



Technical Report

# Veeam Cloud Tier with StorageGRID Object Storage

Ahmad Edalat, Steven Pruchniewski, NetApp  
Adam Bergh, Veeam  
May 2019 | TR-4777

In partnership with



## Abstract

This document describes the steps required to configure Veeam to use NetApp® StorageGRID® as an object storage target. It also explains the Veeam backup procedure in detail, including configuring various settings.

## TABLE OF CONTENTS

<b>1</b>	<b>Introduction.....</b>	<b>3</b>
<b>2</b>	<b>Prerequisites.....</b>	<b>3</b>
<b>3</b>	<b>Veeam Configuration.....</b>	<b>4</b>
<b>4</b>	<b>Tiering Conditions.....</b>	<b>10</b>
4.1	Sealed Backup Chains.....	11
4.2	Backup Job Options.....	11
4.3	Backup Job Options.....	12
<b>5</b>	<b>Tiering Operation.....</b>	<b>15</b>
5.1	Data Tiering.....	15
	<b>Troubleshooting.....</b>	<b>21</b>
	Switching to Maintenance Mode.....	21
	Evacuating Backups from Extents.....	22
	<b>Version History.....</b>	<b>23</b>

## LIST OF TABLES

Table 1)	Types of indexes.....	18
----------	-----------------------	----

## LIST OF FIGURES

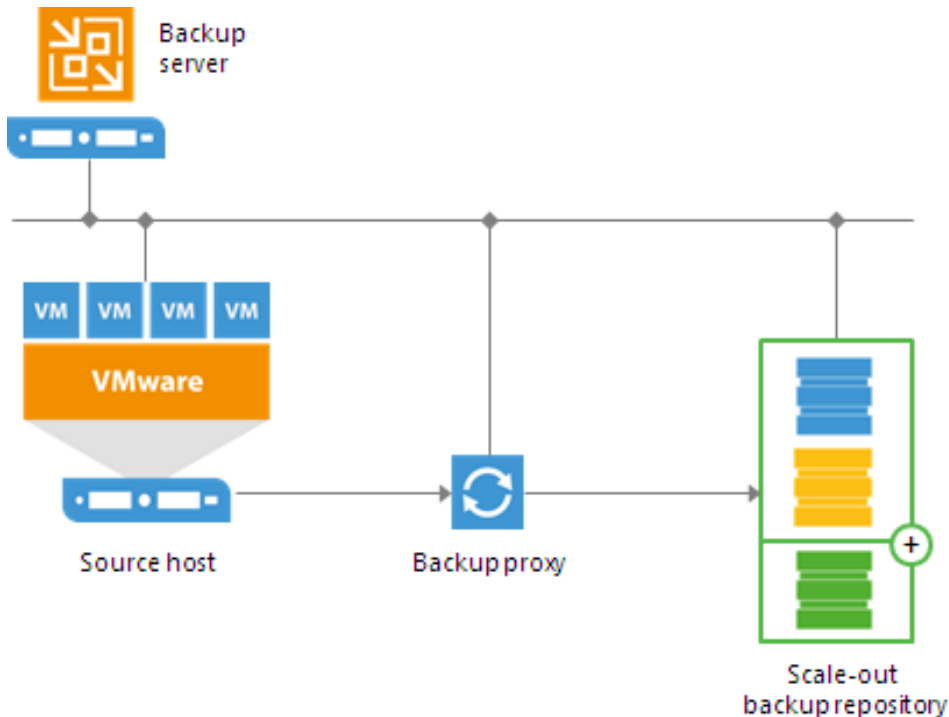
Figure 1)	Overview of Veeam backup procedure (source).....	3
Figure 2)	Storage configuration settings.....	12
Figure 3)	Backup advanced settings.....	13
Figure 4)	Inactive and active backup chains.....	13
Figure 5)	Inactive backup chain available to be offloaded.....	14
Figure 6)	New backup copy job window.....	15
Figure 7)	Manual offload operation.....	16
Figure 8)	Overview of Veeam backup procedure tiering to object storage.....	17
Figure 9)	ArchiveIndex directory folder structure.....	18
Figure 10)	New Scale-out Backup Repository window.....	19
Figure 11)	Creation of full backup.....	20
Figure 12)	Backup files being offloaded onto capacity tier.....	20
Figure 13)	Selecting Maintenance mode.....	22
Figure 14)	Evacuating backups.....	23

## 1 Introduction

Archiving backups is one of the many workloads that can take advantage of the NetApp StorageGRID object storage solution. Veeam uses StorageGRID as secondary storage to enable the Veeam scale-out backup repository (SOBR). Backup data is first moved into primary storage configured by the user. It is then offloaded from primary storage into object storage (or, in Veeam terminology, the capacity tier). Veeam collects data and transfers it to the Netapp StorageGRID repository by running a background activity called SOBR offload, which is executed every 4 hours. The default behavior is to move data that is 30 days old.

The backup and replication software uses a logical grouping of several backup repositories, called a scale-out backup repository. This logical grouping is used to create a pool of storage devices to offload data from primary storage into object storage (Figure 1).

Figure 1) Overview of Veeam backup procedure ([source](#)).



## 2 Prerequisites

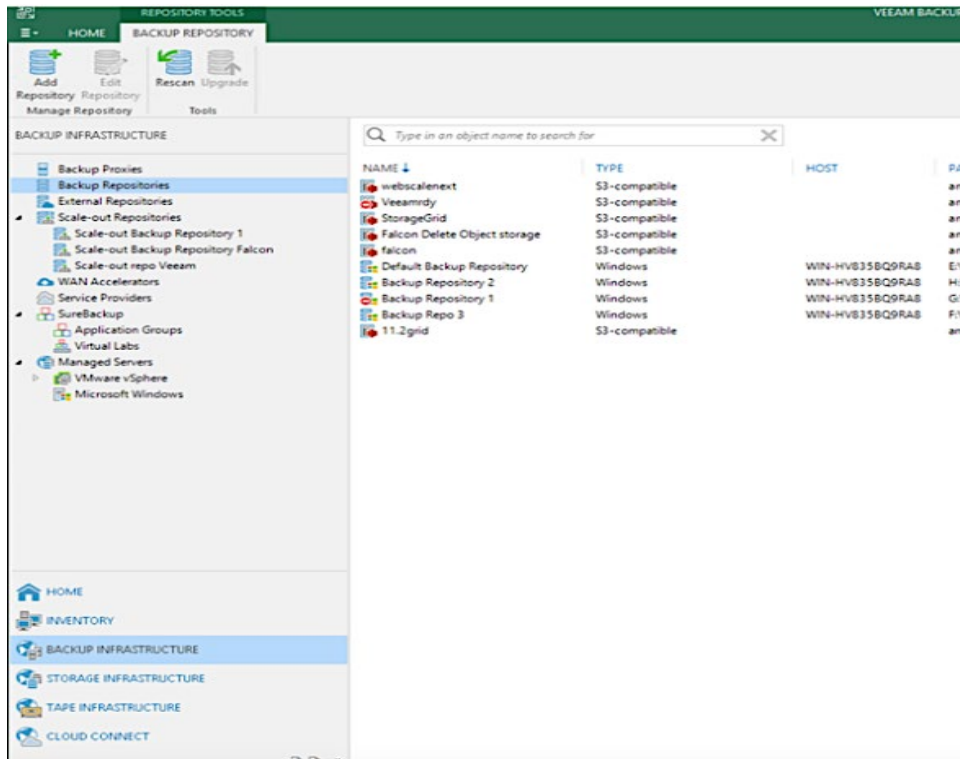
The following list of prerequisites is needed to configure Veeam with StorageGRID:

- StorageGRID 11.1.1.3 or later
  - Configuration: For the SSL certificate, make sure that an object-storage API service-endpoints server certificate is installed. A self-signed SSL is adequate.
  - DNS-configured hostname for S3 endpoint
  - S3 tenant with credentials
  - A bucket configured to store the backed-up data
- Veeam 9.5.4.2399

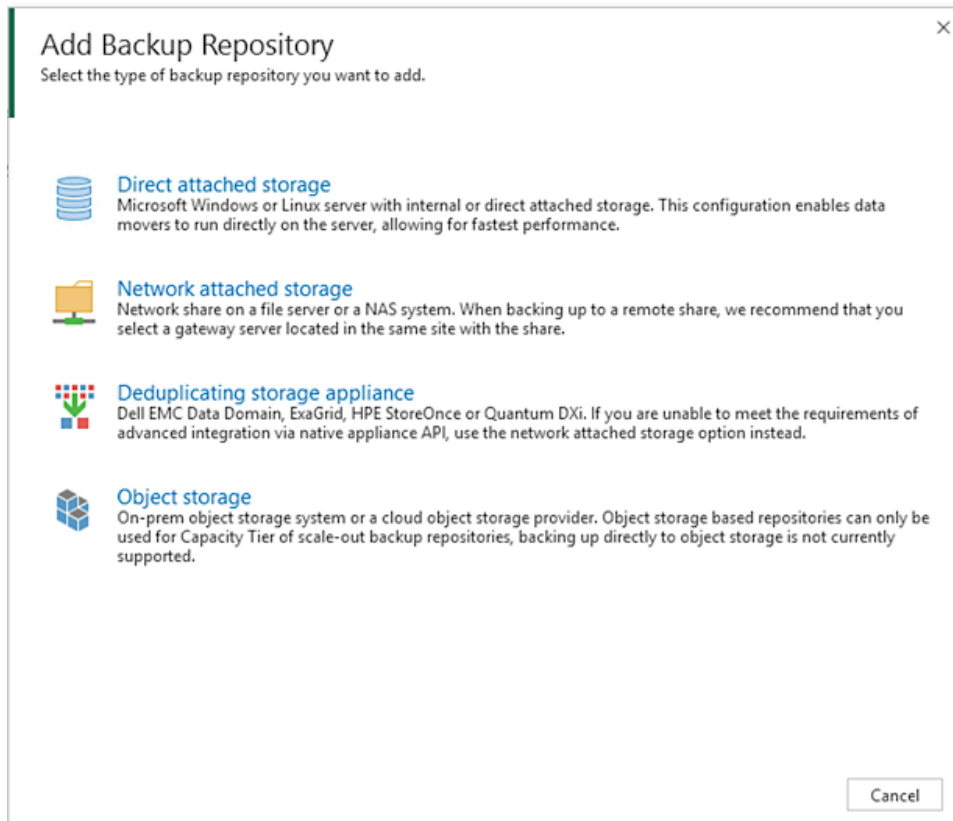
### 3 Veeam Configuration

To configure the Veeam application, complete the following steps:

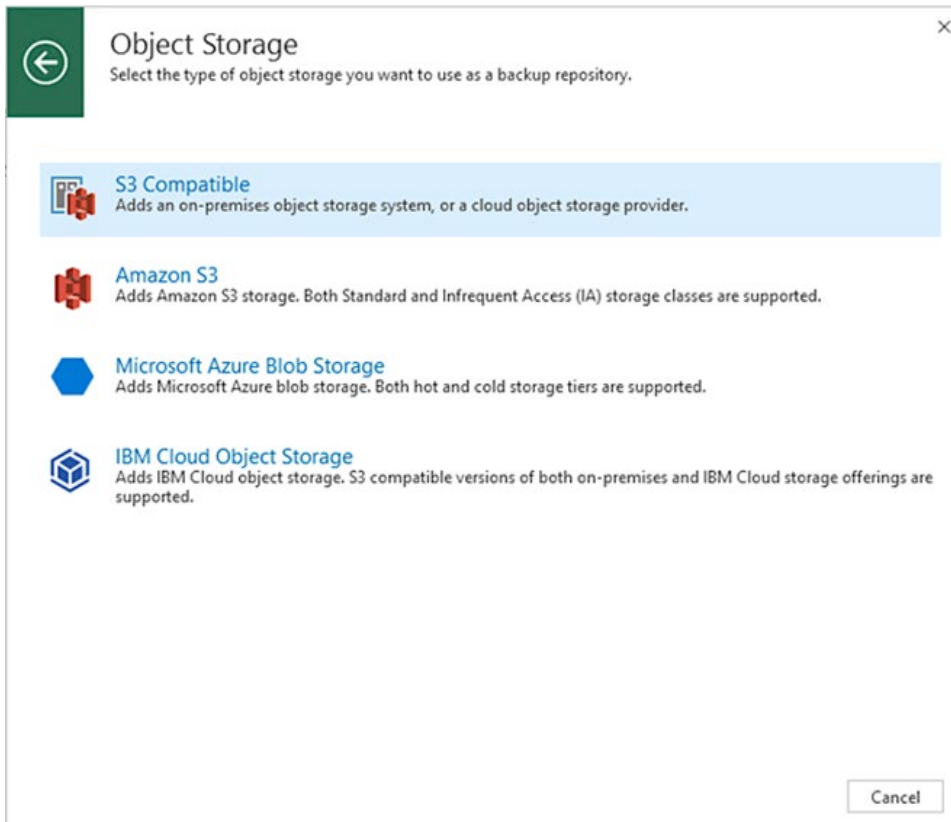
1. Launch the Veeam application. In the left panel, click Backup Infrastructure and select Backup Repositories.



2. Click Add Repository and select Object Storage to set up your object storage account.



3. In the resulting window, select S3 Compatible.



4. Enter a name for your object storage repository. In this example, it is named StorageGrid. Click Next.

5. Provide the service endpoint, the region, and the credentials used for the tenant account. The endpoint must be a host name that can be resolved by DNS and configured with an SSL certificate, as noted in section 2, "Prerequisites." Click Next.

New Object Storage Repository

**Account**  
Specify account to use for connecting to S3 compatible storage system.

Name	Service point: https://webscaledemo.netapp.com:8082
Account	Region: us-east-1
Bucket	Credentials: JDYGCWR60SIT65PBN441 (last edited: less than a day ago) <a href="#">Add...</a> <a href="#">Manage cloud accounts</a>
Summary	<input type="checkbox"/> Use the following gateway server: WIN-HV835BQ9RA8 Select a gateway server to proxy access to the object storage system. If no gateway server is specified, all scale-out backup repository extents must have direct network access to the storage

< Previous   Next >   Finish   Cancel

- Enter the name of the bucket previously created in the StorageGRID Tenant Manager UI or other S3 client.

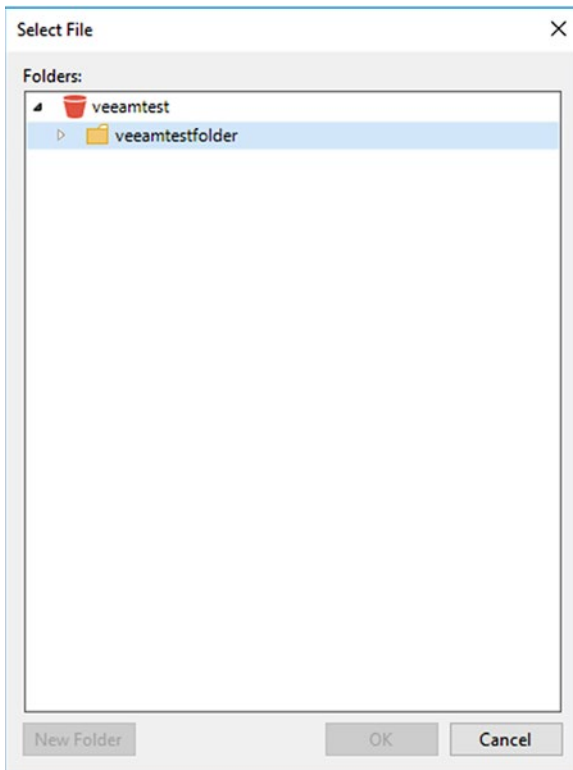
New Object Storage Repository

**Bucket**  
Specify object storage system bucket to use.

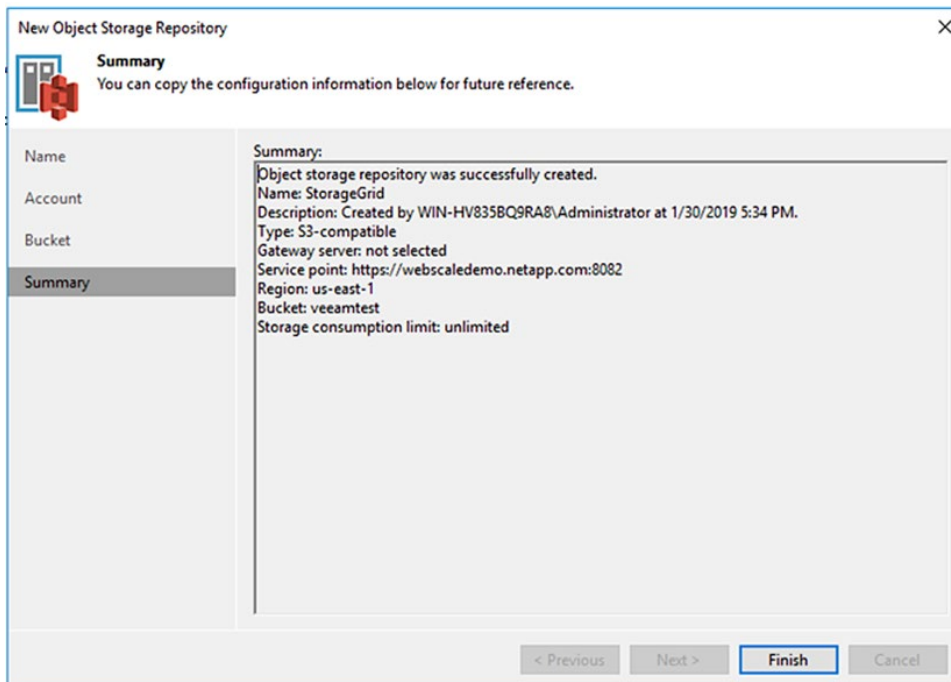
Name	Bucket: veeamtest
Account	Folder: veeamtestfolder <a href="#">Browse...</a>
Bucket	<input type="checkbox"/> Limit object storage consumption to: 10 TB This is a soft limit to help control your cloud storage spend. If the specified limit is exceeded, the already running data offload tasks will be allowed to complete, but no new tasks will start.
Summary	

< Previous   Next >   Finish   Cancel

- Still in the Bucket window, click Browse and create a folder within the bucket.



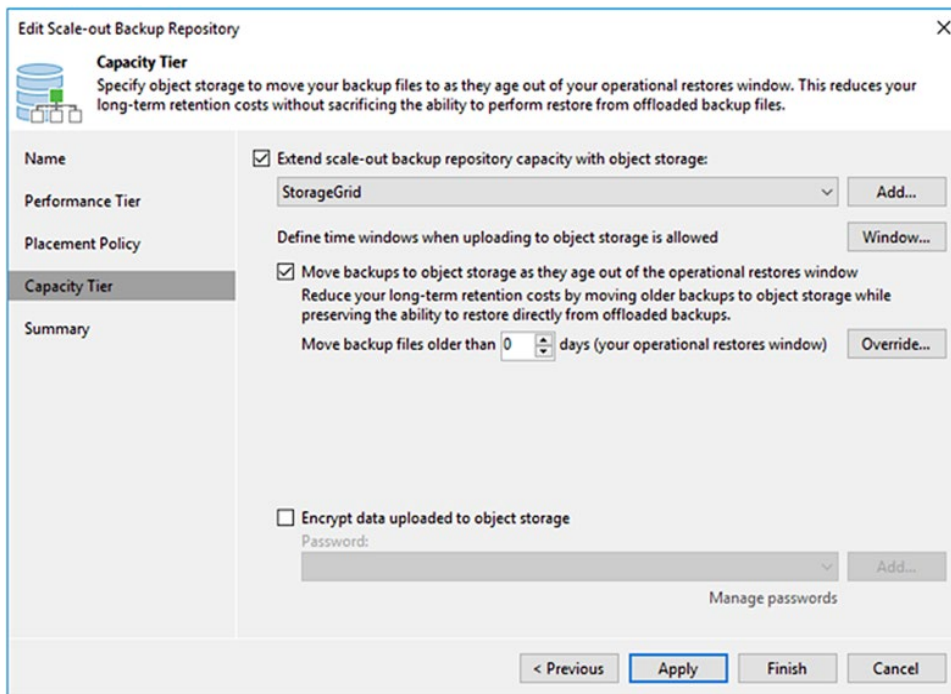
8. Your object storage account is now set up. Click Next to open a summary window showing the details of your account.



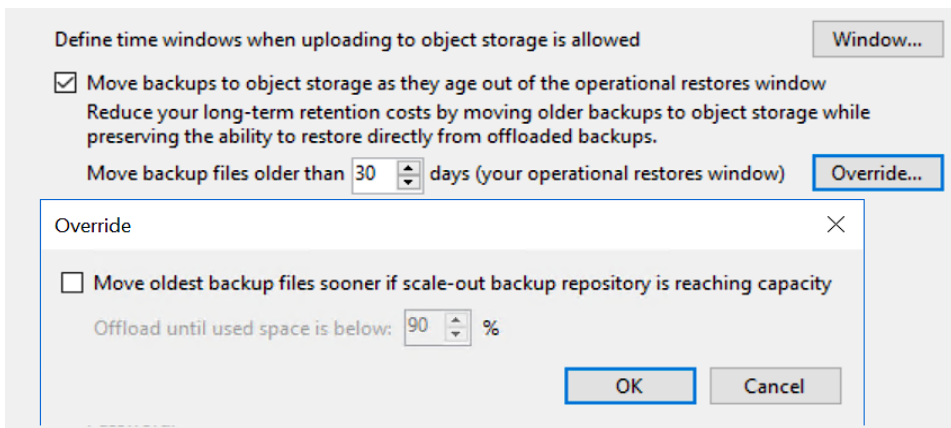
9. Navigate to Scale-Out Repositories and click to add a scale-out repository. Here you combine performance tier and capacity tier. In the prompt window, enter a name for your scale-out repository.







Veeam allows you to move the oldest backup files sooner if the backup repository is reaching capacity.



**Note:** After the backup chains are created and stored on the performance tier (SOBR's standard, non-object extents), they become a subject for tiering conditions, based on the configured policy of the capacity tier.

## 4 Tiering Conditions

The capacity tier employs the automated Offload Job process to handle the validation, verification, and transfer of data to object storage. The Veeam Offload Job process includes the following tasks:

- Verifying the eligibility of the backup chain to be processed and offloaded to the capacity tier (policy-based). Only "sealed" chains are subject to the offload operations.
- Collecting verified backup chains from the extent of each SOBR and sending them directly or through the designated gateway service to StorageGRID.

- Logging session results for further access and review upon request at History > System > '<Name of SOBR Target> Offload'.

## 4.1 Sealed Backup Chains

The concept of “sealed” backup chains is simple, but it is fundamental to the successful use of capacity tier functionality in your environment. Sealed backup chains are backup chains that do not have any ongoing active operations toward them and have no scheduled operations that might require modification of the backup chain's files.

Veeam backup and replication allows the following types of jobs to be configured with SOBR as the target repository:

- Backup jobs (forever forward incremental, forward incremental, and reverse incremental)
- Backup copy jobs (simple or [GFS-enabled](#))
- Backup jobs created by Veeam Agent for Linux 2.0 or later
- Backup jobs created by Veeam Agent for Microsoft Windows 2.0 or later

Inherently, these backup job types are subject to the capacity tier offload operations based on the policy configured, except for the following:

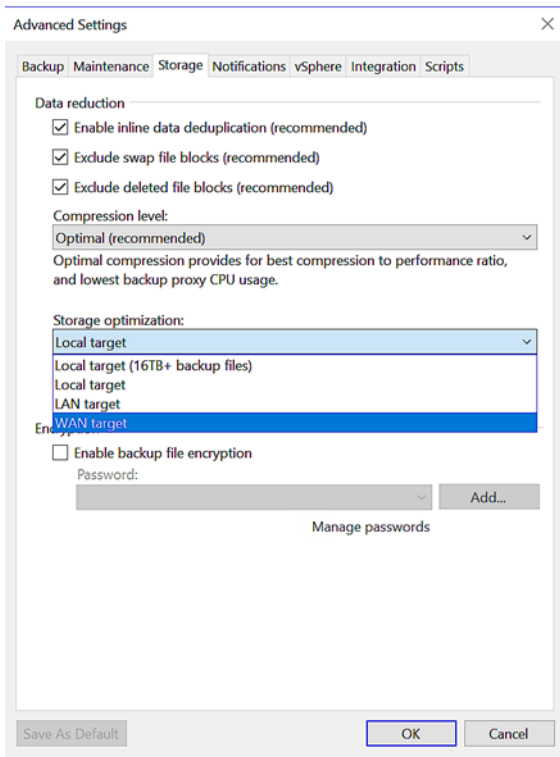
- Backup. Forever forward incremental
- Backup copy. Simple (without GFS policy configured)

## 4.2 Backup Job Options

Configuration settings chosen for the backup job using SOBR with the capacity tier as the target define the type of backup chain as it gets stored in the scale-out backup repository and is further processed by the capacity tier's Offload Job process.

In addition to the type of backup chain configured, the Storage Optimization and Compression settings define the block size for the backup data. The block size in turn defines the number of blocks consumed on the object storage as the backup chains are offloaded to the capacity tier (Figure 2).

Figure 2) Storage configuration settings.



## 4.3 Backup Job Options

Sealing of backup chains directly depends on the type of the selected backup mode.

### Forward Incremental

The main difference between forever forward incremental, which is not eligible for capacity tier offload operations because it is an “active” backup chain, and forward incremental is the creation of a periodic full or synthetic full backup (Figure 3).

Figure 3) Backup advanced settings.

Advanced Settings

Backup Maintenance Storage Notifications vSphere Integration Scripts

Backup mode

☐ **Reverse incremental (slower)**  
Increments are injected into the full backup file, so that the latest backup file is always a full backup of the most recent VM state.

☒ **Incremental (recommended)**  
Increments are saved into new files dependent on previous files in the chain. Best for backup targets with poor random I/O performance.

☒ Create synthetic full backups periodically Days...

Create on: Saturday

☐ Transform previous backup chains into rollbacks  
Converts previous incremental backup chain into rollbacks for the newly created full backup file.

Active full backup

☐ Create active full backups periodically

☐ Monthly on: First Monday Months...

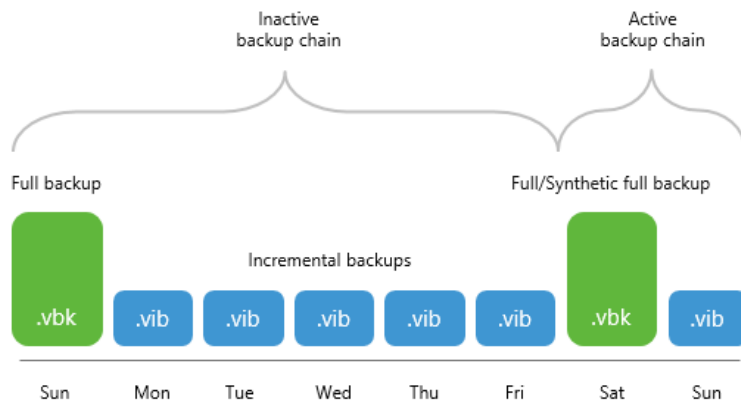
☒ Weekly on selected days: Days...

Saturday

Save As Default OK Cancel

The creation of full and synthetic backups allows the preceding backups and their respectively dependent restore points (incremental backups) to become “sealed” (Figure 4).

Figure 4) Inactive and active backup chains.



The inactive backup chain is considered sealed, and it can be further validated by the Offload Job engine for eligibility to be moved to the capacity tier.

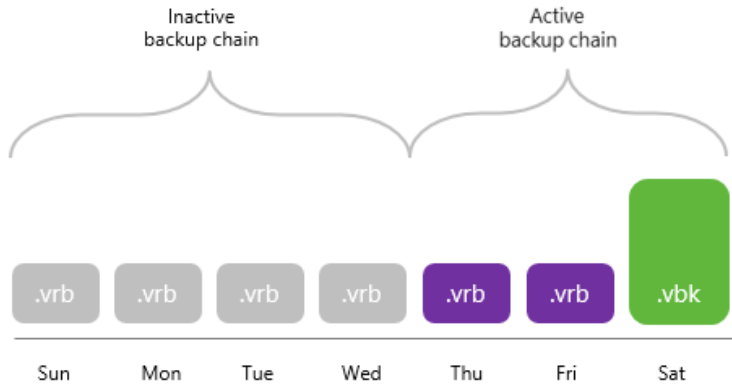
The next time that synthetic full and active full files are successfully created (.vbk), Veeam software seals the previous chain.

## Reverse Incremental

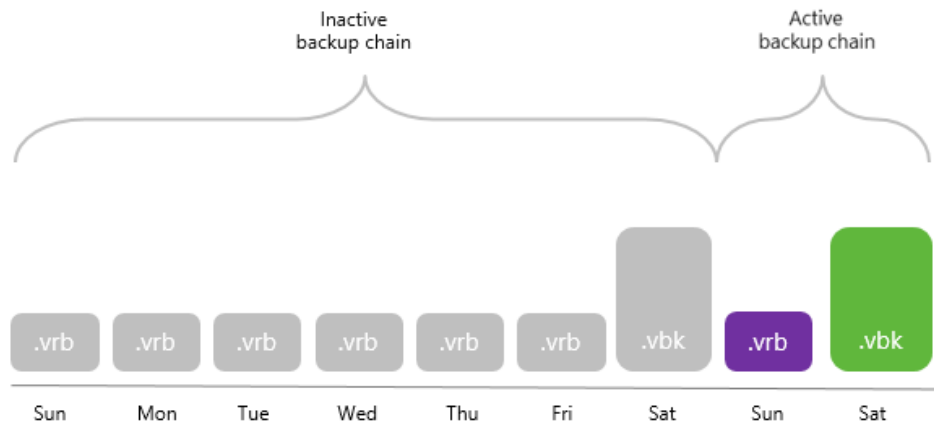
Using this backup type ensures that the changes retrieved through incremental backups are injected into the full backup file.

The latest full backup (.vbk) and the first two dependent incremental backups (.vib) are considered active, while all older incremental files are considered inactive and can be offloaded (Figure 5).

**Figure 5) Inactive backup chain available to be offloaded.**



**Figure A**



**Figure B**

## Backup Copy Jobs

Backup copy jobs with configured retention policies are validated for backup chain activity status based on a simple logic. Only full backup files created as a part of a retention policy can be offloaded to the capacity tier (weekly, monthly, quarterly, or yearly backups).

In Figure 6, the Backup Copy Job retention policy is configured to store four weekly backups, which are subject to offload operations to the capacity tier.

Figure 6) New backup copy job window.

**New Backup Copy Job**

**Target**  
Specify the target backup repository, amount of most recent restore points to keep, and retention policy for full backups. You can use map backup functionality to seed the backup files.

**Job**  
**Objects**  
**Target**  
Data Transfer  
Schedule  
Summary

Backup repository:  
Tenant-B629-Cloud-Repo-01 (Cloud repository)  
25.7 GB free of 40.0 GB [Map backup](#)

Restore points to keep: 7

☒ Keep the following restore points as full backups for archival purposes

Weekly backup: 4 Saturday [Schedule...](#)  
Monthly backup: 0 First Sunday of the month  
Quarterly backup: 0 First Sunday of the quarter  
Yearly backup: 0 First Sunday of the year

☐ Read the entire restore point from source backup instead of synthesizing it from increments

Advanced settings include health check and compact schedule, notifications settings, and automated post-job activity options. [Advanced](#)

< Previous Next > Finish Cancel

## Veeam Agent Jobs

Backup jobs or backup copy jobs sourced from Veeam Agent for Windows or Veeam Agent for Linux are subject to the same policies as those applied to VM backup jobs or VM backup copy jobs.

## 5 Tiering Operation

### 5.1 Data Tiering

Tiering operations can be executed automatically (every 4 hours), or they can be performed manually. The automated operation cannot be disabled.

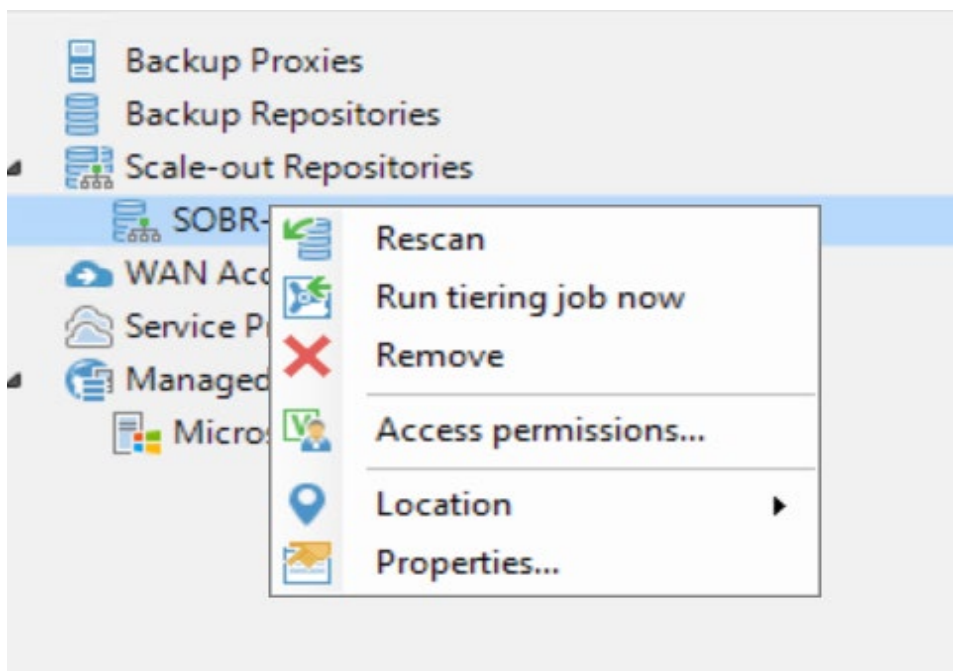
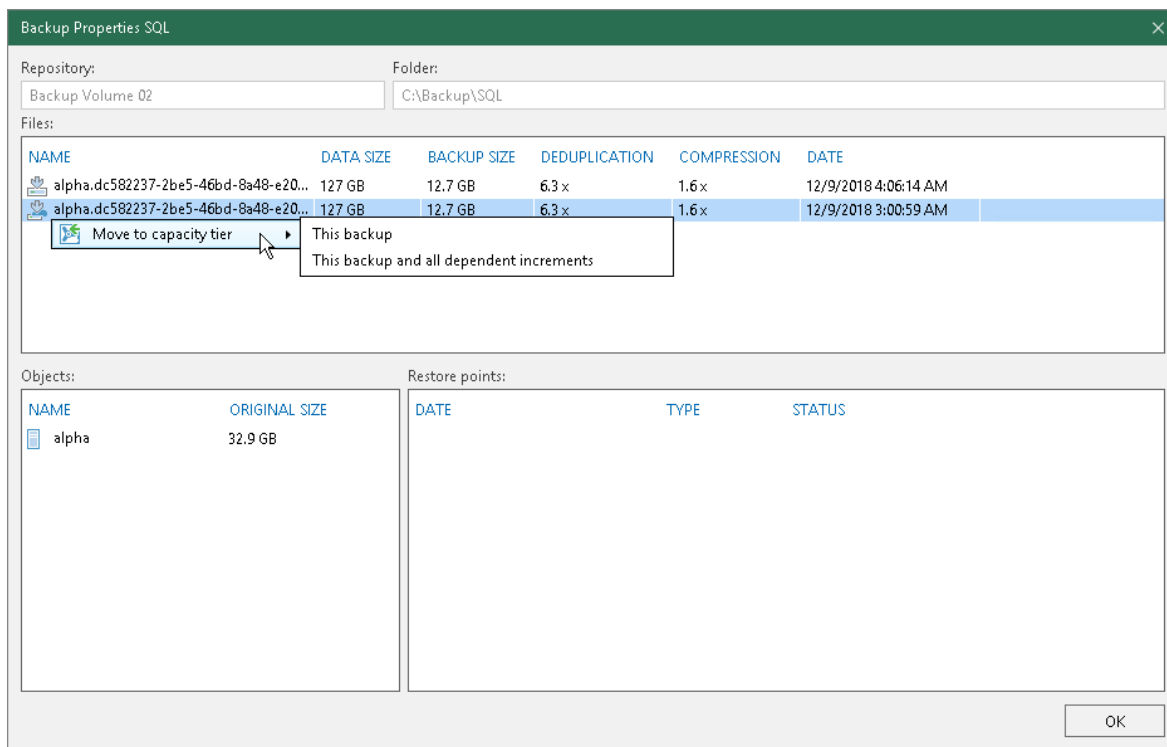
### Manual Offload Operations

Manually initiated offload operations are primarily used to run offload jobs to the capacity tier instead of waiting for the next scheduled automated run. This feature comes handy in scenarios like these:

- The performance tier is getting low on space and an expedited offload operation is desired for the data stored on the performance tier backup chains.
- There is a modified (reduced) restore operational window (the age of backup chains before they are eligible for offloading to the capacity tier).

Figure 7 shows how to initiate a manual offload operation. Select an inactive backup chain from the Backup Properties window.

**Figure 7) Manual offload operation.**





## Data Transfer Considerations

To optimize network traffic flow when working with the capacity tier during the configuration of the StorageGRID object storage repository, NetApp recommends retaining the default settings for the gateway server. You should also make sure that all extents of the scale-out backup repository that are configured with the capacity tier have direct internet access.

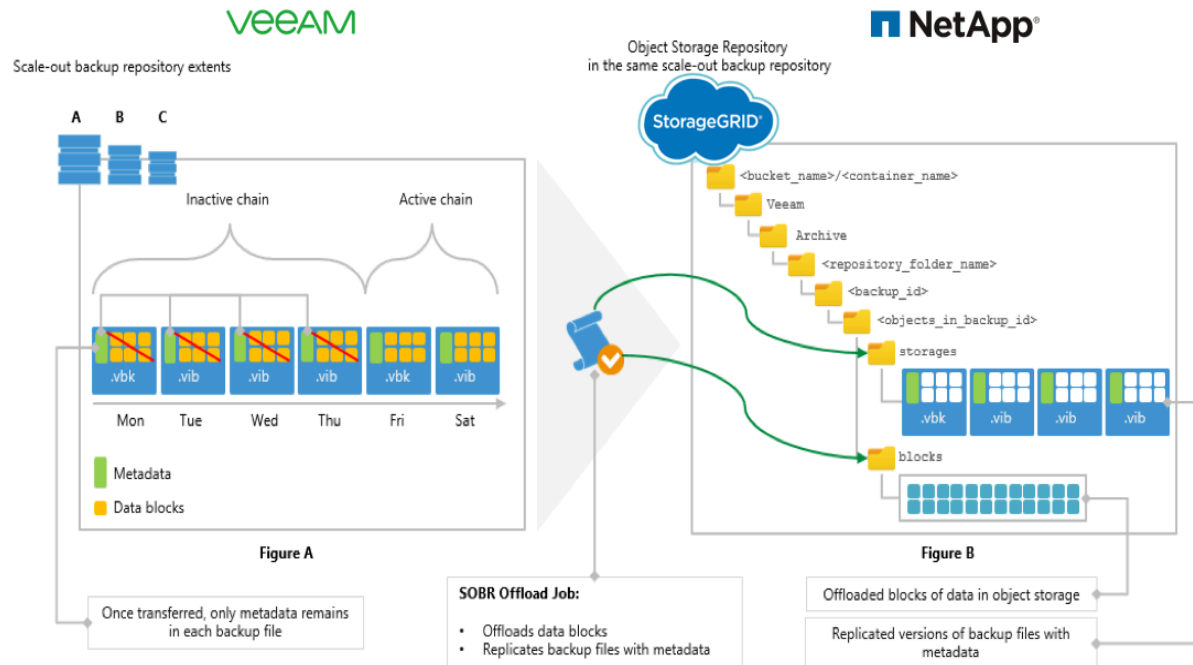
If your organization uses NAT or different types of firewalls and access to the internet is limited, NetApp recommends specifying a gateway server in the configuration for the object storage repository.

## Metadata and Indexes

Metadata files are maintained for each backup file; they contain information about the structure of the files.

The Offload Job process identifies any inactive backup chains that are subject to transfer to the capacity tier. During the data transfer process, metadata is copied along with the backup chain files to the capacity tier. The “dehydrated” files with metadata information about the offloaded backup chain files are also stored on the performance tier (Figure 8).

Figure 8) Overview of Veeam backup procedure tiering to object storage



The system maintains a copy of metadata at both locations. Therefore, it can perform operations like synchronizing backup chains between the SOBR extents and the capacity tier, restoring data back to production systems, and downloading data back to the performance tier. Additionally, metadata files serve as the source of information for indexes created during each offload operation.

Indexes serve the purpose of delivering an optimized solution for data transfer operations. Storing hash information about offloaded blocks and maintaining it for each backup chain ensures that any blocks already transferred to the capacity tier won't be transferred again. This arrangement delivers a balance between cost and efficiency for the whole system.

Index files are stored in the ArchiveIndex directory on the source extents from which the data was offloaded. Figure 9 shows the folder structure and Table 1 lists and describes the kinds of indexes.

Figure 9) ArchiveIndex directory folder structure.

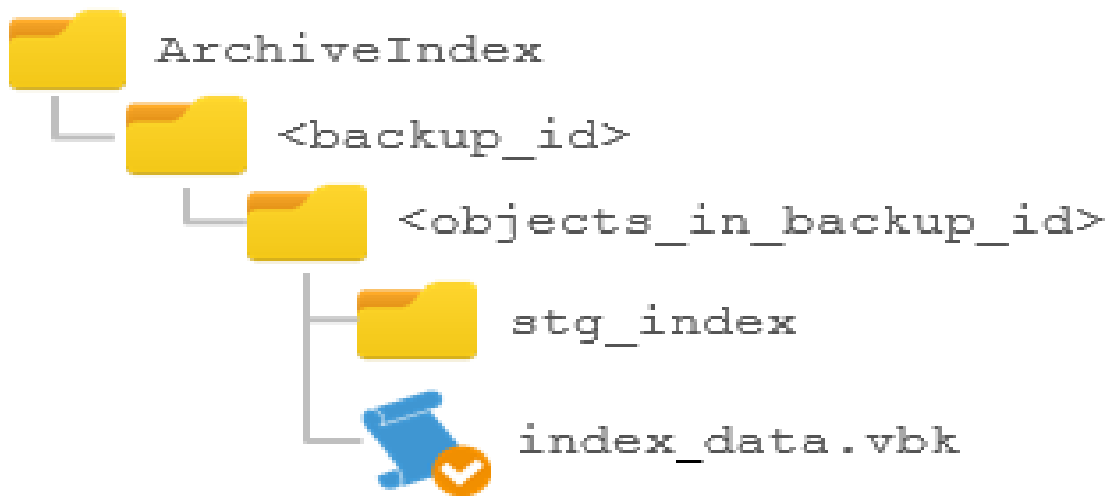


Table 1) Types of indexes.

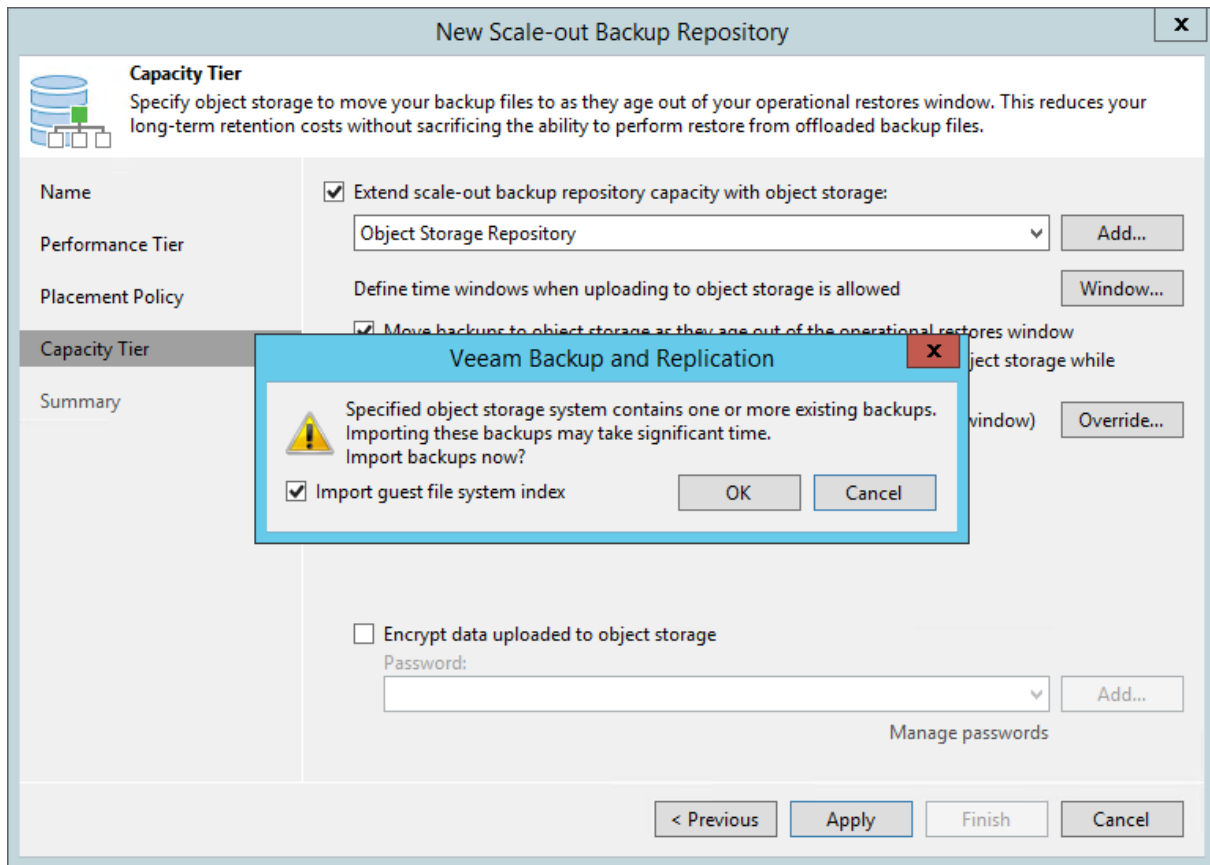
Index	Description
ArchiveIndex	The root directory for keeping indexes. This directory is created in the repository of an extent.
<backup_id>	Contains objects in a backup file.
<objects_in_backup_id>	An identifier of an object in a backup file. If a backup was created using the <a href="#">Per-VM</a> method, each VM is placed in its own directory. If a backup was created as single storage, all the VMs are placed in a single directory.
stg_index	Contains actual indexes of the offloaded backup files (.vbk, .vib, or .vrb).
index_data.vbk	Contains meta information about hash values stored in index files.

Indexes are modified whenever changes are made to the backup chain, and the hash table, consecutively, also needs to be updated. When an index file is rebuilt, the scale-out repository rescan operation initiates the process (Table 1).

## Capacity Tier with Existent Backups

Due to the efficiency of storing indexes and metadata on the performance and capacity tiers, the system prompts to initiate a synchronization process if you add an object repository with backup data already stored on it.

Figure 10) New Scale-out Backup Repository window.



During the synchronization process, Veeam downloads backup files with metadata located in the object storage repository to the extents that are part of a scale-out backup repository that is being added. The placement of the downloaded files is automated and based on resource availability. Disk space availability is the primary metric used during the automated resource-availability placement process.

## Source-Side Deduplication

The availability of metadata and index files allows the blocks of data to be offloaded to the capacity tier only once. If the block has already been transferred to the capacity tier and logged in the index file, the block is not transferred again. This reduces the amount of traffic required for the operations and also reduces general disk space consumption on the StorageGRID bucket.

For example, in a per-VM backup chain consisting of two full and three incremental backups, the incremental backups depend on a single full backup and are considered to be a sealed backup chain. In other words, the creation of the highlighted full backup allowed the previous chain to become inactive (Figure 11)

Figure 11) Creation of full backup.

Backup Properties BKP-2-CLOUD-TIER






Repository:

Cloud repository 1

Folder:

BKP-2-CLOUD-TIER

Files:

NAME	DATA SIZE	BACKUP SIZE	DEDUPLICATION	COMPRESSION	DATE
 vm-10060.b5a4_E20FD2019-03-07T004952.vbk	40.0 GB	21.6 GB	1.2 x	1.6 x	3/7/2019 12:49:52 AM
 vm-10060.b5a4_2A3AD2019-03-06T230635.vib	24.1 MB	30.8 MB	1.0 x	1.0 x	3/6/2019 11:06:35 PM
 vm-10060.b5a4_FE06D2019-03-06T230233.vib	81.1 MB	54.3 MB	1.0 x	1.5 x	3/6/2019 11:02:33 PM
 vm-10060.b5a4_2169D2019-03-06T225604.vib	58.1 MB	43.7 MB	1.0 x	1.4 x	3/6/2019 10:56:04 PM
 vm-10060.b5a4_A123D2019-03-06T224615.vbk	40.0 GB	21.6 GB	1.2 x	1.6 x	3/6/2019 10:46:15 PM

The creation of another full backup effectively seals the previous full backup still residing on the performance tier. The scheduled Offload Job process transfers data to the capacity tier. It also now applies source-side optimizations for data transfers and transfers only the blocks not found on the capacity tier for this backup chain (Figure 12)

Figure 12) Backup files being offloaded onto capacity tier.

SOBR-01-142 Offload

Job progress:

100%

1 of 1 objects

SUMMARY

Duration:

00:57

Processing rate:

14 MB/s

Bottleneck:

Source

DATA

Processed:

236.4 MB (100%)

Read:

236.4 MB

Transferred:

59.4 MB (4x)

STATUS

Success:

1

Warnings:

0

Errors:

0

THROUGHPUT (ALL TIME)

<

By using metadata and index file information about blocks of transferred data and where they exist, the system recognized only 59.4MB of new blocks to be transferred to the capacity tier. That's because the

majority of 21.6GB of the full backup blocks were already present for this backup chain in the object storage repository.

## Intelligent Block Recovery

Based on the availability of metadata files at source extents and the capacity tier, the intelligent block retrieval procedure is applied during the recovery operation. If blocks of data requested are also present on the performance tier, they do not need to be retrieved from the capacity tier to complete the requested operations. Also, the blocks not present on the performance tier are retrieved from object storage.

Not only is this functionality cost and resource effective, it also expedites operations like instant VM recovery, for which the location of blocks of data directly affects the overall speed of recovery.

## 6 Troubleshooting

### 6.1 Switching to Maintenance Mode

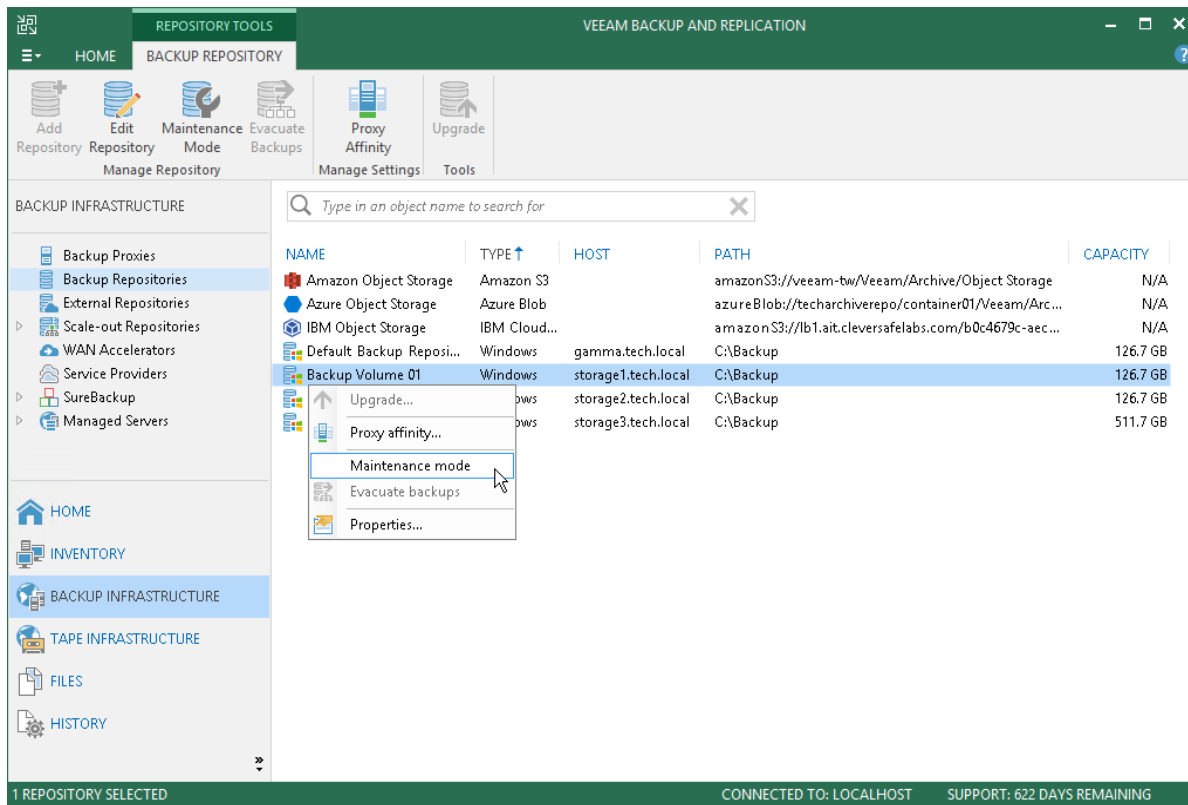
If you need to perform service actions on an extent (for example, to upgrade it or install a patch on it), you can put an extent of the scale-out backup repository in Maintenance mode. You must also put the extent in Maintenance mode before you evacuate backups from this extent.

To put an extent in Maintenance mode, complete the following steps:

1. Open the Backup Infrastructure view.
2. In the inventory pane under Scale-out Repositories, select a scale-out backup repository.
3. In the working area, select the extent and click Maintenance Mode in the ribbon, or right-click the extent and select Maintenance Mode.

To bring the extent back to the normal operational mode, select the extent and click Maintenance Mode in the ribbon, or right-click it and select Maintenance Mode (Figure 13Figure 13).

Figure 13) Selecting Maintenance mode.



## 6.2 Evacuating Backups from Extents

To remove an extent from the scale-out backup repository, you must first evacuate backups from the extent. When you evacuate backups, Veeam Backup & Replication moves backup files from the extent to other extents that belong to the same scale-out backup repository.

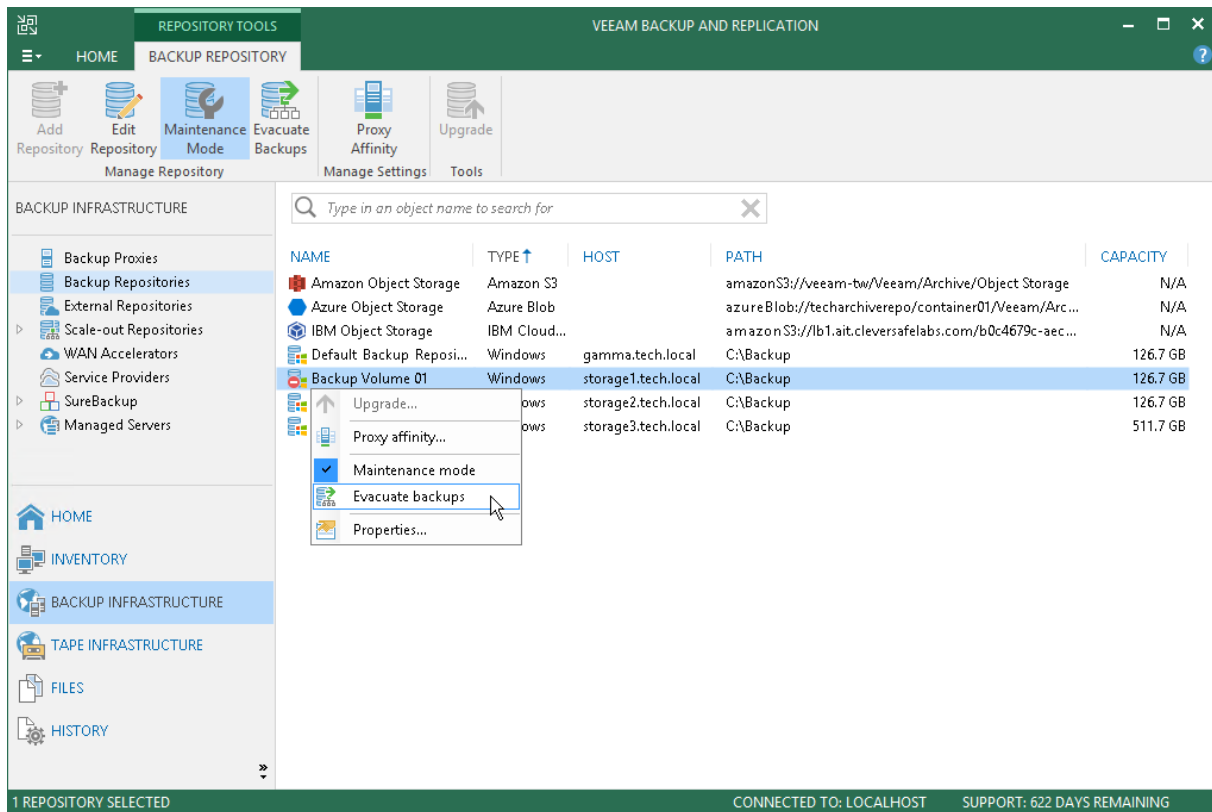
You must put the extent in Maintenance mode before you evacuate backups from it. For more information, see [Switching to Maintenance Mode in the Veeam Help Center](#).

To evacuate backup files from the extent, follow these steps.

1. [Recommended] Stop and disable jobs that are targeted at the extent from which you plan to evacuate backups.
2. Open the Backup Infrastructure view.
3. In the inventory pane under Scale-out Repositories, select a scale-out backup repository.
4. In the working area, select the extent and click Evacuate Backups on the ribbon, or right-click the extent and select Evacuate Backups.
5. If you have disabled jobs, reenable them.

After you evacuate backups, you can remove the extent from the scale-out backup repository. For more information on this process, see [Removing Extents from Scale-Out Repositories in the Veeam Help Center](#) (Figure 14)

Figure 14) Evacuating backups.



## 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 as Veeam Cloud Tier  
<https://www.netapp.com/us/media/sb-3982.pdf>
- NetApp Product Documentation  
<https://docs.netapp.com>

## Version History

Version	Date	Document Version History
Version 1.0	February 5, 2019	Initial draft for the document
Version 1.0.1	February 25, 2019	Edits for comments from Steven Pruchniewski
Version 1.0.2	March 27, 2019	Submitted to Corp Editorial

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.

### **Copyright Information**

Copyright © 2019 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.

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferrable, non-sublicensable, 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).

### **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.