



How to enable StorageGRID in your environment

How to enable StorageGRID in your environment

NetApp
March 05, 2024

Table of Contents

| | |
|---|-----|
| How to enable StorageGRID in your environment | 1 |
| Validated third-party solutions | 2 |
| Validated third-party solutions: Overview | 2 |
| StorageGRID 11.8 validated third-party solutions | 2 |
| StorageGRID 11.7 validated third-party solutions | 4 |
| StorageGRID 11.6 validated third-party solutions | 7 |
| StorageGRID 11.5 validated third-party solutions | 9 |
| StorageGRID 11.4 validated third-party solutions | 11 |
| StorageGRID 11.3 validated third-party solutions | 13 |
| StorageGRID 11.2 validated third-party solutions | 14 |
| Product feature guides | 17 |
| Create Cloud Storage Pool for AWS or Google Cloud | 17 |
| Create Cloud Storage Pool for Azure Blob Storage | 17 |
| Use a Cloud Storage Pool for backup | 18 |
| Configure StorageGRID search integration service | 19 |
| Node Clone | 35 |
| How to use port remap | 38 |
| Tool and application guides | 49 |
| Use Cloudera Hadoop S3A connector with StorageGRID | 49 |
| Use S3cmd to test and demonstrate S3 access on StorageGRID | 55 |
| Vertica Eon mode database using NetApp StorageGRID as communal storage | 56 |
| StorageGRID log analytics using ELK stack | 69 |
| Use Prometheus and Grafana to extend your metrics retention | 75 |
| Datadog SNMP configuration | 91 |
| Use rclone to migrate, PUT, and DELETE objects on StorageGRID | 94 |
| StorageGRID best practices for deployment with Veeam Backup and Replication | 106 |
| Configure Dremio data source with StorageGRID | 117 |
| Procedures and API examples | 121 |
| Test and demonstrate S3 encryption options on StorageGRID | 121 |
| Test and demonstrate S3 object lock on StorageGRID | 124 |
| Example bucket and Group(IAM) policies | 129 |
| NetApp StorageGRID Blogs | 136 |
| NetApp StorageGRID documentation | 137 |
| Legal notices | 138 |
| Copyright | 138 |
| Trademarks | 138 |
| Patents | 138 |
| Privacy policy | 138 |
| Open source | 138 |

How to enable StorageGRID in your environment

Validated third-party solutions

Validated third-party solutions: Overview

NetApp, in collaboration with our partners, has validated these solutions for use with StorageGRID. Review the information in this section to learn what solutions have been validated, and to obtain additional instructions if applicable.

Join forces with NetApp to accelerate portfolio innovation, expand market awareness and increase sales when you create tested, best-of-breed NetApp solutions. [Become an alliance partner today.](#)

StorageGRID 11.8 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.8. If the solution you are looking for is not listed, please contact your NetApp account representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Alluxio
- Apache Kafka
- AWS Mountpoint
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Diskover Data
- Dremio
- eMAM
- FujiFilm Object Archive
- GitHub Enterprise Server
- IBM Filenet
- IBM Spectrum Protect Plus
- Interica

- Komprise
- Microsoft SQL Server Big Data Clusters
- Model9
- Modzy
- Moonwalk Universal
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura
- PixitMedia ngenea
- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Reveille v10 build 220706 or above
- Rubrik CDM
- s3a
- Signiant
- Snowflake
- Spectra Logic On-Prem Glacier
- Splunk Smartstore
- Storage Made Easy
- Trino
- Varnish Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtualica StorageFabric
- Weka v3.10 or later

Third-party solutions validated on StorageGRID with object lock

These solutions have been tested in collaboration with the respective partners.

- Commvault 11 Feature Release 26
- IBM Filenet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 and later

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform
- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.7 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.7. If the solution you are looking for is not listed, please contact your NetApp account representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Alluxio

- Apache Kafka
- AWS Mountpoint
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Diskover Data
- Dremio
- eMAM
- FujiFilm Object Archive
- GitHub Enterprise Server
- IBM Filenet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Microsoft SQL Server Big Data Clusters
- Model9
- Modzy
- Moonwalk Universal
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura
- PixitMedia ngenea
- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Reveille v10 build 220706 or above

- Rubrik CDM
- s3a
- Signiant
- Snowflake
- Spectra Logic On-Prem Glacier
- Splunk Smartstore
- Storage Made Easy
- Trino
- Varnish Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtualica StorageFabric
- Weka v3.10 or later

Third-party solutions validated on StorageGRID with object lock

These solutions have been tested in collaboration with the respective partners.

- Commvault 11 Feature Release 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 and later

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform
- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab

- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.6 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.6. If the solution you are looking for is not listed, please contact your NetApp account representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Alluxio
- Apache Kafka
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Diskover Data
- Dremio
- eMAM
- FujiFilm Object Archive
- GitHub Enterprise Server
- IBM FileNet
- IBM Spectrum Protect Plus
- Interica

- Komprise
- Microsoft SQL Server Big Data Clusters
- Model9
- Modzy
- Moonwalk Universal
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura
- PixitMedia ngenea
- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Reveille v10 build 220706 or above
- Rubrik CDM
- s3a
- Signiant
- Snowflake
- Spectra Logic On-Prem Glacier
- Splunk Smartstore
- Storage Made Easy
- Trino
- Varnish Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtualica StorageFabric
- Weka v3.10 or later

Third-party solutions validated on StorageGRID with object lock

These solutions have been tested in collaboration with the respective partners.

- Commvault 11 Feature Release 26
- IBM Filenet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 and later

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform
- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.5 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.5. If the solution you are looking for is not listed, please contact your NetApp account representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Alluxio

- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- Moonwalk Universal
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura
- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM
- s3a
- Signiant
- Splunk Smartstore
- Trino
- Varnish Enterprise 6.0.4
- Veeam 11
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Vortalica StorageFabric

Third-party solutions validated on StorageGRID with object lock

These solutions have been tested in collaboration with the respective partners.

- OpenText Documentum 21.4
- Veeam 11

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform
- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.4 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.4. If the solution you are looking for is not listed, please contact your NetApp account representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Bridgestor
- Cantemo

- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura
- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM
- Signiant
- Splunk Smartstore
- Varnish Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform

- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.3 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.3. If the solution you are looking for is not listed, please contact your NetApp account representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura

- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- Splunk Smartstore
- Varnish Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform
- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.2 validated third-party solutions

The following third-party solutions have been validated for use with StorageGRID 11.2. If the solution you are looking for is not listed, please contact your NetApp account

representative.

Third-party solutions validated on StorageGRID

These solutions have been tested in collaboration with the respective partners.

- Actifio
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- NICE
- Nasuni
- OpenText Documentum 16.4
- OpenText Media Management 16.5 with CyanGate Cloud
- Panzura
- PoINT Archival Gateway 2.0
- PoINT Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- Splunk Smartstore
- Varnish Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

Third-party solutions supported on StorageGRID

These solutions have been tested.

- Archiware
- Axis Communications
- Congruity360
- DataFrameworks
- EcoDigital DIVA platform
- Encoding.com
- FujiFilm Object Archive
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Reach Engine
- SilverTrak
- SoftNAS
- QStar
- Velasea

Product feature guides

Create Cloud Storage Pool for AWS or Google Cloud

You can use a Cloud Storage Pool if you want to move StorageGRID objects to an external S3 bucket. The external bucket can belong to Amazon S3 (AWS) or Google Cloud.

What you'll need

- StorageGRID 11.6 has been configured.
- You have already set up an external S3 bucket on AWS or Google Cloud.

Steps

1. In the Grid Manager, navigate to **ILM > Storage Pools**.
2. In the Cloud Storage Pools section of the page, select **Create**.

The Create Cloud Storage Pool pop-up appears.

3. Enter a display name.
4. Select **Amazon S3** from the Provider Type drop-down list.

This provider type works for AWS S3 or Google Cloud.

5. Enter the URI for the S3 bucket to be used for the Cloud Storage Pool.

Two formats are allowed:

`https://host:port`

`http://host:port`

6. Enter the S3 bucket name.

The name you specify must exactly match the S3 bucket's name; otherwise, Cloud Storage Pool creation fails. You cannot change this value after the Cloud Storage Pool is saved.

7. Optionally, enter the Access Key ID and the Secret Access Key.
8. Select **Do Not Verify Certificate** from the drop-down.
9. Click **Save**.

Expected result

Confirm that a Cloud Storage Pool has been created for Amazon S3 or Google Cloud.

By Jonathan Wong

Create Cloud Storage Pool for Azure Blob Storage

You can use a Cloud Storage Pool if you want to move StorageGRID objects to an external Azure container.

What you'll need

- StorageGRID 11.6 has been configured.
- You have already set up an external Azure container.

Steps

1. In the Grid Manager, navigate to **ILM > Storage Pools**.
2. In the Cloud Storage Pools section of the page, select **Create**.

The Create Cloud Storage Pool pop-up appears.

3. Enter a display name.
4. Select **Azure Blob Storage** from the Provider Type drop-down list.
5. Enter the URI for the S3 bucket to be used for the Cloud Storage Pool.

Two formats are allowed:

`https://host:port`

`http://host:port`

6. Enter the Azure container name.

The name you specify must exactly match the Azure container name; otherwise, Cloud Storage Pool creation fails. You cannot change this value after the Cloud Storage Pool is saved.

7. Optionally, enter the Azure container's associated account name and account key for authentication.
8. Select **Do Not Verify Certificate** from the drop-down.
9. Click **Save**.

Expected result

Confirm that a Cloud Storage Pool has been created for Azure Blob Storage.

By Jonathan Wong

Use a Cloud Storage Pool for backup

You can create an ILM rule to move objects into a Cloud Storage Pool for backup..

What you'll need

- StorageGRID 11.6 has been configured.
- You have already set up an external Azure container.

Steps

1. In the Grid Manager, navigate to **ILM > Rules > Create**.
2. Enter a description.
3. Enter a criterion to trigger the rule.
4. Click **Next**.

5. Replicate the object to Storage Nodes.
6. Add a placement rule.
7. Replicate the object to the Cloud Storage Pool
8. Click **Next**.
9. Click **Save**.

Expected result

Confirm that the retention diagram shows the objects stored locally in StorageGRID and in a Cloud Storage Pool for backup.

Confirm that, when the ILM rule is triggered, a copy exists in the Cloud Storage Pool and you can retrieve the object locally without doing an object restore.

By Jonathan Wong

Configure StorageGRID search integration service

This guide provides detailed instructions for configuring NetApp StorageGRID 11.6 search integration service with either Amazon OpenSearch Service or on-premises Elasticsearch.

Introduction

StorageGRID supports three types of platform services.

- **StorageGRID CloudMirror replication.** Mirror specific objects from a StorageGRID bucket to a specified external destination.
- **Notifications.** Per-bucket event notifications to send notifications about specific actions performed on objects to a specified external Amazon Simple Notification Service (Amazon SNS).
- **Search integration service.** Send Simple Storage Service (S3) object metadata to a specified Elasticsearch index where you can search or analyze the metadata by using the external service.

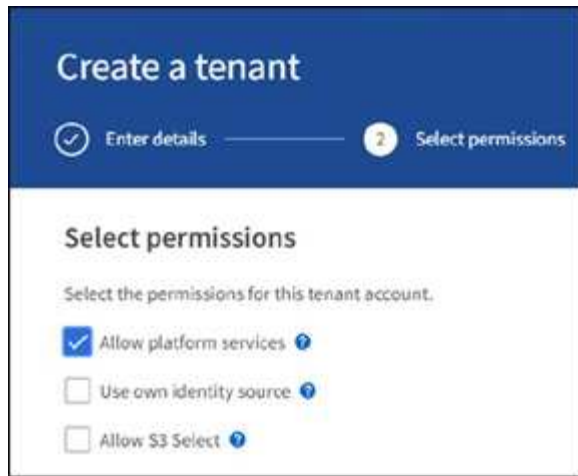
Platform services are configured by the S3 tenant through the Tenant Manager UI. For more information, see [Considerations for using platform services](#).

This document serves as a supplement to the [StorageGRID 11.6 Tenant Guide](#) and provides step by step instructions and examples for the endpoint and bucket configuration for search integration services. The Amazon Web Services (AWS) or on-premises Elasticsearch setup instructions included here are for basic testing or demo purposes only.

Audiences should be familiar with Grid Manager, Tenant Manager, and have access to the S3 browser to perform basic upload (PUT) and download (GET) operations for StorageGRID search integration testing.

Create tenant and enable platform services

1. Create an S3 tenant by using Grid Manager, enter a display name, and select the S3 protocol.
2. On the Permission page, select the Allow Platform Services option. Optionally, select other permissions, if necessary.



3. Set up the tenant root user initial password or, if identify federation is enabled on the grid, select which federated group has root access permission to configure the tenant account.
4. Click Sign In As Root and select Bucket: Create and Manage Buckets.

This takes you to the Tenant Manager page.

5. From Tenant Manager, select My Access Keys to create and download the S3 access key for later testing.

Search integration services with Amazon OpenSearch

Amazon OpenSearch (formerly Elasticsearch) service setup

Use this procedure for a quick and simple setup of the OpenSearch service for testing/demo purposes only. If you are using on-premises Elasticsearch for search integration services, see the section [Search integration services with on premises Elasticsearch](#).



You must have a valid AWS console login, access key, secret access key, and permission to subscribe to the OpenSearch service.

1. Create a new domain using the instructions from [AWS OpenSearch Service Getting Started](#), except for the following:
 - Step 4. Domain name: sgdemo
 - Step 10. Fine-grained access control: deselect the Enable Fine-Grained Access Control option.
 - Step 12. Access policy: select Configure Level Access Policy, select the JSON tab to modify the access policy by using the following example:
 - Replace the highlighted text with your own AWS Identity and Access Management (IAM) ID and user name.
 - Replace the highlighted text (the IP address) with the public IP address of your local computer that you used to access the AWS console.
 - Open a browser tab to <https://checkip.amazonaws.com> to find your public IP.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal":
        {"AWS": "arn:aws:iam:: nnnnnn:user/xyzabc"},
      "Action": "es:*",
      "Resource": "arn:aws:es:us-east-1:nnnnnn:domain/sgdemo/*"
    },
    {
      "Effect": "Allow",
      "Principal": {"AWS": "*"},
      "Action": [
        "es:ESHttp*"
      ],
      "Condition": {
        "IpAddress": {
          "aws:SourceIp": [ "nnn.nnn.nn.n/nn"
          ]
        }
      },
      "Resource": "arn:aws:es:us-east-1:nnnnnn:domain/sgdemo/*"
    }
  ]
}

```

Fine-grained access control

Fine-grained access control provides numerous features to help you keep your data secure. Features include document-level security, field-level security, read-only users, and OpenSearch Dashboards/Kibana tenants. Fine-grained access control requires a master user. [Learn more](#)



☐ Enable fine-grained access control

SAML authentication for OpenSearch Dashboards/Kibana

SAML authentication lets you use your existing identity provider for single sign-on for OpenSearch Dashboards/Kibana. [Learn more](#)



☐ Prepare SAML authentication

To use SAML authentication, you must first enable fine-grained access control.

Amazon Cognito authentication

Enable to use Amazon Cognito authentication for OpenSearch Dashboards/Kibana. Amazon Cognito supports a variety of identity providers for username-password authentication. [Learn more](#)



☐ Enable Amazon Cognito authentication

Access policy

Access policies control whether a request is accepted or rejected when it reaches the Amazon OpenSearch Service domain. If you specify an account, user, or role in this policy, you must sign your requests. [Learn more](#)



Domain access policy

- ☐ Only use fine-grained access control
Allow open access to the domain.
- ☐ Do not set domain level access policy
All requests to the domain will be denied.
- ☒ Configure domain level access policy

Visual editor

JSON

Import policy

Access policy

```
3+  "Statement": [  
4+  {  
5+    "Effect": "Allow",  
6+    "Principal": {  
7+      "AWS": "arn:aws:iam::226190929312:user/ashawn"  
8+    },  
9+    "Action": "es:*",  
10+   "Resource": "arn:aws:es:us-east-1:226190929312:domain/sgdemo/*"  
11+ },  
12+ {  
13+   "Effect": "Allow",  
14+   "Principal": {  
15+     "AWS": "*"   
16+   },  
17+   "Action": [  
18+     "es:ESHttp*"   
19+   ],  
20+   "Condition": {  
21+     "IpAddress": {  
22+       "aws:SourceIp": [  
23+         "216.240.240.0/24"  
24+       ]  
25+     }  
26+   },  
27+   "Resource": "arn:aws:es:us-east-1:226190929312:domain/sgdemo/*"  
28+ }
```

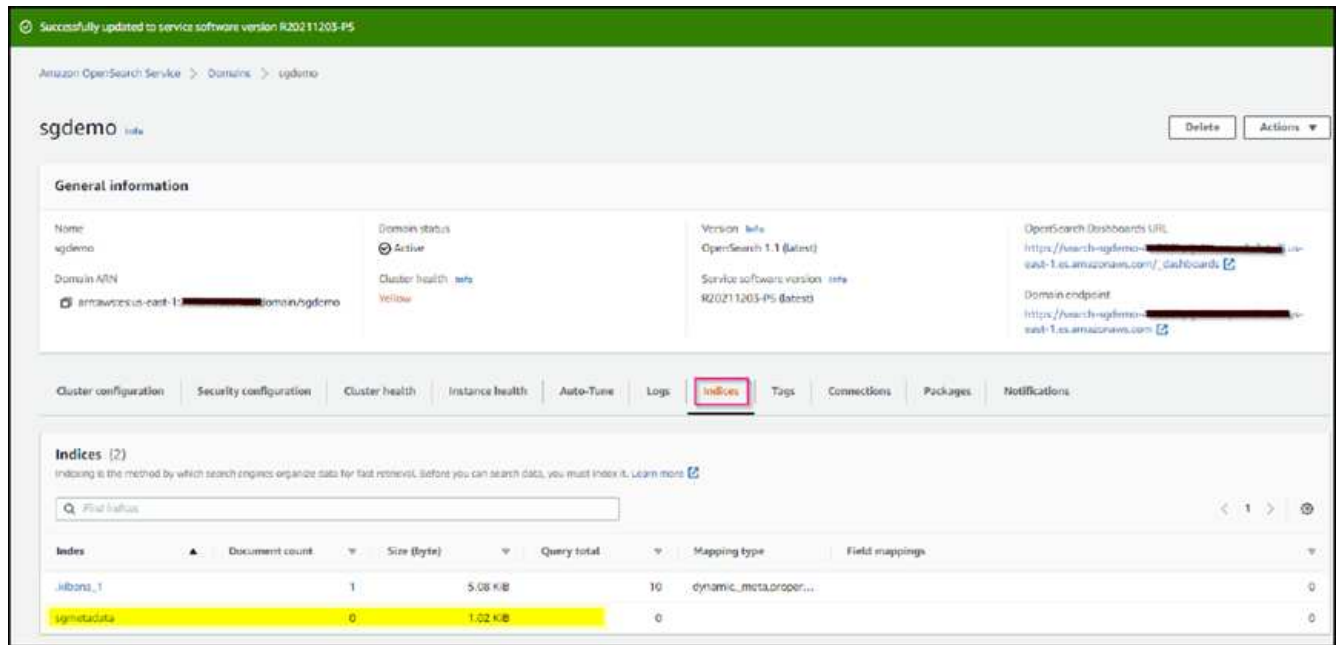

2. Wait 15 to 20 minutes for the domain to become active.



3. Click OpenSearch Dashboards URL to open the domain in a new tab to access the dashboard. If you get an access denied error, verify that the access policy source IP address is correctly set to your computer public IP to allow access to the domain dashboard.
4. On the dashboard welcome page, select Explore On Your Own. From the menu, go to Management → Dev Tools
5. Under Dev Tools → Console , enter `PUT <index>` where you use the index for storing StorageGRID object metadata. We use the index name 'sgmetadata' in the following example. Click the small triangle symbol to execute the PUT command. The expected result displays on the right panel as shown in the following example screenshot.



6. Verify that the index is visible from Amazon OpenSearch UI under sgdomain > Indices.



Platform services endpoint configuration

To configure the platform services endpoints, follow these steps:

1. In Tenant Manager, go to STORAGE(S3) > Platform services endpoints.
2. Click Create Endpoint, enter the following, and then click Continue:
 - Display name example `aws-opensearch`
 - The domain endpoint in the example screenshot under Step 2 of the preceding procedure in the URI field.
 - The domain ARN used in Step 2 of the preceding procedure in the URN field and add `/<index>/_doc` to the end of ARN.

In this example, URN becomes `arn:aws:es:us-east-1:211234567890:domain/sgdemo/sgmedata/_doc`.

Create endpoint

1

Enter details

2

Select authentication type
Optional

3

Verify server
Optional

[Cancel](#)[Continue](#)

3. To access the Amazon OpenSearch sgdomain, choose Access Key as the authentication type and then enter the Amazon S3 access key and secret key. To go the next page, click Continue.

Create endpoint

✓ Enter details

2 Select authentication type Optional

✓ Verify server Optional

Authentication type ?

Select the method used to authenticate connections to the endpoint.

Access Key

Access key ID ?

AKIA[REDACTED]UWO

Secret access key ?

.....

👁

Previous

Continue

- To verify the endpoint, select Use Operating System CA Certificate and Test and Create Endpoint. If verification is successful, an endpoint screen similar to the following figure displays. If verification fails, verify that the URN includes `/<index>/_doc` at the end of the path and the AWS access key and secret key are correct.

Platform services endpoints

A platform services endpoint stores the information StorageGRID needs to use an external resource as a target for a platform service (CloudMirror replication, notifications, or search integration). You must configure an endpoint for each platform service you plan to use.

1 endpoint Create endpoint

Delete endpoint

| <input type="checkbox"/> | Display name ? | Last error ? | Type ? | URI ? | URN ? |
|--------------------------|----------------|--------------|--------|--|--|
| <input type="checkbox"/> | aws-opensearch | | Search | https://search-sgdemo-2-2021-11-24-12-31-us-east-1.es.amazonaws.com/ | arn:aws:es:us-east-1:2[REDACTED]:domain/sgdemo/sgmetadata/_doc |

Search integration services with on premises Elasticsearch

On premises Elasticsearch setup

This procedure is for a quick setup of on premises Elasticsearch and Kibana using docker for testing purposes only. If the Elasticsearch and Kibana server already exists, go to Step 5.

1. Follow this [Docker installation procedure](#) to install docker. We use the [CentOS Docker install procedure](#) in this setup.

```
sudo yum install -y yum-utils
sudo yum-config-manager --add-repo
https://download.docker.com/linux/centos/docker-ce.repo
sudo yum install docker-ce docker-ce-cli containerd.io
sudo systemctl start docker
```

- To start docker after reboot, enter the following:

```
sudo systemctl enable docker
```

- Set the `vm.max_map_count` value to 262144:

```
sysctl -w vm.max_map_count=262144
```

- To keep the setting after reboot, enter the following:

```
echo 'vm.max_map_count=262144' >> /etc/sysctl.conf
```

2. Follow the [Elasticsearch Quick start guide](#) self-managed section to install and run the Elasticsearch and Kibana docker. In this example, we installed version 8.1.



Note down the user name/password and token created by Elasticsearch, you need these to start the Kibana UI and StorageGRID platform endpoint authentication.

Install and run Elasticsearch

1. Install and start [Docker Desktop](#).
2. Run:

```
docker network create elastic
docker pull docker.elastic.co/elasticsearch/elasticsearch:8.1.0
docker run --name es-node01 --net elastic -p 9200:9200 -p 9300:9300 -it
```

When you start Elasticsearch for the first time, the following security configuration occurs automatically:

- [Certificates and keys](#) are generated for the transport and HTTP layers.
- The Transport Layer Security (TLS) configuration settings are written to `elasticsearch.yml`.
- A password is generated for the `elastic` user.
- An enrollment token is generated for Kibana.



You might need to scroll back a bit in the terminal to view the password and enrollment token.

3. Copy the generated password and enrollment token and save them in a secure location. These values are shown only when you start Elasticsearch for the first time. You'll use these to enroll Kibana with your Elasticsearch cluster and log in.



If you need to reset the password for the `elastic` user or other built-in users, run the [elasticsearch-reset-password](#) tool. To generate new enrollment tokens for Kibana or Elasticsearch nodes, run the [elasticsearch-create-enrollment-token](#) tool. These tools are available in the Elasticsearch `bin` directory.

Install and run Kibana

To analyze, visualize, and manage Elasticsearch data using an intuitive UI, install Kibana.

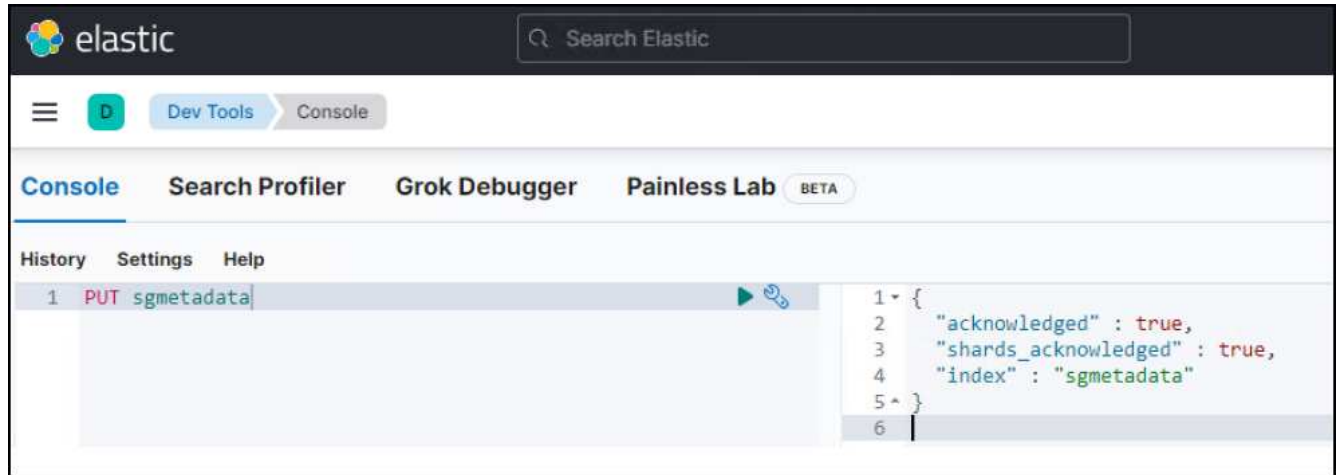
1. In a new terminal session, run:

```
docker pull docker.elastic.co/kibana/kibana:8.1.0
docker run --name kib-01 --net elastic -p 5601:5601 docker.elastic.co/k
```

When you start Kibana, a unique link is output to your terminal.

2. To access Kibana, click the generated link in your terminal.
 - a. In your browser, paste the enrollment token that you copied and click the button to connect your Kibana instance with Elasticsearch.
 - b. Log in to Kibana as the `elastic` user with the password that was generated when you started Elasticsearch.

3. After the Kibana docker container has started, the URL link `https://0.0.0.0:5601` displays in the console. Replace 0.0.0.0 with the server IP address in the URL.
4. Log in to the Kibana UI by using user name `elastic` and the password generated by Elastic in the preceding step.
5. For first time login, on the dashboard welcome page, select Explore On Your Own. From the menu, select Management > Dev Tools.
6. On the Dev Tools Console screen, enter `PUT <index>` where you use this index for storing StorageGRID object metadata. We use the index name `sgmetadata` in this example. Click the small triangle symbol to execute the PUT command. The expected result displays on the right panel as shown in the following example screenshot.



Platform services endpoint configuration

To configure endpoints for platform services, follow these steps:

1. On Tenant Manager, go to STORAGE(S3) > Platform services endpoints
2. Click Create Endpoint, enter the following, and then click Continue:
 - Display name example: `elasticsearch`
 - URI: `https://<elasticsearch-server-ip or hostname>:9200`
 - URN: `urn:<something>:es:::<some-unique-text>/<index-name>/_doc` where the index-name is the name you used on the Kibana console.
Example: `urn:local:es:::sgmd/sgmetadata/_doc`

Create endpoint

1 Enter details

2 Select authentication type
Optional

3 Verify server
Optional

Enter endpoint details

Enter the endpoint's display name, URI, and URN.

Display name ?

URI ?

URN ?

[Cancel](#)[Continue](#)

3. Select Basic HTTP as the authentication type, enter the user name `elastic` and the password generated by the Elasticsearch installation process. To go to the next page, click Continue.

Authentication type ?

Select the method used to authenticate connections to the endpoint.

Basic HTTP

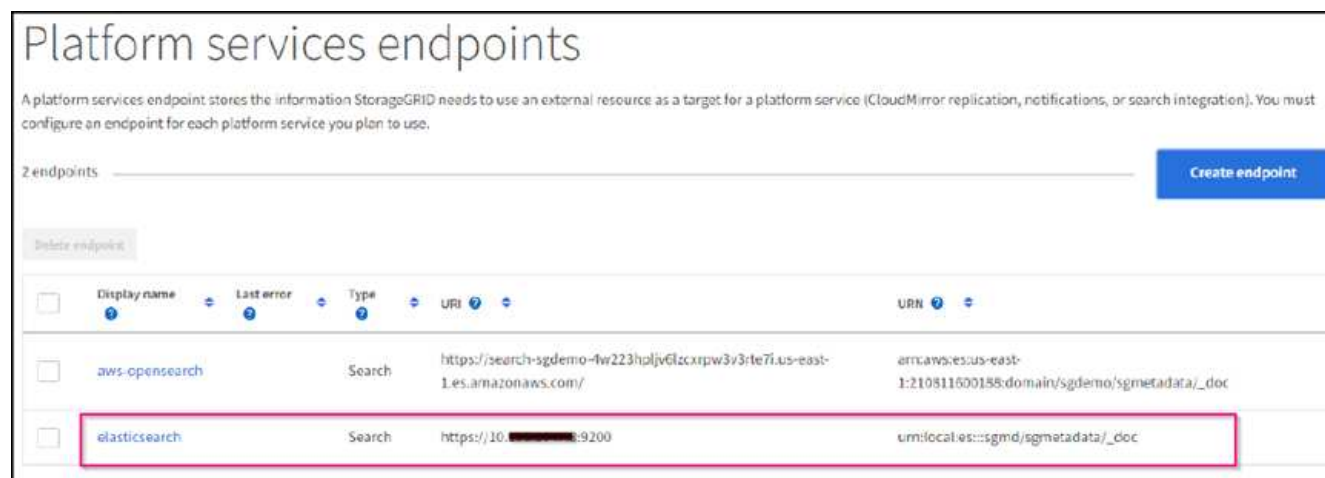
Username ?

Password ?

[Previous](#)[Continue](#)

4. Select Do Not Verify Certificate and Test and Create Endpoint to verify the endpoint. If verification is

successful, an endpoint screen similar to the following screenshot displays. If the verification fails, verify the URN, URI, and username/password entries are correct.



Bucket search integration service configuration

After the platform service endpoint is created, the next step is to configure this service at bucket level to send object metadata to the defined endpoint whenever an object is created, deleted, or its metadata or tags are updated.

You can configure search integration by using Tenant Manager to apply a custom StorageGRID configuration XML to a bucket as follows:

1. In Tenant Manager, go to STORAGE(S3) > Buckets
2. Click Create Bucket, enter the bucket name (for example, sgmetadata-test) and accept the default us-east-1 region.
3. Click Continue > Create Bucket.
4. To bring up the bucket Overview page, click the bucket name, then select Platform Services.
5. Select the Enable Search Integration dialog box. In the provided XML box, enter the configuration XML using this syntax.

The highlighted URN must match the platform services endpoint that you defined. You can open another browser tab to access the Tenant Manager and copy the URN from the defined platform services endpoint.

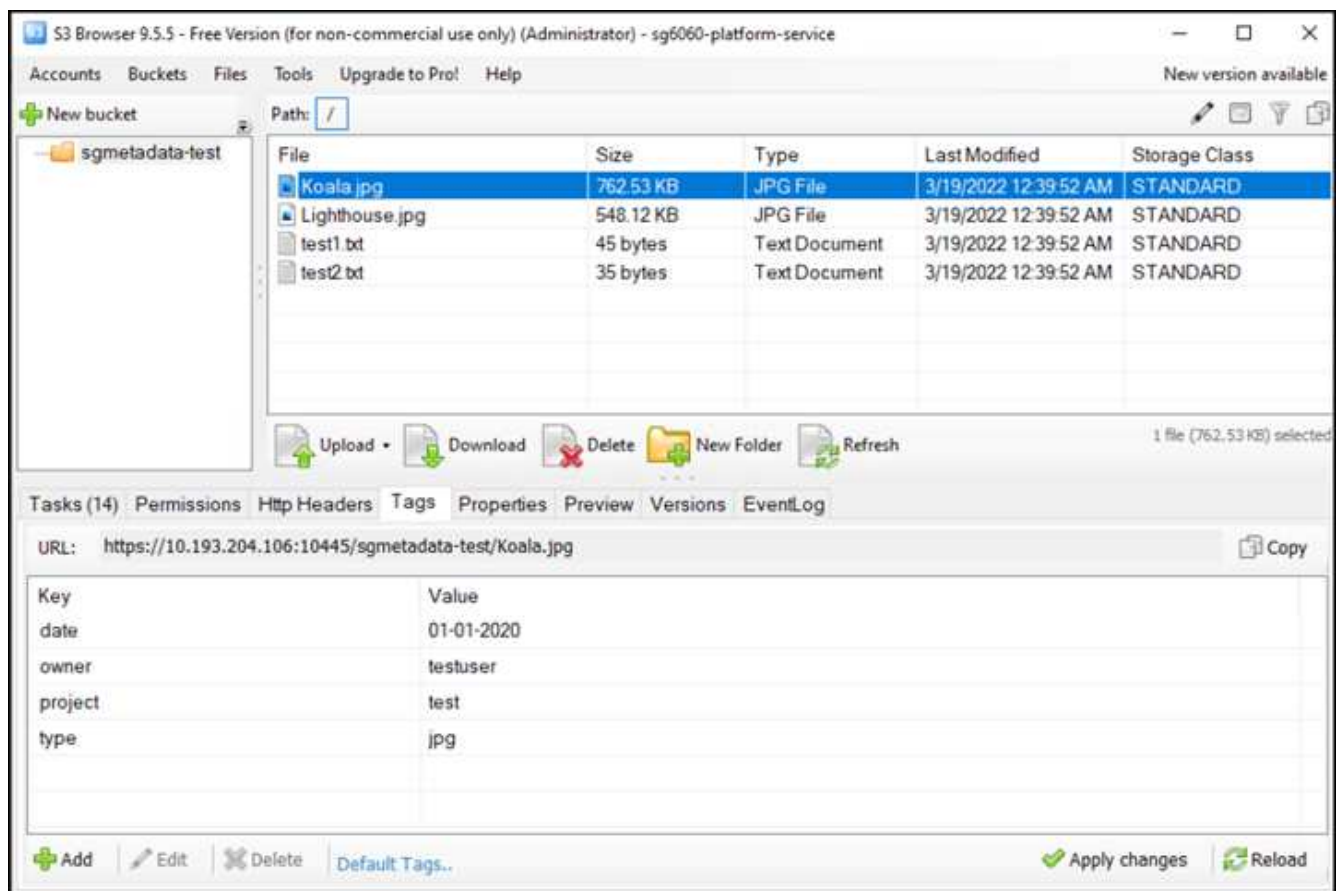
In this example, we used no prefix, meaning that the metadata for every object in this bucket is sent to the Elasticsearch endpoint defined previously.

```

<MetadataNotificationConfiguration>
  <Rule>
    <ID>Rule-1</ID>
    <Status>Enabled</Status>
    <Prefix></Prefix>
    <Destination>
      <Urn> urn:local:es:::sgmd/sgmetadata/_doc</Urn>
    </Destination>
  </Rule>
</MetadataNotificationConfiguration>

```

6. Use S3 Browser to connect to StorageGRID with the tenant access/secret key, upload test objects to sgmetadata-test bucket and add tags or custom metadata to objects.



7. Use the Kibana UI to verify that the object metadata was loaded to sgmetadata's index.
 - a. From the menu, select Management > Dev Tools.
 - b. Paste the sample query to the console panel on the left and click the triangle symbol to execute it.

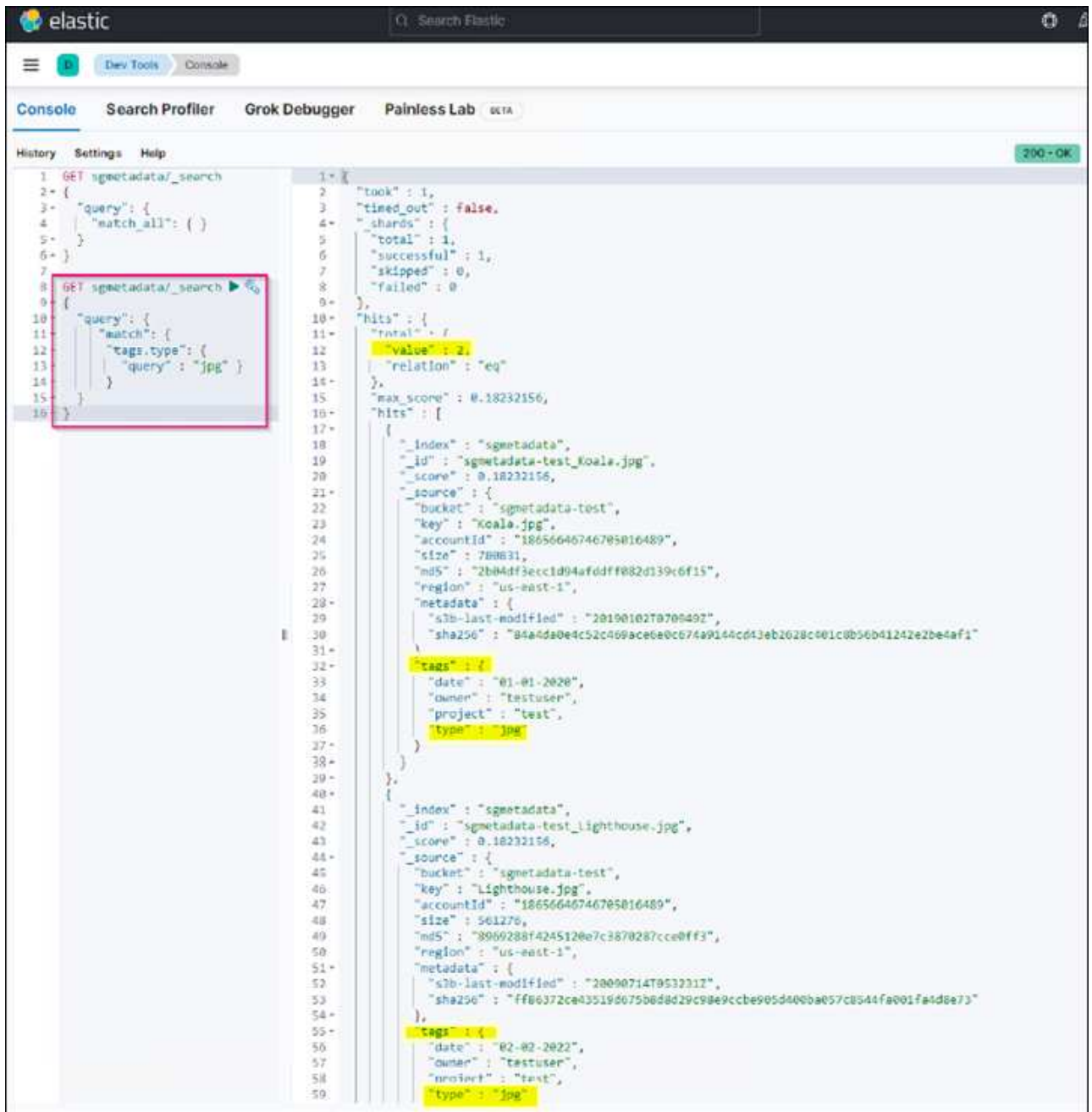
The query 1 sample result in the following example screenshot shows four records. This matches number of objects in the bucket.

```
GET sgmetadata/_search
{
  "query": {
    "match_all": { }
  }
}
```

The screenshot shows the Elastic Search Console interface. On the left, the query is entered in the console: `GET sgmetadata/_search` with a body of `{ "query": { "match_all": { } } }`. On the right, the search results are displayed as a JSON array. The first result is a document with the following fields: `_index`: "sgmetadata", `_id`: "sgmetadata-test_test1.txt", `_score`: 1.0, `_source`: { "bucket": "sgmetadata-test", "key": "test1.txt", "accountId": "18656646746705016489", "size": 45, "md5": "36b194a8ac536f09a7061f024b97211e", "region": "us-east-1", "metadata": { "s3b-last-modified": "20170429T010249Z", "sha256": "6bf95e898615852c94fa701580d9a0399487f4cbe4429e1a1d7d7f4270b10f51" }, "tags": { "owner": "testuser", "project": "test" } }. The second result is a document with the following fields: `_index`: "sgmetadata", `_id`: "sgmetadata-test_Koala.jpg", `_score`: 1.0, `_source`: { "bucket": "sgmetadata-test", "key": "Koala.jpg", "accountId": "18656646746705016489", "size": 780831, "md5": "2b04df3ecc1d94afddff082d139c6f15", "region": "us-east-1", "metadata": { "s3b-last-modified": "20190102T070949Z", "sha256": "84adda0e4c52c409ace6e0c674a9144cd43eb2628c401c8b56b41242e2be4af1" }, "tags": { "date": "01-01-2020", "owner": "testuser", "project": "test", "type": "jpg" } }.

The query 2 sample result in the following screenshot shows two records with tag type jpg.

```
GET sgmetadata/_search
{
  "query": {
    "match": {
      "tags.type": {
        "query" : "jpg" }
      }
    }
  }
}
```



The screenshot shows the Elastic Search Console interface. The left sidebar contains tabs for 'Console', 'Search Profiler', 'Grok Debugger', and 'Painless Lab'. The 'Console' tab is active, displaying a search query and its results. The query is highlighted in a red box in the left pane:

```
GET sgmetadata/_search
{
  "query": {
    "match": {
      "tags.type": {
        "query" : "jpg" }
      }
    }
  }
}
```

The right pane shows the search results in JSON format. The first result is for 'sgmetadata-test_koala.jpg' and the second is for 'sgmetadata-test_lighthouse.jpg'. Both results include metadata and tags. The 'tags' field in the results is highlighted in yellow:

```
{
  "took": 1,
  "timed_out": false,
  "shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 2,
    "value": 2,
    "relation": "eq",
    "max_score": 0.18232156,
    "hits": [
      {
        "_index": "sgmetadata",
        "_id": "sgmetadata-test_koala.jpg",
        "_score": 0.18232156,
        "_source": {
          "bucket": "sgmetadata-test",
          "key": "Koala.jpg",
          "accountId": "18656646746705016489",
          "size": 788631,
          "md5": "2b04df3ecc1d94afddff882d139c6f15",
          "region": "us-east-1",
          "metadata": {
            "s3b-last-modified": "20190102T070949Z",
            "sha256": "84a4da0e4c52c469ace0e0c674a9144cd13eb2628c001c0b56b41242e2be4af1"
          },
          "tags": {
            "date": "01-01-2020",
            "owner": "testuser",
            "project": "test",
            "type": "jpg"
          }
        }
      },
      {
        "_index": "sgmetadata",
        "_id": "sgmetadata-test_lighthouse.jpg",
        "_score": 0.18232156,
        "_source": {
          "bucket": "sgmetadata-test",
          "key": "lighthouse.jpg",
          "accountId": "18656646746705016489",
          "size": 561276,
          "md5": "8969288f4245120e7c3870287cce0ff3",
          "region": "us-east-1",
          "metadata": {
            "s3b-last-modified": "20090714T053231Z",
            "sha256": "ff06372ce43519d075b0d8d29c98e9ccbe965d400ba057c0544fa001fa4d8e73"
          },
          "tags": {
            "date": "02-02-2022",
            "owner": "testuser",
            "project": "test",
            "type": "jpg"
          }
        }
      }
    ]
  }
}
```

Where to find additional information

To learn more about the information that is described in this document, review the following documents and/or websites:

- [What are platform services](#)
- [StorageGRID 11.6 Documentation](#)

By Angela Cheng

Node Clone

Node clone considerations and performance.

Node clone considerations

Node clone can be a faster method for replacing existing appliance nodes for a tech refresh, increase capacity, or increase performance of your StorageGRID system. Node clone can also be useful for converting to node encryption with a KMS, or changing a storage node from DDP8 to DDP16.

- The used capacity of the source node is not relevant to the time required for the clone process to complete. Node clone is a full copy of the node including free space in the node.
- The source and destination appliances must be at the same PGE version
- The destination node must always have larger capacity than the source
 - Make sure the new destination appliance has a larger drive size than the source
 - If the destination appliance has the same size drives and is configured for DDP8, you can configure the destination for DDP16. If the source is already configured for DDP16 then node clone will not be possible.
 - When going from SG5660 or SG5760 appliances to SG6060 appliances be aware that the SG5x60's have 60 capacity drives where the SG6060 only has 58.
- The node clone process requires the source node to be offline to the grid for the duration of the cloning process. If an additional node goes offline during this time client services may be impacted.
- A storage node can only be offline for 15 days. If the cloning process estimate is close to 15 days or will exceed 15 days, use the expansion and decommission procedures.
- For a SG6060 with expansion shelves, you need to add the time for the correct shelf drive size to the time of the base appliance time to get the full clone duration.

Node clone Performance estimates

The following tables contain calculated estimates for node clone duration. Conditions vary so, entries in **BOLD** may risk exceeding the 15 day limit for a node down.

DDP8

SG5612 → Any

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 1 Day | 2 Days | 2.5 Days | 3 Days | 4 Days | 4.5 Days |
| 25GB | 1 Day | 2 Days | 2.5 Days | 3 Days | 4 Days | 4.5 Days |

SG5712 → Any

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 1 Day | 2 Days | 2.5 Days | 3 Days | 4 Days | 4.5 Days |
| 25GB | 1 Day | 2 Days | 2.5 Days | 3 Days | 4 Days | 4.5 Days |

SG5660 → SG5760

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 3 Day | 6 Days | 7 Days | 8.5 Days | 11.5 Days | 13 Days |
| 25GB | 3 Day | 6 Days | 7 Days | 8.5 Days | 11.5 Days | 13 Days |

SG5660 → SG6060

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 2.5 Day | 4.5 Days | 5.5 Days | 6.5 Days | 9 Days | 10 Days |
| 25GB | 2 Day | 4 Days | 5 Days | 6 Days | 8 Days | 9 Days |

SG5760 → SG5760

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 3 Day | 6 Days | 7 Days | 8.5 Days | 11.5 Days | 13 Days |
| 25GB | 3 Day | 6 Days | 7 Days | 8.5 Days | 11.5 Days | 13 Days |

SG5760 → SG6060

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 2.5 Day | 4.5 Days | 5.5 Days | 6.5 Days | 9 Days | 10 Days |
| 25GB | 1.5 Day | 3 Days | 3.5 Days | 4.5 Days | 6 Days | 6.5 Days |

SG6060 → SG6060

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 2.5 Day | 4.5 Days | 5.5 Days | 6.5 Days | 8.5 Days | 9.5 Days |
| 25GB | 1.5 Day | 3 Days | 3.5 Days | 4 Days | 5.5 Days | 6 Days |

DDP16

SG5760 → SG5760

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 3.5 Day | 6.5 Days | 8 Days | 9.5 Days | 12.5 Days | 14 Days |
| 25GB | 3.5 Day | 6.5 Days | 8 Days | 9.5 Days | 12.5 Days | 14 Days |

SG5760 → SG6060

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 2.5 Day | 5 Days | 6 Days | 7.5 Days | 10 Days | 11 Days |
| 25GB | 2 Day | 3.5 Days | 4 Days | 5 Days | 6.5 Days | 7 Days |

SG6060 → SG6060

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 3.5 Day | 5 Days | 6 Days | 7 Days | 9.5 Days | 10.5 Days |
| 25GB | 2 Day | 3 Days | 4 Days | 4.5 Days | 6 Days | 7 Days |

Expansion shelf (add to above SG6060 for each shelf on source appliance)

| Network Interface speed | 4TB Drive size | 8TB Drive size | 10TB Drive size | 12TB Drive size | 16TB Drive size | 18TB Drive size |
|-------------------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 10GB | 3.5 Day | 5 Days | 6 Days | 7 Days | 9.5 Days | 10.5 Days |
| 25GB | 2 Day | 3 Days | 4 Days | 4.5 Days | 6 Days | 7 Days |

By Aron Klein

How to use port remap

You may have a need to remap an incoming or outbound port for multiple reasons. You may be moving from the legacy CLB load balancer service to the current nginx service load balancer endpoint and maintain the same port to reduce the impact to clients, wish to use port 443 for client S3 on an admin node client network, or for firewall restrictions.

Migrate S3 clients from CLB to NGINX with Port ReMap

In releases earlier than StorageGRID 11.3, the included Load Balancer service on the Gateway Nodes is the Connection Load Balancer (CLB). In StorageGRID 11.3, NetApp introduces the NGINX service as a feature rich integrated solution for load balancing HTTP(s) traffic. Because the CLB service remains available in the current release of StorageGRID, you cannot reuse port 8082 in the new load balancer endpoint configuration. To work around this, the 8082 inbound port is remapped to 10443. This makes all HTTPS requests coming into port 8082 on the gateway redirect to port 10443, bypassing the CLB service and instead connecting to the NGINX service. Although the following instructions are for VMware, the PORT_REMAP functionality exists for all installation methods, and you can use a similar process for bare metal deployments and appliances.

VMware virtual machine Gateway Node deployment

The following steps are for a StorageGRID deployment where the Gateway Node or Nodes are deployed in VMware vSphere 7 as VMs using the StorageGRID Open Virtualization Format (OVF). The process entails destructively removing the VM and redeploying the VM with the same name and configuration. Before you power on the VM, change the vAPP property to remap the port, then power on the VM and follow the node recovery process.

Prerequisites

- You are running StorageGRID 11.3 or later
- You have downloaded and have access to the installed StorageGRID version VMware install files.
- You have a vCenter account with permissions to power on/off VMs, change the settings of the VMs and vApps, remove VMs from vCenter, and deploy VMs by OVF.
- You have created a load balancer endpoint
 - The port is configured to the desired redirect port
 - The endpoint SSL certificate is the same as installed for the CLB service in the Configuration/Server Certificates/ Object Storage API Service Endpoints Server Certificate or the client is able to accept a change in certificate.



If your existing certificate is self-signed, you cannot reuse it in the new endpoint. You must generate a new self-signed certificate when creating the endpoint and configure the clients to accept the new certificate.

Destroy the first Gateway Node

To destroy the first Gateway Node, follow these steps:

1. Choose the Gateway Node to start with if the grid contains more than one.
2. Remove the node IPs from all DNS round-robin entities or load balancer pools, if applicable.
3. Wait for Time-to-Live (TTL) and open sessions to expire.
4. Power off the VM node.
5. Remove the VM node from the disk.

Deploy the replacement Gateway Node

To deploy the replacement Gateway Node, follow these steps:

1. Deploy the new VM from OVF, selecting the .ovf, .mf, and .vmdk files from the install package downloaded from the support site:
 - vsphere-gateway.mf
 - vsphere-gateway.ovf
 - NetApp-SG-11.4.0-20200721.1338.d3969b3.vmdk
2. After the VM has been deployed, select it from the list of VMs, select the Configure tab vApp Options.

The screenshot shows the vSphere VM configuration interface. The 'Configure' tab is selected, and the 'vApp Options' sub-tab is active. The left sidebar lists various settings, with 'vApp Options' highlighted. The main content area shows 'OVF Settings' with a 'VIEW OVF ENVIRONMENT' button and an information icon. Below this, there are two rows of settings: 'OVF environment transport' with a value of 'VMware Tools', and 'Installation boot' with a value of 'Disabled'. At the bottom, there is a 'Properties' section with buttons for 'ADD', 'EDIT', 'SET VALUE', and 'DELETE'.

3. Scroll down to the Properties section and select the PORT_REMAP_INBOUND property

| Summary | Monitor | Configure | Permissions | Datastores | Networks | Snapshots | Updates | | | |
|---|---------|----------------------------------|-----------------------|--------------------------------------|----------------|-----------|----------------|---------------------|-----------------------|---|
| <div>Settings</div> <div>VM SDRS Rules</div> <div>vApp Options</div> <div>Alarm Definitions</div> <div>Scheduled Tasks</div> <div>Policies</div> <div>Guest User Mappings</div> | | <input type="radio"/> | ADMIN_IP | Primary Admin IP | 10.193.204.110 | | 0.0.0.0 | Grid Network (eth0) | ip | |
| | | <input type="radio"/> | ADMIN_NETWORK_ESL | Admin network external subnet list | | | | | Admin Network (eth1) | string |
| | | <input type="radio"/> | ADMIN_NETWORK_IP | Admin network IP | 10.193.174.112 | | 0.0.0.0 | | Admin Network (eth1) | ip |
| | | <input type="radio"/> | NODE_TYPE | Node type | | | VM_API_Gateway | | Grid Node Parameters | string["VM_Storage_Node", "VM_min_Node", "VM_API_Gateway", "_Archive_Node"] |
| | | <input type="radio"/> | CLIENT_NETWORK_CONFIG | Client network IP configuration | STATIC | | DISABLED | | Client Network (eth2) | string["DISABLED", "STATIC", "DHCP"] |
| | | <input checked="" type="radio"/> | PORT_REMAP_INBOUND | Inbound port remapping specification | | | | | Advanced | string |
| | | <input type="radio"/> | GRID_NETWORK | Grid network IP configuration | STATIC | | STATIC | | Grid Network | string["STATIC", "DHCP"] |

4. Scroll to the top of the Properties list and click Edit



5. Select the Type tab, confirm that the User Configurable checkbox is selected, and then click Save.

Edit property
Inbound port remapping specification... X

General
Type

☒ Static property

Type
String

User configurable
☒

Length
0 65535

Default value

☐ Dynamic property

Macro
IP address

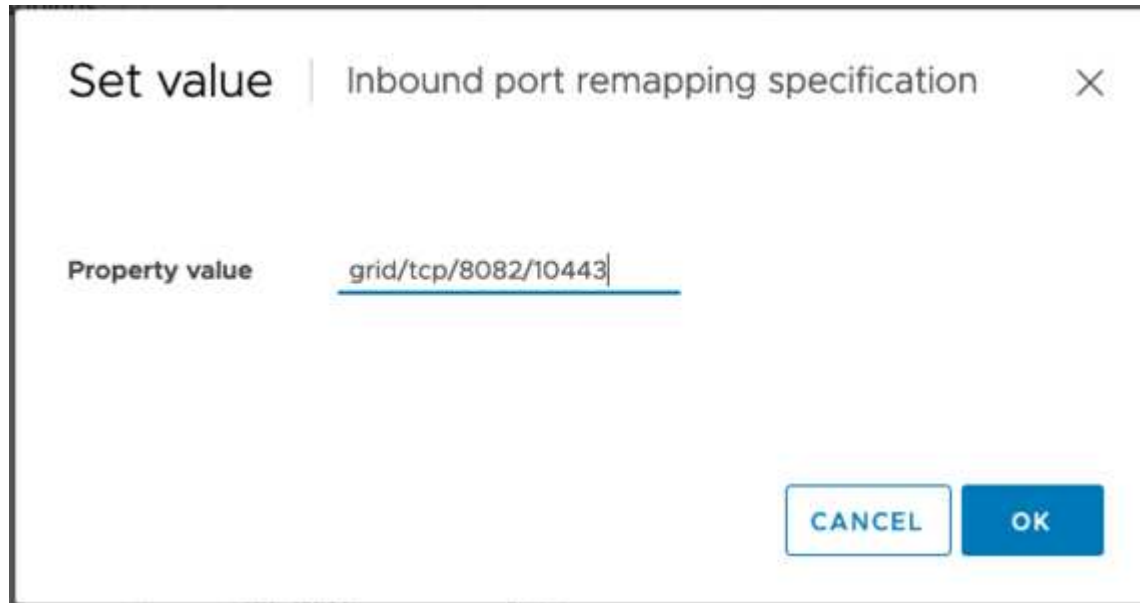
Network
MGMT_564

CANCEL
SAVE

- At the top of the Properties list, with the “PORT_REMAP_INBOUND” property still selected, click Set Value.



- In the Property Value field, enter the network (grid, admin, or client), TCP, the original port (8082), and the new port (10443) with “/” in between each value as depicted following.

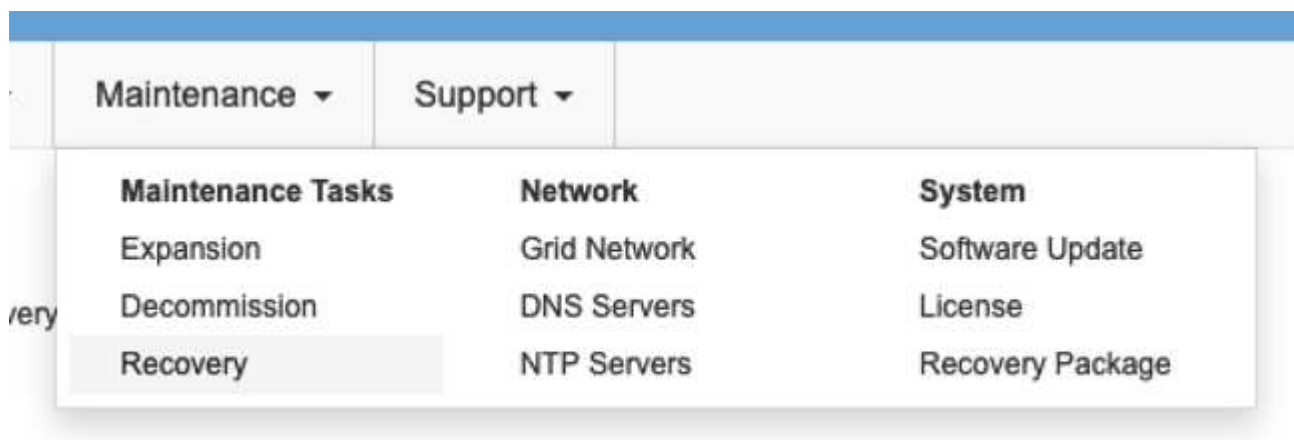


- If you are using multiple networks, use a comma (,) to separate the network strings, for example, grid/tcp/8082/10443,admin/tcp/8082/10443,client/tcp/8082/10443

Recover the Gateway Node

To recover the Gateway Node, follow these steps:

- Navigate to the Maintenance/Recovery section of the Grid Management UI.



2. Power on the VM node and wait for the node to appear in the Maintenance/Recovery Pending Nodes section of the Grid Management UI.

Recovery

Select the failed grid node to recover, enter your provisioning passphrase, and then click Start Recovery to begin the recovery procedure.

Pending Nodes

| Name | IPv4 Address | State | Recoverable |
|-------------------|--------------|-------|-------------|
| No results found. | | | |



For information and directions for node recovery, see the <https://docs.netapp.com/sgws-114/topic/com.netapp.doc.sg-maint/GUID-7E22B1B9-4169-4800-8727-75F25FC0FFB1.html> [Recovery and Maintenance guide]

3. After the node has been recovered, the IP can be included in all DNS round-robin entities, or load balancer pools, if applicable.

Now, any HTTPS sessions on port 8082 go to port 10443

Remap port 443 for client S3 access on an Admin node

The default configuration in the StorageGRID system for an admin node, or HA group containing an Admin node is for port 443 and 80 to be reserved for the management and tenant manager UI's and cannot be used for load balancer endpoints. The solution to this is to use the port remap feature and redirect inbound port 443 to a new port that will be configured as a load balancer endpoint. Once this completed Client S3 traffic will be able to use port 443, the Grid management UI will only be accessible through port 8443, and the Tenant management UI will only be accessible on port 9443. The remap port feature can only be configured at install time of the node. In order to implement a port remap of an active node in the grid, it must be reset to the pre-installed state. This is a destructive procedure that includes a node recovery once the configuration change has been made.

Backup logs and databases

Admin nodes contain audit logs, prometheus metrics, as well as historical information about attributes, alarms, and alerts. Having multiple admin nodes means you have multiple copies of this data. If you do not have multiple admin nodes in your grid, you should make sure to preserve this data to restore after the node has been recovered in the end of this process. If you have another admin node in your grid, you can copy the data from that node during the recovery process. If you do not have another admin node in the grid you can follow these instructions to copy the data before destroying the node.

Copy audit logs

1. Log in to the Admin Node:
 - a. Enter the following command: `ssh admin@grid_node_IP`

- b. Enter the password listed in the `Passwords.txt` file.
- c. Enter the following command to switch to root: `su -`
- d. Enter the password listed in the `Passwords.txt` file.
- e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
- f. Enter the SSH Access Password listed in the `Passwords.txt` file.

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

2. Create the directory to copy all audit log files to a temporary location on a separate grid node lets use `storage_node_01`:
 - a. `ssh admin@storage_node_01_IP`
 - b. `mkdir -p /var/local/tmp/saved-audit-logs`
3. Back on the admin node, stop the AMS service to prevent it from creating a new log file: `service ams stop`
4. Rename the audit.log file so that it does not overwrite the existing file when you copy it to the recovered Admin Node.
 - a. Rename audit.log to a unique numbered file name such as yyyy-mm-dd.txt.1. For example, you can rename the audit log file to 2015-10-25.txt.1

```
cd /var/local/audit/export
ls -l
mv audit.log 2015-10-25.txt.1
```

5. Restart the AMS service: `service ams start`
6. Copy all audit log files: `scp * admin@storage_node_01_IP:/var/local/tmp/saved-audit-logs`

Copy Prometheus data



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

1. Create the directory to copy the prometheus data to a temporary location on a separate grid node, again we will user `storage_node_01`:
 - a. Log in to the storage node:
 - i. Enter the following command: `ssh admin@storage_node_01_IP`
 - ii. Enter the password listed in the `Passwords.txt` file.
 - iii. `mkdir -p /var/local/tmp/prometheus``
2. Log in to the Admin Node:
 - a. Enter the following command: `ssh admin@admin_node_IP`

- b. Enter the password listed in the `Passwords.txt` file.
- c. Enter the following command to switch to root: `su -`
- d. Enter the password listed in the `Passwords.txt` file.
- e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
- f. Enter the SSH Access Password listed in the `Passwords.txt` file.

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

3. From the Admin Node, stop the Prometheus service: `service prometheus stop`
 - a. Copy the Prometheus database from the source Admin Node to the storage node backup location
Node: `/rsync -azh --stats "/var/local/mysql_ibdata/prometheus/data"`
`"storage_node_01_IP:/var/local/tmp/prometheus/"`
4. Restart the Prometheus service on the source Admin Node. `service prometheus start`

Backup historical information

The historical information is stored in a mysql database. In order to dump a copy of the database you will need the user and password from NetApp. If you have another admin node in the grid, this step is not necessary and the database can be cloned from a remaining admin node during the recovery process.

1. Log in to the Admin Node:
 - a. Enter the following command: `ssh admin@admin_node_IP`
 - b. Enter the password listed in the `Passwords.txt` file.
 - c. Enter the following command to switch to root: `su -`
 - d. Enter the password listed in the `Passwords.txt` file.
 - e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
 - f. Enter the SSH Access Password listed in the `Passwords.txt` file.

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

2. Stop StorageGRID services on Admin Node and startup ntp and mysql
 - a. Stop all services: `service servermanager stop`
 - b. restart ntp service: `service ntp start`
 - ..restart mysql service: `service mysql start`
3. Dump mi database to `/var/local/tmp`
 - a. enter the following command: `mysqldump -u username -p password mi > /var/local/tmp/mysql-mi.sql`
4. Copy the mysql dump file to an alternate node, we will use `storage_node_01`:
`scp /var/local/tmp/mysql-mi.sql _storage_node_01_IP:/var/local/tmp/mysql-mi.sql`

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

Rebuild the Admin node

Now that you have a backup copy of all desired data and logs either on another admin node in the grid or stored in a temporary location it is time to reset the appliance so the port remap can be configured.

1. Resetting an appliance returns it to the pre-installed state where it only retains the host name, IP's and network configurations. All data will be lost which is why we made sure to have a backup of any important information.
 - a. enter the following command: `sgareinstall`

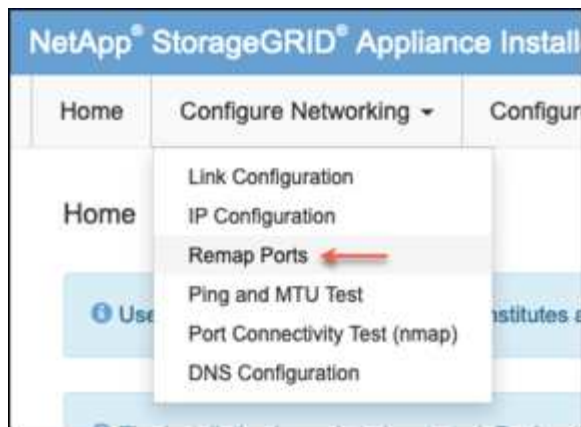
```
root@sg100-01:~ # sgareinstall
WARNING: All StorageGRID Webscale services on this node will be shut
down.
WARNING: Data stored on this node may be lost.
WARNING: You will have to reinstall StorageGRID Webscale to this
node.

After running this command and waiting a few minutes for the node to
reboot,
browse to one of the following URLs to reinstall StorageGRID Webscale
on
this node:

https://10.193.174.192:8443
https://10.193.204.192:8443
https://169.254.0.1:8443

Are you sure you want to continue (y/n)? y
Renaming SG installation flag file.
Initiating a reboot to trigger the StorageGRID Webscale appliance
installation wizard.
```

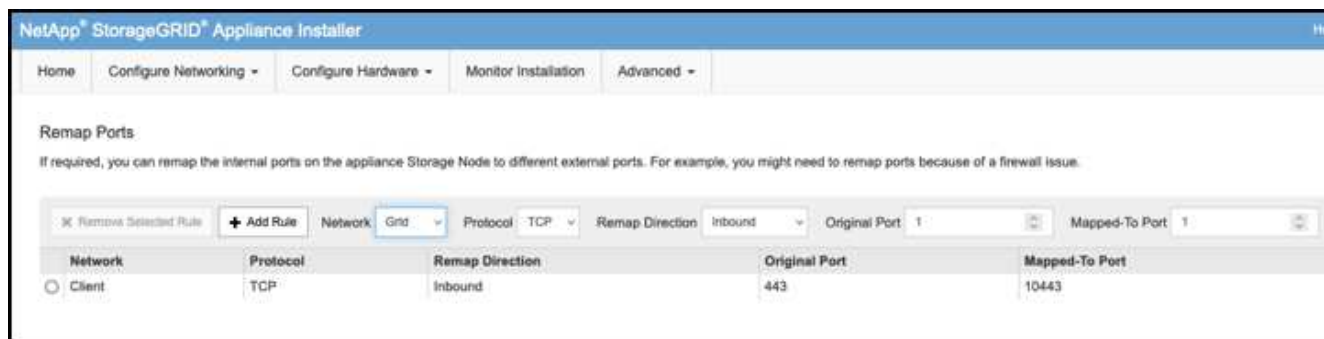
2. After some time has passed the appliance will reboot and you will be able to access the node PGE UI.
3. Browse to the Configure Networking



4. Select the desired network, protocol, direction and ports then click the Add Rule button.



Remap of inbound port 443 on the GRID network will break install, and expansion procedures. It is not recommended to remap port 443 on the GRID network.



5. Once the desired port remaps have been added, you can return to the home tab and click on the Start Installation button.

You can now follow the Admin node recovery procedures in the [product documentation](#)

Restore Databases and logs

Now that the admin node has been recovered, you can restore the metrics, logs, and historical information. If you have another admin node in the grid, follow the [product documentation](#) utilizing the *prometheus-clone-db.sh* and *mi-clone-db.sh* scripts. If this is your only admin node and you chose to backup this data, you can follow the below steps to restore the information.

Copy audit logs back

1. Log in to the Admin Node:
 - a. Enter the following command: `ssh admin@grid_node_IP`
 - b. Enter the password listed in the `Passwords.txt` file.
 - c. Enter the following command to switch to root: `su -`
 - d. Enter the password listed in the `Passwords.txt` file.
 - e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
 - f. Enter the SSH Access Password listed in the `Passwords.txt` file.

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

2. Copy the preserved audit log files to the recovered Admin Node: `scp admin@grid_node_IP:/var/local/tmp/saved-audit-logs/YYYY* .`
3. For security, delete the audit logs from the failed grid node after verifying that they have been copied successfully to the recovered Admin Node.
4. Update the user and group settings of the audit log files on the recovered Admin Node: `chown ams-user:bycast *`

You must also restore any pre-existing client access to the audit share. For more information, see the instructions for administering StorageGRID.

Restore Prometheus metrics



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

1. Log in to the Admin Node:
 - a. Enter the following command: `ssh admin@grid_node_IP`
 - b. Enter the password listed in the `Passwords.txt` file.
 - c. Enter the following command to switch to root: `su -`
 - d. Enter the password listed in the `Passwords.txt` file.
 - e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
 - f. Enter the SSH Access Password listed in the `Passwords.txt` file.

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

2. From the Admin Node, stop the Prometheus service: `service prometheus stop`
 - a. Copy the Prometheus database from the temporary backup location to the admin node: `/rsync -azh --stats "backup_node:/var/local/tmp/prometheus/" "/var/local/mysql_ibdata/prometheus/"`
 - b. verify the data is in the correct path and is complete `ls /var/local/mysql_ibdata/prometheus/data/`
3. Restart the Prometheus service on the source Admin Node. `service prometheus start`

Restore historical information

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

- d. Enter the password listed in the `Passwords.txt` file.
- e. Add the SSH private key to the SSH agent. Enter: `ssh-add`
- f. Enter the SSH Access Password listed in the `Passwords.txt` file.

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

2. Copy the mysql dump file from the alternate node: `scp grid_node_IP_:/var/local/tmp/mysql-mi.sql /var/local/tmp/mysql-mi.sql`
3. Stop StorageGRID services on Admin Node and startup ntp and mysql
 - a. Stop all services: `service servermanager stop`
 - b. restart ntp service: `service ntp start`
..restart mysql service: `service mysql start`
4. Drop the mi database and create a new empty database: `mysql -u username -p password -A mi -e "drop database mi; create database mi;"`
5. restore the mysql database from the database dump: `mysql -u username -p password -A mi < /var/local/tmp/mysql-mi.sql`
6. Restart all other services `service servermanager start`

By Aron Klein

Tool and application guides

Use Cloudera Hadoop S3A connector with StorageGRID

Hadoop has been a favorite of data scientists for some time now. Hadoop allows for the distributed processing of large data sets across clusters of computers using simple programming frameworks. Hadoop was designed to scale up from single servers to thousands of machines, with each machine possessing local compute and storage.

Why use S3A for Hadoop workflows?

As the volume of data has grown over time, the approach of adding new machines with their own compute and storage has become inefficient. Scaling linearly creates challenges for using resources efficiently and managing the infrastructure.

To address these challenges, the Hadoop S3A client offers high-performance I/O against S3 object storage. Implementing a Hadoop workflow with S3A helps you leverage object storage as a data repository and enables you to separate compute and storage, which in turn enables you to scale compute and storage independently. Decoupling compute and storage also enables you to dedicate the right amount of resources for your compute jobs and provide capacity based on the size of your data set. Therefore, you can reduce your overall TCO for Hadoop workflows.

Configure S3A connector to use StorageGRID

Prerequisites

- A StorageGRID S3 endpoint URL, a tenant s3 access key, and a secret key for Hadoop S3A connection testing.
- A Cloudera cluster and root or sudo permission to each host in the cluster to install the Java package.

As of April 2022, Java 11.0.14 with Cloudera 7.1.7 was tested against StorageGRID 11.5 and 11.6. However, the Java version number might be different at the time of a new install.

Install Java package

1. Check the [Cloudera support matrix](#) for the supported JDK version.
2. Download the [Java 11.x package](#) that matches the Cloudera cluster operating system. Copy this package to each host in the cluster. In this example, the rpm package is used for CentOS.
3. Log into each host as root or using an account with sudo permission. Perform the following steps on each host:
 - a. Install the package:

```
$ sudo rpm -Uvh jdk-11.0.14_linux-x64_bin.rpm
```

- b. Check where Java is installed. If multiple versions are installed, set the newly installed version as default:

```
alternatives --config java
```

There are 2 programs which provide 'java'.

| Selection | Command |
|-----------|---------------------------------------|
| +1 | /usr/java/jre1.8.0_291-amd64/bin/java |
| 2 | /usr/java/jdk-11.0.14/bin/java |

Enter to keep the current selection[+], or type selection number: 2

c. Add this line to the end of /etc/profile. The path should match the path of above selection:

```
export JAVA_HOME=/usr/java/jdk-11.0.14
```

d. Run the following command for the profile to take effect:

```
source /etc/profile
```


Cloudera HDFS S3A configuration



Steps











1. From the Cloudera Manager GUI, select Clusters > HDFS, and select Configuration.
2. Under CATEGORY, select Advanced, and scroll down to locate Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml.
3. Click the (+) sign and add following value pairs.

| Name | Value |
|-------------------------------|--|
| fs.s3a.access.key | <tenant s3 access key from StorageGRID> |
| fs.s3a.secret.key | <tenant s3 secret key from StorageGRID> |
| fs.s3a.connection.ssl.enabled | [true or false] (default is https if this entry is missing) |
| fs.s3a.endpoint | <StorageGRID S3 endpoint:port> |
| fs.s3a.impl | org.apache.hadoop.fs.s3a.S3AFileSystem |
| fs.s3a.path.style.access | [true or false] (default is virtual host style if this entry is missing) |

Sample screenshot

Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml 

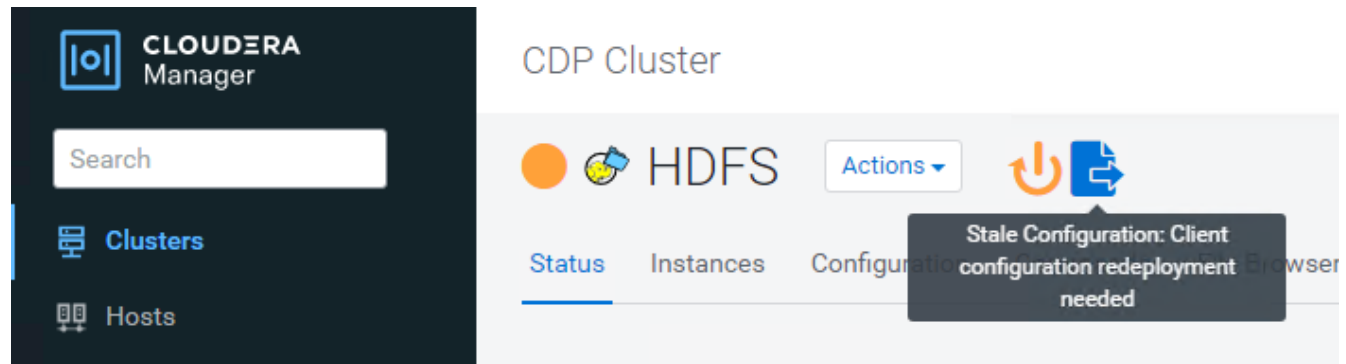
HDFS (Service-Wide)  Undo View as XML 

| | | |
|-------------|---|---|
| Name | fs.s3a.endpoint |   |
| Value | sgdemo.netapp.com:10443 | |
| Description | StorageGRID s3 load balancer endpoint | |
| | <input checked="" type="checkbox"/> Final | |
| Name | fs.s3a.access.key |   |
| Value | OMC[REDACTED]BAN | |
| Description | SG CDP S3 access key | |
| | <input checked="" type="checkbox"/> Final | |
| Name | fs.s3a.secret.key |   |
| Value | mapz[REDACTED]Qfc | |
| Description | SG CDP S3 secret key | |
| | <input checked="" type="checkbox"/> Final | |
| Name | fs.s3a.impl |   |
| Value | org.apache.hadoop.fs.s3a.S3AFileSystem | |
| Description | | |
| | <input checked="" type="checkbox"/> Final | |
| Name | fs.s3a.path.style.access |   |
| Value | true | |
| Description | | |
| | <input checked="" type="checkbox"/> Final | |

Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml Save Changes(CTRL+S)

4. Click the Save Changes button. Select the Stale Configuration icon from the HDFS menu bar, select

Restart Stale Services on the next page, and select Restart Now.



Test S3A connection to StorageGRID

Perform basic connection test

Log into one of the hosts in the Cloudera cluster, and enter `hadoop fs -ls s3a://<bucket-name>/`.

The following example uses path syle with a pre-existing hdfs-test bucket and a test object.

```
[root@ce-n1 ~]# hadoop fs -ls s3a://hdfs-test/
22/02/15 18:24:37 WARN impl.MetricsConfig: Cannot locate configuration:
tried hadoop-metrics2-s3a-file-system.properties,hadoop-
metrics2.properties
22/02/15 18:24:37 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot
period at 10 second(s).
22/02/15 18:24:37 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system started
22/02/15 18:24:37 INFO Configuration.deprecation: No unit for
fs.s3a.connection.request.timeout(0) assuming SECONDS
Found 1 items
-rw-rw-rw-    1 root root      1679 2022-02-14 16:03 s3a://hdfs-test/test
22/02/15 18:24:38 INFO impl.MetricsSystemImpl: Stopping s3a-file-system
metrics system...
22/02/15 18:24:38 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system stopped.
22/02/15 18:24:38 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system shutdown complete.
```

Troubleshooting

Scenario 1

Use an HTTPS connection to StorageGRID and get a `handshake_failure` error after a 15 minute timeout.

Reason: Old JRE/JDK version using outdated or unsupported TLS cipher suite for connection to StorageGRID.

Sample error message

```
[root@ce-n1 ~]# hadoop fs -ls s3a://hdfs-test/
22/02/15 18:52:34 WARN impl.MetricsConfig: Cannot locate configuration:
tried hadoop-metrics2-s3a-file-system.properties,hadoop-
metrics2.properties
22/02/15 18:52:34 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot
period at 10 second(s).
22/02/15 18:52:34 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system started
22/02/15 18:52:35 INFO Configuration.deprecation: No unit for
fs.s3a.connection.request.timeout(0) assuming SECONDS
22/02/15 19:04:51 INFO impl.MetricsSystemImpl: Stopping s3a-file-system
metrics system...
22/02/15 19:04:51 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system stopped.
22/02/15 19:04:51 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system shutdown complete.
22/02/15 19:04:51 WARN fs.FileSystem: Failed to initialize filesystem
s3a://hdfs-test/: org.apache.hadoop.fs.s3a.AWSClietIOException:
doesBucketExistV2 on hdfs: com.amazonaws.SdkClientException: Unable to
execute HTTP request: Received fatal alert: handshake_failure: Unable to
execute HTTP request: Received fatal alert: handshake_failure
ls: doesBucketExistV2 on hdfs: com.amazonaws.SdkClientException: Unable to
execute HTTP request: Received fatal alert: handshake_failure: Unable to
execute HTTP request: Received fatal alert: handshake_failure
```

Resolution: Make sure that JDK 11.x or later is installed and set to default the Java library. Refer to the [Install Java package](#) section for more information.

Scenario 2:

Failed to connect to StorageGRID with error message Unable to find valid certification path to requested target.

Reason: StorageGRID S3 endpoint server certificate is not trusted by Java program.

Sample error message:

```
[root@hdp6 ~]# hadoop fs -ls s3a://hdfs-test/
22/03/11 20:58:12 WARN impl.MetricsConfig: Cannot locate configuration:
tried hadoop-metrics2-s3a-file-system.properties,hadoop-
metrics2.properties
22/03/11 20:58:13 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot
period at 10 second(s).
22/03/11 20:58:13 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system started
22/03/11 20:58:13 INFO Configuration.deprecation: No unit for
fs.s3a.connection.request.timeout(0) assuming SECONDS
22/03/11 21:12:25 INFO impl.MetricsSystemImpl: Stopping s3a-file-system
metrics system...
22/03/11 21:12:25 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system stopped.
22/03/11 21:12:25 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system shutdown complete.
22/03/11 21:12:25 WARN fs.FileSystem: Failed to initialize filesystem
s3a://hdfs-test/: org.apache.hadoop.fs.s3a.AWSClietIOException:
doesBucketExistV2 on hdfs: com.amazonaws.SdkClientException: Unable to
execute HTTP request: PKIX path building failed:
sun.security.provider.certpath.SunCertPathBuilderException: unable to find
valid certification path to requested target: Unable to execute HTTP
request: PKIX path building failed:
sun.security.provider.certpath.SunCertPathBuilderException: unable to find
valid certification path to requested target
```

Resolution: NetApp recommends using a server certificate issued by a known public certificate signing authority to make sure that the authentication is secure. Alternatively, add a custom CA or server certificate to the Java trust store.

Complete the following steps to add a StorageGRID custom CA or server certificate to the Java trust store.

1. Backup the existing default Java cacerts file.

```
cp -ap $JAVA_HOME/lib/security/cacerts
$JAVA_HOME/lib/security/cacerts.orig
```

2. Import the StorageGRID S3 endpoint cert to the Java trust store.

```
keytool -import -trustcacerts -keystore $JAVA_HOME/lib/security/cacerts
-storepass changeit -noprompt -alias sg-lb -file <StorageGRID CA or
server cert in pem format>
```


Troubleshooting tips

1. Increase the hadoop log level to DEBUG.

```
export HADOOP_ROOT_LOGGER=hadoop.root.logger=DEBUG,console
```

2. Execute the command, and direct the log messages to error.log.

```
hadoop fs -ls s3a://<bucket-name>/ &>error.log
```

By Angela Cheng

Use S3cmd to test and demonstrate S3 access on StorageGRID

S3cmd is a free command line tool and client for S3 operations. You can use s3cmd to test and demonstrate s3 access on StorageGRID.

Install and configure S3cmd

To install S3cmd on a workstation or server, download it from [command line S3 client](#). s3cmd is pre-installed on each StorageGRID node as a tool to aid in troubleshooting.

Initial configuration steps

1. s3cmd --configure
2. Provide only access_key and secret_key, for the the rest keep the defaults.
3. Test access with supplied credentials? [Y/n]: n (bypass the test as it will fail)
4. Save settings? [y/N] y
 - a. Configuration saved to '/root/.s3cfg'
5. In .s3cfg make fields host_base and host_bucket empty after the "=" sign :
 - a. host_base =
 - b. host_bucket =



If you specify host_base and host_bucket in step 4, you don't need to specify an endpoint with --host in the CLI. Example:

```
host_base = 192.168.1.91:8082
host_bucket = bucketX.192.168.1.91:8082
s3cmd ls s3://bucketX --no-check-certificate
```

Basic command examples

- Create a bucket:

```
s3cmd mb s3://s3cmdbucket --host=<endpoint>:<port> --no-check-certificate
```

- **List all buckets:**

```
s3cmd ls --host=<endpoint>:<port> --no-check-certificate
```

- **List all buckets and their contents:**

```
s3cmd la --host=<endpoint>:<port> --no-check-certificate
```

- **List objects in a specific bucket:**

```
s3cmd ls s3://<bucket> --host=<endpoint>:<port> --no-check-certificate
```

- **Delete a bucket:**

```
s3cmd rb s3://s3cmdbucket --host=<endpoint>:<port> --no-check-certificate
```

- **Put an object:**

```
s3cmd put <file> s3://<bucket> --host=<endpoint>:<port> --no-check-certificate
```

- **Get an object:**

```
s3cmd get s3://<bucket>/<object> <file> --host=<endpoint>:<port> --no-check-certificate
```

- **Delete an object:**

```
s3cmd del s3://<bucket>/<object> --host=<endpoint>:<port> --no-check-certificate
```

By Aron Klein

Vertica Eon mode database using NetApp StorageGRID as communal storage

This guide describes the procedure to create a Vertica Eon Mode database with communal storage on NetApp StorageGRID.

Introduction

Vertica is an analytic database management software. It is a columnar storage platform designed to handle large volumes of data, which enables very fast query performance in a traditionally intensive scenario. A Vertica database runs in one of the two modes: Eon or Enterprise. You can deploy both modes on-premises or in the cloud.

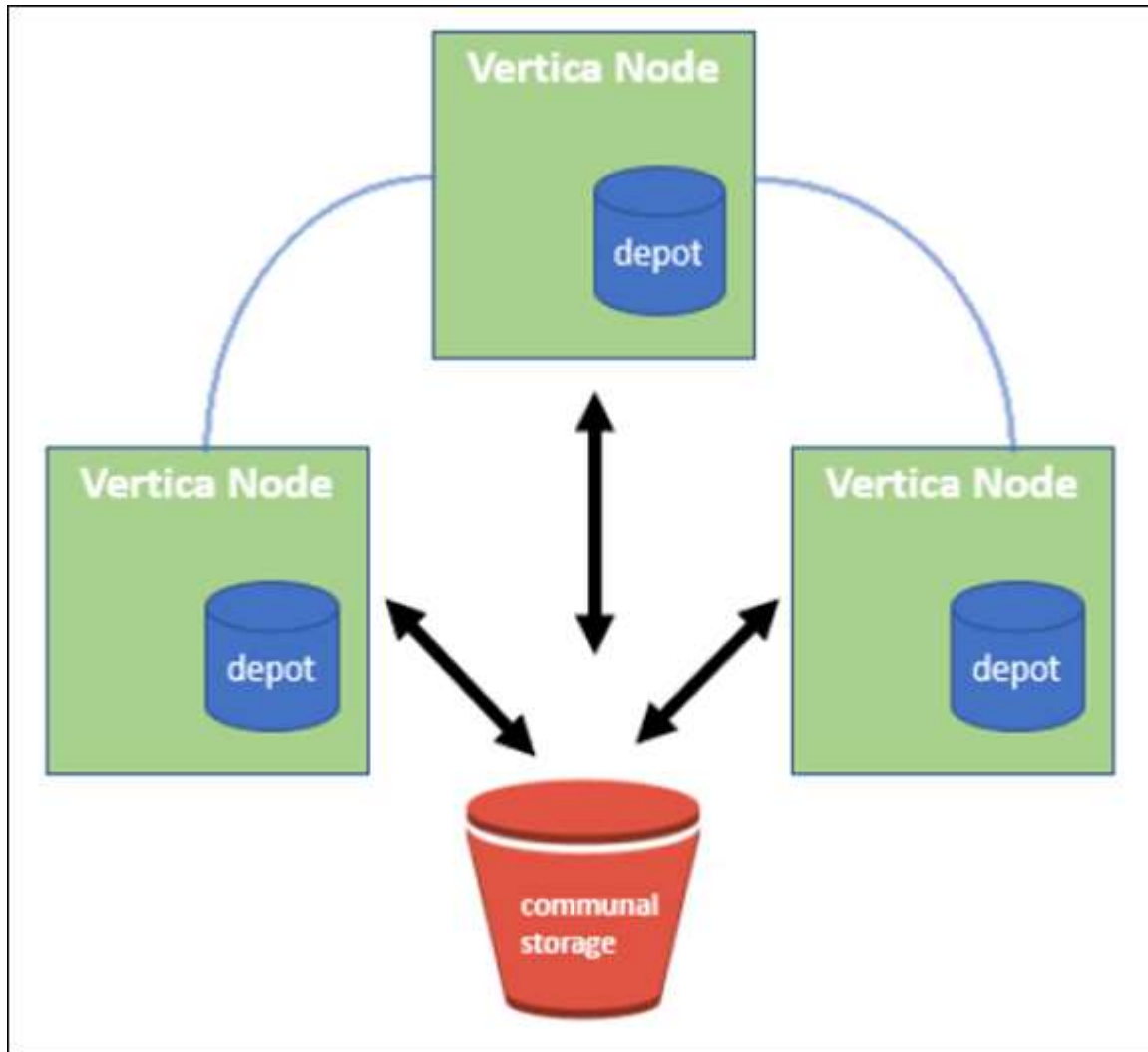
Eon and Enterprise modes primarily differ in where they store data:

- Eon Mode databases use communal storage for their data. This is recommended by Vertica.
- Enterprise Mode databases store data locally in the file system of nodes that make up the database.

Eon Mode architecture

Eon Mode separates the computational resources from the communal storage layer of the database, which allows the compute and storage to scale separately. Vertica in Eon Mode is optimized to address variable workloads and isolate them from one another by using separate compute and storage resources.

Eon Mode stores data in a shared object store called communal storage—an S3 bucket, either hosted on premises or on Amazon S3.



Communal storage

Instead of storing data locally, Eon Mode uses a single communal storage location for all data and the catalog (metadata). Communal storage is the database's centralized storage location, shared among the database nodes.

Communal storage has the following properties:

- Communal storage in the cloud or on-premises object storage is more resilient and less susceptible to data loss due to storage failures than storage on disk on individual machines.
- Any data can be read by any node using the same path.
- Capacity is not limited by disk space on nodes.

- Because data is stored communally, you can elastically scale your cluster to meet changing demands. If the data were stored locally on the nodes, adding or removing nodes would require moving significant amounts of data between nodes to either move it off nodes that are being removed, or onto newly created nodes.

The depot

One drawback of communal storage is its speed. Accessing data from a shared cloud location is slower than reading it from local disk. Also, the connection to communal storage can become a bottleneck if many nodes are reading data from it at once. To improve data access speed, the nodes in an Eon Mode database maintain a local disk cache of data called the depot. When executing a query, the nodes first check whether the data it needs is in the depot. If it is, then it finishes the query by using the local copy of the data. If the data is not in the depot, the node fetches the data from communal storage, and saves a copy in the depot.

NetApp StorageGRID recommendations

Vertica stores database data to object storage as thousands (or millions) of compressed objects (observed size is 200 to 500MB per object). When a user runs database queries, Vertica retrieves the selected range of data from these compressed objects in parallel using the byte-range GET call. Each byte-range GET is approximately 8KB.

During the 10TB database depot off user queries test, 4,000 to 10,000 GET (byte-range GET) requests per second were sent to the grid. When running this test using SG6060 appliances, though the CPU% utilization % per appliance node is low (around 20% to 30%), 2/3 of CPU time is waiting for I/O. A very small percentage (0% to 0.5%) of I/O wait is observed on the SGF6024.

Due to the high demand of small IOPS with very low latency requirements (the average should be less than 0.01 seconds), NetApp recommends using the SFG6024 for object storage services. If the SG6060 is needed for very large database sizes, the customer should work with the Vertica account team on depot sizing to support the actively queried dataset.

For the Admin Node and API Gateway Node, the customer can use the SG100 or SG1000. The choice depends on the number of users' query requests in parallel and database size. If the customer prefers to use a third-party load balancer, NetApp recommends a dedicated load balancer for high performance demand workload. For StorageGRID sizing, consult the NetApp account team.

Other StorageGRID configuration recommendations include:

- **Grid topology.** Do not mix the SGF6024 with other storage appliance models on the same grid site. If you prefer to use the SG6060 for long term archive protection, keep the SGF6024 with a dedicated grid load balancer in its own grid site (either physical or logical site) for an active database to enhance performance. Mixing different models of appliance on same site reduces the overall performance at the site.
- **Data protection.** Use replicate copies for protection. Do not use erasure coding for an active database. The customer can use erasure coding for long term protection of inactive databases.
- **Do not enable grid compression.** Vertica compresses objects before storing to object storage. Enabling grid compression does not further save storage usage and significantly reduces byte-range GET performance.
- **HTTP versus HTTPs S3 endpoint connection.** During the benchmark test, we observed about 5% performance improvement when using an HTTP S3 connection from the Vertica cluster to the StorageGRID load balancer endpoint. This choice should be based on customer security requirements.

Recommendations for a Vertica configuration include:

- **Vertica database default depot settings are enabled (value = 1) for read and write operations.** NetApp strongly recommends keeping these depot settings enabled to enhance performance.
- **Disable streaming limitations.** For configuration details, see the section [Disabling streaming limitations](#).

Installing Eon Mode on-premises with communal storage on StorageGRID

The following sections describe the procedure, in order, to install Eon Mode on-premises with communal storage on StorageGRID. The procedure to configure on-premises Simple Storage Service (S3) compatible object storage is similar to the procedure in the Vertica guide, [Install an Eon Mode Database on-premises](#).

The following setup was used for the functional test:

- StorageGRID 11.4.0.4
- Vertica 10.1.0
- Three virtual machines (VMs) with Centos 7.x OS for Vertica nodes to form a cluster. This setup is for the functional test only, not for the Vertica production database cluster.

These three nodes are set up with a Secure Shell (SSH) key to allow SSH without a password between the nodes within the cluster.

Information required from NetApp StorageGRID

To install Eon Mode on-premises with communal storage on StorageGRID, you must have the following prerequisite information.

- IP address or fully qualified domain name (FQDN) and port number of the StorageGRID S3 endpoint. If you are using HTTPS, use a custom certificate authority (CA) or self-signed SSL certificate implemented on the StorageGRID S3 endpoint.
- Bucket name. It must pre-exist and be empty.
- Access key ID and secret access key with read and write access to the bucket.

Creating an authorization file to access the S3 endpoint

The following prerequisites apply when creating an authorization file to access the S3 endpoint:

- Vertica is installed.
- A cluster is set up, configured, and ready for database creation.

To create an authorization file to access the S3 endpoint, follow these steps:

1. Log in to the Vertica node where you will run `admintools` to create the Eon Mode database.

The default user is `dbadmin`, created during the Vertica cluster installation.

2. Use a text editor to create a file under the `/home/dbadmin` directory.
The file name can be anything you want, for example, `sg_auth.conf`.
3. If the S3 endpoint is using a standard HTTP port 80 or HTTPS port 443, skip the port number. To use HTTPS, set the following values:
 - `awsenablehttps = 1`, otherwise set the value to 0.

◦ `awsauth = <s3 access key ID>:<secret access key>`

◦ `awsendpoint = <StorageGRID s3 endpoint>:<port>`

To use a custom CA or self-signed SSL certificate for the StorageGRID S3 endpoint HTTPS connection, specify the full file path and filename of the certificate. This file must be at the same location on each Vertica node and have read permission for all users. Skip this step if StorageGRID S3 Endpoint SSL certificate is signed by publicly known CA.

- `awscafile = <filepath/filename>`

For example, see the following sample file:

```
awsauth = MNVU4OYFAY2xyz123:03vuO4M4KmdfwffT8nqnBmnMVTr78Gu9wANabcxyz
awsendpoint = s3.england.connectlab.io:10443
awsenablehttps = 1
awscafile = /etc/custom-cert/grid.pem
```



In a production environment, the customer should implement a server certificate signed by a publicly known CA on a StorageGRID S3 load balancer endpoint.

Choosing a depot path on all Vertica nodes

Choose or create a directory on each node for the depot storage path.

The directory you supply for the depot storage path parameter must have the following:

- The same path on all nodes in the cluster (for example, `/home/dbadmin/depot`)
- Be readable and writable by the dbadmin user
- Sufficient storage

By default, Vertica uses 60% of the file system space containing the directory for depot storage. You can limit the size of the depot by using the `--depot-size` argument in the `create_db` command. See [Sizing Your Vertica Cluster for an Eon Mode Database](#) article for general Vertica sizing guidelines or consult with your Vertica account manager.

The `admintools create_db` tool attempts to create the depot path for you if one does not exist.

Creating the Eon on-premises database

To create the Eon on-premises database, follow these steps:

1. To create the database, use the `admintools create_db` tool.

The following list provides a brief explanation of arguments used in this example. See the Vertica document for a detailed explanation of all required and optional arguments.

- `-x <path/filename of authorization file created in “Creating an authorization file to access the S3 endpoint” >`.

The authorization details are stored inside database after successful creation. You can remove this file

to avoid exposing the S3 secret key.

- `--communal-storage-location <s3://storagegrid bucketname>`
- `-s <comma-separated list of Vertica nodes to be used for this database>`
- `-d <name of database to be created>`
- `-p <password to be set for this new database>`.

For example, see the following sample command:

```
admintools -t create_db -x sg_auth.conf --communal-storage
-location=s3://vertica --depot-path=/home/dbadmin/depot --shard
-count=6 -s vertica-vm1,vertica-vm2,vertica-vm3 -d vmart -p
'<password>'
```

Creating a new database takes several minutes duration depending on number of nodes for the database. When creating database for the first time, you will be prompted to accept the License Agreement.

For example, see the following sample authorization file and `create db` command:

```
[dbadmin@vertica-vm1 ~]$ cat sg_auth.conf
awsauth = MNVU4OYFAY2CPKVXVxxxx:03vu04M4KmdfwffT8nqnBmnMVTr78Gu9wAN+xxxx
awsendpoint = s3.england.connectlab.io:10445
awsenablehttps = 1

[dbadmin@vertica-vm1 ~]$ admintools -t create_db -x sg_auth.conf
--communal-storage-location=s3://vertica --depot-path=/home/dbadmin/depot
--shard-count=6 -s vertica-vm1,vertica-vm2,vertica-vm3 -d vmart -p
'xxxxxxxx'
Default depot size in use
Distributing changes to cluster.
  Creating database vmart
  Starting bootstrap node v_vmart_node0007 (10.45.74.19)
  Starting nodes:
    v_vmart_node0007 (10.45.74.19)
  Starting Vertica on all nodes. Please wait, databases with a large
catalog may take a while to initialize.
  Node Status: v_vmart_node0007: (DOWN)
  Node Status: v_vmart_node0007: (DOWN)
  Node Status: v_vmart_node0007: (DOWN)
  Node Status: v_vmart_node0007: (UP)
  Creating database nodes
  Creating node v_vmart_node0008 (host 10.45.74.29)
  Creating node v_vmart_node0009 (host 10.45.74.39)
  Generating new configuration information
  Stopping single node db before adding additional nodes.
```

```

Database shutdown complete
Starting all nodes
Start hosts = ['10.45.74.19', '10.45.74.29', '10.45.74.39']
Starting nodes:
    v_vmart_node0007 (10.45.74.19)
    v_vmart_node0008 (10.45.74.29)
    v_vmart_node0009 (10.45.74.39)
Starting Vertica on all nodes. Please wait, databases with a large
catalog may take a while to initialize.
Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
Node Status: v_vmart_node0007: (UP) v_vmart_node0008: (UP)
v_vmart_node0009: (UP)
Creating depot locations for 3 nodes
Communal storage detected: rebalancing shards

Waiting for rebalance shards. We will wait for at most 36000 seconds.
Installing AWS package
    Success: package AWS installed
Installing ComplexTypes package
    Success: package ComplexTypes installed
Installing MachineLearning package
    Success: package MachineLearning installed
Installing ParquetExport package
    Success: package ParquetExport installed
Installing VFunctions package
    Success: package VFunctions installed
Installing approximate package
    Success: package approximate installed
Installing flextable package
    Success: package flextable installed
Installing kafka package
    Success: package kafka installed
Installing logsearch package
    Success: package logsearch installed
Installing place package
    Success: package place installed
Installing txtindex package
    Success: package txtindex installed
Installing voltagesecure package

```


Success: package voltagesecure installed
Syncing catalog on vmart with 2000 attempts.
Database creation SQL tasks completed successfully. Database vmart created successfully.

| Object size (byte) | Bucket/object key full path |
|--------------------|--|
| 61 | s3://vertica/051/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a07/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a07_0_0.dfs |
| 145 | s3://vertica/2c4/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a3d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a3d_0_0.dfs |
| 146 | s3://vertica/33c/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a1d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a1d_0_0.dfs |
| 40 | s3://vertica/382/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a31/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a31_0_0.dfs |
| 145 | s3://vertica/42f/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a21/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a21_0_0.dfs |
| 34 | s3://vertica/472/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a25/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a25_0_0.dfs |
| 41 | s3://vertica/476/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a2d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a2d_0_0.dfs |
| 61 | s3://vertica/52a/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a5d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a5d_0_0.dfs |

| Object size (byte) | Bucket/object key full path |
|--------------------|---|
| 131 | s3://vertica/5d2/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a19/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a19_0_0.dfs |
| 91 | s3://vertica/5f7/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a11/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a11_0_0.dfs |
| 118 | s3://vertica/82d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a15/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a15_0_0.dfs |
| 115 | s3://vertica/9a2/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a61/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a61_0_0.dfs |
| 33 | s3://vertica/acd/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29_0_0.dfs |
| 133 | s3://vertica/b98/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a4d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a4d_0_0.dfs |
| 38 | s3://vertica/db3/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a49/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a49_0_0.dfs |
| 38 | s3://vertica/eba/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a59/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a59_0_0.dfs |
| 21521920 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000215e2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000215e2.tar |

| Object size (byte) | Bucket/object key full path |
|--------------------|---|
| 6865408 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021602/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021602.tar |
| 204217344 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021610/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021610.tar |
| 16109056 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0.tar |
| 12853248 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021800/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021800.tar |
| 8937984 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a.tar |
| 56260608 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2.tar |
| 53947904 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba.tar |
| 44932608 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de.tar |
| 256306688 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e.tar |

| Object size (byte) | Bucket/object key full path |
|--------------------|---|
| 8062464 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34.tar |
| 20024832 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70.tar |
| 10444 | s3://vertica/metadata/VMart/cluster_config.json |
| 823266 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/chkpt_1.cat.gz |
| 254 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/completed |
| 2958 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/chkpt_1.cat.gz |
| 231 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/completed |
| 822521 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/chkpt_1.cat.gz |
| 231 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/completed |
| 746513 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g14.cat |

| Object size (byte) | Bucket/object key full path |
|--------------------|--|
| 2596 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_3_g3.cat.gz |
| 821065 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_4_g4.cat.gz |
| 6440 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_5_g5.cat |
| 8518 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_8_g8.cat |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |
| 822922 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz |
| 232 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed |
| 822930 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz |
| 755033 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |
| 822922 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz |

| Object size (byte) | Bucket/object key full path |
|--------------------|---|
| 232 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed |
| 822930 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz |
| 755033 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |

Disabling streaming limitations

This procedure is based on the Vertica guide for other on-premises object storage and should be applicable to StorageGRID.

1. After creating the database, disable the `AWSStreamingConnectionPercentage` configuration parameter by setting it to 0.
This setting is unnecessary for an Eon Mode on-premises installation with communal storage. This configuration parameter controls the number of connections to the object store that Vertica uses for streaming reads. In a cloud environment, this setting helps avoid having streaming data from the object store use up all the available file handles. It leaves some file handles available for other object store operations. Due to the low latency of on-premises object stores, this option is unnecessary.
2. Use a `vsq` statement to update the parameter value.
The password is the database password that you set in “Creating the Eon on-premises database”. For example, see the following sample output:

```
[dbadmin@vertica-vm1 ~]$ vsq
Password:
Welcome to vsq, the Vertica Analytic Database interactive terminal.
Type:  \h or \? for help with vsq commands
       \g or terminate with semicolon to execute query
       \q to quit
dbadmin=> ALTER DATABASE DEFAULT SET PARAMETER
AWSStreamingConnectionPercentage = 0; ALTER DATABASE
dbadmin=> \q
```

Verifying depot settings

Vertica database default depot settings are enabled (value = 1) for read and write operations. NetApp strongly

recommends keeping these depot settings enabled to enhance performance.

```
vsq1 -c 'show current all;' | grep -i UseDepot
DATABASE | UseDepotForReads | 1
DATABASE | UseDepotForWrites | 1
```

Loading sample data (optional)

If this database is for testing and will be removed, you can load sample data to this database for testing. Vertica comes with sample dataset, VMart, found under `/opt/vertica/examples/VMart_Schema/` on each Vertica node.

You can find more information about this sample dataset [here](#).

Follow these steps to load the sample data:

1. Log in as dbadmin to one of the Vertica nodes: `cd /opt/vertica/examples/VMart_Schema/`
2. Load sample data to the database and enter the database password when prompted in substeps c and d:
 - a. `cd /opt/vertica/examples/VMart_Schema`
 - b. `./vmart_gen`
 - c. `vsq1 < vmart_define_schema.sql`
 - d. `vsq1 < vmart_load_data.sql`
3. There are multiple predefined SQL queries, you can run some of them to confirm test data are loaded successfully into the database.
For example: `vsq1 < vmart_queries1.sql`

Where to find additional information

To learn more about the information that is described in this document, review the following documents and/or websites:

- [NetApp StorageGRID 11.7 Product Documentation](#)
- [StorageGRID data sheet](#)
- [Vertica 10.1 Product Documentation](#)

Version history

| Version | Date | Document version history |
|-------------|----------------|--------------------------|
| Version 1.0 | September 2021 | Initial release. |

By Angela Cheng

StorageGRID log analytics using ELK stack

With the StorageGRID 11.6 syslog forward feature, you can configure an external syslog server to collect and analyze StorageGRID log messages. ELK (Elasticsearch, Logstash,

Kibana) has become one of the most popular log analytics solutions. Watch the [StorageGRID log analysis using ELK video](#) to view a sample ELK configuration and how it can be used to identify and troubleshoot failed S3 requests.

This article provides sample files of Logstash configuration, Kibana queries, charts and dashboard to give you a quick start for StorageGRID log management and analytics.

Requirements

- StorageGRID 11.6.0.2 or higher
- ELK (Elasticsearch, Logstash and Kibana) 7.1x or higher installed and in operation

Sample files

- [Download the Logstash 7.x sample files package](#)
md5 checksum 148c23d0021d9a4bb4a6c0287464deab
sha256 checksum f51ec9e2e3f842d5a7861566b167a561beb4373038b4e7bb3c8be3d522adf2d6
- [Download the Logstash 8.x sample files package](#)
md5 checksum e11bae3a662f87c310ef363d0fe06835
sha256 checksum 5c670755742cfd5aa723a596ba087e0153a65bcaef3934afdb682f61cd278d

Assumption

Readers are familiar with StorageGRID and ELK terminology and operations.













Instruction

Two sample versions are provided due to differences in names defined by grok patterns.

For example, the SYSLOGBASE grok pattern in Logstash config file defines field names differently depending on the installed Logstash version.

```
match => {"message" => '<{%POSINT:syslog_pri}>{%SYSLOGBASE}
{%GREEDYDATA:msg-details}' }
```




Logstash 7.17 sample

| Field | Value |
|---|-----------------------------|
|  _id | 7C1MaYEBRH8UbfKnIls8 |
|  _index | sgrid2-2022.06.15 |
|  _score | - |
|  _type | _doc |
|  @timestamp | Jun 15, 2022 @ 17:36:46.038 |
|  host | grid2-site2-s1 |
|  logsource | SITE2-S1 |
|  msg-details | Reloading syslog service |
|  pid | 628 |
|  program | update-sysl |
|  syslog_pri | 37 |
|  timestamp | Jun 15 21:36:46 |

Logstash 8.23 sample

Table JSON

 Search field names

| Actions | Field | Value |
|---------|--|---|
| ... |  _id | yuh0iIEBVP6KX4EwqcyU |
| ... |  _index | sglog-2022.06.21 |
| ... |  _score | - |
| ... |  @timestamp | Jun 21, 2022 @ 18:07:45.444 |
| ... |  event.original | <28>Jun 21 22:07:45 SITE2-S3 ADE: syslog messages being dropped |
| ... |  host.hostname | SITE2-S3 |
| ... |  msg-details | syslog messages being dropped |
| ... |  process.name | ADE |
| ... |  syslog_pri | 28 |
| ... |  timestamp | Jun 21 22:07:45 |

Steps

1. Unzip the provided sample based on your installed ELK version.
The sample folder includes two Logstash config samples:
sglog-2-file.conf: this config file outputs StorageGRID log messages to a file on Logstash without data transformation. You can use this to confirm Logstash is receiving StorageGRID messages or to help understand StorageGRID log patterns.
sglog-2-es.conf: this config file transforms StorageGRID log messages using various pattern and filters. It includes example drop statements, which drop messages based on patterns or filter. The output is sent to Elasticsearch for indexing.
Customize the selected config file according to the instruction inside the file.

2. Test the customized config file:

```
/usr/share/logstash/bin/logstash --config.test_and_exit -f <config-file-path/file>
```

If the last line returned is similar to the below line, the config file has no syntax errors:

```
[LogStash::Runner] runner - Using config.test_and_exit mode. Config  
Validation Result: OK. Exiting Logstash
```

3. Copy the customized conf file to the Logstash server's config: /etc/logstash/conf.d
If you have not enabled config.reload.automatic in /etc/logstash/logstash.yml, restart the Logstash service. Otherwise, wait for the config reload interval to elapse.

```
grep reload /etc/logstash/logstash.yml  
# Periodically check if the configuration has changed and reload the  
pipeline  
config.reload.automatic: true  
config.reload.interval: 5s
```

4. Check /var/log/logstash/logstash-plain.log and confirm there are no errors starting Logstash with the new config file.
5. Confirm TCP port is started and listening.
In this example, TCP port 5000 is used.

```
netstat -ntpa | grep 5000  
tcp6          0      0 :::5000          :::*  
LISTEN        25744/java
```

6. From the StorageGRID manager GUI, configure external syslog server to send log messages to Logstash. Refer to the [demo video](#) for details.
7. You need to configure or disable firewall on the Logstash server to allow StorageGRID nodes connection to the defined TCP port.

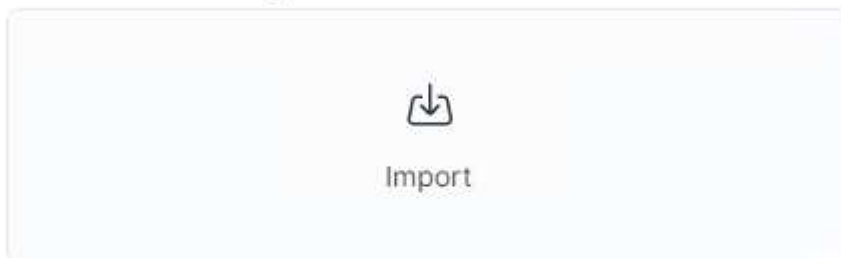
- From Kibana GUI, select Management → Dev Tools. On the Console page, run this GET command to confirm new indices are created on Elasticsearch.

```
GET /_cat/indices/*?v=true&s=index
```

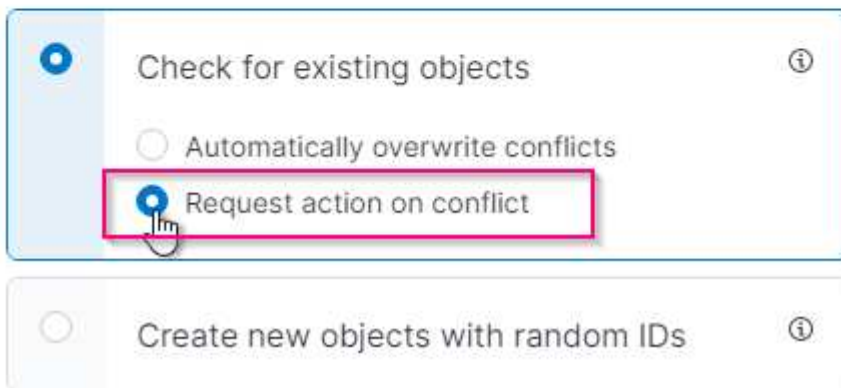
- From Kibana GUI, create index pattern (ELK 7.x) or data view (ELK 8.x).
- From Kibana GUI, enter 'saved objects' in the search box which is located in the top center. On the Saved Objects page, select Import. Under Import options, select 'Request action on conflict'

Import saved objects

Select a file to import



Import options



Import elk<version>-query-chart-sample.ndjson.

When prompted to resolve the conflict, select the index pattern or data view you created in step 8.

Import saved objects

Data Views Conflicts

The following saved objects use data views that do not exist. Please select the data views you'd like re-associated with them. You can [create a new data view](#) if necessary.

| ID | Count | Sample of aff... | New data view |
|--------------------------------------|-------|------------------|---------------|
| 594f91a0-d192-11ec-b30f-09f67aedd1d9 | 2 | | sglog |
| 60cf3620-e5fa-11ec-af71-8f6e980d6eb0 | 1 | | sglog |

The following Kibana objects are imported:

Query

- * audit-msg-s3rq-orlm
- * bycast log s3 related messages
- * loglevel warning or above
- * failed security event

Chart

- * s3 requests count based on bycast.log
- * HTTP status code
- * audit msg breakdown by type
- * average s3 response time

Dashboard

- * S3 request dashboard using the above charts.

You are now ready to perform StorageGRID log analysis using Kibana.

Additional resources

- [syslog101](#)
- [What is the ELK stack](#)

- [Grok patterns list](#)
- [A beginner's guide to Logstash: Grok](#)
- [A practical guide to Logstash: syslog deep dive](#)
- [Kibana guide – Explore the document](#)
- [StorageGRID audit log messages reference](#)

By Angela Cheng

Use Prometheus and Grafana to extend your metrics retention

This technical report provides detailed instructions for configuring NetApp StorageGRID 11.6 with external Prometheus and Grafana services.

Introduction

StorageGRID stores metrics using Prometheus and provides visualizations of these metrics through built in Grafana dashboards. The Prometheus metrics can be accessed securely from StorageGRID by configuring client access certificates and enabling prometheus access for the specified client. Today, the retention of this metric data is limited by the storage capacity of the administration node. To gain a longer duration and an ability to create customized visualizations of these metrics we will deploy a new Prometheus and Grafana server, configure our new server to scrape the metrics from StorageGRID's instance, and build a dashboard with the metrics that are important to us. You can get more information on the Prometheus metrics collected in the [StorageGRID documentation](#).

Federate Prometheus

Lab details

For the purposes of this example, I will be using all virtual machines for StorageGRID 11.6 nodes, and a Debian 11 server. The StorageGRID management interface is configured with a publicly trusted CA certificate. This example will not go through the installation and configuration of the StorageGRID system or Debian linux installation. You can use any Linux flavor you wish that is supported by Prometheus and Grafana. Both Prometheus and Grafana can install as docker containers, build from source, or pre-compiled binaries. In this example I will be installing both Prometheus and Grafana binaries directly on the same Debian server. Download and follow the basic installation instructions from <https://prometheus.io> and <https://grafana.com/grafana/> respectively.

Configure StorageGRID for Prometheus Client access

In order to gain access to StorageGRID's stored prometheus metrics you must generate or upload a client certificate with private key, and enable permission for the client. The StorageGRID management interface must have an SSL certificate. This certificate must be trusted by the prometheus server either by a trusted CA, or manually trusted if it is self-signed. To read more, please visit the [StorageGRID documentation](#).

1. In the StorageGRID management interface, select "CONFIGURATION" on the bottom left hand side, and in the second column under "Security" click on Certificates.
2. On the Certificates page select the "Client" tab and click on the "Add" button.
3. Provide a name for the client that will be granted access and use this certificate. Click on the box under "Permissions", in front of "Allow Prometheus" and click the Continue button.

Add a client certificate

1 Enter details ————— 2 Enter details

Certificate details

Certificate name ?

prometheus

Permissions

☒ Allow prometheus ?

4. If you have a CA signed certificate you can select the radio button for "Upload certificate", but in our case we are going to let storageGRID generate the client certificate by selecting the radio button for "Generate Certificate". The required fields will be displayed to be filled in. Enter the FQDN for the client server, the IP of the server, the subject, and Days valid. Then click the "Generate" button.

Add a client certificate

Enter details

2 Enter details

Certificate type

Upload certificate

Generate certificate

Domain name

prometheus.grid.local

Add another domain

IP

192.168.0.10

Add another IP address

Subject

/CN=Prometheus

Days valid

730

Generate

Previous

Create

Be mindful of the certificate days valid entry as you will need to renew this certificate in both StorageGRID and the Prometheus server before it expires to maintain uninterrupted collection.

1. Download the certificate pem file, and the private key pem file.

77

[Generate](#)

Certificate details

[Download certificate](#)
[Copy certificate PEM](#)

Subject DN: /CN=Prometheus
Serial Number: 72:D9:6E:D7:04:CC:4F:29:66:0A:CA:53:24:79:1B:09:49:3A:BC:56
Issuer DN: /CN=Prometheus
Issued On: 2022-08-22T17:54:33.000Z
Expires On: 2024-08-21T17:54:33.000Z
SHA-1 Fingerprint: 10:47:6E:FD:67:D8:53:E7:6E:E5:D8:8A:DF:BD:45:94:04:53:47:1E
SHA-256 Fingerprint: 74:23:C2:02:3A:D9:08:C0:EE:C1:F8:59:8A:7C:AE:18:AB:80:7D:21:31:F3:EB:AF:BF:4F:9E:C7:90:C9:FA:E7
Alternative Names: DNS:prometheus.grid.local
IP Address:192.168.0.10

Certificate private key ⓘ

⚠ You will not be able to view the certificate private key after you close this dialog. To save the keys for future reference, copy and paste the values to another location.

[Download private key](#)
[Copy private key](#)

```
-----BEGIN RSA PRIVATE KEY-----
MIIEpAIBAAKCAQEA3bIcyIEpMWPk5ritVpMkmIDKLIjaTM3ertq23VcAALwxziaU
...3B1x2uH4ABE7d1eD-471e1B8u7afC-34F81D:07k...2oulT76k0C04...
```



This is the only time you can download the private key, so make sure you do not skip this step.

Prepare the Linux server for Prometheus installation

Before installing Prometheus, I want to get my environment prepared with a Prometheus user, the directory structure, and configure the capacity for the metrics storage location.

1. Create the Prometheus user.

```
sudo useradd -M -r -s /bin/false Prometheus
```

2. Create the directories for Prometheus, client certificate, and metrics data.

```
sudo mkdir /etc/Prometheus /etc/Prometheus/cert /var/lib/Prometheus
```

3. I formatted the disk I am using for metrics retention with an ext4 filesystem.

```
mkfs -t ext4 /dev/sdb
```

4. I then mounted the filesystem to the Prometheus metrics directory.


```
sudo mount -t auto /dev/sdb /var/lib/prometheus/
```

5. Obtain the uuid of the disk you are using for your metrics data.

```
sudo ls -al /dev/disk/by-uuid/  
lrwxrwxrwx 1 root root 9 Aug 18 17:02 9af2c5a3-bfc2-4ec1-85d9-  
ebab850bb4a1 -> ../../sdb
```

6. Adding an entry in /etc/fstab/ making the mount persist across reboots using the uuid of /dev/sdb.

```
/etc/fstab  
UUID=9af2c5a3-bfc2-4ec1-85d9-ebab850bb4a1 /var/lib/prometheus ext4  
defaults 0 0
```

Install and configure Prometheus

Now that the server is ready, I can begin the Prometheus installation and configure the service.

1. Extract the Prometheus installation package

```
tar xzf prometheus-2.38.0.linux-amd64.tar.gz
```

2. Copy the binaries to /usr/local/bin and change the ownership to the prometheus user created earlier

```
sudo cp prometheus-2.38.0.linux-amd64/{prometheus,promtool}  
/usr/local/bin  
sudo chown prometheus:prometheus /usr/local/bin/{prometheus,promtool}
```

3. Copy the consoles and libraries to /etc/prometheus

```
sudo cp -r prometheus-2.38.0.linux-amd64/{consoles,console_libraries}  
/etc/prometheus/
```

4. Copy the client certificate and private key pem files downloaded earlier from StorageGRID to /etc/prometheus/certs

5. Create the prometheus configuration yaml file

```
sudo nano /etc/prometheus/prometheus.yml
```

6. Insert the following configuration. The job name can be anything you wish. Change the "-targets: []" to the

FQDN of the admin node, and if you altered the names of the certificate and private key file names, please update the `tls_config` section to match. then save the file. If your grid management interface, is using a self-signed certificate, download the certificate and place it with the client certificate with a unique name, and in the `tls_config` section add `ca_file: /etc/prometheus/cert/UIcert.pem`

- a. In this example I am collecting all of the metrics that begin with alertmanager, cassandra, node, and storagegrid. You can see more information on the Prometheus metrics in the [StorageGRID documentation](#).

```
# my global config
global:
  scrape_interval: 60s # Set the scrape interval to every 15 seconds.
                        Default is every 1 minute.

scrape_configs:
  - job_name: 'StorageGRID'
    honor_labels: true
    scheme: https
    metrics_path: /federate
    scrape_interval: 60s
    scrape_timeout: 30s
    tls_config:
      cert_file: /etc/prometheus/cert/certificate.pem
      key_file: /etc/prometheus/cert/private_key.pem
    params:
      match[]:
        -
        '{__name__=~"alertmanager_.*|cassandra_.*|node_.*|storagegrid_.*"}'
    static_configs:
      - targets: ['sgdemo-rtp.netapp.com:9091']
```



If your grid management interface is using a self-signed certificate, download the certificate and place it with the client certificate with a unique name. In the `tls_config` section add the certificate above the client certificate and private key lines

```
ca_file: /etc/prometheus/cert/UIcert.pem
```

1. Change the ownership of all files and directories in `/etc/prometheus`, and `/var/lib/prometheus` to the prometheus user

```
sudo chown -R prometheus:prometheus /etc/prometheus/
sudo chown -R prometheus:prometheus /var/lib/prometheus/
```

2. Create a prometheus service file in `/etc/systemd/system`

```
sudo nano /etc/systemd/system/prometheus.service
```

3. Insert the following lines, note the `--storage.tsdb.retention.time=1y` which sets the retention of the metric data to 1 year. Alternatively, you could use `--storage.tsdb.retention.size=300GiB` to base retention on storage limits. This is the only location to set the metrics retention.

```
[Unit]
Description=Prometheus Time Series Collection and Processing Server
Wants=network-online.target
After=network-online.target

[Service]
User=prometheus
Group=prometheus
Type=simple
ExecStart=/usr/local/bin/prometheus \
    --config.file /etc/prometheus/prometheus.yml \
    --storage.tsdb.path /var/lib/prometheus/ \
    --storage.tsdb.retention.time=1y \
    --web.console.templates=/etc/prometheus/consoles \
    --web.console.libraries=/etc/prometheus/console_libraries

[Install]
WantedBy=multi-user.target
```

4. Reload the systemd service to register the new prometheus service. then start and enable the prometheus service.

```
sudo systemctl daemon-reload
sudo systemctl start prometheus
sudo systemctl enable prometheus
```

5. Check the service is running properly

```
sudo systemctl status prometheus
```

- prometheus.service - Prometheus Time Series Collection and Processing Server

Loaded: loaded (/etc/systemd/system/prometheus.service; enabled; vendor preset: enabled)

Active: active (running) since Mon 2022-08-22 15:14:24 EDT; 2s ago

Main PID: 6498 (prometheus)

Tasks: 13 (limit: 28818)

Memory: 107.7M

CPU: 1.143s

CGroup: /system.slice/prometheus.service

└─6498 /usr/local/bin/prometheus --config.file /etc/prometheus/prometheus.yml --storage.tsdb.path /var/lib/prometheus/ --web.console.templates=/etc/prometheus/consoles --web.con>

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.510Z caller=head.go:544 level=info component=tsdb msg="Replaying WAL, this may take a while"

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.816Z caller=head.go:615 level=info component=tsdb msg="WAL segment loaded" segment=0 maxSegment=1

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.816Z caller=head.go:615 level=info component=tsdb msg="WAL segment loaded" segment=1 maxSegment=1

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.816Z caller=head.go:621 level=info component=tsdb msg="WAL replay completed" checkpoint_replay_duration=55.57µs wal_rep>

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.831Z caller=main.go:997 level=info fs_type=EXT4_SUPER_MAGIC

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.831Z caller=main.go:1000 level=info msg="TSDB started"

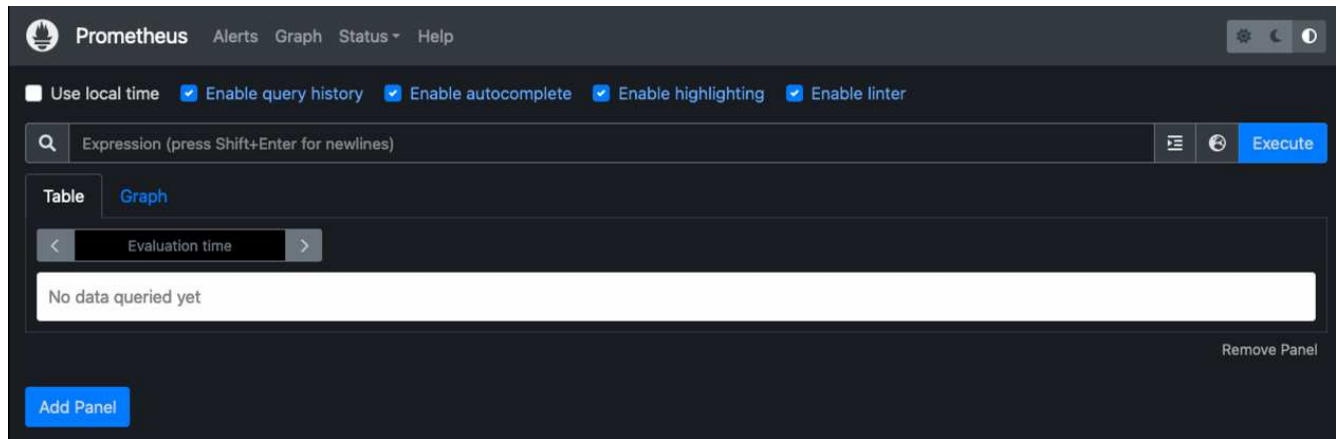
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.831Z caller=main.go:1181 level=info msg="Loading configuration file" filename=/etc/prometheus/prometheus.yml

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.832Z caller=main.go:1218 level=info msg="Completed loading of configuration file" filename=/etc/prometheus/prometheus.y>

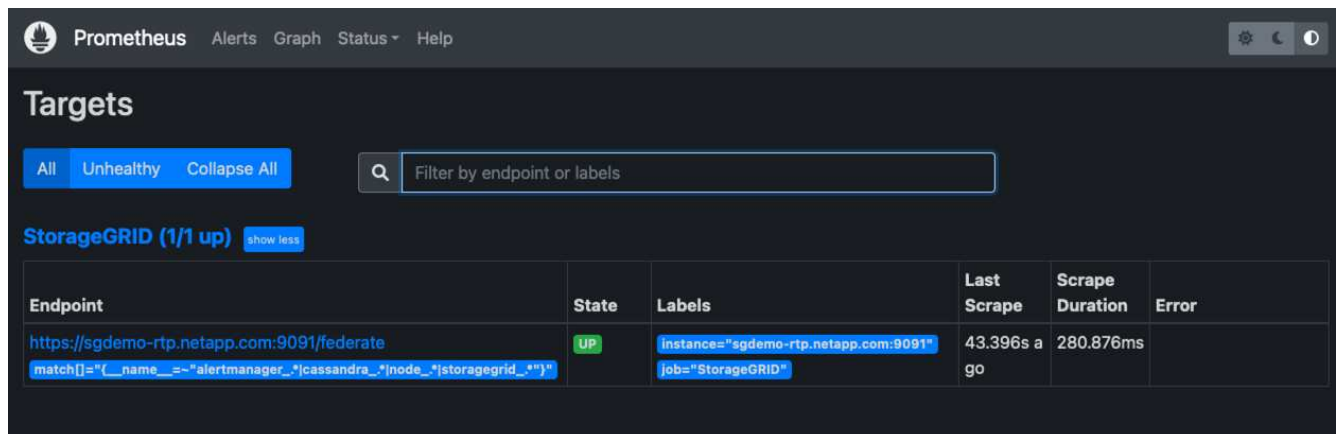
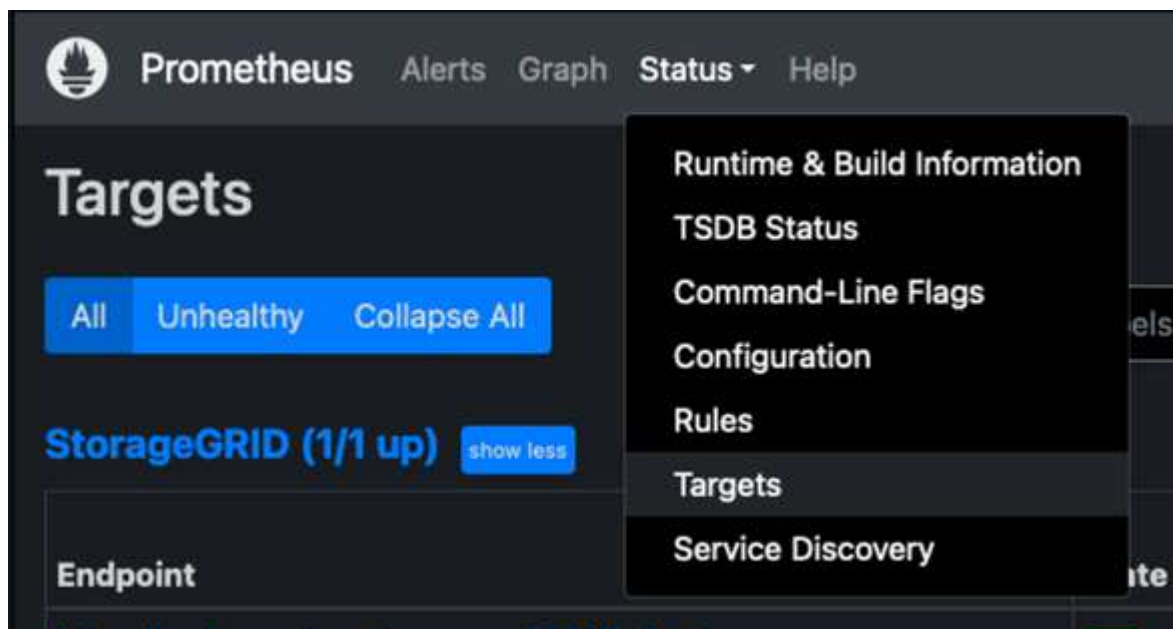
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.832Z caller=main.go:961 level=info msg="Server is ready to receive web requests."

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-22T19:14:24.832Z caller=manager.go:941 level=info component="rule manager" msg="Starting rule manager..."

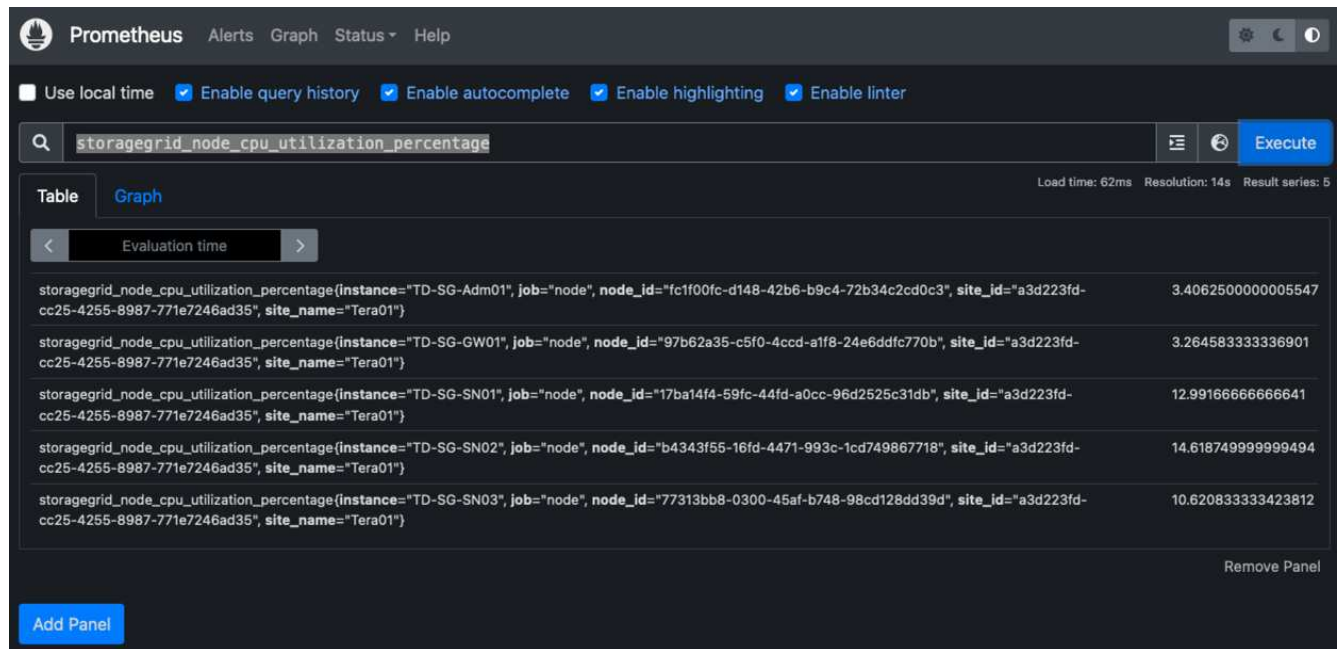
6. You should now be able to browse to the UI of your prometheus server <http://Prometheus-server:9090> and see the UI



- Under "Status" Targets you can see the status of the StorageGRID endpoint we configured in prometheus.yml



- On the Graph page, you can execute a test query and verify the data is successfully being scraped. for example enter "storagegrid_node_cpu_utilization_percentage" into the query bar and click the Execute button.



Install and configure Grafana

Now that prometheus is installed and working, we can move on to installing Grafana and configuring a dashboard

Grafana Instalation

1. Install the latest enterprise edition of Grafana

```
sudo apt-get install -y apt-transport-https
sudo apt-get install -y software-properties-common wget
sudo wget -q -O /usr/share/keyrings/grafana.key
https://packages.grafana.com/gpg.key
```

2. Add this repository for stable releases:

```
echo "deb [signed-by=/usr/share/keyrings/grafana.key]
https://packages.grafana.com/enterprise/deb stable main" | sudo tee -a
/etc/apt/sources.list.d/grafana.list
```

3. After you add the repository.

```
sudo apt-get update
sudo apt-get install grafana-enterprise
```

4. Reload the systemd service to register the new grafana service. then start and enable the Grafana service.

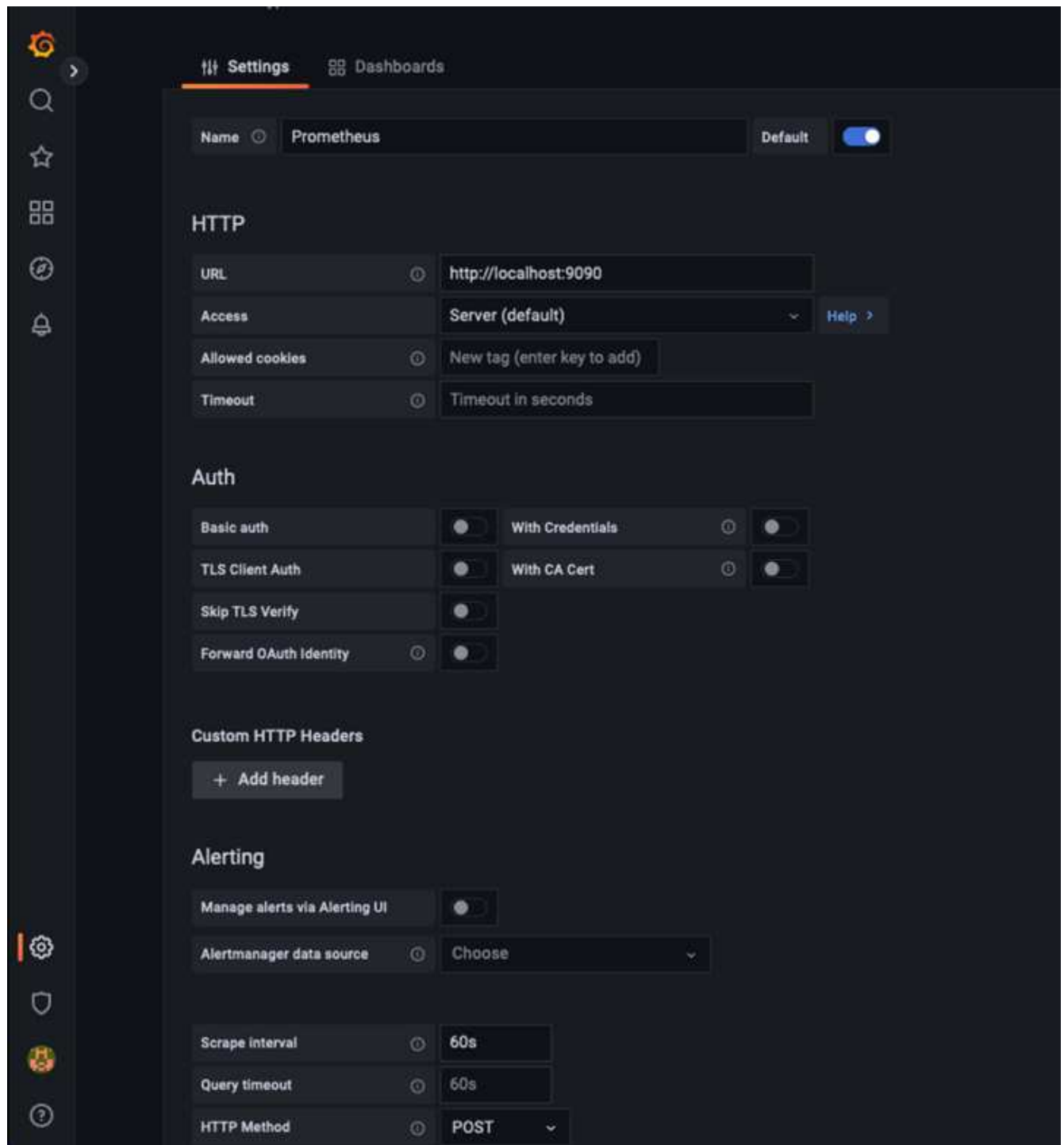
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

5. Grafana is now installed and running. When you open a browser to `HTTP://Prometheus-server:3000` you will be greeted with the Grafana login page.
6. The default login credentials are `admin/admin`, and you should set a new password as it prompts you to.

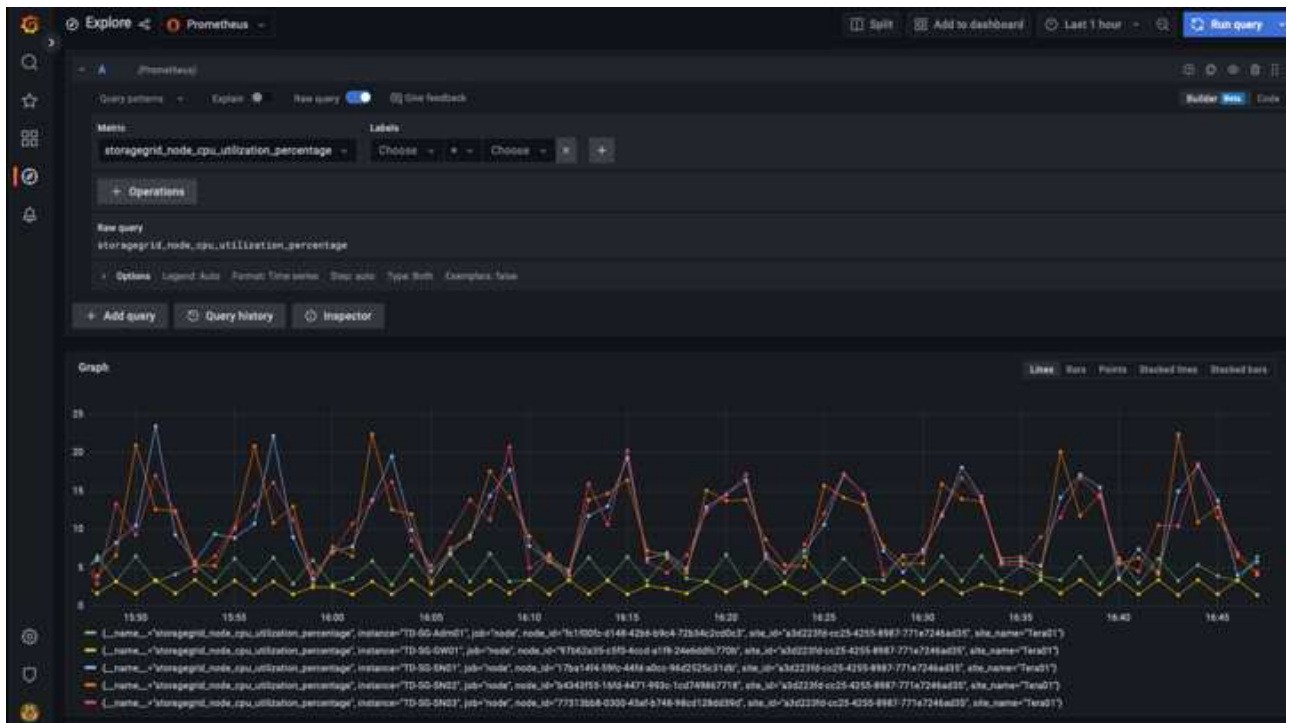
Create a Grafana dashboard for StorageGRID

With Grafana and Prometheus installed and running, now its time to connect the two by creating a data source and build a dashboard

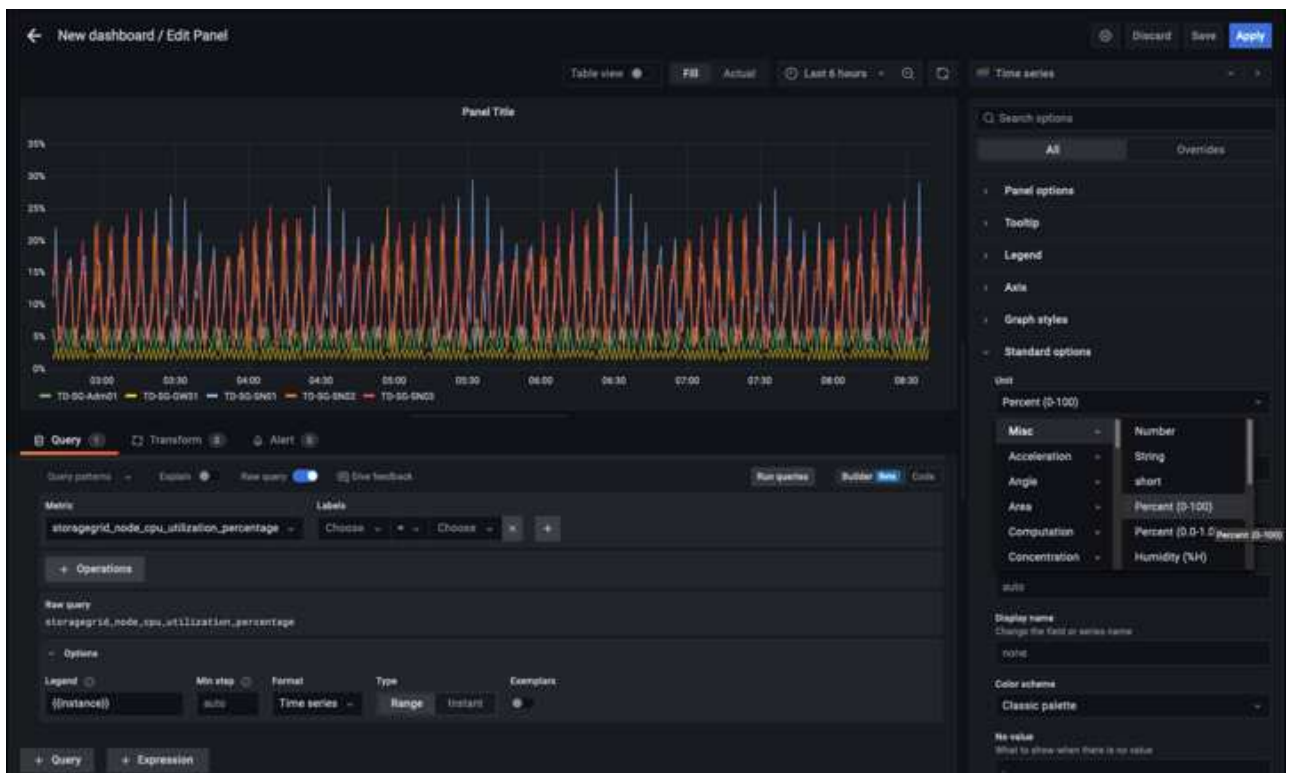
1. On the left hand pane expand "Configuration" and select "Data sources", then click on the "Add Data source" button
2. Prometheus will be one of the top data sources to choose from. If it is not, then use the search bar to locate "Prometheus"
3. Configure the Prometheus source by entering the URL of the prometheus instance, and the scrape interval to match the Prometheus interval. I also disabled the alerting section as I did not configure the alert manager on prometheus.



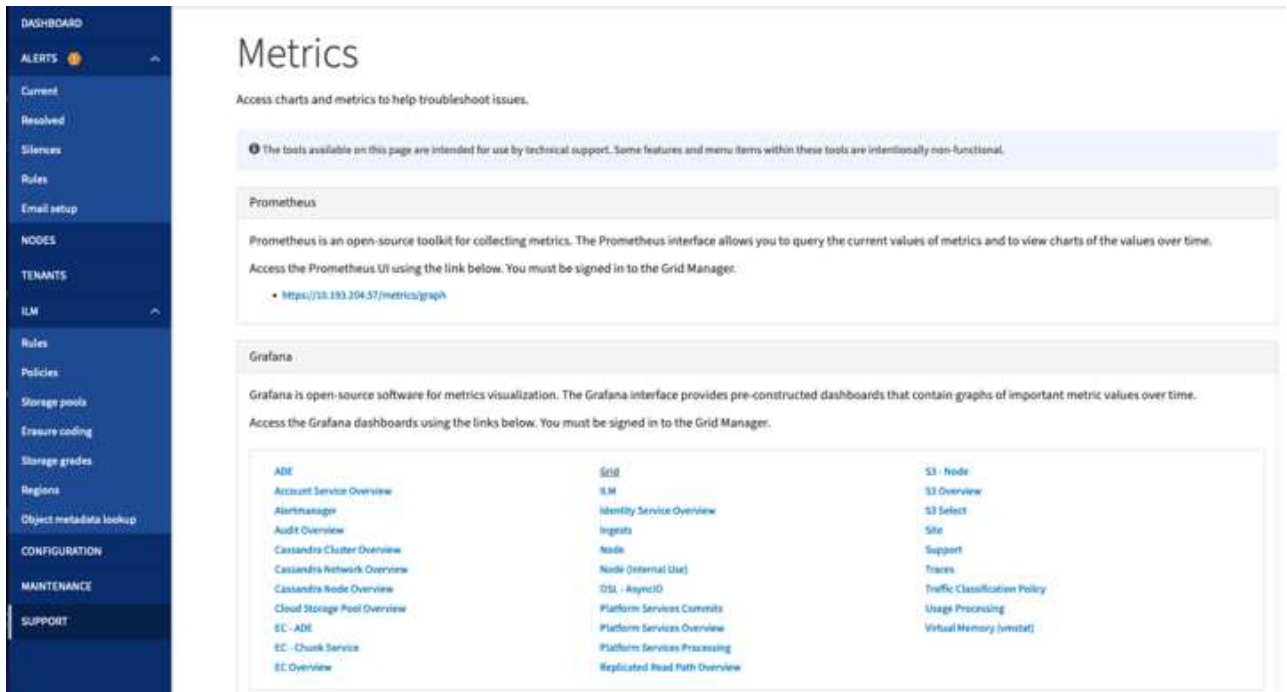
4. With the desired settings entered, scroll down to the bottom and click on "Save & test"
5. After the configuration test is successful, click on the explore button.
 - a. In the explore window you can use the same metric we tested Prometheus with "storagegrid_node_cpu_utilization_percentage", and click the "Run query" button



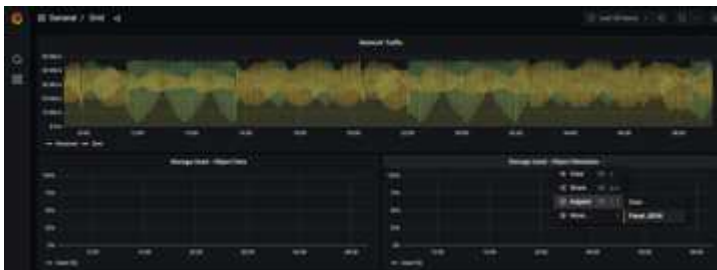
6. Now that we have the data source configured, we can create a dashboard.
 - a. On the left hand pane expand "Dashboards", and select "+ new Dashboard"
 - b. Select "Add a new panel"
 - c. Configure the new panel by selecting a metric, again I will use "storagegrid_node_cpu_utilization_percentage", Enter a title for the panel, expand "Options" at the bottom and for legend change to custom and enter "{instance}" to define the node names", and on the right pane under "Standard options" set "Unit" to "Misc/Percent(0-100)". Then click "Apply" to save the panel to the dashboard.



7. We could continue to build out our dashboard like this for each metric we want, but luckily StorageGRID already has dashboards with panels we can copy into our custom dashboards.
 - a. From the StorageGRID management interface left hand pane, select "Support", and at the bottom of the "Tools" column click on "Metrics".
 - b. Within metrics, I am going to select the "Grid" link on the top of the middle column.



- c. From the Grid dashboard, let's select the "Storage Used - Object Metadata" panel. Click the little down arrow and the end of the panel title to drop down a menu. From this menu select "Inspect" and "Panel JSON".



- d. Copy out the JSON code and close the window.

Inspect: Storage Used - Object Metadata

4 queries with total query time of 549 ms

Data

Stats

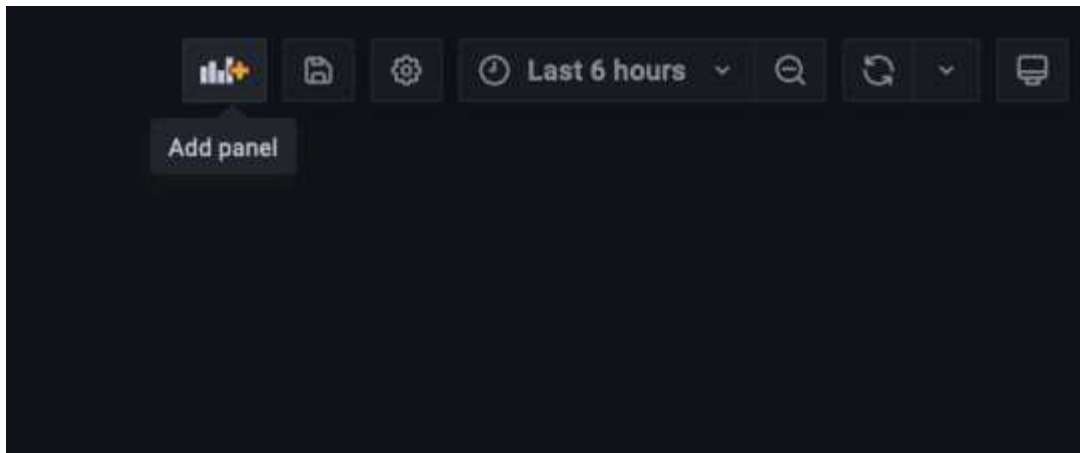
JSON

Select source

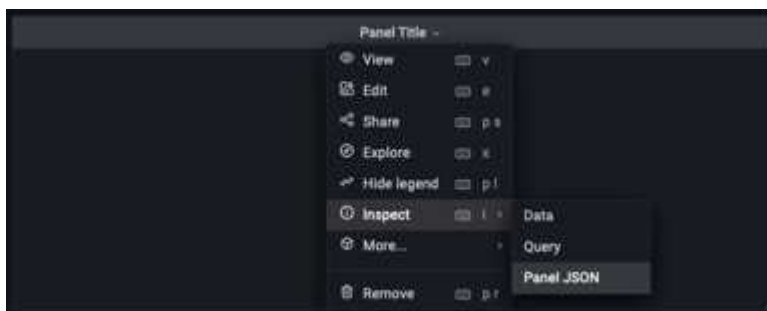
Panel JSON

```
1 {
2   "aliasColors": {},
3   "bars": false,
4   "dashLength": 10,
5   "dashes": false,
6   "datasource": "Prometheus",
7   "decimals": 2,
8   "fill": 1,
9   "fillGradient": 0,
10  "gridPos": {
11    "h": 7,
12    "w": 12,
13    "x": 12,
14    "y": 7
15  },
16  "id": 6,
17  "legend": {
18    "avg": false,
19    "current": false,
20    "max": false,
21    "min": false,
22    "show": true,
23    "total": false,
24    "values": false
25  },
26  "lines": true,
27  "linewidth": 1,
28  "links": [],
29  "nullPointMode": "null",
30  "options": {
31    "alertThreshold": true
32  },
33  "percentage": false,
34  "pointradius": 5,
35  "points": false,
36  "renderer": "flot",
37  "seriesOverrides": [
38    {
39      "alias": "Used",
```

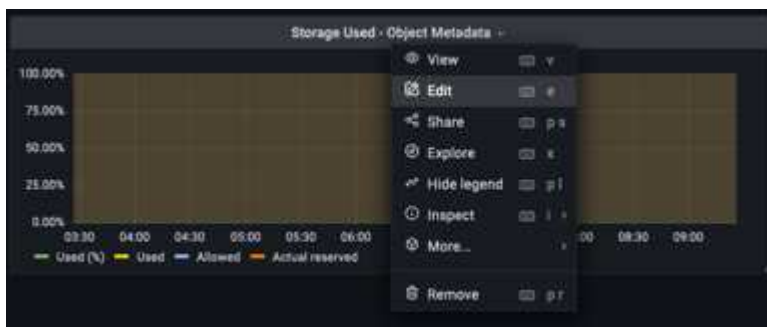
e. In our new dashboard, click on the icon to add a new panel.

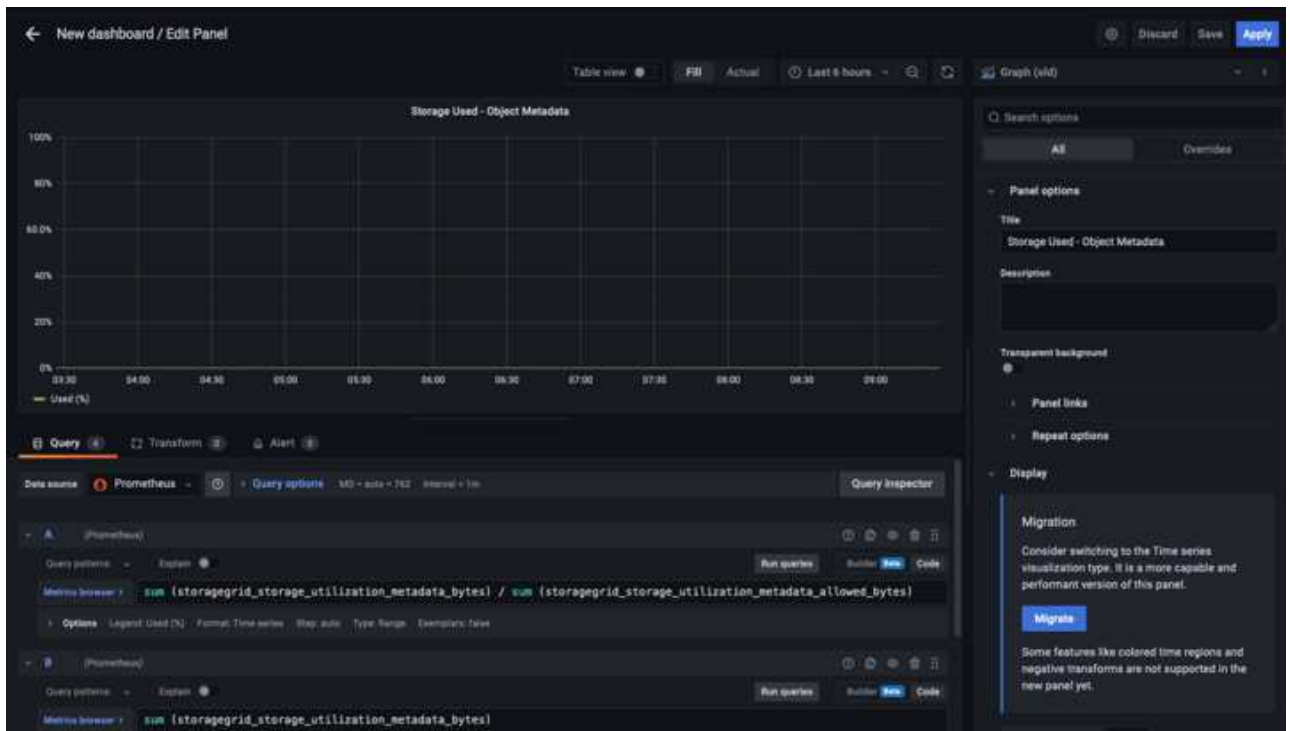


- f. Apply the new panel without making any changes
- g. Just like with the StorageGRID panel, inspect the JSON. Remove all JSON code and replace it with the copied code from the StorageGRID panel.



- h. Edit the new panel, and on the right hand side you will see a Migration message with a "Migrate" button. Click the button and then click the "Apply" button.





8. Once you have all of the panels in place and configured as you like. Save the dashboard by clicking the disk icon in the upper right and give your dashboard a name.

Conclusion

Now we have a Prometheus server with customizable data retention and storage capacity. With this we can continue build out our own dashboards with the metrics that are most relevant to our operations. You can get more information on the Prometheus metrics collected in the [StorageGRID documentation](#).

By Aron Klein

Datadog SNMP configuration

Configure Datadog to collect StorageGRID snmp metrics and traps.

Configure Datadog

Datadog is a monitoring solution providing metrics, visualizations, and alerting. The following configuration was implemented with linux agent version 7.43.1 on an Ubuntu 22.04.1 host deployed local to the StorageGRID system.

Datadog Profile and Trap files Generated from StorageGRID MIB file

Datadog provides a method for converting product MIB files into datadog reference files required to map the SNMP messages.

This StorageGRID yaml file for Datadog Trap resolution mapping generated following the instruction found [here](#).

Place this file in /etc/datadog-agent/conf.d/snmp.d/traps_db/ +

- [Download the trap yaml file](#) +

- **md5 checksum** 42e27e4210719945a46172b98c379517 +
- **sha256 checksum** d0fe5c8e6ca3c902d054f854b70a85f928cba8b7c76391d356f05d2cf73b6887 +

This StorageGRID profile yaml file for Datadog metrics mapping generated following the instruction found [here](#). Place this file in /etc/datadog-agent/conf.d/snmp.d/profiles/ +

- [Download the profile yaml file](#) +
 - **md5 checksum** 72bb7784f4801adda4e0c3ea77df19aa +
 - **sha256 checksum** b6b7fadd33063422a8bb8e39b3ead8ab38349ee0229926eadc8585f0087b8cee +

SNMP Datadog configuration for Metrics

Configuring SNMP for metrics can be managed in two ways. You can configure for auto-discovery by providing a network address range containing the StorageGRID system(s), or define the IP's of the individual devices. The configuration location is different based on the decision made. Auto-discovery is defined in the datadog agent yaml file. Explicit device definitions are configured in the snmp configuration yaml file. Below are examples of each for the same StorageGRID system.

Auto-discovery

configuration located in /etc/datadog-agent/datadog.yaml

```
listeners:
  - name: snmp
snmp_listener:
  workers: 100 # number of workers used to discover devices concurrently
  discovery_interval: 3600 # interval between each autodiscovery in
seconds
  loader: core # use core check implementation of SNMP integration.
recommended
  use_device_id_as_hostname: true # recommended
  configs:
    - network_address: 10.0.0.0/24 # CIDR subnet
      snmp_version: 2
      port: 161
      community_string: 'st0r@gegrid' # enclose with single quote
      profile: netapp-storagegrid
```

Individual devices

/etc/datadog-agent/conf.d/snmp.d/conf.yaml

```

init_config:
  loader: core # use core check implementation of SNMP integration.
recommended
  use_device_id_as_hostname: true # recommended
instances:
- ip_address: '10.0.0.1'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid' # enclose with single quote
- ip_address: '10.0.0.2'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid'
- ip_address: '10.0.0.3'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid'
- ip_address: '10.0.0.4'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid'

```

SNMP configuration for traps

The configuration for SNMP traps is defined in the datadog configuration yaml file `/etc/datadog-agent/datadog.yaml`

```

network_devices:
  namespace: # optional, defaults to "default".
  snmp_traps:
    enabled: true
    port: 9162 # on which ports to listen for traps
    community_strings: # which community strings to allow for v2 traps
      - st0r@gegrid

```

Example StorageGRID SNMP configuration

The SNMP agent in your StorageGRID system is located under the configuration tab, Monitoring column. Enable SNMP and enter the desired information. If you wish to configure traps, select the "Traps Destinations" and Create a destination for the Datadog agent host containing the trap configuration.

SNMP Agent


You can configure SNMP for read-only MIB access and notifications. SNMPv1, SNMPv2c, SNMPv3 are supported. For SNMPv3, only User Security Model (USM) authentication is supported. All nodes in the grid share the same SNMP configuration.

Enable SNMP  ☒

System Contact 

System Location 

lab

Enable SNMP Agent Notifications  ☒

Enable Authentication Traps  ☐

Community Strings

Default Trap Community 

st0r@gegrid

Read-Only Community 

String 1

st0r@gegrid

+

Other Configurations

Agent Addresses (0)

USM Users (0)

Trap Destinations (1)

+ Create

Edit

Remove

| Version | Type | Host | Port | Protocol | Community/USM User |
|-------------------------------|--------|---------------|------|----------|--------------------------------|
| <input type="radio"/> SNMPv2C | Inform | 10.193.92.241 | 9162 | UDP | Default Community: st0r@gegrid |

By Aron Klein

Use rclone to migrate, PUT, and DELETE objects on StorageGRID

rclone is a free command line tool and client for S3 operations. You can use rclone to migrate, copy, and delete object data on StorageGRID. rclone includes the capability to delete buckets even when not empty with a "purge" function as seen in an example below.

Install and configure rclone

To install rclone on a workstation or server, download it from rclone.org.

Initial configuration steps

1. Create the rclone configuration file by either running the config script or manually creating the file.
2. For this example I will use sgdemo for the name of the remote StorageGRID S3 endpoint in the rclone configuration.
 - a. Create the config file ~/.config/rclone/rclone.conf

```
[sgdemo]
type = s3
provider = Other
access_key_id = ABCDEFGH123456789JKL
secret_access_key = 123456789ABCDEFGHIJKLMN0123456789PQRST+V
endpoint = sgdemo.netapp.com
```

- b. Run rclone config

rclone config

```
2023/04/13 14:22:45 NOTICE: Config file
"/root/.config/rclone/rclone.conf" not found - using defaults
No remotes found - make a new one
n) New remote
s) Set configuration password
q) Quit config
n/s/q> n
name> sgdemo
```

Option Storage.

Type of storage to configure.

Enter a string value. Press Enter for the default ("").

Choose a number from below, or type in your own value.

- 1 / lFichier
 \ "fichier"
- 2 / Alias for an existing remote
 \ "alias"
- 3 / Amazon Drive
 \ "amazon cloud drive"
- 4 / Amazon S3 Compliant Storage Providers including AWS,
Alibaba, Ceph, Digital Ocean, Dreamhost, IBM COS, Minio,
SeaweedFS, and Tencent COS
 \ "s3"
- 5 / Backblaze B2
 \ "b2"
- 6 / Better checksums for other remotes
 \ "hasher"
- 7 / Box
 \ "box"
- 8 / Cache a remote
 \ "cache"
- 9 / Citrix Sharefile
 \ "sharefile"
- 10 / Compress a remote
 \ "compress"
- 11 / Dropbox
 \ "dropbox"
- 12 / Encrypt/Decrypt a remote
 \ "crypt"
- 13 / Enterprise File Fabric
 \ "filefabric"
- 14 / FTP Connection

```
\ "ftp"
15 / Google Cloud Storage (this is not Google Drive)
   \ "google cloud storage"
16 / Google Drive
   \ "drive"
17 / Google Photos
   \ "google photos"
18 / Hadoop distributed file system
   \ "hdfs"
19 / Hubic
   \ "hubic"
20 / In memory object storage system.
   \ "memory"
21 / Jottacloud
   \ "jottacloud"
22 / Koofr
   \ "koofr"
23 / Local Disk
   \ "local"
24 / Mail.ru Cloud
   \ "mailru"
25 / Mega
   \ "mega"
26 / Microsoft Azure Blob Storage
   \ "azureblob"
27 / Microsoft OneDrive
   \ "onedrive"
28 / OpenDrive
   \ "opendrive"
29 / OpenStack Swift (Rackspace Cloud Files, Memset Memstore,
   OVH)
   \ "swift"
30 / Pcloud
   \ "pcloud"
31 / Put.io
   \ "putio"
32 / QingCloud Object Storage
   \ "qingstor"
33 / SSH/SFTP Connection
   \ "sftp"
34 / Sia Decentralized Cloud
   \ "sia"
35 / Sugarsync
   \ "sugarsync"
36 / Tardigrade Decentralized Cloud Storage
   \ "tardigrade"
```

```
37 / Transparently chunk/split large files
   \ "chunker"
38 / Union merges the contents of several upstream fs
   \ "union"
39 / Uptobox
   \ "uptobox"
40 / Webdav
   \ "webdav"
41 / Yandex Disk
   \ "yandex"
42 / Zoho
   \ "zoho"
43 / http Connection
   \ "http"
44 / premiumize.me
   \ "premiumizeme"
45 / seafile
   \ "seafile"
```

```
Storage> 4
```

Option provider.

Choose your S3 provider.

Enter a string value. Press Enter for the default ("").

Choose a number from below, or type in your own value.

```
1 / Amazon Web Services (AWS) S3
  \ "AWS"
2 / Alibaba Cloud Object Storage System (OSS) formerly Aliyun
  \ "Alibaba"
3 / Ceph Object Storage
  \ "Ceph"
4 / Digital Ocean Spaces
  \ "DigitalOcean"
5 / Dreamhost DreamObjects
  \ "Dreamhost"
6 / IBM COS S3
  \ "IBMCOS"
7 / Minio Object Storage
  \ "Minio"
8 / Netease Object Storage (NOS)
  \ "Netease"
9 / Scaleway Object Storage
  \ "Scaleway"
10 / SeaweedFS S3
  \ "SeaweedFS"
11 / StackPath Object Storage
  \ "StackPath"
12 / Tencent Cloud Object Storage (COS)
  \ "TencentCOS"
13 / Wasabi Object Storage
  \ "Wasabi"
14 / Any other S3 compatible provider
  \ "Other"
provider> 14
```

```
Option env_auth.  
Get AWS credentials from runtime (environment variables or  
EC2/ECS meta data if no env vars).  
Only applies if access_key_id and secret_access_key is blank.  
Enter a boolean value (true or false). Press Enter for the  
default ("false").  
Choose a number from below, or type in your own value.  
  1 / Enter AWS credentials in the next step.  
    \ "false"  
  2 / Get AWS credentials from the environment (env vars or IAM).  
    \ "true"  
env_auth> 1
```

```
Option access_key_id.  
AWS Access Key ID.  
Leave blank for anonymous access or runtime credentials.  
Enter a string value. Press Enter for the default ("").  
access_key_id> ABCDEFGH123456789JKL
```

```
Option secret_access_key.  
AWS Secret Access Key (password).  
Leave blank for anonymous access or runtime credentials.  
Enter a string value. Press Enter for the default ("").  
secret_access_key> 123456789ABCDEFGHIJKLMN0123456789PQRST+V
```

```
Option region.  
Region to connect to.  
Leave blank if you are using an S3 clone and you don't have a  
region.  
Enter a string value. Press Enter for the default ("").  
Choose a number from below, or type in your own value.  
  / Use this if unsure.  
  1 | Will use v4 signatures and an empty region.  
    \ ""  
  / Use this only if v4 signatures don't work.  
  2 | E.g. pre Jewel/v10 CEPH.  
    \ "other-v2-signature"  
region> 1
```

Option endpoint.

Endpoint for S3 API.

Required when using an S3 clone.

Enter a string value. Press Enter for the default ("").

endpoint> sgdemo.netapp.com

Option location_constraint.

Location constraint - must be set to match the Region.

Leave blank if not sure. Used when creating buckets only.

Enter a string value. Press Enter for the default ("").

location_constraint>

Option acl.

Canned ACL used when creating buckets and storing or copying objects.

This ACL is used for creating objects and if bucket_acl isn't set, for creating buckets too.

For more info visit

<https://docs.aws.amazon.com/AmazonS3/latest/dev/acl-overview.html#canned-acl>

Note that this ACL is applied when server-side copying objects as S3

doesn't copy the ACL from the source but rather writes a fresh one.

Enter a string value. Press Enter for the default ("").

Choose a number from below, or type in your own value.

```
    / Owner gets FULL_CONTROL.
1 | No one else has access rights (default).
  \ "private"
    / Owner gets FULL_CONTROL.
2 | The AllUsers group gets READ access.
  \ "public-read"
    / Owner gets FULL_CONTROL.
3 | The AllUsers group gets READ and WRITE access.
  | Granting this on a bucket is generally not recommended.
  \ "public-read-write"
    / Owner gets FULL_CONTROL.
4 | The AuthenticatedUsers group gets READ access.
  \ "authenticated-read"
    / Object owner gets FULL_CONTROL.
5 | Bucket owner gets READ access.
  | If you specify this canned ACL when creating a bucket,
Amazon S3 ignores it.
  \ "bucket-owner-read"
    / Both the object owner and the bucket owner get FULL_CONTROL
over the object.
6 | If you specify this canned ACL when creating a bucket,
Amazon S3 ignores it.
  \ "bucket-owner-full-control"
acl>
```

Edit advanced config?

y) Yes

n) No (default)

y/n> n


```

-----
[sgdemo]
type = s3
provider = Other
access_key_id = ABCDEFGH123456789JKL
secret_access_key = 123456789ABCDEFGHIJKLMN0123456789PQRST+V
endpoint = sgdemo.netapp.com:443
-----
y) Yes this is OK (default)
e) Edit this remote
d) Delete this remote
y/e/d>

```

Current remotes:

| Name | Type |
|--------|------|
| ==== | ==== |
| sgdemo | s3 |

```

e) Edit existing remote
n) New remote
d) Delete remote
r) Rename remote
c) Copy remote
s) Set configuration password
q) Quit config
e/n/d/r/c/s/q> q

```

Basic command examples

- **Create a bucket:**

```
rclone mkdir remote:bucket
```

```
# rclone mkdir sgdemo:test01
```



Use `--no-check-certificate` if you need to ignore SSL certificates.

- **List all buckets:**

```
rclone lsd remote:
```

```
# rclone lsd sgdemo:
```

- **List objects in a specific bucket:**

```
rclone ls remote:bucket
```

```
# rclone ls sgdemo:test01
```

```
65536 TestObject.0
65536 TestObject.1
65536 TestObject.10
65536 TestObject.12
65536 TestObject.13
65536 TestObject.14
65536 TestObject.15
65536 TestObject.16
65536 TestObject.17
65536 TestObject.18
65536 TestObject.2
65536 TestObject.3
65536 TestObject.5
65536 TestObject.6
65536 TestObject.7
65536 TestObject.8
65536 TestObject.9
33554432 bigobj
  102 key.json
   47 locked01.txt
4294967296 sequential-read.0.0
   15 test.txt
  116 version.txt
```

- **Delete a bucket:**

```
rclone rmdir remote:bucket
```

```
# rclone rmdir sgdemo:test02
```

- **Put an object:**

```
rclone copy filename remote:bucket
```

```
# rclone copy ~/test/testfile.txt sgdemo:test01
```

- **Get an object:**

```
rclone copy remote:bucket/objectname filename
```

```
# rclone copy sgdemo:test01/testfile.txt ~/test/testfileS3.txt
```

- **Delete an object:**

```
rclone delete remote:bucket/objectname
```

```
# rclone delete sgdemo:test01/testfile.txt
```

- **Migrate objects in a bucket**

```
rclone sync source:bucket destination:bucket --progress
```

```
rclone sync source_directory destination:bucket --progress
```

```
# rclone sync sgdemo:test01 sgdemo:clone01 --progress
```

```
Transferred:      4.032 GiB / 4.032 GiB, 100%, 95.484 KiB/s, ETA
0s
Transferred:      22 / 22, 100%
Elapsed time:      1m4.2s
```



Use `--progress` or `-P` to display the progress of the task. Otherwise there is no output.

- **Delete a bucket and all object contents**

```
rclone purge remote:bucket --progress
```

```
# rclone purge sgdemo:test01 --progress
```

```
Transferred:          0 B / 0 B, -, 0 B/s, ETA -  
Checks:          46 / 46, 100%  
Deleted:          23 (files), 1 (dirs)  
Elapsed time:      10.2s
```

```
# rclone ls sgdemo:test01
```

```
2023/04/14 09:40:51 Failed to ls: directory not found
```

By Siegfried Hepp and Aron Klein

StorageGRID best practices for deployment with Veeam Backup and Replication

This guide focuses on the configuration of NetApp StorageGRID and partly Veeam Backup and Replication. This paper is written for storage and network administrators who are familiar with Linux systems and tasked with maintaining or implementing a NetApp StorageGRID system in combination with Veeam Backup and Replication.

Overview

Storage Administrators are looking to manage the growth of their data with solutions that will meet the availability, rapid recovery goals, scale to meet their needs and automate their policy for long-term retention of data. These solutions should also provide protection from loss or malicious attacks. Together, Veeam and NetApp have partnered to create a data protection solution combining Veeam Backup & Recovery with NetApp StorageGRID for on-premises object storage.

Veeam and NetApp StorageGRID provide an easy-to-use solution that work together to help meet the demands of rapid data growth and increasing regulations around the world. Cloud-based object storage is known for its resilience, ability to scale, operational and cost efficiencies that make it a natural choice as a target for your backups. This document will provide guidance and recommendations for the configuration of your Veeam Backup solution and StorageGRID system.

The object workload from Veeam creates a large number of concurrent PUT, DELETE, and LIST operations of small objects. Enabling immutability will add to the number of requests to the object store for setting retention and listing versions. The process of a backup job includes writing objects for the daily change then after the new writes are complete the job will delete any objects based on the retention policy of the backup. The scheduling of backup jobs will almost always overlap. This overlap will result in a large portion of the backup window consisting of 50/50 PUT/DELETE workload on the object store. Making adjustments in Veeam to the number of concurrent operations with the task slot setting, increasing the object size by increasing the backup job block size, reducing the number of objects in the multi-object delete requests, and choosing the maximum time window for the jobs to complete will optimize the solution for performance and cost.

Make sure to read the product documentation for [Veeam Backup and Replication](#) and [StorageGRID](#) before you

begin. Veeam provides calculators for understanding the sizing of the Veeam infrastructure and capacity requirements that should be used prior to sizing your StorageGRID solution. Please always check the Veeam-NetApp validated configurations at the Veeam Ready Program website for [Veeam Ready Object, Object Immutability, and Repository](#).

Veeam configuration

Recommended version

It is always recommended to stay current and apply the latest hotfixes for your Veeam Backup & Replication 12 system. Currently we recommend at a minimum installing Veeam patch P20230718.

S3 Repository configuration

A scale-out backup repository (SOBR) is the capacity tier of S3 object storage. The capacity tier is an extension of the primary repository providing longer data retention periods and a lower cost storage solution. Veeam offers the ability to provide immutability through the S3 Object Lock API. Veeam 12 can use multiple buckets in a scale out repository. StorageGRID does not have a limit for the number of objects or capacity in a single bucket. Using multiple buckets may improve performance when backing up very large datasets where the backup data could get to petabyte scale in objects.

Limiting concurrent tasks may be required depending on the sizing of your specific solution and requirements. The default settings specify one repository task slot for each CPU core and for each task slot a concurrent task slot limit of 64. For example if your server has 2 CPU cores a total of 128 concurrent threads will be used for the object store. This is inclusive of PUT, GET, and batch Delete. It is recommended to select a conservative limit to the task slots to start with and tune this value once Veeam backups have reached a steady state of new backups and expiring backup data. Please work with your NetApp account team to size the StorageGRID system appropriately to meet the desired time windows and performance. Adjusting the number of task slots and the limit of tasks per slot may be required to provide the optimal solution.

Backup job configuration

Veeam backup jobs can be configured with different block size options that should be considered carefully. The default block size is 1MB and with the storage efficiencies Veeam provides with compression and deduplication creates object sizes of approximately 500kB for the initial Full backup and 100-200kB objects for the incremental jobs. We can greatly increase performance and scale down the requirements for the object store by choosing a larger backup block size. Though the larger block size makes great improvements in the object store performance it comes at the cost of potentially increased primary storage capacity requirement due to reduced storage efficiency performance. It is recommended for the backup jobs to be configured with a 4MB block size which creates approximately 2MB objects for the full backups and 700kB-1MB object sizes for incrementals. Customers may consider even configuring backup jobs using 8 MB block size, which can be enabled with assistance from Veeam support.

The implementation of immutable backups makes use of S3 Object Lock on the object store. The immutability option generates an increased number of requests to the object store for listing and retention updates on the objects.

As backup retentions expire the backup jobs will process the deletion of objects. Veeam sends the delete requests to the object store in multi-object delete requests of 1000 objects per request. For small solutions this may need to be adjusted to reduce the number of objects per request. Lowering this value will have the added benefit of more evenly distributing the delete requests across the nodes in the StorageGRID system. It is recommended to use the values in the table below as a starting point in configuring the multi object delete limit. Multiply the value in the table by the number of nodes for the chosen appliance type to get the value for the setting in Veeam. If this value is equal to or greater than 1000 there is no need to adjust the default value. If

this value needs to be adjusted, please work with Veeam support to make the change.

| Appliance Model | S3MultiObjectDeleteLimit per node |
|-----------------|-----------------------------------|
| SG5712 | 34 |
| SG5760 | 75 |
| SG6060 | 200 |



Please work with your NetApp Account team for the recommended configuration based on your specific needs. The Veeam configuration settings recommendations will include:

- Backup job block size = 4MB
- SOBR task slot limit= 2-16
- Multi Object Delete Limit = 34-1000

StorageGRID configuration

Recommended version

NetApp StorageGRID 11.6 or 11.7 with the latest hotfix are the recommended versions for Veeam deployments. Many optimization features were introduced in the StorageGRID 11.6.0.11 and 11.7.0.4 which will be beneficial to Veeam workloads. It is always recommended to stay current and apply the latest hotfixes for your StorageGRID system.

Load balancer and S3 endpoint configuration

Veeam requires the endpoint to be connected via HTTPS only. A non-encrypted connection is not supported by Veeam. The SSL certificate can be a self-signed certificate, private trusted certificate authority, or public trusted certificate authority. To ensure continuous access to the S3 repository it is recommended to use at least two load balancers in an HA configuration. The load balancers can be a StorageGRID provided integrated load balancer service located on every admin node and gateway node or third-party solution such as F5, Kemp, HAproxy, Loadbalancer.org, etc. Using a StorageGRID load balancer will provide the ability to set traffic classifiers (QoS rules) that can prioritize the Veeam workload, or limit Veeam to not impact higher priority workloads on the StorageGRID system.

S3 Bucket

StorageGRID is a secure multi-tenant storage system. It is recommended to create a dedicated tenant for the Veeam workload. A storage quota can be optionally assigned. As a best practice enable “use own identity source”. Secure the tenant root management user with an appropriate password. Veeam Backup 12 requires strong consistency for S3 buckets. StorageGRID offers multiple consistency options configured at the bucket level. For multi-site deployments with Veeam accessing the data from multiple locations, select “strong-global”. If Veeam backups and restores happen at a single site only, consistency level should be set to “strong-site”. For more information on bucket consistency levels please review the [documentation](#). To use StorageGRID for Veeam immutability backups, S3 Object Lock must be enabled globally and configured on the bucket during the bucket creation.

Lifecycle management

StorageGRID supports replication and erasure coding for object level protection across StorageGRID nodes and sites. Erasure Coding requires at least a 200kB object size. The default block size for Veeam of 1MB

produces object sizes that can often be below this 200kB recommended minimum size after Veeam's storage efficiencies. For the performance of the solution, it is not recommended to use an erasure coding profile spanning multiple sites unless the connectivity between the sites is sufficient to not add latency or restrict the bandwidth of the StorageGRID system. In a multi-site StorageGRID system the ILM rule can be configured to store a single copy at each site. For ultimate durability a rule could be configured to store an erasure coded copy at each site. Using two copies local to the Veeam Backup servers is the most recommended implementation for this workload.


Implementation key points

StorageGRID

Ensure Object Lock is enabled on the StorageGRID system if immutability is required. Find the option in the management UI under Configuration/S3 Object Lock.

Configuration > S3 Object Lock

S3 Object Lock

 S3 Object Lock has been enabled for the grid and cannot be disabled.

Enable S3 Object Lock for your entire StorageGRID system if S3 tenant accounts need to satisfy regulatory compliance requirements when saving object data. After this setting is enabled, it cannot be disabled.

Before enabling S3 Object Lock, you must ensure that the default rule in the active ILM policy is compliant. A compliant rule satisfies the requirements of buckets with S3 Object Lock enabled.

- It must create at least two replicated object copies or one erasure-coded copy.
- These copies must exist on Storage Nodes for the entire duration of each line in the placement instructions.
- Object copies cannot be saved in a Cloud Storage Pool.
- Object copies cannot be saved on Archive Nodes.
- At least one line of the placement instructions must start at day 0, using Ingest Time as the reference time.
- At least one line of the placement instructions must be "forever".

☒ Enable S3 Object Lock


Apply

When creating the bucket, select "Enable S3 Object Lock" if this bucket is to be used for immutability backups. This will automatically enable bucket versioning. Leave default retention disabled as Veeam will set object retention explicitly. Versioning and S3 Object Lock should not be selected if Veeam isn't creating immutable backups.

Manage object settings Optional

Object versioning

Enable object versioning if you want to store every version of each object in this bucket. You can then retrieve previous versions of an object as needed.

 Object versioning has been enabled automatically because this bucket has S3 Object Lock enabled.

☒ Enable object versioning

S3 Object Lock

S3 Object Lock allows you to specify retention and legal hold settings for the objects ingested into a bucket. If you want to use S3 Object Lock, you must enable this setting when you create the bucket. You cannot add or disable S3 Object Lock after a bucket is created.

If S3 Object Lock is enabled, object versioning is enabled for the bucket automatically and cannot be suspended.

☒ Enable S3 Object Lock

Default retention 

Automatically protect new objects put into this bucket from being deleted or overwritten.

☒ Disable

☐ Enable

Once the bucket is created go to the details page of the bucket created. Select the consistency level.

Buckets > veeam12

veeam12

Region: us-east-1
 S3 Object Lock: Enabled
 Date created: 2023-09-21 08:01:38 GMT
 Object count: 0

[View bucket contents in Experimental S3 Console](#)

[Delete objects in bucket](#) [Delete bucket](#)

Bucket options [Bucket access](#) [Platform services](#)

| | | |
|--------------------------|--------------------------------|---|
| Consistency level | Read-after-new-write (default) | ▼ |
| Last access time updates | Disabled | ▼ |
| Object versioning | Enabled | ▼ |
| S3 Object Lock | Enabled | ▼ |

Veeam requires strong consistency for S3 buckets. So, for multi-site deployments with Veeam accessing the data from multiple locations, select “strong-global”. If Veeam backups and restores happen at a single site only, consistency level should be set to “strong-site”. Save the changes.

Bucket options [Bucket access](#) [Platform services](#)

Consistency level Read-after-new-write (default) ▲

Change the consistency control for operations performed on the objects in the bucket. Consistency levels provide a balance between the availability of objects and the consistency of those objects across different Storage Nodes and sites.

In general, use the **Read-after-new-write** consistency level for your buckets. Then, if objects do not meet availability or consistency requirements, change the client application's behavior, or set the Consistency-Control header for an individual API request, which overrides the bucket setting.

☐ All
Provides the highest guarantee of consistency. All nodes receive the data immediately, or the request will fail.

☒ Strong-global
Guarantees read-after-write consistency for all client requests across all sites.

☐ Strong-site
Guarantees read-after-write consistency for all client requests within a site.

☐ Read-after-new-write (default)
Provides read-after-write consistency for new objects and eventual consistency for object updates. Offers high availability and data protection guarantees. Recommended for most cases.

☐ Available
Provides eventual consistency for both new objects and object updates. For S3 buckets, use only as required (for example, for a bucket that contains log values that are rarely read, or for HEAD or GET operations on keys that do not exist). Not supported for FabricPool buckets.

[Save changes](#)

Last access time updates Disabled ▼

StorageGRID provides an integrated load balancer service on every admin node and dedicated gateway nodes. One of the many advantages of using this load balancer is the ability to configure Traffic Classification

Policies (QoS). Though these are mainly used for limiting an applications impact on other client workloads or prioritizing a workload over others, they also provide a bonus of additional metrics collection to assist in monitoring.

In the configuration tab, select “Traffic Classification” and create a new policy. Name the rule and select either the bucket(s) or tenant as the type. Enter the name(s) of the bucket(s) or tenant. If QoS is required, set a limit, but for most implementations, we just want to add the monitoring benefits this provides so do not set a limit.

Create a traffic classification policy

You can create traffic classification policies to monitor the network traffic for specific buckets, tenants, IP addresses, subnets, or load balancer endpoints. You can optionally limit this traffic based on bandwidth, number of concurrent requests, or the request rate.

✓ Enter policy name

—

✓ Add matching rules

—

✓ Set limits

—

4 Review the policy

Review the policy

Policy name:

Veeam

Description:

Policy to monitor Veeam bucket traffic

Matching rules

| Type ? | Match value ? | Inverse match ? |
|--------|-----------------|-----------------|
| Bucket | <div>test</div> | No |

Veeam

Depending on the model and quantity of StorageGRID appliances it may be necessary to select and configure a limit to the number of concurrent operations on the bucket.

New Object Storage Repository

Name
Type in a name and description for this object storage repository.

Name:
Object storage repository 1

Description:
Created by SRV92\Administrator at 2/3/2021 8:15 AM.

☒ Limit concurrent tasks to: 2

Use this setting to limit the maximum number of tasks that can be processed concurrently in cases when your object storage is overloaded or cannot keep up with the number of API requests issued by multiple object storage offload tasks.

< Previous Next > Finish Cancel

Follow the Veeam documentation on backup job configuration in the Veeam console to start the wizard. After adding VMs select the SOBR repository.

Edit Backup Job vm backup 4mb

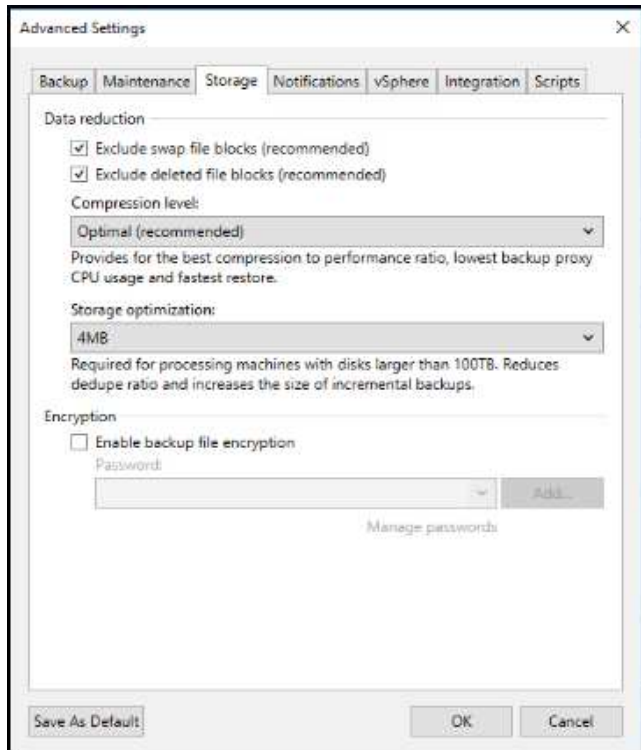
Storage
Specify processing proxy server to be used for source data retrieval, backup repository to store the backup files produced by this job and customize advanced job settings if required.

Name:
Virtual Machines

Storage:
Backup proxy: Automatic selection
Backup repository: baremetal 4mb (Created by MUCCBC\phaensel at 14.03.2023 15:21.)
N/A
Retention policy: 30 days
☒ Keep certain full backups longer for archival purposes
6 weekly, 3 monthly
☐ Configure secondary destinations for this job
Copy backups produced by this job to another backup repository, or tape. We recommend to make at least one copy of your backups to a different storage device that is located off-site.
Advanced job settings include backup mode, compression and deduplication, block size, notification settings, automated post-job activity and other settings.

< Previous Next > Finish Cancel

Click Advanced settings and change storage optimization settings to 4 MB or larger. Compression and deduplication shall be enabled. Change guest settings according to your requirements and configure the backup job schedule.



Monitoring StorageGRID

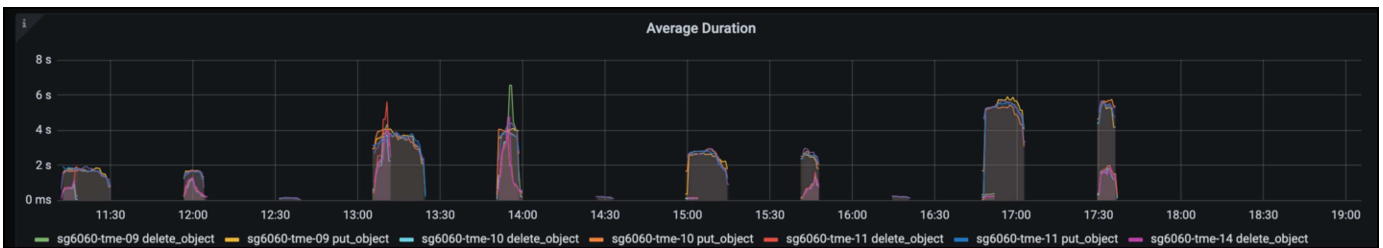
To get the full picture of how Veeam and StorageGRID are performing together you will need to wait until the retention time of the first backups have expired. Up until this point the Veeam workload consists primarily of PUT operations and no DELETES have occurred. Once there is backup data expiring and cleanups are occurring you can now see the full consistent usage in the object store and adjust the settings in Veeam if needed.

StorageGRID provides convenient charts to monitor the operation of the system located in the Support tab Metrics page. The primary dashboards to look at will be the S3 Overview, ILM, and Traffic Classification Policy if a policy was created. In the S3 Overview dashboard you will find information on the S3 operation rates, latencies, and request responses.

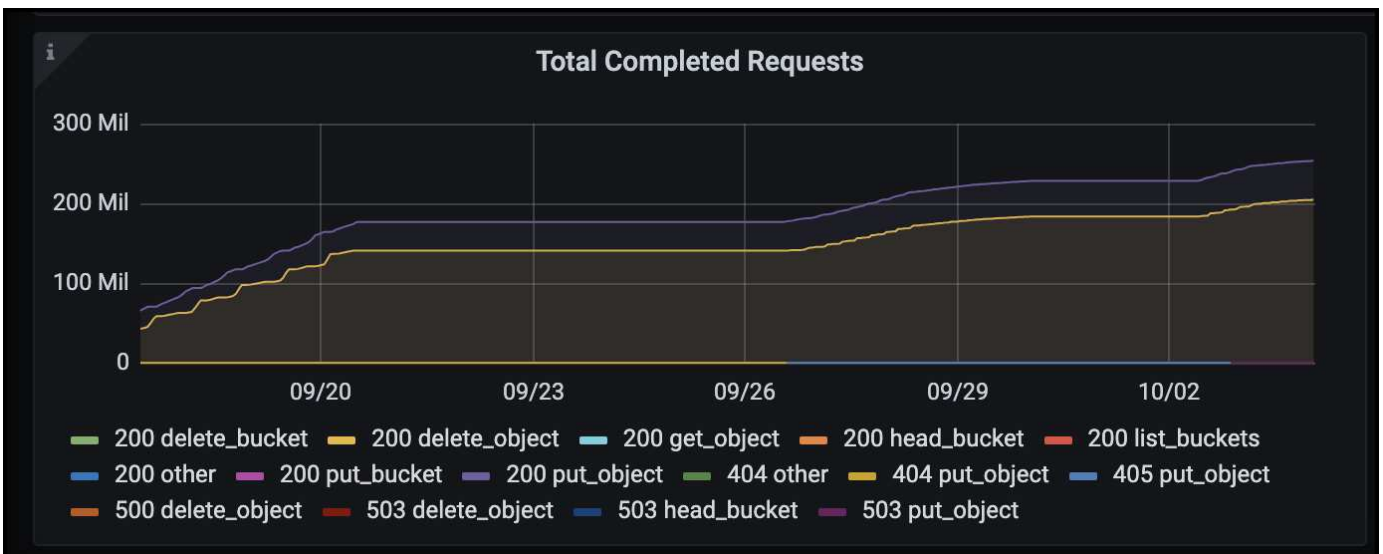
Looking at the S3 rates and active requests you can see how much of the load each node is handling and the overall number of requests by type.



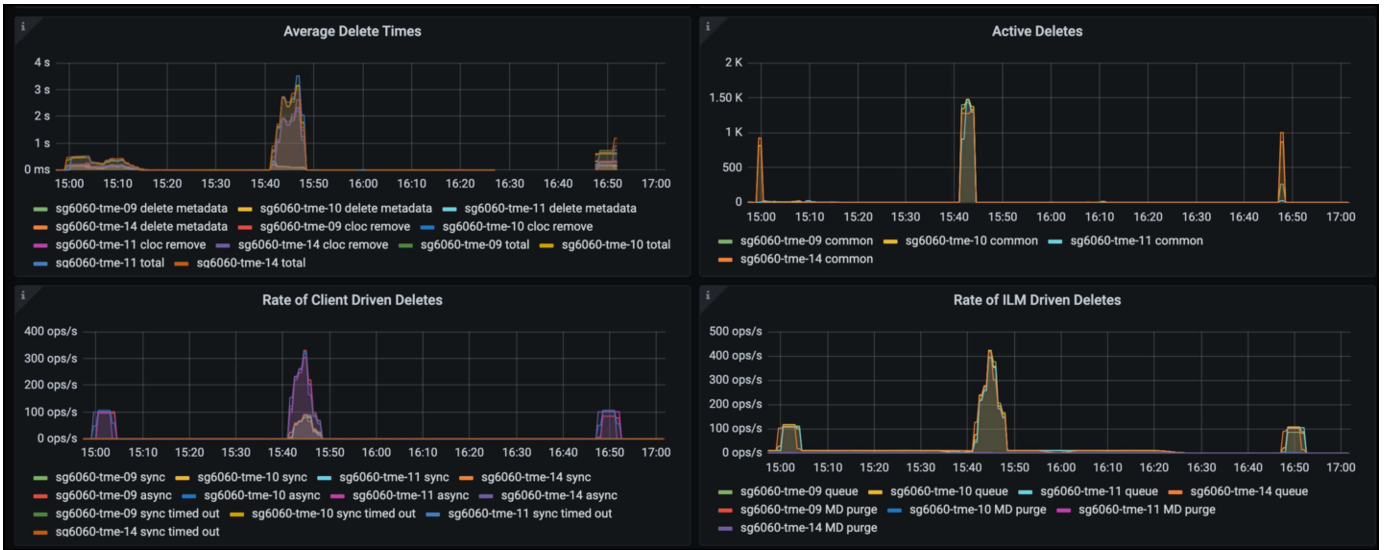
The Average Duration chart shows the average time each node is taking for each request type. This is the average latency of the request and may be a good indicator that additional tuning may be required, or there is room for the StorageGRID system to take on more load.



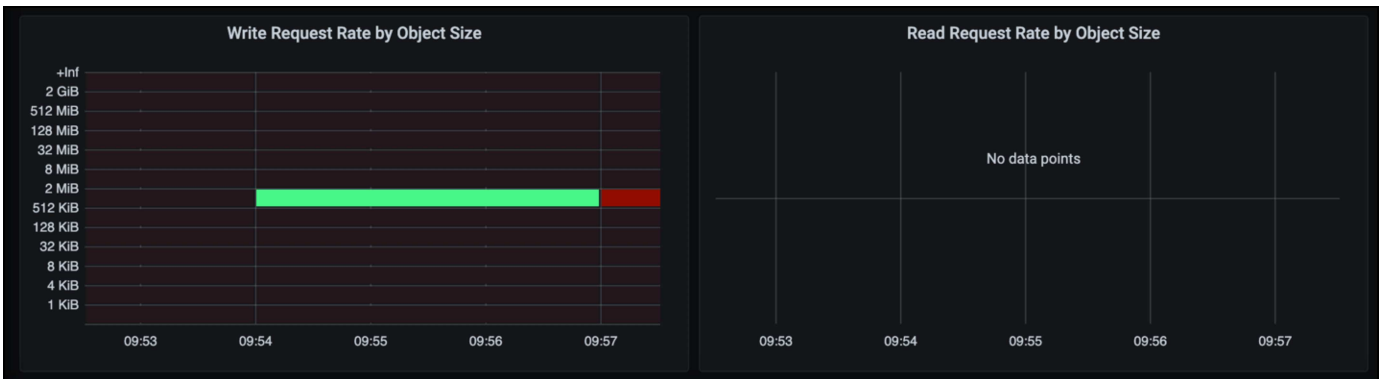
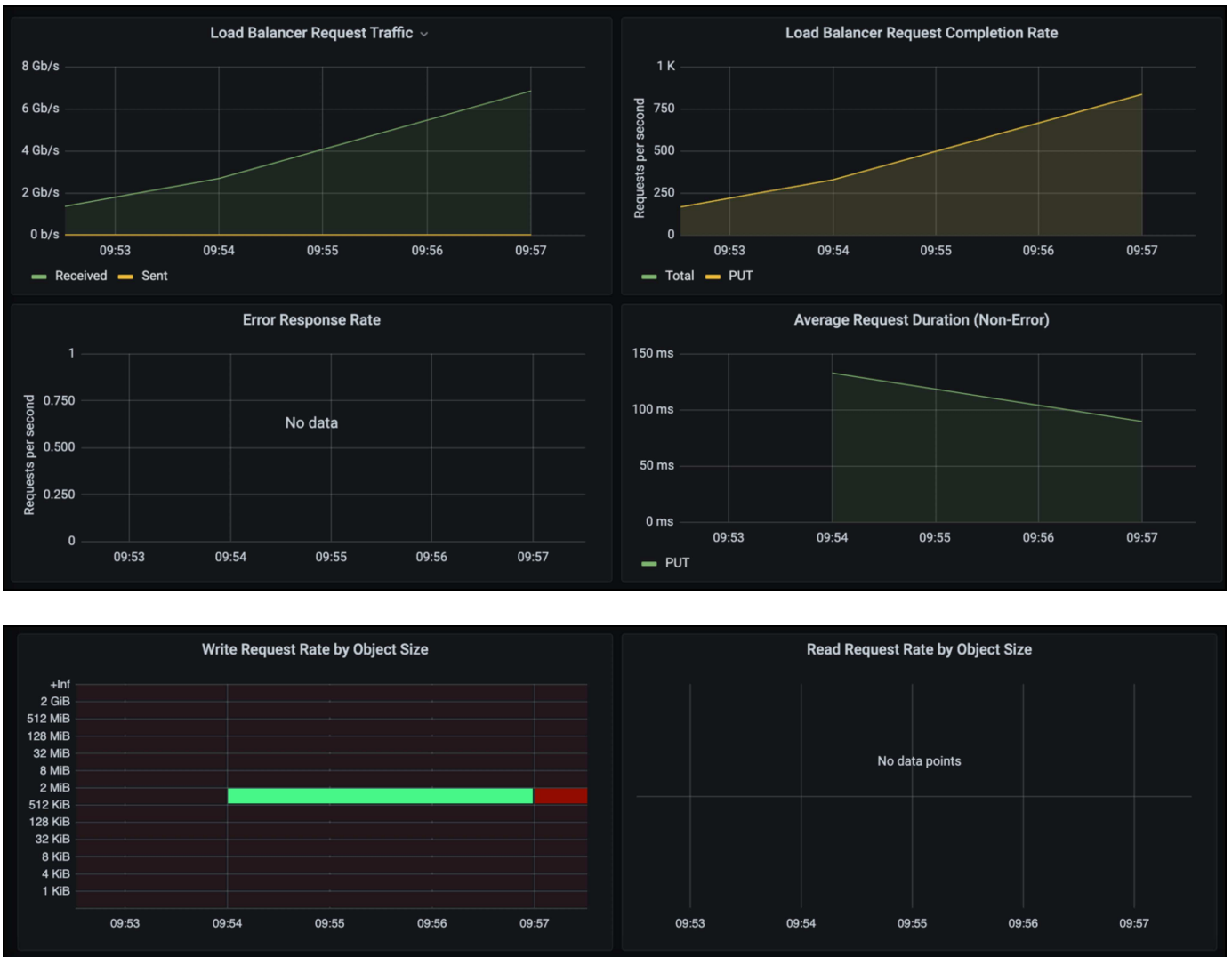
In the Total Completed Requests chart, you can see the requests by type and response codes. If you see responses other than 200 (Ok) for the responses this may indicate an issue like the StorageGRID system is getting heavily loaded sending 503 (Slow Down) responses and some additional tuning may be necessary, or the time has come to expand the system for the increased load.



In the ILM Dashboard you can monitor the Delete performance of your StorageGRID system. StorageGRID uses a combination of synchronous and asynchronous deletes on each node to try and optimize the overall performance for all requests.



With a Traffic Classification Policy, we can view metrics on the load balancer Request throughput, rates, duration, as well as the object sizes Veeam is sending and receiving.



Where to find additional information

To learn more about the information that is described in this document, review the following documents and/or websites:

- [NetApp StorageGRID 11.7 Product Documentation](#)
- [Veeam Backup and Replication](#)

By Oliver Haensel and Aron Klein

Configure Dremio data source with StorageGRID

Dremio supports a variety of data sources, including cloud-based or on-premises object storage. You can configure Dremio to use StorageGRID as object storage data source.

Configure Dremio data source

Prerequisites

- A StorageGRID S3 endpoint URL, a tenant s3 access key ID, and secret access key.
- StorageGRID configuration recommendation: disable compression (disabled by default).
Dremio uses byte range GET to fetch different byte ranges from within the same object concurrently during query. Typical size for byte-range requests is 1MB. Compressed object degrades byte-range GET performance.

Dremio guide

[Connecting to Amazon S3 - Configuring S3-Compatible Storage.](#)

Instruction

1. On Dremio Datasets page, click + sign to add a source, select 'Amazon S3'.
2. Enter a name for this new data source, StorageGRID S3 tenant access key ID and secret access key.
3. Check the box 'Encrypt connection' if using https for connection to StorageGRID S3 endpoint.
If using self-signed CA cert for this s3 endpoint, follow Dremio guide instruction to add this CA cert into Dremio server's <JAVA_HOME>/jre/lib/security

Sample screenshot


General

Advanced Options

Reflection Refresh

Metadata

Privileges



Amazon S3 Source

Name

parquet-1tb

Authentication

☒ AWS Access Key
 ☐ EC2 Metadata
 ☐ AWS Profile
 ☐ No Authentication

All or allowlisted (if specified) buckets associated with this access key or IAM role to assume (if specified) will be available.

AWS Access Key

XXXXXXXXXXXXXXXXXXXX

AWS Access Secret

.....


IAM Role to Assume

☒ Encrypt connection

Public Buckets

Buckets

No public buckets added

 Add bucket

- Click 'Advanced Options', check 'Enable compatibility mode'
- Under Connection properties, click + Add Properties and add these s3a properties.
- fs.s3a.connection.maximum default is 100. If your s3 datasets include large Parquet files with 100 or more columns, must enter a value greater than 100. Refer to Dremio guide for this setting.

| Name | Value |
|---------------------------|--------------------------------|
| fs.s3a.endpoint | <StorageGRID S3 endpoint:port> |
| fs.s3a.path.style.access | true |
| fs.s3a.connection.maximum | <a value greater than 100> |

Sample screenshot

118

General

Advanced Options

Reflection Refresh
Metadata
Privileges

☒ Enable asynchronous access when possible
☒ Enable compatibility mode
☐ Apply requester-pays to S3 requests
☒ Enable file status check
☐ Enable partition column inference

Root Path

Server side encryption key ARN

Default CTAS Format

PARQUET

Connection Properties

| Name | Value | |
|--|--|---|
| <input type="text" value="fs.s3a.path.style.access"/> | <input type="text" value="true"/> | × |
| <input type="text" value="fs.s3a.endpoint"/> | <input type="text" value="sgdemo.netapp.com"/> | × |
| <input type="text" value="fs.s3a.connection.maximum"/> | <input type="text" value="1000"/> | × |

+ Add property

Allowlisted buckets

No allowlisted buckets added


+ Add bucket


Cache Options

☒ Enable local caching when possible


Max percent of total available cache space to use when possible

- Configure other Dremio options as per your organization or application requirements.
 - Click the Save button to create this new data source.
 - Once StorageGRID data source is added successfully, a list of buckets will be displayed on the left panel.
- Sample screenshot**







Datasets




hdp-user



Spaces (0)







No spaces yet


[Add space](#)

Sources





Object Storage (2)




StorageGRID


0

StorageGRID


Name ↑




apache-hive




cdp-cluster




cdp-tera




databrick-tpcds




delta-lake



dcluster-tpcds



dremio-10g-csv



dremio-csv

By Angela Cheng

Procedures and API examples

Test and demonstrate S3 encryption options on StorageGRID

StorageGRID and the S3 API offer a number of different ways to encrypt your data at rest. To learn more, see [Review StorageGRID encryption methods](#).

This guide will demonstrate the S3 API encryption methods.

Server Side Encryption (SSE)

SSE allows the client to store an object and encrypt it with a unique key that is managed by StorageGRID. When the object is requested, the object is decrypted by the key stored in storageGRID.

SSE Example

- PUT an object with SSE

```
aws s3api put-object --bucket <bucket> --key <file> --body "<file>"  
--server-side-encryption AES256 --endpoint-url https://s3.example.com
```

- HEAD the object to verify encryption

```
aws s3api head-object --bucket <bucket> --key <file> --endpoint-url  
https://s3.example.com
```

```
{  
  "AcceptRanges": "bytes",  
  "LastModified": "2022-05-02T19:03:03+00:00",  
  "ContentLength": 47,  
  "ETag": "\"82e8bfb872e778a4687a26e6c0b36bc1\"",  
  "ContentType": "text/plain",  
  "ServerSideEncryption": "AES256",  
  "Metadata": {}  
}
```

- GET the object

```
aws s3api get-object --bucket <bucket> --key <file> <file> --endpoint  
-url https://s3.example.com
```

Server Side Encryption with Customer provided keys (SSE-C)

SSE allows the client to store an object and encrypt it with a unique key that is provided by the client with the object. When the object is requested, the same key must be provided in order to decrypt and return the object.

SSE-C Example

- For testing or demonstration purposes you can create an encryption key
 - Create an encryption key

```
openssl enc -aes-128-cbc -pass pass:secret -P`
```

```
salt=E9DBB6603C7B3D2A  
key=23832BAC16516152E560F933F261BF03  
iv =71E87C0F6EC3C45921C2754BA131A315
```

- Put an object with the generated key

```
aws s3api put-object --bucket <bucket> --key <file> --body "file" --sse  
-customer-algorithm AES256 --sse-customer-key  
23832BAC16516152E560F933F261BF03 --endpoint-url https://s3.example.com
```

- Head the object

```
aws s3api head-object --bucket <bucket> --key <file> --sse-customer  
-algorithm AES256 --sse-customer-key 23832BAC16516152E560F933F261BF03  
--endpoint-url https://s3.example.com
```

```
{  
  "AcceptRanges": "bytes",  
  "LastModified": "2022-05-02T19:20:02+00:00",  
  "ContentLength": 47,  
  "ETag": "\"f92ef20ab87e0e13951d9bee862e9f9a\"",  
  "ContentType": "binary/octet-stream",  
  "Metadata": {},  
  "SSECustomerAlgorithm": "AES256",  
  "SSECustomerKeyMD5": "rjGuMdjLpPV1eRuotNaPMQ=="  
}
```



If you do not provide the encryption key, you will receive an error "An error occurred (404) when calling the HeadObject operation: Not Found"

- Get the object

```
aws s3api get-object --bucket <bucket> --key <file> <file> --sse
--customer-algorithm AES256 --sse-customer-key
23832BAC16516152E560F933F261BF03 --endpoint-url https://s3.example.com
```



If you do not provide the encryption key, you will receive an error "An error occurred (InvalidRequest) when calling the GetObject operation: The object was stored using a form of Server Side Encryption. The correct parameters must be provided to retrieve the object."

Bucket Server Side Encryption (SSE-S3)

SSE-S3 allows the client to define a default encryption behavior for all objects stored in a bucket. The objects are encrypted with a unique key that is managed by StorageGRID. When the object is requested, the object is decrypted by the key stored in storageGRID.

Bucket SSE-S3 Example

- Create a new bucket and set a default encryption policy
 - Create new bucket

```
aws s3api create-bucket --bucket <bucket> --region us-east-1
--endpoint-url https://s3.example.com
```

- Put bucket encryption

```
aws s3api put-bucket-encryption --bucket <bucket> --server-side
--encryption-configuration '{"Rules":
[{"ApplyServerSideEncryptionByDefault": {"SSEAlgorithm":
"AES256"}}]}' --endpoint-url https://s3.example.com
```

- Put an object in the bucket

```
aws s3api put-object --bucket <bucket> --key <file> --body "file"
--endpoint-url https://s3.example.com
```

- Head the object

```
aws s3api head-object --bucket <bucket> --key <file> --endpoint-url
https://s3.example.com
```

```
{
  "AcceptRanges": "bytes",
  "LastModified": "2022-05-02T20:16:23+00:00",
  "ContentLength": 47,
  "ETag": "\"82e8bfb872e778a4687a26e6c0b36bc1\"",
  "ContentType": "binary/octet-stream",
  "ServerSideEncryption": "AES256",
  "Metadata": {}
}
```

- GET the object

```
aws s3api get-object --bucket <bucket> --key <file> <file> --endpoint
-url https://s3.example.com
```

By Aron Klein

Test and demonstrate S3 object lock on StorageGRID

Object Lock provides a WORM model to prevent objects from being deleted or overwritten. StorageGRID implementation of object lock is Cohasset assessed to help meet regulatory requirements, supporting legal hold and compliance mode for object retention, and default bucket retention policies.

This guide will demonstrate the S3 Object Lock API.

Legal hold

- Object Lock legal hold is a simple on/off status applied to an object.

```
aws s3api put-object-legal-hold --bucket <bucket> --key <file> --legal
-hold Status=ON --endpoint-url https://s3.company.com
```

- Verify it with a GET operation.

```
aws s3api get-object-legal-hold --bucket <bucket> --key <file>
--endpoint-url https://s3.company.com
```

```
{
  "LegalHold": {
    "Status": "ON"
  }
}
```

- Turn legal hold off

```
aws s3api put-object-legal-hold --bucket <bucket> --key <file> --legal
--hold Status=OFF --endpoint-url https://s3.company.com
```

- Verify it with a GET operation.

```
aws s3api get-object-legal-hold --bucket <bucket> --key <file>
--endpoint-url https://s3.company.com
```

```
{
  "LegalHold": {
    "Status": "OFF"
  }
}
```

Compliance mode

- The object retention is done with a retain until timestamp.

```
aws s3api put-object-retention --bucket <bucket> --key <file>
--retention '{"Mode":"COMPLIANCE", "RetainUntilDate": "2025-06-
10T16:00:00"}' --endpoint-url https://s3.company.com
```

- Verify the retention status

```
aws s3api get-object-retention --bucket <bucket> --key <file> --endpoint
-url https://s3.company.com
+
```

```
{
  "Retention": {
    "Mode": "COMPLIANCE",
    "RetainUntilDate": "2025-06-10T16:00:00+00:00"
  }
}
```

Default retention

- Set the retention period in days and years verses a retain until date defined with the per object api.

```
aws s3api put-object-lock-configuration --bucket <bucket> --object-lock
-configuration '{"ObjectLockEnabled": "Enabled", "Rule": {
"DefaultRetention": { "Mode": "COMPLIANCE", "Days": 10 }}}' --endpoint
-url https://s3.company.com
```

- Verify the retention status

```
aws s3api get-object-lock-configuration --bucket <bucket> --endpoint-url
https://s3.company.com
```

```
{
  "ObjectLockConfiguration": {
    "ObjectLockEnabled": "Enabled",
    "Rule": {
      "DefaultRetention": {
        "Mode": "COMPLIANCE",
        "Days": 10
      }
    }
  }
}
```

- Put an object in the bucket

```
aws s3api put-object --bucket <bucket> --key <file> --body "file"
--endpoint-url https://s3.example.com
```

- The retention duration set on the bucket is converted to a retention timestamp on the object.


```
aws s3api get-object-retention --bucket <bucket> --key <file> --endpoint-url https://s3.company.com
```

```
{
  "Retention": {
    "Mode": "COMPLIANCE",
    "RetainUntilDate": "2022-03-02T15:22:47.202000+00:00"
  }
}
```

Test deleting an object with a defined retention

Object Lock is built on top of versioning. The retention is defined on a version of the object. If an attempt is made to delete an object with a retention defined, and no version is specified, a delete marker is created as the current version of the object.

- Delete the object with retention defined

```
aws s3api delete-object --bucket <bucket> --key <file> --endpoint-url https://s3.example.com
```

- List the objects in the bucket

```
aws s3api list-objects --bucket <bucket> --endpoint-url https://s3.example.com
```

- Notice the object is not listed.
- List versions to see the delete marker, and the original locked version

```
aws s3api list-object-versions --bucket <bucket> --prefix <file> --endpoint-url https://s3.example.com
```

```
{
  "Versions": [
    {
      "ETag": "\"82e8bfb872e778a4687a26e6c0b36bc1\"",
      "Size": 47,
      "StorageClass": "STANDARD",
      "Key": "file.txt",
      "VersionId":
"RDVDMjYwMTQtQkNDQS0xMUVDLThGOEUtNjQ3NTAwQzAxQTkl",
      "IsLatest": false,
      "LastModified": "2022-04-15T14:46:29.734000+00:00",
      "Owner": {
        "DisplayName": "Tenant01",
        "ID": "56622399308951294926"
      }
    }
  ],
  "DeleteMarkers": [
    {
      "Owner": {
        "DisplayName": "Tenant01",
        "ID": "56622399308951294926"
      },
      "Key": "file01.txt",
      "VersionId":
"QjVDQzgZOTAtQ0FGNi0xMUVDLThFMzgtQ0RGMjAwQjk0MjMl",
      "IsLatest": true,
      "LastModified": "2022-05-03T15:35:50.248000+00:00"
    }
  ]
}
```

- Delete the locked version of the object

```
aws s3api delete-object --bucket <bucket> --key <file> --version-id
"<VersionId>" --endpoint-url https://s3.example.com
```

An error occurred (AccessDenied) when calling the DeleteObject operation: Access Denied

By Aron Klein

Example bucket and Group(IAM) policies

Here are examples of bucket policies and group policies(IAM Policies).

Group Policies (IAM)

Home Directory style bucket access

This group policy will only allow users to access objects in the bucket named the users username.

```
"Statement": [
  {
    "Sid": "AllowListBucketOfASpecificUserPrefix",
    "Effect": "Allow",
    "Action": "s3:ListBucket",
    "Resource": "arn:aws:s3:::home",
    "Condition": {
      "StringLike": {
        "s3:prefix": "${aws:username}/*"
      }
    }
  },
  {
    "Sid": "AllowUserSpecificActionsOnlyInTheSpecificUserPrefix",
    "Effect": "Allow",
    "Action": "s3:*Object",
    "Resource": "arn:aws:s3:::home/??/${aws:username}/*"
  }
]
```

Deny object lock bucket creation

This group policy will restrict users from creating a bucket with object lock enabled on the bucket.



This policy is not enforced in the StorageGRID UI, it is only enforced by S3 API.

```

{
  "Statement": [
    {
      "Action": "s3:*",
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::*"
    },
    {
      "Action": [
        "s3:PutBucketObjectLockConfiguration",
        "s3:PutBucketVersioning"
      ],
      "Effect": "Deny",
      "Resource": "arn:aws:s3:::*"
    }
  ]
}

```

Object lock retention limit

This Bucket policy will restrict Object-Lock retention duration to 10 days or less

```

{
  "Version": "2012-10-17",
  "Id": "CustSetRetentionLimits",
  "Statement": [
    {
      "Sid": "CustSetRetentionPeriod",
      "Effect": "Deny",
      "Principal": "*",
      "Action": [
        "s3:PutObjectRetention"
      ],
      "Resource": "arn:aws:s3:::testlock-01/*",
      "Condition": {
        "NumericGreaterThan": {
          "s3:object-lock-remaining-retention-days": "10"
        }
      }
    }
  ]
}

```

Restrict users from deleting objects by versionID

This group policy will restrict users from deleting versioned objects by versionID

```
{
  "Statement": [
    {
      "Action": [
        "s3:DeleteObjectVersion"
      ],
      "Effect": "Deny",
      "Resource": "arn:aws:s3:::*"
    },
    {
      "Action": "s3:*",
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::*"
    }
  ]
}
```

This bucket policy will restrict a user(identified by userID "56622399308951294926") from deleting versioned objects by versionID

```

{
  "Statement": [
    {
      "Action": [
        "s3:DeleteObjectVersion"
      ],
      "Effect": "Deny",
      "Resource": "arn:aws:s3:::verdeny/*",
      "Principal": {
        "AWS": [
          "56622399308951294926"
        ]
      }
    },
    {
      "Action": "s3:*",
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::verdeny/*",
      "Principal": {
        "AWS": [
          "56622399308951294926"
        ]
      }
    }
  ]
}

```

Restrict bucket to single user with read-only access

This policy allows a single user to have read-only access to a bucket and explicitly denys access to all other users. Grouping the Deny statements at the top of the policy is a good practice for faster evaluation.

```

{
  "Statement": [
    {
      "Sid": "Deny non user1",
      "Effect": "Deny",
      "NotPrincipal": {
        "AWS":
"urn:sgws:identity::34921514133002833665:user/user1"
      },
      "Action": [
        "s3:*"
      ],
      "Resource": [
        "urn:sgws:s3::bucket1",
        "urn:sgws:s3::bucket1/*"
      ]
    },
    {
      "Sid": "Allow user1 read access to bucket bucket1",
      "Effect": "Allow",
      "Principal": {
        "AWS":
"urn:sgws:identity::34921514133002833665:user/user1"
      },
      "Action": [
        "s3:GetObject",
        "s3:ListBucket"
      ],
      "Resource": [
        "urn:sgws:s3::bucket1",
        "urn:sgws:s3::bucket1/*"
      ]
    }
  ]
}

```

Restrict a group to single subdirectory (prefix) with read-only access

This policy allows members of the group to have read-only access to a subdirectory (prefix) within a bucket. The bucket name is "study" and the subdirectory is "study01".

```

{
  "Statement": [
    {
      "Sid": "AllowUserToSeeBucketListInTheConsole",

```

```

        "Action": [
            "s3:ListAllMyBuckets"
        ],
        "Effect": "Allow",
        "Resource": [
            "arn:aws:s3:::*"
        ]
    },
    {
        "Sid": "AllowRootAndstudyListingOfBucket",
        "Action": [
            "s3:ListBucket"
        ],
        "Effect": "Allow",
        "Resource": [
            "arn:aws:s3::: study"
        ],
        "Condition": {
            "StringEquals": {
                "s3:prefix": [
                    "",
                    "study01/"
                ],
                "s3:delimiter": [
                    "/"
                ]
            }
        }
    },
    {
        "Sid": "AllowListingOfstudy01",
        "Action": [
            "s3:ListBucket"
        ],
        "Effect": "Allow",
        "Resource": [
            "arn:aws:s3:::study"
        ],
        "Condition": {
            "StringLike": {
                "s3:prefix": [
                    "study01/*"
                ]
            }
        }
    }
},

```



```
{
  {
    "Sid": "AllowAllS3ActionsInstudy01Folder",
    "Effect": "Allow",
    "Action": [
      "s3:Getobject"
    ],
    "Resource": [
      "arn:aws:s3:::study/study01/*"
    ]
  }
}
```

NetApp StorageGRID Blogs

You can find some great NetApp StorageGRID blogs here:

- May 10: [Lab on demand is your best sales tool for StorageGRID](#)
- May 24: [Modernize your analytics workloads with NetApp and Alluxio](#)
- May 26: [StorageGRID: storing and managing the on-premises backup and replication data](#)
- June 9: [Use Cloudera Hadoop S3A connector with StorageGRID](#)
- July 26: [Check out the growing list of validated partner solutions for StorageGRID](#)
- Aug 5: [NetApp StorageGRID earns Common Criteria security certification](#)
- Aug 16: [Integrating StorageGRID with the open-source ELK stack to enhance customer experience](#)
- Aug 17: [It all starts with Object Locking... Building a S3 storage ecosystem for critical backup applications](#)
- Aug 23: [Build your data lake on StorageGRID](#)
- Sep 1: [Take these Metrics and Graph it](#)
- Sep 19: [DataLock and Ransomware Protection Support for StorageGRID](#)
- Sep 26: [NetApp StorageGRID for service providers](#)
- Oct 5: [Defrost your data on StorageGRID for Snowflake](#)
- Oct 5: [NetApp Cloud Insights adds StorageGRID gallery dashboards](#)
- Nov 7: [StorageGRID and ONTAP S3 support: Differences, similarities, and integration](#)
- Nov 23: [Explainable AI with MLOps powered by NetApp and Modzy](#)
- Dec 6: [StorageGRID achieves KPMG compliance certification](#)
- Jan 16: [StorageGRID renews NF203 and ISO/IEC 25051 compliance certification](#)
- Jan 18: [StorageGRID S3 Object Lock validated for Veritas NetBackup](#)
- Feb 14: [What do chocolate, skiing, watches, and mainframes have in common?](#)
- Mar 14: [How to back up Epic Systems EHR databases with one command in a 3:2:1-compliant architecture](#)
- Mar 30: [Use BlueXP to protect Epic EHR with a 3:2:1 -compliant backup policy](#)
- Mar 30: [Mountpoint for Amazon S3 alpha release with StorageGRID](#)
- May 16: [What's new in the StorageGRID object storage family](#)
- May 16: [Introducing StorageGRID 11.7 and the new all-flash object storage appliance SGF6112](#)
- Aug 30: [Mountpoint for Amazon S3 File System is Now GA](#)
- Sep 1: [Leveraging Cloud Insights to Monitor and Collect Logs Using Fluent Bit](#)
- Oct 17: [Moving on from Hadoop: Modernizing Data Analytics with Dremio and StorageGRID](#)
- Nov 7: [Spectra Logic On-Prem Glacier with StorageGRID](#)
- Dec 12: [Big data analytics on StorageGRID: Dremio performs 23 times faster than Apache Hive](#)
- Feb 2: [Announcing the StorageGRID + lakeFS Solution Brief](#)
- Feb 16: [Introducing StorageGRID 11.8: Enhanced security, simplicity, and user experience](#)
- Feb 16: [Introducing StorageGRID 11.8](#)

NetApp StorageGRID documentation

You can find the complete documentation for each NetApp StorageGRID release here:

- [StorageGRID appliances](#)
- [StorageGRID 11.8](#)
- [StorageGRID 11.7](#)
- [StorageGRID 11.6](#)
- [StorageGRID 11.5](#)
- [StorageGRID 11.4](#)
- [StorageGRID 11.3](#)
- [StorageGRID 11.2](#)

Legal notices

Legal notices provide access to copyright statements, trademarks, patents, and more.

Copyright

<https://www.netapp.com/company/legal/copyright/>

Trademarks

NETAPP, the NETAPP logo, and the marks listed on the NetApp Trademarks page are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.

<https://www.netapp.com/company/legal/trademarks/>

Patents

A current list of NetApp owned patents can be found at:

<https://www.netapp.com/pdf.html?item=/media/11887-patentspage.pdf>

Privacy policy

<https://www.netapp.com/company/legal/privacy-policy/>

Open source

Notice files provide information about third-party copyright and licenses used in NetApp software.

https://library.netapp.com/ecm/ecm_download_file/2879263

https://library.netapp.com/ecm/ecm_download_file/2881511

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.