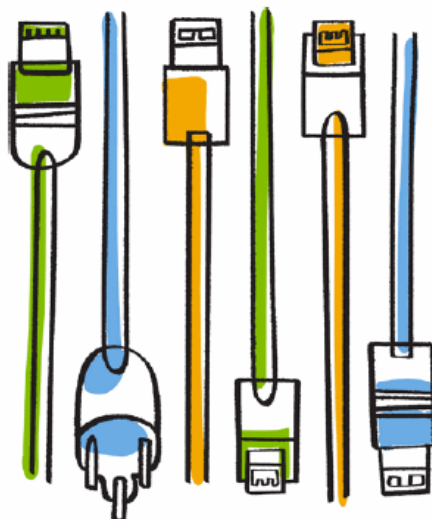




NetApp® AltaVault® Cloud Integrated Storage 4.2

Deployment Guide



NetApp, Inc.
495 East Java Drive
Sunnyvale, CA 94089
U.S.

Telephone: +1 (408) 822-6000
Fax: +1 (408) 822-4501
Support telephone: +1(888) 463-8277
Web: www.netapp.com
Feedback: doccomments@netapp.com

Part number: 215-11001_A0
April 2016

Contents

Contents	3
Chapter 1 - Deploying AltaVault appliance.....	5
What is AltaVault appliance?	5
Deployment types	7
Deployment guidelines	7
Information you must provide for deployments	9
Deployment steps	9
Obtaining documentation and release notes	10
Chapter 2 - Disaster recovery.....	11
Disaster recovery preparations	11
Exporting the configuration file	12
Disaster recovery testing	12
Suspending replication at the production site.....	12
Enabling AltaVault for a disaster recovery test	13
Data restoration for disaster recovery testing	14
Performing post-DR testing activities	14
Disaster recovery	15
Enabling AltaVault for disaster recovery	15
Data restoration for disaster recovery.....	15
Chapter 3 - Data prepopulation	17
Prepopulating data	17
Using prepopulation with Amazon Glacier cloud storage.....	20
Automatic prepopulation	20
Chapter 4 - Migrating data to a new cloud	21
Data migration overview	21
Setting up cloud credentials and performing the migration	22
Chapter 5 - Peer monitoring configuration	23
Peer monitoring benefits.....	23
Configuring appliance monitoring.....	23

Configuring REST API access	24
Specifying the API access code	24
Chapter 6 - Best practices for Amazon Glacier	25
Amazon Glacier overview	25
When to select S3 versus Glacier for cloud storage	26
Optimizing data movement to and from Amazon Glacier.....	26
Protecting data to Amazon Glacier.....	26
Recovering data from Amazon Glacier	27
AltaVault appliance best practices for EMC NetWorker for Amazon Glacier	27
AltaVault appliance best practices for IBM Spectrum Protect for Amazon Glacier	28
AltaVault appliance best practices for Veritas NetBackup for Amazon Glacier	30
AltaVault appliance best practices for Veritas Backup Exec for Amazon Glacier	32
AltaVault appliance best practices for Veeam backup and replication for Amazon Glacier	32
Chapter 7 - Implementing Veritas OST with AltaVault.....	35
Veritas OpenStorage (OST) overview	35
OST implemented features with AltaVault.....	35
Key benefits.....	37
Network planning	37
Recommended Deployments	39
AltaVault OST Deployment Actions	39
Summary of AltaVault OST best practices	49
Copyright information.....	51
Trademark information.....	53
How to send your comments.....	55
Index	57

CHAPTER 1 Deploying AltaVault appliance

This chapter provides an overview of AltaVault cloud integrated storage and deployment options. It includes the following sections:

- [“What is AltaVault appliance?” on page 5](#)
- [“Deployment types” on page 7](#)
- [“Deployment guidelines” on page 7](#)
- [“Deployment steps” on page 9](#)
- [“Obtaining documentation and release notes” on page 10](#)

What is AltaVault appliance?

AltaVault appliance is a disk-to-disk data backup and archive storage optimization system with unique cloud storage integration. AltaVault integrates seamlessly with your existing backup and archive technologies and cloud storage provider application program interfaces (APIs) to provide rapid replication of data to the cloud for off site storage and rapid retrieval.

AltaVault is a replacement for tape, virtual tape library (VTL), and disk-to-disk technology. AltaVault becomes the backup target for the enterprise. Rather than writing to tape, disk-to-disk, or VTL, a backup server writes its backups to AltaVault.

AltaVault is an inexpensive solution to storing large numbers of backups, without the cost and maintenance of a secondary data center. It is like having a tape library, a vaulting system, a large number of backups or archives in an off site storage facility, and a secondary data center in one appliance.

The application servers, email servers, and file servers connect to the backup server, typically through their backup agents. AltaVault easily integrates into your existing backup infrastructure. You can use your existing backup software, such as Veritas NetBackup, Veritas Backup Exec, or IBM Spectrum Protect. AltaVault acts as a storage target for your existing infrastructure. AltaVault appears to the backup server as a shared disk, using SMB (Server Message Block), OST (OpenStorage), or Network File System (NFS) protocols. AltaVault supports SMB, OST, and NFS protocols.

When it is time for a backup, the backup server contacts the backup client. Next, the backup server contacts the backup media (in this case, AltaVault) and starts writing an image of the clients or objects it is backing up.

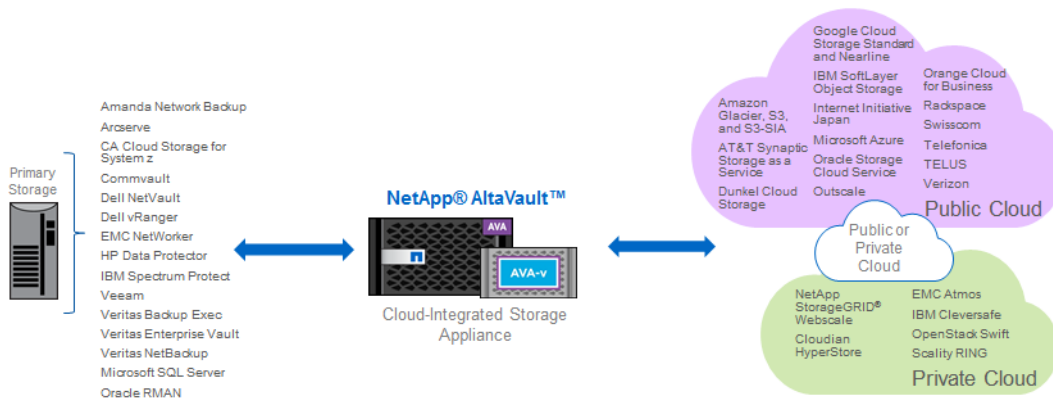
When you back up to AltaVault, it performs inline (real-time) deduplication of the backup data and replicates data into the cloud. AltaVault uses the local disk to store enough data for recovery of most recent backups. Such a mechanism provides LAN performance for the most likely restores. This deduplication process uses variable segment length inline deduplication plus compression, which is superior to other techniques such as fixed block. AltaVault deduplication level typically ranges between 10 and 30x. Deduplication performance depends on the incoming data type so turn off encryption and compression in the backup applications. Use the native encryption and deduplication in AltaVault to get higher data reduction rates than other typical software products.

AltaVault writes a copy of the data into the cloud storage provider. After AltaVault fills the capacity of its cache, it removes the least recently used data and replaces it with new incoming data. This process is called *eviction*. Evicted data can be recalled from the cloud transparently without user interaction in typical configurations, with most clouds. Amazon Glacier, which is also supported, changes the workflow to be less transparent.

AltaVault also optimizes restores from the cloud because it recalls only deduplicated data (which is not in the local cache) from the cloud. So if the customer is getting 10x deduplication, for example, and he or she needs to restore 10 TB of data, AltaVault needs only about 1 TB to restore. Over a 100-Mb line, this results in a time saving of days.

Data moves from the backup client to the backup server, to AltaVault, and then to the cloud. When you restore data, data moves from the cache in AltaVault, in which it is expanded to its original size to the backup server and to the backup client. If the data is not local, it moves from the cloud to AltaVault, to the backup server, and to the backup client.

The figure below shows the backup applications and cloud providers that AltaVault Cloud Integrated Storage supports.



Deployment types

AltaVault appliance is a disk-to-disk data storage optimization system with unique cloud storage integration.

There are three types of AltaVault deployments:

- Physical hardware appliance: AltaVault appliance is available in the AVA400 and AVA800 models.
- Virtual appliance: AltaVault virtual appliance is a virtual machine hosted package.
You can use VMware ESXi servers, Veritas KVM, or Microsoft® Hyper-V, to create a virtual machine (VM) and install the AltaVault-v software on the VM.
AltaVault-v is available in the following models for both, VMware or Microsoft Hyper-V:
 - AVA-v2
 - AVA-v8
 - AVA-v16
 - AVA-v32
- Cloud-based virtual appliance:
Amazon Machine Images (AMI) for the following models:
 - AVA-c4: The AVA-c4 is designed and optimized for deployment only on DS3 instances.
 - AVA-c8
 - AVA-c16Microsoft Azure model AVA-c4.

Note: Downgrades are not supported.

Deployment guidelines

Use the following guidelines to deploy a physical AltaVault appliance:

- AltaVault is supported with the backup applications and cloud storage providers identified by the IMT (interoperability matrix tool)
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.

- Use the following table to make a comparison of using AltaVault in backup versus cold storage mode:

Modes	Pros	Cons
Backup mode	• Allows access to the most recent backups on cache. • Allows global deduplication of all data received by AltaVault, leading to higher deduplication rates. • Provides the highest ingest performance achieved from backup applications. • Maximizes data movement efficiency of the WAN through deduplication of data. • Cache expansion capability via add on shelves allows for growth as needed by the business.	Limits the amount of cloud capacity managed (up to 960TB in the cloud).
Cold storage mode	Allows access to far greater cloud capacity (10PB of storage, based on 1.333 billion files of 100MB average file size). Provides expansive long term storage in just one head controller unit. Reduces computing requirements of AltaVault because limited deduplication and compression is performed.	Limits network and WAN performance, dependent on average, file size of objects sent to AltaVault. No expansion capability with shelves. Restores are always from the cloud provider.

Note: Cold storage mode applies to the AVA400 and virtual models.

- You can configure AltaVault folder shares to help describe a policy target.

For example, you can configure a backup application to direct critical system backups to point to a critical folder on one AltaVault data connection, while noncritical backups might be directed by a backup application to point to a non-critical folder on another AltaVault data connection. This method helps balance priorities of data over the network and organize data for recovery in case of a disaster.

- If possible, organize your backup policies so that generations of the same data arrive at the same AltaVault unit.

For example, if you are backing up a Windows server farm to multiple AltaVault appliances, operating system backups are likely to have the best deduplication rates when grouped together to the same AltaVault. File and application server backups obtain better deduplication when grouped together, because similar data might be stored in each location.

- If you are choosing to move from one provider to another, you can use the cloud agility feature.

Using a few CLI commands, you designate the new cloud bucket and data is systematically copied from the old provider to the new one. For more information, see [Chapter 4, “Data migration overview.”](#)

- AltaVault exports its configuration to a file called altavault_config_(HOSTNAME)_(DATETIME).tgz.

NetApp recommends that you store the configuration file in different physical locations. The configuration file contains information about the configuration, including the encryption key. Alternatively, you can just export the encryption key alone.

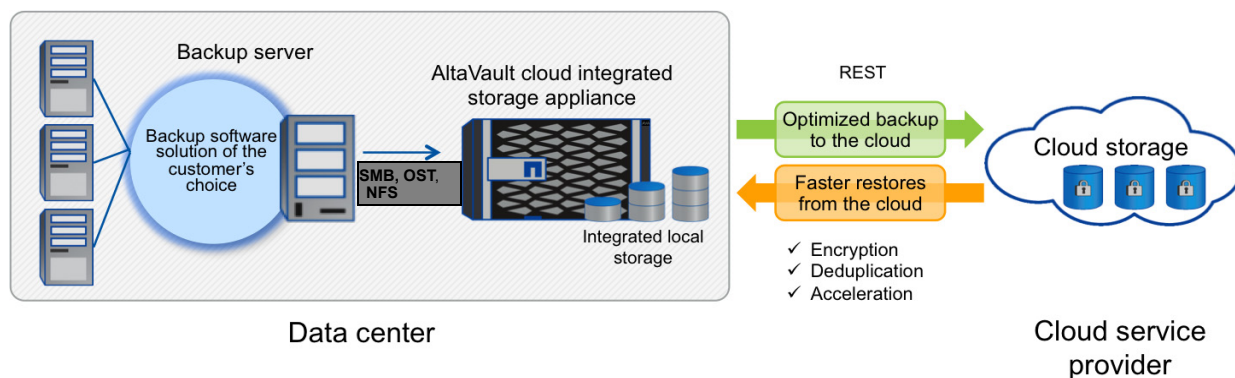
Note: To access the encrypted data, you need an encryption key. If you lose the encryption key, AltaVault cannot reconstitute the encrypted data.

- You can deploy each AltaVault to only one cloud storage provider at a time.

If an AltaVault must back up to a different cloud storage provider than the one configured, you must clear the AltaVault cache before reconfiguring the new cloud storage provider credentials. All existing data associated with the previous cloud storage provider remains, and you can recover it using AltaVault-v if necessary.

Information you must provide for deployments

The image below provides deployment information about the backup application, network, AltaVault, and Cloud Storage Provider.



For a list of information required for deployments, see the *Configuration Worksheet* in *Appendix A* of the *NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Physical Appliances*.

Deployment steps

Use the following table to guide your AltaVault deployment:

Deployment steps	Reference
Installation and configuration	
AltaVault physical appliances	<i>NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Physical Appliances:</i>
Site preparation	• Chapter 1 - Appliance setup preparation
Installing AltaVault	• Chapter 2- Installing the controller chassis and shelves
Connecting the Network	• Chapter 3 - Connecting the system
AltaVault virtual appliances	<i>NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Virtual Appliances:</i>
	• Chapter 1 - Installing and Configuring AltaVault on Microsoft Hyper-V • Chapter 2 - Installing and Configuring AltaVault Virtual Appliance on VMware ESXi • Chapter 3 - Installing and Configuring AltaVault Virtual Appliance on Linux KVM

Deployment steps	Reference
AltaVault cloud-based appliances	<i>NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Cloud Appliances:</i> Chapter 2 - Deploying an Amazon Machine Image Chapter 4 - Deploying a Microsoft Azure virtual machine
Running the appliance GUI Wizard	<i>NetApp AltaVault Cloud Integrated Storage User's Guide:</i> Chapter 2 Using the AltaVault configuration wizard
Configuring the data interface	<i>NetApp AltaVault Cloud Integrated Storage User's Guide:</i> Chapter 4 section - Modifying data interfaces
Configuring SMB, OST, or NFS shares	<i>NetApp AltaVault Cloud Integrated Storage User's Guide, Chapter 3:</i> • Configuring SMB • Configuring NFS • Configuring OST
Configuring email settings	<i>NetApp AltaVault Cloud Integrated Storage User's Guide:</i> Chapter 6 section - Configuring email settings
Exporting configurations	<i>NetApp AltaVault Cloud Integrated Storage User's Guide:</i> Chapter 2 section - Using the export configuration wizard option
Working with S3 Bucket Policies (Amazon only)	<i>NetApp AltaVault Cloud Integrated Storage User's Guide:</i> Appendix B - Amazon AWS IAM and S3 bucket policies
Configuring security settings Key Management Interoperability Protocol (KMIP)	<i>NetApp AltaVault Cloud Integrated Storage User's Guide:</i> Chapter 7 section - Configuring KMIP

Obtaining documentation and release notes

To obtain the most current versions of all NetApp documentation, including Release Notes, perform a search from the NetApp Support site at <https://mysupport.netapp.com>.

The AltaVault documentation set includes the following:

- *NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Physical Appliances*
- [*SAS Disk Shelves Installation and Service Guide for DS4243, DS2246, DS4486, and DS4246*](#)
- *NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Virtual Appliances*
- *NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Cloud Appliances*
- *NetApp AltaVault Cloud Integrated Storage Deployment Guide*
- *NetApp AltaVault Cloud Integrated Storage User's Guide*
- *NetApp AltaVault Cloud Integrated Storage Command-Line Interface Reference Guide*

CHAPTER 2 Disaster recovery

Disaster recovery is the process of recovering the technology infrastructure critical to an organization after a natural or man-made disaster. AltaVault supports disaster recovery by enabling you to retrieve your data in case of a failure.

This chapter includes the following sections:

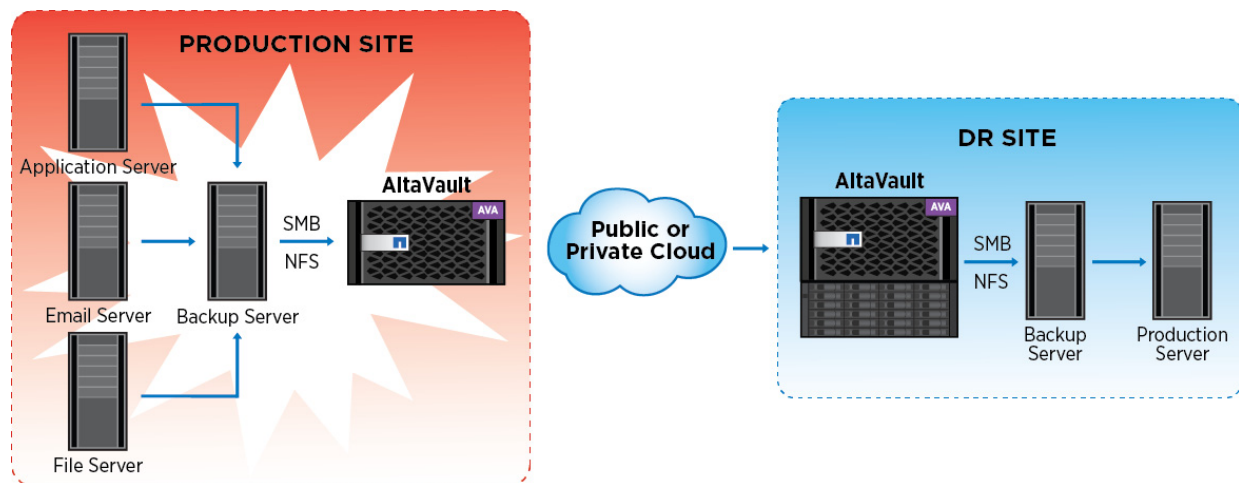
- “Disaster recovery preparations” on page 11
- “Disaster recovery testing” on page 12
- “Disaster recovery” on page 15

Disaster recovery preparations

You can enable AltaVault at the disaster recovery site to access backups that originated from an AltaVault at the affected data center. Depending on the data size, you can also use AltaVault-v at the recovery site.

Note: You do not need a license to restore data in read-only mode in AltaVault. You can download AltaVault-v for free from the NetApp Support site at <https://mysupport.netapp.com> and use it to recover your data.

For example, consider a data center with AltaVault located at the Production Site (site A), as shown in the figure below.



The backup site is the DR (Site B), located in a different physical location (such as different city, country, or continent). If there is a disaster at Site A, the data still resides in the cloud. Site B contains a passive AltaVault that is not powered on.

You can also use AltaVault-v at Site B, depending on the size of the data that you need to restore. AltaVault-v can store data up to 32 TB. NetApp recommends that you use an appliance in the disaster recovery site (Site B) that has the same or greater local storage capacity as the affected AltaVault (in Site A). If the appliances at the two sites do not match, you can still initiate the recovery process; however, it recovers only as much data as the size of the storage on AltaVault at the disaster recovery site. If the recovery process attempts to bring back more data than the disaster recovery AltaVault can handle, then the recovery process might fail.

For details about AltaVault sizes, see the *NetApp AltaVault Cloud Integrated Storage Installation and Service Guide for Physical Appliances*.

Exporting the configuration file

To prepare for disaster recovery, export your current configuration file from AltaVault at Site A, `altavault_config_(HOSTNAME)_(DATETIME).tgz`, and store it in a safe place, such as with your business continuity plans. For details on exporting your configuration file, see the *NetApp AltaVault Cloud Integrated Storage User's Guide*.

Disaster recovery testing

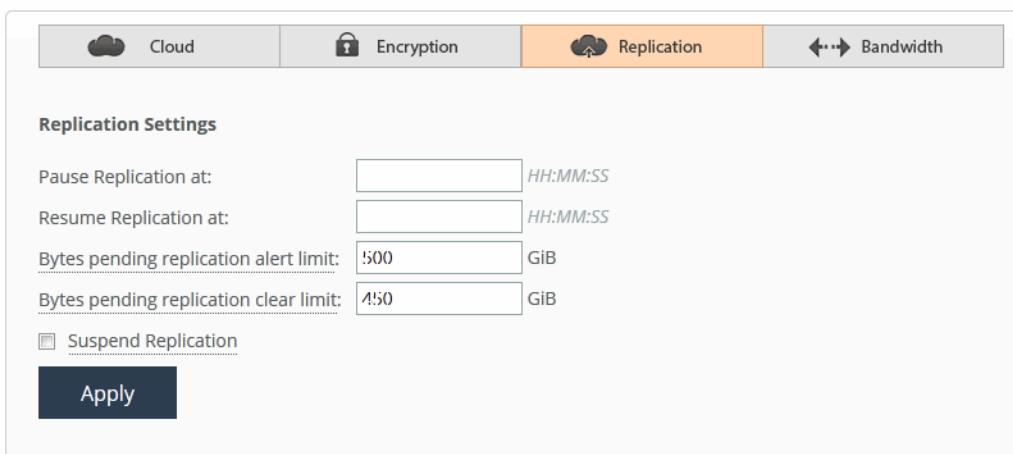
If you are restoring data for disaster recovery testing, you must first disable replication on AltaVault at site A and then restore your data at site B.

Suspending replication at the production site

To suspend replication

1. On the original production AltaVault appliance, suspend replication by selecting Storage > Cloud Settings.

Cloud Settings



Cloud	Encryption	Replication	Bandwidth
Replication Settings			
Pause Replication at:		HH:MM:SS	
Resume Replication at:		HH:MM:SS	
Bytes pending replication alert limit:	500	GiB	
Bytes pending replication clear limit:	450	GiB	
☐ Suspend Replication			
Apply			

2. Select the Replication tab.

3. To suspend replication, select the check box, Suspend Replication.
4. Click **Apply**.
5. Click **Save** to save your configuration.

Enabling AltaVault for a disaster recovery test

To recover your configuration to an AltaVault at site B

1. Log in to the appliance with the admin account using the serial console. If you are using a virtual AltaVault appliance, use the VM console to log in.
2. You can now configure the management IP address using the Configuration Wizard.
3. Use the IP address to connect to the appliance UI using a browser.
4. Log in to the appliance as `admin`. The default password is `password`.
5. Choose Settings > Setup Wizard.
6. Start the Import Configuration wizard.
7. Import to AltaVault in Site B, the configuration exported from the appliance in Site A:

Use the checkbox, Import Shared Data Only, if you are importing settings to a new AltaVault appliance. If you have configured the encryption key with an encryption key passphrase on the AltaVault in Site A, enter that passphrase into the encryption key passphrase field. After the import occurs, ensure that the new appliance in Site B uses the same cloud provider credentials, bucket name, and encryption key that Site A uses.

For more information on wizard related steps, see the *NetApp AltaVault Cloud Integrated Storage User's Guide, Chapter 2, Using the AltaVault configuration wizard*.
8. To test disaster recovery from a secondary site while the primary site is still alive, connect to the CLI by logging in as `admin` using SSH, and enter the following commands:

```
amnesiac > enable
amnesiac # configure terminal
amnesiac (config) # no service enable
amnesiac (config) # datastore format local
amnesiac (config) # replication dr-test enable
amnesiac (config) # service enable
amnesiac (config) # show service
```

The recovery process for both testing and an actual recovery can take anywhere from a few seconds to a few hours, depending on the backup(s) being restored. During the recovery process, the system communicates with the cloud provider and recovers all the namespace files that existed before the failure. The duration of this process depends on how many files you stored on AltaVault before the failure. Enter the **show service** command to determine the date and time until which the data store has been replicated.

After your service restarts, you can browse to your share and see your files. Because the recovery process downloads only the namespace and metadata, initial file access might be slow, because AltaVault downloads all of the data from the cloud.

Data restoration for disaster recovery testing

The process for restoring data from AltaVault for a disaster recovery test is similar to the process for restoring data from AltaVault under normal conditions. However, because none of the data is local after recovering the AltaVault configuration on a new appliance, AltaVault must recover the data back from cloud storage when requested by the backup application or user accessing the files.

Data is recovered serially as designated by read requests of the files from the SMB shares, OST shares, or NFS mounts. To speed up the recovery process, you can prepopulate data to the cache, which allows AltaVault to request recovery of data segments in parallel. For more information, see [“Prepopulating data” on page 17](#).

Performing post-DR testing activities

After completing a disaster recovery test, perform the following steps to clean up the disaster recovery environment, and to re-enable operations on the production AltaVault appliance.

To revert back to original state after the DR test is complete

1. When disaster recovery testing is complete, the AltaVault appliance used for DR testing is no longer needed. To revert the appliance back to an initial, non-configured state for future DR tests and to release control of the cloud bucket, enter the following commands:

```
amnesiac > enable
amnesiac # configure terminal
amnesiac (config) # no service enable
amnesiac (config) # datastore format local
amnesiac (config) # reset factory
```

Note: The AltaVault appliance shuts down as the final action in the reset factory command.

2. On the original production AltaVault appliance, re-enable replication by selecting Storage > Cloud Settings.

Cloud Settings ?

The screenshot shows the 'Cloud Settings' page with the 'Replication' tab active. The 'Replication Settings' section contains the following fields and values:

Field	Value	Unit
Pause Replication at:		HH:MM:SS
Resume Replication at:		HH:MM:SS
Bytes pending replication alert limit:	500	GiB
Bytes pending replication clear limit:	450	GiB

The 'Suspend Replication' checkbox is checked. An 'Apply' button is located at the bottom of the settings section.

3. Select the **Replication** tab.
4. To resume replication, clear the check box, **Suspend Replication**, and click **Apply**.

5. If pending data is awaiting to be replicated to the cloud, review the Reports > Back-End Throughput graph to confirm that data has resumed replication to cloud storage.

Disaster recovery

Disaster recovery actions must be performed as follows to enable a new AltaVault appliance for recovery of data backed up by the original AltaVault appliance lost in the disaster.

Enabling AltaVault for disaster recovery

To recover your configuration to an AltaVault at site B

1. Log in to the appliance using the admin account using the serial console. If you are using a virtual AltaVault appliance, use the VM console.
2. You can now configure the management IP address using the Configuration Wizard.
3. Use the IP address to connect to the appliance UI using a browser.
4. Log in to the appliance as `admin`. The default password is `password`.
5. Choose Settings > Setup Wizard.
6. Start the Import Configuration wizard.
7. Import to AltaVault in Site B, the configuration exported from the appliance in Site A:
 Use the checkbox, Import Shared Data Only, if you are importing settings to a new AltaVault appliance. If you have configured the encryption key with an encryption key passphrase on the AltaVault in Site A, enter that passphrase into the encryption key passphrase field. After the import occurs, ensure that the new appliance in Site B uses the same cloud provider credentials, bucket name, and encryption key that Site A uses.
 For more information on wizard related steps, see the *NetApp AltaVault Cloud Integrated Storage User's Guide, Chapter 2, Using the AltaVault configuration wizard*.
8. To perform disaster recovery after a lost primary site, connect to the CLI by logging in as `admin` using SSH, and enter the following CLI commands:

```
amnesiac > enable
amnesiac # configure terminal
amnesiac (config) # no service enable
amnesiac (config) # datastore format local
amnesiac (config) # replication recovery enable
amnesiac (config) # service enable
amnesiac (config) # show service
```

Data restoration for disaster recovery

The process for restoring data from AltaVault after a disaster is similar to the process for restoring data from AltaVault under normal conditions. However, because none of the data is local after recovering the AltaVault configuration on a new appliance, AltaVault must recover the data back from cloud storage when requested by the backup application or user accessing the files.

Data is recovered serially as designated by read requests of the files from the SMB shares, OST shares, or NFS mounts. To speed up the recovery process, you can prepopulate data to the cache, which allows AltaVault to request recovery of data segments in parallel. For more information, see [“Prepopulating data” on page 17](#).

CHAPTER 3 Data prepopulation

This chapter provides information for the data population process for the AltaVault appliance. It includes the following sections:

- [“Prepopulating data” on page 17](#)
- [“Using prepopulation with Amazon Glacier cloud storage” on page 20](#)
- [“Automatic prepopulation” on page 20](#)

Prepopulating data

You can retrieve the backup data from the cloud and populate AltaVault with it locally so that AltaVault has a local copy of the target data (which improves file access performance) either using the Management Console or using the command-line interface.

NetApp recommends that you use the AltaVault prepopulation process because it is a more efficient way of restoring data from the cloud than using the backup application directly. Although, it might seem longer (because this is a step that occurs before AltaVault restores data through the backup application), the prepopulation process improves restore times. It eliminates sporadic read operations for restore and uses sequential reads, thereby warming the AltaVault cache much more quickly.

For details on prepopulating data using the AV Management Console, see the *NetApp AltaVault Cloud Integrated Storage User's Guide*.

To prepopulate data using the command-line interface

1. Connect to the AltaVault command-line interface using SSH.
2. Enter the following command:

```
amnesiac (config) # datastore prepop { [num-days <number-of-days>] | [start-date *] [end-date *]  
| [pattern <pattern>] }
```

The following table shows the parameter options:

Parameter	Description
num-days <number-of-days>	Specify the number of last-modified days to start data retrieval (from the present date to the number of days you specify).
start-date <start-date>	Specify the date from which the data retrieval should start. The system prepopulates the files modified on or before this date.
end-date <end-date>	Specify the date on which the data retrieval should end. Stop prepopulating files on or after this date.
pattern <pattern>	Filters the data retrieved by the pattern you specify. The pattern specified contains a required internal share name created on AltaVault, one or more optional subfolder names from the external share name visible to the user, and finally a required regular expression describing the file or files to be prepopulated. The asterisk (*) symbol with the regular expression matches all characters.

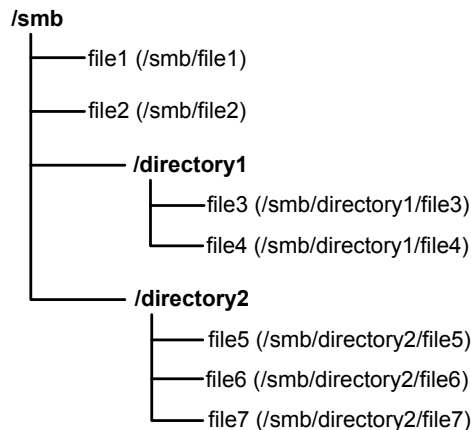
To view the current status of prepopulation, enter the following command:

```
amnesiac (config) # show datastore prepap
```

Example 1: Pattern-based data store prepopulation

This example explains pattern-based data store prepopulation. Consider the directory structure example shown in the figure below.

Directory structure



The following table shows different examples of the `datastore prepap` command for this directory structure:

Command	Description
<code>datastore prepap pattern smb/f*</code>	Populates only file1 and file2.
<code>datastore prepap pattern smb/*</code>	Populates all of the files (file1 through file7) with directory1 and directory2.

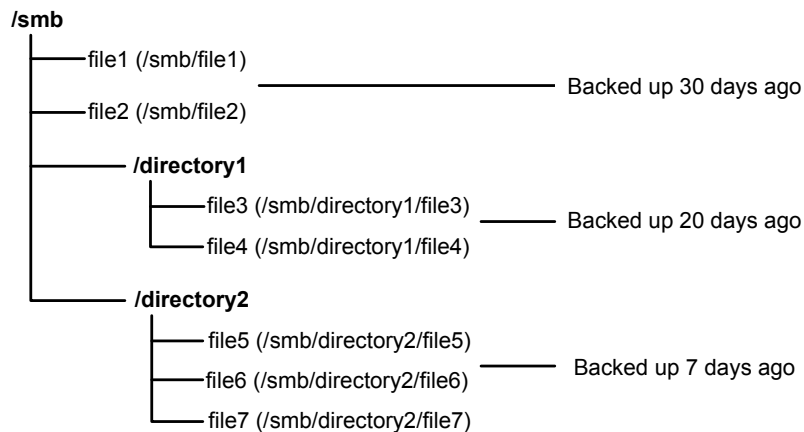
Command	Description
<code>datastore prepop pattern smb/directory1/*</code>	Populates only file3 and file4.
<code>datastore prepop pattern smb/directory1/file3 /smb/directory2/file7</code>	Populates only file3 and file7.

The datastore prepop command operates from the local pathname for each SMB share created.

Example 2: Time-based data store prepopulation

This example explains time-based data store prepopulation. Consider the directory structure example shown in the figure below.

Directory structure



To obtain the most recent files backed up, enter the following command on the AltaVault command-line interface:

```
datastore prepop num-days 7
```

This command fetches data that is seven days old from the cloud.

Example 3: Prepopulating from backups

In this example, assume that:

- All full backups are stored in a directory called fulls.
- All full backups for Host A are stored in a subdirectory called hostA.

To prepopulate all backups for Host A that occurred for a 24-hour duration starting on 2016-01-01 (YYYY-MM-DD), enter the following command:

```
amnesiac (config) # datastore prepop pattern fulls/hostA/*.img start-date 2016-01-01 end-date 2016-01-02 num-threads 64
```

To prepopulate all backups for Host A that occurred in the past 30 days (from the current time), enter the following command:

```
amnesiac (config) # datastore prepop num-days 30 pattern fulls/hostA/*.img num-threads 64
```

After this process finishes, you can initiate a restore process using the restore feature of the backup application. For details about how to restore your backups, refer to the relevant documentation for your backup application.

Using prepopulation with Amazon Glacier cloud storage

When you use Amazon Glacier as the cloud storage provider, it takes approximately four to five hours for data to be available for download, after you send the initial request to the cloud. Due to this delay, if data is not available on the local cache, it cannot be paged back from the cloud on demand.

In such cases, you must first restore the files to be read from the cloud to the local cache on AltaVault using either the prepopulation GUI or CLI commands. After the data is restored from the cloud, it can be read from the local cache.

Automatic prepopulation

You can also use settings in AltaVault to automatically trigger prepopulation of a file when you try to read the file and find that data must be restored from the cloud. For example, a backup application attempt to restore data from files stored in Glacier will fail but trigger an automatic prepopulation request. If the backup application can be configured to retry the restore operation after the 3-5 hour wait period for data, the retry will succeed without any additional user intervention to perform prepopulation. For assistance setting up automatic prepopulation, contact technical support.

Note: Enabling automatic prepopulation settings can trigger the prepopulation of entire files from the cloud upon read failures. This can result in restore charges related to recovering the data back from Glacier, as well as possible eviction of backup data in order to place the recovered data on cache. Enable automatic prepopulation only after careful consideration.

CHAPTER 4 Migrating data to a new cloud

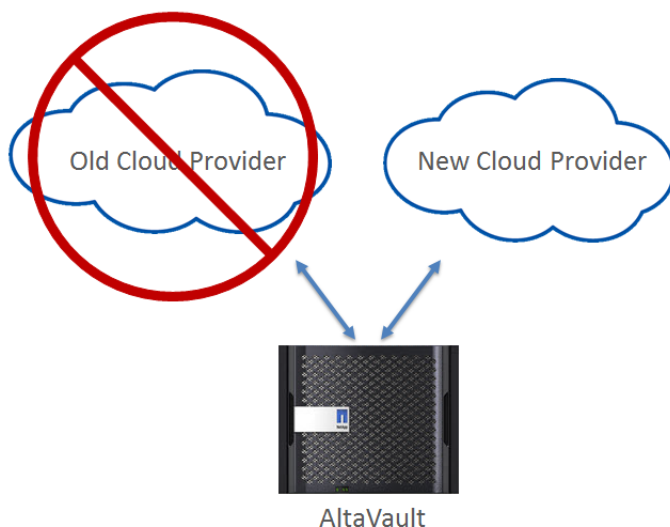
This chapter describes how to migrate data to a new cloud. It includes the following sections:

- [“Data migration overview” on page 21](#)
- [“Setting up cloud credentials and performing the migration” on page 22](#)

Data migration overview

AltaVault writes the deduplicated, compressed, and encrypted data to a private or public cloud storage provider. When business requirements dictate that data be migrated to a new cloud (for example, migrating from public cloud to public cloud, private cloud to public cloud, or public cloud to private cloud), the data that resides in that cloud storage must be relocated to a new cloud storage through data migration.

AltaVault implements the cloud migration feature called Cloud Agility to address this requirement. For an example of the three commands used for cloud migration, see [“Setting up cloud credentials and performing the migration” on page 22](#). The first two commands provide the cloud credentials and storage target of the new cloud storage service. The third command begins the migration. AltaVault copies all of the data from the first cloud storage to the second cloud storage, as seen in the figure below.



AltaVault acts as a data replicator during the migration. As the data flows from the existing cloud, through AltaVault, and then on to the new cloud, AltaVault does not reprocess the data. Therefore, no data is evicted from the AltaVault cache during this process; the data simply flows through the networking components of the appliance. AltaVault also continues to accept data from backup applications, so no interruption to backup schedules occurs during migration. Replication is suspended during data migration. If the backup application fills the AV cache, then AV will reject writes and fail backup operations until replication can be resumed following migration completion. Restores while migration is occurring may also fail if the needed deduplicated data segments from the old cloud have not yet been migrated.

When the data migration process completes, AltaVault automatically updates the cloud storage provider credentials to the new provider and resumes replication of any pending data that was queued during the migration process.

Setting up cloud credentials and performing the migration

Use the following commands to set up the cloud credentials and perform the migration. The command to set up the authentication type may be different depending on the provider you use.

To set cloud credentials:

```
replication migrate-to provider type <provider-name> bucket-name <bucket-name> hostname <host-name>
port <port-value>
replication migrate-to auth type <authentication-type> acc-key-id <access-key> secret-acc-key
<secret-key>
replication migrate-to proxy hostname <hostname> [port<port number>] [username <username>]
[password <password>]
```

To start cloud migration:

```
replication migrate-to enable [num-threads <value-1-to-128>]
replication migrate-to proxy enable
```

To monitor cloud migration:

```
show replication migrate-to estimate
show replication migrate-to proxy
```

Note: Cloud Agility is not supported from Amazon Glacier to any other cloud storage provider.

If migrating from Amazon S3 to Amazon Glacier, use the command:

```
replication s3-to-glacier
```

CHAPTER 5 Peer monitoring configuration

This chapter describes how to configure the peer monitoring feature in AltaVault. It includes the following sections:

- [“Peer monitoring benefits” on page 23](#)
- [“Configuring appliance monitoring” on page 23](#)
- [“Configuring REST API access” on page 24](#)
- [“Specifying the API access code” on page 24](#)

Peer monitoring benefits

AltaVault provides a capability called peer monitoring to allow multiple AltaVault appliances to be monitored under one centralized AltaVault appliance. For large deployments, this can significantly reduce the amount of administration required to review the utilization and status of the AltaVault appliances.

The benefits of monitoring peer appliances are:

- Centralizes management
- Improves storage visibility in large or multi-office configurations

Configuring appliance monitoring

Any AltaVault can monitor a peer AltaVault. After you configure REST API access and add the API access code for the peer appliance, the Appliance Monitoring report enables you to view the health status, disk space, and cloud service status of AltaVault.

The monitoring appliance probes the monitored peer appliances every 60 seconds by default.

To configure appliance monitoring

1. Enable REST-based access on the monitored appliance.
2. Generate the API access code on the monitored appliance.
3. Enter the API access code on the monitoring appliance.

Configuring REST API access

AltaVault uses REST APIs that you can access to set up peer appliance monitoring.

When you add an appliance to be monitored by AltaVault, you must generate an API access code to enable authenticated communication between the monitoring appliance and the monitored peer appliance.

For details on configuring REST API access, see the *NetApp AltaVault Cloud Integrated Storage User's Guide*, Chapter 7, *Configuring appliance monitoring*.

Specifying the API access code

After you generate the REST API access code on the monitored appliance, you must enter the code in the monitoring appliance to authenticate the monitored appliance.

For details on specifying the API access code in the monitoring appliance, see the *NetApp AltaVault Cloud Integrated Storage User's Guide*, Chapter 7, *Configuring appliance monitoring*.

CHAPTER 6 Best practices for Amazon Glacier

This chapter describes best practices for Amazon Glacier. It includes the following sections:

- [“Amazon Glacier overview” on page 25](#)
- [“When to select S3 versus Glacier for cloud storage” on page 26](#)
- [“AltaVault appliance best practices for EMC NetWorker for Amazon Glacier” on page 27](#)
- [“AltaVault appliance best practices for IBM Spectrum Protect for Amazon Glacier” on page 28](#)
- [“AltaVault appliance best practices for Veritas NetBackup for Amazon Glacier” on page 30](#)
- [“AltaVault appliance best practices for Veritas Backup Exec for Amazon Glacier” on page 32](#)
- [“AltaVault appliance best practices for Veeam backup and replication for Amazon Glacier” on page 32](#)

Amazon Glacier overview

AltaVault appliances can use Amazon Glacier as a cloud storage target to significantly reduce costs of storing cold or infrequently accessed data for long periods of time. Typical use cases for Amazon Glacier include storing completed project files, archiving audio, video, or medical records, and storing archived email or other documents that need to be retained for legal hold, security, or audit compliance.

AltaVault uses Amazon Glacier storage through the Amazon S3 to Glacier API. Data processed by AltaVault travels to Amazon S3 first, where it is then migrated automatically to Amazon Glacier after a default period (typically less than 24 hours). Certain metadata information about the migrated data is always maintained on Amazon S3 and is not migrated to Amazon Glacier, but is a small fraction of the overall dataset size.

When AltaVault makes a data request from Amazon Glacier, a 3-5 hour wait period occurs while the data is retrieved. After this period completes, the data objects are extracted from Amazon Glacier. The flow of the data segments work in reverse; for example, a copy of the data is placed back on Amazon S3 from Amazon Glacier, and then restored onto the AltaVault cache immediately after it arrives on Amazon S3. The data placed on Amazon S3 is held temporarily, and is automatically cleaned up after a 24 hour period.

It costs more to restore data from Glacier than S3 because of its two stage cost structure, which entails a retrieval fee and a data transfer fee. AltaVault takes advantage of the [complimentary 5% retrieval](#) amount that the Glacier service offers users each month, meaning that the additional data segments can be recovered at minimal cost. This design means that by default, AltaVault performs flow control of data back from Amazon Glacier at a potentially slower rate than what the WAN connection allows. In the event of large restores or restores that must flow at a faster rate than the free rate, this flow control can be increased or disabled on AltaVault to allow faster retrieve of data. The trade off is [higher retrieval costs](#).

When to select S3 versus Glacier for cloud storage

AltaVault appliances can only store data in one cloud storage target, so users must select S3 or Amazon Glacier when deploying an AltaVault appliance. As a result, a typical evaluation of a customer environment includes a discussion about whether S3 or Glacier is more appropriate as the cloud storage target for AltaVault appliance data. The bullet points below represent rough guidelines for deciding on the appropriate Amazon storage service for you to use.

Select S3 if you meet the following criteria:

- You perform frequent recoveries of data, of which some data may be from older, less frequently accessed backups or archives
- Your recovery time objective (RTO) window is small and cannot be satisfied within the Glacier retrieval window
- Your data retention rates are not long enough to take advantage of the cost benefit that Glacier provides

Select Glacier if you meet the following criteria:

- You perform infrequent recoveries of data, and store data typically for long term compliance and audit objectives
- You do not require retrievals of data to be performed within a specific RTO window
- You require a larger cloud storage target for storing cold data at the lowest possible cost

Optimizing data movement to and from Amazon Glacier

AltaVault appliances can be configured in backup or cold storage mode for use with Amazon Glacier. As most use cases for Amazon Glacier are typically for long term storage of inactive data, AltaVault is generally configured in cold storage mode to maximize cloud storage. Be aware that when using AltaVault in cold storage mode, the appliance cache is used for holding a greater amount of deduplicated data in Amazon Glacier, which results in a lower amount of cache being available to hold data for immediate recovery needs. For more information on the two appliance modes, see the table in the section, [“Deployment guidelines” on page 7](#).

Protecting data to Amazon Glacier

Protecting backup and critical production servers using Glacier also adds additional considerations. In most data protection scenarios backup servers would be protected the same way as any other production server backup data, and stored in the same storage target but under a different set of retention requirements and typically a different target location. During disaster recovery (DR), backup and critical production server recoveries are the first processes that occur, and thus retrieval of the backup server backup is critical in the DR process. If you use a cloud storage target like S3, then this would typically not pose a problem in terms of the DR time frame. However, if you use Amazon Glacier, then this would most likely incur a heavy delay that most businesses would find unacceptable from an RTO perspective. The appropriate approach for handling backup and critical production server protection if using Amazon Glacier would be to have the backup and critical production server backups go to a separate share on AltaVault which is pinned. AltaVault pinned data is always held on local cache and is never evicted, allowing administrators to have immediate access to these server backups immediately and without any delays associated with accessing data stored in Amazon Glacier.

AltaVault can also be configured to delay the migration of data from Amazon S3 to Amazon Glacier. By default, data is typically migrated within a 24 hour window once it arrives on Amazon S3. Using the AltaVault CLI command, `replication migration-delay`, Amazon S3 can be instructed to maintain the data for a longer period of time (set in days), until the data is moved to Amazon Glacier. However, it is important to understand that while this can extend the retention on Amazon S3 for new data segments, iterations of the same data segments sent to AltaVault (such as during subsequent full backups) will not result in the migration delay being extended for that deduplicated data segment already residing on Amazon S3. It is likely that over time, only a subset of a backup will reside on Amazon S3 awaiting migration, but a majority of the data will reside in Amazon Glacier since it hasn't changed. Refer to the *NetApp AltaVault Cloud Integrated Storage Command-Line Interface Reference Guide* for more details.

Recovering data from Amazon Glacier

When recovering data from Amazon Glacier, NetApp recommends that all the data segments related to a restore or retrieve by the backup application be present on the AltaVault cache first. This avoids error or retry conditions by the backup or archive application while it waits on Amazon Glacier to send the data to the AltaVault cache. To recover the data segments associated with the files to the AltaVault cache, use the AltaVault prepopulation feature. This limits the time penalty for data retrieval to a single 3 to 5 hour waiting period. If several, separate retrieve requests are made for individual data segments (for example, when a backup application restores files sequentially, where each file is not in the AltaVault cache), this can slow down the overall retrieval process from Amazon Glacier and the overall recovery. For more information on data prepopulation, see [“Prepopulating data” on page 17](#).

The discovery of which files have data which must be retrieved from Amazon Glacier to the AltaVault cache varies by application. Recommendations are highlighted for specific applications in the remainder of this chapter.

AltaVault appliance best practices for EMC NetWorker for Amazon Glacier

The following process allows you to stage a restore operation, identifying the files required for the restore and migrate the files from Glacier to AltaVault.

To identify which file objects will be needed for a Networker restore

1. Stage a *saveset* restore operation, by issuing the following command from the Networker command line:

```
Networker User > Operation > Save Set Recover > Source Client (oak-cs.cb2k3r2.com)
Save Set Name: D:\ (number of versions: 2)
Version Date:
[x] 6/13/2013 11:06 AM 38765393 browsable
[ ] 6/13/2013 11:49 AM 55531031 browsable

Required Volumes: oak_cs.cb2k3r2.com.001.RO
```

2. Assume that you want the saveset from "6/13/2013 11:06 AM" and go to the Windows command prompt and run the following commands:

```
cd "c:\Program Files\Legato\nsr\bin"

mminfo -otc -v -q name=D:\
volume      type      client      date        time        size  ssid      fl  lvl  name
oak_cs.cb2k3r2.com.001  adv_file  oak-cs.cb2k3r2.com 6/13/2013 11:06:53 AM 38 GB 4273605193 cb manual D:\
oak_cs.cb2k3r2.com.001.RO  adv_file  oak-cs.cb2k3r2.com 6/13/2013 11:06:53 AM 38 GB 4273605193 cb manual D:\
oak_cs.cb2k3r2.com.001  adv_file  oak-cs.cb2k3r2.com 6/13/2013 11:49:49 AM 55 GB 4256830541 cb manual D:\
oak_cs.cb2k3r2.com.001.RO  adv_file  oak-cs.cb2k3r2.com 6/13/2013 11:49:49 AM 55 GB 4256830541 cb manual D:\
```

- Using the SSID value from the output of the above, find the name of the corresponding file on the disk:

```
mminfo -q "ssid=4273605193" -r "ssid(53)"
8d5688a2-00000006-feba0a49-51ba0a49-00030e00-a8d346b6
```

- Now that the filename needed for restore or retrieve has been identified, locate the file on the AltaVault appliance share or mount.

Linux:

```
find <AVAmntpointname> -name "8d5688a2-00000006-feba0a49-51ba0a49-0003"
< AVAmntpointname >/networker/66/76/notes/8d5688a2-00000006-feba0a49-51ba0a49-00030e00-
a8d346b6
< AVAmntpointname >/networker/66/76/8d5688a2-00000006-feba0a49-51ba0a49-00030e00-a8d346b6
```

Windows:

Map AltaVault share path as Windows mapped network drive (Z: in this example)

```
Z:\>dir "8d5688a2-00000006-feba0a49-51ba0a49-00030e00-a8d346b6" /s
```

Volume in drive Z is LKWZE

Volume Serial Number is 009A-9A03

Directory of Z:\66\76

```
06/13/2013  11:19 AM      39,695,762,716  8d5688a2-00000006-feba0a49-51ba0a49-00030e00-a8d346b6
1 File(s)  39,695,762,716 bytes
```

Directory of Z:\66\76\notes

```
06/13/2013  11:19 AM                236  8d5688a2-00000006-feba0a49-51ba0a49-00030e00-a8d346b6
1 File(s)                236 bytes
```

Total Files Listed:

```
2 File(s)  39,695,762,952 bytes
```

```
0 Dir(s)   499,903,440,459,776 bytes free
```

Note: The mminfo query for the session id (SSID) should list out multiple files if the backup spans multiple media files. You need to list out all levels of the backup through mminfo. To get the SSID, you can also go to Network Administration > Media > Disk Volumes > Show Save Sets or Media > Save Sets and query them.

- Prepopulate the file or files identified in Step 4 using the Prepopulation GUI as described in [“Prepopulating data” on page 17](#), and wait until the files migrate from Glacier to the AltaVault cache.
- Initiate your restore from *Networker User* as you normally would to complete the recovery.

AltaVault appliance best practices for IBM Spectrum Protect for Amazon Glacier

The following process allows you to identify AltaVault volumes and initiate a Spectrum Protect client restore.

To identify what volumes from AltaVault required for a Spectrum Protect client restore

- On the Spectrum Protect server administrative command line issue the following SELECT statement to identify the object ID for the file you want to restore:

```
select * from backups where node_name='<tsmclientnodename>' and ll_name='<filenamestore>'
select * from backups where node_name='CLIENT1' and LL_NAME='MYFILE.TXT'
```

```

      NODE_NAME: CLIENT1
    FILESPACE_NAME: \\CLIENT1\s$
      FILESPACE_ID: 1
        STATE: ACTIVE_VERSION
        TYPE: FILE
        HL_NAME: \BACKUP\SET2\
        LL_NAME: MYFILE.TXT
      OBJECT_ID: 1062
    BACKUP_DATE: 2013-04-23 20:02:38.000000
  DEACTIVATE_DATE:
    OWNER:
    CLASS_NAME: DEFAULT

      NODE_NAME: CLIENT1
    FILESPACE_NAME: \\CLIENT1\s$
      FILESPACE_ID: 1
        STATE: ACTIVE_VERSION
        TYPE: FILE
        HL_NAME: \BACKUP\SET3\
        LL_NAME: MYFILE.TXT
      OBJECT_ID: 6786
    BACKUP_DATE: 2013-04-23 20:06:19.000000
  DEACTIVATE_DATE:
    OWNER:
    CLASS_NAME: DEFAULT

```

2. From the list of file versions above, identify the version needed (for this example, the version from 20:06:19 on 4/23/2013), and then issue the following command using the OBJECT_ID value from the output.

Note: Note that if the *bitfile* is part of a super-bitfile, rerun the below command against the super-bitfile OBJECT_ID below.

```

show bfo 6786
Bitfile Object: 6786
  Active
**Sub-bitfile 6786 is stored in the following aggregate(s)
  Super-bitfile: 6783, Offset: 2000, Length 694, Deduped: F

show bfo 6783
Bitfile Object: 6783
**Super-bitfile 6783 contains following aggregated bitfiles,
Bitfile Id, offset, length, active state or owner, link bfid
6783          0          671          Active
6784          671          663          Active
6785         1334          666          Active
6786         2000          694          Active
6787         2694         1406          Active
6788         4100          676          Active
6789         4776          679          Active
.....
**Sub-bitfile 6783 is stored in the following aggregate(s)
  Super-bitfile: 6783, Offset: 0, Length 671, Deduped: F

**Disk Bitfile Entry
  Bitfile Type: PRIMARY  Storage Format: 22
  Logical Size: 25880969  Physical Size: 25886720  Number of Segments: 1,
  Deleted: False
  Storage Pool ID: 4  Volume ID: 3  Volume Name: H:\TSMVOLS\AVAAVOL001.BFS

```

3. Prepopulate the volume identified in Step 2, using the Prepopulation GUI as described in [“Prepopulating data” on page 17](#), and wait until the files migrate from Glacier to the AltaVault cache.
4. Initiate your restore from the Spectrum Protect client as you normally would to complete the recovery.

To identify all the volumes related to a AltaVault based storage pool

1. Issue the following Spectrum Protect administrative SELECT command, using the appropriate storage pool name that points to AltaVault:

```
select volume_name from volumes where stgpool_name='<AVASTGPOOLNAME>'
```

```
select volume_name from volumes where stgpool_name='AVACOPYPOOL'
```

```
VOLUME_NAME: \\AltaVault-01\TSM\00000002.BFS
```

```
VOLUME_NAME: \\AltaVault-01\TSM\00000003.BFS
```

```
VOLUME_NAME: \\AltaVault-01\TSM\00000004.BFS
```

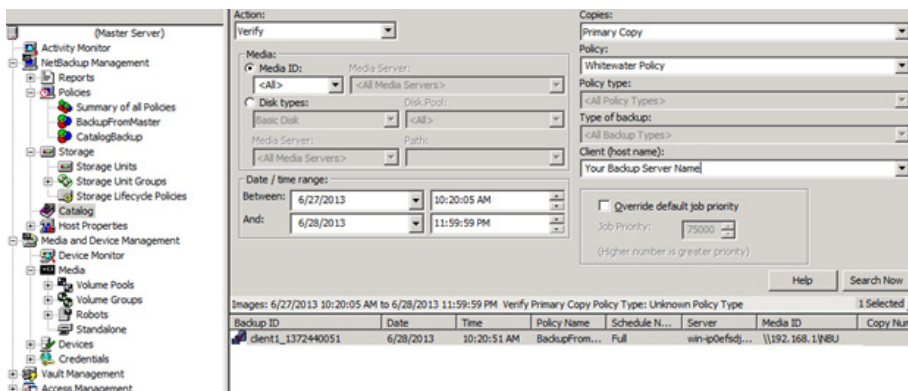
2. Prepopulate the volumes identified using the Prepopulation GUI as described in [“Prepopulating data” on page 17](#), and wait until the files migrated from Glacier to the AltaVault cache.
3. Perform primary Spectrum Protect storage pool recovery as appropriate if the storage pool above is a copy storage pool.
4. Initiate your restore from the Spectrum Protect client as you normally would to complete the recovery.

AltaVault appliance best practices for Veritas NetBackup for Amazon Glacier

The NetBackup Catalog maintains an inventory of the backups and can be used to identify which media volumes are required for restore.

To identify which media volumes are required for restore

1. Go to the Catalog menu item from the NetBackup GUI and search for the backup from which you want to restore data.
2. Use the filter criteria to select the policy that goes to AltaVault, the client you want to recover data from, and the date and time range.



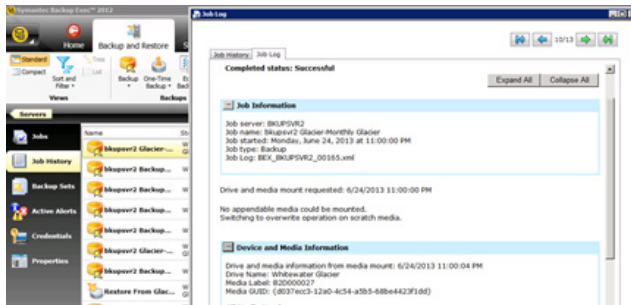
3. In the results field at the bottom of the page, identify the Backup_ID.
This corresponds to a portion of the file name of the volume created on AltaVault for the backup.
4. Prepopulate the identified volume identified in step 2, using the Prepopulation GUI as described in [“Prepopulating data” on page 17](#), and wait until the files migrate from Glacier to the AltaVault cache.
5. Initiate your restore from the NetBackup GUI as you normally would to complete the recovery.

AltaVault appliance best practices for Veritas Backup Exec for Amazon Glacier

The Backup Exec job activity page maintains inventory of the backups, which can be used to identify which media volumes are required for restore. Those volumes can then be located by referring to the Backup-To-Disk-Folder or Storage target and locating the volume from the corresponding AltaVault SMB share.

To identify which medium volumes are required for restore for Backup Exec 2012:

1. Select the Backup and Restore tab
2. Select the server from which you want to restore objects.
3. Select the Job History tab from the left side.
4. Select the backup job you want to restore from the list of backups that display.
5. In the Job Log page that displays, select the Job Log tab and find the media volumes used.



Note: Backup Exec users sending data to Glacier should run the following AltaVault CLI command to ensure that media recycling of volumes on AltaVault appliances occurs correctly with respect to Glacier storage: `megastore keep-bkf-local enable`.

6. Prepopulate the file or files identified in Step 5 using the Prepopulation GUI as described in [“Prepopulating data” on page 17](#), and wait until the files migrate from Glacier to the AltaVault cache.

AltaVault appliance best practices for Veeam backup and replication for Amazon Glacier

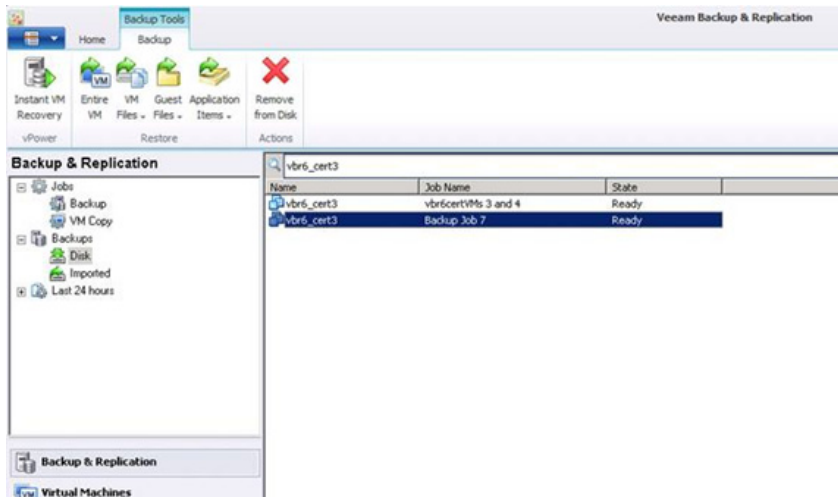
For Veeam Backup & Replication, folders are named after their respective job numbers in Veeam, as are the files.

Note: Note that all of the files and folders are time stamped automatically starting from Veeam 5.

To locate the backup required for prepopulation

1. Locate which backup job to which the VM belongs.

2. In this case, the example is VM is *vbr6_cert3*. In Veeam, you can locate the backup job by searching the backup database. In the figure below, you can see that *vbr6_cert3* has been backed up by a job named *Backup Job 7*.



3. After identifying the job name, prepopulate the most recent *.vbk file* plus all subsequent *.vib* files, and the job metadata file with the *.vbm* extension, using the Prepopulation GUI as described in “[Prepopulating data](#)” on [page 17](#), and wait until the files migrated from Glacier to the AltaVault cache.

Note: There is no need to prepopulate older backup chains. Veeam time stamps the backup job files to make it easy to identify.

The figure below shows an example of how to prepopulate the current Veeam backup files needed to restore the most recent version of the VM named `vbr6_cert3`.



4. Initiate your restore from the Veeam Backup and Replication GUI as you normally would to complete the recovery.

CHAPTER 7 **Implementing Veritas OST with AltaVault**

This chapter provides information for implementing OST for the AltaVault appliance. It includes the following sections:

- [“Veritas OpenStorage \(OST\) overview” on page 35](#)
- [“OST implemented features with AltaVault” on page 35](#)
- [“Key benefits” on page 37](#)
- [“Network planning” on page 37](#)
- [“Recommended Deployments” on page 39](#)
- [“Summary of AltaVault OST best practices” on page 49](#)

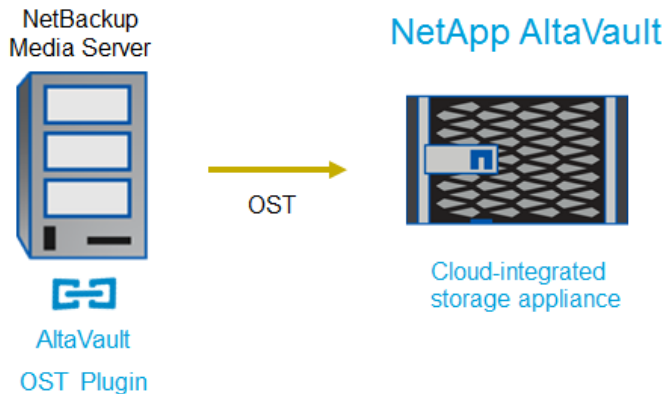
Veritas OpenStorage (OST) overview

Veritas OST is an API that delivers optimized backup integration between NetBackup and storage appliances such as NetApp AltaVault. Starting with AltaVault version 4.2, AltaVault can now leverage the OST framework to provide NetBackup the ability to efficiently stream backups to AltaVault, as well as manage lifecycle policy of copies of backups stored on AltaVault and in the cloud.

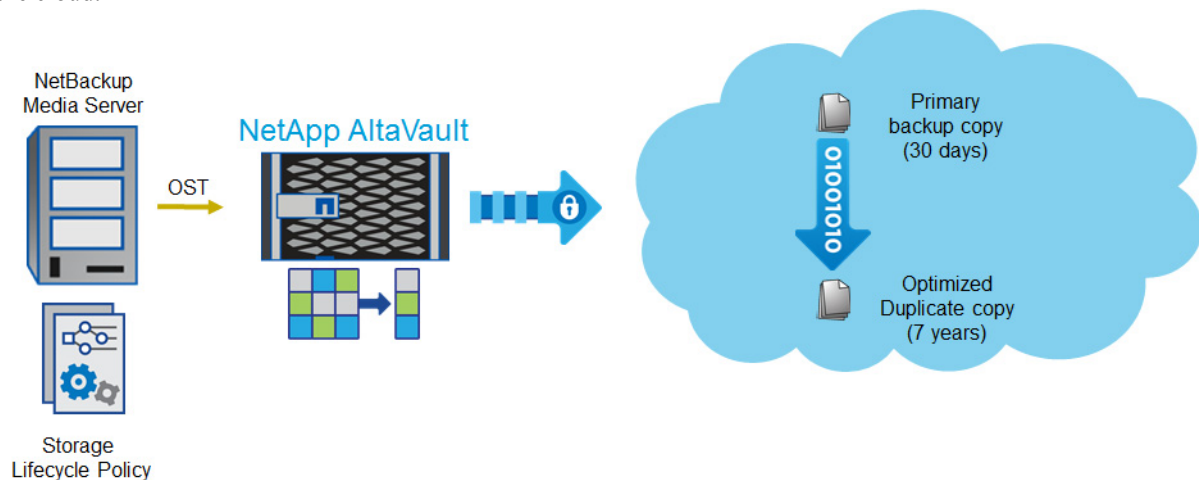
OST implemented features with AltaVault

AltaVault support for OST consists of two features:

- AltaVault supports the OST protocol, which is a communications protocol between NetBackup and AltaVault. The OST protocol not only allows AltaVault to receive backup data from NetBackup, but also allows NetBackup to have increased visibility and control of backups stored in AltaVault, as well as in cloud storage that AltaVault is connected to. To enable OST support, the AltaVault OST Plug-in is installed on NetBackup media servers which will communicate and deliver backup data to AltaVault. In addition, AltaVault version 4.2 or higher is required, as it contains the necessary OST protocol components to establish OST communication from NetBackup media servers that have the AltaVault OST Plug-in.



- AltaVault supports OST optimized duplication of backups triggered by NetBackup storage lifecycle policies (SLP). By using OST, NetBackup has visibility and control of backups stored on AltaVault and in the cloud. This enables NetBackup administrators to leverage NetBackup SLP to ensure that backup copies are appropriately maintained for short term and long term recovery. For example, an IT administrator may be required to create a month end duplicate copy of a backup from a current weekly backup policy, and store those monthly backups for 7 years. By using AltaVault with OST and SLP, optimized duplicates of those backups can now be requested by NetBackup to AltaVault, which will be dynamically created based on previous backup copies stored by AltaVault in cloud storage. A minimum amount of processing by AltaVault or NetBackup is required, because AltaVault and OST coordinate the duplication of the metadata between the original backup and duplicate copy created in the cloud.



Key benefits

- Availability of data where you need it. With OST, NetBackup administrators can now ensure that critical datasets that are needed for short term use are locally available on the AltaVault appliance cache. Optimized duplicate copies for long term recovery and compliance can be managed and driven via secondary retention separately in the cloud. AltaVault intelligently manages the retention of the two copies in accordance with SLP.
- Off-site replication validation. Prior to OST support, NetBackup administrators could not determine via the management console if backups to AltaVault had indeed been protected to off-site cloud storage. The addition of OST now allows NetBackup to have visibility into backups stored on AltaVault, and consequently when backups complete replication to cloud storage. This provides piece of mind, and simplified backup management.
- Enhanced optimization of resources end to end. AltaVault intelligently deduplicates, compresses and encrypts NetBackup data to cloud storage, reducing the costs of transmission and storage while simultaneously easing management of backups. With the addition of OST, NetBackup can now further extend optimization techniques to additional duplicate copies with minimal impact on NetBackup media servers, LAN and WAN networks, and cloud storage. AltaVault efficiently creates copies using information of backups already protected in cloud storage, generating minimal bandwidth and data consumption across all resources for these types of operations.

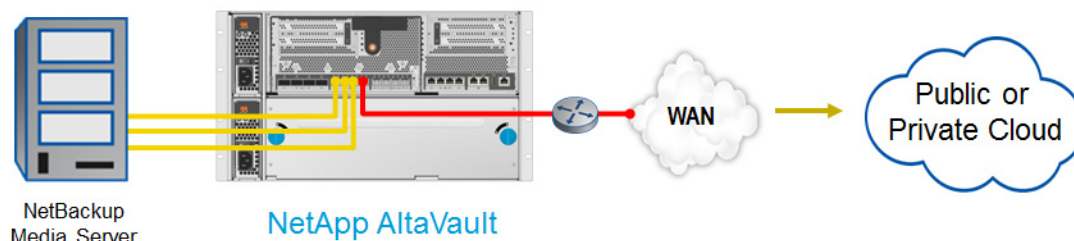
Network planning

Each AltaVault physical appliance provides four (4) 10GbE connections as well as four (4) 1GbE connections to use with backup infrastructure and to connect to cloud storage. Large environments will typically use multiple 10GbE connections for receiving backups from NetBackup, although it is acceptable to use a combination of 10GbE and 1GbE connections depending on the network infrastructure and locations of NetBackup media servers.

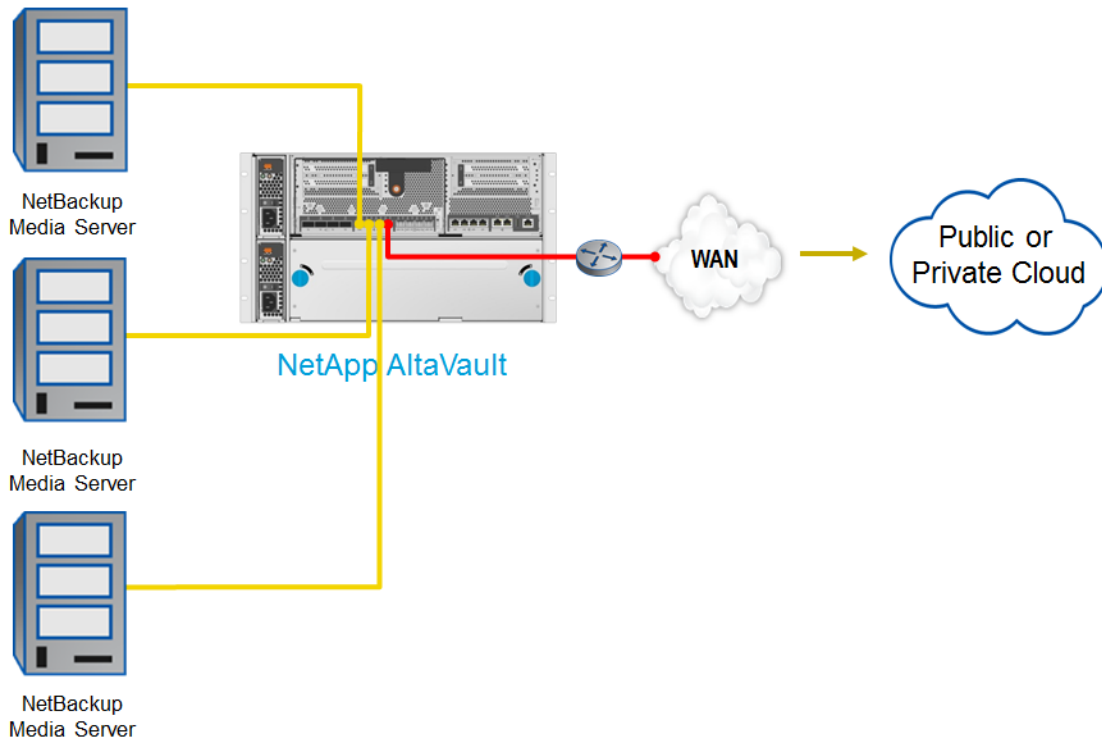
In addition to connecting AltaVault to the NetBackup infrastructure, AltaVault will also need to connect to a public or private cloud storage destination. Depending on your network and WAN resources, use one of the 10GbE or 1GbE interfaces. Many cloud providers offer mechanisms to deliver 10GbE, such as Amazon Direct Connect, which can help ensure data flow of replicated backups to cloud is not constricted by smaller WAN connections. Refer to the AltaVault Performance Tuning Guide for more details about AltaVault networking and configuration.

Example topologies:

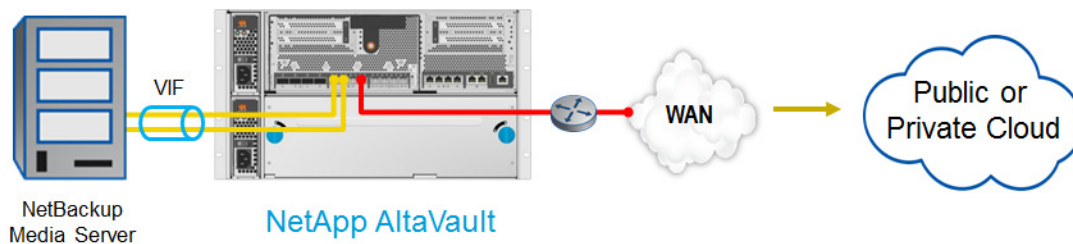
- Single media server to AltaVault using multiple 10GbE connections.



- Multiple media servers to AltaVault, each using a single 10GbE connection.



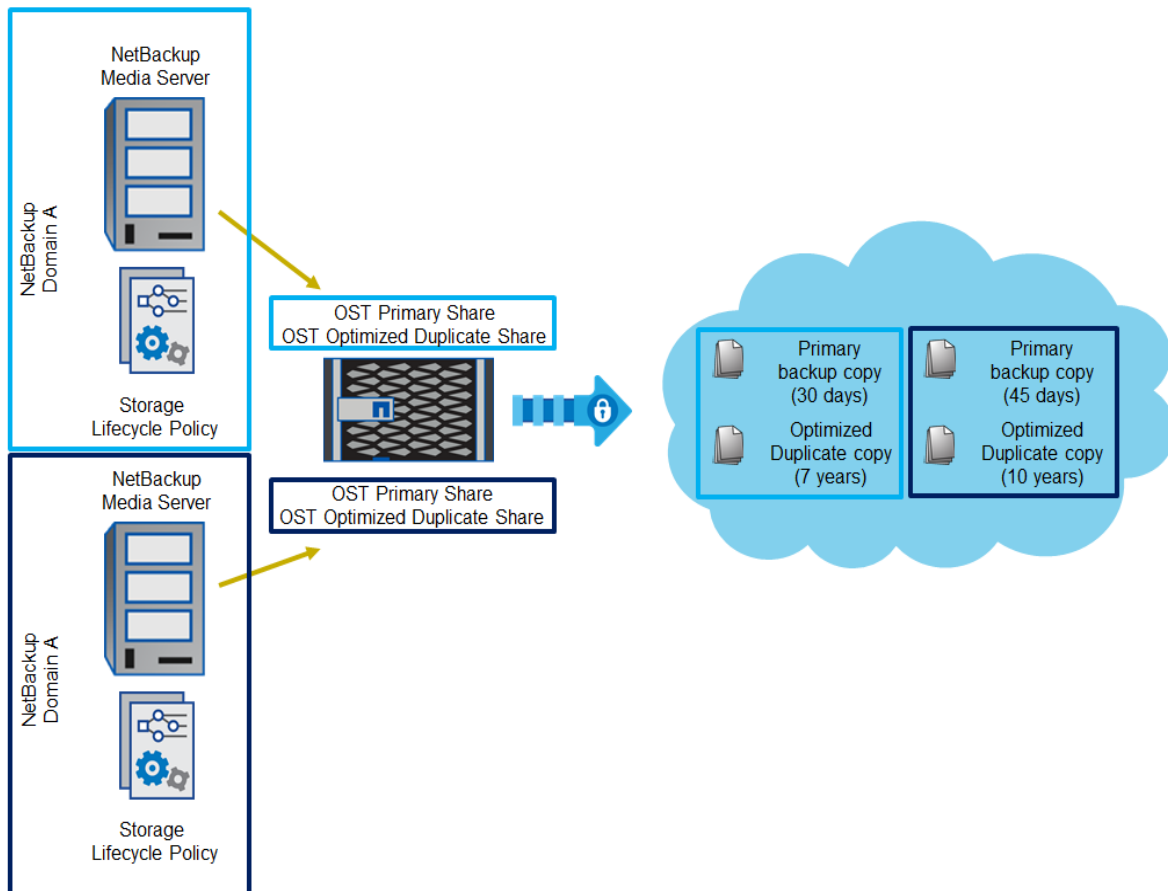
- Failover network connectivity/Load balancing via VIF.



Since optimized duplication operations with AltaVault are dependent on cloud storage, optimized duplication cannot currently be performed between AltaVault appliances at this time to achieve off-site replication and warm disaster recovery (DR). You can however configure a secondary AltaVault appliance in disaster recovery mode, to allow it to quickly be ready to take over effective operations from the primary AltaVault appliance should it become unavailable. This secondary AltaVault appliance can then take ownership of optimized duplicate backups in the cloud and with NetBackup at the secondary site, manage or recover them as needed. For more details about performing AltaVault DR, refer to Chapter 2—*Disaster Recovery*.

Recommended Deployments

Because a NetBackup master server can perform oversight and manage operations across the backup storage infrastructure, multiple NetBackup media servers using OST can coordinate optimized duplication tasks with an AltaVault appliance to deliver an enterprise scale data protection solution backed by cloud storage. In addition, multiple NetBackup master servers, along with their respective NetBackup media servers using OST, may also use one AltaVault appliance. It is recommended however that for manageability, separate OST shares created by AltaVault be provided to each NetBackup environment.



AltaVault OST Deployment Actions

1. Install the AltaVault OST Plug-in on NetBackup media server.
2. Configure the required OST shares on AltaVault.
3. Run the NetBackup Storage Server Wizard.
4. Create a Storage Lifecycle Policy.

Install AltaVault OST Plug-in

To install the AltaVault OST Plug-in on the NetBackup Media Server, refer to the installation instructions in Appendix D in the *NetApp AltaVault Cloud Integrated Storage User's Guide*.

Configure the Required OST shares on AltaVault

To perform OST operations, at least one regular OST share must be created on AltaVault. To perform optimized duplication operations, at least one regular OST share and one cloud OST share must be created on AltaVault. Each share must also have a local user associated with that share, which will be provided to NetBackup when creating the Storage Server and related disk pool and storage unit resources. Refer to the configuration instructions in Chapter 3 in the *NetApp AltaVault Cloud Integrated Storage User's Guide*.

Run the NetBackup Storage Server Wizard

The NetBackup Storage Server Wizard configures the necessary storage devices, including the storage server, disk pool, and storage unit, enabling OST communications to AltaVault. Since each share created on AltaVault is associated with one storage server, disk pool and storage unit, you will perform the following procedure once per share you wish to enable with NetBackup OST.

1. Start the NetBackup Storage Server Wizard. On the Add Storage Server page, fill in the following fields as shown below. Click **Next**.

Storage Server Configuration Wizard
Add Storage Server
Provide storage server details.

Select a media server that has the vendor's OpenStorage plug-in installed.
NetBackup uses this media server to determine the storage server capabilities.

Media server:

Storage server type:
NetAppAltaVault

Storage server name:

Enter storage server credentials

User name:

Password:

Confirm password:

< Back Next > Cancel Help

- Media Server
Specify the media server that has the AltaVault OST Plug-in installed.
- Storage server type
Manually type in the value **NetAppAltaVault**. The value does not appear in the drop down list.
- Storage server name

Enter in a name to call this storage server.

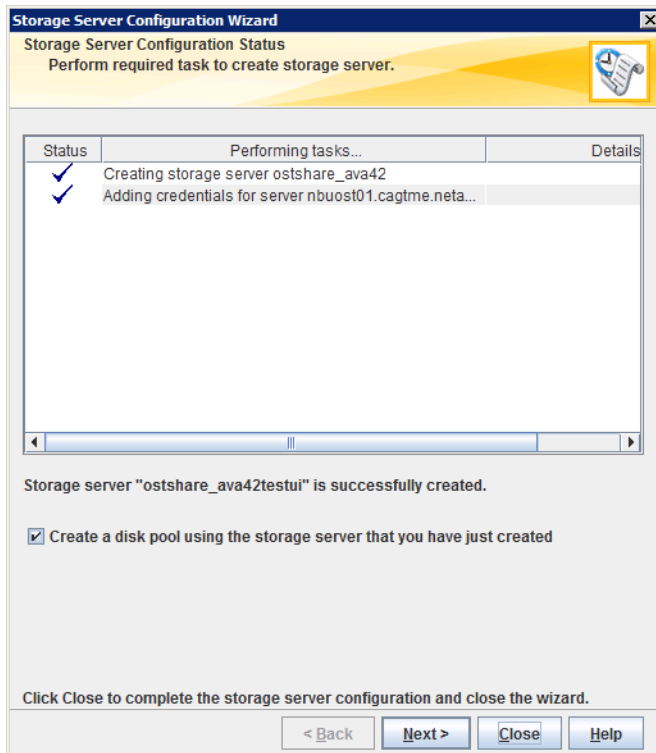
- User name

Provide the name of the OST user associated with the AltaVault OST share. Refer to the AltaVault OST page for the credentials.

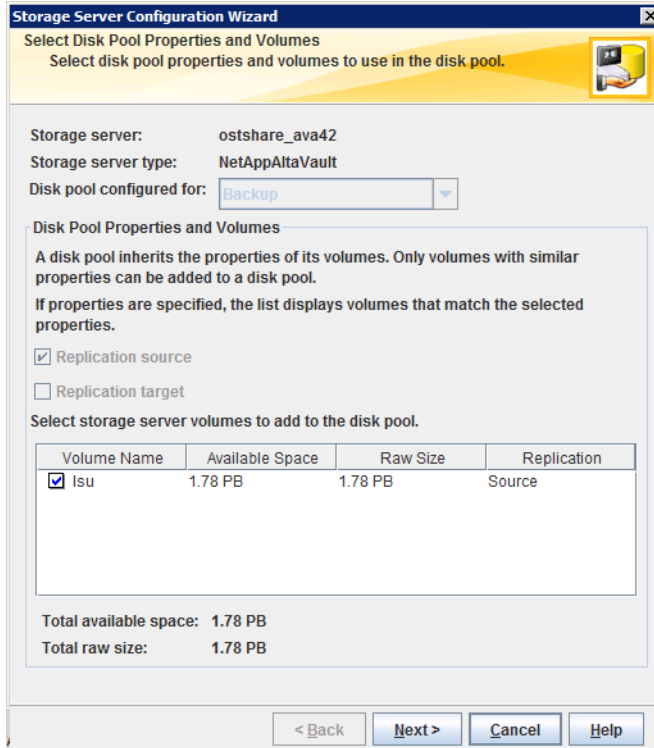
- Password

Provide the password of the OST user associated with the AltaVault OST share. Refer to the AltaVault OST page for the credentials.

2. Complete the storage server setup wizard. Ensure the **Create a disk pool using the storage server you have just created** checkbox is selected. Click **Next**,



3. In the Disk Pool and Volumes page, select the appropriate pool type, and select the volume to add to the disk pool by clicking the check box next to the logical name of the AltaVault volume in the table. Click **Next**.



Storage Server Configuration Wizard
Select Disk Pool Properties and Volumes
Select disk pool properties and volumes to use in the disk pool.

Storage server: ostshare_ava42
Storage server type: NetAppAltaVault
Disk pool configured for: Backup

Disk Pool Properties and Volumes
A disk pool inherits the properties of its volumes. Only volumes with similar properties can be added to a disk pool.
If properties are specified, the list displays volumes that match the selected properties.

☒ Replication source
☐ Replication target

Select storage server volumes to add to the disk pool.

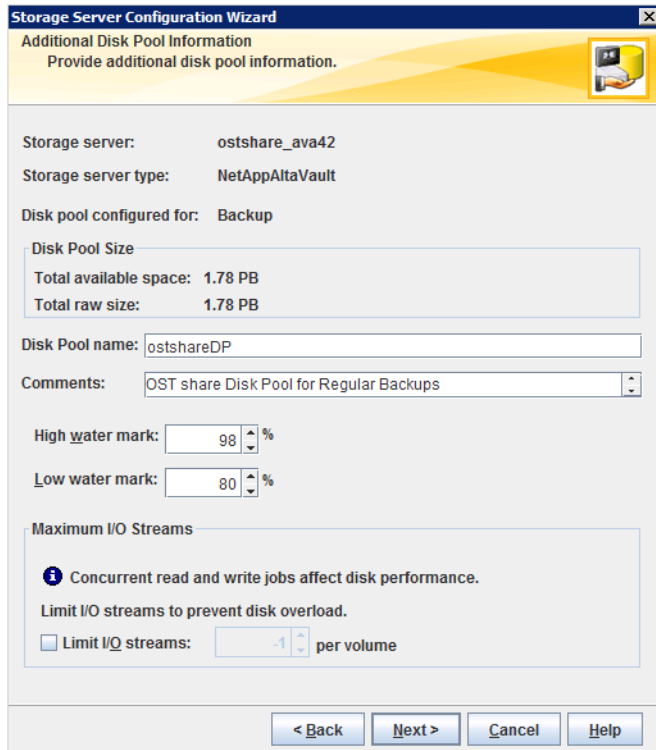
Volume Name	Available Space	Raw Size	Replication
☒ Isu	1.78 PB	1.78 PB	Source

Total available space: 1.78 PB
Total raw size: 1.78 PB

< Back Next > Cancel Help

- **Replication Source**
Use this type when configuring a disk pool for AltaVault regular OST shares.
- **Replication Target**
Use this type when configuring a disk pool for AltaVault cloud OST shares.

4. In the Additional Disk Pool Information page, enter a disk pool name. Do not select the Limit I/O streams checkbox. Click Next.



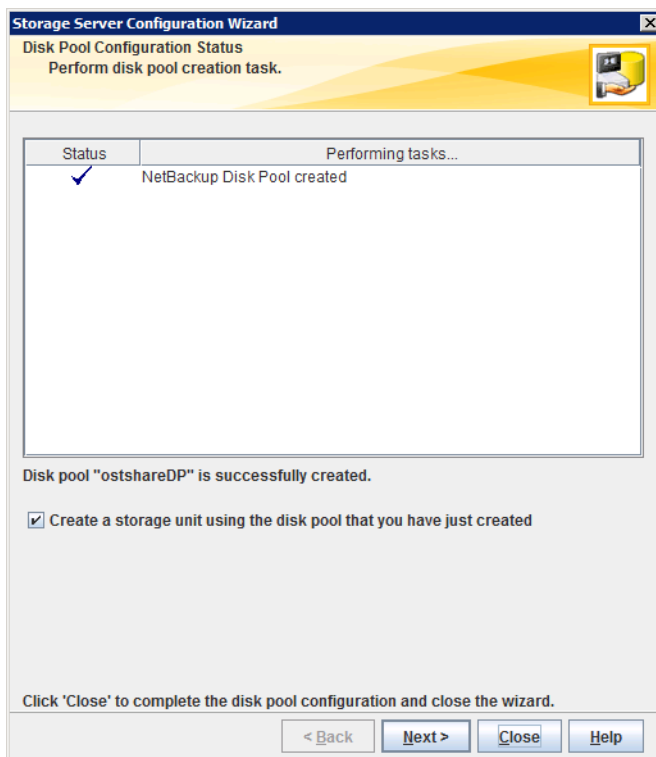
The screenshot shows the 'Storage Server Configuration Wizard' window, specifically the 'Additional Disk Pool Information' step. The window has a yellow header bar with the title 'Storage Server Configuration Wizard' and a sub-header 'Additional Disk Pool Information' with the instruction 'Provide additional disk pool information.' Below the header, the configuration details are as follows:

- Storage server: ostshare_ava42
- Storage server type: NetAppAltaVault
- Disk pool configured for: Backup
- Disk Pool Size section: Total available space: 1.78 PB, Total raw size: 1.78 PB
- Disk Pool name: ostshareDP (entered in the text field)
- Comments: OST share Disk Pool for Regular Backups (selected in the dropdown)
- High water mark: 98 % (spin box)
- Low water mark: 80 % (spin box)
- Maximum I/O Streams section: Contains an information icon and text 'Concurrent read and write jobs affect disk performance. Limit I/O streams to prevent disk overload.' Below this is a checkbox 'Limit I/O streams:' which is unchecked, followed by a spin box showing '-1' and the text 'per volume'.

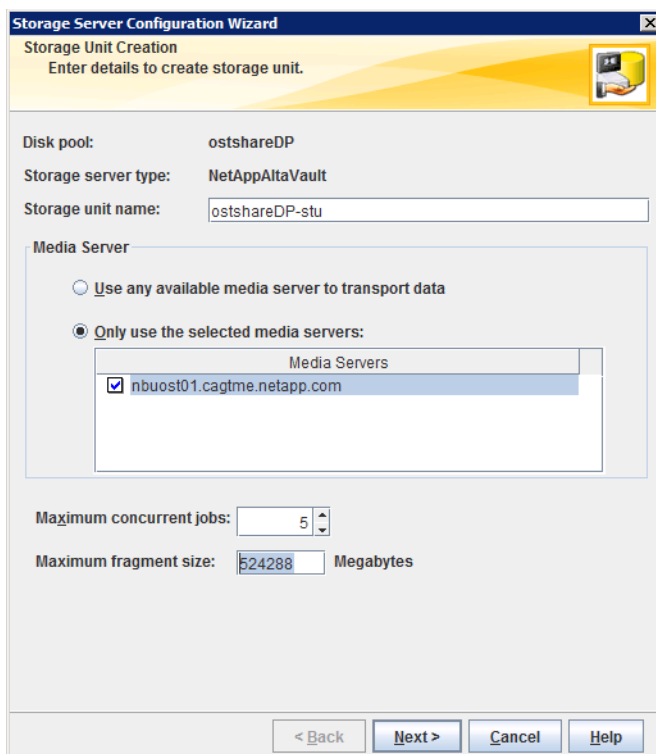
At the bottom of the window are four buttons: '< Back', 'Next >', 'Cancel', and 'Help'.

- Disk Pool name
Provide the name of the disk pool you are creating.
- High water mark
Leave this value at its default value of 98 percent.
- Low water mark
Leave this value at its default value of 80 percent.
- Limit I/O streams
Leave this checkbox unselected for best AltaVault performance.

- When the disk pool task is done, select the checkbox **Create a storage unit using the disk pool that you have just created**. Click **Next**.



- In the Storage Unit Creation page, fill in the fields as follows. Click **Next**.



- Storage unit name

Provide a name for the storage unit you are creating.

- Only use the selected media servers

Identify the NetBackup media servers which are enabled with the AltaVault OST Plug-in to perform data movement of backups.

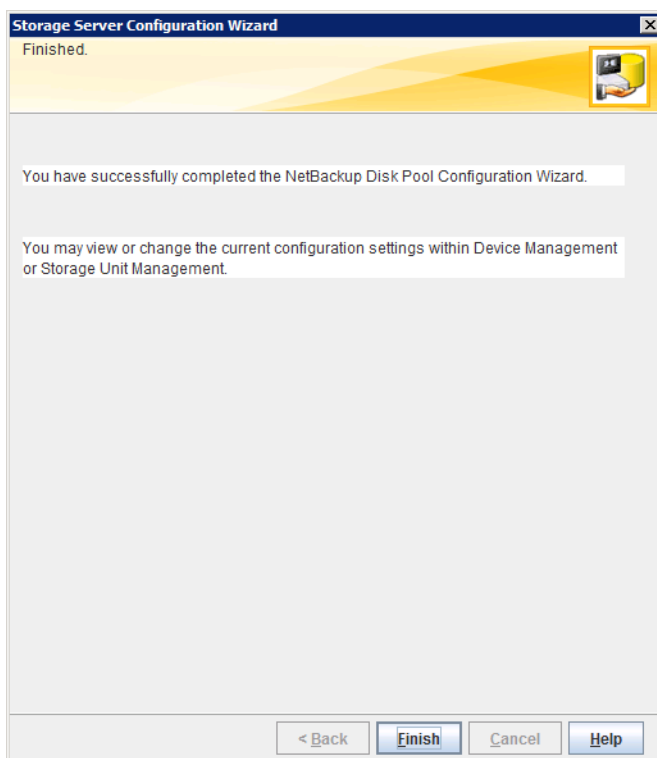
- Maximum concurrent jobs

This establishes how many writers are allowed by data protection clients. Tune performance by setting a value higher than 1. The value will depend on your available resources and infrastructure environment. Adjust the number of streams accordingly based on your observed performance.

- Maximum fragment size

The fragment size is the maximum-size object that NetBackup can create to store backups. AltaVault recommends that a fragment size should be no larger than 20480MB (20GB).

7. The Storage Server Configuration Wizard completes. Click **Finish** to close.

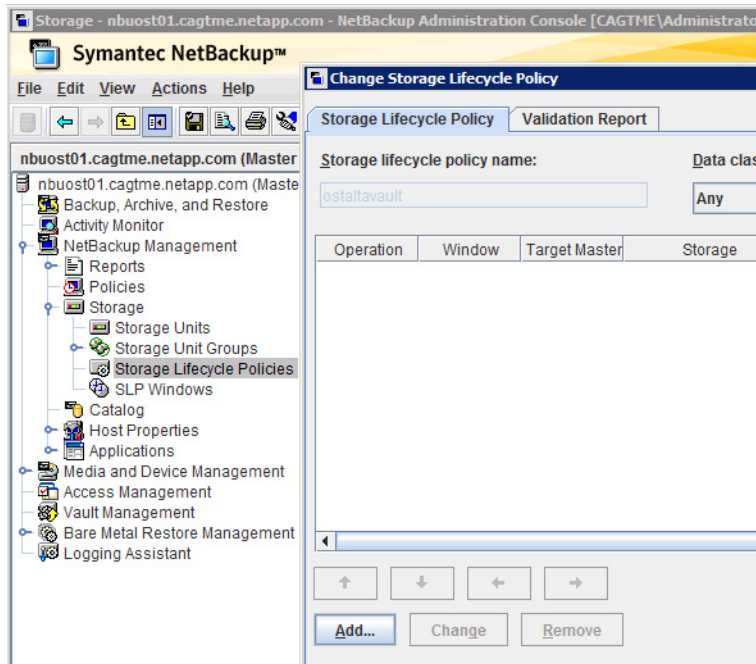


Create a Storage Lifecycle Policy

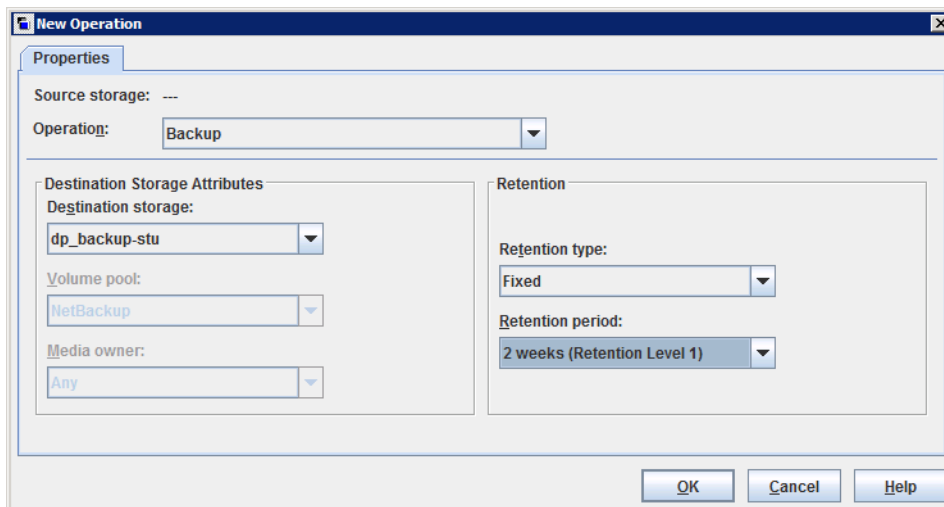
A storage lifecycle policy (SLP) allows NetBackup to maintain backup copies that meet retention requirements of the business. For example, weekly backups may be required to be held for 1 month, while monthly backups may be required to be held for 7 years. You will need to create at least one backup operation in an SLP representing the primary backup to AltaVault. Additional optimized duplication operations can be created as tiered operations under a backup operation to reflect the longer term retentions of backups according to your business needs.

Create a backup operation

1. Open the Storage Lifecycle Policies page in the NetBackup Administration Console, and click Add to add a backup step to the SLP.



2. In the page that appears, fill in the fields as follows. Click **OK** to complete the backup operation creation.



- **Operation**
Select Backup to identify a backup job that will be sent to the AltaVault OST share.
- **Destination Storage**
Select the storage unit configured to use the regular AltaVault OST share.
- **Retention type**
Specify your retention requirements.
- **Retention period**

Specify your retention time frame to maintain the backup. Backups after this time elapse expire from AltaVault and will be deleted.

Create an optimized duplication operation

1. Select the backup operation created previously, and then click **Add** to add an optimized duplication operation.

Change Storage Lifecycle Policy

Storage Lifecycle Policy | Validation Report

Storage lifecycle policy name:

Data classification:

Priority for secondary operations: (higher number is greater priority)

Operation	Window	Target Master	Storage	Volume Pool	Media Owner	Retention Ty...	Retention P...	Alternate Re...	Preserve m...
Backup	--	--	dp_backup-stu	--	--	Fixed	2 weeks	--	No

State of secondary operation processing

☒ Active
☐ Postponed

☐ Until

To find impact on Policies associated with this SLP due to change in configuration click here.

2. In the page that appears, fill in the fields as follows. Note that the operation type should default to Duplication if you properly selected the backup operation in the previous step. Click **OK** to complete the optimized duplication operation creation

- **Operation**
Select Duplication to identify a duplication operation will be performed to a cloud based AltaVault OST share.
- **Destination Storage**
Select the storage unit configured to use the regular AltaVault OST share.
- **Retention type**
Specify your retention requirements.
- **Retention period**
Specify your retention time frame to maintain the backup. Backups after this time elapse expire from AltaVault and will be deleted.
- **Alternate read server**
You can specify an additional Media Server if you wish to read duplicated backups from a separate media server. This media server must also have the AltaVault OST Plug-in installed.

An example SLP is shown below after creation of the backup and duplication operations.

Change Storage Lifecycle Policy

Storage Lifecycle Policy | Validation Report

Storage lifecycle policy name: Data classification: Priority for secondary operations: (higher number is greater priority)

Operation	Window	Target Master	Storage	Volume Pool	Media Owner	Retention Ty...	Retention P...	Alternate Re...	Preserve m...
Backup	--	--	dp_back...	--	--	Fixed	2 weeks	--	No
Duplication Default_24x...	--	--	dp_dupli...	--	--	Fixed	1 year	nbuost01.ca...	No

State of secondary operation processing

☒ Active

☐ Postponed

☐ Until

To find impact on Policies associated with this SLP due to change in configuration click here.

[Validate Across Backup Policies](#)

OK Cancel Help

Summary of AltaVault OST best practices

Practice	NetApp Recommendation	Notes
Disk pool maximum I/O streams	Do not enable and provide a value for maximum I/O streams.	Use maximum concurrent jobs under the Storage Unit settings instead.
Assigning media servers to a storage unit	Manually select the appropriate media servers with the AltaVault OST Plug-in installed.	
Maximum concurrent jobs	Set appropriately according to your available media server resources.	Tune this setting rather than the maximum I/O streams value at the disk pool level. Set to 0 if you are switching cloud providers, or are experiencing intermittent cloud connectivity issues.
Maximum fragment size	20480 MB.	This can be left at the default if desired by your business, but is not recommended to be set lower than 20480 MB.
NetBackup Catalog backups	AltaVault OST can be used for NetBackup catalog backup operations.	

Copyright information

Copyright © 1994-2016 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.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark information

Trademark information

NetApp, the NetApp logo, Go Further, Faster, AltaVault, ASUP, AutoSupport, Campaign Express, Cloud ONTAP, Clustered Data ONTAP, Customer Fitness, Data ONTAP, DataMotion, Fitness, Flash Accel, Flash Cache, Flash Pool, FlashRay, FlexArray, FlexCache, FlexClone, FlexPod, FlexScale, FlexShare, FlexVol, FPolicy, GetSuccessful, LockVault, Manage ONTAP, Mars, MetroCluster, MultiStore, NetApp Insight, OnCommand, ONTAP, ONTAPI, RAID DP, RAID-TEC, SANtricity, SecureShare, Simplicity, Simulate ONTAP, Snap Creator, SnapCenter, SnapCopy, SnapDrive, SnapIntegrator, SnapLock, SnapManager, SnapMirror, SnapMover, SnapProtect, SnapRestore, Snapshot, SnapValidator, SnapVault, StorageGRID, Tech OnTap, Unbound Cloud, and WAFL and other names are trademarks or registered trademarks of NetApp, Inc., in the United States, and/or other countries. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such. A current list of NetApp trademarks is available on the web at <http://www.netapp.com/us/legal/netapptmlist.aspx>.

How to send your comments

You can help us to improve the quality of our documentation by sending us your feedback.

Your feedback is important in helping us to provide the most accurate and high-quality information. If you have suggestions for improving this document, send us your comments by email to doccomments@netapp.com. To help us direct your comments to the correct division, include in the subject line the product name, version, and operating system.

You can also contact us in the following ways:

- NetApp, Inc., 495 East Java Drive, Sunnyvale, CA 94089 U.S.
- Telephone: +1 (408) 822-6000
- Fax: +1 (408) 822-4501
- Support telephone: +1 (888) 463-8277

Index

A

- AltaVault
 - appliance deployments 7
 - cloud appliance 7
 - definition 5
 - deployment guidelines 7
 - installations 9
 - overview 5
 - physical appliance 7
- AltaVault configuration wizard 10
- Amazon
 - AWS IAM 10
 - Amazon Machine Image 10
- Amazon S3
 - migrating to Amazon Glacier 22
- AMI 7
- AVA400 7
- AVA-c16 7
- AVA-c4 7
 - Azure 7
- AVA-c8 7
- AVA-v16 7
- AVA-v2 7
- AVA-v32 7
- AVA-v8 7

B

- backup applications 6

C

- cloud
 - agility 21
 - migration 21, 22
 - set credentials 22
 - settings 14
 - supported providers 6
- configuration
 - file export 12
 - recovery 13
- Configuring email settings 10
- Configuring NFS 10
- Configuring SMB 10
- Connecting the Network 9

D

- data
 - eviction 6
 - migration 21
 - restoration 12

- restoration for disaster recovery 15
- data restoration
 - disaster recovery testing 14
- deployment
 - guidelines 7
 - steps 9
- disaster recovery
 - definition 11
 - preparation 12
 - testing 12
 - testing activities 14
- downgrades 7

E

- ESXi 7
- eviction of data 6
- Exporting configurations 10

G

- Glacier 22

K

- KVM 7

M

- Microsoft Azure virtual machine 10
- Microsoft Hyper-V 7, 9
- migration
 - data 21
 - monitor cloud 22
- models
 - virtual 7
- Modifying data interfaces 10

N

- Network File System (NFS) protocols 5
- NFS 10
- NFS protocols 10

O

- OpenStorage (OST) 35
- OST best practices 49
- OST deployments 39
- OST features 35

R

- Related reading 10
- REST API 24
- restoring data 12

Index

S

S3 Bucket Policies 10

site preparation 9

SLP 45

SMB 5, 10

storage lifecycle policy (SLP) 45

V

Veritas KVM 7

Veritas OST 35

VMware 7

VMware ESXi 9