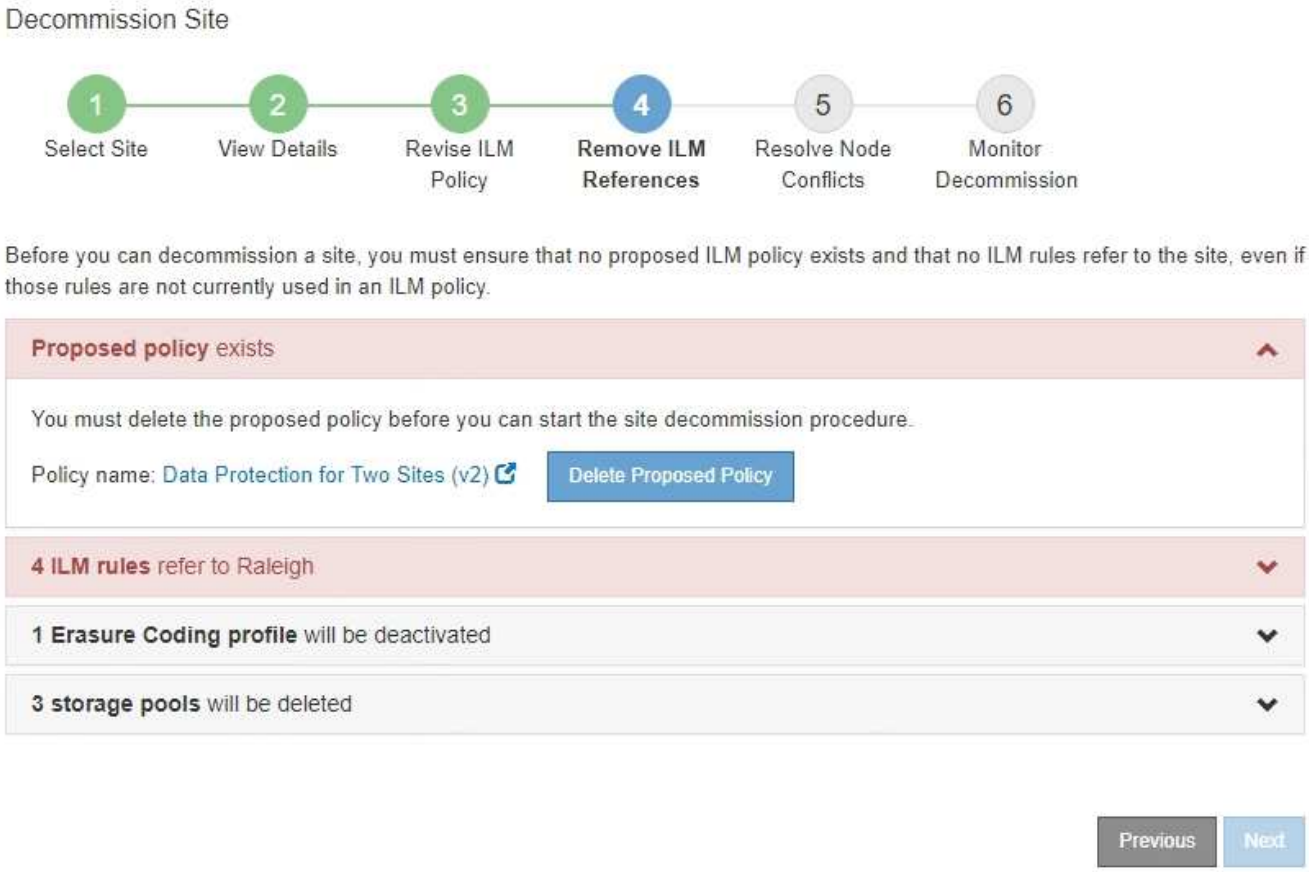
**About this task**

You are prevented from starting the site decommission procedure in these cases:

- A proposed ILM policy exists. If you have a proposed policy, you must delete it.
- Any ILM rule refers to the site, even if that rule is not used in any ILM policy. You must delete or edit all rules that refer to the site.

**Steps**

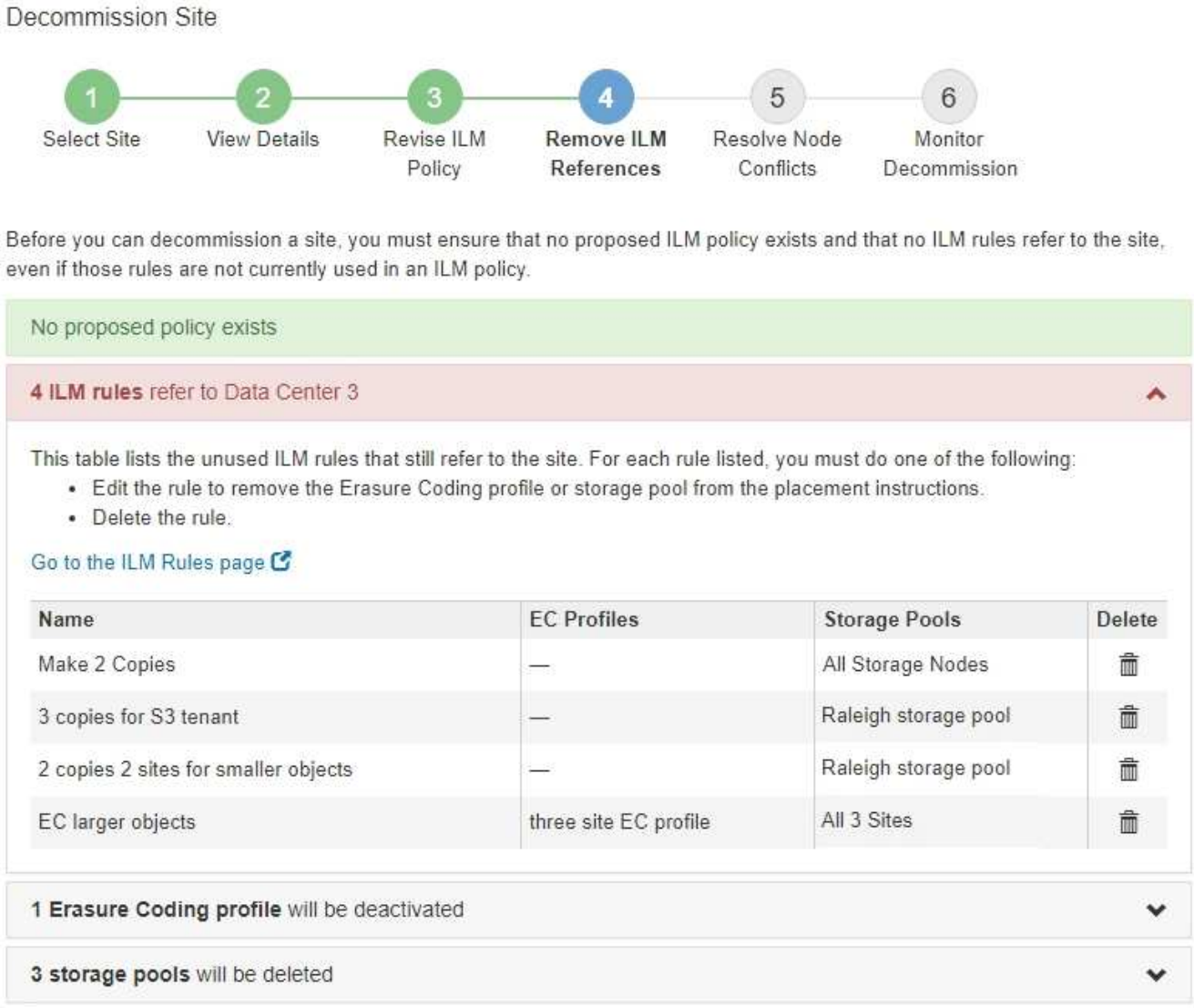
1. If a proposed policy is listed, remove it.




a. Select **Delete Proposed Policy**.

b. Select **OK** in the confirmation dialog box.

2. Determine whether any unused ILM rules refer to the site.




Any ILM rules that are listed still refer to the site but are not used in any policy. In the example:

- The stock **Make 2 Copies** rule uses the system-default **All Storage Nodes** storage pool, which uses the All Sites site.
- The unused **3 copies for S3 tenant** rule refers to the **Raleigh** storage pool.
- The unused **2 copy 2 sites for smaller objects** rule refers to the **Raleigh** storage pool.
- The unused **EC larger objects** rules uses the Raleigh site in the **All 3 Sites** Erasure Coding profile.
- If no ILM rules are listed, select **Next** to go to **Step 5 (Resolve Node Conflicts)**.

  Step 5: Resolve Node Conflicts (and start decommission)

  When StorageGRID decommissions the site, it will automatically deactivate any unused Erasure Coding profiles that refer to the site, and it will automatically delete any unused storage pools that refer to the site. The system-default All Storage Nodes storage pool is removed because it uses the All Sites site.

- If one or more ILM rules are listed, go to the next step.

3. Edit or delete each unused rule:

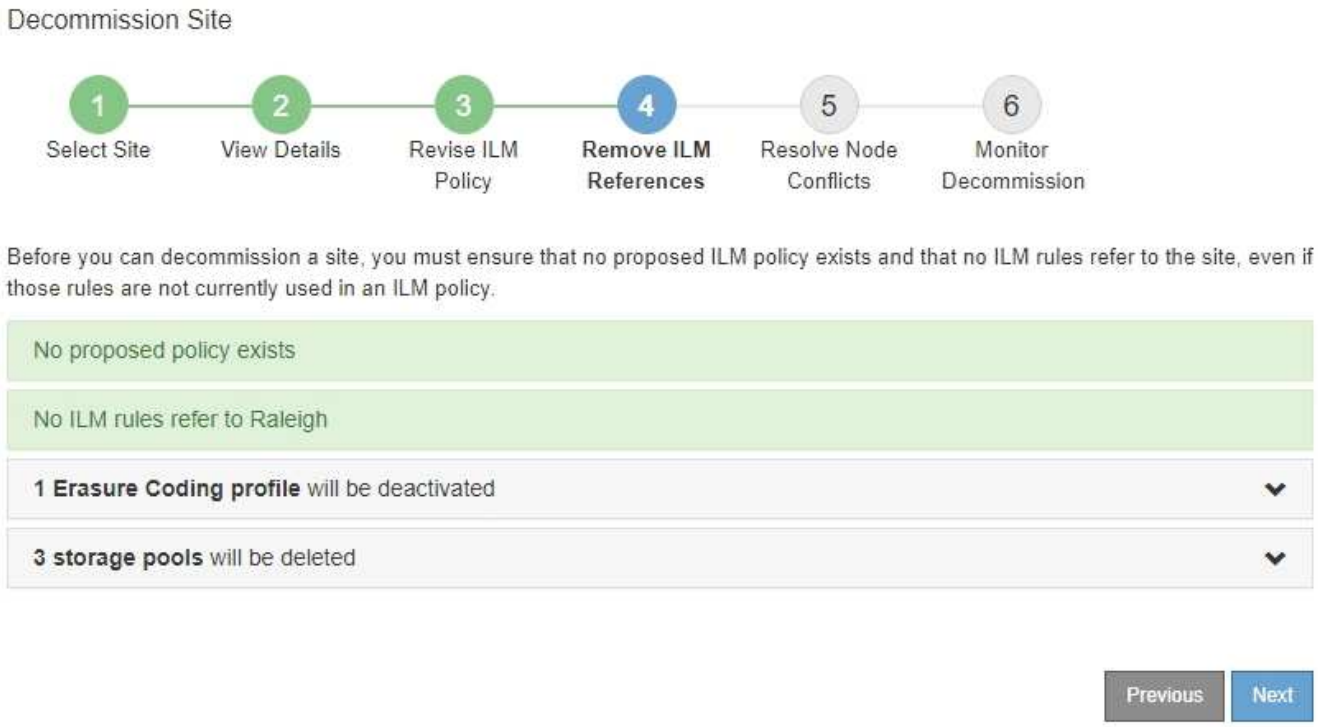
- To edit a rule, go the ILM Rules page and update all placements that use an Erasure Coding profile or storage pool that refers to the site. Then, return to **Step 4 (Remove ILM References)**.

  For details, see the instructions for managing objects with information lifecycle management.

- To delete a rule, select the trash can icon and select **OK**.

  You must delete the stock **Make 2 Copies** rule before you can decommission a site.

4. Confirm that no proposed ILM policy exists, no unused ILM rules refer to the site, and the **Next** button is enabled.




5. Select **Next**.

   Any remaining storage pools and Erasure Coding profiles that refer to the site will become invalid when the site is removed. When StorageGRID decommissions the site, it will automatically deactivate any unused Erasure Coding profiles that refer to the site, and it will automatically delete any unused storage pools that refer to the site. The system-default All Storage Nodes storage pool is removed because it uses the All Sites site.

   Step 5 (Resolve Node Conflicts) appears.

#### Step 5: Resolve Node Conflicts (and start decommission)

From Step 5 (Resolve Node Conflicts) of the Decommission Site wizard, you can determine if any nodes in your StorageGRID system are disconnected or if any nodes at the selected site belong to a high availability (HA) group. After any node conflicts are resolved, you start the decommission procedure from this page.

You must ensure that all nodes in your StorageGRID system are in the correct state, as follows:

- All nodes in your StorageGRID system must be connected (✔).

> If you are performing a disconnected site decommission, all nodes at the site you are removing must be disconnected, and all nodes at all other sites must be connected.

- No node at the site you are removing can have an interface that belongs to a high availability (HA) group.

If any node is listed for Step 5 (Resolve Node Conflicts), you must correct the issue before you can start the decommission.

Before starting the site decommission procedure from this page, review the following considerations:

- You must allow adequate time for the decommission procedure to complete.

> Moving or deleting object data from a site might take days, weeks, or even months, depending on the amount of data at the site, the load on your system, network latencies, and the nature of the required ILM changes.

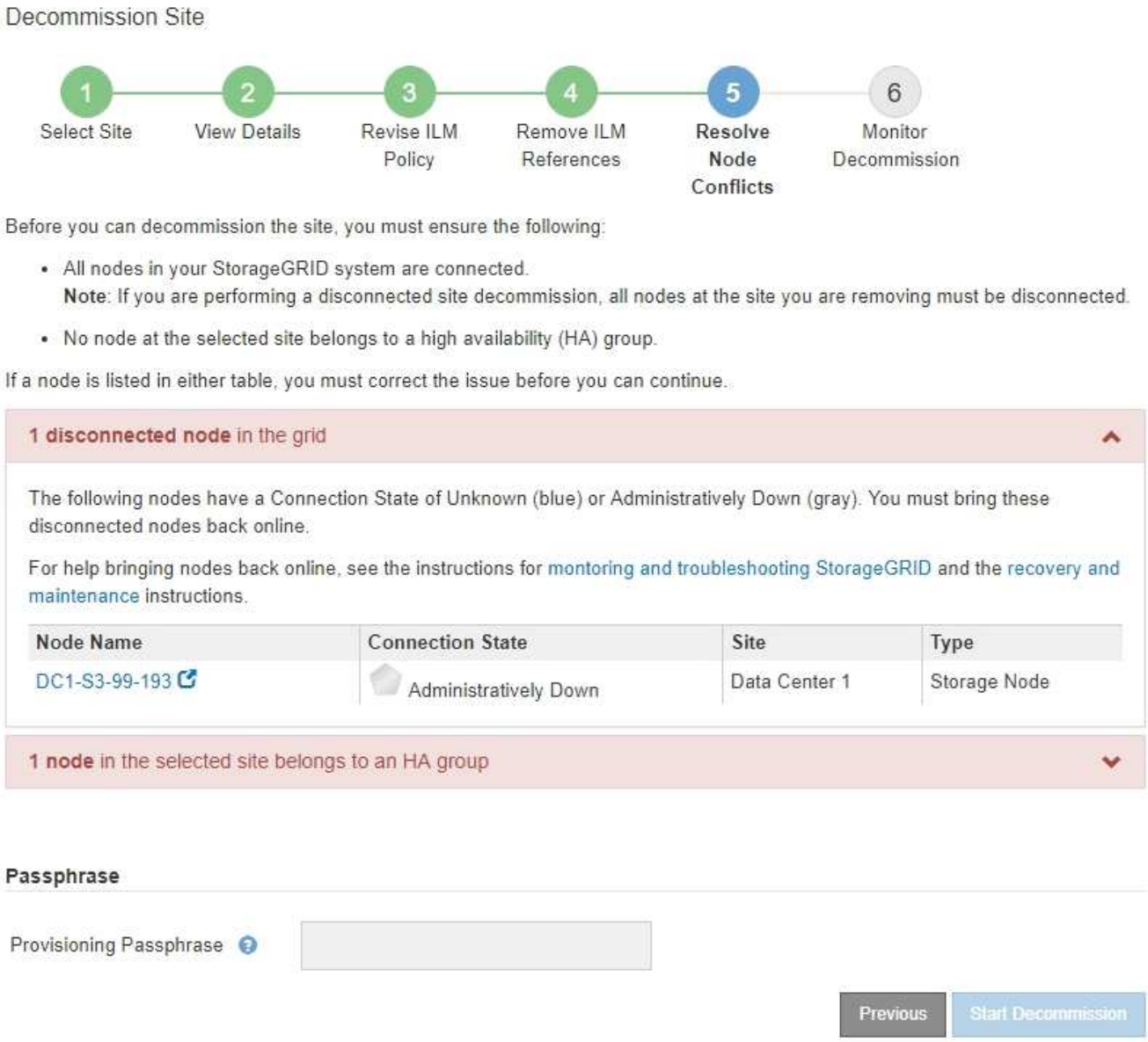
- While the site decommission procedure is running:
  - You cannot create ILM rules that refer to the site being decommissioned. You also cannot edit an existing ILM rule to refer to the site.
  - You cannot perform other maintenance procedures, such as expansion or upgrade.

> If you need to perform another maintenance procedure during a connected site decommission, you can pause the procedure while the Storage Nodes are being removed. The **Pause** button is enabled during the “Decommissioning Replicated and Erasure Coded Data” stage.

  - If you need to recover any node after starting the site decommission procedure, you must contact support.

**Steps**

1. Review the disconnected nodes section of Step 5 (Resolve Node Conflicts) to determine if any nodes in your StorageGRID system have a Connection State of Unknown (?) or Administratively Down ().

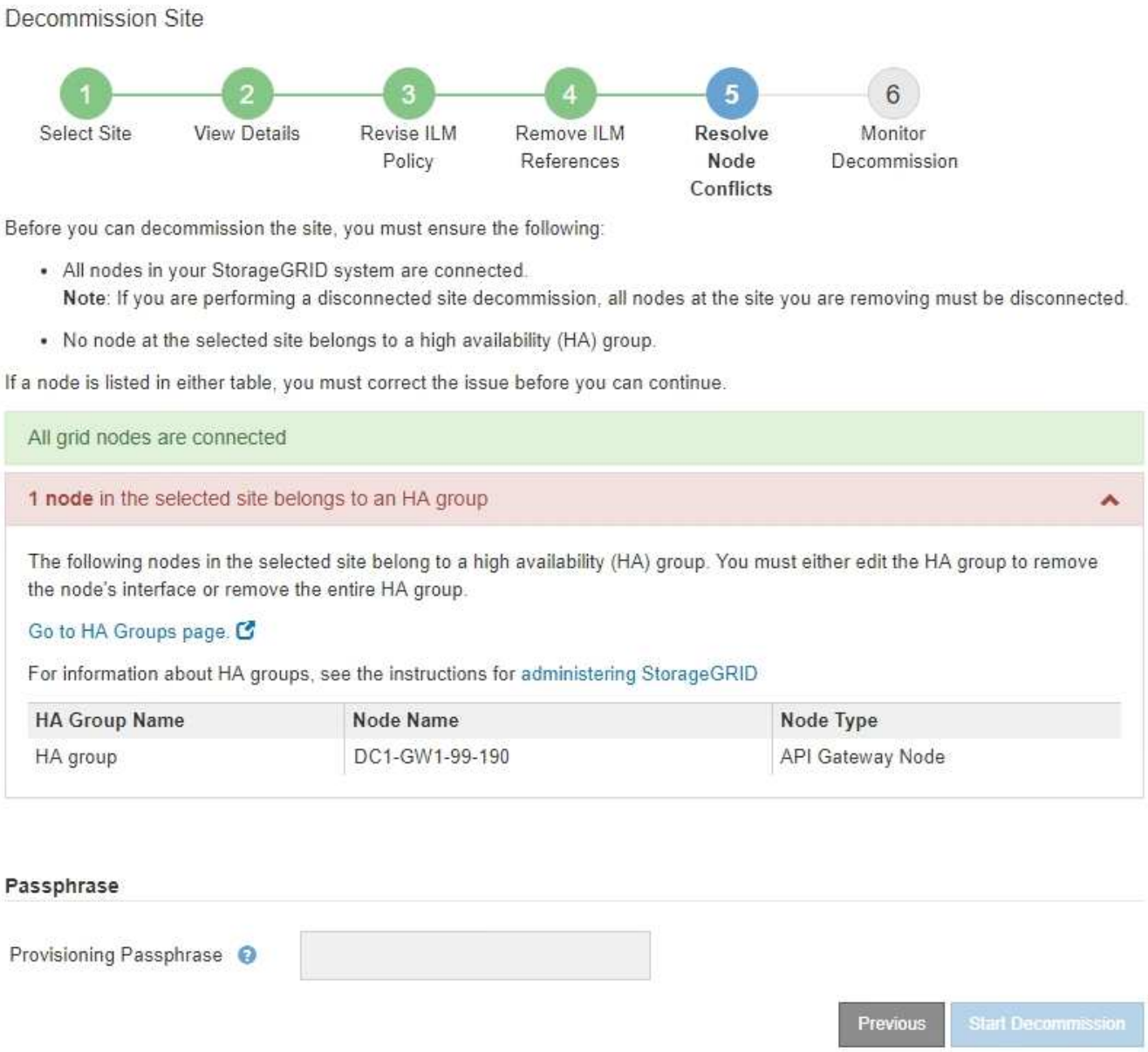



2. If any nodes are disconnected, bring them back online.

   See the instructions for monitoring and troubleshooting StorageGRID and the grid node procedures. Contact technical support if you need assistance.

3. When all disconnected nodes have been brought back online, review the HA groups section of Step 5 (Resolve Node Conflicts).

   This table lists any nodes at the selected site that belong to a high availability (HA) group.

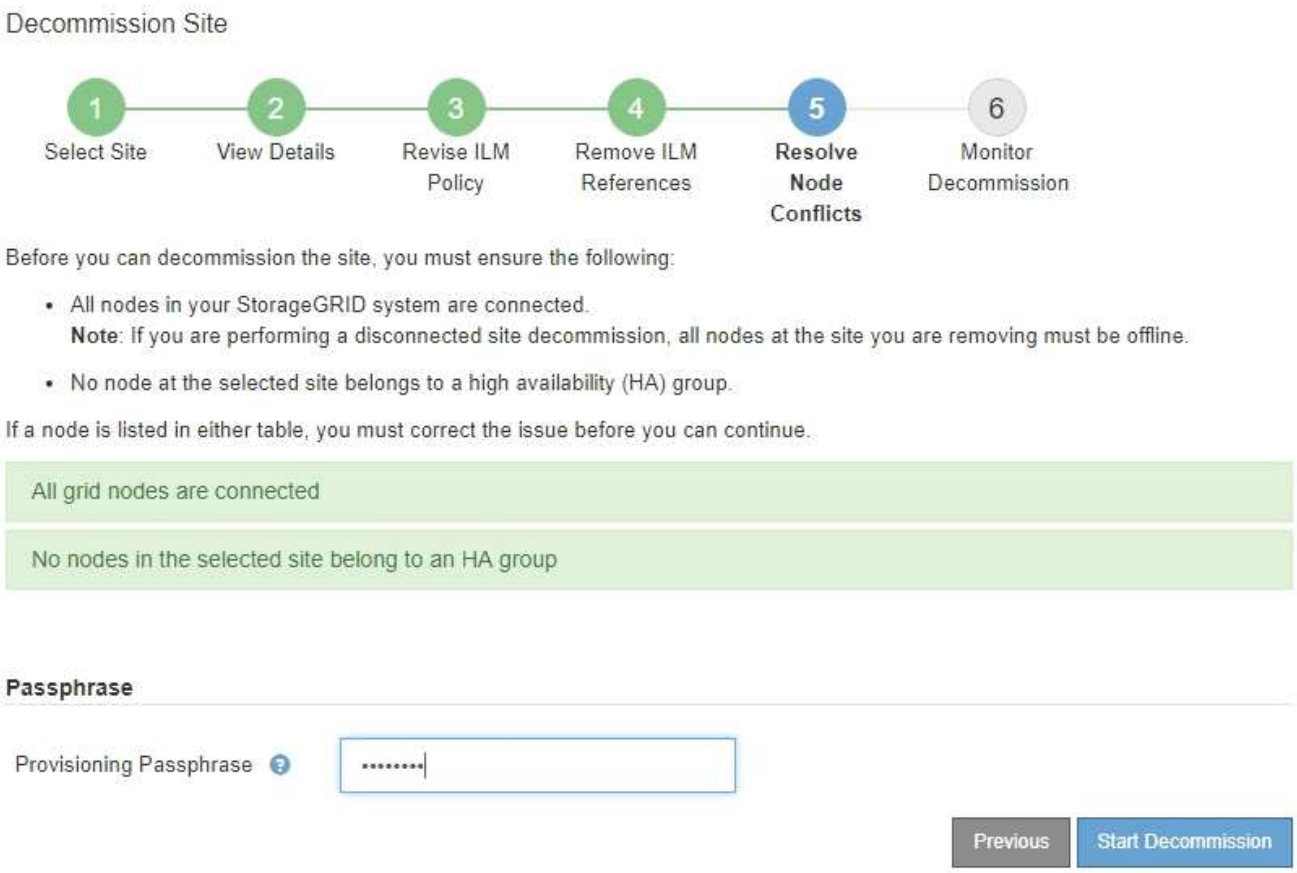



4. If any nodes are listed, do either of the following:
   - Edit each affected HA group to remove the node interface.
   - Remove an HA group that only includes nodes from this site. See the instructions for administering StorageGRID.

   If all nodes are connected and no nodes in the selected site are used in an HA group, the **Provisioning Passphrase** field is enabled.

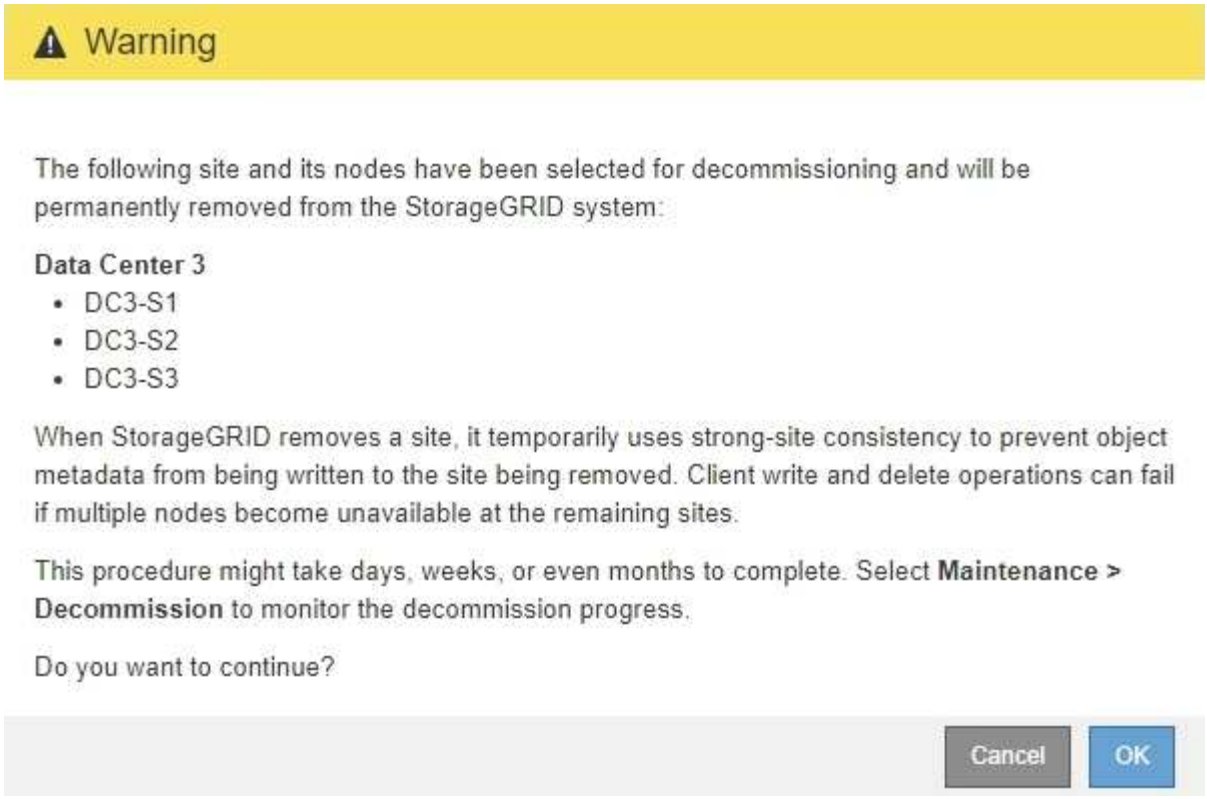
5. Enter the provisioning passphrase.

   The **Start Decommission** button becomes enabled.



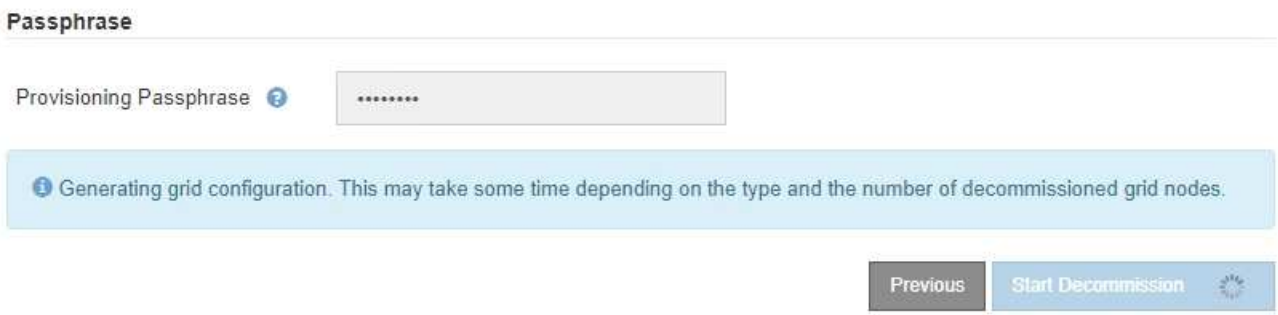

6. If you are ready to start the site decommission procedure, select **Start Decommission**.

   A warning lists the site and nodes that will be removed. You are reminded that it might take days, weeks, or even months to completely remove the site.




7. Review the warning. If you are ready to begin, select **OK**.

   A message appears as the new grid configuration is generated. This process might take some time, depending on the type and number of decommissioned grid nodes.




   When the new grid configuration has been generated, Step 6 (Monitor Decommission) appears.

   The **Previous** button remains disabled until the decommission is complete.

**Related information**

Monitor & troubleshoot

Grid node procedures

Administer StorageGRID

#### Step 6: Monitor Decommission

From Step 6 (Monitor Decommission) of the Decommission Site page wizard, you can monitor the progress as the site is removed.

**About this task**

When StorageGRID removes a connected site, it removes nodes in this order:

1. Gateway Nodes
2. Admin Nodes
3. Storage Nodes

When StorageGRID removes a disconnected site, it removes nodes in this order:

1. Gateway Nodes
2. Storage Nodes
3. Admin Nodes

Each Gateway Node or Admin Node might only require a few minutes or an hour to remove; however, Storage Nodes might take days or weeks.

**Steps**

1. As soon as a new Recovery Package has been generated, download the file.

> Download the Recovery Package as soon as possible to ensure you can recover your grid if something goes wrong during the decommission procedure.

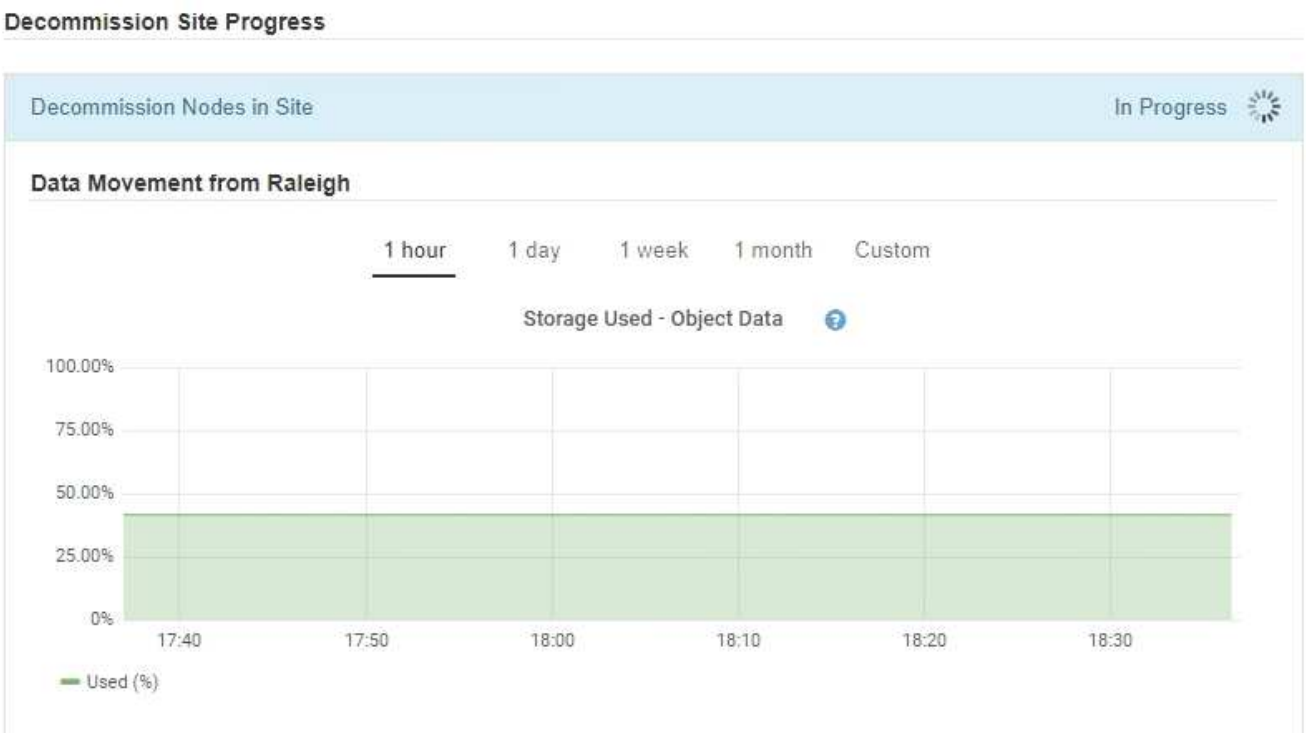
   a. Select the link in the message, or select **Maintenance** > **System** > **Recovery Package**.
   b. Download the `.zip` file.

      See the instructions for downloading the Recovery Package.

> The Recovery Package file must be secured because it contains encryption keys and passwords that can be used to obtain data from the StorageGRID system.

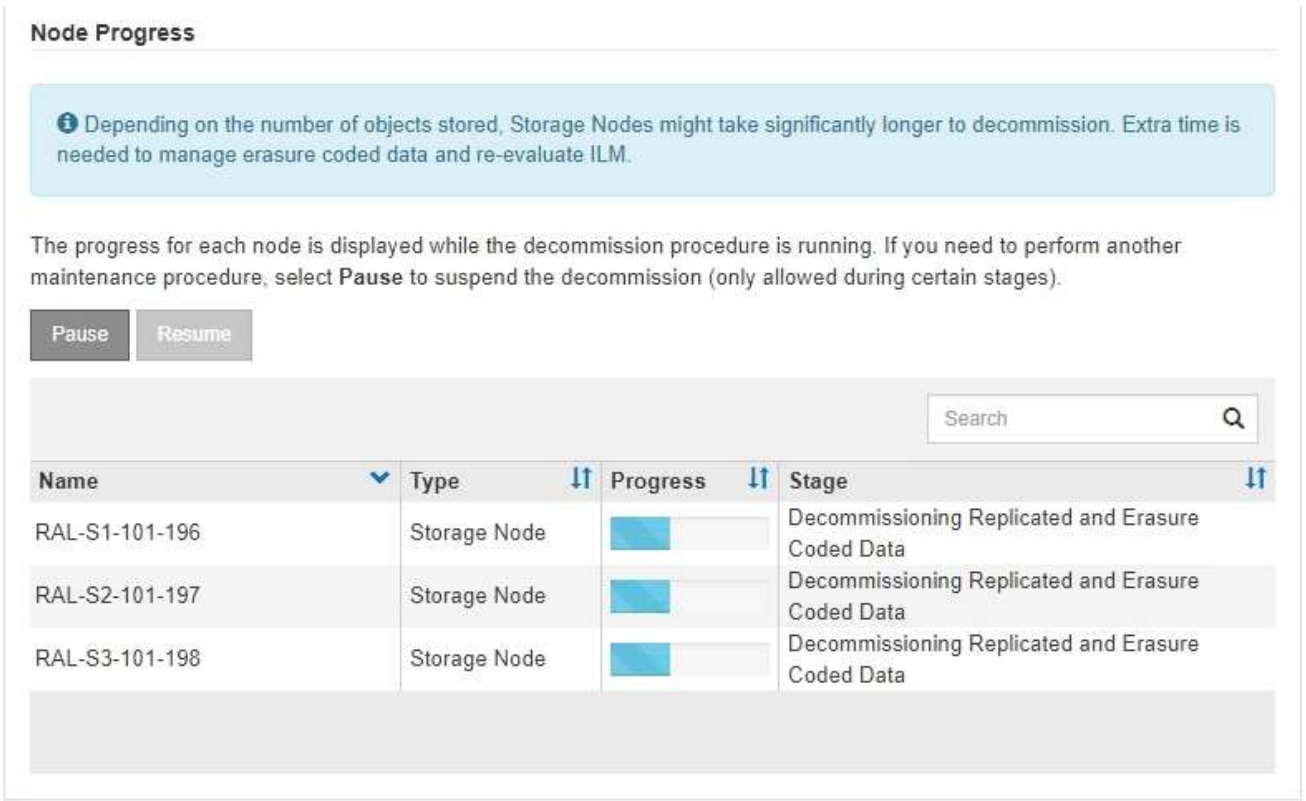
2. Using the Data Movement chart, monitor the movement of object data from this site to other sites.

   Data movement started when you activated the new ILM policy in Step 3 (Revise ILM Policy). Data movement will occur throughout the decommission procedure.




3. In the Node Progress section of the page, monitor the progress of the decommission procedure as nodes are removed.

   When a Storage Node is removed, each node goes through a series of stages. Although most of these stages occur quickly or even imperceptibly, you might need to wait days or even weeks for other stages to complete, based on how much data needs to moved. Additional time is required to manage erasure-coded data and re-evaluate ILM.



If you are monitoring the progress of a connected site decommission, refer to this table to understand the decommission stages for a Storage Node:

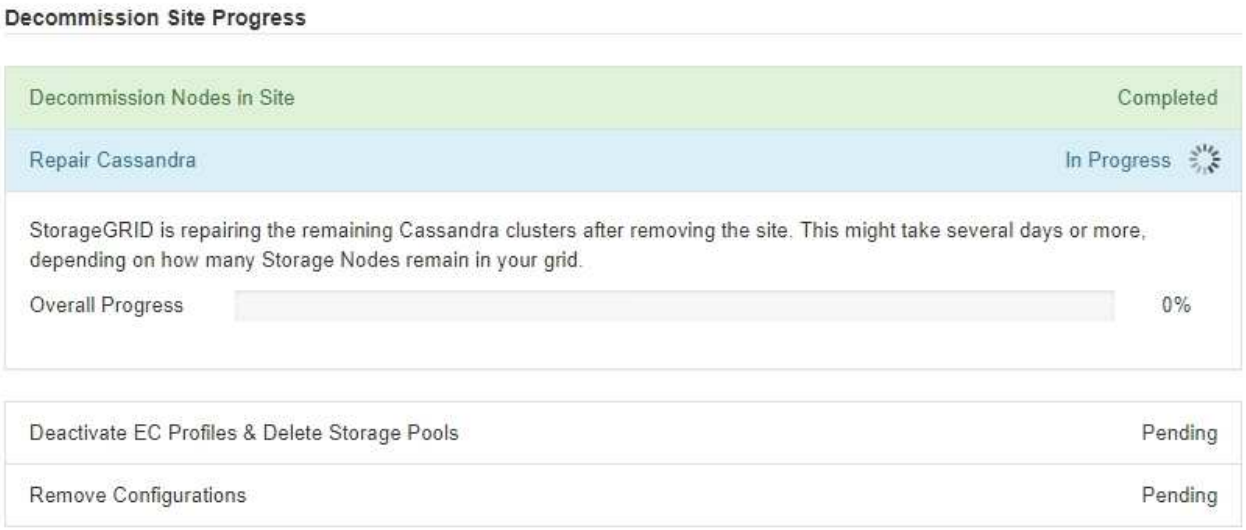
| Stage | Estimated duration |
|---|---|
| Pending | Minute or less |
| Wait for Locks | Minutes |
| Prepare Task | Minute or less |
| Marking LDR Decommissioned | Minutes |
| Decommissioning Replicated and Erasure Coded Data | Hours, days, or weeks based on the amount of data<br><br>**Note**: If you need to perform other maintenance activities, you can pause the site decommission during this stage. |
| LDR Set State | Minutes |
| Flush Audit Queues | Minutes to hours, based on the number of messages and network latency. |
| Complete | Minutes |

If you are monitoring the progress of a disconnected site decommission, refer to this table to understand the decommission stages for a Storage Node:

| Stage | Estimated duration |
|---|---|
| Pending | Minute or less |
| Wait for Locks | Minutes |
| Prepare Task | Minute or less |
| Disable External Services | Minutes |
| Certificate Revocation | Minutes |
| Node Unregister | Minutes |
| Storage Grade Unregister | Minutes |
| Storage Group Removal | Minutes |
| Entity Removal | Minutes |

| Stage | Estimated duration |
| --- | --- |
| Complete | Minutes |

4. After all nodes have reached the Complete stage, wait for the remaining site decommission operations to complete.
   - During the **Repair Cassandra** step, StorageGRID makes any necessary repairs to the Cassandra clusters that remain in your grid. These repairs might take several days or more, depending on how many Storage Nodes remain in your grid.

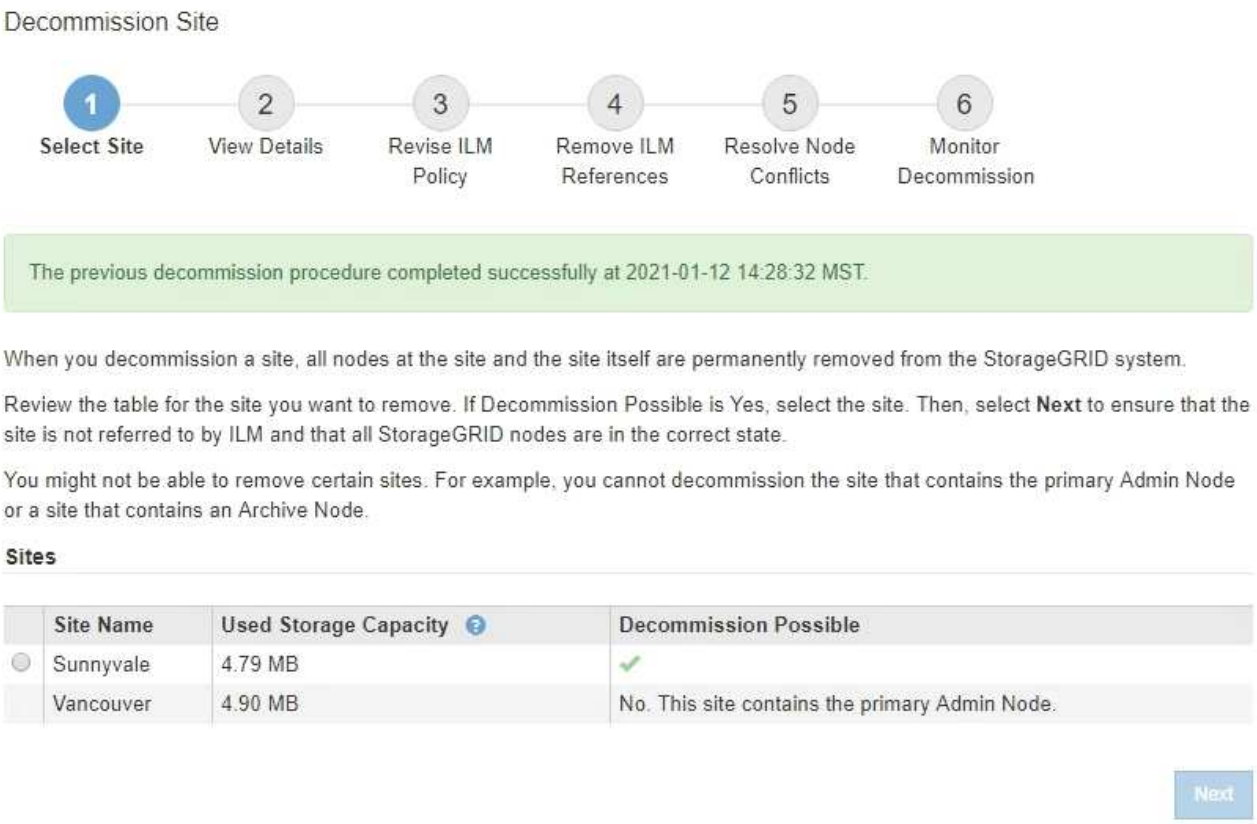



   - During the **Deactivate EC Profiles & Delete Storage Pools** step, the following ILM changes are made:
     - Any Erasure Coding profiles that referred to the site are deactivated.
     - Any Storage Pools that referred to the site are deleted.

> The system-default All Storage Nodes storage pool is also removed because it uses the All Sites site.

   - Finally, during the **Remove Configuration** step, any remaining references to the site and its nodes are removed from the rest of the grid.

5. When the decommission procedure has completed, the Decommission Site page shows a success message, and the removed site is no longer shown.




**After you finish**

Complete these tasks after you complete the site decommission procedure:

- Ensure that the drives of all Storage Nodes in the decommissioned site are wiped clean. Use a commercially available data wiping tool or service to permanently and securely remove data from the drives.
- If the site included one or more Admin Nodes and single sign-on (SSO) is enabled for your StorageGRID system, remove all relying party trusts for the site from Active Directory Federation Services (AD FS).
- After the nodes have been gracefully powered off automatically as part of the connected site decommission procedure, remove the associated virtual machines.

**Related information**

Downloading the Recovery Package

## Network maintenance procedures

You can configure the list of subnets on the Grid Network or update IP addresses, DNS servers, or NTP servers for your StorageGRID system.

**Choices**

- Updating subnets for the Grid Network
- Configuring IP addresses

- Configuring DNS servers
- Configuring NTP servers
- Restoring network connectivity for isolated nodes

### Updating subnets for the Grid Network

StorageGRID maintains a list of the network subnets used to communicate between grid nodes on the Grid Network (eth0). These entries include the subnets used for the Grid Network by each site in your StorageGRID system as well as any subnets used for NTP, DNS, LDAP, or other external servers accessed through the Grid Network gateway. When you add grid nodes or a new site in an expansion, you might need to update or add subnets to the Grid Network.

#### What you'll need

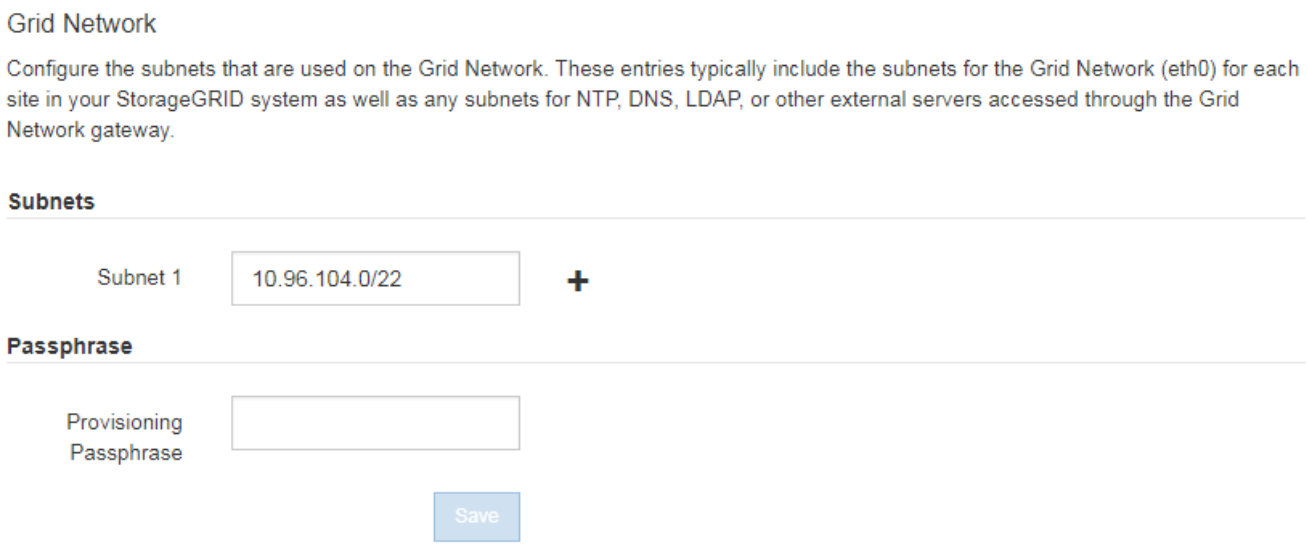
- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have the network addresses, in CIDR notation, of the subnets you want to configure.

#### About this task

If you are performing an expansion activity that includes adding a new subnet, you must add the new Grid subnet before you start the expansion procedure.

#### Steps

1. Select **Maintenance** > **Network** > **Grid Network**.




2. In the Subnets list, click the plus sign to add a new subnet in CIDR notation.

   For example, enter `10.96.104.0/22`.

3. Enter the provisioning passphrase, and click **Save**.

   The subnets you have specified are configured automatically for your StorageGRID system.

### Configuring IP addresses

## You can perform network configuration by configuring IP addresses for grid nodes using the Change IP tool.

You must use the Change IP tool to make most changes to the networking configuration that was initially set during grid deployment. Manual changes using standard Linux networking commands and files might not propagate to all StorageGRID services, and might not persist across upgrades, reboots, or node recovery procedures.

If you want to change the Grid Network IP address for all nodes in the grid, use the special procedure for grid-wide changes.

Changing IP addresses for all nodes in the grid

If you are making changes to the Grid Network Subnet List only, use the Grid Manager to add or change the network configuration. Otherwise, use the Change IP tool if the Grid Manager is inaccessible due to a network configuration issue, or you are performing both a Grid Network routing change and other network changes at the same time.

The IP change procedure can be a disruptive procedure. Parts of the grid might be unavailable until the new configuration is applied.

**Ethernet interfaces**

The IP address assigned to eth0 is always the grid node's Grid Network IP address. The IP address assigned to eth1 is always the grid node's Admin Network IP address. The IP address assigned to eth2 is always the grid node's Client Network IP address.

Note that on some platforms, such as StorageGRID appliances, eth0, eth1, and eth2 might be aggregate interfaces composed of subordinate bridges or bonds of physical or VLAN interfaces. On these platforms, the **SSM** > **Resources** tab might show the Grid, Admin, and Client network IP address assigned to other interfaces in addition to eth0, eth1, or eth2.

**DHCP**

You can only set up DHCP during the deployment phase. You cannot set up DHCP during configuration. You must use the IP address change procedures if you want to change IP addresses, subnet masks, and default gateways for a grid node. Using the Change IP tool will cause DHCP addresses to become static.

**High availability (HA) groups**

- You cannot change the Client network IP address outside the subnet of an HA group configured on the Client network interface.
- You cannot change the Client network IP address to the value of an existing virtual IP address assigned by an HA group configured on the Client network interface.
- You cannot change the Grid network IP address outside the subnet of an HA group configured on the Grid network interface.
- You cannot change the Grid network IP address to the value of an existing virtual IP address assigned by an HA group configured on the Grid network interface.

**Choices**

- Changing a node's network configuration
- Adding to or changing subnet lists on the Admin Network
- Adding to or changing subnet lists on the Grid Network
- Linux: Adding interfaces to an existing node
- Changing IP addresses for all nodes in the grid

#### Changing a node's network configuration

You can change the network configuration of one or more nodes using the Change IP tool. You can change the configuration of the Grid Network, or add, change, or remove the Admin or Client Networks.

**What you'll need**

You must have the `Passwords.txt` file.

**About this task**

**Linux:** If you are adding a grid node to the Admin Network or Client Network for the first time, and you did not previously configure ADMIN_NETWORK_TARGET or CLIENT_NETWORK_TARGET in the node configuration file, you must do so now.

See the StorageGRID installation instructions for your Linux operating system.

**Appliances:** On StorageGRID appliances, if the Client or Admin Network was not configured in the StorageGRID Appliance Installer during the initial installation, the network cannot be added by using only the Change IP tool. First, you must place the appliance in maintenance mode, configure the links, return the appliance to normal operating mode, and then use the Change IP tool to modify the network configuration. See the procedure for configuring network links in the installation and maintenance instructions for your appliance.

You can change the IP address, subnet mask, gateway, or MTU value for one or more nodes on any network.

You can also add or remove a node from a Client Network or from an Admin Network:

- You can add a node to a Client Network or to an Admin Network by adding an IP address/subnet mask on that network to the node.
- You can remove a node from a Client Network or from an Admin Network by deleting the IP address/subnet mask for the node on that network.

  Nodes cannot be removed from the Grid Network.

IP address swaps are not allowed. If you must exchange IP addresses between grid nodes, you must use a temporary intermediate IP address.

If single sign-on (SSO) is enabled for your StorageGRID system and you are changing the IP address of an Admin Node, be aware that any relying party trust that was configured using the Admin Node's IP address (instead of its fully qualified domain name, as recommended) will become invalid. You will no longer be able to sign in to the node. Immediately after changing the IP address, you must update or reconfigure the node's relying party trust in Active Directory Federation Services (AD FS) with the new IP address. See the instructions for administering StorageGRID.

Any changes you make to the network using the Change IP tool are propagated to the installer firmware for the StorageGRID appliances. That way, if StorageGRID software is reinstalled on an appliance, or if an appliance is placed into maintenance mode, the networking configuration will be correct.

**Steps**

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.

2. Start the Change IP tool by entering the following command: `change-ip`
3. Enter the provisioning passphrase at the prompt.

   The main menu appears.

```
Welcome to the StorageGRID IP Change Tool.

Selected nodes: all

1:   SELECT NODES to edit
2:   EDIT IP/mask, gateway and MTU
3:   EDIT admin network subnet lists
4:   EDIT grid network subnet list
5:   SHOW changes
6:   SHOW full configuration, with changes highlighted
7:   VALIDATE changes
8:   SAVE changes, so you can resume later
9:   CLEAR all changes, to start fresh
10:  APPLY changes to the grid
0:   Exit

Selection:
```

4. Optionally select **1** to choose which nodes to update. Then select one of the following options:
   - **1**: Single node — select by name
   - **2**: Single node — select by site, then by name
   - **3**: Single node — select by current IP
   - **4**: All nodes at a site
   - **5**: All nodes in the grid

   **Note:** If you want to update all nodes, allow "all" to remain selected.

   After you make your selection, the main menu appears, with the **Selected nodes** field updated to reflect your choice. All subsequent actions are performed only on the nodes displayed.

5. On the main menu, select option **2** to edit IP/mask, gateway, and MTU information for the selected nodes.
   a. Select the network where you want to make changes:

- **1**: Grid network
- **2**: Admin network
- **3**: Client network
- **4**: All networks After you make your selection, the prompt shows the node name, network name (Grid, Admin, or Client), data type (IP/mask, Gateway, or MTU), and current value.

Editing the IP address, prefix length, gateway, or MTU of a DHCP-configured interface will change the interface to static. When you select to change an interface configured by DHCP, a warning is displayed to inform you that the interface will change to static.

Interfaces configured as `fixed` cannot be edited.

a. To set a new value, enter it in the format shown for the current value.

b. To leave the current value unchanged, press **Enter**.

c. If the data type is `IP/mask`, you can delete the Admin or Client Network from the node by entering **d** or **0.0.0.0/0**.

d. After editing all nodes you want to change, enter **q** to return to the main menu.

Your changes are held until cleared or applied.

6. Review your changes by selecting one of the following options:
   - **5**: Shows edits in output that is isolated to show only the changed item. Changes are highlighted in green (additions) or red (deletions), as shown in the example output:

```
================================================================================
Site: RTP
================================================================================
username-x Grid    IP      [    172.16.0.239/21 ]: 172.16.0.240/21
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Admin   IP      [    10.224.0.244/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.245/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.240/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.241/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.242/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.243/21 ]: 0.0.0.0/0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
Press Enter to continue
```

   - **6**: Shows edits in output that displays the full configuration. Changes are highlighted in green (additions) or red (deletions).

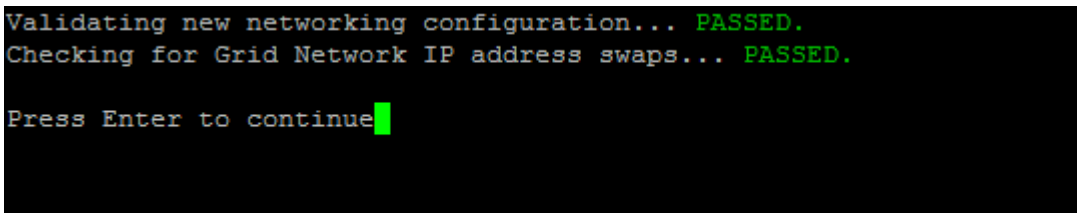
Certain command line interfaces might show additions and deletions using strikethrough formatting. Proper display depends on your terminal client supporting the necessary VT100 escape sequences.

7. Select option **7** to validate all changes.

   This validation ensures that the rules for the Grid, Admin, and Client Networks, such as not using overlapping subnets, are not violated.

   In this example, validation returned errors.

   In this example, validation passed.




8. Once validation passes, choose one of the following options:
   - **8**: Save unapplied changes.

     This option allows you to quit the Change IP tool and start it again later, without losing any unapplied changes.

   - **10**: Apply the new network configuration.
9. If you selected option **10**, choose one of the following options:
   - **apply**: Apply the changes immediately and automatically restart each node if necessary.

     If the new network configuration does not require any physical networking changes, you can select **apply** to apply the changes immediately. Nodes will be restarted automatically, if necessary. Nodes that need to be restarted will be displayed.

   - **stage**: Apply the changes the next time the nodes are restarted manually.

     If you need to make physical or virtual networking configuration changes for the new network configuration to function, you must use the **stage** option, shut down the affected nodes, make the necessary physical networking changes, and restart the affected nodes. If you select **apply** without first making these networking changes, the changes will usually fail.

     If you use the **stage** option, you must restart the node as soon as possible after staging to minimize disruptions.

- **cancel**: Do not make any network changes at this time.

  If you were unaware that the proposed changes require nodes to be restarted, you can defer the changes to minimize user impact. Selecting **cancel** returns you to the main menu and preserves your changes so you can apply them later.

  When you select **apply** or **stage**, a new network configuration file is generated, provisioning is performed, and nodes are updated with new working information.

  During provisioning, the output displays the status as updates are applied.

  ```
  Generating new grid networking description file...

  Running provisioning...

  Updating grid network configuration on Name
  ```

After applying or staging changes, a new Recovery Package is generated as a result of the grid configuration change.

10. If you selected **stage**, follow these steps after provisioning is complete:
    a. Make the physical or virtual networking changes that are required.

       **Physical networking changes**: Make the necessary physical networking changes, safely shutting down the node if necessary.

       **Linux**: If you are adding the node to an Admin Network or Client Network for the first time, ensure that you have added the interface as described in “Adding interfaces to an existing node.”

    b. Restart the affected nodes.
11. Select **0** to exit the Change IP tool after your changes are complete.
12. Download a new Recovery Package from the Grid Manager.
    a. Select **Maintenance** > **System** > **Recovery Package**.
    b. Enter the provisioning passphrase.

**Related information**

Linux: Adding interfaces to an existing node

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

SG100 & SG1000 services appliances

SG6000 storage appliances

SG5700 storage appliances

Administer StorageGRID

#### Adding to or changing subnet lists on the Admin Network

## You can add, delete, or change the subnets in the Admin Network Subnet List of one or more nodes.

### What you'll need

- You must have the `Passwords.txt` file.

You can add, delete, or change subnets to all nodes on the Admin Network Subnet List.

### Steps

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.
2. Start the Change IP tool by entering the following command: `change-ip`
3. Enter the provisioning passphrase at the prompt.

   The main menu appears.

```
Welcome to the StorageGRID IP Change Tool.

Selected nodes: all

1:   SELECT NODES to edit
2:   EDIT IP/mask, gateway and MTU
3:   EDIT admin network subnet lists
4:   EDIT grid network subnet list
5:   SHOW changes
6:   SHOW full configuration, with changes highlighted
7:   VALIDATE changes
8:   SAVE changes, so you can resume later
9:   CLEAR all changes, to start fresh
10:  APPLY changes to the grid
0:   Exit

Selection:
```

4. Optionally, limit the networks/nodes on which operations are performed. Choose one of the following:
   - Select the nodes to edit by choosing **1**, if you want to filter on specific nodes on which to perform the operation. Select one of the following options:
     - **1**: Single node (select by name)
     - **2**: Single node (select by site, then by name)
     - **3**: Single node (select by current IP)
     - **4**: All nodes at a site
     - **5**: All nodes in the grid

- **0**: Go back

- Allow “all” to remain selected. After the selection is made, the main menu screen appears. The Selected nodes field reflects your new selection, and now all operations selected will only be performed on this item.

5. On the main menu, select the option to edit subnets for the Admin Network (option **3**).
6. Choose one of the following:

   - Add a subnet by entering this command: `add CIDR`
   - Delete a subnet by entering this command: `del CIDR`
   - Set the list of subnets by entering this command: `set CIDR`

   > For all commands, you can enter multiple addresses using this format: `add CIDR, CIDR`

   Example: `add 172.14.0.0/16, 172.15.0.0/16, 172.16.0.0/16`

   > You can reduce the amount of typing required by using “up arrow” to recall previously typed values to the current input prompt, and then edit them if necessary.

   The example input below shows adding subnets to the Admin Network Subnet List:

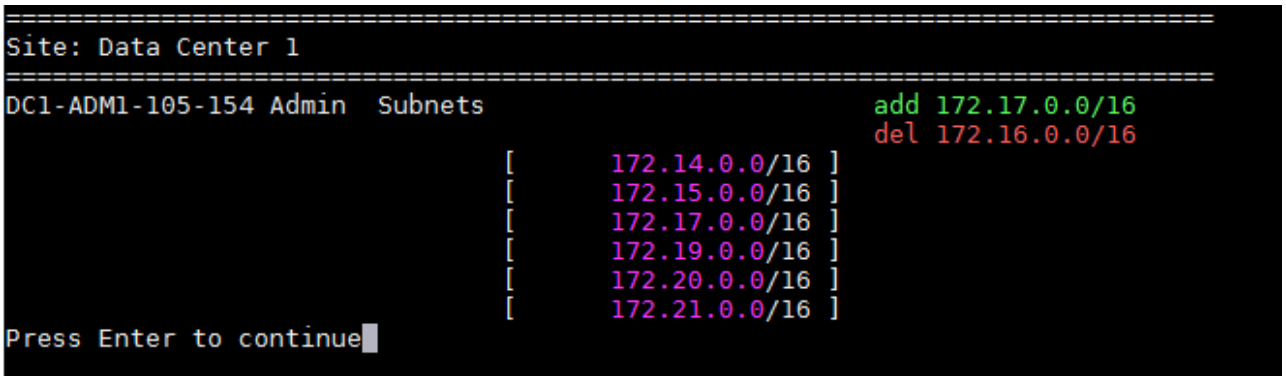
```
Editing: Admin Network Subnet List for node DK-10-224-5-20-G1

Press <enter> to use the list as shown
Use up arrow to recall a previously typed value, which you can then edit
Use 'add <CIDR> [, <CIDR>]' to add subnets <CIDR> [, <CIDR>] to the list
Use 'del <CIDR> [, <CIDR>]' to delete subnets <CIDR> [, <CIDR>] from the list
Use 'set <CIDR> [, <CIDR>]' to set the list to the given list
Use q to complete the editing session early and return to the previous menu

DK-10-224-5-20-G1
  10.0.0.0/8
  172.19.0.0/16
  172.21.0.0/16
  172.20.0.0/16

[add/del/set/quit <CIDR>, ...]: add 172.14.0.0/16, 172.15.0.0/16
```

7. When ready, enter **q** to go back to the main menu screen. Your changes are held until cleared or applied.

   > If you selected any of the "all" node selection modes in step 2, you must press **Enter** (without **q**) to get to the next node in the list.

8. Choose one of the following:

   - Select option **5** to show edits in output that is isolated to show only the changed item. Changes are highlighted in green (additions) or red (deletions), as shown in the example output below:




- Select option **6** to show edits in output that displays the full configuration. Changes are highlighted in green (additions) or red (deletions). **Note:** Certain terminal emulators might show additions and deletions using strikethrough formatting.

  When you attempt to change the subnet list, the following message is displayed:

  ```
  CAUTION: The Admin Network subnet list on the node might contain /32
  subnets derived from automatically applied routes that are not
  persistent. Host routes (/32 subnets) are applied automatically if
  the IP addresses provided for external services such as NTP or DNS
  are not reachable using default StorageGRID routing, but are
  reachable using a different interface and gateway. Making and
  applying changes to the subnet list will make all automatically
  applied subnets persistent. If you do not want that to happen, delete
  the unwanted subnets before applying changes. If you know that all
  /32 subnets in the list were added intentionally, you can ignore this
  caution.
  ```

  If you did not specifically assign the NTP and DNS server subnets to a network, StorageGRID creates a host route (/32) for the connection automatically. If, for example, you would rather have a /16 or /24 route for outbound connection to a DNS or NTP server, you should delete the automatically created /32 route and add the routes you want. If you do not delete the automatically created host route, it will be persisted after you apply any changes to the subnet list.

  > Although you can use these automatically discovered host routes, in general you should manually configure the DNS and NTP routes to ensure connectivity.

9. Select option **7** to validate all staged changes.

   This validation ensures that the rules for the Grid, Admin, and Client Networks are followed, such as using overlapping subnets.

10. Optionally, select option **8** to save all staged changes and return later to continue making changes.

    This option allows you to quit the Change IP tool and start it again later, without losing any unapplied changes.

11. Do one of the following:

- Select option **9** if you want to clear all changes without saving or applying the new network configuration.
- Select option **10** if you are ready to apply changes and provision the new network configuration. During provisioning, the output displays the status as updates are applied as shown in the following sample output:

```
Generating new grid networking description file...

Running provisioning...

Updating grid network configuration on Name
```

12. Download a new Recovery Package from the Grid Manager.
    a. Select **Maintenance** > **System** > **Recovery Package**.
    b. Enter the provisioning passphrase.

**Related information**

Configuring IP addresses

#### Adding to or changing subnet lists on the Grid Network

## You can use the Change IP tool to add or change subnets on the Grid Network.

**What you'll need**

- You have the `Passwords.txt` file.

**About this task**

You can add, delete, or change subnets in the Grid Network Subnet List. Changes will affect routing on all nodes in the grid.

> If you are making changes to the Grid Network Subnet List only, use the Grid Manager to add or change the network configuration. Otherwise, use the Change IP tool if the Grid Manager is inaccessible due to a network configuration issue, or you are performing both a Grid Network routing change and other network changes at the same time.

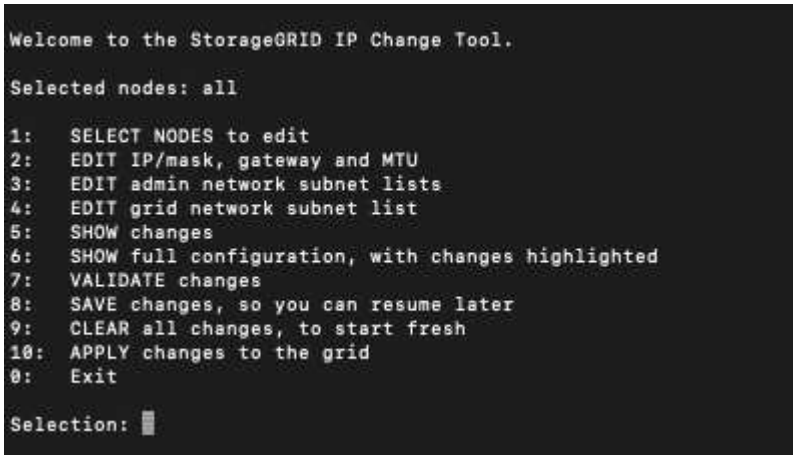
**Steps**

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@`*`primary_Admin_Node_IP`*
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Start the Change IP tool by entering the following command: `change-ip`
3. Enter the provisioning passphrase at the prompt.

The main menu appears.




4. On the main menu, select the option to edit subnets for the Grid Network (option **4**).

   Changes to the Grid Network Subnet List are grid-wide.

5. Choose one of the following:

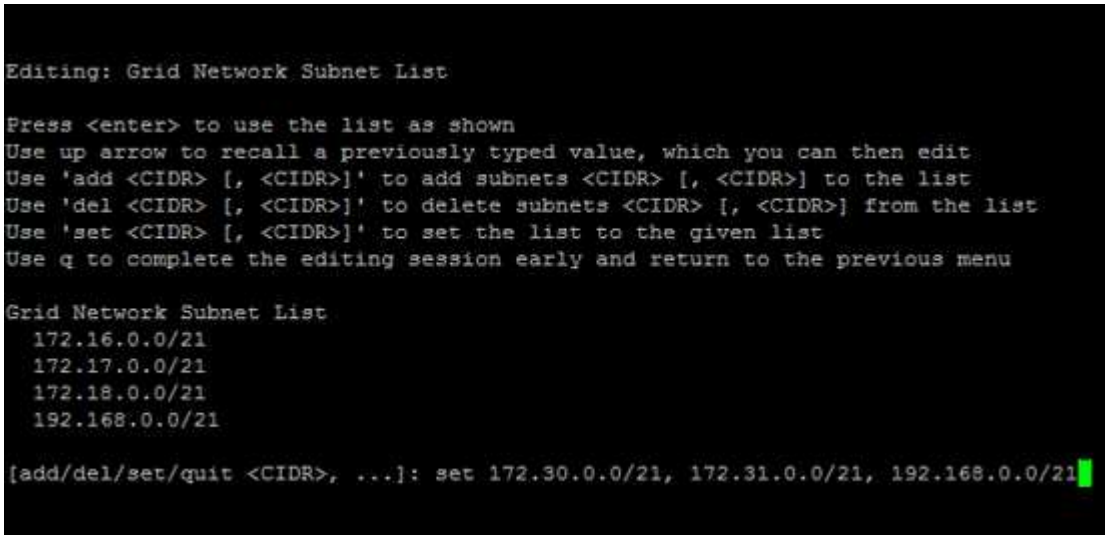
   - Add a subnet by entering this command: `add CIDR`
   - Delete a subnet by entering this command: `del CIDR`
   - Set the list of subnets by entering this command: `set CIDR`

     For all commands, you can enter multiple addresses using this format: `add CIDR, CIDR`

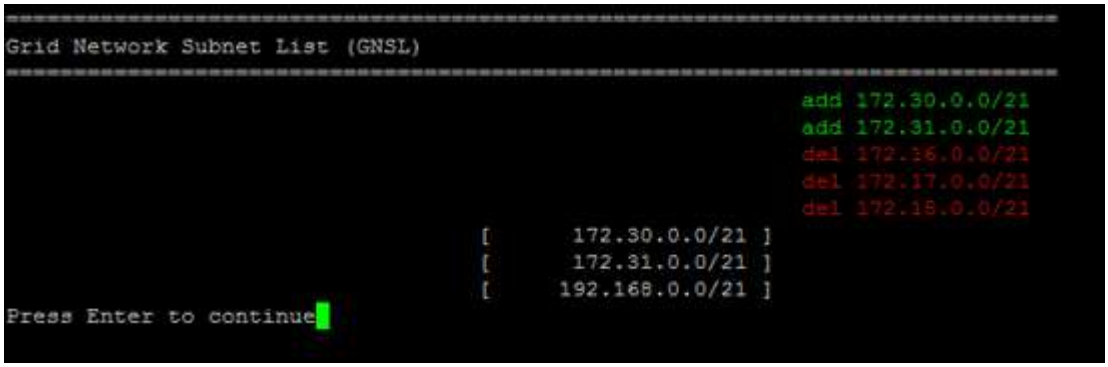
     Example: `add 172.14.0.0/16, 172.15.0.0/16, 172.16.0.0/16`

     You can reduce the amount of typing required by using “up arrow” to recall previously typed values to the current input prompt, and then edit them if necessary.

     The example input below shows setting subnets for the Grid Network Subnet List:



6. When ready, enter **q** to go back to the main menu screen. Your changes are held until cleared or applied.
7. Choose one of the following:
   - Select option **5** to show edits in output that is isolated to show only the changed item. Changes are highlighted in green (additions) or red (deletions), as shown in the example output below:




   - Select option **6** to show edits in output that displays the full configuration. Changes are highlighted in green (additions) or red (deletions).

> Certain command line interfaces might show additions and deletions using strikethrough formatting.

8. Select option **7** to validate all staged changes.

   This validation ensures that the rules for the Grid, Admin, and Client Networks are followed, such as using overlapping subnets.

9. Optionally, select option **8** to save all staged changes and return later to continue making changes.

   This option allows you to quit the Change IP tool and start it again later, without losing any unapplied changes.

10. Do one of the following:
    - Select option **9** if you want to clear all changes without saving or applying the new network configuration.
    - Select option **10** if you are ready to apply changes and provision the new network configuration. During provisioning, the output displays the status as updates are applied as shown in the following sample output:

```
Generating new grid networking description file...

Running provisioning...

Updating grid network configuration on Name
```

11. If you selected option **10** when making Grid Network changes, select one of the following options:
    - **apply**: Apply the changes immediately and automatically restart each node if necessary.

      If the new network configuration will function simultaneously with the old network configuration without

any external changes, you can use the **apply** option for a fully automated configuration change.

- **stage**: Apply the changes the next time the nodes are restarted.

  If you need to make physical or virtual networking configuration changes for the new network configuration to function, you must use the **stage** option, shut down the affected nodes, make the necessary physical networking changes, and restart the affected nodes.

  > If you use the **stage** option, you must restart the node as soon as possible after staging to minimize disruptions.

- **cancel**: Do not make any network changes at this time.

  If you were unaware that the proposed changes require nodes to be restarted, you can defer the changes to minimize user impact. Selecting **cancel** returns you to the main menu and preserves your changes so you can apply them later.

After applying or staging changes, a new Recovery Package is generated as a result of the grid configuration change.

12. If configuration is stopped due to errors, the following options are available:
    - To abort the IP change procedure and return to the main menu, enter **a**.
    - To retry the operation that failed, enter **r**.
    - To continue to the next operation, enter **c**.

      The failed operation can be retried later by selecting option **10** (Apply Changes) from the main menu. The IP change procedure will not be complete until all operations have completed successfully.

    - If you had to manually intervene (to reboot a node, for example) and are confident that the action the tool thinks has failed was actually completed successfully, enter **f** to mark it as successful and move to the next operation.
13. Download a new Recovery Package from the Grid Manager.
    a. Select **Maintenance** > **System** > **Recovery Package**.
    b. Enter the provisioning passphrase.

    > The Recovery Package file must be secured because it contains encryption keys and passwords that can be used to obtain data from the StorageGRID system.

**Related information**

Configuring IP addresses

#### Linux: Adding interfaces to an existing node

## If you want to add an interface to a Linux-based node that you did not install initially, you must use this procedure.

If you did not configure ADMIN_NETWORK_TARGET or CLIENT_NETWORK_TARGET in the node configuration file on the Linux host during installation, use this procedure to add the interface. For more information about the node configuration file, see the StorageGRID installation instructions for your Linux operating system.

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

You perform this procedure on the Linux server hosting the node that needs the new network assignment, not inside the node. This procedure only adds the interface to the node; a validation error occurs if you attempt to specify any other network parameters.

To provide addressing information, you must use the Change IP tool. See the information about changing a node's network configuration.

Changing a node's network configuration

**Steps**

1. Log in to the Linux server hosting the node that needs the new network assignment.
2. Edit the node configuration file at `/etc/storagegrid/nodes/node-name.conf`.

   Do not specify any other network parameters, or a validation error will result.

   a. Add the new network target.

   ```
   CLIENT_NETWORK_TARGET = bond0.3206
   ```

   b. Optional: Add a MAC address.

   ```
   CLIENT_NETWORK_MAC = aa:57:61:07:ea:5c
   ```

3. Run the node validate command: `sudo storagegrid node validate node-name`
4. Resolve all validation errors.
5. Run the node reload command: `sudo storagegrid node reload node-name`

**Related information**

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

Changing a node's network configuration

#### Changing IP addresses for all nodes in the grid

If you need to change the Grid Network IP address for all nodes in the grid, you must follow this special procedure. You cannot do a grid-wide Grid Network IP change using the procedure to change individual nodes.

**What you'll need**

- You must have the `Passwords.txt` file.

**About this task**

To ensure that the grid starts up successfully, you must make all the changes at once.

> This procedure applies to the Grid Network only. You cannot use this procedure to change IP addresses on the Admin or Client Networks.

If you want to change the IP addresses and MTU for the nodes at one site only, follow the instructions for changing a node's network configuration.

**Steps**

1. Plan ahead for changes that you need to make outside of the Change IP tool, such as changes to DNS or NTP, and changes to the single sign-on (SSO) configuration, if used.

   > If the existing NTP servers will not be accessible to the grid on the new IP addresses, add the new NTP servers before you perform the change-ip procedure.

   > If the existing DNS servers will not be accessible to the grid on the new IP addresses, add the new DNS servers before you perform the change-ip procedure.

   > If SSO is enabled for your StorageGRID system and any relying party trusts were configured using Admin Node IP addresses (instead of fully qualified domain names, as recommended), be prepared to update or reconfigure these relying party trusts in Active Directory Federation Services (AD FS) immediately after you change IP addresses. See the instructions for administering StorageGRID.

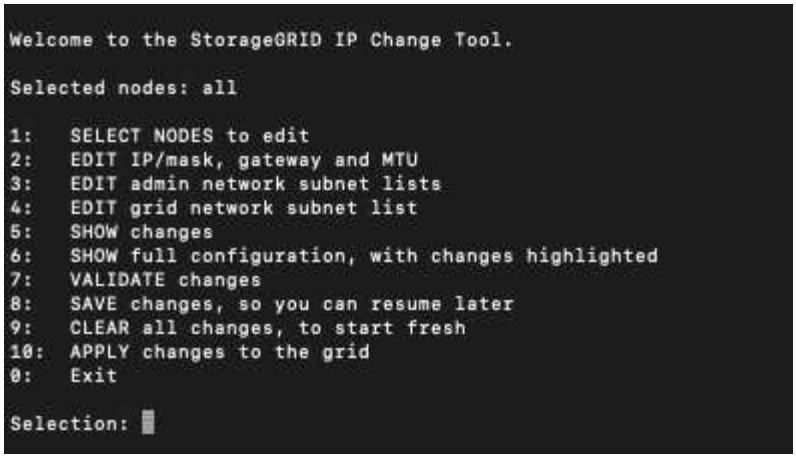
   > If necessary, add the new subnet for the new IP addresses.

2. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.

3. Start the Change IP tool by entering the following command: `change-ip`
4. Enter the provisioning passphrase at the prompt.

   The main menu appears. By default, the `Selected nodes` field is set to `all`.




5. On the main menu, select **2** to edit IP/subnet mask, gateway, and MTU information for all the nodes.
   a. Select **1** to make changes to the Grid Network.

      After you make your selection, the prompt shows the node names, Grid Network name, data type (IP/mask, Gateway, or MTU), and current values.

      Editing the IP address, prefix length, gateway, or MTU of a DHCP-configured interface will change the interface to static. A warning is displayed before each interface configured by DHCP.

      Interfaces configured as `fixed` cannot be edited.

   b. To set a new value, enter it in the format shown for the current value.
   c. After editing all nodes you want to change, enter **q** to return to the main menu.

      Your changes are held until cleared or applied.

6. Review your changes by selecting one of the following options:
   - **5**: Shows edits in output that is isolated to show only the changed item. Changes are highlighted in green (additions) or red (deletions), as shown in the example output:

```
================================================================================
Site: RTP
================================================================================
username-x Grid    IP      [    172.16.0.239/21 ]: 172.16.0.240/21
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Grid    MTU     [              1400 ]: 9000
username-x Admin   IP      [    10.224.0.244/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.245/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.240/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.241/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.242/21 ]: 0.0.0.0/0
username-x Admin   IP      [    10.224.0.243/21 ]: 0.0.0.0/0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   Gateway [         10.224.0.1 ]: 0.0.0.0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
username-x Admin   MTU     [              1400 ]: 0
Press Enter to continue
```

- **6**: Shows edits in output that displays the full configuration. Changes are highlighted in green (additions) or red (deletions).

> Certain command line interfaces might show additions and deletions using strikethrough formatting. Proper display depends on your terminal client supporting the necessary VT100 escape sequences.

7. Select option **7** to validate all changes.

   This validation ensures that the rules for the Grid Network, such as not using overlapping subnets, are not violated.

   In this example, validation returned errors.

```
Validating new networking configuration... FAILED.

  DK-10-224-5-20-G1: The admin subnet 172.18.0.0/16 overlaps the 172.18.0.0/21 grid network.
  DK-10-224-5-22-S1: Duplicate Grid IP 172.16.5.18 (also in use by DK-10-224-5-21-ADM1)

You must correct these errors before you can apply any changes.
Checking for Grid Network IP address swaps... PASSED.

Press Enter to continue
```

   In this example, validation passed.

```
Validating new networking configuration... PASSED.
Checking for Grid Network IP address swaps... PASSED.

Press Enter to continue
```

8. Once validation passes, select **10** to apply the new network configuration.
9. Select **stage** to apply the changes the next time the nodes are restarted.

> You must select **stage**. Do not perform a rolling restart, either manually or by selecting **apply** instead of **stage**; the grid will not start up successfully.

10. After your changes are complete, select **0** to exit the Change IP tool.
11. Shut down all nodes simultaneously.

> The entire grid must be shut down at once, so that all nodes are down at the same time.

12. Make the physical or virtual networking changes that are required.
13. Verify that all grid nodes are down.
14. Power on all nodes.
15. Once the grid starts up successfully:
    a. If you added new NTP servers, delete the old NTP server values.
    b. If you added new DNS servers, delete the old DNS server values.
16. Download the new Recovery Package from the Grid Manager.
    a. Select **Maintenance** > **System** > **Recovery Package**.
    b. Enter the provisioning passphrase.

**Related information**

Administer StorageGRID

Changing a node's network configuration

Adding to or changing subnet lists on the Grid Network

Shutting down a grid node

### Configuring DNS servers

You can add, remove, and update domain name system (DNS) servers, so that you can use fully qualified domain name (FQDN) hostnames rather than IP addresses.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the IP addresses of the DNS servers to configure.

**About this task**

Specifying DNS server information allows you to use fully qualified domain name (FQDN) hostnames rather than IP addresses for email or SNMP notifications and AutoSupport. Specifying at least two DNS servers is recommended.

Provide between two to six IP addresses for DNS servers. In general, select DNS servers that each site can access locally in the event of network islanding. This is to ensure an islanded site continues to have access to the DNS service. After configuring the grid-wide DNS server list, you can further customize the DNS server list for each node.

### Modifying the DNS configuration for a single grid node

If the DNS server information is omitted or incorrectly configured, a DNST alarm is triggered on each grid node's SSM service. The alarm clears when DNS is configured correctly and the new server information has reached all grid nodes.

**Steps**

1. Select **Maintenance** > **Network** > **DNS Servers**.
2. In the Servers section, add update, or remove DNS server entries, as necessary.

   The best practice is to specify at least two DNS servers per site. You can specify up to six DNS servers.

3. Click **Save**.

#### Modifying the DNS configuration for a single grid node

## Rather than configure the Domain Name System (DNS) globally for the entire deployment, you can run a script to configure DNS differently for each grid node.

In general, you should use the **Maintenance** > **Network** > **DNS Servers** option on the Grid Manager to configure DNS servers. Only use the following script if you need to use different DNS servers for different grid nodes.

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.

   e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
   f. Enter the SSH Access Password listed in the `Passwords.txt` file.
2. Log in to the node you want to update with a custom DNS configuration: `ssh node_IP_address`
3. Run the DNS setup script: `setup_resolv.rb.`

   The script responds with the list of supported commands.

```
Tool to modify external name servers

available commands:
  add search <domain>
               add a specified domain to search list
               e.g.> add search netapp.com
  remove search <domain>
               remove a specified domain from list
               e.g.> remove search netapp.com
  add nameserver <ip>
               add a specified IP address to the name server list
               e.g.> add nameserver 192.0.2.65
  remove nameserver <ip>
               remove a specified IP address from list
               e.g.> remove nameserver 192.0.2.65
  remove nameserver all
               remove all nameservers from list
  save         write configuration to disk and quit
  abort        quit without saving changes
  help         display this help message


Current list of name servers:
    192.0.2.64
Name servers inherited from global DNS configuration:
    192.0.2.126
    192.0.2.127
Current list of search entries:
    netapp.com

Enter command [`add search <domain>|remove search <domain>|add
nameserver <ip>`]
              [`remove nameserver <ip>|remove nameserver
all|save|abort|help`]
```

4. Add the IPv4 address of a server that provides domain name service for your network: `add <nameserver IP_address>`
5. Repeat the `add nameserver` command to add name servers.
6. Follow instructions as prompted for other commands.
7. Save your changes and exit the application: `save`
8. Close the command shell on the server: `exit`
9. For each grid node, repeat the steps from logging into the node through closing the command shell.
10. When you no longer require passwordless access to other servers, remove the private key from the SSH

agent. Enter: `ssh-add -D`

### Configuring NTP servers

You can add, update, or remove network time protocol (NTP) servers to ensure that data is synchronized accurately between grid nodes in your StorageGRID system.

#### What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.
- You must have the IPv4 addresses of the NTP servers to configure.

#### About this task

The StorageGRID system uses the network time protocol (NTP) to synchronize time between all grid nodes in the grid.

At each site, at least two nodes in the StorageGRID system are assigned the primary NTP role. They synchronize to a suggested minimum of four, and a maximum of six, external time sources and with each other. Every node in the StorageGRID system that is not a primary NTP node acts as an NTP client and synchronizes with these primary NTP nodes.

The external NTP servers connect to the nodes to which you previously assigned Primary NTP roles. For this reason, specifying at least two nodes with Primary NTP roles is recommended.

> Make sure that at least two nodes at each site can access at least four external NTP sources. If only one node at a site can reach the NTP sources, timing issues will occur if that node goes down. In addition, designating two nodes per site as primary NTP sources ensures accurate timing if a site is isolated from the rest of the grid.

The specified external NTP servers must use the NTP protocol. You must specify NTP server references of Stratum 3 or better to prevent issues with time drift.

> When specifying the external NTP source for a production-level StorageGRID installation, do not use the Windows Time (W32Time) service on a version of Windows earlier than Windows Server 2016. The time service on earlier versions of Windows is not sufficiently accurate and is not supported by Microsoft for use in high-accuracy environments, such as StorageGRID.

Support boundary to configure the Windows Time service for high-accuracy environments

If you encounter problems with the stability or availability of the NTP servers originally specified during installation, you can update the list of external NTP sources that the StorageGRID system uses by adding additional servers, or updating or removing existing servers.

#### Steps

1. Select **Maintenance** > **Network** > **NTP Servers**.
2. In the Servers section, add update, or remove NTP server entries, as necessary.

   You should include at least 4 NTP servers, and you can specify up to 6 servers.

3. In the **Provisioning Passphrase** text box, enter the provisioning passphrase for your StorageGRID

system and click **Save**.

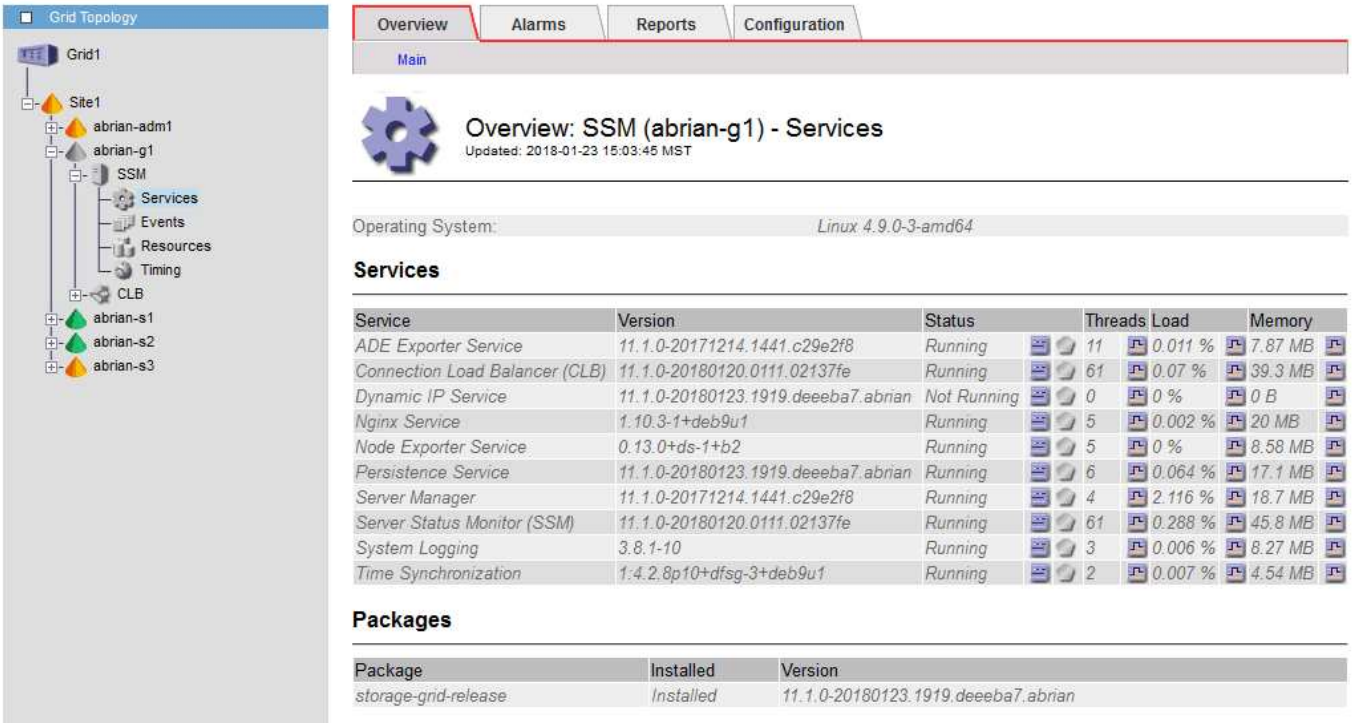
The status of the procedure is displayed at the top of the page. The page is disabled until the configuration updates are complete.

> If all of your NTP servers fail the connection test after you save the new NTP servers, do not proceed. Contact technical support.

### Restoring network connectivity for isolated nodes

Under certain circumstances, such as site- or grid-wide IP address changes, one or more groups of nodes might not be able to contact the rest of the grid.

In the Grid Manager (**Support** > **Tools** > **Grid Topology**), if a node is gray, or if a node is blue with many of its services showing a status other than Running, you should check for node isolation.




Some of the consequences of having isolated nodes include the following:

- If multiple nodes are isolated, you might not be able to sign in to or access the Grid Manager.
- If multiple nodes are isolated, the storage usage and quota values shown on the Dashboard for the Tenant Manager might be out of date. The totals will be updated when network connectivity is restored.

To resolve the isolation issue, you run a command line utility on each isolated node or on one node in a group (all nodes in a subnet that does not contain the primary Admin Node) that is isolated from the grid. The utility provides the nodes with the IP address of a non-isolated node in the grid, which allows the isolated node or group of nodes to contact the entire grid again.

> If the multicast Domain Name System (mDNS) is disabled in the networks, the command line utility might have to be run on each isolated node.

**Steps**

1. Access the node and check `/var/local/log/dynip.log` for isolation messages.

   For example:

   ```
   [2018-01-09T19:11:00.545] UpdateQueue - WARNING -- Possible isolation,
   no contact with other nodes.
   If this warning persists, manual action may be required.
   ```

   If you are using the VMware console, it will contain a message that the node might be isolated.

   On Linux deployments, isolation messages would appear in `/var/log/storagegrid/node/<nodename>.log` files.

2. If the isolation messages are recurring and persistent, run the following command:

   `add_node_ip.py <address\>`

   where `<address\>` is the IP address of a remote node that is connected to the grid.

   ```
   # /usr/sbin/add_node_ip.py  10.224.4.210

   Retrieving local host information
   Validating remote node at address 10.224.4.210
   Sending node IP hint for 10.224.4.210 to local node
   Local node found on remote node. Update complete.
   ```

3. Verify the following for each node that was previously isolated:
   - The node's services have started.
   - The status of the Dynamic IP Service is "Running" after you run the `storagegrid-status` command.
   - In the Grid Topology tree, the node no longer appears disconnected from the rest of the grid.

> If running the `add_node_ip.py` command does not solve the problem, there could be other networking issues that need to be resolved.

## Host-level and middleware procedures

Some maintenance procedures are specific to Linux or VMware deployments of StorageGRID, or are specific to other components of the StorageGRID solution.

#### Linux: Migrating a grid node to a new host

You can migrate StorageGRID nodes from one Linux host to another to perform host maintenance (such as OS patching and reboot) without impacting the functionality or availability of your grid.

You migrate one or more nodes from one Linux host (the “source host”) to another Linux host (the “target host”). The target host must have previously been prepared for StorageGRID use.

> You can use this procedure only if you planned your StorageGRID deployment to include migration support.

To migrate a grid node to a new host, both of the following conditions must be true:

- Shared storage is used for all per-node storage volumes
- Network interfaces have consistent names across hosts

> In a production deployment, do not run more than one Storage Node on a single host. Using a dedicated host for each Storage Node provides an isolated failure domain.

Other types of nodes, such as Admin Nodes or Gateway Nodes, can be deployed on the same host. However, if you have multiple nodes of the same type (two Gateway Nodes, for example), do not install all instances on the same host.

For more information, see “Node migration requirements” in the StorageGRID installation instructions for your Linux operating system.

**Related information**

Deploying new Linux hosts

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

#### Linux: Exporting the node from the source host

## Shut down the grid node and export it from the source Linux host.

Run the following command on the source Linux host.

1. Obtain the status of all nodes currently running on the source host.

```
sudo storagegrid node status all
```

```
Name Config-State Run-State

DC1-ADM1 Configured Running

DC1-ARC1 Configured Running

DC1-GW1 Configured Running

DC1-S1 Configured Running

DC1-S2 Configured Running

DC1-S3 Configured Running
```

2. Identify the name of the node you want to migrate, and stop it if its Run-State is `Running`.

```
sudo storagegrid node stop DC1-S3
```

```
Stopping node DC1-S3

Waiting up to 630 seconds for node shutdown
```

3. Export the node from the source host.

```
sudo storagegrid node export DC1-S3
```

```
Finished exporting node DC1-S3 to /dev/mapper/sgws-dc1-s3-var-local.

Use 'storagegrid node import /dev/mapper/sgws-dc1-s3-var-local' if you

want to import it again.
```

4. Take note of the `import command suggested in the output of the `export` command.

   You will run this command on the target host in the next step.

#### Linux: Importing the node on the target host

After exporting the node from the source host, you import and validate the node on the target Linux host. Validation confirms that the node has access to the same block storage and network interface devices as it had on the source host.

Run the following command on the target Linux host.

1. Import the node on the target host.

```
sudo storagegrid node import /dev/mapper/sgws-dc1-s3-var-local
```

```
Finished importing node DC1-S3 from /dev/mapper/sgws-dc1-s3-var-local.

You should run 'storagegrid node validate DC1-S3'
```

2. Validate the node configuration on the new host.

```
sudo storagegrid node validate DC1-S3
```

```
Confirming existence of node DC1-S3... PASSED

Checking configuration file /etc/storagegrid/nodes/DC1-S3.conf for node DC1-
```

```
S3... PASSED

Checking for duplication of unique values... PASSED
```

3. If any validation errors occur, address them before starting the migrated node.

   For troubleshooting information, see the StorageGRID installation instructions for your Linux operating system.

**Related information**

Install Red Hat Enterprise Linux or CentOS

Install Ubuntu or Debian

#### Linux: Starting the migrated node

After you validate the migrated node, you start the node by running a command on the target Linux host.

**Steps**

1. Start the node on the new host.

```
sudo storagegrid node start DC1-S3
Starting node DC1-S3
```

2. In the Grid Manager, verify that the status of the node is green with no alarms raised against it.

   > Verifying that the status of the node is green ensures that the migrated node has fully restarted and rejoined the grid. If the status is not green, do not migrate any additional nodes so that you will not have more than one node out of service.

   If you are unable to access the Grid Manager, wait for 10 minutes, then run the following command:

   `sudo storagegrid node status` ***`node-name`***

   Confirm that the migrated node has a Run-State of `Running`.

### Archive Node maintenance for TSM middleware

Archive Nodes might be configured to target either tape through a TSM middleware server or the cloud through the S3 API. Once configured, an Archive Node's target cannot be changed.

If the server hosting the Archive Node fails, replace the server and follow the appropriate recovery procedure.

#### Fault with archival storage devices

If you determine that there is a fault with the archival storage device that the Archive Node is accessing through Tivoli Storage Manager (TSM), take the Archive Node offline to limit the number of alarms displayed in the StorageGRID system. You can then use the administrative tools of the TSM server or the storage device,

or both, to further diagnose and resolve the problem.

### Taking the Target component offline

Before undertaking any maintenance of the TSM middleware server that might result in it becoming unavailable to the Archive Node, take the Target component offline to limit the number of alarms that are triggered if the TSM middleware server becomes unavailable.

#### What you'll need

You must be signed in to the Grid Manager using a supported browser.

#### Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select **Archive Node** > **ARC** > **Target** > **Configuration** > **Main**.
3. Change the value of Tivoli Storage Manager State to **Offline**, and click **Apply Changes**.
4. After maintenance is complete, change the value of Tivoli Storage Manager State to **Online**, and click **Apply Changes**.

#### Tivoli Storage Manager administrative tools

The dsmadmc tool is the administrative console for the TSM middleware server that is installed on the Archive Node. You can access the tool by typing `dsmadmc` at the command line of the server. Log in to the administrative console using the same administrative user name and password that is configured for the ARC service.

The `tsmquery.rb` script was created to generate status information from dsmadmc in a more readable form. You can run this script by entering the following command at the command line of the Archive Node:
`/usr/local/arc/tsmquery.rb status`

For more information about the TSM administrative console dsmadmc, see the *Tivoli Storage Manager for Linux: Administrator's Reference*.

#### Object permanently unavailable

When the Archive Node requests an object from the Tivoli Storage Manager (TSM) server and the retrieval fails, the Archive Node retries the request after an interval of 10 seconds. If the object is permanently unavailable (for example, because the object is corrupted on tape), the TSM API has no way to indicate this to the Archive Node, so the Archive Node continues to retry the request.

When this situation occurs, an alarm is triggered, and the value continues to increase. To see the alarm, select **Support** > **Tools** > **Grid Topology**. Then, select **Archive Node** > **ARC** > **Retrieve** > **Request Failures**.

If the object is permanently unavailable, you must identify the object and then manually cancel the Archive Node's request as described in the procedure, Determining if objects are permanently unavailable.

A retrieval can also fail if the object is temporarily unavailable. In this case, subsequent retrieval requests should eventually succeed.

If the StorageGRID system is configured to use an ILM rule that creates a single object copy and that copy cannot be retrieved, the object is lost and cannot be recovered. However, you must still follow the procedure to determine if the object is permanently unavailable to "clean up" the StorageGRID system, to cancel the Archive Node's request, and to purge metadata for the lost object.

### Determining if objects are permanently unavailable

You can determine if objects are permanently unavailable by making a request using the TSM administrative console.

#### What you'll need

- You must have specific access permissions.
- You must have the `Passwords.txt` file.
- You must know the IP address of an Admin Node.

#### About this task

This example is provided for your information only; this procedure cannot help you identify all failure conditions that might result in unavailable objects or tape volumes. For information about TSM administration, see TSM Server documentation.

#### Steps

1. Log in to an Admin Node:
   a. Enter the following command: `ssh admin@Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
2. Identify the object or objects that could not be retrieved by the Archive Node:
   a. Go to the directory containing the audit log files: `cd /var/local/audit/export`

      The active audit log file is named audit.log. Once a day, the active `audit.log` file is saved, and a new `audit.log` file is started. The name of the saved file indicates when it was saved, in the format `yyyy-mm-dd.txt`. After a day, the saved file is compressed and renamed, in the format `yyyy-mm-dd.txt.gz`, which preserves the original date.

   b. Search the relevant audit log file for messages indicating that an archived object could not be retrieved. For example, enter: `grep ARCE audit.log | less -n`

      When an object cannot be retrieved from an Archive Node, the ARCE audit message (Archive Object Retrieve End) displays ARUN (archive middleware unavailable) or GERR (general error) in the result field. The following example line from the audit log shows that the ARCE message terminated with the result ARUN for CBID 498D8A1F681F05B3.

      ```
      [AUDT:[CBID(UI64):0x498D8A1F681F05B3][VLID(UI64):□20091127][RSLT(FC32
      ):ARUN][AVER(UI32):7]
      [ATIM(UI64):1350613602969243][ATYP(FC32):ARCE][ANID(UI32):13959984][A
      MID(FC32):ARCI]
      [ATID(UI64):4560349751312520631]]
      ```

      For more information see the instructions for understanding audit messages.

   c. Record the CBID of each object that had a request failure.

      You might also want to record the following additional information used by the TSM to identify objects saved by the Archive Node:

- **File Space Name**: Equivalent to the Archive Node ID. To find the Archive Node ID, select **Support** > **Tools** > **Grid Topology**. Then, select **Archive Node** > **ARC** > **Target** > **Overview**.
- **High Level Name**: Equivalent to the volume ID assigned to the object by the Archive Node. The volume ID takes the form of a date (for example, `20091127`), and is recorded as the VLID of the object in archive audit messages.
- **Low Level Name**: Equivalent to the CBID assigned to an object by the StorageGRID system.

d. Log out of the command shell: `exit`

3. Check the TSM server to see if the objects identified in step 2 are permanently unavailable:

a. Log in to the administrative console of the TSM server: `dsmadmc`

Use the administrative user name and password that are configured for the ARC service. Enter the user name and password in the Grid Manager. (To see the user name, select **Support** > **Tools** > **Grid Topology**. Then, select **Archive Node** > **ARC** > **Target** > **Configuration**.)

b. Determine if the object is permanently unavailable.

For example, you might search the TSM activity log for a data integrity error for that object. The following example shows a search of the activity log for the past day for an object with CBID `498D8A1F681F05B3`.

```
> query actlog begindate=-1 search=276C14E94082CC69
12/21/2008 05:39:15 ANR0548W Retrieve or restore
failed for session 9139359 for node DEV-ARC-20 (Bycast ARC)
processing file space /19130020 4 for file /20081002/
498D8A1F681F05B3 stored as Archive - data
integrity error detected. (SESSION: 9139359)
>
```

Depending on the nature of the error, the CBID might not be recorded in the TSM activity log. You might need to search the log for other TSM errors around the time of the request failure.

c. If an entire tape is permanently unavailable, identify the CBIDs for all objects stored on that volume: `query content TSM_Volume_Name`

where `TSM_Volume_Name` is the TSM name for the unavailable tape. The following is an example of the output for this command:

```
 > query content TSM-Volume-Name
Node Name     Type Filespace  FSID Client's Name for File Name
------------- ---- ---------- ---- --------------------------
DEV-ARC-20    Arch /19130020  216  /20081201/ C1D172940E6C7E12
DEV-ARC-20    Arch /19130020  216  /20081201/ F1D7FBC2B4B0779E
```

The `Client's Name for File Name` is the same as the Archive Node volume ID (or TSM "high level name") followed by the object's CBID (or TSM "low level name"). That is, the `Client's Name for File Name` takes the form `/Archive Node volume ID /CBID`. In the first line of the

example output, the `Client's Name for File Name` is `/20081201/ C1D172940E6C7E12`.

Recall also that the `Filespace` is the node ID of the Archive Node.

You will need the CBID of each object stored on the volume and the node ID of the Archive Node to cancel the retrieval request.

4. For each object that is permanently unavailable, cancel the retrieval request and issue a command to inform the StorageGRID system that the object copy was lost:

> Use the ADE Console with caution. If the console is used improperly, it is possible to interrupt system operations and corrupt data. Enter commands carefully, and only use the commands documented in this procedure.

   a. If you are not already logged in to the Archive Node, log in as follows:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.
   b. Access the ADE console of the ARC service: `telnet localhost 1409`
   c. Cancel the request for the object: `/proc/BRTR/cancel -c CBID`

      where `CBID` is the identifier of the object that cannot be retrieved from the TSM.

      If the only copies of the object are on tape, the "bulk retrieval" request is canceled with a message, "1 requests canceled". If copies of the object exist elsewhere in the system, the object retrieval is processed by a different module so the response to the message is "0 requests canceled".

   d. Issue a command to notify the StorageGRID system that an object copy has been lost and that an additional copy must be made: `/proc/CMSI/Object_Lost CBID node_ID`

      where `CBID` is the identifier of the object that cannot be retrieved from the TSM server, and `node_ID` is the node ID of the Archive Node where the retrieval failed.

      You must enter a separate command for each lost object copy: entering a range of CBIDs is not supported.

      In most cases, the StorageGRID system immediately begins to make additional copies of object data to ensure that the system's ILM policy is followed.

      However, if the ILM rule for the object specified that only one copy be made and that copy has now been lost, the object cannot be recovered. In this case running the `Object_Lost` command purges the lost object's metadata from the StorageGRID system.

      When the `Object_Lost` command completes successfully, the following message is returned:

```
CLOC_LOST_ANS returned result 'SUCS'
```

> The `/proc/CMSI/Object_Lost` command is only valid for lost objects that are stored on Archive Nodes.

e. Exit the ADE Console: `exit`

f. Log out of the Archive Node: `exit`

5. Reset the value of Request Failures in the StorageGRID system:
   a. Go to **Archive Node** > **ARC** > **Retrieve** > **Configuration**, and select **Reset Request Failure Count**.
   b. Click **Apply Changes**.

**Related information**

Administer StorageGRID

Review audit logs

### VMware: Configuring a virtual machine for automatic restart

If the virtual machine does not restart after VMware vSphere Hypervisor is restarted, you might need to configure the virtual machine for automatic restart.

You should perform this procedure if you notice that a virtual machine does not restart while you are recovering a grid node or performing another maintenance procedure.

**Steps**

1. In the VMware vSphere Client tree, select the virtual machine that is not started.
2. Right-click the virtual machine, and select **Power on**.
3. Configure VMware vSphere Hypervisor to restart the virtual machine automatically in future.

## Grid node procedures

You might need to perform procedures on a specific grid node. While you can perform a few of these procedures from Grid Manager, most of the procedures require you to access Server Manager from the node's command line.

Server Manager runs on every grid node to supervise the starting and stopping of services and to ensure that services gracefully join and leave the StorageGRID system. Server Manager also monitors the services on every grid node and will automatically attempt to restart any services that report faults.

> You should access Server Manager only if technical support has directed you to do so.

> You must close the current command shell session and log out after you are finished with Server Manager. Enter: `exit`

**Choices**

- Viewing Server Manager status and version
- Viewing current status of all services
- Starting Server Manager and all services

- Restarting Server Manager and all services
- Stopping Server Manager and all services
- Viewing current status of a service
- Stopping a service
- Placing an appliance into maintenance mode
- Forcing a service to terminate
- Starting or restarting a service
- Removing port remaps
- Removing port remaps on bare metal hosts
- Rebooting a grid node
- Shutting down a grid node
- Powering down a host
- Powering off and on all nodes in the grid
- Using a DoNotStart file
- Troubleshooting Server Manager

### Viewing Server Manager status and version

For each grid node, you can view the current status and version of Server Manager running on that grid node. You can also obtain the current status of all services running on that grid node.

#### What you'll need

You must have the `Passwords.txt` file.

#### Steps

1. Log in to the grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. View the current status of Server Manager running on the grid node: **`service servermanager status`**

   The current status of Server Manager running on the grid node is reported (running or not). If Server Manager's status is `running`, the time it has been running since last it was started is listed. For example:

   ```
   servermanager running for 1d, 13h, 0m, 30s
   ```

   This status is the equivalent of the status shown in the header of the local console display.

3. View the current version of Server Manager running on a grid node: **`service servermanager version`**

   The current version is listed. For example:

   ```
   11.1.0-20180425.1905.39c9493
   ```

4. Log out of the command shell: **`exit`**

### Viewing current status of all services

You can view the current status of all services running on a grid node at any time.

#### What you'll need

You must have the `Passwords.txt` file.

#### Steps

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

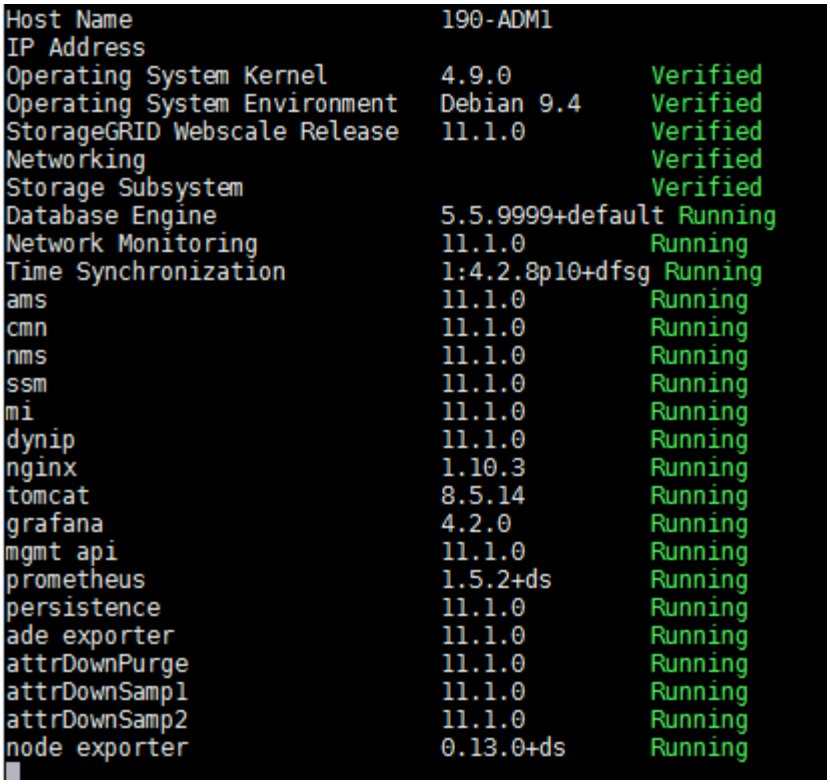
   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. View the status of all services running on the grid node: `storagegrid-status`

   For example, the output for the primary Admin Node shows the current status of the AMS, CMN, and NMS services as Running. This output is updated immediately if the status of a service changes.




3. Return to the command line, press **Ctrl+C**.
4. Optionally, view a static report for all services running on the grid node: `/usr/local/servermanager/reader.rb`

   This report includes the same information as the continuously updated report, but it is not updated if the status of a service changes.

5. Log out of the command shell: `exit`

### Starting Server Manager and all services

You might need to start Server Manager, which also starts all services on the grid node.

#### What you'll need

You must have the `Passwords.txt` file.

#### About this task

Starting Server Manager on a grid node where it is already running results in a restart of Server Manager and all services on the grid node.

#### Steps

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

2. Start Server Manager: `service servermanager start`
3. Log out of the command shell: `exit`

### Restarting Server Manager and all services

You might need to restart server manager and all services running on a grid node.

#### What you'll need

You must have the `Passwords.txt` file.

#### Steps

1. Log in to the grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Restart Server Manager and all services on the grid node: `service servermanager restart`

   Server Manager and all services on the grid node are stopped and then restarted.

   > Using the `restart` command is the same as using the `stop` command followed by the `start` command.

3. Log out of the command shell: `exit`

### Stopping Server Manager and all services

Server Manager is intended to run at all times, but you might need to stop Server Manager and all services running on a grid node.

#### What you'll need

You must have the `Passwords.txt` file.

#### About this task

The only scenario that requires you to stop Server Manager while keeping the operating system running is when you need to integrate Server Manager to other services. If there is a requirement to stop the Server Manager for servicing of the hardware or reconfiguration of the server, the entire server should be halted.

#### Steps

1. Log in to the grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`

b. Enter the password listed in the `Passwords.txt` file.

c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

2. Stop Server manager and all services running on the grid node: `service servermanager stop`

   Server Manager and all services running on the grid node are gracefully terminated. Services can take up to 15 minutes to shut down.

3. Log out of the command shell: `exit`

#### Viewing current status of a service

You can view the current status of a services running on a grid node at any time.

##### What you'll need

You must have the `Passwords.txt` file.

##### Steps

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. View the current status of a service running on a grid node: `**service *servicename* status** The current status of the requested service running on the grid node is reported (running or not). For example:

   ```
   cmn running for 1d, 14h, 21m, 2s
   ```

3. Log out of the command shell: **`exit`**

#### Stopping a service

Some maintenance procedures require you to stop a single service while keeping other services on the grid node running. Only stop individual services when directed to do so by a maintenance procedure.

##### What you'll need

You must have the `Passwords.txt` file.

##### About this task

When you use these steps to "administratively stop" a service, Server Manager will not automatically restart

the service. You must either start the single service manually or restart Server Manager.

If you need to stop the LDR service on a Storage Node, be aware that it might take a while to stop the service if there are active connections.

**Steps**

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Stop an individual service: `service servicename stop`

   For example:

   ```
   service ldr stop
   ```

   Services can take up to 11 minutes to stop.

3. Log out of the command shell: `exit`

**Related information**

Forcing a service to terminate

### Placing an appliance into maintenance mode

You must place the appliance into maintenance mode before performing specific maintenance procedures.

**What you'll need**

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission. For details, see the instructions for administering StorageGRID.

**About this task**

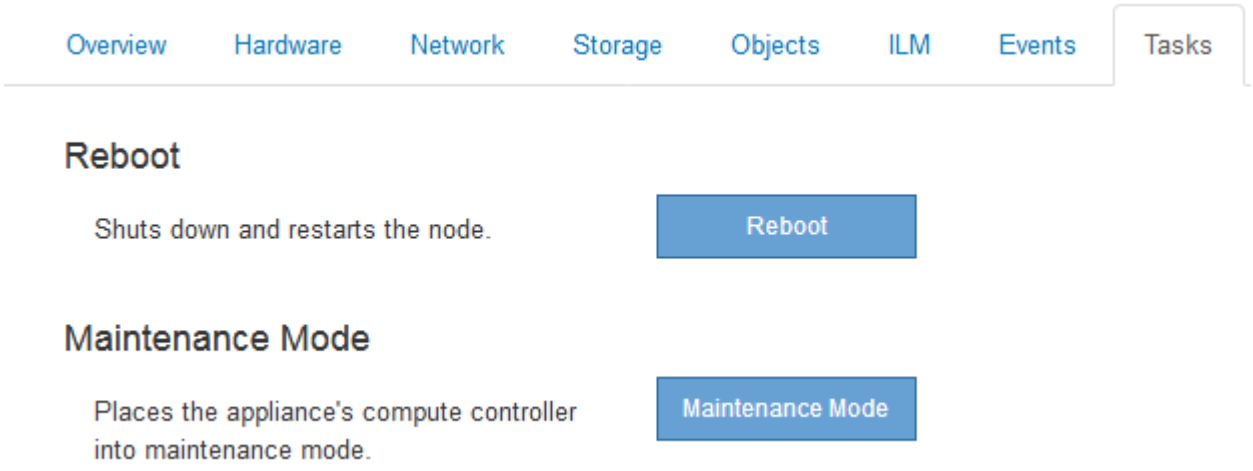
Placing a StorageGRID appliance into maintenance mode might make the appliance unavailable for remote access.

The password and host key for a StorageGRID appliance in maintenance mode remain the same as they were when the appliance was in service.

**Steps**

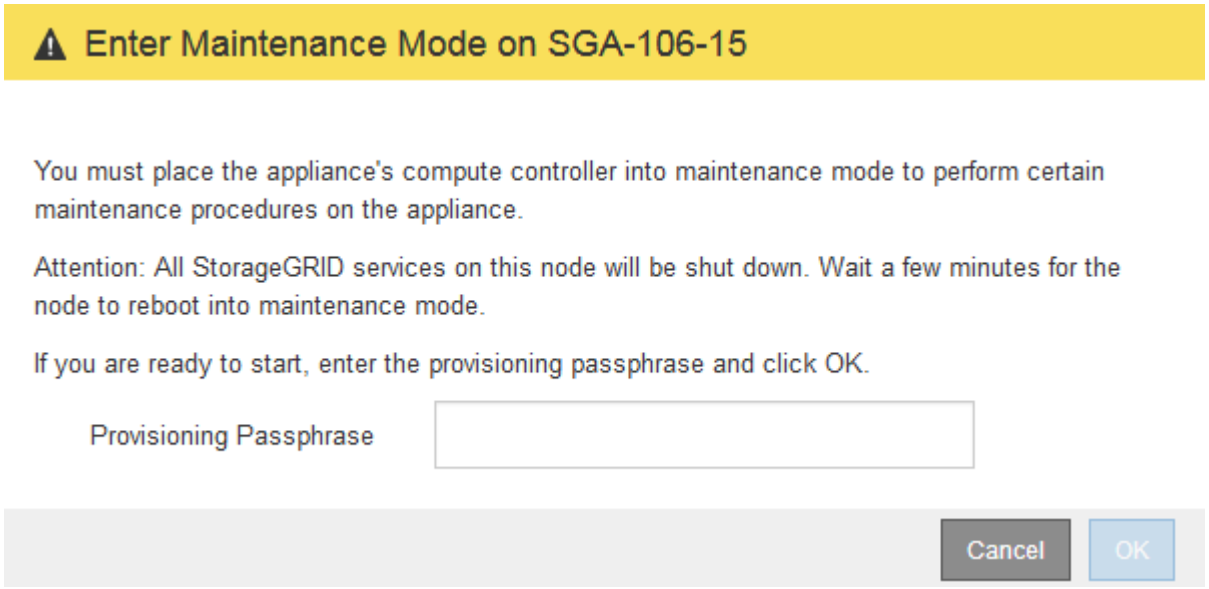
1. From the Grid Manager, select **Nodes**.

2. From the tree view of the Nodes page, select the appliance Storage Node.
3. Select **Tasks**.



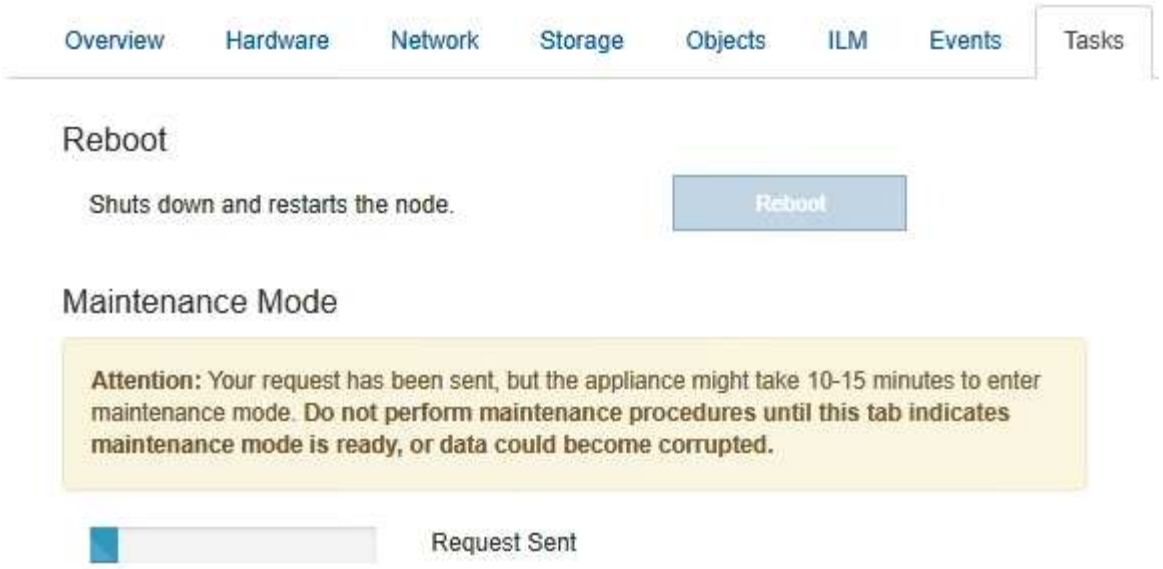

4. Select **Maintenance Mode**.

   A confirmation dialog box appears.

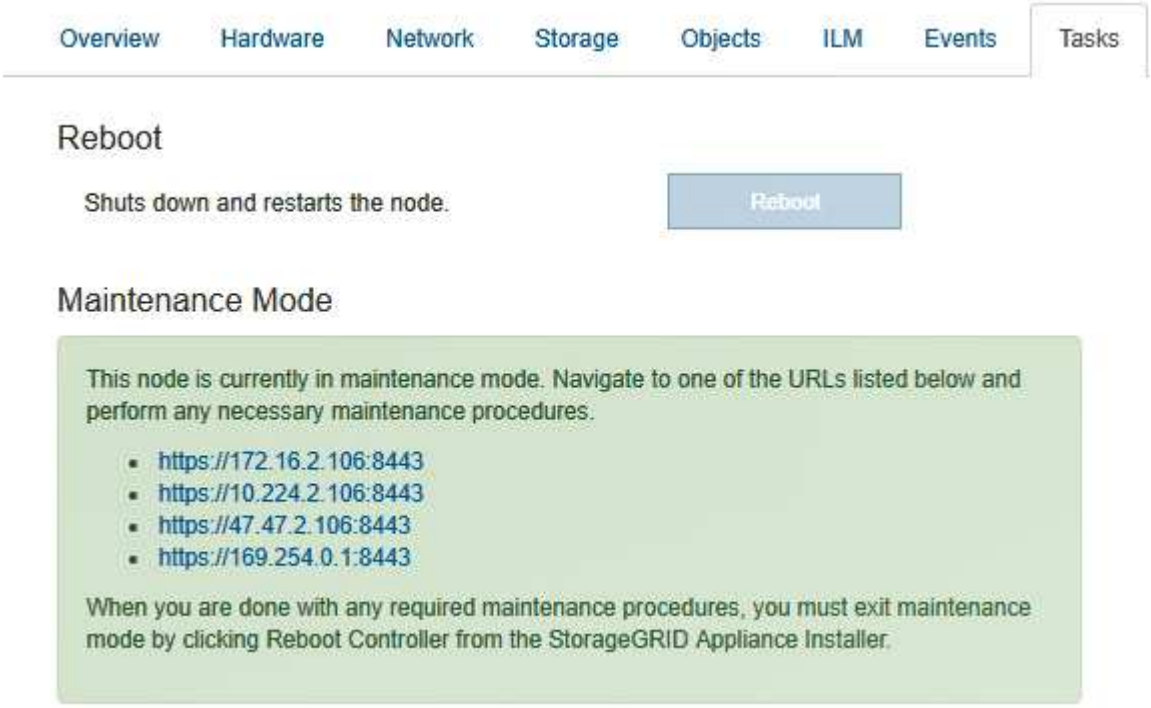



5. Enter the provisioning passphrase, and select **OK**.

   A progress bar and a series of messages, including "Request Sent","Stopping StorageGRID", and "Rebooting", indicate that the appliance is completing the steps for entering maintenance mode.




When the appliance is in maintenance mode, a confirmation message lists the URLs you can use to access the StorageGRID Appliance Installer.



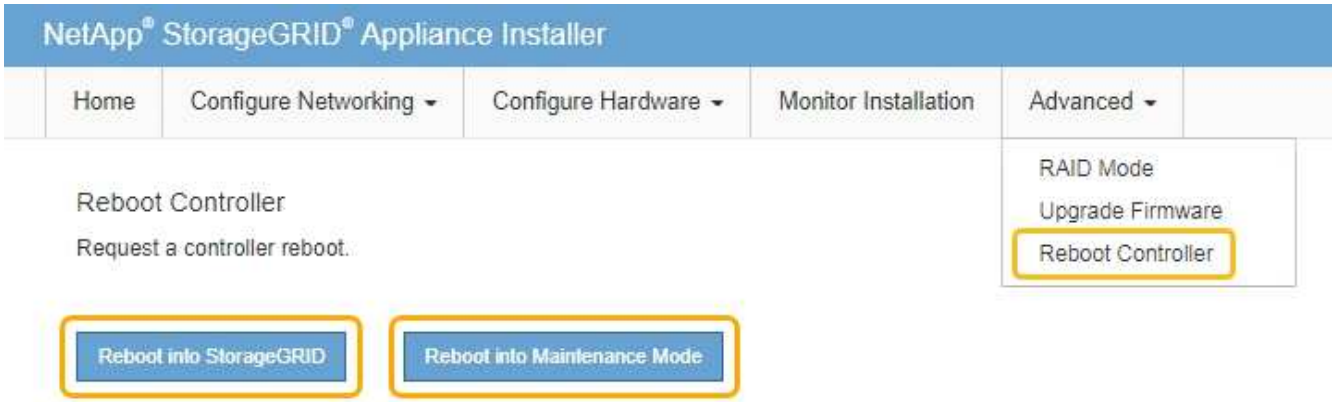

6. To access the StorageGRID Appliance Installer, browse to any of the URLs displayed.

   If possible, use the URL containing the IP address of the appliance's Admin Network port.

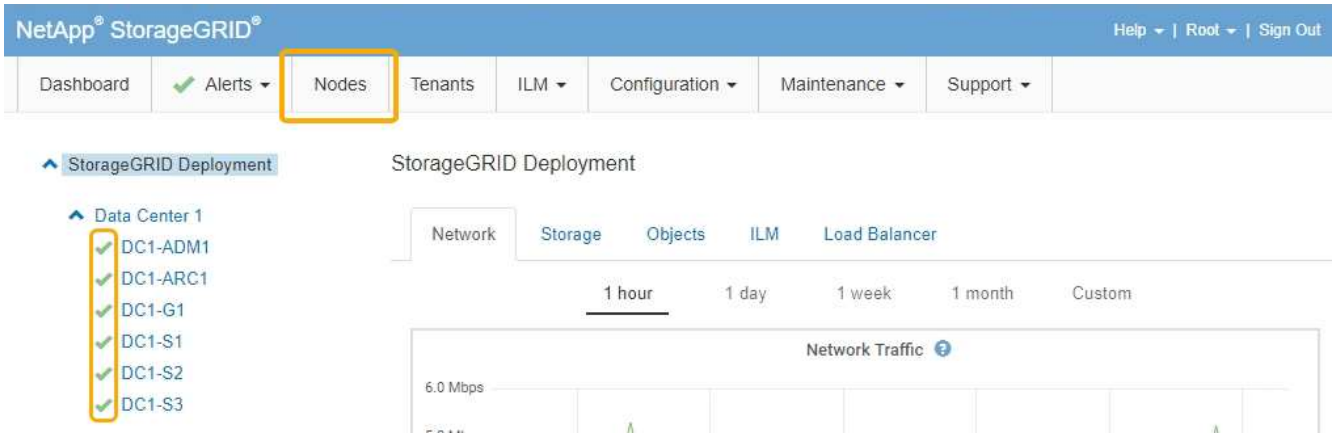
   > Accessing `https://169.254.0.1:8443` requires a direct connection to the local management port.

7. From the StorageGRID Appliance Installer, confirm that the appliance is in maintenance mode.

8. Perform any required maintenance tasks.
9. After completing maintenance tasks, exit maintenance mode and resume normal node operation. From the StorageGRID Appliance Installer, select **Advanced** > **Reboot Controller**, and then select **Reboot into StorageGRID**.




It can take up to 20 minutes for the appliance to reboot and rejoin the grid. To confirm that the reboot is complete and that the node has rejoined the grid, go back to the Grid Manager. The **Nodes** tab should display a normal status ✔ for the appliance node, indicating that no alerts are active and the node is connected to the grid.




### Forcing a service to terminate

If you need to stop a service immediately, you can use the `force-stop` command.

#### What you'll need

You must have the `Passwords.txt` file.

#### Steps

1. Log in to the grid node:

a. Enter the following command: `ssh admin@grid_node_IP`

b. Enter the password listed in the `Passwords.txt` file.

c. Enter the following command to switch to root: `su -`

d. Enter the password listed in the `Passwords.txt` file.

When you are logged in as root, the prompt changes from `$` to `#`.

2. Manually force the service to terminate: `service servicename force-stop`

   For example:

   ```
   service ldr force-stop
   ```

   The system waits 30 seconds before terminating the service.

3. Log out of the command shell: `exit`

#### Starting or restarting a service

## You might need to start a service that has been stopped, or you might need to stop and restart a service.

#### What you'll need

You must have the `Passwords.txt` file.

#### Steps

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Decide which command to issue, based on whether the service is currently running or stopped.

   - If the service is currently stopped, use the `start` command to start the service manually: `service servicename start`

     For example:

     ```
     service ldr start
     ```

   - If the service is currently running, use the `restart` command to stop the service and then restart it: `service servicename restart`

For example:

```
service ldr restart
```

> Using the `restart` command is the same as using the `stop` command followed by the `start` command. You can issue `restart` even if the service is currently stopped.

3. Log out of the command shell: `exit`

### Removing port remaps

If you want to configure an endpoint for the Load Balancer service, and you want to use a port that has already been configured as the Mapped-To Port of a port remap, you must first remove the existing port remap, or the endpoint will not be effective. You must run a script on each Admin Node and Gateway Node that has conflicting remapped ports to remove all of the node's port remaps.

> This procedure removes all port remaps. If you need to keep some of the remaps, contact technical support.

For information about configuring load balancer endpoints, see the instructions for administering StorageGRID.

> If the port remap provides client access, the client should be reconfigured to use a different port configured as an load balancer endpoint if possible, to avoid loss of service, Otherwise, removing the port mapping will result in loss of client access and should be scheduled appropriately.

> This procedure does not work for a StorageGRID system deployed as a container on bare metal hosts. See the instructions for removing port remaps on bare metal hosts.

**Steps**

1. Log in to the node.
   a. Enter the following command: `ssh -p 8022 admin@node_IP`

      Port 8022 is the SSH port of the base OS, while port 22 is the SSH port of the Docker container running StorageGRID.

   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

      When you are logged in as root, the prompt changes from `$` to `#`.

2. Run the following script: `remove-port-remap.sh`
3. Reboot the node.

Follow the instructions for rebooting a grid node.

4. Repeat these steps on each Admin Node and Gateway Node that has conflicting remapped ports.

**Related information**

Administer StorageGRID

Rebooting a grid node

Removing port remaps on bare metal hosts

### Removing port remaps on bare metal hosts

If you want to configure an endpoint for the Load Balancer service, and you want to use a port that has already been configured as the Mapped-To Port of a port remap, you must first remove the existing port remap, or the endpoint will not be effective. If you are running StorageGRID on bare metal hosts, follow this procedure instead of the general procedure for removing port remaps. You must edit the node configuration file for each Admin Node and Gateway Node that has conflicting remapped ports to remove all of the node's port remaps and restart the node.

This procedure removes all port remaps. If you need to keep some of the remaps, contact technical support.

For information about configuring load balancer endpoints, see the instructions for administering StorageGRID.

This procedure can result in temporary loss of service as nodes are restarted.

**Steps**

1. Log in to the host supporting the node. Log in as root or with an account that has sudo permission.
2. Run the following command to temporarily disable the node: `sudo storagegrid node stop node-name`
3. Using a text editor such as vim or pico, edit the node configuration file for the node.

   The node configuration file can be found at `/etc/storagegrid/nodes/node-name.conf`.

4. Locate the section of the node configuration file that contains the port remaps.

   See the last two lines in the following example.

```
ADMIN_NETWORK_CONFIG = STATIC
ADMIN_NETWORK_ESL = 10.0.0.0/8, 172.19.0.0/16, 172.21.0.0/16
ADMIN_NETWORK_GATEWAY = 10.224.0.1
ADMIN_NETWORK_IP = 10.224.5.140
ADMIN_NETWORK_MASK = 255.255.248.0
ADMIN_NETWORK_MTU = 1400
ADMIN_NETWORK_TARGET = eth1
ADMIN_NETWORK_TARGET_TYPE = Interface
BLOCK_DEVICE_VAR_LOCAL = /dev/sda2
CLIENT_NETWORK_CONFIG = STATIC
CLIENT_NETWORK_GATEWAY = 47.47.0.1
CLIENT_NETWORK_IP = 47.47.5.140
CLIENT_NETWORK_MASK = 255.255.248.0
CLIENT_NETWORK_MTU = 1400
CLIENT_NETWORK_TARGET = eth2
CLIENT_NETWORK_TARGET_TYPE = Interface
GRID_NETWORK_CONFIG = STATIC
GRID_NETWORK_GATEWAY = 192.168.0.1
GRID_NETWORK_IP = 192.168.5.140
GRID_NETWORK_MASK = 255.255.248.0
GRID_NETWORK_MTU = 1400
GRID_NETWORK_TARGET = eth0
GRID_NETWORK_TARGET_TYPE = Interface
NODE_TYPE = VM_API_Gateway
<strong>PORT_REMAP = client/tcp/8082/443</strong>
<strong>PORT_REMAP_INBOUND = client/tcp/8082/443</strong>
```

5. Edit the PORT_REMAP and PORT_REMAP_INBOUND entries to remove port remaps.

```
PORT_REMAP =
PORT_REMAP_INBOUND =
```

6. Run the following command to validate your changes to the node configuration file for the node: `sudo storagegrid node validate node-name`

   Address any errors or warnings before proceeding to the next step.

7. Run the following command to restart the node without port remaps: `sudo storagegrid node start node-name`
8. Log in to the node as admin using the password listed in the `Passwords.txt` file.
9. Verify that the services start correctly.
   a. View a listing of the statuses of all services on the server:`sudo storagegrid-status`

      The status is updated automatically.

b. Wait until all services have a status of either Running or Verified.

c. Exit the status screen: `Ctrl+C`

10. Repeat these steps on each Admin Node and Gateway Node that has conflicting remapped ports.

### Rebooting a grid node

You can reboot a grid node from the Grid Manager or from the node's command shell.

#### About this task

When you reboot a grid node, the node shuts down and restarts. All services are restarted automatically.

If you plan to reboot Storage Nodes, note the following:

- If an ILM rule specifies an ingest behavior of Dual commit or the rule specifies Balanced and it is not possible to immediately create all required copies, StorageGRID immediately commits any newly ingested objects to two Storage Nodes on the same site and evaluates ILM later. If you want to reboot two or more Storage Nodes on a given site, you might not be able to access these objects for the duration of the reboot.
- To ensure you can access all objects while a Storage Node is rebooting, stop ingesting objects at a site for approximately one hour before rebooting the node.

#### Related information

Administer StorageGRID

#### Choices

- Rebooting a grid node from the Grid Manager
- Rebooting a grid node from the command shell

##### Rebooting a grid node from the Grid Manager

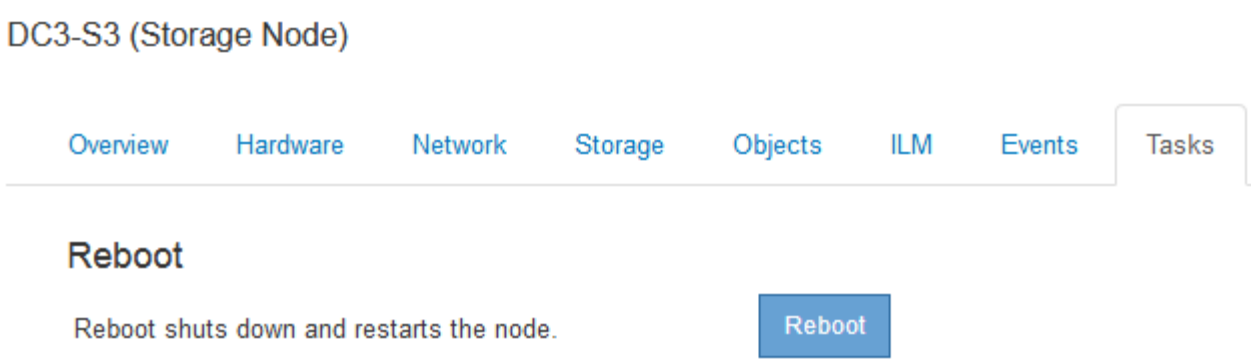
Rebooting a grid node from the Grid Manager issues the `reboot` command on the target node.

#### What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have the Maintenance or Root Access permission.
- You must have the provisioning passphrase.

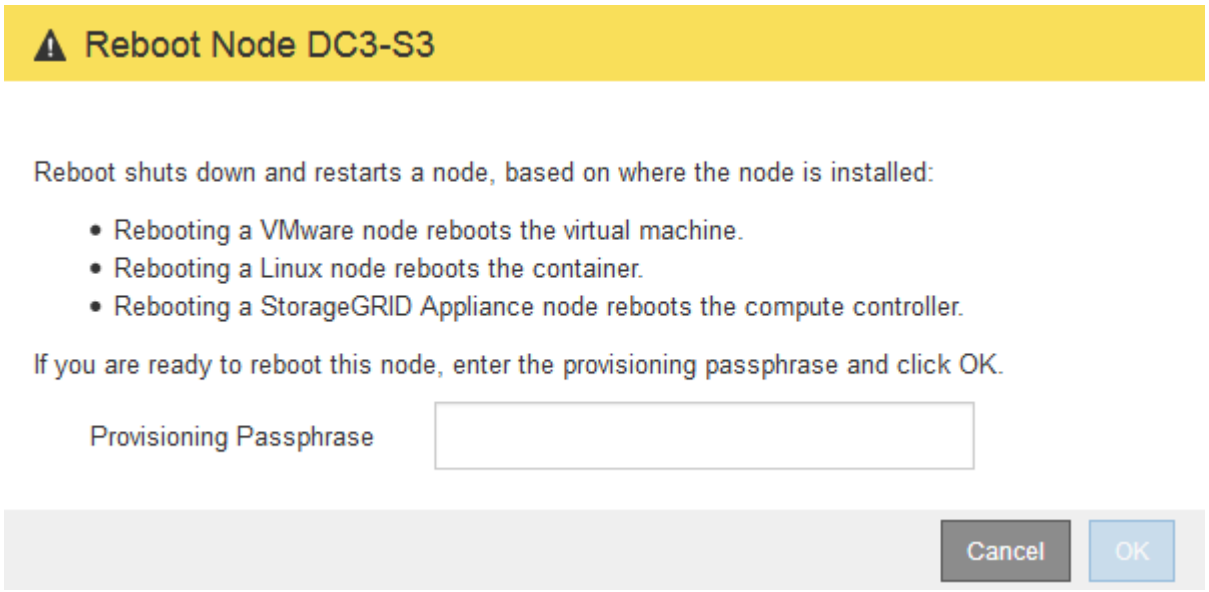
#### Steps

1. Select **Nodes**.
2. Select the grid node you want to reboot.
3. Select the **Tasks** tab.




4. Click **Reboot**.

   A confirmation dialog box appears.




> If you are rebooting the primary Admin Node, the confirmation dialog box reminds you that your browser's connection to the Grid Manager will be lost temporarily when services are stopped.

5. Enter the provisioning passphrase, and click **OK**.
6. Wait for the node to reboot.

   It might take some time for services to shut down.

   When the node is rebooting, the gray icon (Administratively Down) appears on the left side of the Nodes page. When all services have started again, the icon changes back to its original color.

#### Rebooting a grid node from the command shell

If you need to monitor the reboot operation more closely or if you are unable to access the Grid Manager, you can log into the grid node and run the Server Manager reboot command from the command shell.

#### What you'll need

- You must have the `Passwords.txt` file.

**Steps**

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Optionally, stop services: `service servermanager stop`

   Stopping services is an optional, but recommended step. Services can take up to 15 minutes to shut down, and you might want to log in to the system remotely to monitor the shutdown process before you reboot the node in the next step.

3. Reboot the grid node: `reboot`

4. Log out of the command shell: `exit`

### Shutting down a grid node

You can shut down a grid node from the node's command shell.

**What you'll need**

- You must have the `Passwords.txt` file.

**About this task**

Before performing this procedure, review these considerations:

- In general, you should not shut down more than one node at a time to avoid disruptions.
- Do not shut down a node during a maintenance procedure unless explicitly instructed to do so by the documentation or by technical support.
- The shutdown process is based on where the node is installed, as follows:
  - Shutting down a VMware node shuts down the virtual machine.
  - Shutting down a Linux node shuts down the container.
  - Shutting down a StorageGRID appliance node shuts down the compute controller.
- If you plan to shut down Storage Nodes, note the following:
  - If an ILM rule specifies an ingest behavior of Dual commit or the rule specifies Balanced and it is not possible to immediately create all required copies, StorageGRID immediately commits any newly ingested objects to two Storage Nodes on the same site and evaluates ILM later. If you want to shut down two or more Storage Nodes on a given site, you might not be able to access these objects for the duration of the shutdown.
  - To ensure you can access all objects when a Storage Node is shut down, stop ingesting objects at a site for approximately one hour before shutting down the node.

**Steps**

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Stop all services: `service servermanager stop`

   Services can take up to 15 minutes to shut down, and you might want to log in to the system remotely to monitor the shutdown process.

3. Log out of the command shell: `exit`

After being shut down you can power off the grid node.

Powering down a host

**Related information**

Administer StorageGRID

**Powering down a host**

Before you power down a host, you must stop services on all grid nodes on that host.

**Steps**

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Stop all services running on the node: `service servermanager stop`

   Services can take up to 15 minutes to shut down, and you might want to log in to the system remotely to monitor the shutdown process.

3. Repeat steps 1 and 2 for each node on the host.
4. If you have a Linux host:

   a. Log in to the host operating system.

   b. Stop the node: `storagegrid node stop`

   c. Shut down the host operating system.

5. If the node is running on a VMware virtual machine or it is an appliance node, issue the shutdown command: `shutdown -h now`

   Perform this step regardless of the outcome of the `service servermanager stop` command.

   > After you issue the `shutdown -h now` command on an appliance node, you must power cycle the appliance to restart the node.

   For the appliance, this command shuts down the controller, but the appliance is still powered on. You must complete the next step.

6. If you are powering down an appliance node:
   - For the SG100 or SG1000 services appliance
     i. Turn off the power to the appliance.
     ii. Wait for the blue power LED to turn off.
   - For the SG6000 appliance
     i. Wait for the green Cache Active LED on the back of the storage controller to turn off.

        This LED is on when cached data needs to be written to the drives. You must wait for this LED to turn off before you turn off power.

     ii. Turn off the power to the appliance, and wait for the blue power LED to turn off.
   - For the SG5700 appliance
     i. Wait for the green Cache Active LED on the back of the storage controller to turn off.

        This LED is on when cached data needs to be written to the drives. You must wait for this LED to turn off before you turn off power.

     ii. Turn off the power to the appliance, and wait for all LED and seven-segment display activity to stop.
7. Log out of the command shell: `exit`

**Related information**

SG100 & SG1000 services appliances

SG6000 storage appliances

SG5700 storage appliances

### Powering off and on all nodes in the grid

You might need to shut down your entire StorageGRID system, for example, if you are moving a data center. These steps provide a high-level overview of the recommended sequence for performing a controlled shutdown and startup.

When you power off all nodes in a site or grid, you will not be able to access ingested objects while the Storage Nodes are offline.

#### Stopping services and shutting down grid nodes

# Before you can power off a StorageGRID system, you must stop all services running on each grid node, and then shut down all VMware virtual machines, Docker containers, and StorageGRID appliances.

### About this task

If possible, you should stop services on the grid nodes in this order:

- Stop services on Gateway Nodes first.
- Stop services on the primary Admin Node last.

This approach allows you to use the primary Admin Node to monitor the status of the other grid nodes for as long as possible.

> If a single host includes more than one grid node, do not shut down the host until you have stopped all of the nodes on that host. If the host includes the primary Admin Node, shut down that host last.

> If required, you can migrate nodes from one Linux host to another to perform host maintenance without impacting the functionality or availability of your grid.

Linux: Migrating a grid node to a new host

### Steps

1. Stop all client applications from accessing the grid.
2. Log in to each Gateway Node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
3. Stop all services running on the node: `service servermanager stop`

   Services can take up to 15 minutes to shut down, and you might want to log in to the system remotely to monitor the shutdown process.
4. Repeat the previous two steps to stop the services on all Storage Nodes, Archive Nodes, and non-primary Admin Nodes.

   You can stop the services on these nodes in any order.

   > If you issue the `service servermanager stop` command to stop the services on an appliance Storage Node, you must power cycle the appliance to restart the node.
5. For the primary Admin Node, repeat the steps for logging into the node and stopping all services on the

node.

6. For nodes that are running on Linux hosts:
   a. Log in to the host operating system.
   b. Stop the node: `storagegrid node stop`
   c. Shut down the host operating system.
7. For nodes that are running on VMware virtual machines and for appliance Storage Nodes, issue the shutdown command: `shutdown -h now`

   Perform this step regardless of the outcome of the `service servermanager stop` command.

   For the appliance, this command shuts down the compute controller, but the appliance is still powered on. You must complete the next step.

8. If you have appliance nodes:
   - For the SG100 or SG1000 services appliance
     i. Turn off the power to the appliance.
     ii. Wait for the blue power LED to turn off.
   - For the SG6000 appliance
     i. Wait for the green Cache Active LED on the back of the storage controller to turn off.

        This LED is on when cached data needs to be written to the drives. You must wait for this LED to turn off before you turn off power.

     ii. Turn off the power to the appliance, and wait for the blue power LED to turn off.
   - For the SG5700 appliance
     i. Wait for the green Cache Active LED on the back of the storage controller to turn off.

        This LED is on when cached data needs to be written to the drives. You must wait for this LED to turn off before you turn off power.

     ii. Turn off the power to the appliance, and wait for all LED and seven-segment display activity to stop.
9. If required, log out of the command shell: `exit`

   The StorageGRID grid has now been shut down.

**Related information**

SG100 & SG1000 services appliances

SG6000 storage appliances

SG5700 storage appliances

#### Starting up the grid nodes

Follow this sequence to start up the grid nodes after a complete shutdown.

If the entire grid has been shut down for more than 15 days, you must contact technical support before starting up any grid nodes. Do not attempt the recovery procedures that rebuild Cassandra data. Doing so might result in data loss.

**About this task**

If possible, you should power on the grid nodes in this order:

- Apply power to Admin Nodes first.
- Apply power to Gateway Nodes last.

If a host includes multiple grid nodes, the nodes will come back online automatically when you power on the host.

**Steps**

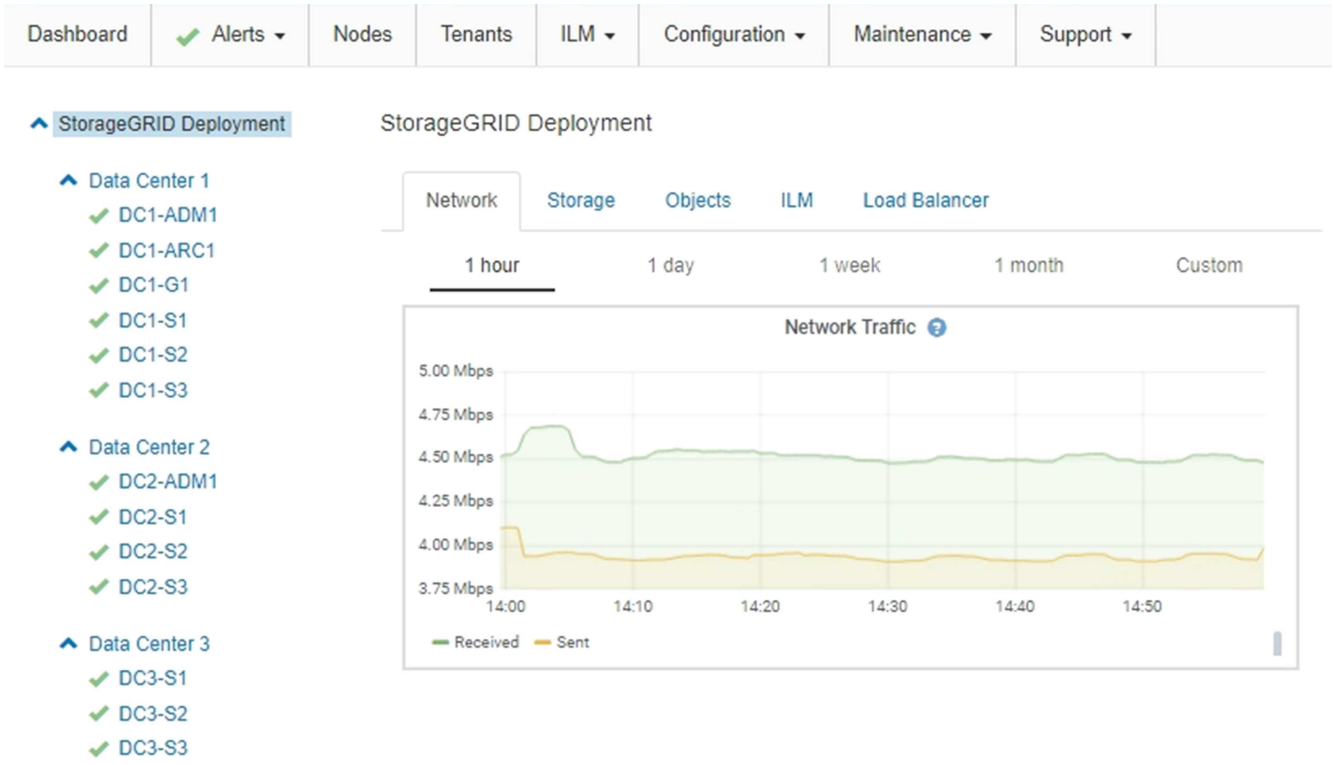
1. Power on the hosts for the primary Admin Node and any non-primary Admin Nodes.

   You will not be able to log in to the Admin Nodes until the Storage Nodes have been restarted.

2. Power on the hosts for all Archive Nodes and Storage Nodes.

   You can power on these nodes in any order.

3. Power on the hosts for all Gateway Nodes.
4. Sign into the Grid Manager.
5. Click **Nodes**, and monitor the status of the grid nodes. Verify that all nodes return to “green” status.



### Using a DoNotStart file

If you are performing various maintenance or configuration procedures under the direction of technical support, you might be asked to use a DoNotStart file to prevent services from starting when Server Manager is started or restarted.

> You should add or remove a DoNotStart file only if technical support has directed you to do so.

To prevent a service from starting, place a DoNotStart file in the directory of the service you want to prevent from starting. At start-up, Server Manager looks for the DoNotStart file. If the file is present, the service (and any services dependent on it) is prevented from starting. When the DoNotStart file is removed, the previously stopped service will start on the next start or restart of Server Manager. Services are not automatically started when the DoNotStart file is removed.

The most efficient way to prevent all services from restarting is to prevent the NTP service from starting. All services are dependent on the NTP service and cannot run if the NTP service is not running.

#### Adding a DoNotStart file for a service

You can prevent an individual service from starting by adding a DoNotStart file to that service's directory on a grid node.

**What you'll need**

You must have the `Passwords.txt` file.

**Steps**

1. Log in to the grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Add a DoNotStart file: `touch /etc/sv/service/DoNotStart`

   where `service` is the name of the service to be prevented from starting. For example,

   ```
   touch /etc/sv/ldr/DoNotStart
   ```

   A DoNotStart file is created. No file content is needed.

   When Server Manager or the grid node is restarted, Server Manager restarts, but the service does not.

3. Log out of the command shell: `exit`

#### Removing a DoNotStart file for a service

When you remove a DoNotStart file that is preventing a service from starting, you must start that service.

**What you'll need**

You must have the `Passwords.txt` file.

**Steps**

1. Log in to the grid node:
   a. Enter the following command: `ssh admin@grid_node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
2. Remove the DoNotStart file from the service directory: `rm /etc/sv/service/DoNotStart`

   where `service` is the name of the service. For example,

   ```
   rm /etc/sv/ldr/DoNotStart
   ```
3. Start the service: `service servicename start`
4. Log out of the command shell: `exit`

### Troubleshooting Server Manager

Technical support might direct you to troubleshooting tasks to determine the source of Server Manager-related problems.

#### Accessing the Server Manager log file

If a problem arises when using Server Manager, check its log file.

Error messages related to Server Manager are captured in the Server Manager log file, which is located at: `/var/local/log/servermanager.log`

Check this file for error messages regarding failures. Escalate the issue to technical support if required. You might be asked to forward log files to technical support.

#### Service with an error state

If you detect that a service has entered an error state, attempt to restart the service.

**What you'll need**

You must have the `Passwords.txt` file.

**About this task**

Server Manager monitors services and restarts any that have stopped unexpectedly. If a service fails, Server Manager attempts to restart it. If there are three failed attempts to start a service within five minutes, the service enters an error state. Server Manager does not attempt another restart.

**Steps**

1. Log in to the grid node:

   a. Enter the following command: `ssh admin@grid_node_IP`

   b. Enter the password listed in the `Passwords.txt` file.

   c. Enter the following command to switch to root: `su -`

   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.

2. Confirm the error state of the service: `service servicename status`

   For example:

   ```
   service ldr status
   ```

   If the service is in an error state, the following message is returned: `servicename in error state`. For example:

   ```
   ldr in error state
   ```

   > If the service status is `disabled`, see the instructions for removing a DoNotStart file for a service.

3. Attempt to remove the error state by restarting the service: `service servicename restart`

   If the service fails to restart, contact technical support.

4. Log out of the command shell: `exit`

**Related information**

Removing a DoNotStart file for a service

## Appliance node cloning

You can clone an appliance node in StorageGRID to use an appliance of newer design or increased capabilities. Cloning transfers all information on the existing node to the new appliance, provides a hardware-upgrade process that is easy to perform, and provides an alternative to decommissioning and expansion for replacing appliances.

## How appliance node cloning works

Appliance node cloning lets you easily replace an existing appliance node (source) in your grid with a compatible appliance (target) that is part of the same logical StorageGRID site. The process transfers all data to the new appliance, placing it in service to replace the old appliance node and leaving the old appliance in a pre-install state.

### Why clone an appliance node?

You can clone an appliance node if you need to:

- Replace appliances that are reaching end-of-life.
- Upgrade existing nodes to take advantage of improved appliance technology.
- Increase grid storage capacity without changing the number of Storage Nodes in your StorageGRID system.
- Improve storage efficiency, such as by changing the RAID mode from DDP-8 to DDP-16, or to RAID-6.
- Efficiently implement node encryption to allow the use of external key management servers (KMS).

### Which StorageGRID network is used?

Cloning transfers data from the source node directly to the target appliance over any of the three StorageGRID networks. The Grid Network is typically used, but you can also use the Admin Network or the Client Network if the source appliance is connected to these networks. Choose the network to use for cloning traffic that provides the best data-transfer performance without degrading StorageGRID network performance or data availability.

When you install the replacement appliance, you must specify temporary IP addresses for StorageGRID connection and data transfer. Since the replacement appliance will be part of the same networks as the appliance node it replaces, you must specify temporary IP addresses for each of these networks on the replacement appliance.

### Target appliance compatibility

Replacement appliances must be the same type as the source node they are replacing and both must be part of the same logical StorageGRID site.

- A replacement services appliance can be different than the Admin Node or Gateway Node it is replacing.
  - You can clone an SG100 source node appliance to an SG1000 services target appliance to give the Admin Node or Gateway Node greater capability.
  - You can clone an SG1000 source node appliance to an SG100 services target appliance to redeploy the SG1000 for a more demanding application.

    For example, if an SG1000 source node appliance is being used as an Admin Node and you want to use it as a dedicated load-balancing node.

  - Replacing an SG1000 source node appliance with an SG100 services target appliance reduces the maximum speed of the network ports from 100-GbE to 25-GbE.
  - The SG100 and SG1000 appliances have different network connectors. Changing the appliance type might require replacing the cables or SFP modules.

- A replacement storage appliance must have equal or greater capacity than the Storage Node it is replacing.
  - If the target storage appliance has the same number of drives as the source node, the drives in the target appliance must have the same capacity (in TB) or larger.
  - If the number of standard drives installed in a target storage appliance is less than the number of drives in the source node, due to installation of solid-state drives (SSDs), the overall storage capacity of the standard drives in the target appliance (in TB) must meet or exceed the total functional drive capacity of all drives in the source Storage Node.

    For example, when cloning an SG5660 source Storage Node appliance with 60 drives to an SG6060 target appliance with 58 standard drives, larger drives should be installed in the SG6060 target appliance before cloning to maintain storage capacity. (The two drive slots containing SSDs in the target appliance are not included in the total appliance-storage capacity.)

    However, if a 60-drive SG5660 source node appliance is configured with SANtricity Dynamic Disk Pools DDP-8, configuring a 58-drive same-size-drive SG6060 target appliance with DDP-16 might make the SG6060 appliance a valid clone target due to its improved storage efficiency.

    You can view information about the current RAID mode of the source appliance node on the **Nodes** page in Grid Manager. Select the **Storage** tab for the appliance.

#### What information is not cloned?

The following appliance configurations do not transfer to the replacement appliance during cloning. You must configure them during initial set up of the replacement appliance.

- BMC interface
- Network links
- Node encryption status
- SANtricity System Manager (for Storage Nodes)
- RAID mode (for Storage Nodes)

#### What issues prevent cloning?

If any of the following issues are encountered while cloning, the cloning process halts and an error message is generated:

- Wrong network configuration
- Lack of connectivity between the source and target appliances
- Source and target appliance incompatibility
- For Storage Nodes, a replacement appliance of insufficient capacity

You must resolve each issue for cloning to continue.

### Considerations and requirements for appliance node cloning

Before cloning an appliance node, you must understand the considerations and requirements.

#### Hardware requirements for the replacement appliance

Ensure that the replacement appliance meets the following criteria:

- The source node (appliance being replaced) and the target (new) appliance must be the same type of appliance:
  - You can only clone an Admin Node appliance or a Gateway Node appliance to a new services appliance.
  - You can only clone a Storage Node appliance to a new storage appliance.
- For Admin Node or Gateway Node appliances, the source node appliance and the target appliance do not need to be the same type of appliance; however, changing the appliance type might require replacing the cables or SFP modules.

  For example, you can replace a SG1000 node appliance with a SG100 or replace a SG100 appliance with a SG1000 appliance.

- For Storage Node appliances, the source node appliance and the target appliance do not need to be the same type of appliance; however, the target appliance must have the same or greater storage capacity as the source appliance.

  For example, you can replace a SG5600 node appliance with a SG5700 or a SG6000 appliance.

Contact your StorageGRID sales representative for help choosing compatible replacement appliances to clone specific appliance nodes in your StorageGRID installation.

#### Preparing to clone an appliance node

You must have the following information before you clone an appliance node:

- Obtain a temporary IP address for the Grid Network from your network administrator for use with the target appliance during initial installation. If the source node belongs to an Admin Network or Client Network, obtain temporary IP addresses for these networks.

  Temporary IP addresses are normally on the same subnet as the source node appliance being cloned and are not needed after cloning completes. The source and target appliances must both connect to the primary Admin Node of your StorageGRID to establish a cloning connection.

- Determine which network to use for cloning data-transfer traffic that provides the best data-transfer performance without degrading StorageGRID network performance or data availability.

  > Using the 1-GbE Admin Network for clone data transfer results in slower cloning.

- Determine if node encryption using a key management server (KMS) will be used on the target appliance, so that you can enable node encryption during initial target appliance installation before cloning. You can check if node encryption is enabled on the source appliance node as described in appliance installation.

  The source node and target appliance can have different node-encryption settings. Data decryption and encryption is performed automatically during data transfer and when the target node restarts and joins the grid.

  - SG100 & SG1000 services appliances
  - SG5600 storage appliances

    - SG5700 storage appliances
    - SG6000 storage appliances
- Determine if the RAID mode on the target appliance should be changed from its default setting, so you can specify this information during initial target appliance installation before cloning. You can view information about the current RAID mode of the source appliance node on the **Nodes** page in Grid Manager. Select the **Storage** tab for the appliance.

  The source node and target appliance can have different RAID settings.

- Plan for sufficient time to complete the node cloning process. Several days might be required to transfer data from an operational Storage Node to a target appliance. Schedule cloning at a time that minimizes the impact to your business.
- You should only clone one appliance node at a time. Cloning can prevent you from performing other StorageGRID maintenance functions at the same time.
- After you have cloned an appliance node, you can use the source appliance that was returned to a pre-install state as the target to clone another compatible node appliance.

### Appliance node cloning procedure

The cloning process might take several days to transfer data between the source node (appliance being replaced) and the target (new) appliance.

#### What you'll need

- You have installed the compatible target appliance into a cabinet or rack, connected all cables, and applied power.
- You have verified that the StorageGRID Appliance Installer version on the replacement appliance matches the software version of your StorageGRID system, upgrading the StorageGRID Appliance Installer firmware, if necessary.
- You have configured the target appliance, including configuring StorageGRID connections, SANtricity System Manager (storage appliances only), and the BMC interface.
    - When configuring StorageGRID connections, use the temporary IP addresses.
    - When configuring network links, use the final link configuration.

> Leave the StorageGRID Appliance Installer open after you complete initial target appliance configuration. You will return to the target appliance's installer page after you start the node cloning process.

- You have optionally enabled node encryption for the target appliance.
- You have optionally set the RAID mode for the target appliance (storage appliances only).
- Considerations and requirements for appliance node cloning

  SG100 & SG1000 services appliances

  SG5600 storage appliances

  SG5700 storage appliances

  SG6000 storage appliances

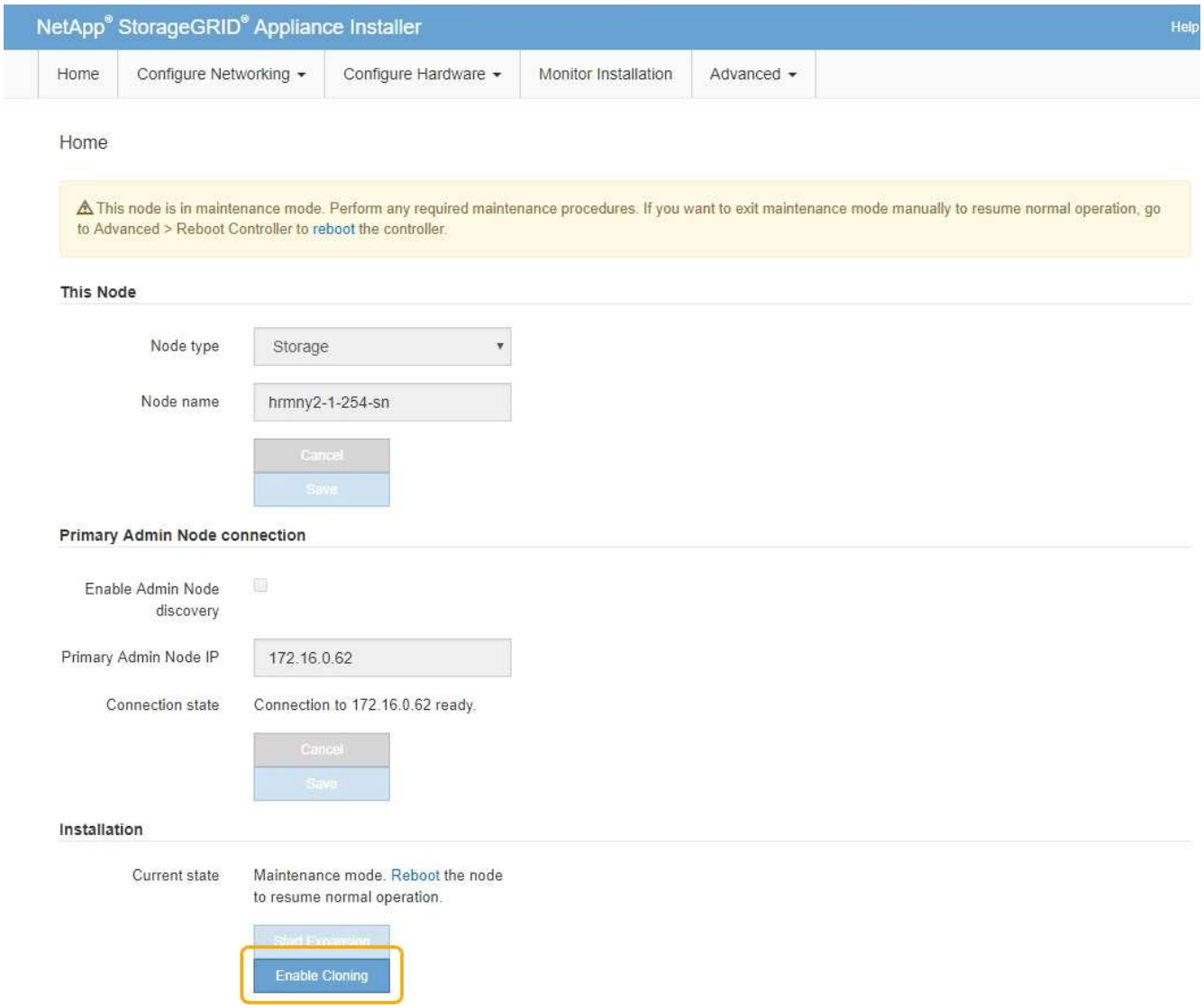
You should clone only one appliance node at a time to maintain StorageGRID network performance and data availability.

**Steps**

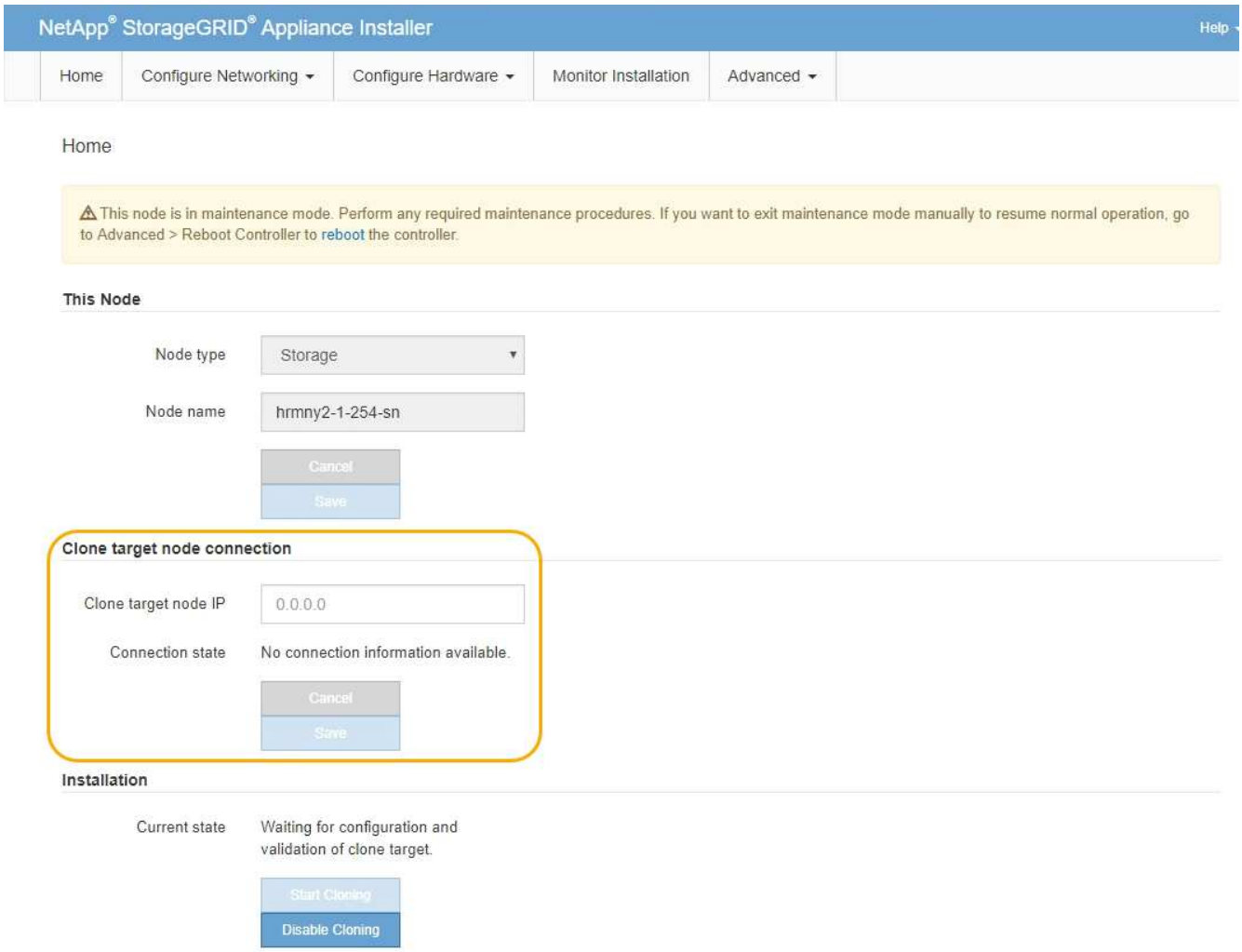
1. Place the source node you are cloning into maintenance mode.

   Placing an appliance into maintenance mode

2. From the StorageGRID Appliance Installer on the source node, in the Installation section of the Home page, select **Enable Cloning**.




   The Primary Admin Node connection section is replaced with the Clone target node connection section.

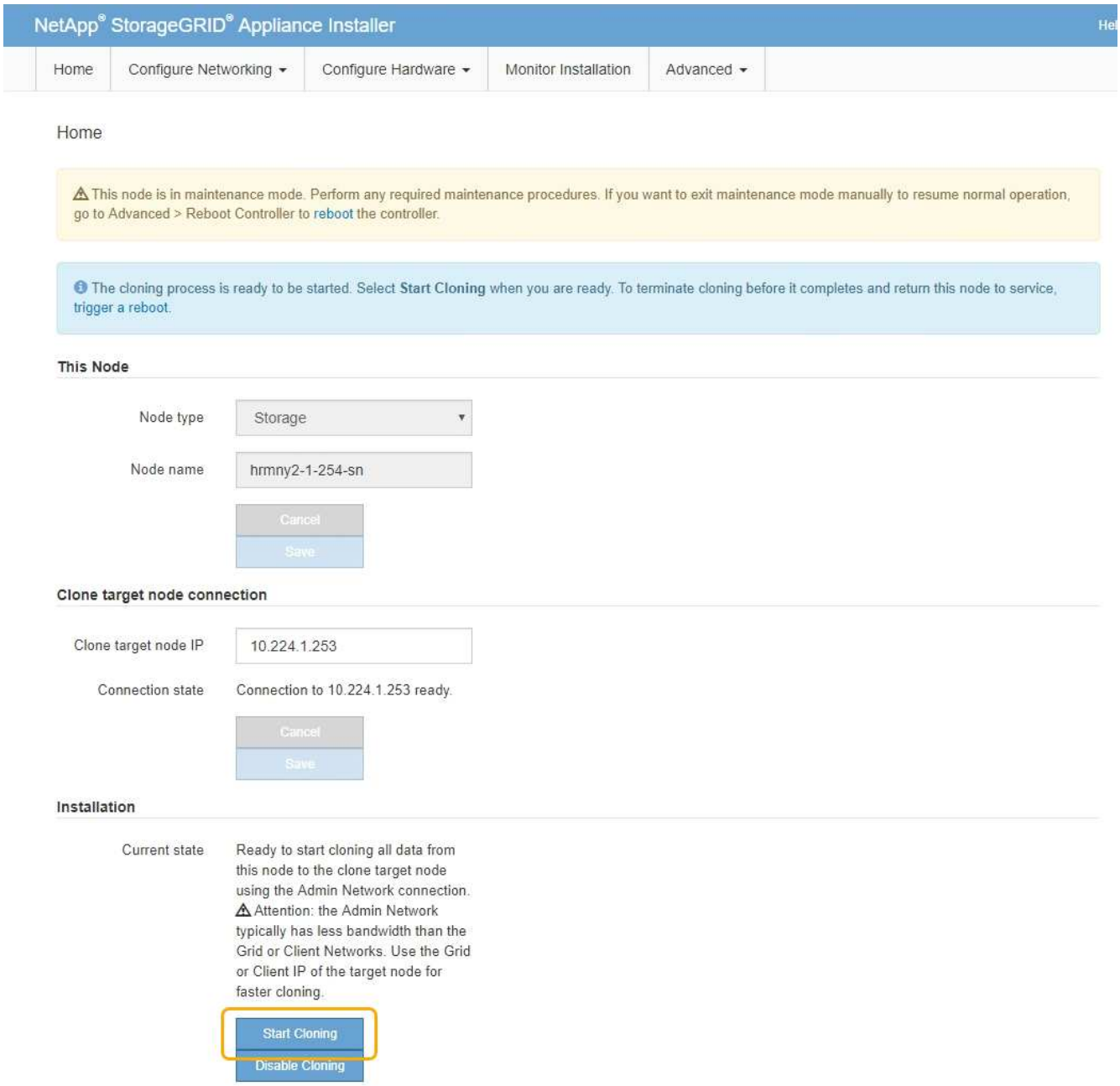



3. For **Clone target node IP**, enter the temporary IP address assigned to the target node for the network to use for clone data-transfer traffic, and then select **Save**.

   Typically, you enter the IP address for the Grid Network, but if you need to use a different network for clone data-transfer traffic, enter the IP address of the target node on that network.

   > Using the 1-GbE Admin Network for clone data transfer results in slower cloning.

   After the target appliance is configured and validated, in the Installation section, **Start Cloning** is enabled on the source node.



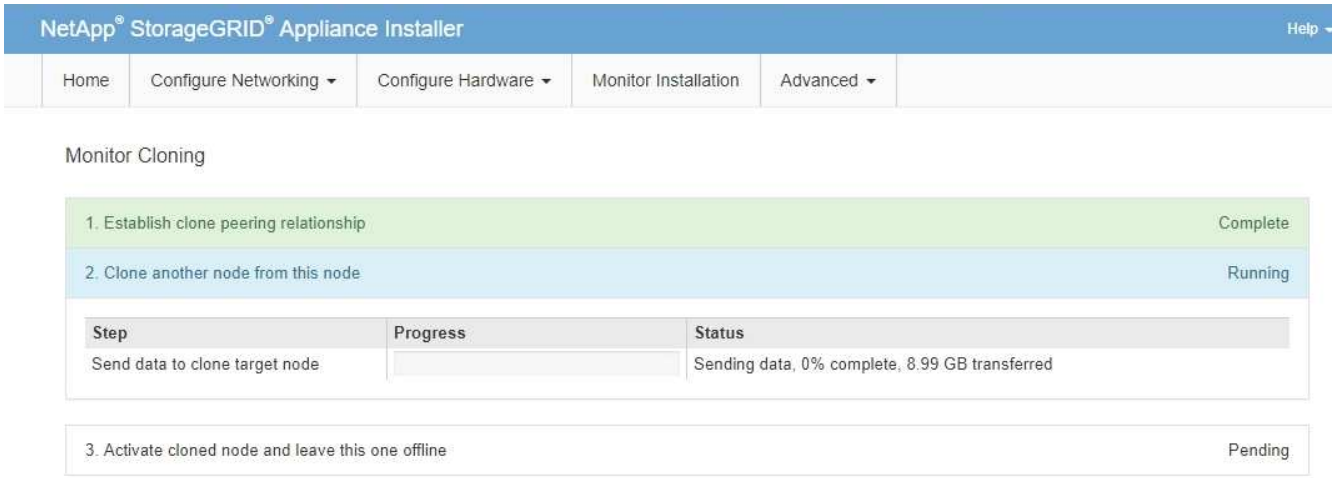

If issues exist that prevent cloning, **Start Cloning** is not enabled and issues that you must resolve are listed as the **Connection state**. These issues are listed on the StorageGRID Appliance Installer Home page of both the source node and the target appliance. Only one issue displays at a time and the state automatically updates as conditions change. Resolve all cloning issues to enable **Start Cloning**.

When **Start Cloning** is enabled, the **Current state** indicates the StorageGRID network that was selected for cloning traffic, along with information about using that network connection.

Considerations and requirements for appliance node cloning

4. Select **Start Cloning** on the source node.
5. Monitor the cloning progress using the StorageGRID Appliance Installer on either the source or target node.

   The StorageGRID Appliance Installer on both the source and target nodes indicates the same status.




The Monitor Cloning page provides detailed progress for each stage of the cloning process:

- **Establish clone peering relationship** shows the progress of cloning set up and configuration.
- **Clone another node from this node** shows the progress of data transfer. (This part of the cloning process can take several days to complete.)
- **Activate cloned node and leave this one offline** shows the progress of transferring control to the target node and placing the source node in a pre-install state, after data transfer is complete.

6. If you need to terminate the cloning process and return the source node to service before cloning is complete, on the source node go to the StorageGRID Appliance Installer Home page and select **Advanced** > **Reboot Controller**, and then select **Reboot into StorageGRID**.

   If the cloning process is terminated:

   - The source node exits maintenance mode and rejoins StorageGRID.
   - The target node remains in the pre-install state. To restart cloning the source node, start the cloning process again from step 1.

When cloning successfully completes:

- The source and target nodes swap IP addresses:
  - The target node now uses the IP addresses originally assigned to the source node for the Grid, Admin, and Client Networks.
  - The source node now uses the temporary IP address initially assigned to the target node.
- The target node exits maintenance mode and joins StorageGRID, replacing the source node.
- The source appliance is in a pre-installed state, as if you had prepared it for reinstallation.

Preparing an appliance for reinstallation (platform replacement only)

> If the appliance does not rejoin the grid, go to the StorageGRID Appliance Installer Home page for the source node, select **Advanced** > **Reboot Controller**, and then select **Reboot into Maintenance Mode**. After the source node reboots in maintenance mode, repeat the node cloning procedure.

User data remains on the source appliance as a recovery option if an unexpected issue occurs with the target node. After the target node has successfully rejoined StorageGRID, user data on the source appliance is

outdated and is no longer needed. If desired, ask StorageGRID Support to clear the source appliance to destroy this data.

You can:

- Use the source appliance as a target for additional cloning operations: no additional configuration is required. This appliance already has the temporary IP address assigned that were originally specified for the first clone target.
- Install and set up the source appliance as a new appliance node.
- Discard the source appliance if it is no longer of use with StorageGRID.

**Copyright information**

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

**Trademark information**

NETAPP, the NETAPP logo, and the marks listed at http://www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.