Technical Report

NetApp StorageGRID with Rubrik CDM configuration guide

Implementing NetApp StorageGRID as an archival location for Rubrik backups

Jonathan Wong and Steve Pruchniewski, NetApp
Ben Kendall, Rubrik
September 2021 | TR-4812

In partnership with

Abstract

This report describes the configuration, best practices, and guidelines for NetApp® StorageGRID® as an archive endpoint for Rubrik.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Scope

Rubrik Cloud Data Management (CDM) is a backup and recovery solution for enterprise workloads. As customers set out to modernize their data management and data protection practices, Rubrik's CDM platform is an increasingly attractive solution. For longer-term backup retention, data can be archived from Rubrik to the public cloud or to object-based storage such as NetApp® StorageGRID®.

This document is a reference for configuring NetApp StorageGRID as an archival target. It also covers configuring a backup policy in Rubrik, configuring a S3 tenant in StorageGRID, and an information lifecycle management (ILM) policy in StorageGRID, as well as sizing and performance guidelines for StorageGRID with Rubrik.

The recommendations in this document are suggested guidelines. When designing a solution, it is important to consider as many aspects of the backup environment as possible.

This document is intended for NetApp employees and partners who are familiar with StorageGRID concepts and terminology, and who also have technical familiarity with backup solutions.

Rubrik Cloud Data Management

Rubrik's CDM platform organizes and protects data on the premises, at the edge, and in the cloud. Users can easily automate backup jobs by configuring SLAs and archiving data to public or private clouds, including NetApp StorageGRID object-based storage.

Rubrik scales linearly, so customers are not hampered by forklift upgrades. Deduplication, compression, and other data services scale in line with the cluster to maximize efficiency and savings. Furthermore, Rubrik simplifies data recovery by allowing "Google-like" search functionality to recover only the database, applications, or files needed.

In addition, Rubrik is built on an API-first architecture that is fully featured and easy to learn, allowing customers to automate their data workflows.

NetApp StorageGRID

NetApp StorageGRID is a software-defined object storage solution that supports industry-standard object APIs such as Amazon S3 API and OpenStack Swift API. StorageGRID uses intelligent, policy-driven data management to store, protect, and preserve data, enabling you to create metadata-driven object lifecycle policies to optimize durability, performance, cost, and location across multiple geographies.

In addition, NetApp StorageGRID is built as a scale-out, node-based architecture that gives you the agility to increase capacity and performance on demand across your sites. Scaling is asymmetrical and as simple as adding more storage nodes, which are available as an appliance, software only/bare metal, or virtual machine.

Reasons to choose StorageGRID

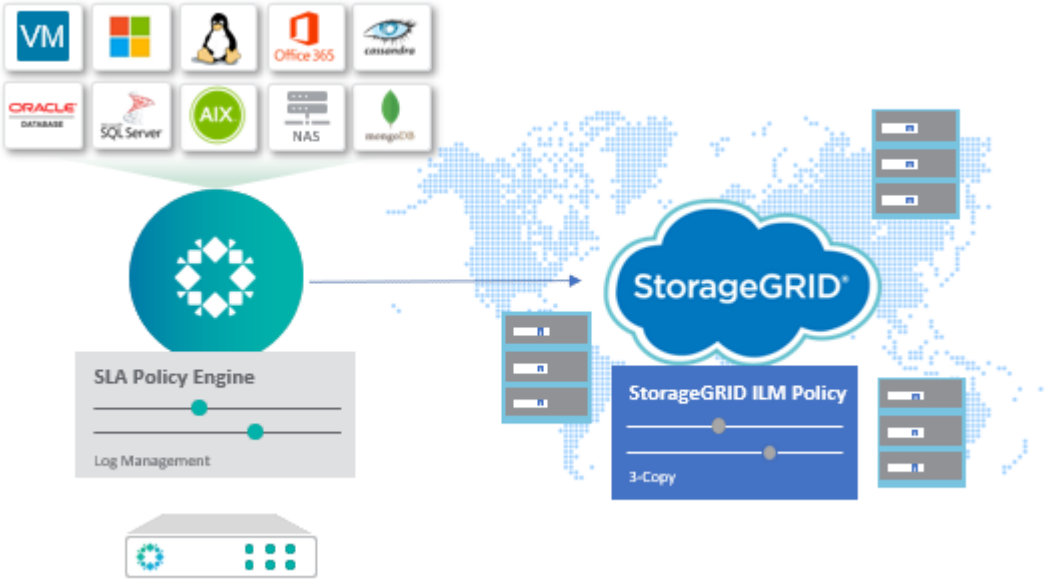
By reducing complexity and providing flexible, policy-driven data management, Rubrik delivers simplicity and efficiency. At the same time, given the popularity and ever-expanding use case for object storage, customers are choosing NetApp StorageGRID for similar reasons: policy-driven data management, simplicity, and flexibility.

NetApp and Rubrik are delivering newly integrated solutions that offer policy-based simplicity, cost-effective scale, and cloud mobility.

StorageGRID makes a superior archive endpoint in the following scenarios:

- You operate in a hybrid cloud model to optimize costs and avoid vendor lock-in. You can deploy locally or globally with no proprietary hardware lock in.
- You want a system that can grow organically. Refresh, expand, and migrate non-disruptively without the dependencies that other systems have.
- You require granular data protection to comply with data sovereignty and other regulatory compliance requirements. Dynamically balance data durability, performance, cost, and location with the industry's leading lifecycle policy engine.
- You want to archive data and leverage object storage for other use cases. Easily stand up other workflows with our multi-tenant capabilities and hybrid cloud capabilities.

Figure 1) NetApp StorageGRID with Rubrik as archival location.



Requirements

Although StorageGRID and Rubrik have been implemented in the field as far back as StorageGRID 10.4 and Rubrik 2.3, this document validates the current General Availability (GA) versions:

- StorageGRID 11.5.0
- Rubrik CDM 6.0.0-p1-12566

As prerequisites, the following must be configured on StorageGRID:

- FQDN for S3 endpoint; for example, s3.company.com
- API service endpoint server certificates
- Commercial trusted SSL is preferred
- Self-signed is acceptable but requires Rubrik Support to enable
- S3 tenant created with S3 keys (Swift API is not supported for this solution.)
- Load balancer endpoint
- Disable grid encryption and compression (default setting). Rubrik already sends data encrypted and compressed.

Rubrik

Rubrik can be deployed on a hardware appliance, software at the edge, software in the cloud, and on third-party industry platforms.

For this guide, a minimum configuration was deployed on VMware:

- One Rubrik Edge

StorageGRID

StorageGRID can be deployed on a hardware appliance, software only/bare-metal servers, or a hypervisor. For this test, a minimum configuration was deployed on a VMware hypervisor:

- One admin node
- Three storage nodes

HTTP and HTTPS support

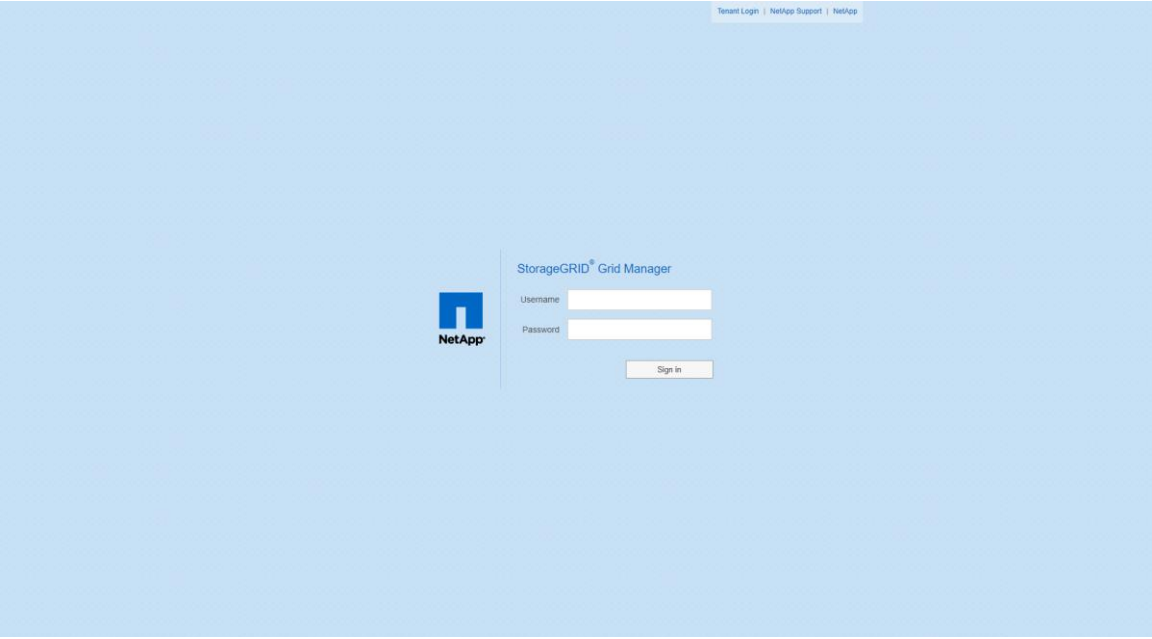
StorageGRID supports both HTTP and HTTPS. HTTPS is enabled by default. HTTP is recommended only for lab environments and is disabled by default.

Customers can choose to use standard ports with a third-party load balancer or create a load-balancer endpoint (11.3 and later). For details, see Configuring Load Balancer Endpoints in the StorageGRID 11.5 Documentation Center.

Configuring StorageGRID with Rubrik

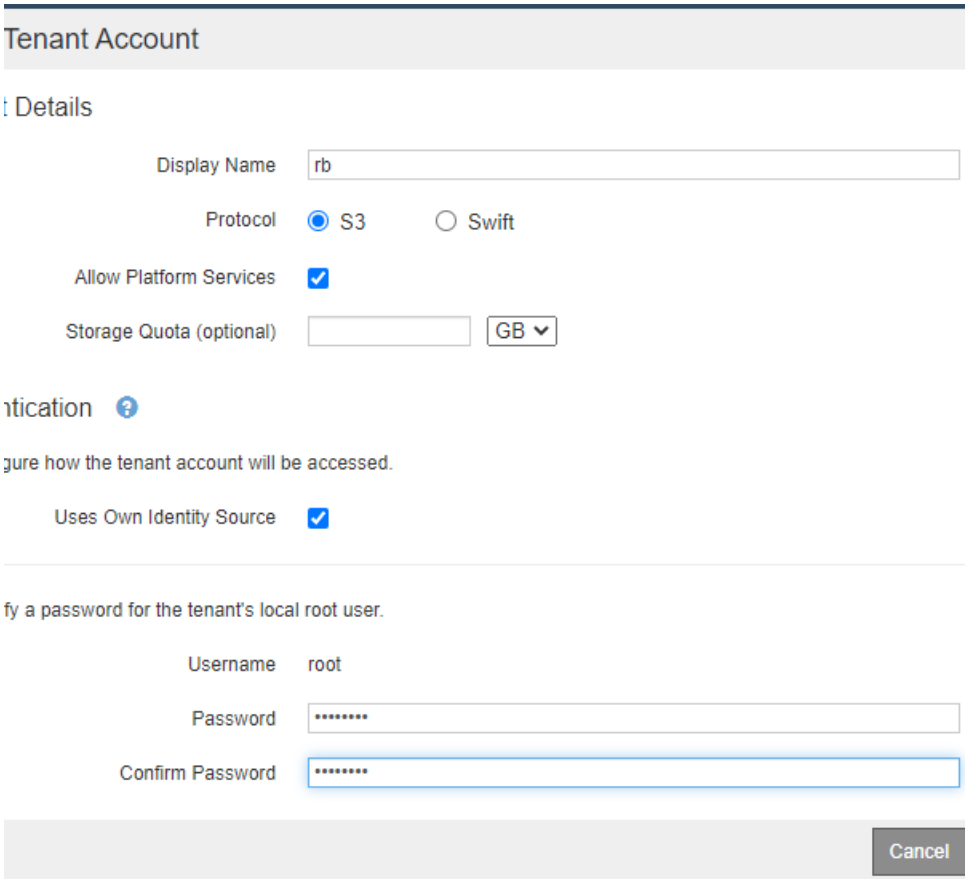
Configuring an S3 tenant on StorageGRID

1. Log in to the StorageGRID Grid Manager Interface.

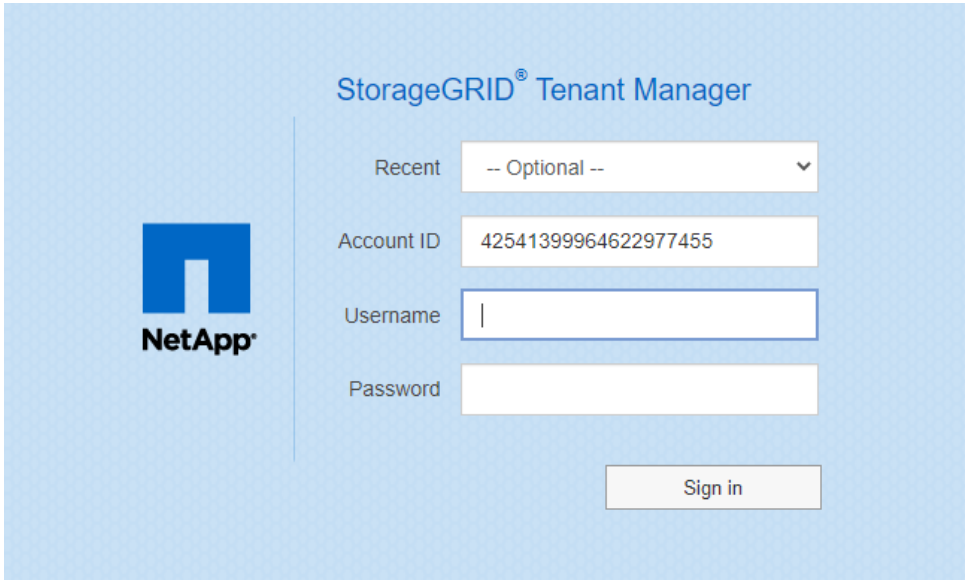


2. Click Tenants on the top menu bar and select + Create.
3. Create a tenant account:

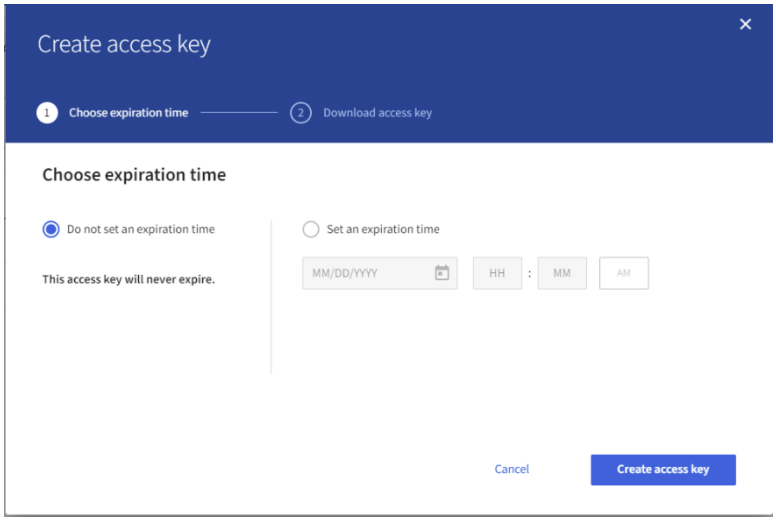
a. Enter a display name.
b. Select S3 as the protocol.
c. Enter a password for the tenant's local root user.



4. After the tenant account is created, select Sign In next to the tenant to access the Tenant Manager Interface (TMI) and log in.



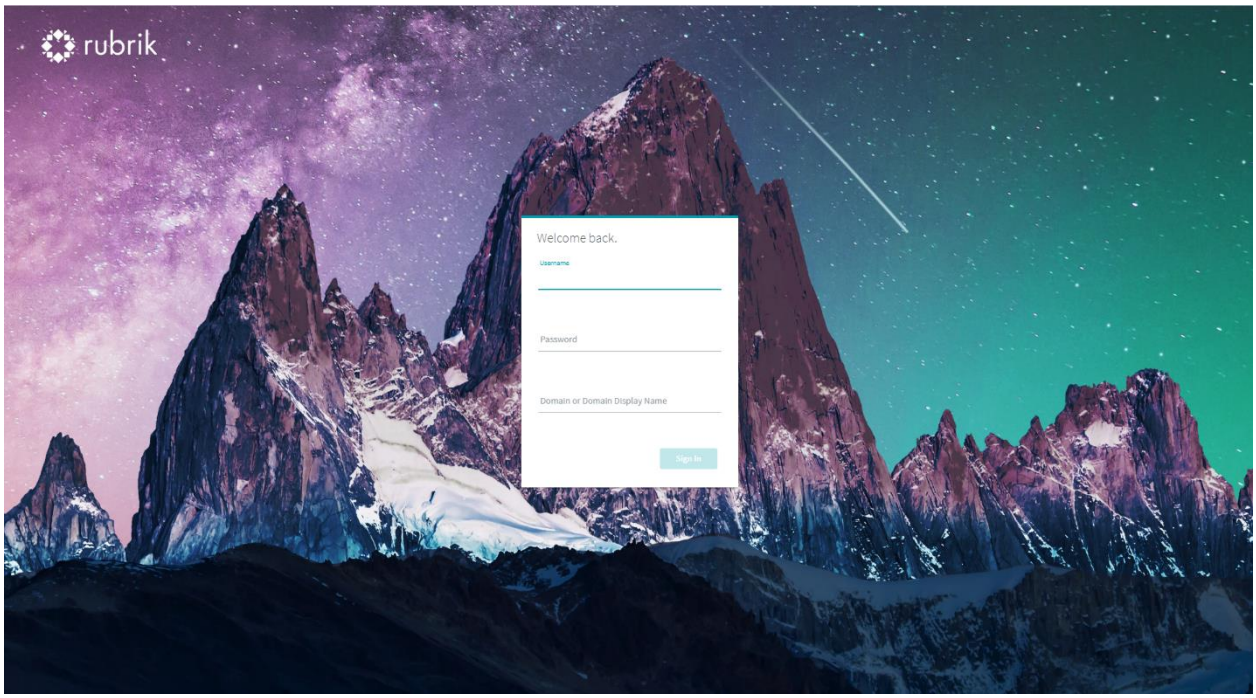
5. In the TMI, go to S3 → My Access Keys, and click Create Key.
6. Step through the process and be sure to download your Access Key and Secret Access Key.



The StorageGRID S3 Access and Secret Access Keys are now ready to be used with Rubrik.

Configuring StorageGRID as an archival location

1. Log in to the Rubrik Cluster using an account with Admin permission.



2. Click the cog symbol at the top right in the Rubrik GUI. Under System Configuration, select Archival Locations.
3. Click the + to create a new archival location and then follow these configuration steps:
   a. Set Archival Type as Object Store and set Object Store Vendor as S3 Compatible (StorageGRID).

b. Enter an S3 Access Key and Secret Access Key.

c. Set Host Name to the FQDN of the S3 endpoint. If not using standard HTTPS port 443, enter the port of the load balancer endpoint:

d. Enter a bucket prefix, for example, sg.

e. Rubrik creates multiple buckets named `prefix-rubrik-x`.

f. Select the number of buckets to be used.

g. Rubrik stores all data for a source (VM, database, etc.) in a single bucket along with the metadata that allows Rubrik to validate and perform recovery. NetApp recommends creating no more than one bucket per Rubrik Archive.

h. Generate an RSA key for encryption.

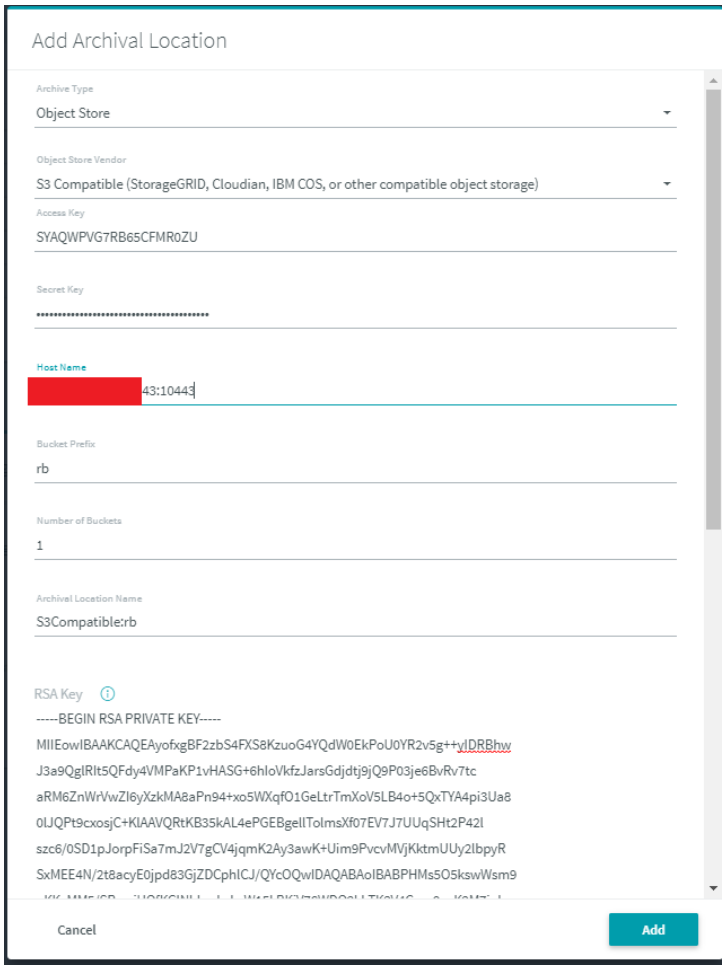
- Run the following command on a secure computer that has the OpenSSL toolkit:

```
openssl genrsa -out rubrik_encryption_key.pem 2048
```

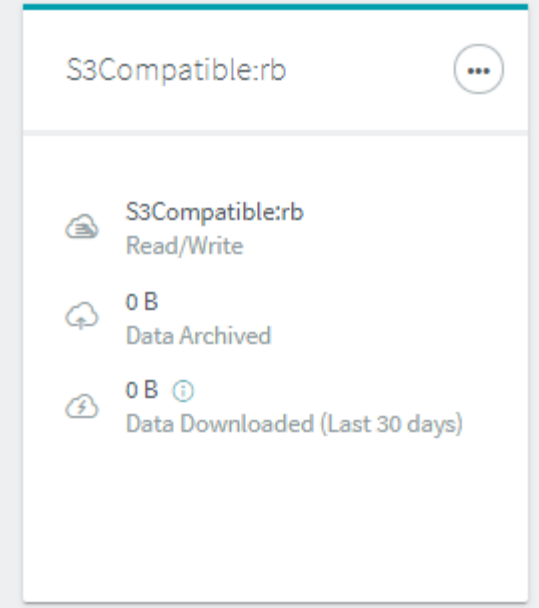
- Paste the RSA key into the window.

Rubrik uses this key to encrypt the archival data.

i. Click Add.



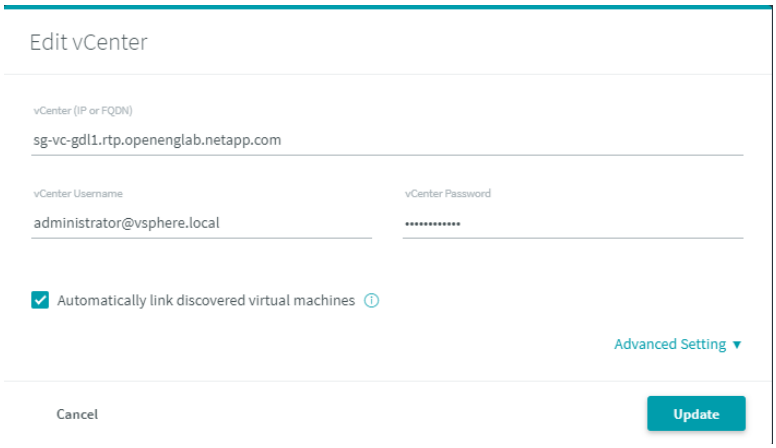
4. The StorageGRID S3 object store is added to the Rubrik Cluster as an archival location and is available to be used by Rubrik SLAs.



Configuring VM backup and restore

Rubrik supports backup for different enterprise workloads such as databases and virtual machines. Rubrik also supports different hypervisors – VMware vSphere (ESXi), Microsoft Hyper-V, and Nutanix AHV. For this test, we connected Rubrik to a vCenter Server for VM backup.

1. Click the cog symbol at the top right in the Rubrik GUI to add the vCenter Servers.
2. In the Rubrik GUI click, the + to add vCenter and then follow these configuration steps:
   a. Enter a vCenter IP or FQDN.
   b. Enter a vCenter username and password.



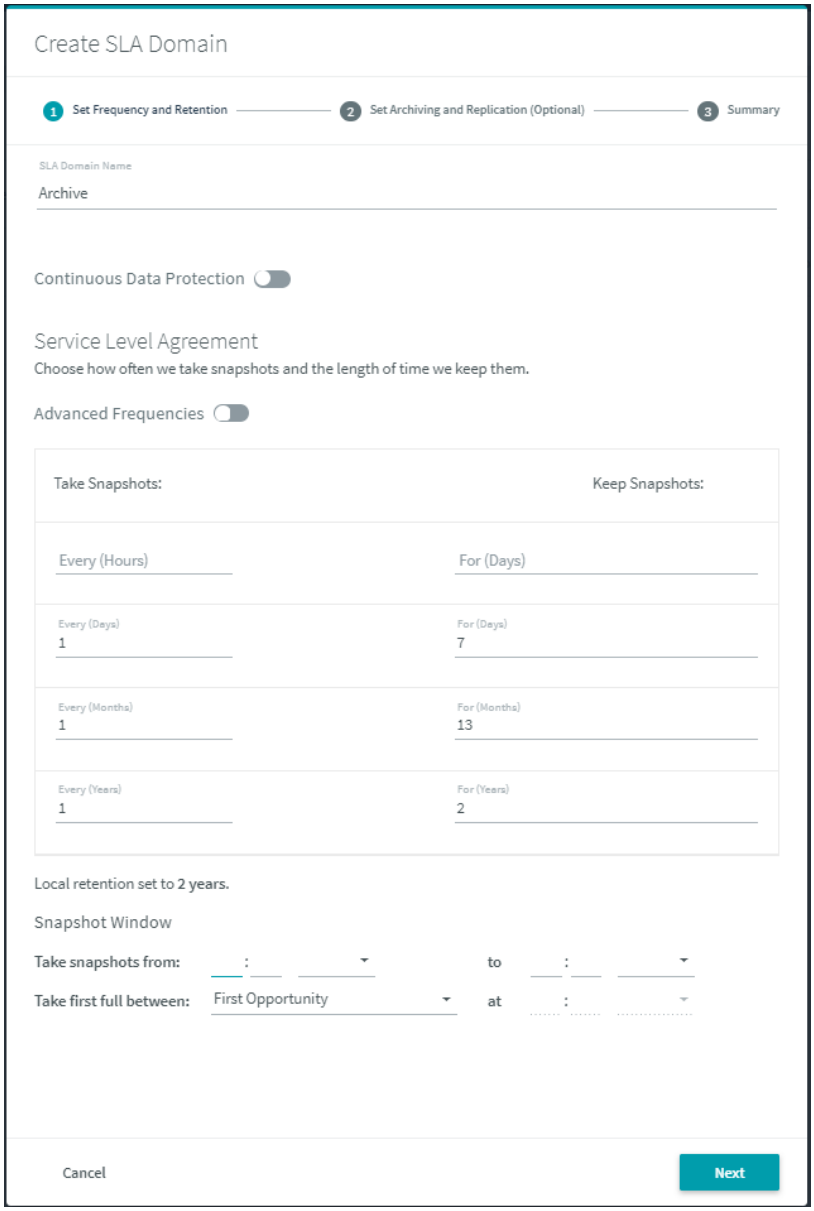
3. When the vCenter has been added, Rubrik makes an inventory of VMs, hosts, and folders.

Creating an SLA

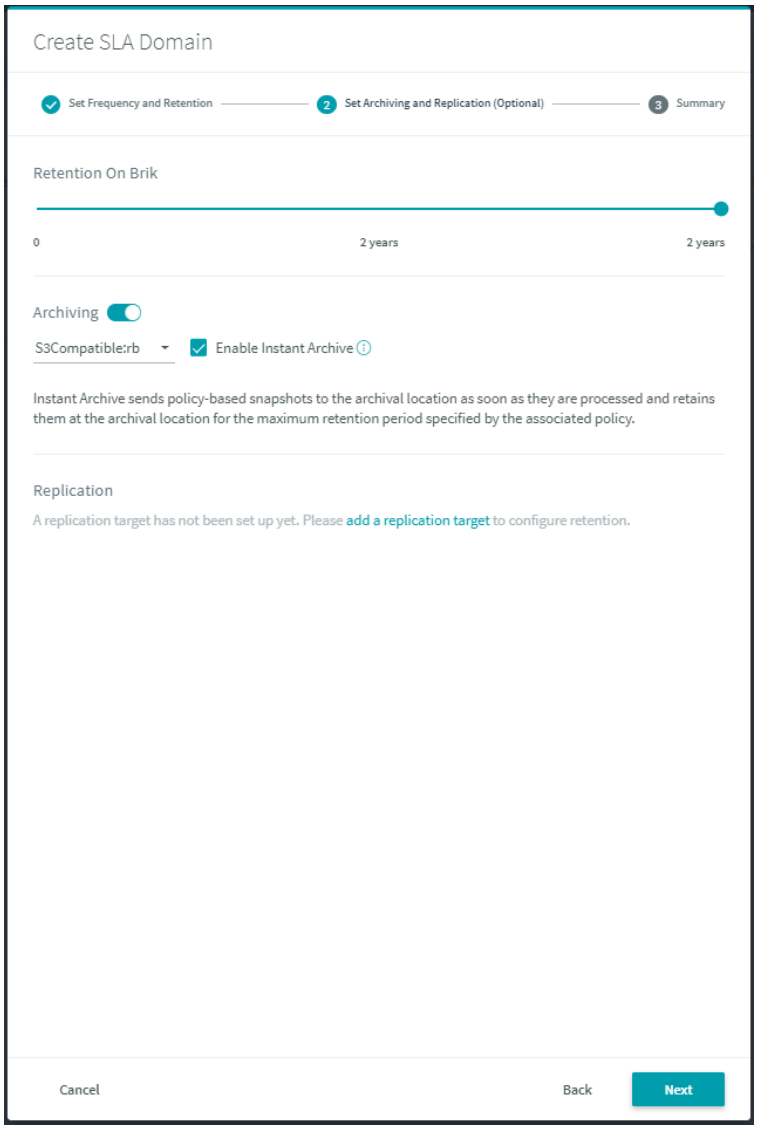
Rubrik SLAs define protection levels for workloads composed of snapshot protection and retention, replication, and archiving. For this test, we configured an SLA targeted at archiving to StorageGRID.

1. From the left-side menu, select SLA Domains > Local Domains.
2. In the Rubrik GUI, click the + to create a new SLA domain and then follow these configuration steps:

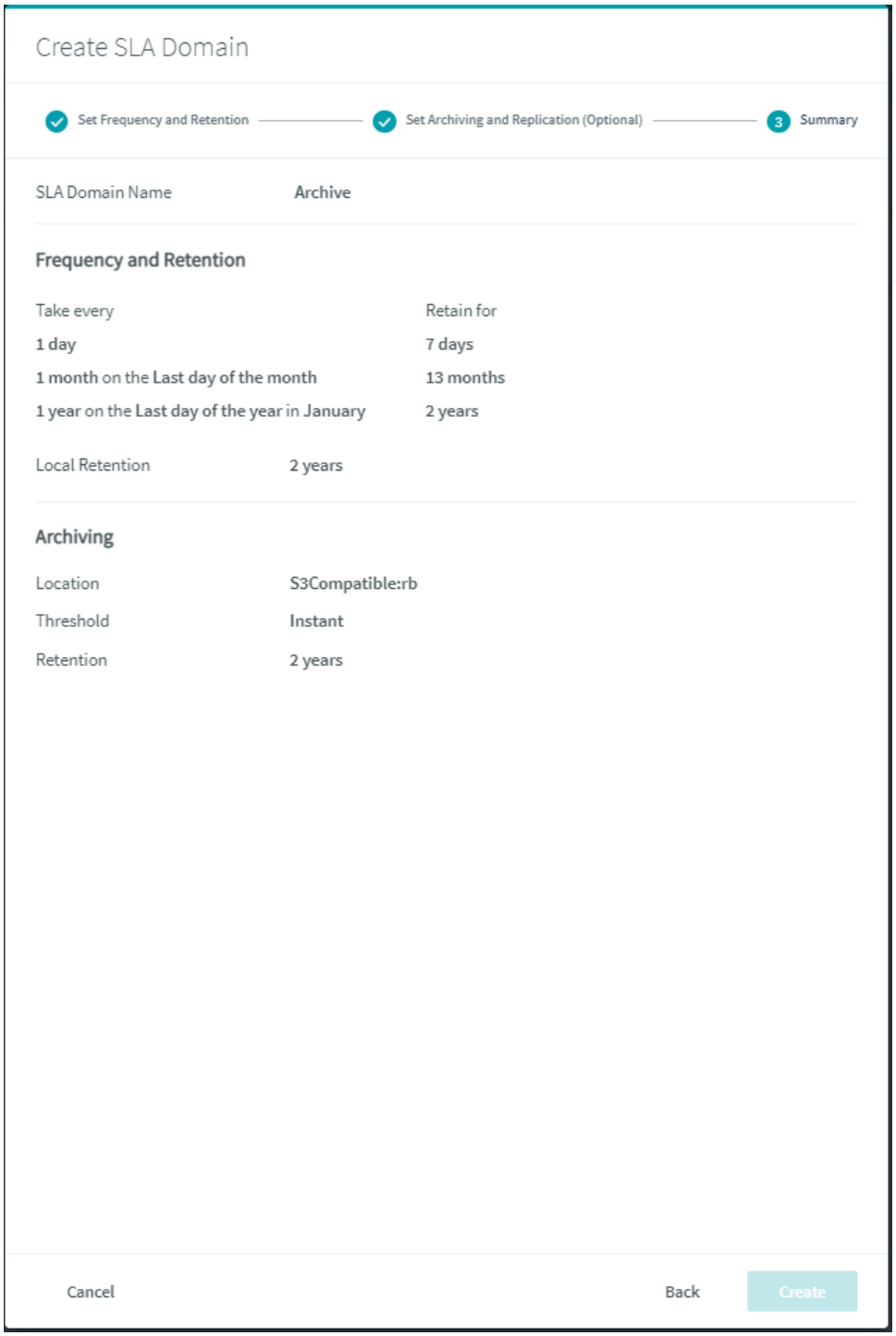
a. Enter an SLA domain name.
b. Specify snapshot protection and retention.



3. Click Next.
4. Enable Archiving and select the configured StorageGRID archival target.

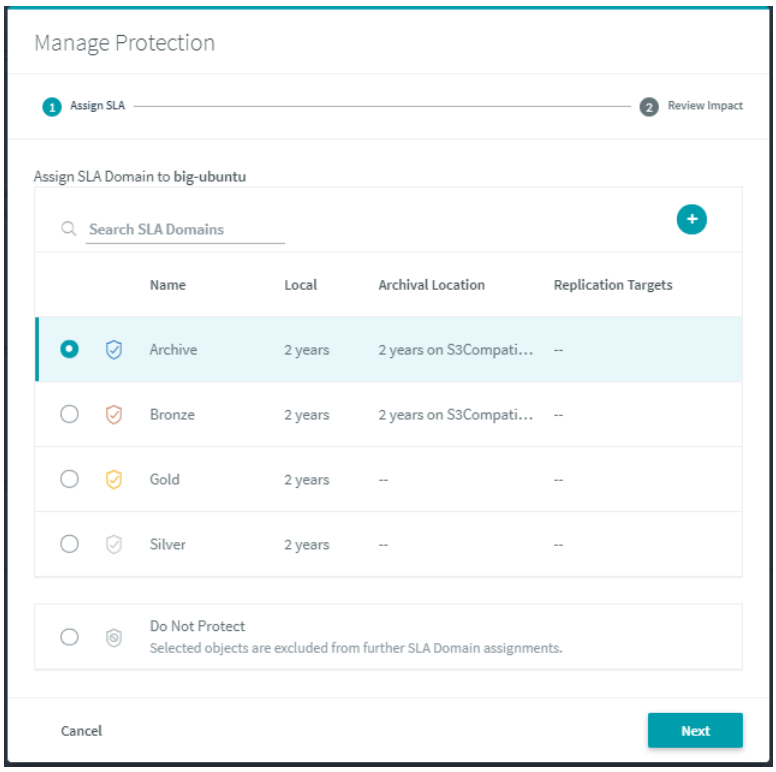


5. Verify the settings for SLA frequency, retention, and archiving.



Selecting a VM to archive to StorageGRID

1. From the left-side menu, select Virtual Machines → vSphere VMs.
2. In the Rubrik UI, select a VM and click Manage Protection.
3. Select an SLA domain and click Next.



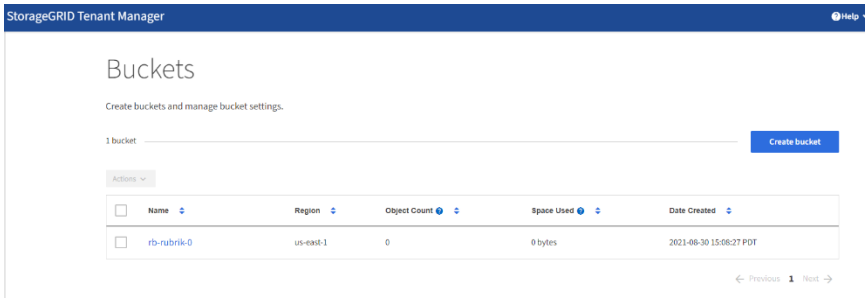
4. Review the protection settings and click Submit

Creating an ILM rule to protect the Rubrik archive data

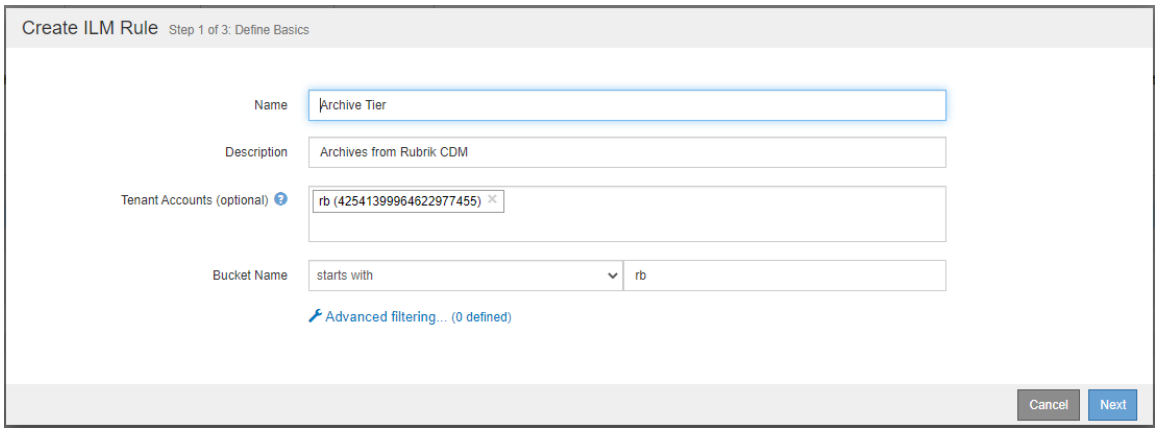
In StorageGRID, an ILM rule contains the instructions for placing objects in the system over time. Rules perform actions based on matching criteria such as bucket name or user-specified metadata fields.

In this test, we replicated our archival data to three sites and retained the replications for a year before transitioning to erasure coding.

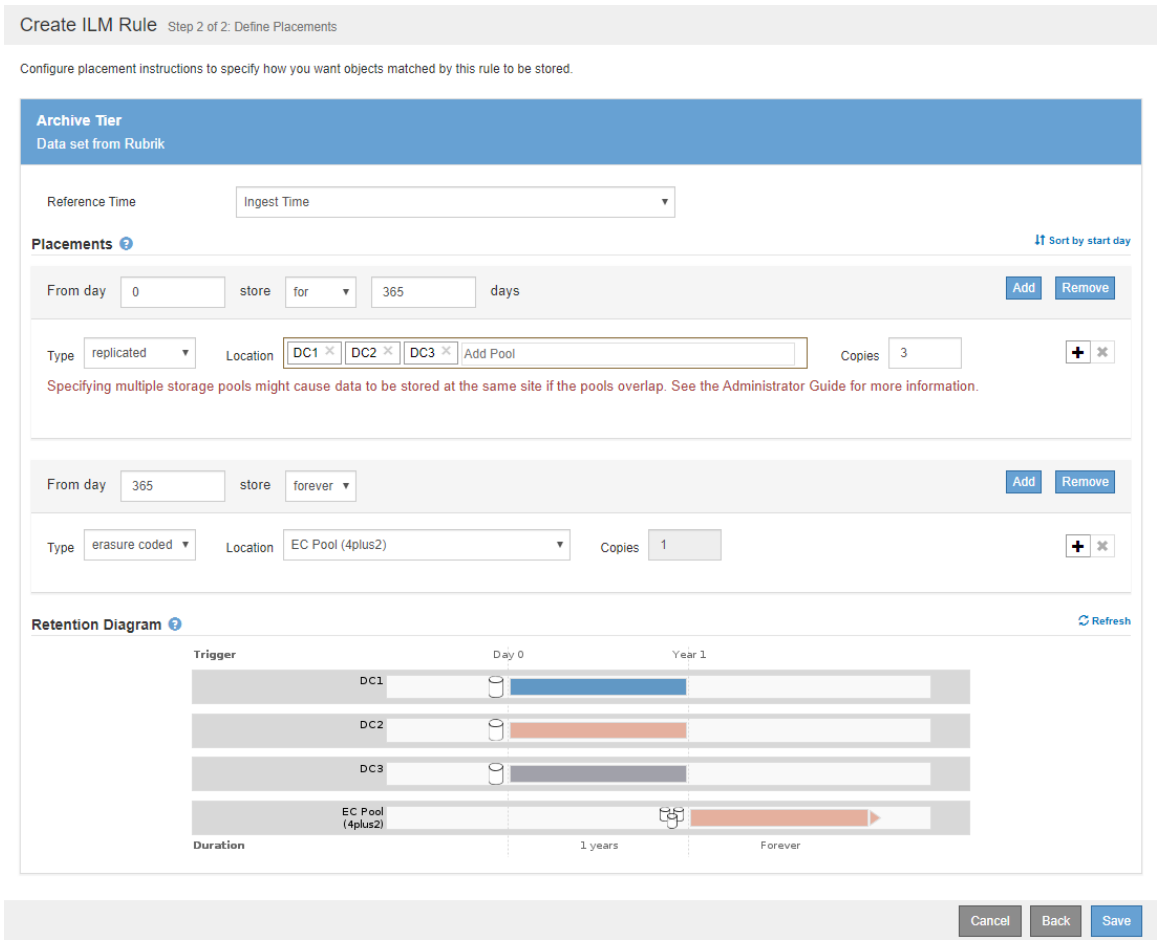
1. Log in to the StorageGRID Grid Management Interface (GMI).
2. When StorageGRID is configured as a Rubrik archival location, buckets are automatically created based on the Rubrik configurations set in section 3.2. To see the buckets, follow these steps.
   a. In the StorageGRID GMI, go to Tenants → Sign In (next to the Rubrik Tenant) to access the Tenant Manager.
   b. Enter the Rubrik tenant root account and password.
   c. Go to S3 → Buckets.



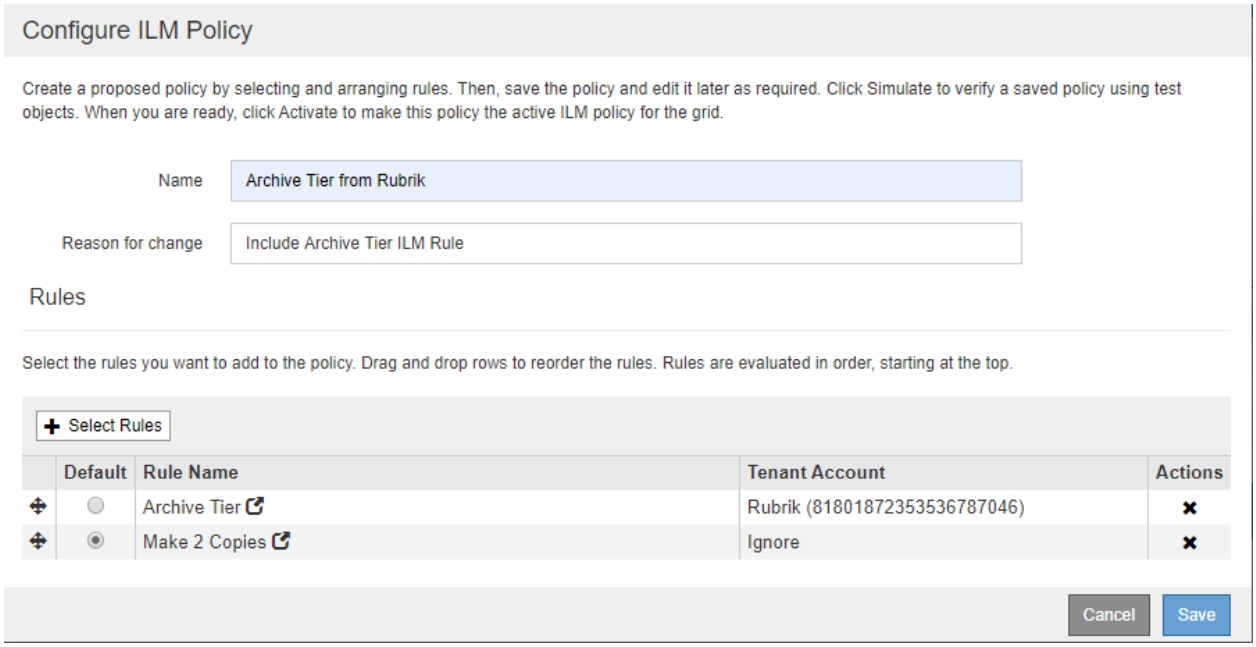
3. Create an ILM rule to manage the data. In the StorageGRID GMI, go to ILM → ILM Rules → + Create.
   a. Enter a name.
   b. Enter a description.
   c. Specify criteria that the Tenant Account must be the configured Rubrik tenant.
   d. Specify criteria that the Bucket Name must start with `prefix-rubrik`. (Prefix configured in section 3.2.)
   e. Click Next.



4. Example configuration of an ILM placement rule:
   a. Replicate three copies across three data centers for geo distribution.
   b. Transition to erasure coding after a year to reduce storage space.
   c. Verify that the retention diagram is correct.
   d. Click Save.



5. Go to ILM → Policies → + Create Proposed Policy:
   a. Enter a name.
   b. Enter a reason for the change.
   c. Select the rules to apply.
   d. Click Apply.
   e. Select a default rule.
   f. Click Save.

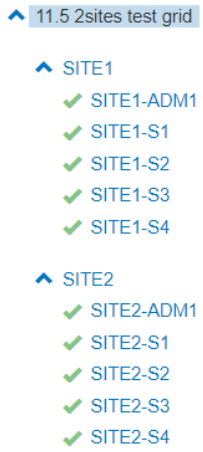


6. Activate the policy.

Multisite configuration (optional)

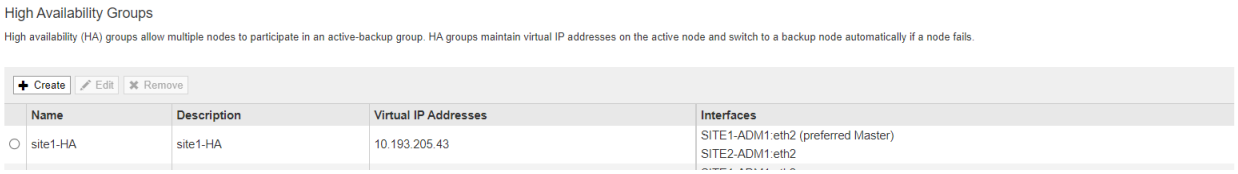
It is a common deployment architecture to have a primary data center and a secondary data center both deploying Rubrik and StorageGRID. The secondary data center ensures backups can still be archived and restored during a full primary data center failover. To create a basic configuration for a two-site deployment, follow these steps:

1. Deploy a two-site StorageGRID grid with a primary Admin Node in the primary data center and a non-primary Admin Node in the secondary data center.

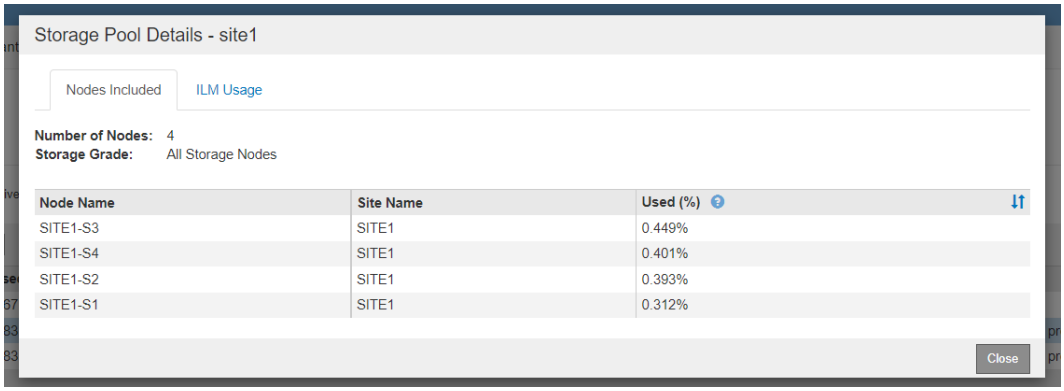


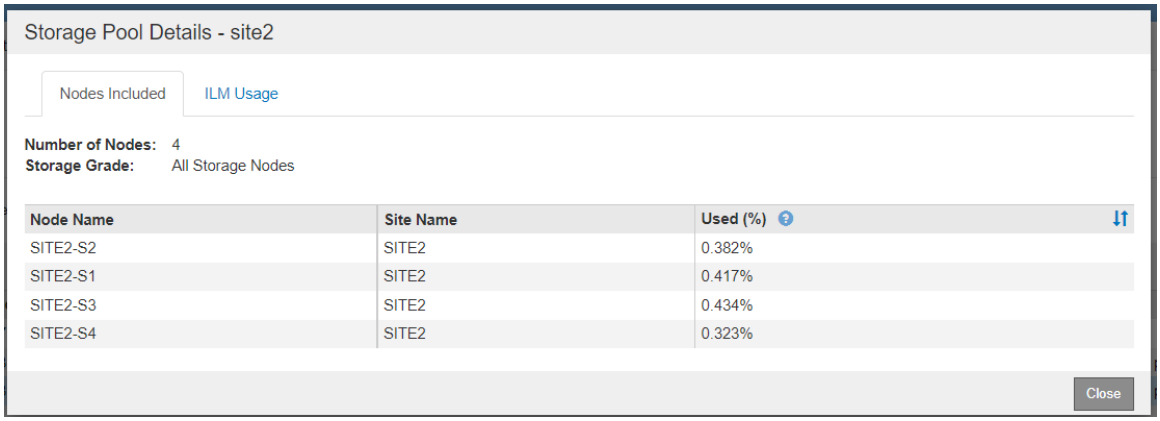
The Admin Nodes act as the S3 endpoint to provide load balancing to the Storage Nodes. S3 endpoints can also be provided by dedicated Gateway Nodes.

2. To enable connectivity to StorageGRID during data center failures, configure a high availability group. This ties the two S3 endpoints into an active-backup group with a virtual IP (VIP).

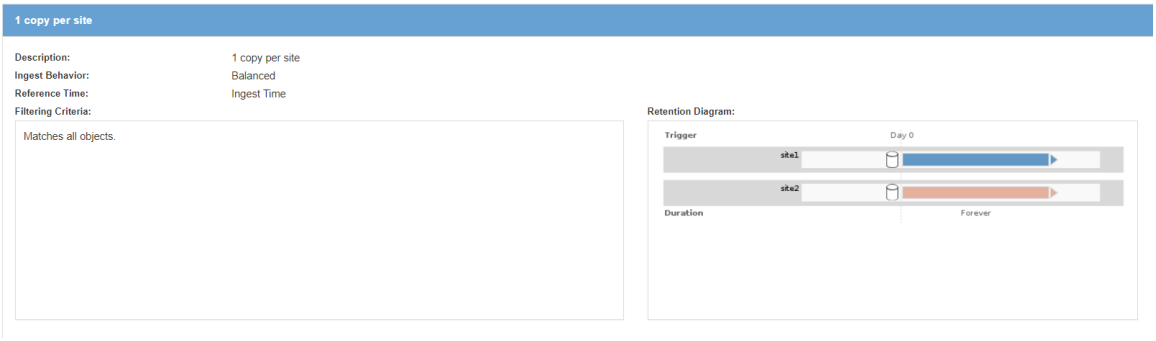


3. The VIP is the IP address that Rubrik connects to when configuring the archival location. During a primary data center loss, the VIP reroutes to the S3 endpoint in the secondary data center transparently.
4. Make sure that the backups archived to StorageGRID are replicated into both primary and secondary data centers. That way, during a data center failure, the object data is still available.
5. Create two storage pools; one for the primary data center and another for the secondary. Storage pools are logical groupings of nodes used to define object placement.
   a. Go to ILM → Storage Pools → + Create
   b. Follow the steps and create a storage pool for each data center.





6. Create an ILM rule specifying object data to be stored in both sides. In this example, a one copy in each site rule was created.
   a. Set Ingest Behavior to Balanced (default). This makes sure that when a site is lost, object data can still be written to StorageGRID.



7. The above configuration example results in the following:
   a. A virtual S3 endpoint IP that during primary data center loss reroutes traffic to the secondary data center automatically.
   b. When backups are archived to StorageGRID, one copy is made per data center. If the primary data center fails, a replicated copy still exists in the secondary data center ready for retrieval.
   c. Both are transparent to Rubrik. The archival location continues to use the same IP and bucket name.

Configuration recommendations

Rubrik already encrypts and compresses its data. Therefore, NetApp recommends disabling encryption and compression on StorageGRID.

Rubrik is not versioning aware, so versioning on StorageGRID buckets should be disabled.

Rubrik has not been validated with Cloud Storage Pools. If customers want to use capacity from the public cloud, you can configure Rubrik to support additional cloud storage targets.

StorageGRID performs better with larger objects, and you can increase part size to 128MB or more to increase throughput.

Rubrik, as of the current release tested in this document, performs nonexistent HEADs as part of the connectivity check. NetApp recommends changing the bucket consistency of the StorageGRID buckets to

`available` to enable greater availability during node or site loss. For more information, read about the StorageGRID consistency controls. For information about how to change the consistency level, see the instructions in the StorageGRID documentation center.

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 documentation
  http://docs.netapp.com/sgws-115/index.jsp
  https://www.netapp.com/us/media/ds-3613.pdf
- TR-6773 StorageGRID Performance 11.5
  https://fieldportal.netapp.com/content/205463
- StorageGRID Solution Designer
  https://fieldportal.netapp.com/content/593648
- Rubrik resources
  https://www.rubrik.com/resources/
- Rubrik support articles (requires login)
  https://support.rubrik.com/s/article/000003591
  https://support.rubrik.com/s/article/000003058
  https://support.rubrik.com/s/article/000002692
- Rubrik SDK
  https://build.rubrik.com/
  https://github.com/rubrikinc

Version history

Version	Date	Document version history
Version 1.0	December 2019	Initial release.
Version 1.1	May 2020	Updated sizing and performance section
Version 1.2	September 2021	• Updated to the latest template. • Updated screenshots to latest Rubrik and StorageGRID version • Included information on multisite deployments Updated “Configuration recommendations” section

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

TR-4812-0921

NetApp