



Configure client connections

StorageGRID 11.7

NetApp
March 05, 2024

Table of Contents

- Configure client connections 1
 - Configure S3 and Swift client connections: Overview 1
 - Use S3 setup wizard 4
 - Manage HA groups 14
 - Manage load balancing 24
 - Configure S3 endpoint domain names 36
 - Summary: IP addresses and ports for client connections 38

Configure client connections

Configure S3 and Swift client connections: Overview

As a grid administrator, you manage the configuration options that control how S3 and Swift client applications connect to your StorageGRID system to store and retrieve data.

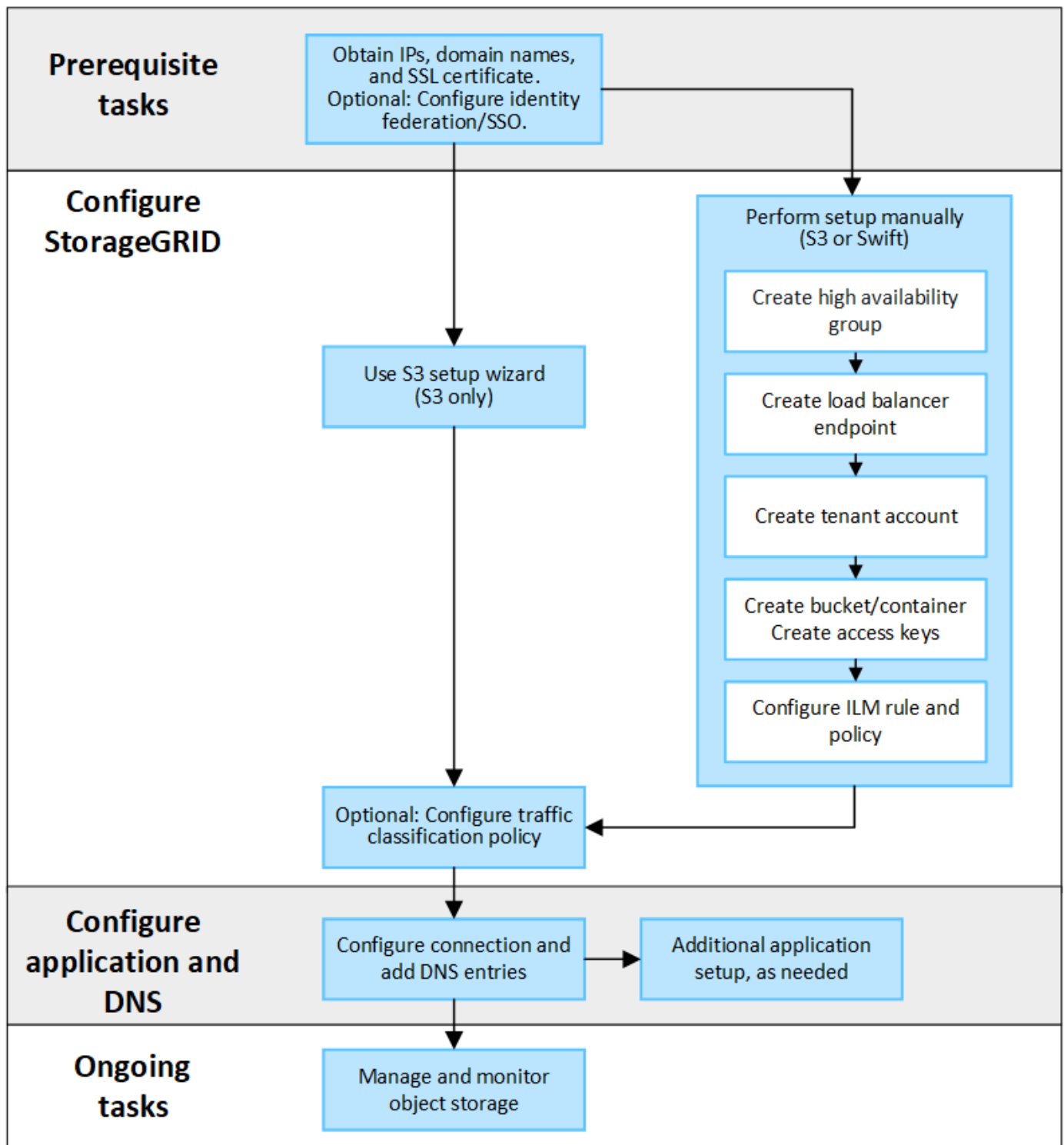


Support for Swift client applications has been deprecated and will be removed in a future release.

Configuration workflow

As shown in the workflow diagram, there are four primary steps for connecting StorageGRID to any S3 or Swift application:

1. Perform prerequisite tasks in StorageGRID, based on how the client application will connect to StorageGRID.
2. Use StorageGRID to obtain the values the application needs to connect to the grid. You can either use the S3 setup wizard or configure each StorageGRID entity manually.
3. Use the S3 or Swift application to complete the connection to StorageGRID. Create DNS entries to associate IP addresses to any domain names you plan to use.
4. Perform ongoing tasks in the application and in StorageGRID to manage and monitor object storage over time.



Information needed to attach StorageGRID to a client application

Before you can attach StorageGRID to an S3 or Swift client application, you must perform configuration steps in StorageGRID and obtain certain value.

What values do I need?

The following table shows the values you must configure in StorageGRID and where those values are used by the S3 or Swift application and the DNS server.

Value	Where value is configured	Where value is used
Virtual IP (VIP) addresses	StorageGRID > HA group	DNS entry
Port	StorageGRID > Load balancer endpoint	Client application
SSL certificate	StorageGRID > Load balancer endpoint	Client application
Server name (FQDN)	StorageGRID > Load balancer endpoint	<ul style="list-style-type: none"> • Client application • DNS entry
S3 access key ID and secret access key	StorageGRID > Tenant and bucket	Client application
Bucket/Container name	StorageGRID > Tenant and bucket	Client application

How do I get these values?

Depending on your requirements, you can do either of the following to obtain the information you need:

- **Use the [S3 setup wizard](#).** The S3 setup wizard helps you to quickly configure the required values in StorageGRID and outputs one or two files that you can use when you configure the S3 application. The wizard guides you through the required steps and helps to make sure your settings conform to StorageGRID best practices.



If you are configuring an S3 application, using the S3 setup wizard is recommended unless you know you have special requirements or your implementation will require significant customization.

- **Use the [FabricPool setup wizard](#).** Similar to the S3 setup wizard, the FabricPool setup wizard helps you to quickly configure required values and outputs a file that you can use when you configure a FabricPool cloud tier in ONTAP.



If you plan to use StorageGRID as the object storage system for a FabricPool cloud tier, using the FabricPool setup wizard is recommended unless you know you have special requirements or your implementation will require significant customization.

- **Configure items manually.** If you are connecting to a Swift application (or you are connecting to an S3 application and prefer not to use the S3 setup wizard), you can obtain the required values by performing the configuration manually. Follow these steps:
 1. Configure the high availability (HA) group you want to use for the S3 or Swift application. See [Configure high availability groups](#).
 2. Create the load balancer endpoint that the S3 or Swift application will use. See [Configure load balancer endpoints](#).
 3. Create the tenant account that the S3 or Swift application will use. See [Create a tenant account](#).
 4. For an S3 tenant, sign in to the tenant account, and generate an access key ID and secret access key

for each user that will access the application. See [Create your own access keys](#).

5. Create one or more S3 buckets or Swift containers within the tenant account. For S3, see [Create S3 bucket](#). For Swift, use the [PUT container request](#).
6. To add specific placement instructions for the objects belonging to the new tenant or bucket/container, create a new ILM rule and activate a new ILM policy to use that rule. See [Create ILM rule](#) and [Create ILM policy](#).

Use S3 setup wizard

Use S3 setup wizard: Considerations and requirements

You can use the S3 setup wizard to configure StorageGRID as the object storage system for an S3 application.

When to use the S3 setup wizard

The S3 setup wizard guides you through each step of configuring StorageGRID for use with an S3 application. As part of completing the wizard, you download files that you can use to enter values into the S3 application. Use the wizard to configure your system more quickly and to make sure your settings conform to StorageGRID best practices.

If you have the Root access permission, you can complete the S3 setup wizard when you start using the StorageGRID Grid Manager, or you can access and complete the wizard at any later time. Depending on your requirements, you can also configure some or all of the required items manually and then use the wizard to assemble the values that an S3 application needs.

Before using the wizard

Before using the wizard, confirm you have completed these prerequisites.

Obtain IP addresses and set up VLAN interfaces

If you will configure a high availability (HA) group, you know which nodes the S3 application will connect to and which StorageGRID network will be used. You also know which values to enter for the subnet CIDR, gateway IP address, and virtual IP (VIP) addresses.

If you plan to use a virtual LAN to segregate the traffic from the S3 application, you have already configured the VLAN interface. See [Configure VLAN interfaces](#).

Configure identity federation and SSO

If you plan to use identity federation or single sign-on (SSO) for your StorageGRID system, you have enabled these features. You also know which federated group should have root access for the tenant account that the S3 application will use. See [Use identity federation](#) and [Configure single sign-on](#).

Obtain and configure domain names

You know which fully qualified domain name (FQDN) to use for StorageGRID. Domain name server (DNS) entries will map this FQDN to the virtual IP (VIP) addresses of the HA group that you create using the wizard.

If you plan to use S3 virtual hosted-style requests, you should have [configured S3 endpoint domain names](#). Using virtual hosted-style requests is recommended.

Review load balancer and security certificate requirements

If you plan to use the StorageGRID load balancer, you have reviewed the general considerations for load balancing. You have the certificates you will upload or the values you need to generate a certificate.

If you plan to use an external (third-party) load balancer endpoint, you have the fully qualified domain name (FQDN), port, and certificate for that load balancer.

Configure any grid federation connections

If you want to allow the S3 tenant to clone account data and replicate bucket objects to another grid using a grid federation connection, confirm the following before starting the wizard:

- You have [configured the grid federation connection](#).
- The status of the connection is **Connected**.
- You have Root access permission.

Access and complete the S3 setup wizard

You can use the S3 setup wizard to configure StorageGRID for use with an S3 application. The setup wizard provides the values the application needs to access a StorageGRID bucket and to save objects.

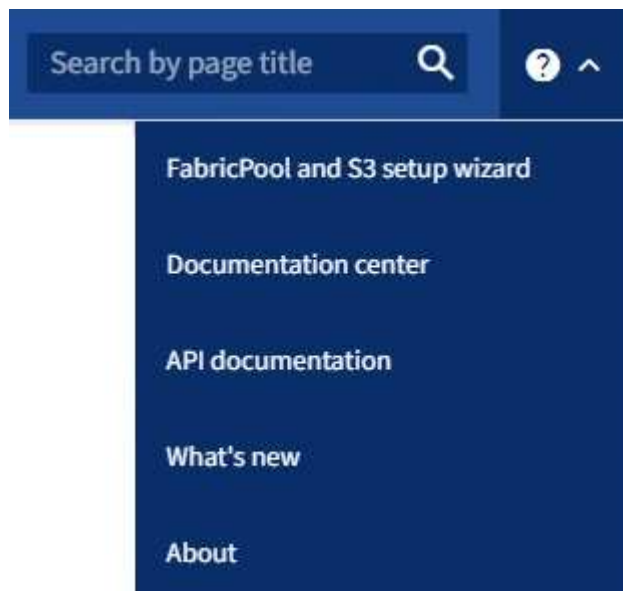
Before you begin

- You have the [Root access permission](#).
- You have reviewed the [considerations and requirements](#) for using the wizard.

Access the wizard

Steps

1. Sign in to the Grid Manager using a [supported web browser](#).
2. If the **FabricPool and S3 setup wizard** banner appears on the dashboard, select the link in the banner. If the banner no longer appears, select the help icon from the header bar in the Grid Manager and select **FabricPool and S3 setup wizard**.



3. In the S3 application section of the FabricPool and S3 setup wizard page, select **Configure now**.

Step 1 of 6: Configure HA group

An HA group is a collection of nodes that each contain the StorageGRID Load Balancer service. An HA group can contain Gateway Nodes, Admin Nodes, or both.

You can use an HA group to help keep the S3 data connections available. If the active interface in the HA group fails, a backup interface can manage the workload with little impact to S3 operations.

For details about this task, see [Manage high availability groups](#).

Steps

1. If you plan to use an external load balancer, you don't need to create an HA group. Select **Skip this step** and go to [Step 2 of 6: Configure load balancer endpoint](#).
2. To use the StorageGRID load balancer, you can create a new HA group or use an existing HA group.

Create HA group

- a. To create a new HA group, select **Create HA group**.
- b. For the **Enter details** step, complete the following fields.

Field	Description
HA group name	A unique display name for this HA group.
Description (optional)	The description of this HA group.

- c. For the **Add interfaces** step, select the node interfaces you want to use in this HA group.

Use the column headers to sort the rows, or enter a search term to locate interfaces more quickly.

You can select one or more nodes, but you can select only one interface for each node.

- d. For the **Prioritize interfaces** step, determine the Primary interface and any backup interfaces for this HA group.

Drag rows to change the values in the **Priority order** column.

The first interface in the list is the Primary interface. The Primary interface is the active interface unless a failure occurs.

If the HA group includes more than one interface and the active interface fails, the virtual IP (VIP) addresses move to the first backup interface in the priority order. If that interface fails, the VIP addresses move to the next backup interface, and so on. When failures are resolved, the VIP addresses move back to the highest priority interface available.

- e. For the **Enter IP addresses** step, complete the following fields.

Field	Description
Subnet CIDR	<p>The address of the VIP subnet in CIDR notation — an IPv4 address followed by a slash and the subnet length (0-32).</p> <p>The network address must not have any host bits set. For example, 192.16.0.0/22.</p>
Gateway IP address (optional)	If the S3 IP addresses used to access StorageGRID aren't on the same subnet as the StorageGRID VIP addresses, enter the StorageGRID VIP local gateway IP address. The local gateway IP address must be within the VIP subnet.
Virtual IP address	<p>Enter at least one and no more than ten VIP addresses for the active interface in the HA group. All VIP addresses must be within the VIP subnet.</p> <p>At least one address must be IPv4. Optionally, you can specify additional IPv4 and IPv6 addresses.</p>

- f. Select **Create HA group** and then select **Finish** to return to the S3 setup wizard.
- g. Select **Continue** to go to the load balancer step.

Use existing HA group

- a. To use an existing HA group, select the HA group name from the **Select an HA group**.
- b. Select **Continue** to go to the load balancer step.

Step 2 of 6: Configure load balancer endpoint

StorageGRID uses a load balancer to manage the workload from client applications. Load balancing maximizes speed and connection capacity across multiple Storage Nodes.

You can use the StorageGRID Load Balancer service, which exists on all Gateway and Admin Nodes, or you can connect to an external (third-party) load balancer. Using the StorageGRID load balancer is recommended.

For details about this task, see [Considerations for load balancing](#).

To use the StorageGRID Load Balancer service, select the **StorageGRID load balancer** tab and then create or select the load balancer endpoint you want to use. To use an external load balancer, select the **External load balancer** tab and provide details about the system you have already configured.

Create endpoint

Steps

1. To create a load balancer endpoint, select **Create endpoint**.
2. For the **Enter endpoint details** step, complete the following fields.

Field	Description
Name	A descriptive name for the endpoint.
Port	<p>The StorageGRID port you want to use for load balancing. This field defaults to 10433 for the first endpoint you create, but you can enter any unused external port. If you enter 80 or 443, the endpoint is configured only on Gateway Nodes, because these ports are reserved on Admin Nodes.</p> <p>Note: Ports used by other grid services aren't permitted. See the Network port reference.</p>
Client type	Must be S3 .
Network protocol	<p>Select HTTPS.</p> <p>Note: Communicating with StorageGRID without TLS encryption is supported but not recommended.</p>

3. For the **Select binding mode** step, specify the binding mode. The binding mode controls how the endpoint is accessed—using any IP address or using specific IP addresses and network interfaces.

Option	Description
Global (default)	<p>Clients can access the endpoint using the IP address of any Gateway Node or Admin Node, the virtual IP (VIP) address of any HA group on any network, or a corresponding FQDN.</p> <p>Use the Global setting (default) unless you need to restrict the accessibility of this endpoint.</p>
Virtual IPs of HA groups	<p>Clients must use a virtual IP address (or corresponding FQDN) of an HA group to access this endpoint.</p> <p>Endpoints with this binding mode can all use the same port number, as long as the HA groups you select for the endpoints don't overlap.</p>
Node interfaces	Clients must use the IP addresses (or corresponding FQDNs) of selected node interfaces to access this endpoint.
Node type	Based on the type of node you select, clients must use either the IP address (or corresponding FQDN) of any Admin Node or the IP address (or corresponding FQDN) of any Gateway Node to access this endpoint.

4. For the Tenant access step, select one of the following:

Field	Description
Allow all tenants (default)	All tenant accounts can use this endpoint to access their buckets.
Allow selected tenants	Only the selected tenant accounts can use this endpoint to access their buckets.
Block selected tenants	The selected tenant accounts can't use this endpoint to access their buckets. All other tenants can use this endpoint.

5. For the **Attach certificate** step, select one of the following:

Field	Description
Upload certificate (recommended)	Use this option to upload a CA-signed server certificate, certificate private key, and optional CA bundle.
Generate certificate	Use this option to generate a self-signed certificate. See Configure load balancer endpoints for details of what to enter.
Use StorageGRID S3 and Swift certificate	Use this option only if you have already uploaded or generated a custom version of the StorageGRID global certificate. See Configure S3 and Swift API certificates for details.

6. Select **Finish** to return to the S3 setup wizard.
7. Select **Continue** to go to the tenant and bucket step.



Changes to an endpoint certificate can take up to 15 minutes to be applied to all nodes.

Use existing load balancer endpoint

Steps

1. To use an existing endpoint, select its name from the **Select a load balancer endpoint**.
2. Select **Continue** to go to the tenant and bucket step.

Use external load balancer

Steps

1. To use an external load balancer, complete the following fields.

Field	Description
FQDN	The fully qualified domain name (FQDN) of the external load balancer.
Port	The port number that the S3 application will use to connect to the external load balancer.

Field	Description
Certificate	Copy the server certificate for the external load balancer and paste it into this field.

2. Select **Continue** to go to the tenant and bucket step.

Step 3 of 6: Create tenant and bucket

A tenant is an entity that can use S3 applications to store and retrieve objects in StorageGRID. Each tenant has its own users, access keys, buckets, objects, and a specific set of capabilities. You must create the tenant before you can create the bucket that the S3 application will use to store its objects.

A bucket is a container used to store a tenant's objects and object metadata. Although some tenants might have many buckets, the wizard helps you to create a tenant and a bucket in the quickest and easiest way. You can use the Tenant Manager later to add any additional buckets you need.

You can create a new tenant for this S3 application to use. Optionally, you can also create a bucket for the new tenant. Finally, you can allow the wizard to create the S3 access keys for the tenant's root user.

For details about this task, see [Create tenant account](#) and [Create S3 bucket](#).

Steps

1. Select **Create tenant**.
2. For the Enter details steps, enter the following information.

Field	Description
Name	A name for the tenant account. Tenant names don't need to be unique. When the tenant account is created, it receives a unique, numeric account ID.
Description (optional)	A description to help identify the tenant.
Client type	The type of client protocol this tenant will use. For the S3 setup wizard, S3 is selected and the field is disabled.
Storage quota (optional)	If you want this tenant to have a storage quota, a numerical value for the quota and the units.

3. Select **Continue**.
4. Optionally, select any permissions you want this tenant to have.



Some of these permissions have additional requirements. For details, select the help icon for each permission.

Permission	If selected...
Allow platform services	The tenant can use S3 platform services such as CloudMirror. See Manage platform services for S3 tenant accounts .
Use own identity source	The tenant can configure and manage its own identity source for federated groups and users. This option is disabled if you have configured SSO for your StorageGRID system.
Allow S3 Select	<p>The tenant can issue S3 SelectObjectContent API requests to filter and retrieve object data. See Manage S3 Select for tenant accounts.</p> <p>Important: SelectObjectContent requests can decrease load-balancer performance for all S3 clients and all tenants. Enable this feature only when required and only for trusted tenants.</p>
Use grid federation connection	<p>The tenant can use a grid federation connection.</p> <p>Selecting this option:</p> <ul style="list-style-type: none"> • Causes this tenant and all tenant groups and users added to the account to be cloned from this grid (the <i>source grid</i>) to the other grid in the selected connection (the <i>destination grid</i>). • Allows this tenant to configure cross-grid replication between corresponding buckets on each grid. <p>See Manage the permitted tenants for grid federation.</p> <p>Note: You can only select Use grid federation connection when you are creating a new S3 tenant; you can't select this permission for an existing tenant.</p>

- If you selected **Use grid federation connection**, select one of the available grid federation connections.
- Define root access for the tenant account, based on whether your StorageGRID system uses [identity federation](#), [single sign-on \(SSO\)](#), or both.

Option	Do this
If identity federation is not enabled	Specify the password to use when signing into the tenant as the local root user.
If identity federation is enabled	<ol style="list-style-type: none"> 1. Select an existing federated group to have Root access permission for the tenant. 2. Optionally, specify the password to use when signing in to the tenant as the local root user.
If both identity federation and single sign-on (SSO) are enabled	Select an existing federated group to have Root access permission for the tenant. No local users can sign in.

7. If you want the wizard to create the access key ID and secret access key for the root user, select **Create root user S3 access key automatically**.



Select this option if the only user for the tenant will be the root user. If other users will use this tenant, use Tenant Manager to configure keys and permissions.

8. Select **Continue**.

9. For the Create bucket step, optionally create a bucket for the tenant's objects. Otherwise, select **Create tenant without bucket** to go to the [download data step](#).



If S3 Object Lock is enabled for the grid, the bucket created in this step doesn't have S3 Object Lock enabled. If you need to use an S3 Object Lock bucket for this S3 application, select **Create tenant without bucket**. Then, use Tenant Manager to [create the bucket](#) instead.

- a. Enter the name of the bucket that the S3 application will use. For example, `S3-bucket`.



You can't change the bucket name after creating the bucket.

- b. Select the **Region** for this bucket.


Use the default region (us-east-1) unless you expect to use ILM in the future to filter objects based on the bucket's region.

- c. Select **Enable object versioning** if you want to store each version of each object in this bucket.
- d. Select **Create tenant and bucket** and go to the download data step.

Step 4 of 6: Download data

In the download data step, you can download one or two files to save the details of what you just configured.

Steps

1. If you selected **Create root user S3 access key automatically**, do one or both of the following:
 - Select **Download access keys** to download a `.csv` file containing the tenant account name, access key ID, and secret access key.
 - Select the copy icon () to copy the access key ID and secret access key to the clipboard.
2. Select **Download configuration values** to download a `.txt` file containing the settings for the load balancer endpoint, tenant, bucket, and the root user.
3. Save this information to a secure location.



Don't close this page until you have copied both access keys. The keys will not be available after you close this page. Make sure to save this information in a secure location because it can be used to obtain data from your StorageGRID system.

4. If prompted, select the checkbox to confirm that you have downloaded or copied the keys.
5. Select **Continue** to go to the ILM rule and policy step.

Step 5 of 6: Review ILM rule and ILM policy for S3

Information lifecycle management (ILM) rules control the placement, duration, and ingest behavior of all objects in your StorageGRID system. The ILM policy included with StorageGRID makes two replicated copies of all objects. This policy is in effect until you create a new proposed policy and activate it.

Steps

1. Review the information provided on the page.
2. If you want to add specific instructions for the objects belonging to the new tenant or bucket, create a new rule and a new policy. See [Create ILM rule](#) and [Create ILM policy: Overview](#).
3. Select **I have reviewed these steps and understand what I need to do**.
4. Select the checkbox to indicate that you understand what to do next.
5. Select **Continue** to go to **Summary**.

Step 6 of 6: Review summary

Steps

1. Review the summary.
2. Make note of the details in the next steps, which describe the additional configuration that might be needed before you connect to the S3 client. For example, selecting **Sign in as root** takes you to the Tenant Manager, where you can add tenant users, create additional buckets, and update bucket settings.
3. Select **Finish**.
4. Configure the application using the file you downloaded from StorageGRID or the values you obtained manually.

Manage HA groups

Manage high availability (HA) groups: Overview

You can group the network interfaces of multiple Admin and Gateway Nodes into a high availability (HA) group. If the active interface in the HA group fails, a backup interface can manage the workload.

What is an HA group?

You can use high availability (HA) groups to provide highly available data connections for S3 and Swift clients or to provide highly available connections to the Grid Manager and the Tenant Manager.

Each HA group provides access to the shared services on the selected nodes.

- HA groups that include Gateway Nodes, Admin Nodes, or both provide highly available data connections for S3 and Swift clients.
- HA groups that include only Admin Nodes provide highly available connections to the Grid Manager and the Tenant Manager.
- An HA group that includes only SG100 or SG1000 appliances and VMware-based software nodes can provide highly available connections for [S3 tenants that use S3 Select](#). HA groups are recommended when using S3 Select, but not required.

How do you create an HA group?

1. You select a network interface for one or more Admin Nodes or Gateway Nodes. You can use a Grid Network (eth0) interface, Client Network (eth2) interface, VLAN interface, or an access interface you have added to the node.



You can't add an interface to an HA group if it has a DHCP-assigned IP address.

2. You specify one interface to be the Primary interface. The Primary interface is the active interface unless a failure occurs.
3. You determine the priority order for any Backup interfaces.
4. You assign one to 10 virtual IP (VIP) addresses to the group. Clients applications can use any of these VIP addresses to connect to StorageGRID.

For instructions, see [Configure high availability groups](#).

What is the active interface?

During normal operation, all of the VIP addresses for the HA group are added to the Primary interface, which is the first interface in the priority order. As long as the Primary interface remains available, it is used when clients connect to any VIP address for the group. That is, during normal operation, the Primary interface is the “active” interface for the group.

Similarly, during normal operation, any lower priority interfaces for the HA group act as “backup” interfaces. These backup interfaces aren't used unless the Primary (currently active) interface becomes unavailable.

View the current HA group status of a node

To see if a node is assigned to an HA group and determine its current status, select **NODES > node**.

If the **Overview** tab includes an entry for **HA groups**, the node is assigned to the HA groups listed. The value after the group name is the current status of the node in the HA group:

- **Active:** The HA group is currently being hosted on this node.
- **Backup:** The HA group is not currently using this node; this is a backup interface.
- **Stopped:** The HA group can't be hosted on this node because the High Availability (keepalived) service has been stopped manually.
- **Fault:** The HA group can't be hosted on this node because of one or more of the following:
 - The Load Balancer (nginx-gw) service is not running on the node.
 - The node's eth0 or VIP interface is down.
 - The node is down.

In this example, the primary Admin Node has been added to two HA groups. This node is currently the active interface for the Admin clients group and a backup interface for the FabricPool clients group.

DC1-ADM1 (Primary Admin Node)

Overview
Hardware
Network
Storage
Load balancer
Tasks

Node information

Name:
DC1-ADM1

Type:
Primary Admin Node

ID:
ce00d9c8-8a79-4742-bdef-c9c658db5315

Connection state:
Connected

Software version:
11.6.0 (build 20211207.1804.614bc17)

HA groups:

Admin clients (Active)

FabricPool clients (Backup)

IP addresses:

172.16.1.225 - eth0 (Grid Network)

10.224.1.225 - eth1 (Admin Network)

47.47.0.2, 47.47.1.225 - eth2 (Client Network)

Show additional IP addresses

What happens when the active interface fails?

The interface that currently hosts the VIP addresses is the active interface. If the HA group includes more than one interface and the active interface fails, the VIP addresses move to the first available backup interface in the priority order. If that interface fails, the VIP addresses move to the next available backup interface, and so on.

Failover can be triggered for any of these reasons:

- The node on which the interface is configured goes down.
- The node on which the interface is configured loses connectivity to all other nodes for at least 2 minutes.
- The active interface goes down.
- The Load Balancer service stops.
- The High Availability service stops.



Failover might not be triggered by network failures external to the node that hosts the active interface. Similarly, failover is not triggered by the services for the Grid Manager or the Tenant Manager.

The failover process generally takes only a few seconds and is fast enough that client applications should experience little impact and can rely on normal retry behaviors to continue operation.

When failure is resolved and a higher priority interface becomes available again, the VIP addresses are automatically moved to the highest priority interface that is available.

How are HA groups used?

You can use high availability (HA) groups to provide highly available connections to StorageGRID for object data and for administrative use.

- An HA group can provide highly available administrative connections to the Grid Manager or the Tenant Manager.
- An HA group can provide highly available data connections for S3 and Swift clients.
- An HA group that contains only one interface allows you to provide many VIP addresses and to explicitly set IPv6 addresses.

An HA group can provide high availability only if all nodes included in the group provide the same services. When you create an HA group, add interfaces from the types of nodes that provide the services you require.

- **Admin Nodes:** Include the Load Balancer service and enable access to the Grid Manager or the Tenant Manager.
- **Gateway Nodes:** Include the Load Balancer service.

Purpose of HA group	Add nodes of this type to the HA group
Access to Grid Manager	<ul style="list-style-type: none">• Primary Admin Node (Primary)• Non-primary Admin Nodes <p>Note: The primary Admin Node must be the Primary interface. Some maintenance procedures can only be performed from the primary Admin Node.</p>
Access to Tenant Manager only	<ul style="list-style-type: none">• Primary or non-primary Admin Nodes
S3 or Swift client access — Load Balancer service	<ul style="list-style-type: none">• Admin Nodes• Gateway Nodes
S3 client access for S3 Select	<ul style="list-style-type: none">• SG100 or SG1000 appliances• VMware-based software nodes <p>Note: HA groups are recommended when using S3 Select, but not required.</p>

Limitations of using HA groups with Grid Manager or Tenant Manager

If a Grid Manager or Tenant Manager service fails, HA group failover is not triggered.

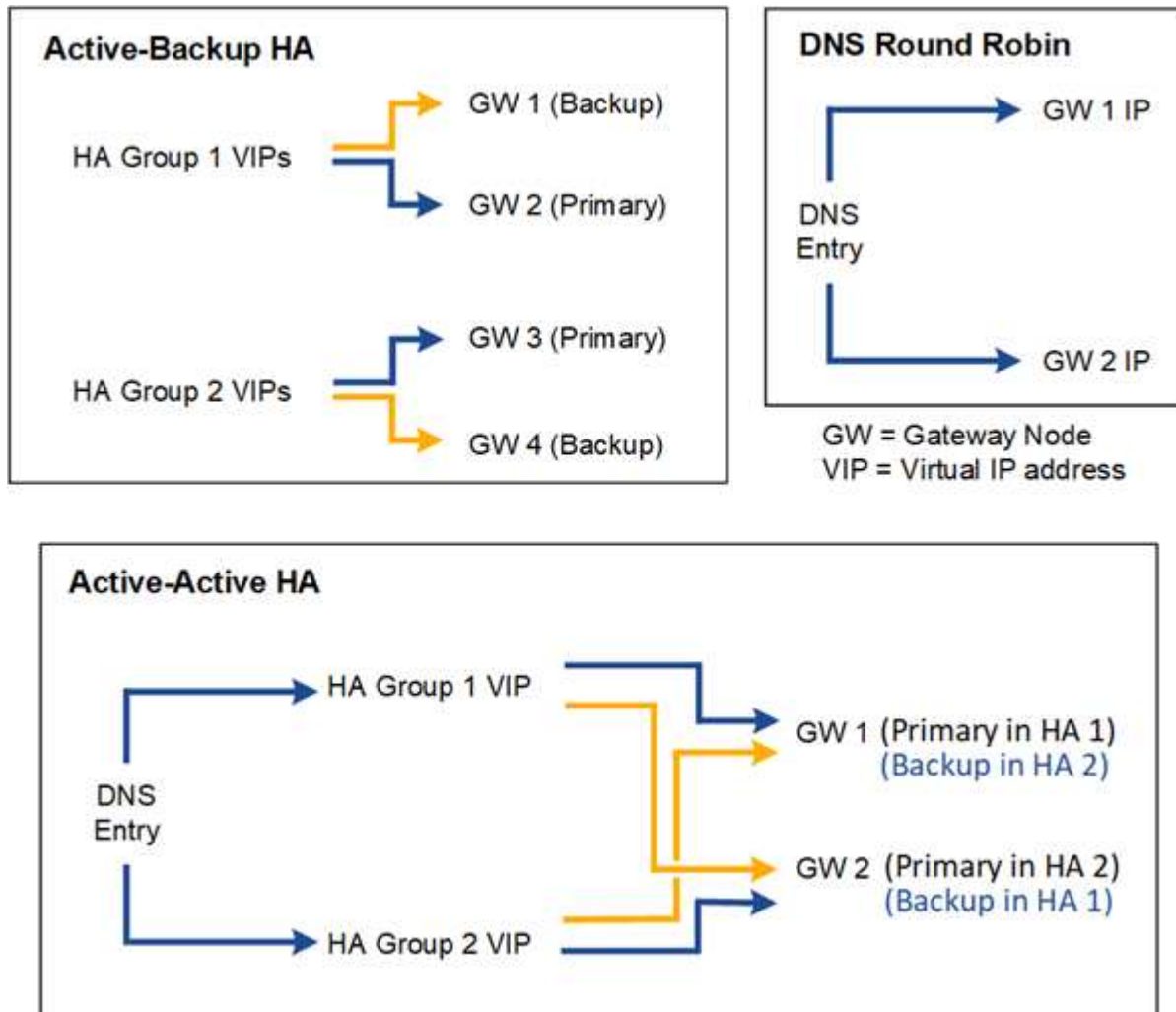
If you are signed in to the Grid Manager or the Tenant Manager when failover occurs, you are signed out and must sign in again to resume your task.

Some maintenance procedures can't be performed when the primary Admin Node is unavailable. During failover, you can use the Grid Manager to monitor your StorageGRID system.

Configuration options for HA groups

The following diagrams provide examples of different ways you can configure HA groups. Each option has advantages and disadvantages.

In the diagrams, blue indicates the primary interface in the HA group and yellow indicates the backup interface in the HA group.



The table summarizes the benefits of each HA configuration shown in the diagram.

Configuration	Advantages	Disadvantages
Active-Backup HA	<ul style="list-style-type: none">• Managed by StorageGRID with no external dependencies.• Fast failover.	<ul style="list-style-type: none">• Only one node in an HA group is active. At least one node per HA group will be idle.

Configuration	Advantages	Disadvantages
DNS Round Robin	<ul style="list-style-type: none"> • Increased aggregate throughput. • No idle hosts. 	<ul style="list-style-type: none"> • Slow failover, which could depend on client behavior. • Requires configuration of hardware outside of StorageGRID. • Needs a customer-implemented health check.
Active-Active HA	<ul style="list-style-type: none"> • Traffic is distributed across multiple HA groups. • High aggregate throughput that scales with the number of HA groups. • Fast failover. 	<ul style="list-style-type: none"> • More complex to configure. • Requires configuration of hardware outside of StorageGRID. • Needs a customer-implemented health check.

Configure high availability groups

You can configure high availability (HA) groups to provide highly available access to the services on Admin Nodes or Gateway Nodes.

Before you begin

- You are signed in to the Grid Manager using a [supported web browser](#).
- You have the Root access permission.
- If you plan to use a VLAN interface in an HA group, you have created the VLAN interface. See [Configure VLAN interfaces](#).
- If you plan to use an access interface for a node in an HA group, you have created the interface:
 - **Red Hat Enterprise Linux or CentOS (before installing the node):** [Create node configuration files](#)
 - **Ubuntu or Debian (before installing the node):** [Create node configuration files](#)
 - **Linux (after installing the node):** [Linux: Add trunk or access interfaces to a node](#)
 - **VMware (after installing the node):** [VMware: Add trunk or access interfaces to a node](#)

Create a high availability group

When you create a high availability group, you select one or more interfaces and organize them in priority order. Then, you assign one or more VIP addresses to the group.

An interface must be for a Gateway Node or an Admin Node to be included in an HA group. An HA group can only use one interface for any given node; however, other interfaces for the same node can be used in other HA groups.

Access the wizard

Steps

1. Select **CONFIGURATION > Network > High availability groups**.
2. Select **Create**.

Enter details for the HA group

Steps

1. Provide a unique name for the HA group.
2. Optionally, enter a description for the HA group.
3. Select **Continue**.

Add interfaces to the HA group


Steps

1. Select one or more interfaces to add to this HA group.













Use the column headers to sort the rows, or enter a search term to locate interfaces more quickly.

Add interfaces to the HA group

Select one or more interfaces for this HA group. You can select only one interface for each node.



Total interface count: 4

	Node 	Interface  	Site  	IPv4 subnet 	Node type  
<input type="checkbox"/>	DC1-ADM1-104-96	eth0 	DC1	10.96.104.0/22	Primary Admin Node
<input type="checkbox"/>	DC1-ADM1-104-96	eth2 	DC1	—	Primary Admin Node
<input type="checkbox"/>	DC2-ADM1-104-103	eth0 	DC2	10.96.104.0/22	Admin Node
<input type="checkbox"/>	DC2-ADM1-104-103	eth2 	DC2	—	Admin Node

0 interfaces selected

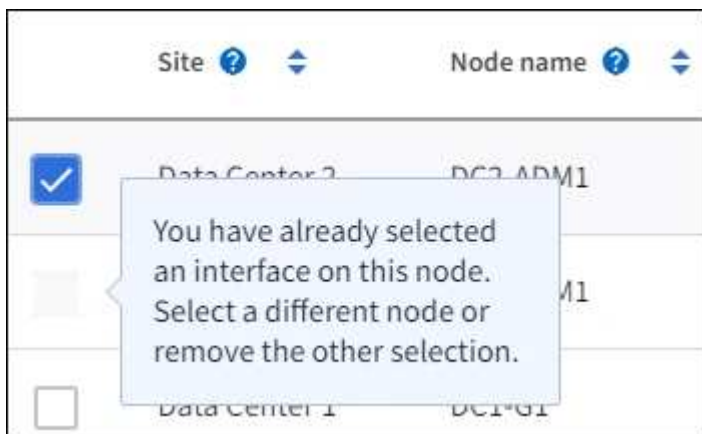


After creating a VLAN interface, wait up to 5 minutes for the new interface to appear in the table.

Guidelines for selecting interfaces

- You must select at least one interface.
- You can select only one interface for a node.
- If the HA group is for HA protection of Admin Node services, which include the Grid Manager and the Tenant Manager, select interfaces on Admin Nodes only.
- If the HA group is for HA protection of S3 or Swift client traffic, select interfaces on Admin Nodes, Gateway Nodes, or both.
- If you select interfaces on different types of nodes, an informational note appears. You are reminded that if a failover occurs, services provided by the previously active node might not be available on the newly active node. For example, a backup Gateway Node can't provide HA protection of Admin Node services. Similarly, a backup Admin Node can't perform all of the maintenance procedures that the primary Admin Node can provide.

- If you can't select an interface, its checkbox is disabled. The tool tip provides more information.



- You can't select an interface if its subnet value or gateway conflicts with another selected interface.
- You can't select a configured interface if it does not have a static IP address.

2. Select **Continue**.

Determine the priority order

If the HA group includes more than one interface, you can determine which is the Primary interface and which are the Backup (failover) interfaces. If the Primary interface fails, the VIP addresses move to the highest priority interface that is available. If that interface fails, the VIP addresses move to the next highest priority interface that is available, and so on.

Steps

1. Drag rows in the **Priority order** column to determine the Primary interface and any Backup interfaces.

The first interface in the list is the Primary interface. The Primary interface is the active interface unless a failure occurs.

Determine the priority order

Determine the primary interface and the backup (failover) interfaces for this HA group. Drag and drop rows or select the arrows.

Priority order ?	Node	Interface ?	Node type ?
1 (Primary interface)	DC1-ADM1-104-96	eth2	Primary Admin Node
2	DC2-ADM1-104-103	eth2	Admin Node



If the HA group provides access to the Grid Manager, you must select an interface on the primary Admin Node to be the Primary interface. Some maintenance procedures can only be performed from the primary Admin Node.

2. Select **Continue**.

Enter IP addresses

Steps

1. In the **Subnet CIDR** field, specify the VIP subnet in CIDR notation—an IPv4 address followed by a slash and the subnet length (0-32).

The network address must not have any host bits set. For example, 192.16.0.0/22.



If you use a 32-bit prefix, the VIP network address also serves as the gateway address and the VIP address.

Enter details for the HA group

Subnet CIDR ⓘ
Specify the subnet in CIDR notation. The optional gateway IP and all VIPs must be in this subnet.

IPv4 address followed by a slash and the subnet length (0-32)

Gateway IP address (optional) ⓘ
Optionally specify the IP address of the gateway, which must be in the subnet. If the subnet address length is 32, the gateway IP address is automatically set to the subnet IP.

Virtual IP address ⓘ
Specify at least 1 and no more than 10 virtual IPs for the HA group. All virtual IPs must be in the same subnet. If the subnet length is 32, only one VIP is allowed, which is automatically set to the subnet/gateway IP.

[Add another IP address](#)

2. Optionally, if any S3, Swift, administrative or tenant clients will access these VIP addresses from a different subnet, enter the **Gateway IP address**. The gateway address must be within the VIP subnet.

Client and admin users will use this gateway to access the virtual IP addresses.

3. Enter at least one and no more than ten VIP addresses for the active interface in the HA group. All VIP addresses must be within the VIP subnet and all will be active at the same time on the active interface.

You must provide at least one IPv4 address. Optionally, you can specify additional IPv4 and IPv6 addresses.

4. Select **Create HA group** and select **Finish**.

The HA Group is created, and you can now use the configured virtual IP addresses.



Wait up to 15 minutes for changes to an HA group to be applied to all nodes.

Next steps

If you will use this HA group for load balancing, create a load balancer endpoint to determine the port and network protocol and to attach any required certificates. See [Configure load balancer endpoints](#).

Edit a high availability group

You can edit a high availability (HA) group to change its name and description, add or remove interfaces, change the priority order, or add or update virtual IP addresses.

For example, you might need to edit an HA group if you want to remove the node associated with a selected interface in a site or node decommission procedure.

Steps

1. Select **CONFIGURATION > Network > High availability groups**.

The High availability groups page shows all existing HA groups.

2. Select the checkbox for the HA group you want to edit.
3. Do one of the following, based on what you want to update:
 - Select **Actions > Edit virtual IP address** to add or remove VIP addresses.
 - Select **Actions > Edit HA group** to update the group's name or description, add or remove interfaces, change the priority order, or add or remove VIP addresses.
4. If you selected **Edit virtual IP address**:
 - a. Update the virtual IP addresses for the HA group.
 - b. Select **Save**.
 - c. Select **Finish**.
5. If you selected **Edit HA group**:
 - a. Optionally, update the group's name or description.
 - b. Optionally, select or clear the checkboxes to add or remove interfaces.



If the HA group provides access to the Grid Manager, you must select an interface on the primary Admin Node to be the Primary interface. Some maintenance procedures can only be performed from the primary Admin Node

- c. Optionally, drag rows to change the priority order of the Primary interface and any Backup interfaces for this HA group.
- d. Optionally, update the virtual IP addresses.
- e. Select **Save** and then select **Finish**.



Wait up to 15 minutes for changes to an HA group to be applied to all nodes.

Remove a high availability group

You can remove one or more high availability (HA) groups at a time.



You can't remove an HA group if it is bound to a load balancer endpoint. To delete an HA group, you must remove it from any load balancer endpoints that use it.

To prevent client disruptions, update any affected S3 or Swift client applications before you remove an HA group. Update each client to connect using another IP address, for example, the virtual IP address of a different HA group or the IP address that was configured for an interface during installation.

Steps

1. Select **CONFIGURATION > Network > High availability groups**.
2. Review the **Load balancer endpoints** column for each HA group you want to remove. If any load balancer endpoints are listed:
 - a. Go to **CONFIGURATION > Network > Load balancer endpoints**.
 - b. Select the checkbox for the endpoint.
 - c. Select **Actions > Edit endpoint binding mode**.
 - d. Update the binding mode to remove the HA group.
 - e. Select **Save changes**.
3. If no load balancer endpoints are listed, select the checkbox for each HA group you want to remove.
4. Select **Actions > Remove HA group**.
5. Review the message and select **Delete HA group** to confirm your selection.

All HA groups you selected are removed. A green success banner appears on the High availability groups page.

Manage load balancing

Considerations for load balancing

You can use load balancing to handle ingest and retrieval workloads from S3 and Swift clients.

What is load balancing?

When a client application saves or retrieves data from a StorageGRID system, StorageGRID uses a load balancer to manage the ingest and retrieval workload. Load balancing maximizes speed and connection capacity by distributing the workload across multiple Storage Nodes.

The StorageGRID Load Balancer service is installed on all Admin Nodes and all Gateway Nodes and provides Layer 7 load balancing. It performs Transport Layer Security (TLS) termination of client requests, inspects the requests, and establishes new secure connections to the Storage Nodes.

The Load Balancer service on each node operates independently when forwarding client traffic to the Storage Nodes. Through a weighting process, the Load Balancer service routes more requests to Storage Nodes with higher CPU availability.



Although the StorageGRID Load Balancer service is the recommended load balancing mechanism, you might want to integrate a third-party load balancer instead. For information, contact your NetApp account representative or refer to [TR-4626: StorageGRID third-party and global load balancers](#).

How many load balancing nodes do I need?

As a general best practice, each site in your StorageGRID system should include two or more nodes with the Load Balancer service. For example, a site might include two Gateway Nodes or both an Admin Node and a Gateway Node. Make sure that there is adequate networking, hardware, or virtualization infrastructure for each load-balancing node, whether you are using SG100 or SG1000 services appliances, bare metal nodes, or virtual machine (VM) based nodes.

What is a load balancer endpoint?

A load balancer endpoint defines the port and the network protocol (HTTPS or HTTP) that incoming and outgoing client application requests will use to access those nodes that contain the Load Balancer service. The endpoint also defines the client type (S3 or Swift), the binding mode, and optionally a list of allowed or blocked tenants.

To create a load balancer endpoint, either select **CONFIGURATION > Network > Load balancer endpoints** or complete the FabricPool and S3 setup wizard. For instructions:

- [Configure load balancer endpoints](#)
- [Use the S3 setup wizard](#)
- [Use the FabricPool setup wizard](#)

Considerations for the port

The port for a load balancer endpoint defaults to 10433 for the first endpoint you create, but you can specify any unused external port between 1 and 65535. If you use port 80 or 443, the endpoint will use the Load Balancer service on Gateway Nodes only. These ports are reserved on Admin Nodes. If you use the same port for more than one endpoint, you must specify a different binding mode for each endpoint.

Ports used by other grid services aren't permitted. See the [Network port reference](#).

Considerations for the network protocol

In most cases, the connections between client applications and StorageGRID should use Transport Layer Security (TLS) encryption. Connecting to StorageGRID without TLS encryption is supported but not recommended, especially in production environments. When you select the network protocol for the StorageGRID load balancer endpoint, you should select **HTTPS**.

Considerations for load balancer endpoint certificates

If you select **HTTPS** as the network protocol for the load balancer endpoint, you must provide a security certificate. You can use any of these three options when you create the load balancer endpoint:

- **Upload a signed certificate (recommended).** This certificate can be signed by either a publicly trusted or a private certificate authority (CA). Using a publicly trusted CA server certificate to secure the connection is the best practice. In contrast to generated certificates, certificates signed by a CA can be rotated nondisruptively, which can help avoid expiration issues.

You must obtain the following files before you create the load balancer endpoint:

- The custom server certificate file.
- The custom server certificate private key file.
- Optionally, a CA bundle of the certificates from each intermediate issuing certificate authority.

- **Generate a self-signed certificate.**
- **Use the global StorageGRID S3 and Swift certificate.** You must upload or generate a custom version of this certificate before you can select it for the load balancer endpoint. See [Configure S3 and Swift API certificates](#).

What values do I need?

To create the certificate, you must know all of the domain names and IP addresses that S3 or Swift client applications will use to access the endpoint.

The **Subject DN** (Distinguished Name) entry for the certificate must include the fully qualified domain name that the client application will use for StorageGRID. For example:

```
Subject DN:  
/C=Country/ST=State/O=Company, Inc./CN=s3.storagegrid.example.com
```

As required, the certificate can use wildcards to represent the fully qualified domain names of all Admin Nodes and Gateway Nodes running the Load Balancer service. For example, `*.storagegrid.example.com` uses the `*` wildcard to represent `adm1.storagegrid.example.com` and `gn1.storagegrid.example.com`.

If you plan to use S3 virtual hosted-style requests, the certificate must also include an **Alternative Name** entry for each [S3 endpoint domain name](#) you have configured, including any wildcard names. For example:

```
Alternative Name: DNS:*.s3.storagegrid.example.com
```



If you use wildcards for domain names, review the [Hardening guidelines for server certificates](#).

You must also define a DNS entry for each name in the security certificate.

How do I manage expiring certificates?



If the certificate used to secure the connection between the S3 application and StorageGRID expires, the application might temporarily lose access to StorageGRID.

To avoid certificate expiration issues, follow these best practices:

- Carefully monitor any alerts that warn of approaching certificate expiration dates, such as the **Expiration of load balancer endpoint certificate** and **Expiration of global server certificate for S3 and Swift API** alerts.
- Always keep the StorageGRID and S3 application's versions of the certificate in sync. If you replace or renew the certificate used for a load balancer endpoint, you must replace or renew the equivalent certificate used by the S3 application.
- Use a publicly signed CA certificate. If you use a certificate signed by a CA, you can replace soon-to-expire certificates nondisruptively.
- If you have generated a self-signed StorageGRID certificate and that certificate is about to expire, you must manually replace the certificate in both StorageGRID and in the S3 application before the existing certificate expires.

Considerations for the binding mode

The binding mode lets you control which IP addresses can be used to access a load balancer endpoint. If an endpoint uses a binding mode, client applications can only access the endpoint if they use an allowed IP address or its corresponding fully qualified domain name (FQDN). Client applications using any other IP address or FQDN can't access the endpoint.

You can specify any of the following binding modes:

- **Global** (default): Client applications can access the endpoint using the IP address of any Gateway Node or Admin Node, the virtual IP (VIP) address of any HA group on any network, or a corresponding FQDN. Use this setting unless you need to restrict the accessibility of an endpoint.
- **Virtual IPs of HA groups**. Client applications must use a virtual IP address (or corresponding FQDN) of an HA group.
- **Node interfaces**. Clients must use the IP addresses (or corresponding FQDNs) of selected node interfaces.
- **Node type**. Based on the type of node you select, clients must use either the IP address (or corresponding FQDN) of any Admin Node or the IP address (or corresponding FQDN) of any Gateway Node.

Considerations for tenant access

Tenant access is an optional security feature that lets you control which StorageGRID tenant accounts can use a load balancer endpoint to access their buckets. You can allow all tenants to access an endpoint (default), or you can specify a list of the allowed or blocked tenants for each endpoint.

You can use this feature to provide better security isolation between tenants and their endpoints. For example, you might use this feature to ensure that the top-secret or highly classified materials owned by one tenant remain completely inaccessible to other tenants.



For the purpose of access control, the tenant is determined from the access keys used in the client request, if no access keys are provided as part of the request (such as with anonymous access) the bucket owner is used to determine the tenant.

Tenant access example

To understand how this security feature works, consider the following example:

1. You have created two load balancer endpoints, as follows:
 - **Public** endpoint: Uses port 10443 and allows access to all tenants.
 - **Top secret** endpoint: Uses port 10444 and allows access to the **Top secret** tenant only. All other tenants are blocked from accessing this endpoint.
2. The `top-secret.pdf` is in a bucket owned by the **Top secret** tenant.

To access the `top-secret.pdf`, a user in the **Top secret** tenant can issue a GET request to `https://w.x.y.z:10444/top-secret.pdf`. Because this tenant is allowed to use the 10444 endpoint, the user can access the object. However, if a user belonging to any other tenant issues the same request to the same URL, they receive an immediate Access Denied message. Access is denied even if the credentials and signature are valid.

CPU availability

The Load Balancer service on each Admin Node and Gateway Node operates independently when forwarding S3 or Swift traffic to the Storage Nodes. Through a weighting process, the Load Balancer service routes more requests to Storage Nodes with higher CPU availability. Node CPU load information is updated every few minutes, but weighting might be updated more frequently. All Storage Nodes are assigned a minimal base weight value, even if a node reports 100% utilization or fails to report its utilization.

In some cases, information about CPU availability is limited to the site where the Load Balancer service is located.

Configure load balancer endpoints

Load balancer endpoints determine the ports and network protocols S3 and Swift clients can use when connecting to the StorageGRID load balancer on Gateway and Admin Nodes.



Support for Swift client applications has been deprecated and will be removed in a future release.

Before you begin

- You are signed in to the Grid Manager using a [supported web browser](#).
- You have the Root access permission.
- You have reviewed the [considerations for load balancing](#).
- If you previously remapped a port you intend to use for the load balancer endpoint, you have [removed the port remap](#).
- You have created any high availability (HA) groups you plan to use. HA groups are recommended, but not required. See [Manage high availability groups](#).
- If the load balancer endpoint will be used by [S3 tenants for S3 Select](#), it must not use the IP addresses or FQDNs of any bare-metal nodes. Only SG100 or SG1000 appliances and VMware-based software nodes are allowed for the load balancer endpoints used for S3 Select.
- You have configured any VLAN interfaces you plan to use. See [Configure VLAN interfaces](#).
- If you are creating an HTTPS endpoint (recommended), you have the information for the server certificate.



Changes to an endpoint certificate can take up to 15 minutes to be applied to all nodes.

- To upload a certificate, you need the server certificate, the certificate private key, and optionally, a CA bundle.
- To generate a certificate, you need all of the domain names and IP addresses that S3 or Swift clients will use to access the endpoint. You must also know the subject (Distinguished Name).
- If you want to use the StorageGRID S3 and Swift API certificate (which can also be used for connections directly to Storage Nodes), you have already replaced the default certificate with a custom certificate signed by an external certificate authority. See [Configure S3 and Swift API certificates](#).

Create a load balancer endpoint

Each load balancer endpoint specifies a port, a client type (S3 or Swift), and a network protocol (HTTP or HTTPS).

Access the wizard

Steps

1. Select **CONFIGURATION > Network > Load balancer endpoints**.
2. Select **Create**.

Enter endpoint details

Steps

1. Enter details for the endpoint.

Field	Description
Name	A descriptive name for the endpoint, which will appear in the table on the Load balancer endpoints page.
Port	<p>The StorageGRID port you want to use for load balancing. This field defaults to 10433 for the first endpoint you create, but you can enter any unused external port between 1 and 65535.</p> <p>If you enter 80 or 443, the endpoint is configured only on Gateway Nodes. These ports are reserved on Admin Nodes.</p>
Client type	The type of client application that will use this endpoint, either S3 or Swift .
Network protocol	<p>The network protocol that clients will use when connecting to this endpoint.</p> <ul style="list-style-type: none">• Select HTTPS for secure, TLS encrypted communication (recommended). You must attach a security certificate before you can save the endpoint.• Select HTTP for less secure, unencrypted communication. Use HTTP only for a non-production grid.

2. Select **Continue**.

Select a binding mode

Steps

1. Select a binding mode for the endpoint to control how the endpoint is accessed—using any IP address or using specific IP addresses and network interfaces.

Option	Description
Global (default)	<p>Clients can access the endpoint using the IP address of any Gateway Node or Admin Node, the virtual IP (VIP) address of any HA group on any network, or a corresponding FQDN.</p> <p>Use the Global setting (default) unless you need to restrict the accessibility of this endpoint.</p>

Option	Description
Virtual IPs of HA groups	<p>Clients must use a virtual IP address (or corresponding FQDN) of an HA group to access this endpoint.</p> <p>Endpoints with this binding mode can all use the same port number, as long as the HA groups you select for the endpoints don't overlap.</p>
Node interfaces	Clients must use the IP addresses (or corresponding FQDNs) of selected node interfaces to access this endpoint.
Node type	Based on the type of node you select, clients must use either the IP address (or corresponding FQDN) of any Admin Node or the IP address (or corresponding FQDN) of any Gateway Node to access this endpoint.



If more than one endpoint uses the same port, StorageGRID uses this priority order to decide which endpoint to use: **Virtual IPs of HA groups** > **Node interfaces** > **Node type** > **Global**.

- If you selected **Virtual IPs of HA groups**, select one or more HA groups.
- If you selected **Node interfaces**, select one or more node interfaces for each Admin Node or Gateway Node that you want to associate with this endpoint.
- If you selected **Node type**, select either Admin Nodes, which includes both the primary Admin Node and any non-primary Admin Nodes, or Gateway Nodes.

Control tenant access

Steps

- For the **Tenant access** step, select one of the following:

Field	Description
Allow all tenants (default)	<p>All tenant accounts can use this endpoint to access their buckets.</p> <p>You must select this option if you have not yet created any tenant accounts. After you add tenant accounts, you can edit the load balancer endpoint to allow or block specific accounts.</p>
Allow selected tenants	Only the selected tenant accounts can use this endpoint to access their buckets.
Block selected tenants	The selected tenant accounts can't use this endpoint to access their buckets. All other tenants can use this endpoint.

- If you are creating an **HTTP** endpoint, you don't need to attach a certificate. Select **Create** to add the new load balancer endpoint. Then, go to [After you finish](#). Otherwise, select **Continue** to attach the certificate.

Attach certificate

Steps

1. If you are creating an **HTTPS** endpoint, select the type of security certificate you want to attach to the endpoint.

The certificate secures the connections between S3 and Swift clients and the Load Balancer service on Admin Node or Gateway Nodes.

- **Upload certificate.** Select this option if you have custom certificates to upload.
- **Generate certificate.** Select this option if you have the values needed to generate a custom certificate.
- **Use StorageGRID S3 and Swift certificate.** Select this option if you want to use the global S3 and Swift API certificate, which can also be used for connections directly to Storage Nodes.

You can't select this option unless you have replaced the default S3 and Swift API certificate, which is signed by the grid CA, with a custom certificate signed by an external certificate authority. See [Configure S3 and Swift API certificates](#).

2. If you aren't using the StorageGRID S3 and Swift certificate, upload or generate the certificate.

Upload certificate

a. Select **Upload certificate**.

b. Upload the required server certificate files:

- **Server certificate**: The custom server certificate file in PEM encoding.
- **Certificate private key**: The custom server certificate private key file (.key).



EC private keys must be 224 bits or larger. RSA private keys must be 2048 bits or larger.

- **CA bundle**: A single optional file containing the certificates from each intermediate issuing certificate authority (CA). The file should contain each of the PEM-encoded CA certificate files, concatenated in certificate chain order.

c. Expand **Certificate details** to see the metadata for each certificate you uploaded. If you uploaded an optional CA bundle, each certificate displays on its own tab.

- Select **Download certificate** to save the certificate file or select **Download CA bundle** to save the certificate bundle.

Specify the certificate file name and download location. Save the file with the extension .pem.

For example: storagegrid_certificate.pem

- Select **Copy certificate PEM** or **Copy CA bundle PEM** to copy the certificate contents for pasting elsewhere.

d. Select **Create**.

The load balancer endpoint is created. The custom certificate is used for all subsequent new connections between S3 and Swift clients and the endpoint.

Generate certificate

a. Select **Generate certificate**.

b. Specify the certificate information:

Field	Description
Domain name	One or more fully qualified domain names to include in the certificate. Use an * as a wildcard to represent multiple domain names.
IP	One or more IP addresses to include in the certificate.
Subject (optional)	X.509 subject or distinguished name (DN) of the certificate owner. If no value is entered in this field, the generated certificate uses the first domain name or IP address as the subject common name (CN).
Days valid	Number of days after creation that the certificate expires.

Field	Description
Add key usage extensions	<p>If selected (default and recommended), key usage and extended key usage extensions are added to the generated certificate.</p> <p>These extensions define the purpose of the key contained in the certificate.</p> <p>Note: Leave this checkbox selected unless you experience connection problems with older clients when certificates include these extensions.</p>

c. Select **Generate**.

d. Select **Certificate details** to see the metadata for the generated certificate.

- Select **Download certificate** to save the certificate file.

Specify the certificate file name and download location. Save the file with the extension `.pem`.

For example: `storagegrid_certificate.pem`

- Select **Copy certificate PEM** to copy the certificate contents for pasting elsewhere.

e. Select **Create**.

The load balancer endpoint is created. The custom certificate is used for all subsequent new connections between S3 and Swift clients and this endpoint.

After you finish

Steps

1. If you use a DNS, ensure that the DNS includes a record to associate the StorageGRID fully qualified domain name (FQDN) to each IP address that clients will use to make connections.

The IP address you enter in the DNS record depends on whether you are using an HA group of load-balancing nodes:

- If you have configured an HA group, clients will connect to the virtual IP addresses of that HA group.
- If you aren't using an HA group, clients will connect to the StorageGRID Load Balancer service using the IP address of a Gateway Node or Admin Node.

You must also ensure that the DNS record references all required endpoint domain names, including any wildcard names.

2. Provide S3 and Swift clients with the information needed to connect to the endpoint:

- Port number
- Fully qualified domain name or IP address
- Any required certificate details

View and edit load balancer endpoints

You can view details for existing load balancer endpoints, including the certificate metadata for a secured endpoint. You can also change an endpoint's name or binding mode and update any associated certificates.

You can't change the service type (S3 or Swift), the port, or the protocol (HTTP or HTTPS).

- To view basic information for all load balancer endpoints, review the table on the Load balancer endpoints page.

Create

Actions

Search...

Total endpoints count: 1

<input type="checkbox"/>	Name	Port	Network protocol	Binding mode	Certificate expiration
<input type="checkbox"/>	S3 load balancer endpoint	10443	HTTPS	Global	Jun 12th, 2024

- To view all details about a specific endpoint, including certificate metadata, select the endpoint's name in the table.

S3 load balancer endpoint

Port: 10443

Client type: S3

Network protocol: HTTPS

Binding mode: Global

Endpoint ID: 3d02c126-9437-478c-8b24-08384401d3cb

Remove

Binding mode

Certificate

Tenant access (2 allowed)

You can select a different binding mode or change IP addresses for the current binding mode.

Edit binding mode


Binding mode: Global

This endpoint uses the Global binding mode. Unless there are one or more overriding endpoints for the same port, clients can access this endpoint using the IP address of any Gateway Node, any Admin Node, or the virtual IP of any HA group on any network.

- To edit an endpoint, use the **Actions** menu on the Load balancer endpoints page or the details page for a specific endpoint.



After editing an endpoint, you might need to wait up to 15 minutes for your changes to be applied to all nodes.

Task	Actions menu	Details page
Edit endpoint name	a. Select the checkbox for the endpoint. b. Select Actions > Edit endpoint name . c. Enter the new name. d. Select Save .	a. Select the endpoint name to display the details. b. Select the edit icon  . c. Enter the new name. d. Select Save .
Edit endpoint binding mode	a. Select the checkbox for the endpoint. b. Select Actions > Edit endpoint binding mode . c. Update the binding mode as required. d. Select Save changes .	a. Select the endpoint name to display the details. b. Select Edit binding mode . c. Update the binding mode as required. d. Select Save changes .
Edit endpoint certificate	a. Select the checkbox for the endpoint. b. Select Actions > Edit endpoint certificate . c. Upload or generate a new custom certificate or begin using the global S3 and Swift certificate, as required. d. Select Save changes .	a. Select the endpoint name to display the details. b. Select the Certificate tab. c. Select Edit certificate . d. Upload or generate a new custom certificate or begin using the global S3 and Swift certificate, as required. e. Select Save changes .
Edit tenant access	a. Select the checkbox for the endpoint. b. Select Actions > Edit tenant access . c. Choose a different access option, select or remove tenants from the list, or do both. d. Select Save changes .	a. Select the endpoint name to display the details. b. Select the Tenant access tab. c. Select Edit tenant access . d. Choose a different access option, select or remove tenants from the list, or do both. e. Select Save changes .

Remove load balancer endpoints

You can remove one or more endpoints using the **Actions** menu, or you can remove a single endpoint from the details page.



To prevent client disruptions, update any affected S3 or Swift client applications before you remove a load balancer endpoint. Update each client to connect using a port assigned to another load balancer endpoint. Be sure to update any required certificate information as well.

- To remove one or more endpoints:
 - a. From the Load balancer page, select the checkbox for each endpoint you want to remove.

- b. Select **Actions** > **Remove**.
- c. Select **OK**.
- To remove one endpoint from the details page:
 - a. From the Load balancer page, select the endpoint name.
 - b. Select **Remove** on the details page.
 - c. Select **OK**.

Configure S3 endpoint domain names

To support S3 virtual-hosted-style requests, you must use the Grid Manager to configure the list of S3 endpoint domain names that S3 clients connect to.



Using an IP address for an endpoint domain name is unsupported. Future releases will prevent this configuration.

Before you begin

- You are signed in to the Grid Manager using a [supported web browser](#).
- You have [specific access permissions](#).
- You have confirmed that a grid upgrade is not in progress.



Don't make any changes to the domain name configuration when a grid upgrade is in progress.

About this task

To enable clients to use S3 endpoint domain names, you must do all of the following:

- Use the Grid Manager to add the S3 endpoint domain names to the StorageGRID system.
- Ensure that the [certificate the client uses for HTTPS connections to StorageGRID](#) is signed for all domain names that the client requires.

For example, if the endpoint is `s3.company.com`, you must ensure that the certificate used for HTTPS connections includes the `s3.company.com` endpoint and the endpoint's wildcard Subject Alternative Name (SAN): `*.s3.company.com`.

- Configure the DNS server used by the client. Include DNS records for the IP addresses that clients use to make connections, and ensure that the records reference all required S3 endpoint domain names, including any wildcard names.



Clients can connect to StorageGRID using the IP address of a Gateway Node, an Admin Node, or a Storage Node, or by connecting to the virtual IP address of a high availability group. You should understand how client applications connect to the grid so you include the correct IP addresses in the DNS records.

Clients that use HTTPS connections (recommended) to the grid can use either of these certificates:

- Clients that connect to a load balancer endpoint can use a custom certificate for that endpoint. Each load balancer endpoint can be configured to recognize different S3 endpoint domain names.

- Clients that connect to a load balancer endpoint or directly to a Storage Node can customize the global S3 and Swift API certificate to include all required S3 endpoint domain names.



If you don't add S3 endpoint domain names and the list is empty, support for S3 virtual-hosted-style requests is disabled.

Add an S3 endpoint domain name

Steps

1. Select **CONFIGURATION > Network > S3 endpoint domain names**.
2. Enter the domain name in the **Domain name 1** field. Select **Add another domain name** to add more domain names.
3. Select **Save**.
4. Ensure that the server certificates that clients use match the required S3 endpoint domain names.
 - If clients connect to a load balancer endpoint that uses its own certificate, [update the certificate associated with the endpoint](#).
 - If clients connect to a load balancer endpoint that uses the global S3 and Swift API certificate or directly to Storage Nodes, [update the global S3 and Swift API certificate](#).
5. Add the DNS records required to ensure that endpoint domain name requests can be resolved.

Result

Now, when clients use the endpoint `bucket.s3.company.com`, the DNS server resolves to the correct endpoint and the certificate authenticates the endpoint as expected.

Rename an S3 endpoint domain name

If you change a name used by S3 applications, virtual-hosted-style requests will fail.


Steps

1. Select **CONFIGURATION > Network > S3 endpoint domain names**.
2. Select the domain name field you want to edit and make the necessary changes.
3. Select **Save**.
4. Select **Yes** to confirm your change.

Delete an S3 endpoint domain name

If you remove a name used by S3 applications, virtual-hosted-style requests will fail.

Steps

1. Select **CONFIGURATION > Network > S3 endpoint domain names**.
2. Select the delete icon  next to the domain name.
3. Select **Yes** to confirm the deletion.

Related information

- [Use S3 REST API](#)
- [View IP addresses](#)

- [Configure high availability groups](#)

Summary: IP addresses and ports for client connections

To store or retrieve objects, S3 and Swift client applications connect to the Load Balancer service, which is included on all Admin Nodes and Gateway Nodes, or to the Local Distribution Router (LDR) service, which is included on all Storage Nodes.

Client applications can connect to StorageGRID using the IP address of a grid node and the port number of the service on that node. Optionally, you can create high availability (HA) groups of load-balancing nodes to provide highly available connections that use virtual IP (VIP) addresses. If you want to connect to StorageGRID using a fully qualified domain name (FQDN) instead of an IP or VIP address, you can configure DNS entries.

This table summarizes the different ways that clients can connect to StorageGRID and the IP addresses and ports that are used for each type of connection. If you have already created load balancer endpoints and high availability (HA) groups, see [Where to find IP addresses](#) to locate these values in the Grid Manager.

Where connection is made	Service that client connects to	IP address	Port
HA group	Load Balancer	Virtual IP address of an HA group	Port assigned to the load balancer endpoint
Admin Node	Load Balancer	IP address of the Admin Node	Port assigned to the load balancer endpoint
Gateway Node	Load Balancer	IP address of the Gateway Node	Port assigned to the load balancer endpoint
Storage Node	LDR	IP address of Storage Node	Default S3 ports: <ul style="list-style-type: none">• HTTPS: 18082• HTTP: 18084 Default Swift ports: <ul style="list-style-type: none">• HTTPS: 18083• HTTP: 18085

Example URLs

To connect a client application to the Load Balancer endpoint of an HA group of Gateway Nodes, use a URL structured as shown below:

```
https://VIP-of-HA-group:LB-endpoint-port
```

For example, if the virtual IP address of the HA group is 192.0.2.5 and the port number of the load balancer endpoint is 10443, then an application could use the following URL to connect to StorageGRID:

Where to find IP addresses

1. Sign in to the Grid Manager using a [supported web browser](#).
2. To find the IP address of a grid node:
 - a. Select **NODES**.
 - b. Select the Admin Node, Gateway Node, or Storage Node to which you want to connect.
 - c. Select the **Overview** tab.
 - d. In the Node Information section, note the IP addresses for the node.
 - e. Select **Show more** to view IPv6 addresses and interface mappings.

You can establish connections from client applications to any of the IP addresses in the list:

- **eth0**: Grid Network
- **eth1**: Admin Network (optional)
- **eth2**: Client Network (optional)



If you are viewing an Admin Node or a Gateway Node and it is the active node in a high availability group, the virtual IP address of the HA group is shown on eth2.

3. To find the virtual IP address of a high availability group:
 - a. Select **CONFIGURATION > Network > High availability groups**.
 - b. In the table, note the virtual IP address of the HA group.
4. To find the port number of a Load Balancer endpoint:
 - a. Select **CONFIGURATION > Network > Load balancer endpoints**.
 - b. Note the port number for the endpoint you want to use.



If the port number is 80 or 443, the endpoint is configured only on Gateway Nodes, because those ports are reserved on Admin Nodes. All other ports are configured on both Gateway Nodes and Admin Nodes.

- c. Select the name of the endpoint from the table.
- d. Confirm that the **Client type** (S3 or Swift) matches the client application that will use the endpoint.

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.