Manage load balancing

StorageGRID

NetApp
September 04, 2024

Table of Contents

Manage load balancing

Manage load balancing: Overview

You can use the StorageGRID load balancing functions to handle ingest and retrieval workloads from S3 and Swift clients. Load balancing maximizes speed and connection capacity by distributing the workloads and connections across multiple Storage Nodes.

You can load balance client workloads in the following ways:

- Use the Load Balancer service, which is installed on Admin Nodes and Gateway Nodes. The Load Balancer service provides Layer 7 load balancing and performs TLS termination of client requests, inspects the requests, and establishes new secure connections to the Storage Nodes. This is the recommended load balancing mechanism.

  See How load balancing works - Load Balancer service.

- Use the deprecated Connection Load Balancer (CLB) service, which is installed on Gateway Nodes only. The CLB service provides Layer 4 load balancing and supports link costs.

  See How load balancing works - CLB service (deprecated).

- Integrate a third-party load balancer. Contact your NetApp account representative for details.

How load balancing works - Load Balancer service

The Load Balancer service distributes incoming network connections from client applications to Storage Nodes. To enable load balancing, you must configure load balancer endpoints using the Grid Manager.

You can configure load balancer endpoints only for Admin Nodes or Gateway Nodes, since these node types contain the Load Balancer service. You cannot configure endpoints for Storage Nodes or Archive Nodes.

Each load balancer endpoint specifies a port, a network protocol (HTTP or HTTPS), a client type (S3 or Swift), and a binding mode. HTTPS endpoints require a server certificate. Binding modes allow you to restrict the accessibility of endpoint ports to:

- The virtual IP addresses (VIPs) of specific high availability (HA) groups
- Specific network interfaces of specific Admin and Gateway Nodes

Port considerations

Clients can access any of the endpoints you configure on any node running the Load Balancer service, with two exceptions: ports 80 and 443 are reserved on Admin Nodes, so endpoints configured on these ports support load balancing operations only on Gateway Nodes.

If you have remapped any ports, you cannot use the same ports to configure load balancer endpoints. You can create endpoints using remapped ports, but those endpoints will be remapped to the original CLB ports and service, not the Load Balancer service. Follow the steps in Remove port remaps.

The CLB service is deprecated.

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.

What you'll need

- You are signed in to the Grid Manager using a supported web browser.
- You have the Root access permission.
- 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.

    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`. See Configure S3 API endpoint domain names.

Create a load balancer endpoint

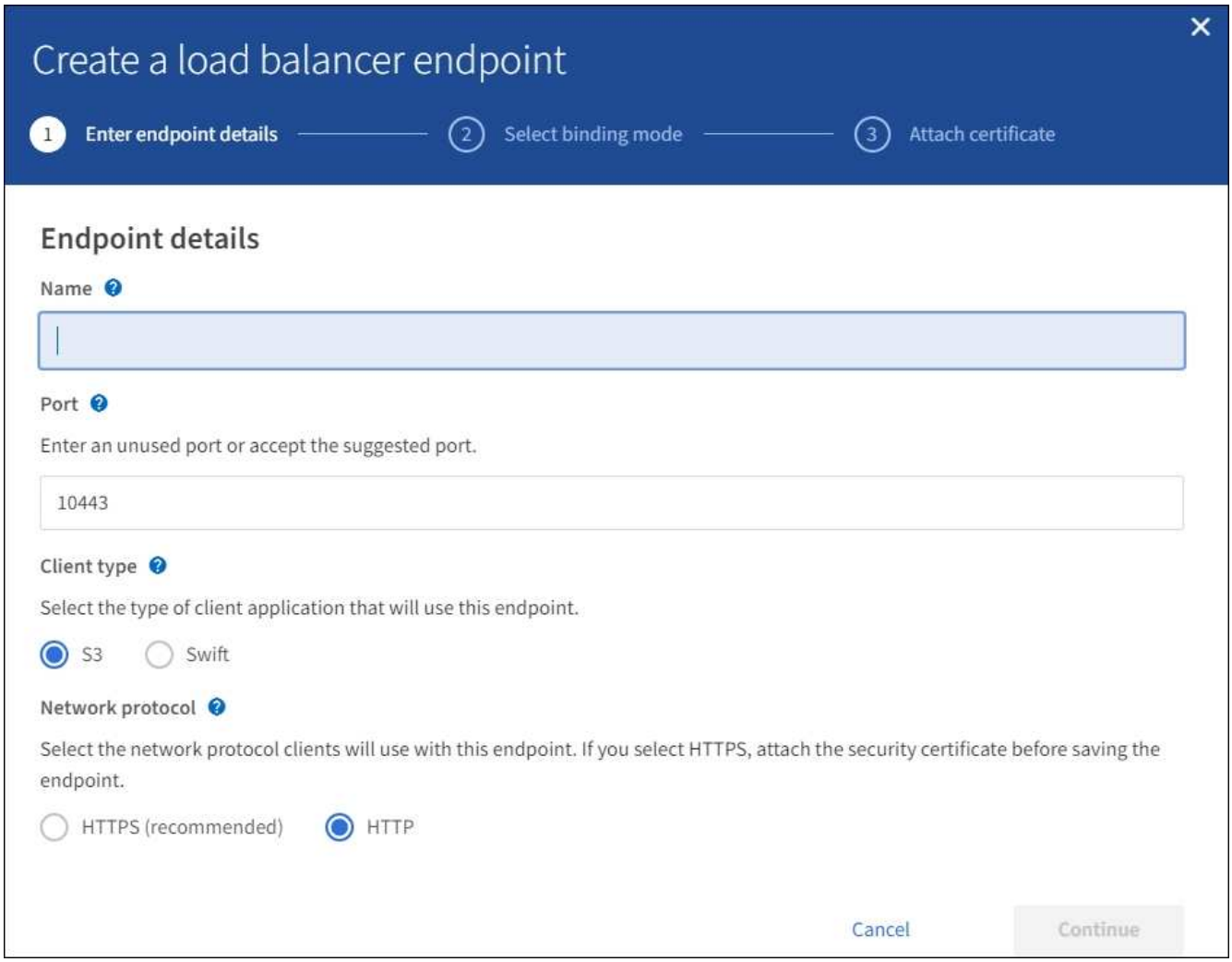
Each load balancer endpoint specifies a port, a client type (S3 or Swift), and a network protocol (HTTP or HTTPS).

Access the wizard

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

Enter endpoint details

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.

Field	Description
Port	The port clients will use to connect to the Load Balancer service on Admin Nodes and Gateway Nodes. Accept the suggested port number or enter any external port that is not used by another grid service. Enter a value between 1 and 65535. If you enter **80** or **443**, the endpoint is configured only on Gateway Nodes. These ports are reserved on Admin Nodes. See the Networking guidelines for information about external ports.
Client type	The type of client application that will use this endpoint, either **S3** or **Swift**.
Network protocol	The network protocol that clients will use when connecting to this endpoint. • 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**.

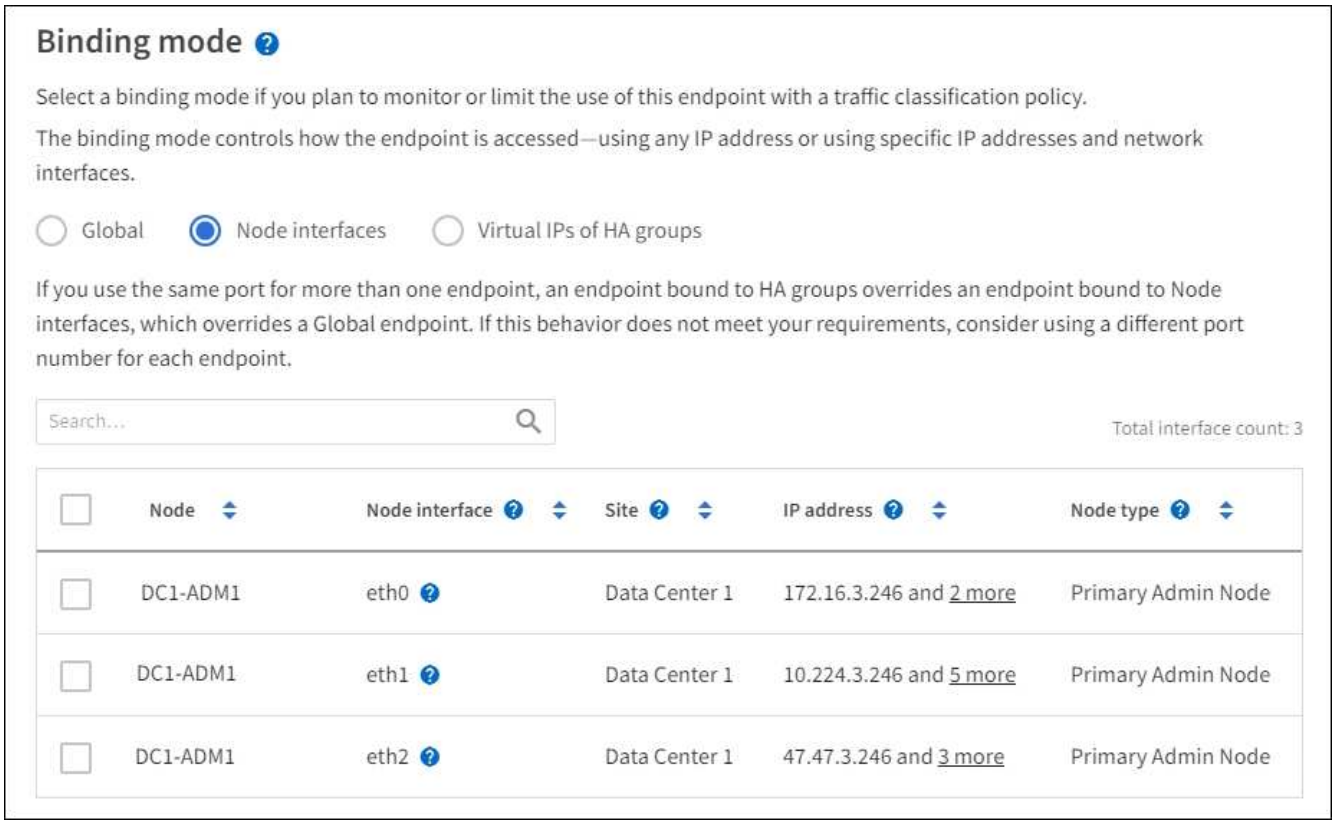
Select the binding mode

1. Select a binding mode for the endpoint to control how the endpoint is accessed.

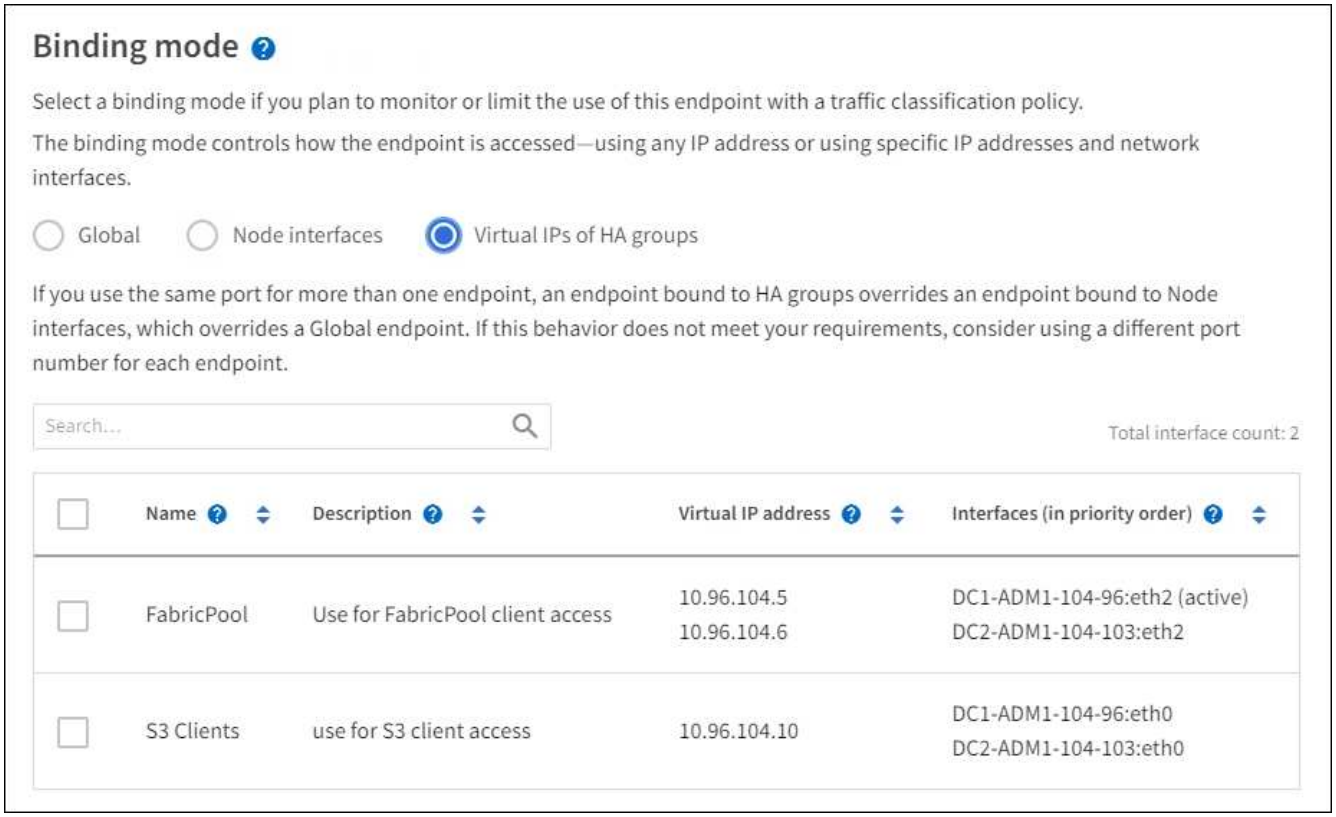
Option	Description
Global (default)	Clients can access the endpoint using a fully qualified domain name (FQDN), the IP address of any Gateway Node or Admin Node, or the virtual IP address of any HA group on any network. Use the **Global** setting (default) unless you need to restrict the accessibility of this endpoint.
Node interfaces	Clients must use the IP address of a selected node and network interface to access this endpoint.
Virtual IPs of HA groups	Clients must use a virtual IP address of an HA group to access this endpoint. Endpoints with this binding mode can all use the same port number, as long as the HA groups you select for the endpoints do not overlap. Endpoints with this mode can all use the same port number as long as the interfaces you select for the endpoints do not overlap.

If you use the same port for more than one endpoint, an endpoint using **Virtual IPs of HA groups** mode overrides an endpoint using **Node interfaces** mode, which overrides an endpoint using **Global** mode.

2. 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.



3. If you selected **Virtual IPs of HA groups**, select one or more HA groups.



4. If you are creating an **HTTP** endpoint, you do not 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

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 cannot 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 are not 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:

- **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**: X.509 subject or distinguished name (DN) of the certificate owner.
- **Days valid**: Number of days after creation that the certificate expires.

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

1. If you use a domain name system (DNS), ensure that the DNS includes a record to associate the StorageGRID fully qualified domain name 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 are not using an HA group, clients will connect to the StorageGRID Load Balancer service using the IP address of any Gateway Node or Admin Node.

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

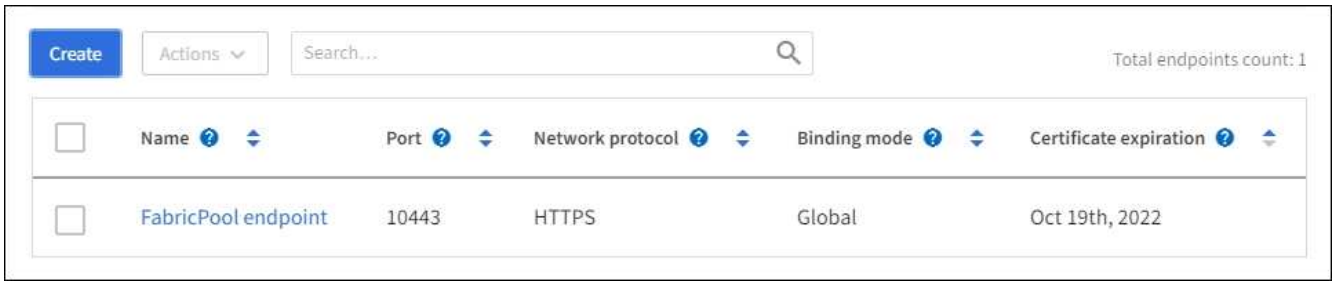
2. Provide S3 and Swift clients with the information needed to connect to the endpoint:
   - Port number
   - Fully qualified domain name or IP address
   - Any required certificate details

View and edit load balancer endpoints

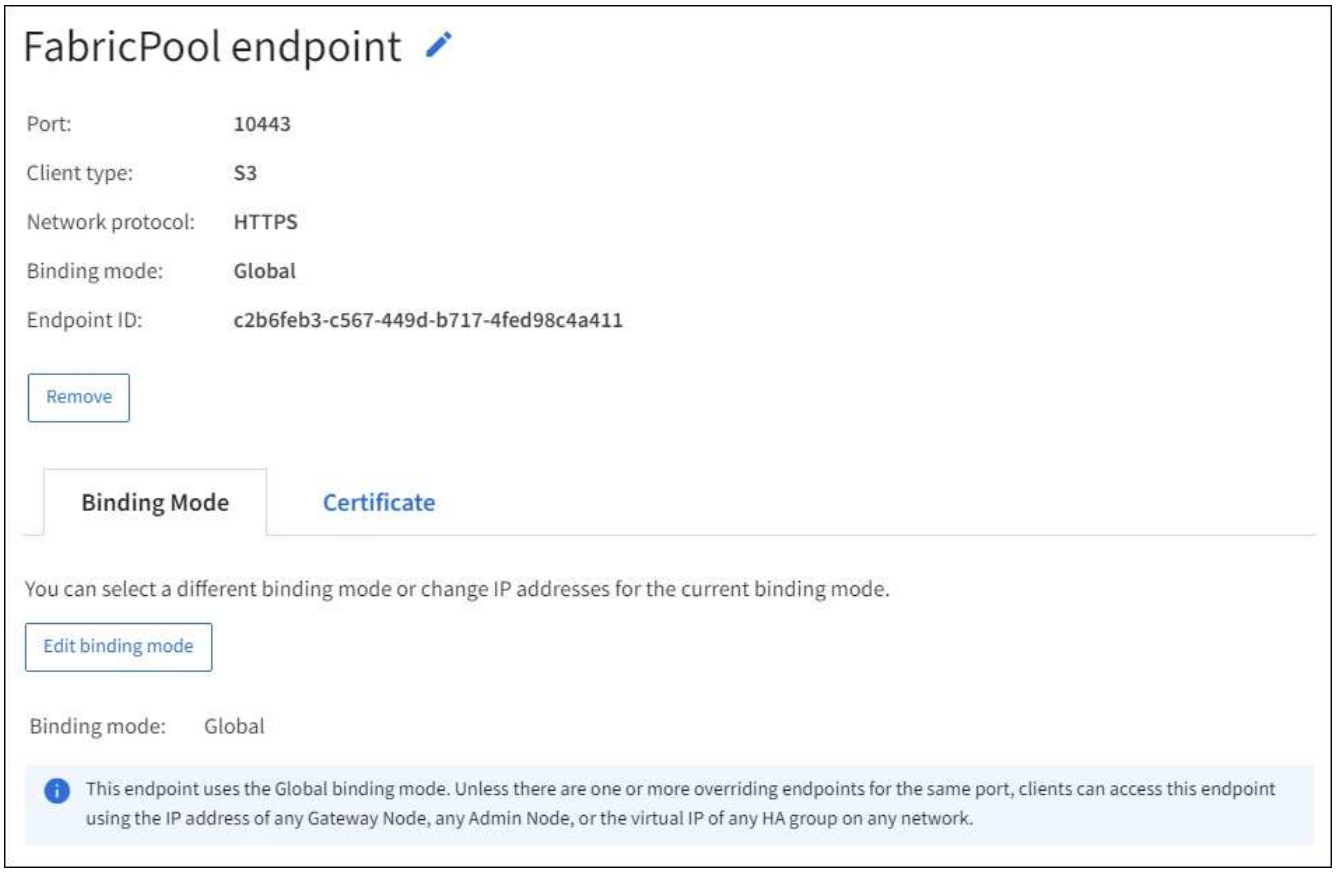
You can view details for existing load balancer endpoints, including the certificate metadata for a secured endpoint. You can also change an endpoint's name or binding mode and update any associated certificates.

You cannot 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.



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



- 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 check box 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 check box 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**.

Task	Actions menu	Details page
Edit endpoint certificate	a. Select the check box 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**.

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.

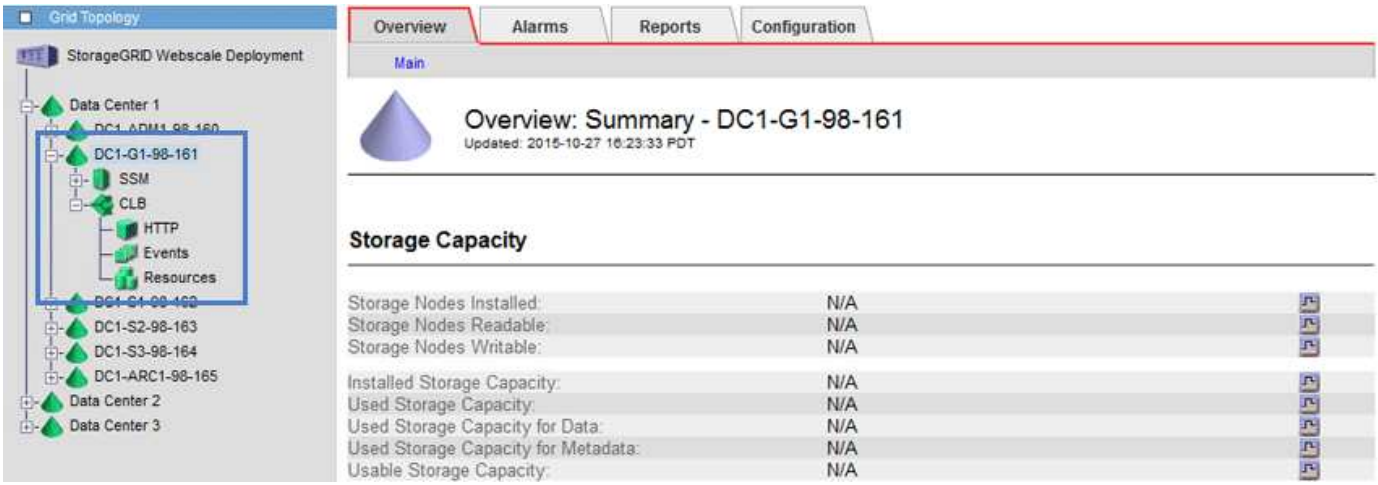
- To remove one or more endpoints:
  a. From the Load balancer page, select the check box for each endpoint you want to remove.
  b. Select **Actions** > **Remove**.
  c. Select **OK**.
- To remove one endpoint from the details page:
  a. From the Load balancer page. select the endpoint name.
  b. Select **Remove** on the details page.
  c. Select **OK**.

How load balancing works - CLB service (deprecated)

The Connection Load Balancer (CLB) service on Gateway Nodes is deprecated. The Load Balancer service is now the recommended load balancing mechanism.

The CLB service uses Layer 4 load balancing to distribute incoming TCP network connections from client applications to the optimal Storage Node based on availability, system load, and the administrator-configured link cost. When the optimal Storage Node is chosen, the CLB service establishes a two-way network connection and forwards the traffic to and from the chosen node. The CLB does not consider the Grid Network configuration when directing incoming network connections.

To view information about the CLB service, select **SUPPORT** > **Tools** > **Grid topology**, and then expand a Gateway Node until you can select **CLB** and the options below it.



If you choose to use the CLB service, you should consider configuring link costs for your StorageGRID system.

- What link costs are
- Update link costs

Copyright information

Trademark information