



TECHNICAL WHITE PAPER

Using StorageGrid as a Rubrik Archival Location

Steve Wharton
May 2020
RWP-0520

TABLE OF CONTENTS

3 INTRODUCTION

- 3 Audience
- 3 Objectives
- 3 The Rubrik Approach

4 ARCHITECTURE AND COMPONENTS

- 4 High Level Architecture
 - 4 StorageGRID
 - 4 Rubrik

5 HOW IT WORKS

- 5 Adding StorageGrid as an archive
- 8 Backup (workflow)
- 9 Recovery (workflow)

9 BEST PRACTICES

- 9 StorageGRID compression and encryption
- 10 StorageGRID certificates
- 10 StorageGRID object versioning
- 10 StorageGRID Cloud Storage Pools
- 10 Rubrik Archive Reader/Writer Connection and Rubrik Replication
- 10 Useful hotfixes for StorageGRID 11.2

10 WHAT'S NOT SUPPORTED

11 CONCLUSION

12 SOURCES AND NOTES

- 12 Further Resources

12 VERSION HISTORY

INTRODUCTION

This paper describes how NetApp StorageGRID can be leveraged with Rubrik CDM as a long term archive for Rubrik backups.

AUDIENCE

This document is intended for Sales Engineers, Field Consultants, Professional Services, IT Managers, Partner Engineering and customers who want to understand how to configure Rubrik and StorageGRID together.

OBJECTIVES

The document will walk through the configuration steps to be taken on Rubrik and StorageGRID and include any caveats or considerations for deployment.

THE RUBRIK APPROACH

As customers set out to modernize their data management and data protection practices Rubrik's Cloud Data Management (CDM) platform has increasingly become an attractive solution.

Rubrik simplifies backup and recovery for virtualized, physical and NAS environments. It removes the complexity of installing and managing backup software running across siloed hardware components by integrating data orchestration, catalog management, and deduplicated storage into a single, scale-out fabric. It offers simple data management of near term data held upon the Rubrik platform and also longer term backup data that has been archived off the Rubrik Platform to Public Cloud or Object Storage such as NetApp StorageGRID.

Rubrik scales linearly so that customers are not hampered by forklift upgrades. Deduplication, compression, and other data services scale in-line with the cluster to maximize efficiency and savings. Instant recovery is delivered without rehydration or additional storage provisioning. With Rubrik, lengthy backup setup times, broken job scheduling, and uncertain recoveries are relics of the past.

The Rubrik Cloud Data Management Platform is available as a plug and play appliance, as software on approved third-party partner platforms or as virtual appliances running in ROBO or on Public Cloud.

By reducing complexity and providing flexible, policy-driven data management Rubrik has delivered simplicity and efficiency. At the same time, with the popularity and the ever expanding use cases for object storage, customers have chosen NetApp's StorageGRID object storage for similar reasons: policy-driven data management, simplicity and flexibility.

It's no surprise that customers have been implementing Rubrik with StorageGRID together in production environments for several years. The combined benefits of Rubrik and StorageGRID provide ease of use with enterprise-grade data protection that can be implemented from a single datacenter to global multi-site infrastructures.

ARCHITECTURE AND COMPONENTS

NetApp StorageGRID is a software-defined object storage solution that supports industry-standard object APIs such as the Amazon Simple Storage Service (Amazon S3) API. Users can create multiple service levels with metadata-driven object lifecycle policies, optimizing durability, protection, performance, cost, and location across multiple geographies.

As requirements change, users can adjust policies and realign their data landscape. Enterprises have the flexibility of deploying StorageGRID as an enterprise-grade turnkey solution or as a software-only solution on physical or virtual servers, using heterogeneous storage underneath. With nearly two decades of enterprise deployments, customers can depend on StorageGRID for data protection and integrity.

With Rubrik and NetApp StorageGRID, you can now automate data lifecycle management through Rubrik's simple control plane while using StorageGRID as a cloud-scale object-based archive target. Both Rubrik and StorageGRID index file metadata to enable global file-level searching for instant access to massive amounts of unstructured data. Users can quickly find and restore virtual machines, databases, files, and more, regardless of location—in Rubrik, in StorageGRID, and even in a public cloud.

HIGH LEVEL ARCHITECTURE

STORAGEGRID

StorageGRID has multiple deployment options:

- **Physical Appliances**
Compute, networking, and storage in a single chassis—a building block for enterprise-grade object storage that's easy to deploy.
- **VMware Based**
Deploy StorageGRID on virtual machines (VMs) on VMware.
- **Docker Based**
Deploy on RHEL/CentOS servers in Docker Containers.

In each scenario the architecture contains 3 types of components:

- 1 or more Admin nodes (Primary Admin node plus a secondary if required)
- 1 or more Gateway/API/Load Balancer nodes
- 3 or more Storage Nodes

The Admin node is required for configuration and administration of the StorageGRID. It provides services for the web GUI/CLI, system configuration and audit logs. Each GRID requires at least one Admin node. However, once a GRID has been configured, the Admin node is not required for the StorageGRID's continued operation.

StorageGRID includes a basic load balancer called the API Gateway Node. In cases where customers require advanced features and customization, a third-party load balancer, either commercial or open source can be incorporated into the StorageGRID architecture.

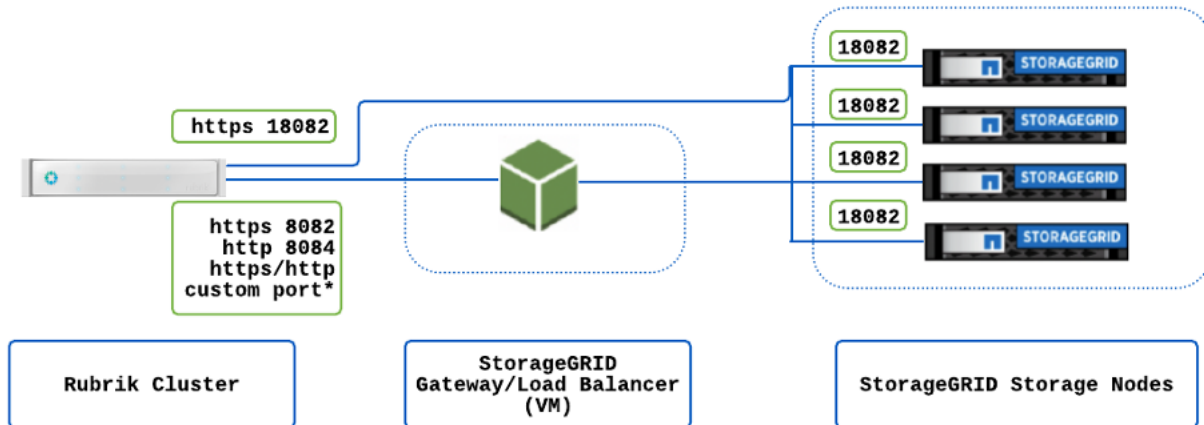
A Storage Node includes the services and processes required to store, move, verify, and retrieve object data and metadata on disk. StorageGRID protects objects using replication and erasure coding.

RUBRIK

Rubrik connects to the StorageGRID cluster using the S3 protocol. The connection is typically made over https (although http is possible if it had been enabled on StorageGRID). If the connection is made via the StorageGRID gateway/load-balancer VM, the connections take place on ports 8082 for https and 8084 for http. The connection can also be made directly to the Storage nodes using https on port 18082.

Should a customer wish to use network ports other than the default one, they can configure `PORT_REMAP`¹ or in version 11.3 of StorageGRID configure Load-Balancer endpoints². As mentioned above, it is also possible to configure third party load-balancers though this will not be covered in this document.

Connecting Rubrik to StorageGRID using S3



HOW IT WORKS

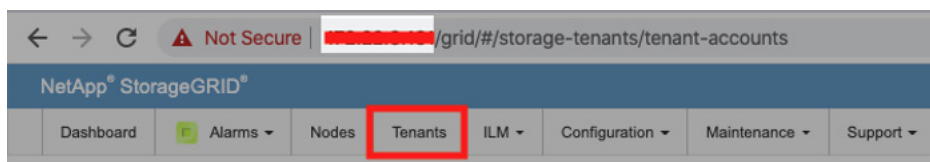
Before configuring StorageGRID as a Rubrik Archive it's necessary to have DNS entries (A record and PTR) configured for the S3 endpoints. The Fully Qualified Domain Name will point to either

- StorageGRID Gateway(s).
- All the Storage Nodes (round-robin).
- Load balancer endpoints defined on the StorageGRID gateway (StorageGRID 11.3).

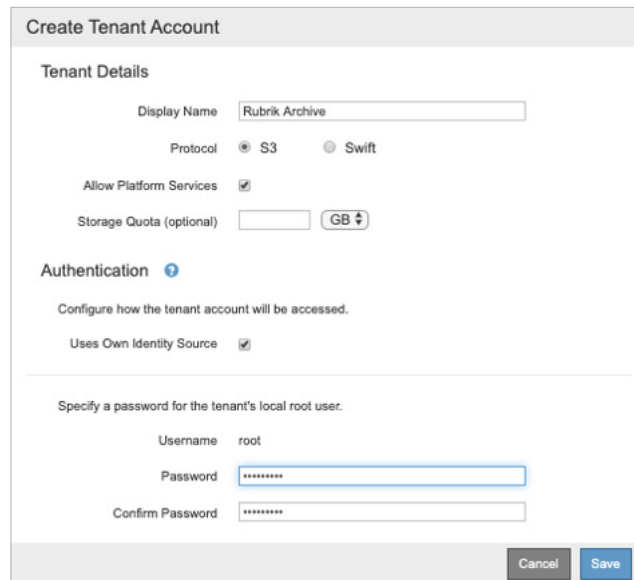
When configuring HTTPS connections to StorageGRID running versions earlier than 11.3 it will be necessary to create a self-signed certificate³. This certificate will then need to be installed on each node in the Rubrik cluster by Rubrik Support. If the StorageGRID cluster is running version 11.3, which is the current GA version at time of writing this document, this step may not always be required although is considered best practice.

ADDING STORAGEGRID AS AN ARCHIVE

1. Log into StorageGRID as root for the GRID and click on TENANTS



2. Create a tenant using the options S3, Allow Platform Services, Uses Own Identity Source and specify a root password for the tenant.



3. Click finish, then find the tenant you've just created and click on "sign-in" and enter the root password for the tenant.

	Rubrik Archive	46020354566164449797	S3	Sign in
---	----------------	----------------------	----	---------

4. Click on S3, My Credentials and Create an access key. In StorageGRID 11.3 the default is that the key will never expire. Click SAVE, then download the information as a CSV. Click Finish once the CSV has been downloaded.

Tenant Account Name	Access Key ID	Secret Access Key
Rubrik Archive	VZ504MXRQQN9S9XMNFY	2lanIJQV8BdMib0uz2ritUiXbyoToPskUOnYuFWf

5. Click on S3, Buckets, then Create Bucket. Create a bucket with the format of **bucketprefix-rubrik-0**

The bucketprefix can be something descriptive and will be used later when adding the bucket to Rubrik. The **-rubrik-** must be part of the bucket name. When adding multiple buckets to a Rubrik archive, the first bucket is number 0, the next 1 etc..

NOTE: Whilst its possible to create more than one bucket per Rubrik Archive Location, the current best practice to create a single bucket.

6. On a secure computer, start a shell session or open a terminal window. The computer must have the OpenSSL toolkit installed. For most Linux and Unix distributions, the standard operating system packages include the OpenSSL toolkit. The OpenSSL toolkit can also be downloaded and installed on Windows computers. At the command prompt type "openssl genrsa -out rubrik_encryption_key.pem 2048" – this will create a private RSA key in the current directory. This key will be used by Rubrik to encrypt/decrypt the data sent to/from the StorageGRID Object Store.

NOTE: If the RSA key created does not start with '-----BEGIN RSA PRIVATE KEY-----' and end with '-----END RSA PRIVATE KEY-----' upgrade your version of openssl

NOTE: The RSA key must be stored securely. In the event of a catastrophic disaster, it will not be possible to recover data from the StorageGRID to another Rubrik cluster without the encryption key.

7. Now log into the Rubrik cluster using an account with admin permission

Click the 'gear icon' in the top right hand corner of the Rubrik GUI, select Archive Locations then click on the + symbol in the top right corner.

Select "Object Store" from the drop down menu and enter the Access key and Secret key that was created earlier in step 4.

For hostname, enter the URL for the S3 endpoint you plan to use including the appropriate port number. For example

<https://mys3endpoint.domainname.com:8082>

Enter the **bucketprefix** that was used earlier in step 5 and specify the number of buckets - typically 1.

If you wish you can change the default description for the "Archive Location"

You can now paste in the RSA key created in step 6, and click the ADD button.

Add Archival Location

Archival Type
Object Store

Object Store Vendor
S3 Compatible (StorageGRID, Cloudian, IBM COS, or other compatible object storage)

Access Key
VZ504MXRQQN9S9XMNFY

Secret Key

Host Name
https://s3[redacted]:8082

Bucket Prefix
archive

Number of Buckets
1

Archival Location Name
S3Compatible:archive

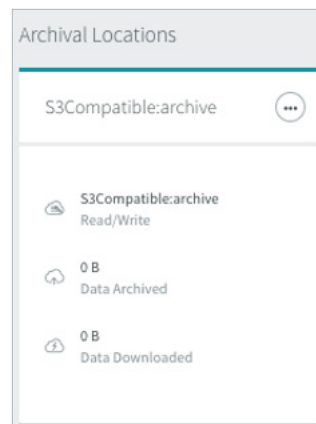
RSA Key ⓘ
-----BEGIN RSA PRIVATE KEY-----
MIIEowIBAAKCAQEA1qlv4hoitdFE4SD873f3V9niJuPFCEHWJiRGesUp0HTjDP3Z
PIQC21N7Q3gxjMFRaRSXoz6mLBvPtYPmfFPdWzDDCTw18WCpcPiWfM6tPBmkRWMP
zY4wkFO+XfWZgA5Ic9iidWBlwofl9yOkSR8nCTosJHEqekLvRtIbFrJOX99gcil
ISYzzum9vd042rJtiOG/l9VNeVDTIMLgc+68dbA65puFFoC8+BzLoPZPIIIDQ8yG
MIPD/S1cejof393Z8IUQt11/BMSilJoDUz9whK1+WZLDixlYIo5fzZSe3Dfve1L
FGljLxWWrzQc476XekzMjpavQf3oNBeESpJWFwIDAQABAolBAQC1cUR7xgGe2MG2
VSTWrv1nvWhJ3O53O0bauHDGZw8f/xzHo23P2wfrIVYoOlkHk4PfoSNHPbYxxFO+

Cancel Add

NOTE: If the bucket(s) on StorageGRID have not been pre-created using the format bucketprefix-rubrik-bucketnumber then the archive may not be added. The Rubrik GUI will remain on the “Add Archive Location” window and the message below will appear at the bottom of the window.

“The specified location constraint is not valid. For more information about Regions, see How to Select a Region for Your Buckets. (Service: Amazon S3; Status Code: 400; Error Code: InvalidLocationConstraint; Request ID: 1587123821706974; S3 Extended Request ID: null)”

8. The Archive should now be listed on the Archival Locations Dashboard and be ready to use.



BACKUP (WORKFLOW)

Backup's performed by Rubrik are controlled by SLA's. You can configure as many SLA's as required in order to meet your needs. That SLA may be assigned to multiple types of data sources - VMware VM's, NAS shares, SQL databases etc.

There are two parts to an SLA - the first controls the frequency of protection and retention periods, the second determines how long data resides locally on Rubrik and whether it should reside in other locations such as Cloud storage or another Rubrik cluster.

NOTE: The “Retention on Brik” period will vary based upon the Recovery Time Objective you have for the objects under protection by that Rubrik SLA. For example you may determine that most urgent restore requests relate to data that's <10 days old. So set that threshold accordingly.

It's also possible to configure Rubrik to Instantly Archive but also keep a copy on Rubrik until it meets the “Retention On Brik” age.

Create SLA Domain

SLA Domain Name

General Purpose

Continuous Data Protection ☐

Advanced Configuration ☐

Service Level Agreement

Choose how often we take snapshots and the length of time we keep them.

Take Snapshots:

Keep Snapshots:

Every (Hours)

1

For (Days)

31

Every (Days)

1

For (Days)

31

Every (Months)

1

For (Months)

3

Every (Years)

1

For (Years)

7

Local retention set to 10 days.

Snapshot Window

Take snapshots from: to

Take first full between: First Opportunity at

Remote Settings

Cancel

Create

Create SLA Domain

Remote Storage Configuration

Retention On Brik

0 10 days 7 years

Archival ☐

S3Compatible:archive ☒ Enable Instant Archive

Archival starts immediately, and is retained on the archival location for 7 years .

Replication

A replication target has not been set up yet. Please [add a replication target](#) to configure retention.

SLA Domain Creation

Cancel

Create

Objects ingested by Rubrik are stored space efficiently and distributed across the nodes. The data is then indexed at the file level. The “Remote settings” of the SLA control The “on-Brik” retention, archival & replication.

RECOVERY (WORKFLOW)

When recovering data via Rubrik, it is not necessary for the user to understand whether that data exists locally on Rubrik or within the StorageGRID archive bucket. The customer just selects the version of the object that’s required and Rubrik will identify where that data resides and initiate the recovery process. Depending on the method of recovery selected by the customer the data may be streamed directly to the client or staged upon the Rubrik cluster.

There are also heuristics that can be configured by Rubrik Support that may be used to affect the Cloud Storage egress versus download performance for certain recoveries.

BEST PRACTICES

STORAGEGRID COMPRESSION AND ENCRYPTION

Data stored in a Rubrik Archive is space-efficient, encrypted and compressed. It is recommended that StorageGRID’s Grid wide configuration settings for compression and encryption are left in their default settings of:

COMPRESSION=OFF
 ENCRYPTION=OFF

This avoids placing an unnecessary workload upon the StorageGRID nodes and will also reduce the response time for RANGE READ requests issued by the Rubrik Cluster when performing certain recovery tasks⁴.

STORAGEGRID CERTIFICATES

On versions of StorageGRID < 11.3, when using HTTPS archive connectivity, it is necessary to install StorageGRID certificates upon each node in the attaching Rubrik Cluster. This requires the assistance of Rubrik support. In version 11.3 of StorageGRID, this may not always be required although it is still considered best practice.

These storage certificates are created using the make-certificate utility on the StorageGRID Admin node⁵.

The FQDN to be used when using the make-certificate command can be either the FQDN of the StorageGRID API Gateway connected to Rubrik using network port 8082 or the FQDN of the StorageGRID Storage nodes when connecting to Rubrik using network ports 18082.

Issues which may be caused by the certificate not being present or by having expired could include Archive Location disconnection, not being able to add the Archive Location or not being able to upload rubrik_cluster_lock.txt at connection time.

STORAGEGRID OBJECT VERSIONING

Rubrik performs its own versioning and validation of data stored in the archive. It is not beneficial to configure StorageGRID object versioning on S3 buckets being used by Rubrik⁶.

STORAGEGRID CLOUD STORAGE POOLS

Cloud Storage Pools (essentially the ability to tier data off on-prem StorageGRID into public cloud) are not supported on buckets being used as a Rubrik Archive.

RUBRIK ARCHIVE READER/WRITER CONNECTION AND RUBRIK REPLICATION

In the event of a catastrophic failure of a Rubrik cluster that has been Archiving data to StorageGRID, it is possible to connect another Rubrik system to that archive and recover the information stored in the archive. The Rubrik System may be connected as an Archive Reader and if necessary promoted to an Archive Writer - taking ownership of that Archive.

Data stored in the Archive is encrypted using a customer provided RSA key. In order to enable an alternative Rubrik system to connect to that Archive Location, the customer must use the same key.

NOTE: It is not possible to configure a Rubrik Cluster as an Archive Reader/Writer if that Cluster has already been in a Replication Relationship with the original Rubrik Cluster. Rubrik 5.0.1 prevents such a configuration being made. Earlier Rubrik versions may allow such a configuration but it is not supported. ***Consult the appropriate version of Rubrik User Guide and Release notes for more information.***

USEFUL HOTFIXES FOR STORAGEGRID 11.2

0035954: Fix retrievals of compressed multipart object with thousands of segments

0036375: Fix high memory usage while reading compressed object

WHAT'S NOT SUPPORTED

Cloud Storage Pools are not supported on buckets being used as a Rubrik Archive.

Connecting a “new” Rubrik cluster to an existing StorageGrid Archive as an Archive Reader/Writer if the original Rubrik cluster was in a Replication relationship with the “new” Rubrik Cluster.

CONCLUSION

Many organizations need long-term retention of their backup data, whether because of specific business needs, legal reasons or to comply with various types of regulations. Historically, this was achieved by offloading backups to tape on a regular basis, and sending these tapes offsite with the help of a specialized third-party service provider.

The tape offsite process is somewhat inefficient and many customers today wish to remove manual handling steps when recovering from their long-term data. Because of this many organizations have chosen to replace tape with cloud storage. This introduces operational simplicity as well as additional benefits such as high-availability, durability and streamlined recoveries.

Some enterprises that are willing to use such cloud storage cannot choose the public cloud due to a variety of reasons such as data sovereignty concerns, or legal complexity. In such cases, private object storage solutions like NetApp StorageGRID meet such requirements, and at the same time bring the same benefits as public cloud storage on-premises.

Compared to legacy data management systems, Rubrik significantly simplifies the protection, recovery and retention of data on-premises, and in public and private cloud storage.

SOURCES AND NOTES

- 1 <http://docs.netapp.com/sgws-113/topic/com.netapp.doc.sg-install-vmw/GUID-99A029C8-F7FB-4027-A5DB-CAAACF3F39E7.html>
- 2 <http://docs.netapp.com/sgws-112/topic/com.netapp.doc.sg-install-vmw/GUID-99A029C8-F7FB-4027-A5DB-CAAACF3F39E7.html>
- 3 <http://docs.netapp.com/sgws-112/topic/com.netapp.doc.sg-admin/GUID-E1AF31C7-BDA2-495C-ABFE-C3A45A12B026.html>
- 4 <https://docs.netapp.com/sgws-113/topic/com.netapp.doc.sg-admin/GUID-A8365935-D6DA-4CBD-B5B6-4269527303DB.html>
- 5 <http://docs.netapp.com/sgws-112/topic/com.netapp.doc.sg-admin/GUID-E1AF31C7-BDA2-495C-ABFE-C3A45A12B026.html>
- 6 <https://docs.netapp.com/sgws-113/topic/com.netapp.doc.sg-s3/GUID-53596498-9334-44DB-A4CE-DFEC28CF21FF.html>

FURTHER RESOURCES

<https://www.netapp.com/us/media/tr-4812.pdf>

<https://support.rubrik.com> (login required)

VERSION HISTORY

Version	Date	Summary of Changes
1.0	May 2020	Initial Release



Global HQ

1001 Page Mill Rd., Building 2
Palo Alto, CA 94304
United States

1-844-4RUBRIK
inquiries@rubrik.com
www.rubrik.com

Rubrik, the Multi-Cloud Data Control™ Company, enables enterprises to maximize value from data that is increasingly fragmented across data centers and clouds. Rubrik delivers a single, policy-driven platform for data recovery, governance, compliance, and cloud mobility. For more information, visit www.rubrik.com and follow [@rubrikInc](https://twitter.com/rubrikInc) on Twitter. © 2020 Rubrik. Rubrik is a registered trademark of Rubrik, Inc. Other marks may be trademarks of their respective owners.

20200504_v1