VERITAS

AWS Cloud Storage with Veritas NetBackup

Long-Term Retention Solution.

White Paper | May 2022

Contents

Revision History . . . 3

Introduction . . . 4

- Executive Summary . . . 4
- Scope . . . 4
- Target Audience . . . 4

Solution Value . . . 4

Solution Key Featuares . . . 5

- Data Insight . . . 5
- Storage Efficiencies . . . 5
- Security . . . 5
- Protection. . . . 5
- Migration Options . . . 5

Solution Architecture Overview . . . 6

Solution Components . . . 7

- Data Insight . . . 7
- NetBackup . . . 8
  - Deduplication . . . 9
  - Traditional Duplication (Without Deduplication) . . . 10

AWS . . . 10

- S3 Storage Classes . . . 10
- Network Connectivity . . . 12
- AWS Snowball and AWS Snowball Edge . . . 12

Solution Integration Flow . . . 13

- Identification of Data to Send to the Cloud . . . 13
- Optimized Duplication (Deduplication) Data Flow to Amazon S3 Storage Classes . . . 14
- Access Appliance Optimized Duplication (Deduplication) Data Flow to AWS Cloud Storage . . . 16
- Traditional Duplication Data Flow . . . 17
- Migration to the Cloud . . . 19

Disaster Recovert in the Cloud . . . 20

Best Practices and Recommendations. . . . 20

- Privacy Laws. . . . 20

Contents

Compression 21
Deduplication 21
AWS Snowball/AWS Snowball Edge 21
NetBackup Retrieval Attributes 21
Conclusion 21
References 22
Appendix 22
AWS Security and Access Keys 24
Creation of NetBackup Storage Server, Disk Pools and Storage Units for MSDP Local Storage and MSDP-C Cloud . . . 26
Creation of the Protection Plan 29
Validation of Setup 30

Revision History

Date	Update
Rev 1.01 November 2019	Initial version
Rev 2.0 26 January 2021	Updates relating to NetBackp 8.3, NetBackup 9.0 and Appliances

Introduction

Executive Summary

More companies are venturing to the public cloud as an option for long-term data retention and/or to safeguard their data from on-premises failures, attacks and disasters. Cost, security and insight into what data can be sent to the cloud have been some of the main requirements. The Veritas suite of products alleviates some of these issues and concerns. Veritas NetBackup™, for instance, has encryption features for data at rest and in motion for security. It also has storage efficiency technologies such as deduplication and compression to reduce egress and ingress costs to and from the cloud. The use of Veritas Data Insight provides valuable insights to identify which data should remain on-premises and which is ideal to move to the cloud. NetBackup supports sending and retrieving data to and from the different storage classes offered by Amazon Web Services (AWS), with varying costs, availability and performance for long-term retention and preservation of an organization's critical digital assets. In addition, NetBackup supports AWS Snowball and AWS Snowball Edge for large-scale data migration or initial seeding of data to the public cloud. Veritas products and AWS cloud storage and services work together to optimize data protection, reduce cost and minimize risks and liability.

Scope

The purpose of this document is to provide technical details to assist in understanding AWS cloud storage with NetBackup as a solution for long-term retention of backup data. It describes the components of this solution, its value and some best practices. We recommend you refer to Veritas product documentation or AWS documentation for installation, configuration and administration of each of the products discussed in this white paper. NOTE: *This document is updated periodically, so if you downloaded a local copy, please get the latest version from this* link.

Target Audience

This document is targeted for customers, partners and Veritas field personnel interested in learning more about the AWS cloud storage and services with NetBackup solution for long-term retention. It provides a technical overview of this solution and highlights some best practices.

Solution Value

The Veritas portfolio of data management and protection products with AWS cloud storage and services provides advantages that include:

- Valuable insights—Organizations tend to blindly back up data or not remove stale or orphaned data from their primary storage. Data Insight provides a view of digital assets residing on file system storage for assistance in data placement and lifecycle. With new regulatory requirements imposed by numerous countries, it is crucial to gain an understanding of your data, its value and liability.
- Reduced risk—NetBackup allows for encryption of data prior to transmission to AWS cloud storage. Data maintains its encrypted form in AWS cloud storage for data at rest. For data in flight, NetBackup uses the Secure Sockets Layer (SSL) protocol for data transfers between NetBackup and cloud storage for enhanced security.
- Minimized cost—With NetBackup deduplication and compression features, the amount of data sent or retrieved from the cloud is reduced, minimizing overall costs. Also, support for sending data to Amazon Simple Storage Service (Amazon S3), including Amazon S3 Glacier and Amazon S3 Glacier Deep Archive storage classes, results in additional cost savings.

Solution Key Features

The key features companies look for in an off-premises, long-term retention solution include insights, security, storage efficiency and migration path. AWS cloud storage services with NetBackup provides these features to assist customers in preserving their most valued data.

Data Insight

One of the challenges organizations often face is deciding which data to send to the public cloud. Data Insight has the ability to scan and classify file system sources such as filers, SharePoint, Documentum repositories and cloud storage. It classifies the data into certain categories such as ownership, age, size, activity and access patterns so administrators can identify data they can archive or tier to cheaper storage, enforce security, perform information lifecycle management and risk analysis. The ability to identify areas of risk, value and ROT (redundant, obsolete, trivial) in data improves operational efficiency, reduces storage cost and minimizes risk and liability.

Storage Efficiencies

Support for storage efficiency is one of the main factors to consider when choosing and purchasing a long-term retention storage platform solution. The ability to maximize storage space helps reduce overall cost. Backup images stored in AWS cloud storage can be deduplicated using NetBackup Media Server Deduplication Pool (MSDP) technology. Data is sent to Amazon S3 storage classes using the S3 protocol.

NetBackup also supports compression prior to sending data to AWS cloud storage. Compression improves storage utilization by reducing the number of bits required to represent data. The type of data defines the degree to which a file can be compressed. Data types that compresses well include text files or unstripped binaries. Data that is already compressed and stripped binaries are not good candidates for compression. You can find detailed information on NetBackup compression attributes in the NetBackup Cloud Administrators Guide and the NetBackup Deduplication Guide.

Security

NetBackup has security features that protect all NetBackup components and operations at different security implementation levels such as the data center and the enterprise. For further details on NetBackup's security implementations and levels, refer to the NetBackup Security and Encryption Guide.

For enhanced security, NetBackup also offers encryption of data. Any encryption done by NetBackup is maintained on the cloud storage target. For more information on how NetBackup conducts encryption specifically for cloud storage servers, refer to the NetBackup Cloud Administrators Guide. When using NetBackup deduplication technology, there is encryption for deduplicated data that is separate and different from NetBackup policy-based encryption. For more information on implementation, refer to the NetBackup Deduplication Guide. Additional security that is employed for this solution is the requirement to use access keys and credentials when configuring AWS as a cloud storage destination. Data transfers to and from AWS can also be secured by enabling the NetBackup SSL feature.

Protection

As part of automated disaster recovery (DR) in the cloud, NetBackup will send deduplicated data to a cloud storage target in addition to the metadata. This "self-describing image" feature allows for recovery of images in the cloud when the on-premises NetBackup catalog is not available due to corruption, power outage and network or other issues. You can reconstruct data and metadata in the cloud or in a different data center by instantiating a NetBackup instance from the AWS Marketplace and attaching the existing bucket to the new NetBackup domain.

Migration Options

For customers that are starting to migrate to the cloud and need to move their "initial data set" or replace their tapes with AWS cloud storage, NetBackup supports AWS Snowball and AWS Snowball Edge. Using these devices with NetBackup provides an easy path to the cloud when dealing with large datasets.

Solution Architecture Overview

Figure 1 depicts a high-level overview of the Veritas data protection solution with AWS cloud storage for long-term retention. An integral part of this solution is the use of Data Insight to provide information on the data residing on-premises to help make decisions about the lifecycle of the data. A report generated by Data Insight assists users in making informed decisions about which data should to store on-premises or send to the cloud and/or define the storage lifecycle policies of the data. For instance, the lifecycle of backup data can first reside on-premises on a NetBackup Appliance or Build Your Own Server (BYOS) for the short term, then move to Veritas Access Appliances for mid-term retention and finally to the AWS cloud for long-term retention and DR. NetBackup supports several Amazon S3 storage classes such as S3 Standard, S3 Standard Infrequent Access, S3 Standard One Zone Infrequent Access, S3 Glacier, S3 Glacier Vault and S3 Glacier Deep Archive. These classes of storage differ in terms of cost, usage, restore time, availability and other services.

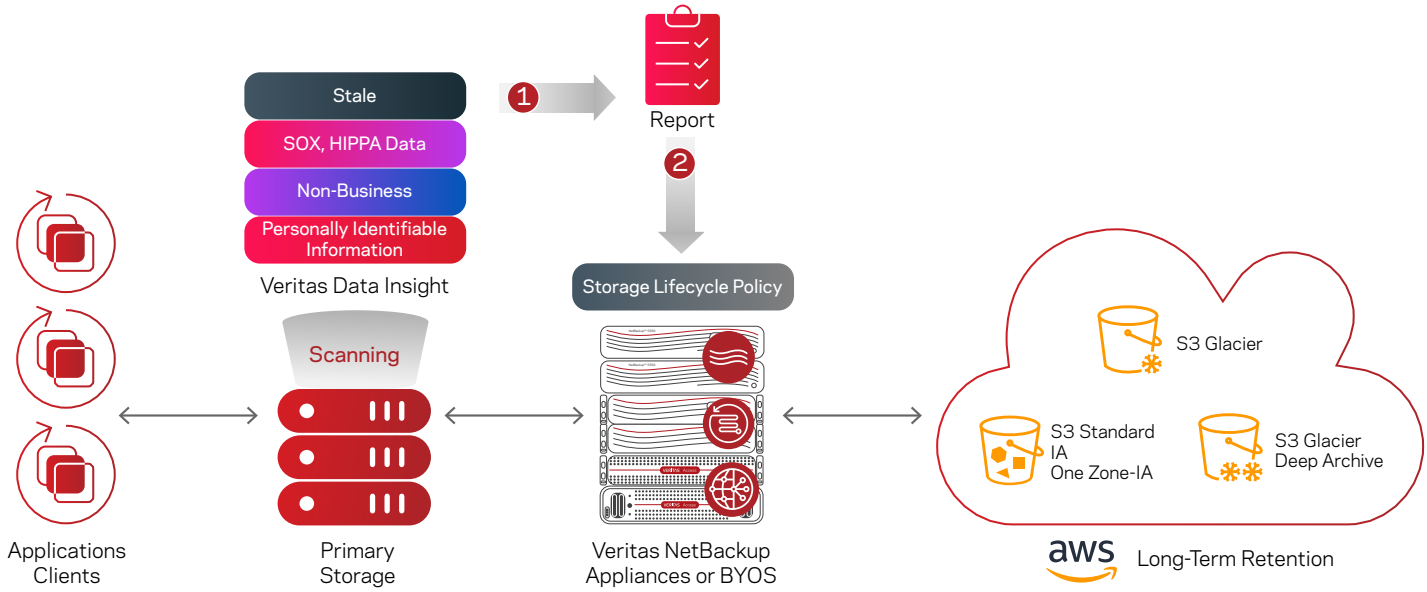


Figure 1. A high-level solution overview of NetBackup with AWS cloud storage classes.

For migration of data from tape or initial seeding of the cloud with on-premises backup data, AWS offers shippable physical storage devices, AWS Snowball and AWS Snowball Edge. NetBackup supports these devices to provide customers a migration path to AWS cloud storage. In general, as shown in Figure 2, AWS Snowball and AWS Snowball Edge devices are placed on-premises and communicate with NetBackup either via Amazon S3 or the Network File System (NFS) protocols, depending on which AWS Snowball device is used. Once data is transferred to the AWS storage device, it is then shipped to the AWS data center via courier. After its arrival, AWS migrates the data to the destined S3 Standard storage class specified by NetBackup during cloud storage server configuration.

For migration of data from tape or initial seeding of the cloud with on-premises backup data, AWS offers shippable physical storage devices, AWS Snowball and AWS Snowball Edge. NetBackup supports these devices to provide customers a migration path to AWS cloud storage. In general, as shown in Figure 2, AWS Snowball and AWS Snowball Edge devices are placed on-premises and communicate with NetBackup either via Amazon S3 or the Network File System (NFS) protocols, depending on which AWS Snowball device is used. Once data is transferred to the AWS storage device, it is then shipped to the AWS data center via courier. After its arrival, AWS migrates the data to the destined S3 Standard storage class specified by NetBackup during cloud storage server configuration.

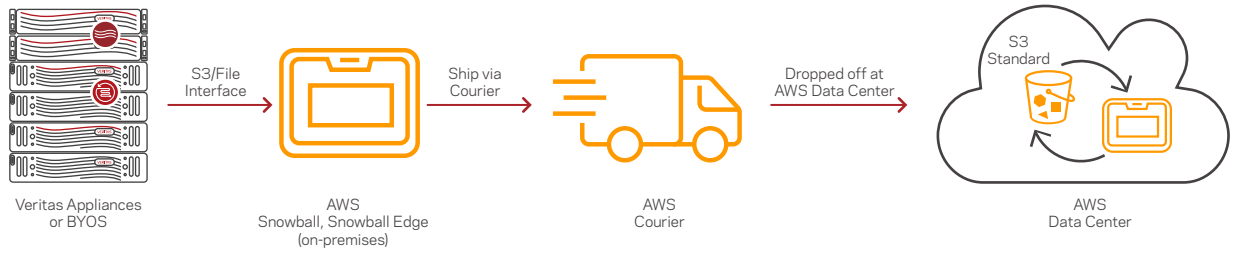


Figure 2. An overview of the data migration flow with NetBackup, AWS Snowball and AWS Snowball Edge.

Solution Components

To get a better understanding of how AWS cloud storage services work with NetBackup for long-term data retention, the following sections explain all the involved solution components in further detail. The main components discussed include Data Insight, NetBackup, Amazon S3 storage classes, AWS Snowball and AWS Snowball Edge.

Data Insight

Data Insight provides visibility by connecting to various file system data sources such as NetApp, Dell EMC, SharePoint and Microsoft OneDrive and scanning the file systems to collect metadata, including information such as data owner, file access date, creation date and file type. It also has the ability to conduct classification of the data into certain categories such as ownership, age, size, activity, stale, non-business, user risk, type and data patterns. This process enables administrators to identify data that can be archived or tiered to cheaper storage or the public cloud. Architecturally, Data Insight consists of four components:

1. **Management Server**—Hosts the web user interface that provides reporting, remediation and administration.
2. **Collector Node**—Scans data sources and collects information on the data (such as size and access control lists) and data access events (such as creation, access and modifications)
3. **Indexer Node**—Acts as the database to store and index the data collected from the Collector Node.
4. **Classification Server**—Retrieves content from the data sources, looks for patterns that matches the user-defined or preconfigured policies and sends the assigned tags to the database.

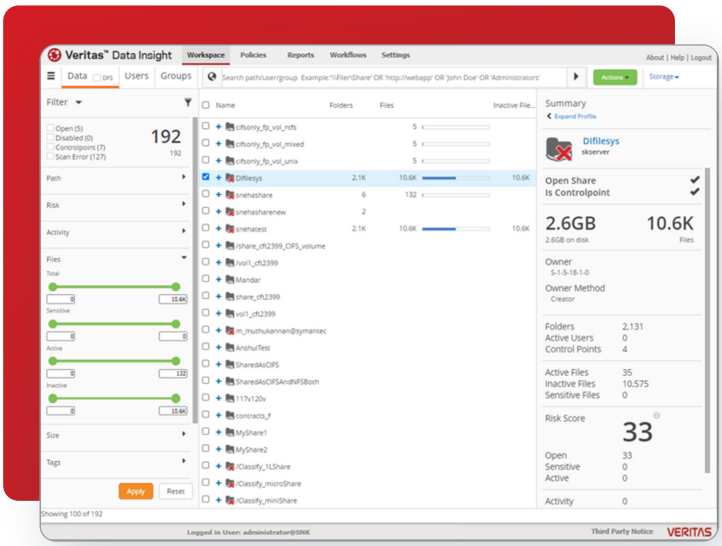


Figure 3. An example of Data Insight's web GUI.

Depending on the number and size of data sources, you can install Data Insight on single-, two- or three-tier architectures for better distribution of the workload. An example of Data Insight's web graphical user interface (GUI) is shown in Figure 3. The GUI displays information such as number of inactive files, where your files reside, file extensions and age. You can inspect the reports generated by Data Insight and use them to define the "Backup Selection List" within NetBackup backup policies. Storage lifecycle policies are used to specify the data that can be kept on-premises and/or duplicated to the AWS cloud. Knowledge of what type of information is stored in different data sources allows organizations to make more informed decisions about what to do with the data for storage optimization, security, compliance, archival and long-term retention. For more information on Data Insight, refer to the technical white paper Intelligent Data Migration to the Cloud.

NetBackup

NetBackup provides protection for a wide variety of data and platforms such as operating systems, virtual systems, databases and applications, files and all kinds of content. It has many features to speed up backups, snapshot management, backup automation and provide insights on where active and inactive backups are located. It has the ability to back up data to tape, storage area network (SAN), network-attached storage (NAS) and public or private clouds. Schedules, retention periods and the ability to tier to different types of storage are defined in policies or storage lifecycle policies (SLPs).

A typical NetBackup environment consists of three components:

1. **Primary (master) Server** —Manages and controls backup and recovery activities and hosts the catalog that contains policies and schedules, metadata about the backup jobs and media, device and image metadata information.
2. **Media Server**—Writes client data as backup images to varying types of storage such as local disks, tape, NAS, SAN and the cloud, and later restores the data to the client as instructed by the primary server.
3. **Clients**—NetBackup client components are installed on hosts that have the data to be backed up and are responsible for sending and receiving data to and from the media server for backup and recovery.

The primary and media server components can be in one system or distributed to several servers, depending on the number of clients and backup workload. For a small environment, the primary and media server typically can exist in one server, and for a large environment, there is one dedicated primary server and several media servers. NetBackup can also be configured in a multi-domain where there are separate instances of primary servers in different locations. In a multi-domain environment, each NetBackup domain is independent but can be centrally managed by the NetBackup OpsCenter, a web-based console for managing, monitoring and reporting on NetBackup operations. Figure 4 illustrates a sample configuration of a set of clients with a dedicated primary server and one media server.

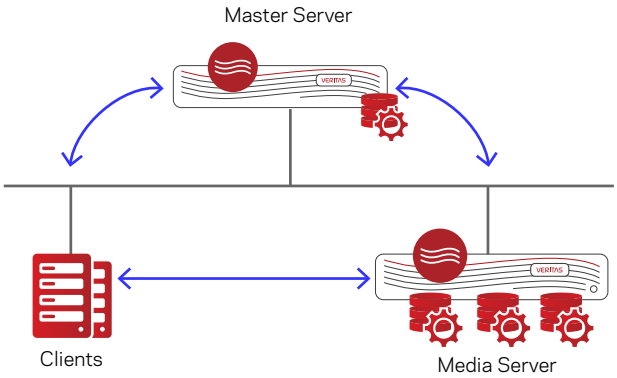


Figure 4. An example of a configuration with a dedicated .primary server and one media server protecting several clients

NetBackup is very flexible in its deployment. You can deployed it on an all Veritas Appliances, on commodity servers (also known as Build Your Own Servers or BYOS) or on a mixture of both. It also can be run on as on-premises and cloud-based virtual machines (VMs) or in "containers" when using the Flex 5340 Appliance and the Flex Scale Appliance. Following are the highlights of each of these deployment options:

- NetBackup Appliances Solutions—Purpose-built, highly tuned, scalable and resilient integrated appliances for NetBackup components. These appliances address the most demanding backup and recovery requirements of enterprises.
  - NetBackup 5250—For small and moderate workloads with a maximum of 442 TB storage capacity.
  - NetBackup 5340—For demanding workloads and those requiring higher usable capacity that can scale up to 2160 TB. The 5340 Appliances have high availability (HA) configurations that include an additional node to continue operations should the active node fail.
  - NetBackup Flex—NetBackup Flex supports container technology, allowing you to create multiple containers with different roles of primary or media servers in one Appliance. It also has support to create multiple domains in one Appliance. The appeal of NetBackup Flex is its multi-tenant capability and the ease of deploying a full NetBackup environment with multiple, independent versions of NetBackup quickly. NetBackup Flex is available in the 5340 or 5150 Appliances.
  - NetBackup Flex Scale—A hyperconverged, scalable solution built using commodity hardware. Flex Scale runs NetBackup services in a containerized fashion with efficient storage management. Designed for enterprise customers seeking a "pay as you grow" type of architecture.

- **NetBackup BYOS**—You can deploy NetBackup components on commodity servers running on a Linux or Windows platform. For a full list of supported platform versions, refer to the NetBackup Software Compatibility List. In production, the minimum requirement for a primary server is 4 cores and 16 GB of memory. Each media server has a 4 GB minimum memory requirement and clients require a minimum of 512 MB. Other than the hardware and platform differences, there is a difference in the maximum MSDP capacity that can be set up on a single server in BYOS. The MSDP capacity for BYOS is limited to 250 TB per server for systems configured with Red Hat Enterprise Linux (RHEL), Windows Server and SUSE Linux and 64 TB for others.
- **NetBackup Virtual Appliances**—NetBackup can also run on virtual appliances; however this deployment is appropriate mostly for remote offices. Implementing NetBackup on VMs provides a simple deployment and minimizes capital expenditures. For more information on supported hypervisors, see the NetBackup Software Compatibility List.
- NetBackup in Cloud AWS Marketplace—NetBackup is also available for automated deployment in the AWS Marketplace as a Bring Your Own License (BYOL) version.

The Access Appliance is part of the Veritas Appliance portfolio; however, it mainly acts as an on-premises, mid-term or long-term retention storage target for NetBackup. For those seeking to extend their on-premises, disk-based storage platform for faster recovery times, control and/or simplicity, the Access Appliance is a turnkey storage solution designed for high capacity and cost optimization. The Access Appliance model 3340 is composed of two clustered nodes and one main storage shelf and up to three additional expansion storage shelves. It can scale up to 2,800 TB of usable space.

There are two ways to store backup images in the public cloud from NetBackup:

1. **Deduplication** (optimized duplication)—Deduplication using NetBackup MSDP deduplication technology.
2. **Without deduplication** (traditional duplication)—Backup images are duplicated to the public cloud from NetBackup.

Deduplication

Backup images are generally ideal for deduplication because the probability of encountering duplicated blocks of data is higher when compared to other data types such as encrypted data. The deduplication ratio defines how well the data can be deduplicated. The higher the ratio, the more space you save. Deciding whether to deduplicate your data or not depends on several factors: data type, data change rate, retention period and backup policy. For instance, encrypted data is inherently unique and will not benefit from any deduplication savings. Data that has a high change rate will not take advantage of the savings long enough to justify the overhead imposed by deduplication. In the context of backup images, daily full backups will have higher deduplication ratios when compared with incremental or differential backups.

NetBackup MSDP is Veritas proprietary deduplication technology. Therefore, if the data is first placed in an MSDP and then duplicated to another storage platform that does not support MSDP technology, NetBackup would rehydrate deduplicated data prior to sending data to the storage platform. Rehydration involves putting the backup image back to a non-deduplicated form. NetBackup allows for inline deduplication of backup images on either the client or media server. The difference between client-side or media server deduplication is where the deduplication occurs. For a client-side deduplication target, the backup data is first deduplicated on the client before being sent to target storage. Client-side deduplication uses available resources on clients and reduces network traffic because deduplicated data is sent over the network. In either scenario, the backup images are placed in an MSDP.

Architecturally, NetBackup MSDP deduplication is composed of the following main components or services:

- **Deduplication Plug-In**—Separate the data into segments or chunks. Use a hash algorithm to calculate fingerprints to identify each unique segment. Compare incoming data fingerprints with the fingerprints of existing data.
- **Deduplication Engine** (spoold)—Manage and store the fingerprint database and metadata, store unique segments or use a reference or pointer to the data already stored and conduct integrity checks.
- **Deduplication Manager** (spad)—Maintains the configuration, controls and dispatches the internal processes, security and events handling.

NetBackup MSDP uses SHA-2 (SHA256) for the hash algorithm. The chunk segment size unit used to compute fingerprints is by default a fixed length of 128 KB or configurable to a variable-length size based on the chunk boundary. NetBackup MSDP also compresses deduplicated data for further storage efficiency. Furthermore, there is an option to encrypt deduplicated data. Both compression and encryption (if enabled) are performed after the fingerprint is calculated and prior to sending the data to the target storage. For more information on the architecture of NetBackup MSDP deduplication technology, refer to the NetBackup Deduplication Guide.

MSDP supports sending deduplicated data without rehydration directly to the AWS cloud, which is referred to as MSDP cloud (MSDP-C) tiering. Configuring an appliance or BYOS NetBackup as an MSDP-C storage server allows data to be sent via the S3 protocol to one local storage target and one or more cloud storage targets. However, there is support only for a combined capacity of 1.2 PB. In addition, client-side deduplication is supported, so you can send deduplicated data directly from the client to the cloud using MSDP-C.

Traditional Duplication (Without Deduplication)

In some cases, deduplication of backup images is not ideal. Backups that have a strict time limit for restores, have a high rate of change or are encrypted are not good candidates for deduplication. For these types of data or backup, it is best to send images to the public cloud without deduplication. Data is sent to the AWS cloud from NetBackup using the S3 protocol and stored in an Amazon S3 storage class. NetBackup traditional duplication to the AWS cloud supports S3 lifecycle policies, so data can transition from Amazon S3 Standard to Standard Infrequent Access and/or S3 Glacier.

AWS

AWS has a wide range of cloud storage services for backup, archival and long-term retention use cases. NetBackup supports several of the Amazon S3 storage classes, as indicated in the NetBackup Hardware and Cloud Storage Compatibility List. There are also different connectivity options provided by AWS based on desired transmission performance and cost. For movement of large-scale data to the cloud, AWS offers AWS Snowball and AWS Snowball Edge, on-premises compute and storage devices that include physical and logical security to securely send large amounts of data to the cloud.

A NetBackup Open Storage Technology (OST) plug-in is necessary to send or retrieve data to and from the AWS cloud storage target. OST is an Application Programming Interface (API) developed by Veritas to enable the development or creation of plug-ins to connect and manage third-party vendor storage platforms from within NetBackup. OST plug-ins were developed to interact with S3-compatible cloud providers. The OST cloud plug-ins are by default installed on NetBackup media and primary servers to send data to AWS cloud storage classes.

S3 Storage Classes

NetBackup can back up and retrieve data from the following Amazon S3 storage classes:

- **Amazon S3 Standard (S3 Standard)**—For frequently accessed data
- **Amazon S3 Standard Infrequent Access (S3 Standard-IA)**—For infrequently accessed data
- **Amazon S3 Standard One Zone Infrequent Access (S3 One Zone-IA)**—For infrequently accessed data that does not require the availability and resilience of S3 Standard or S3 Standard-IA. In this class, data is stored and available only in one zone
- **Amazon S3 Glacier (S3 Glacier)**—Mainly for archival or long-term retention of data. S3 Glacier Vault is available for data archival or long-term retention with specific lock policies such as WORM (Write Once, Read Many) or legal holds
- **Amazon S3 Glacier Deep Archive (S3 Glacier Deep Archive)**—The cheapest of the Amazon S3 storage classes and priced to compete with tape or "cheap and deep" on-premises storage platforms. Best for long-term retention of data

The Amazon S3 storage classes above are available with NetBackup in several regions. For up-to-date region availability, please refer to the NetBackup Hardware and Cloud Storage Compatibility List. Here are some of the available regions:

- Asia Pacific—Mumbai, Seoul, Singapore, Sydney, Tokyo, Bahrain
- Canada—Central
- China—Beijing, Ningxia, Hong Kong
- Europe—Frankfurt, Ireland, London, Paris
- South America—Sao Paulo
- U.S. East—N. Virginia, Ohio
- U.S. West—N. California, Oregon

For government agencies that require more stringent security and compliance features, NetBackup also supports the above storage classes within the GovCloud for region US West 1 (Gov FIPS 140-2, Gov Non-FIPS). When using these storage classes in the GovCloud, SSL is required, and encryption is offered.

Deciding to which AWS cloud storage classes you should send backup images is usually based on several factors, including restore time, cost and NetBackup feature support. Table 1 provides a comparison of each of the AWS cloud storage classes based on these factors. Refer to the AWS website for more information on Amazon S3 storage classes.

Table 1. Comparison of Amazon S3 Storage Classes

Amazon S3 Storage Class	Restore Time with NetBackup	Cost**	IPv6 with NetBackup	MSDP-C	Accelerator	Minimum Storage Duration
S3 Standard	Milliseconds	$$$$	Supported	Supported	Supported	Not Applicable
S3 Standard Infrequent Access	Milliseconds	$$$	Supported	Supported	Supported	30 days
S3 Glacier	With MSDP-C: ▪ Expedited: 1–5 minutes ▪ Standard: 3–5 hours ▪ Bulk: 5–12 hours Without MSDP-C: ▪ Standard: 3–5 hours	$$	Not Supported	Supported	Not Supported	90 days
S3 Glacier Vault	Standard: 3–5 hrs	$$ (about 2% Glacier)	Not Supported	Not Supported	Not Supported	90 days
S3 Glacier Deep Archive	With MSDP-C: ▪ Standard: within 12 hours ▪ Bulk: within 48 hours Without MSDP-C: ▪ Standard: within 12 hours	$	Not Supported	Supported	Not Supported	180 days

*** The "$$" represents a simple cost view. For exact pricing, please refer to AWS Pricing.*

Network Connectivity

The speed of backups and restores is highly dependent on the type of network connectivity. AWS offers three different types of connectivity to AWS regions from on-premises data centers:

1. **Internet**—Basic Internet connectivity and speeds offered by Internet service providers (ISPs).
2. AWS Direct Connect—Private, dedicated connectivity between an on-premises environment to an AWS region. AWS Direct Connect addresses some of the standard Internet challenges such as speed (bandwidth and throughput), network congestion and/or contention. It supports two bandwidth levels: 1 GbE and 10 GbE fiber-optic connection.
3. AWS Virtual Private Network (VPN)—Secure, private tunnel from an on-premises network to the AWS global network. It consists of the following two services:
   1) AWS Client VPN—Allows secure access to AWS resources via an on-premises network. An endpoint is configured to set a secure Transport Layer Security (TLS) VPN session.
   2) **AWS Site-to-Site VPN**—Enables a secure connection from an on-premises network to the AWS Virtual Private Cloud (VPC).

AWS Snowball and AWS Snowball Edge

For organizations that must send large amounts of data to Amazon S3 storage classes regularly to initially seed data and/or for data migration, AWS developed physical storage devices, AWS Snowball and AWS Snowball Edge. These devices are deployed on-premises and are required to be in the same region as the destination bucket in the AWS cloud. NetBackup connects to these devices and is configured as a cloud storage server. Organizations can duplicate regular backup data (live) or data residing on tapes (old) or secondary storage to these devices using NetBackup storage lifecycle policies. The data is transferred to these devices on-premises using AWS Snowball and AWS Snowball Edge tools. Once the transfer is complete, the devices are shipped by courier to the AWS data center, where data will be uploaded to the destined Amazon S3 bucket. If duplicating live regular backup data, then policies are suspended during the physical transport and resumed once data is available in the cloud or another AWS storage device is available on-premises. **NOTE**: *NetBackup currently only supports sending data to the Amazon S3 Standard storage class with these devices. It is also not supported with MSDP-C*. For information on how to configure NetBackup with the AWS Snowball and AWS Snowball Edge devices, refer to the NetBackup Cloud Administrator's Guide and the AWS website. Table 2 highlights the differences between these devices. Refer to the AWS website for more information or for the latest specifications of AWS Snowball and AWS Snowball Edge.

Table 2. Main Differences between AWS Snowball and AWS Snowball Edge

Features	AWS Snowball	AWS Snowball Edge
Raw Capacity	50 TB (U.S. Region only) 80 TB	100 TB - Storage optimized 42 TB plus 7.68 TB dedicated NVMe SSD - Compute optimized
Usable Capacity	42 TB (U.S. Region only) 72 TB	80 TB - Storage optimized 39.5 TB- Compute optimized **NOTE**: *NetBackup does not support Clustered Snowball Edge*.
Network Connector	RJ45, SFP+, SFP+ (with optic connector)	RJ45, SFP+, SFP+ (with optic connector), QSFP
Tools	Snowball client—Transfers data and encrypts to or from Snowball. Need to download it from the Amazon resources site and install it on a separate server. Amazon S3 Adapter—Transfers data to and from AWS Snowball using the Amazon S3 Rest API. Need to download and install it on a separate server. Minimum server specifications dedicated to run tools—16 core CPU and 16 GB of memory. **NOTE**: *Running multiple instances of the S3 Adapter or client requires 7 GB of memory.*	Snowball Client—Used to unlock AWS Snowball Edge. Need to download it from the Amazon resources site and install it on a separate server. Amazon S3 Adapter—Transfers and encrypts data to and from AWS Snowball Edge using the Amazon S3 REST API. Installed on the device by default. File Interface (NFS v3/v4/v4.1)—Allows transfers and encryption of data into a bucket on AWS Snowball using an NFS mount point. Installed on the device by default.

Solution Integration Flow

All data goes through a lifecycle from being created, read, modified, moved to other tiers of storage and eventually expired or deleted once it is no longer of use. When data is actively used, it resides in primary storage and is backed up to secondary storage for data protection. When data or backup data is infrequently accessed, organizations move it to cheaper storage on-premises and/or off-premises. Understanding data in terms of usage, age, type and whether it contains personally identifiable information (PII), is non-business or is subject to regulatory compliance is crucial in determining where to move that data across storage platforms or different tiers of storage on-premises or off-premises.

NetBackup manages the lifecycle of data using storage lifecycle policies (SLPs). NetBackup backup policies and/or SLPs define the path or flow of data. You can define backup policies to send data either to a single target or to an SLP. An SLP defines the lifecycle objectives of the data from backup to duplication to varying storage types and/or replication to different domains. As previously mentioned, data is sent to Amazon S3 storage classes using the S3 protocol with the S3 OST cloud plug-in installed on the media server. The solution integration flows described in this section include:

- Identification of data to send to the public cloud using Data Insight
- Optimized duplication (deduplication) to the cloud using MSDP-C
- Optimized duplication (deduplication) to the cloud from an Access Appliance
- Traditional duplication to the cloud
- Migration to the cloud

Identification of Data to Send to the Cloud

Identification of what data to send to the cloud can be done via visual inspection or if applicable, you can use Data Insight to do a more granular, user-defined filtered search to identify data you can send to the cloud. As shown in Figure 5, items within data sources such as SMB shares or Microsoft SharePoint are scanned and classified based on certain preconfigured filters or user-defined criteria such as PII, non-business and activity. You can generate a report either via the Data Insight web GUI or through APIs.

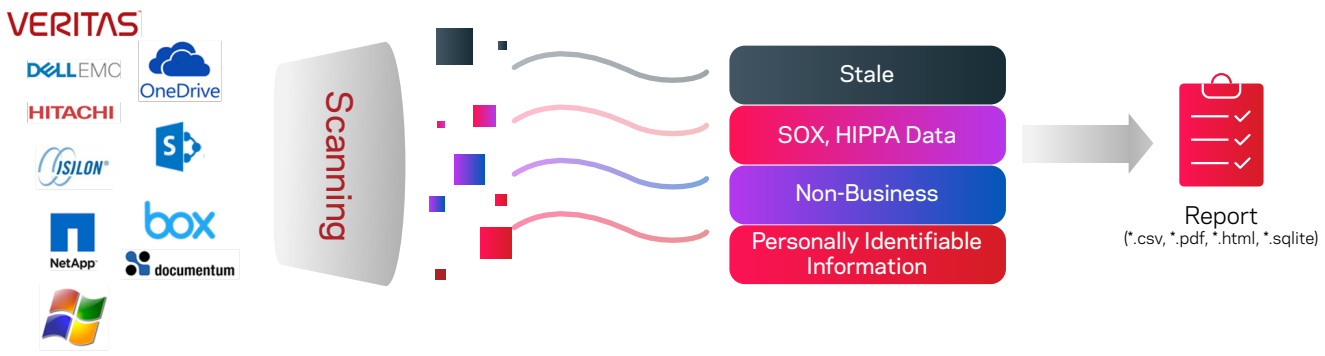


Figure 5. An overview of Data Insight's report generation flow.

You can create reports in comma-separated variable (CSV), PDF or HTML format using the GUI or via the APIs. Users who require the report in SQLite format to conduct more advanced SQL queries need to use the APIs. A sample snippet output in CSV format generated to find inactive files is shown in Figure 6. The contents of this report include a header that describes the type of data in the report such as share name, size, last accessed time, file path, creation time or time on disk. The extracted data is in the subsequent lines.

```
File Group,File Server \ Web Application,Share Name,Size (MB),Last Accessed Time,File (Path),DFS Path,Creation Time,Last Modified Time,On Disk Size (MB),

All Files,winserver.example.com,Demo_share,0.0,Fri Mar 31 12:50:37 PDT 2017,\\winserver.example.com\Demo_share\finance\2012\expense.txt:Zone.Identifier,,"Dec 31, 1969, 4:00 PM""Dec 31, 1969, 4:00 PM",0.0,

All Files,winserver.example.com,Demo_share,0.0,Fri Mar 31 12:50:34 PDT 2017,\\winserver.example.com\Demo_share\finance\2012\claims.txt:Zone.Identifier,,"Dec 31, 1969, 4:00 PM""Dec 31, 1969, 4:00 PM",0.0,

All Files,winserver.example.com,Demo_share,0.0,Fri Mar 31 12:50:27 PDT 2017,\\winserver.example.com\Demo_share\finance\2011\expense.txt:Zone.Identifier,,"Dec 31, 1969, 4:00 PM""Dec 31, 1969, 4:00 PM",0.0,
```

*Figure 6. A sample snippet of a *.csv report of old data.*

As shown in Figure 7, you would need to further filter this report using scripts to pull the file paths and then manually enter them into the NetBackup backup policy "Backup Selection List" attribute using the NetBackup administration console. Alternatively, you can also create an "inclusion" script or "exclusion" list file and feed it into the NetBackup commands using bpplinclude and bpsetconfig, respectively. Once the backup policy is defined, it can be targeted to an SLP to first do a backup to a local storage target followed by a duplication to a cloud target. Parameters in the SLP determine at what point the second copy is made as well as the retention policies for all copies in the SLP. The backup policy attribute "Policy Storage" is modified to use the desired storage lifecycle policy.

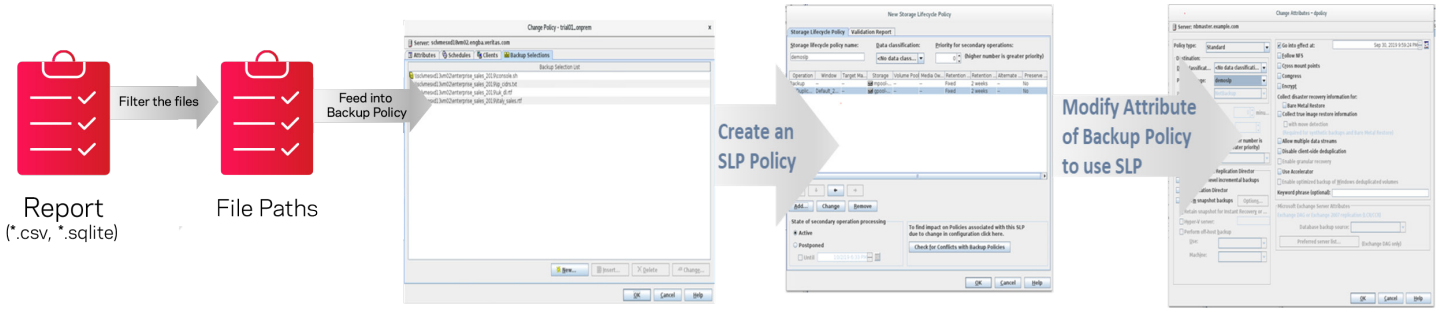


Figure 7. An overview of the Data Insight integration flow with NetBackup.

Optimized Duplication (Deduplication) Data Flow to Amazon S3 Storage Classes

Defining an MSDP-C storage server is required when sending deduplicated data to any of the Amazon S3 storage classes supported by NetBackup. The data paths and restore time differ, depending on the target Amazon S3 storage class. Below are example paths for each of the Amazon S3 storage classes supported by NetBackup with MSDP-C. Although client-side deduplication to AWS S3 storage classes is allowed, the most common path to cloud storage is when data from clients is initially backed up and deduplicated to an MSDP local storage for short-term retention and duplicated to the cloud for long-term retention via MSDP-C. The data can be restored from any of the copies that reside on-premises or in the public cloud. You send data to an Amazon S3 storage class using the S3 protocol with the S3 OST cloud plug-in installed on the media server, as previously discussed.

Figure 8 illustrates an example data path from a NetBackup client to an Amazon S3 Standard, Standard-IA and Standard One Zone-IA storage class. In this example, the client data is backed up and deduplicated to an MSDP local storage target residing on a media server for short- or mid-term retention (Copy 1). This deduplicated data is sent to AWS via MSDP-C, which uploads the unique data to the Amazon S3 Standard, Standard-IA or Standard One Zone-IA storage class for long-term retention (Copy 2). For restores, data goes through a similar path but in the reverse direction. By default, restores use images from Copy 1 unless specifically restored from different copies.

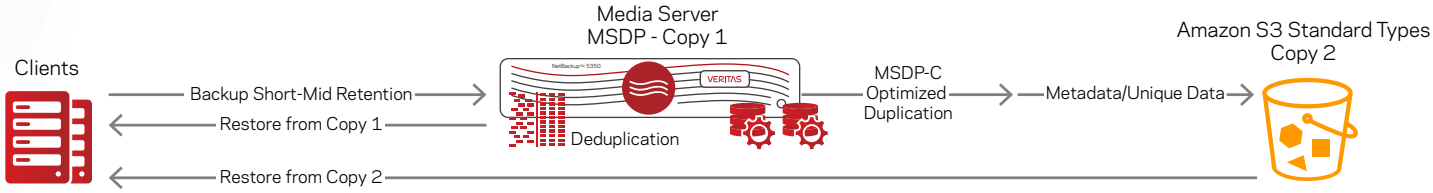


Figure 8. An overview of the data flow from NetBackup to and from Amazon S3 Standard, Standard IA and Standard One Zone IA storage classes.

When writing to an Amazon S3 Glacier storage class as shown in Figure 9, the metadata and data less than 256K are written to an S3 Standard storage class and regular data is written to an S3 Glacier storage class. In this example, client data is first backed up and deduplicated to MSDP local storage (Copy 1) on a media server and then duplicated to S3 Glacier (Copy 2) via MSDP-C using the S3 protocol. Restoration of data from S3 Glacier requires a "warming" step where the data is recalled in a temporary area (Reduced Redundancy Storage) within the AWS cloud prior to the data being retrieved. The media server would poll the status of the warming operation as part of the restore job. Once all fragments of the data are warmed for recovery, the media server would then restore the data and stream it to out to the client. AWS offers three types of retrieval operations for the S3 Glacier storage class:

1. Expedited—Completes within 1-5 minutes
2. Standard—Completes within 3-5 hours
3. Bulk—Completes within 5-12 hours

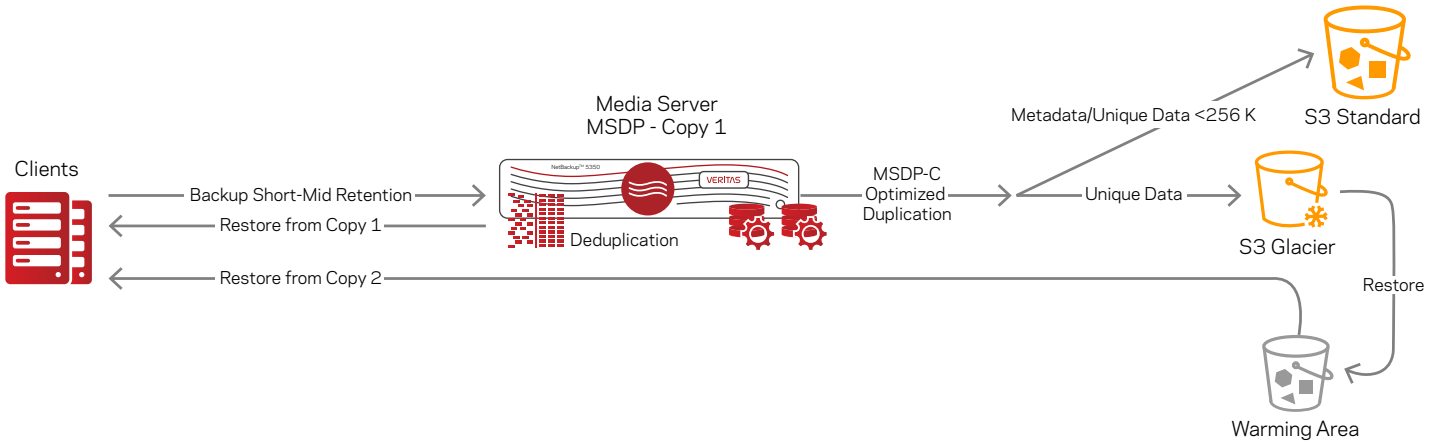


Figure 9. An overview of the data flow from NetBackup to and from an Amazon S3 Glacier storage class.

With NetBackup, the default retrieval type is "bulk." If you prefer to use the other types, then a GLACIER_RETRIEVAL file containing one of the strings "expedited" or "standard" is placed in the /usr/openv/netbackup/bin directory of the primary server. **NOTE**: *The restore in the warming area is done in a serial fashion, so if image consists of multiple fragments, the restore time may then be longer*.

As in the previous example, when sending data to the Amazon S3 Glacier Deep Archive storage class pictured in Figure 10, the client data is backed up and deduplicated on a media server and placed on MSDP local storage and then duplicated to the AWS cloud. Metadata and unique data less than 256K are sent to an S3 Standard storage class and larger-sized unique data is sent to an S3 Glacier Deep Archive storage class. A warming area within AWS is in the path during restores. When using an S3 Glacier Deep Archive storage class, supported retrieval types are:

- Standard—Completes within 12 hours
- Bulk—Completes within 48 hours
  Bulk is the default type unless the GLACIER_RETRIEVAL file in /usr/openv/netbackup/bin directory of the primary server contains the string "standard." NetBackup requests a restore from the S3 Deep Archive storage class and checks or polls the full availability of the data prior to retrieval of the data.

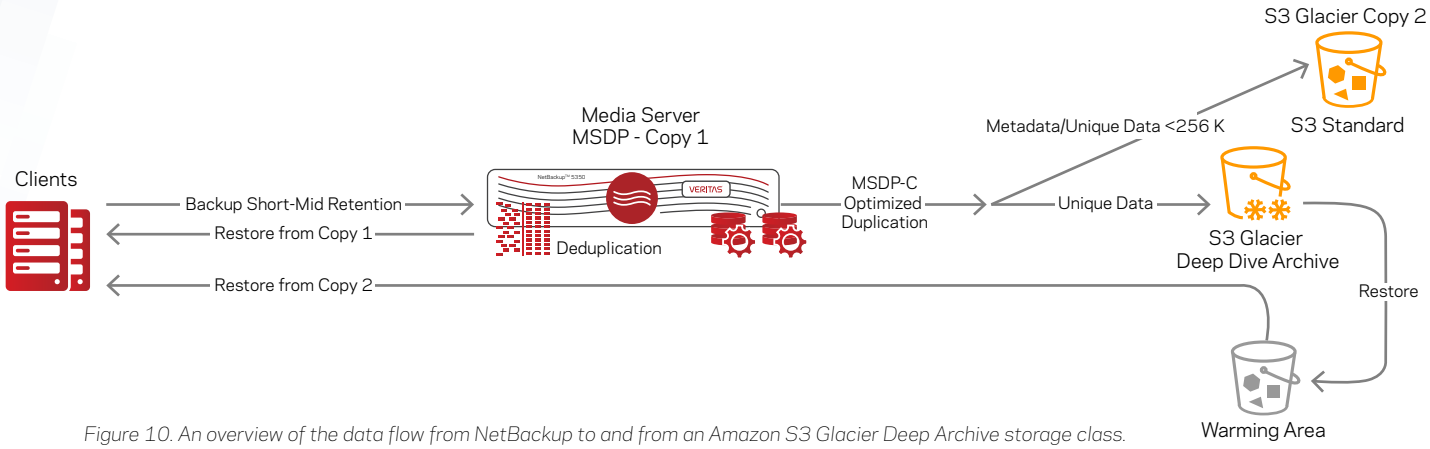


Figure 10. An overview of the data flow from NetBackup to and from an Amazon S3 Glacier Deep Archive storage class.

In addition to the MSDP meta-data, MSDP separates the backup image into container files and adds a header for each container. Thus, there are two files for each MSDP container: a data container file holding unique data with fingerprint and a header file for each container. If encryption is enabled, there is additional information relating to the keys and header for the keys.

Access Appliance Optimized Duplication (Deduplication) Data Flow to AWS Cloud Storage

In some scenarios, the Access Appliance is in the data path for on-premises mid-term or long-term retention prior to sending it off-premises. When using Access data deduplication with NetBackup, you can duplicate data to the cloud using NetBackup SLP policies. The SLP would specify to duplicate the data from the Access Appliance to the AWS cloud target. Figure 11 provides a view of this approach to sending data to the cloud from Access data deduplication. In this example, deduplicated data is sent to the Access Appliance from the media server and then it does an optimized duplication to the cloud using MSDP-C. The role of the media server during the optimized duplication to the cloud is to control and orchestrate the transfer between the Access Appliance and AWS cloud storage. The actual I/O is between the Access Appliance and AWS cloud storage. A restore can be done either from the Access Appliance (Copy 2) or from the AWS cloud (Copy 3). By default, Copy 1 is used for restores unless it is specified to restore from the different copies.

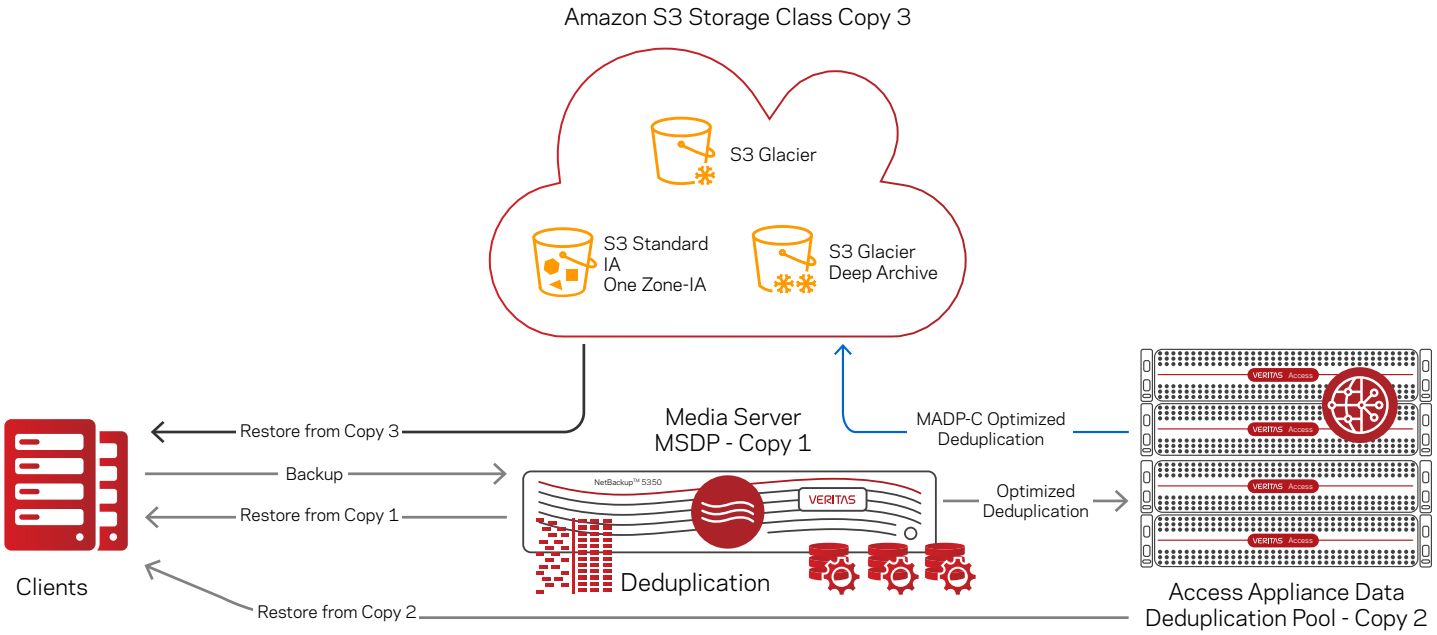


Figure 11. An overview of the Access Appliance data deduplication path to Amazon S3 storage classes via MSDP-C.

Traditional Duplication Data Flow

For data that does not benefit from deduplication, backup data is duplicated from the media server to Amazon S3 storage classes using the NetBackup S3 OST plug-in installed by default on the media server. Figures 12 to 14 show examples of the data flow from NetBackup to and from the different Amazon S3 storage classes.

The traditional flow for sending data to Amazon S3 Standard, Standard-IA and Standard One Zone IA storage classes involves the data being initially stored in an advanced disk (Copy 1) on a media server for short- to mid-term retention and then copied to the Amazon S3 storage class (Copy 2) for long-term retention. Restores can be done from either the advanced disk (Copy 1, the default) or from Amazon S3 storage classes (Copy 2), as shown in Figure 12.

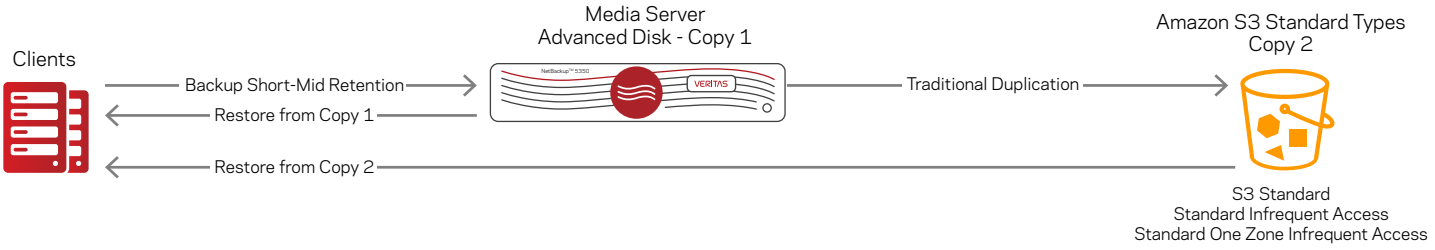


Figure 12. An overview of the data flow using NetBackup traditional duplication to and from Amazon S3 Standard storage classes.

When sending data to S3 Glacier or S3 Glacier Deep Archive, metadata is stored in an S3 Standard storage class and data is stored in an S3 Glacier or S3 Glacier Deep Archive storage class, as shown in Figure 13. Restoration for any of the Glacier types of S3 storage classes requires pulling the data to an Amazon warming area prior to retrieval from NetBackup. For traditional duplication, the retrieval type supported for both is standard, however the S3 Glacier storage class completes retrieval within 3–5 hours whereas Glacier Deep Archive storage classes complete within 12 hours.

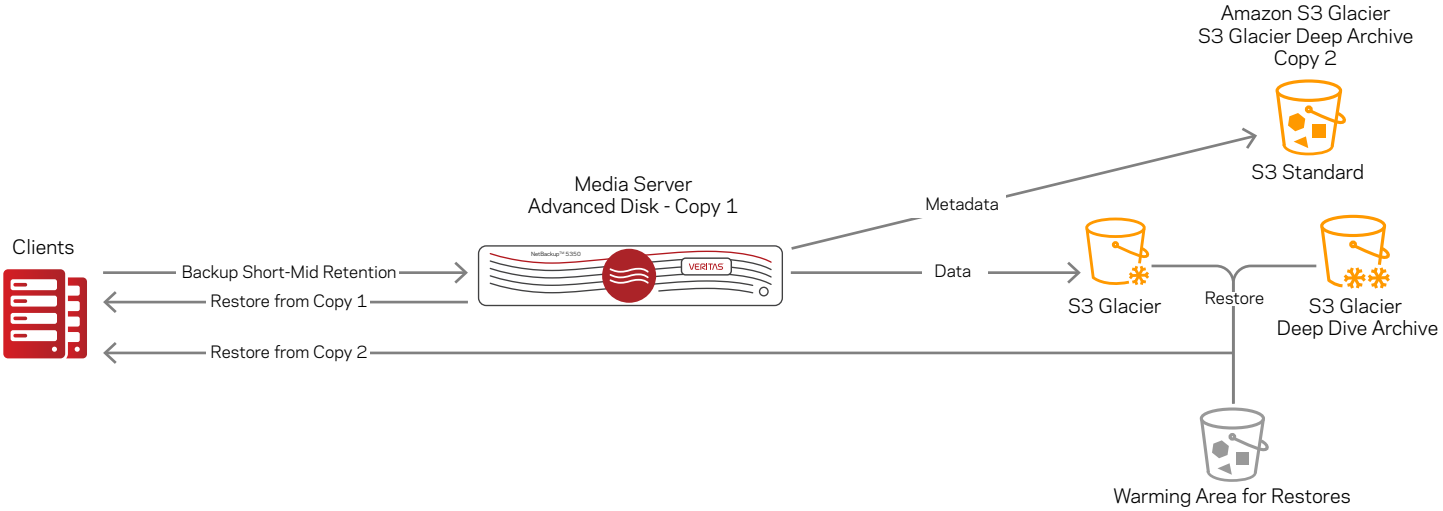


Figure 13. An overview of the data flow using NetBackup traditional duplication to and from Amazon S3 Glacier and Glacier Deep Archive storage classes.

If there is a requirement to store data with certain immutable or lock policies such as WORM or retention for compliance or regulatory rules, Amazon S3 Glacier Vault should be used. The vault lock policies are defined and managed within AWS after the S3 Glacier Vault is created as a cloud storage target within NetBackup. As shown in Figure 14, both the metadata and data are stored in the S3 Glacier Vault with vault lock policies applied. In addition, metadata is stored in an S3 Standard storage class. As with the S3 Glacier storage class, restores from S3 Glacier Vault require placing the data in a warming area prior to restoring it from NetBackup. However, unlike S3 Glacier, the retrieval type supported by NetBackup with S3 Glacier Vault is standard, so the restoration takes a minimum of 3 to 5 hours. The NetBackup media server checks the status of the restore in the warming area and pulls the data once it is fully available.
NOTE: *The data is available for retrieval from the AWS warming location for a maximum of 24 hours.*

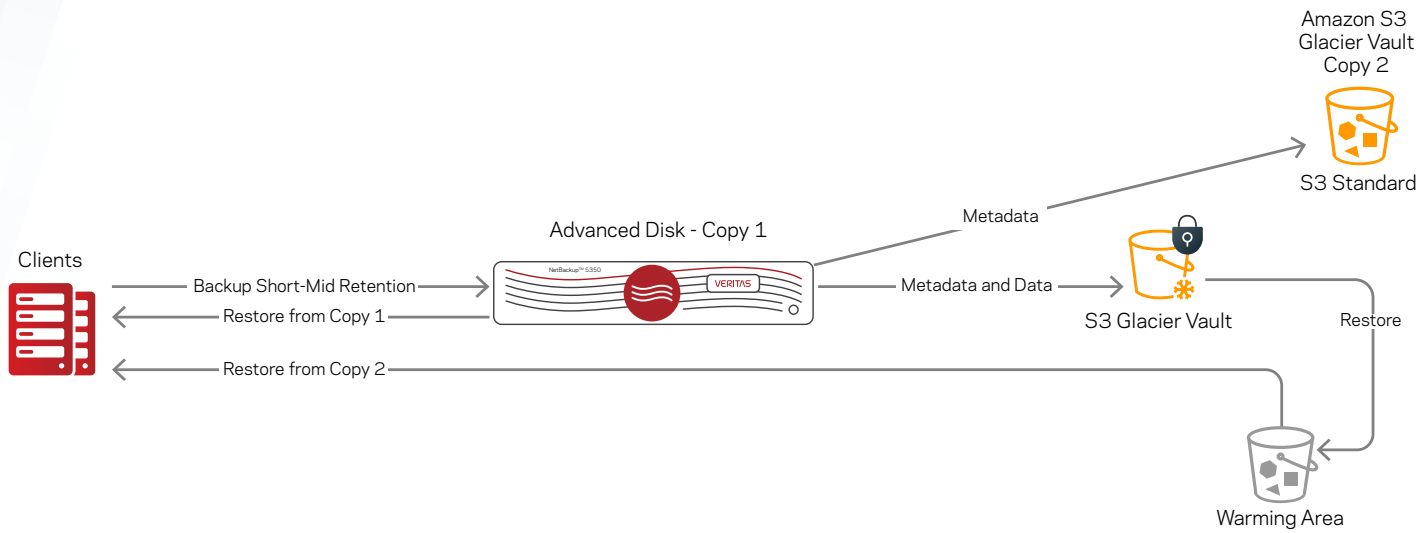
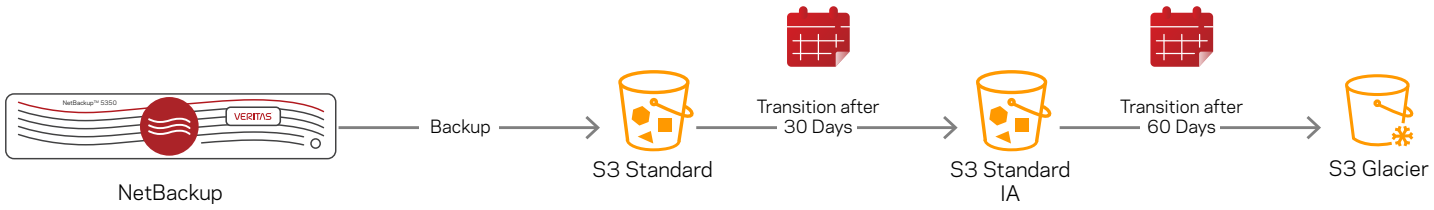


Figure 14. An overview of the data flow using NetBackup traditional duplication to and from Amazon S3 Glacier Vault.

Duplicated backup images and associated header information are stored in a directory structure. Each directory contains the image properties, the block map file and the actual backup image. The header directory contains the header information, the properties of the header information and the block map file for the header. The S3 OST cloud plug-in breaks the backup image up into configurable fixed object sizes. With encryption, there is one key per OST image, so there will be additional directories and objects related to the keys.

Additionally, with traditional duplication, NetBackup supports tiering data (not metadata) across the different levels of Amazon S3 object storage classes using Amazon S3 SLPs. For instance, tiering data from S3 Standard to Glacier or S3 Standard to S3 Standard Infrequent Access is available and supported by NetBackup. However, Amazon SLPs are only supported for instances where Accelerator (the feature that provides full backups with reduction of the backup window and compute and storage resources of an incremental backup) and MSDP-C are not used to send data directly to the AWS cloud. So, for data that does not require deduplication, NetBackup has parameters such as UPLOAD_CLASS, TRANSITION_TO_STANDARD_IA_AFTER and TRANSITION_TO_GLACIER_AFTER to integrate with AWS's lifecycle cloud tiering. For example, backup data can first reside in an S3 Standard storage class and then after 30 days transition to S3 Standard Infrequent Access and then after 60 days transition to Glacier, as illustrated in Figure 15. **NOTE**: *If transitioning to Glacier, check to make sure Glacier is supported for the region to which the bucket belongs. Also, Amazon S3 SLPs are not supported with NetBackup Accelerator.*



***NOTE:** Amazon Storage Lifecycle Policies not supported with MSDP-C and Accelerator

Figure 15. An overview of NetBackup data retention with Amazon S3 storage lifecycle policies (SLPs).

Migration to the Cloud

Migration to the cloud or initial seeding of data into the cloud is simple when using NetBackup with AWS Snowball and AWS Snowball Edge. These devices are placed on-premises to handle large-scale data transfers quickly without overloading the corporate network. NOTE: *This process is not currently supported with MSDP-C.* As illustrated in Figure 16, the process involves the following:

1. From the NetBackup administration console, configure a cloud storage server and create Amazon S3 Standard class as the target storage. Create the bucket in the same region the AWS Snowball or AWS Snowball Edge is available.
2. On the AWS management console, create an import job specifying the bucket created in the previous step and indicate the region and shipping information.
3. The AWS Snowball or AWS Snowball Edge is shipped by courier to the data center where the NetBackup domain resides.
4. The device is connected to the same network as the NetBackup domain. As previously discussed, the tools required for transfers using AWS Snowball are downloaded from the AWS resources website onto a server and the tools for AWS Snowball Edge are already on the device. Once the devices are configured, a cloud storage server is created in NetBackup with the device's properties such as IP address and type of transfer: S3 or Filesystem (NFS). A backup policy and/or SLP is defined to transfer data to AWS Snowball or AWS Snowball Edge.
5. After all the data is transferred, the policy is postponed or deactivated in NetBackup and the device is disconnected and shipped back to the AWS data center via courier.
6. Once the device is received at the AWS data center, the data is transferred to the destined S3 Standard class (bucket). You can view or monitor the progress of the transfer on the AWS console.
7. After all the data has been uploaded, the backup policy and/or SLP can be resumed on NetBackup.

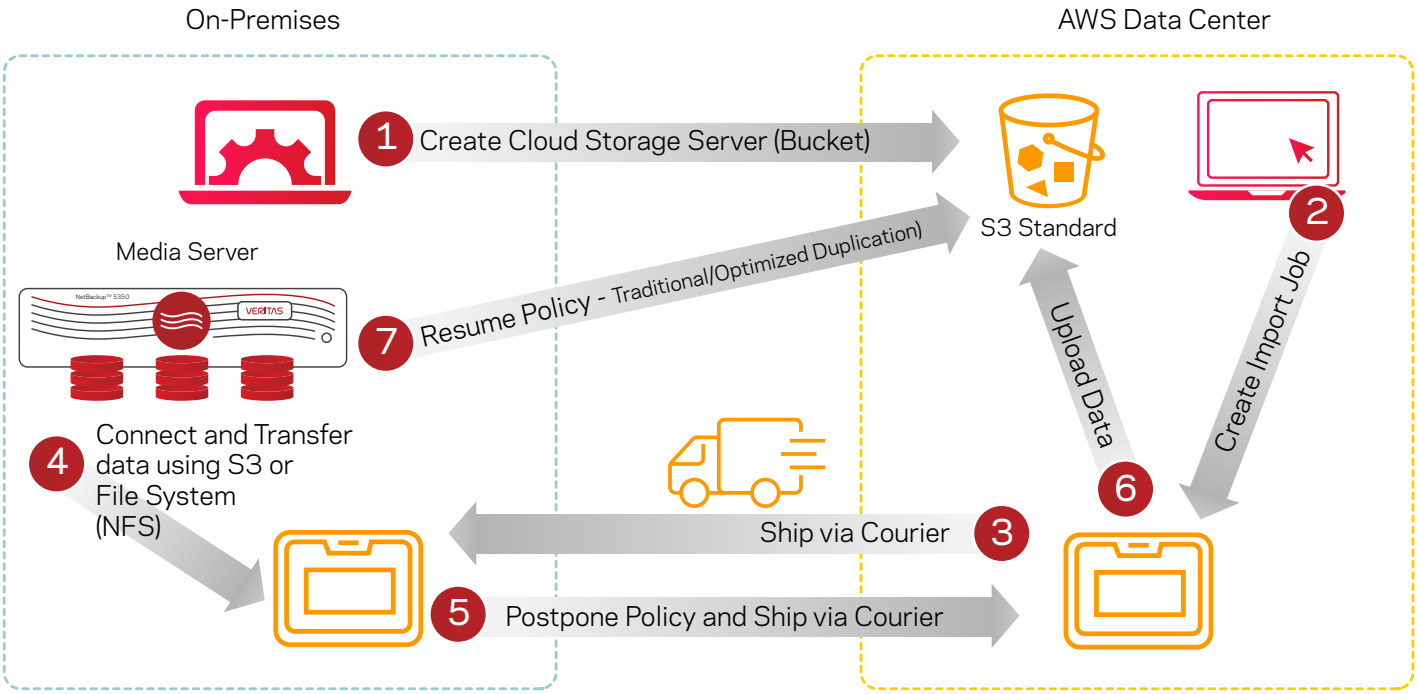


Figure 16. An overview of the migration to the cloud data flow using NetBackup with AWS Snowball or AWS Snowball Edge.

Disaster Recovery in the Cloud

In scenarios where the on-premises NetBackup environment is lost due to a power outage, natural disaster or catastrophic event, a feature called "Image Sharing" allows for recovery of the backup images in the cloud. Using MSDP-C with this feature allows for creation of "self-describing" images where both the data and the metadata of the backup image are sent to the AWS cloud. As shown in Figure 17, when the on-premises NetBackup environment become unavailable, the recovery flow is as follows:

1. A NetBackup instance(s) is created on AWS.
2. The image is read from the Amazon S3 Standard cloud storage class and imported to re-create the NetBackup catalog.
3. The image(s) is recovered within the cloud.

You can manage this feature via the command line or the web UI. For a list of workloads Image Sharing supports, please refer to the NetBackup Hardware and Cloud Storage Compatibility List and see to the "Support for Workloads with Image Sharing in the Cloud" section. **NOTE**: *This feature is qualified for data stored only in an Amazon S3 Standard storage class (bucket). It is not supported with S3 Glacier or S3 Glacier Deep Archive. Also, from the NetBackup instances in the cloud, the S3 Standard bucket is configured as "read-only."*

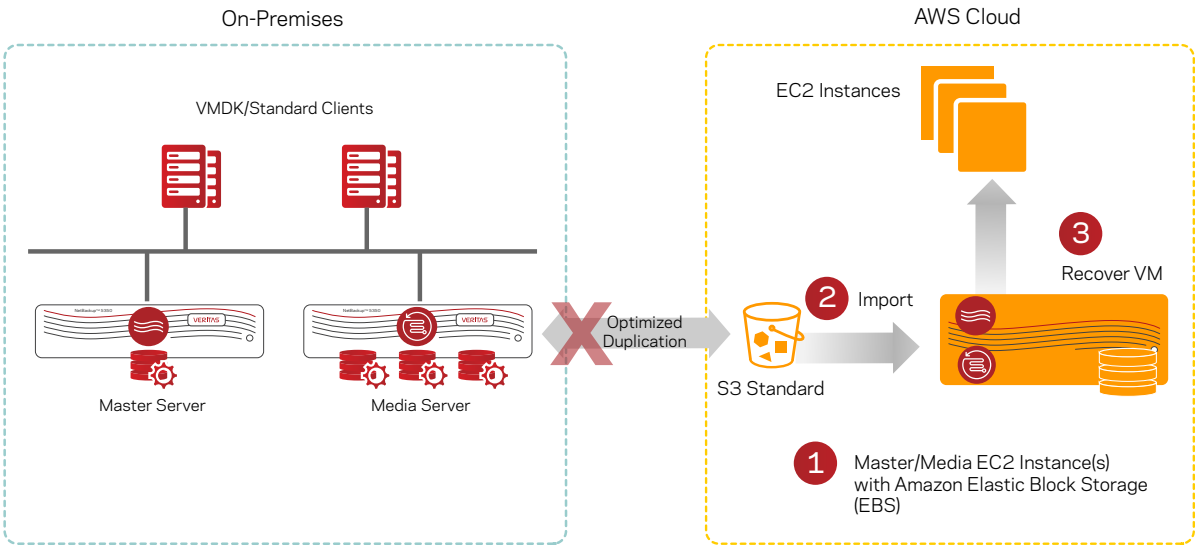


Figure 17. An overview of automated disaster recovery in the AWS cloud using NetBackup Image Sharing.

Best Practices and Recommendations

The following best practices are important in creating an optimum deployment. This section highlights some best practices related to using AWS cloud storage as a long-term retention solution for NetBackup. Some of these best practices are covered in-depth in the following documentation:

- NetBackup Cloud Administrator's Guide
- AWS Snowball and Snowball Edge Best Practices

Privacy Laws

Certain countries have employed the European Union's General Data Protection Regulation (GDPR) and privacy laws specifically related to data stored in the public cloud. Although NetBackup stores data in a backup image format, the data can still be read or scanned for PII, confidential and business-critical information. Use tools such as Data Insight to decide what information is best stored on-premises versus in the public cloud to reduce your organization's liability and violation of these laws. Also be aware of restoring the data in the cloud when using NetBackup's automated DR, "self-describing images." For example, if data is backed up in the Europe region, the NetBackup cloud instances and data restored in the cloud should be in the same region or adhere to your organization's GDPR policies. Use of NetBackup encryption would also help address some of the concerns relating to privacy.

Compression

For better storage utilization, using NetBackup compression might be an option when deduplication is not ideal and the data type being backed up is compressible. Although compression can reduce the size of a backup, it also can consume server resources. As a best practice, the media server should be sized appropriately for compression. For detailed information on NetBackup compression attributes and considerations, refer to the NetBackup Administration Guide, Volume I, and for information on compression for cloud storage targets and deduplication, refer to the NetBackup Cloud Administrators Guide and the NetBackup Deduplication Guide, respectively.

Deduplication

If more than 250 TB of MSDP is required, use the NetBackup Appliances as opposed to the BYOS version. NetBackup BYOS has a MSDP size limitation of 250 TB, whereas the size of MSDP on an Appliance can be up to 323 TB or 1.2 PB, depending on the model. The maximum MSDP capacity depends on the NetBackup Appliance used as a media server.

For MSDP-C, multiple buckets are supported per storage server. By default, the storage server defines cache properties in the contentrouter.cfg file, with a total of 1 TB. Therefore, as a rule of thumb, configure 1 TB of disk storage per cloud target when configuring the storage server.

AWS Snowball/AWS Snowball Edge

Following are some of the best practices when using AWS Snowball and/or AWS Snowball Edge:

- For AWS Snowball Edge, use the S3 Adapter to transfer data as oppose to the File Interface. Performance using the S3 Adapter is significantly better compared to the File Interface.
- Keep the copies on-premises until the data has been fully imported into Amazon S3 storage and has been verified.
- Create the bucket within the NetBackup administration console to conform to the bucket naming convention.

NetBackup Retrieval Attributes

There are several NetBackup attributes that are best to use when sending to or restoring data from Glacier, Glacier Vault and/or Glacier Deep Archive. For instance, it is best to use the option True Image Restore whenever possible. Enabling this option significantly reduces the retrieval or restore of data from hours to minutes in some cases. With True Image Restore, NetBackup keeps track of information and only restores the files that were present at the time of the last backup as opposed to restoring all files.

Another useful attribute is RETRIEVAL RETENTION PERIOD to implement with Glacier and Glacier Deep Archive only. Set this value within the NetBackup storage server to a minimum of three days. Setting the value to days allows the data to be kept within the warming area for scenarios where it might take several retries to pull data due to network issues, contention and other failures.

Conclusion

Veritas products along with AWS cloud storage services provide a long-term retention solution for data protection and management of data. AWS offers varying storage classes at different levels of availability, cost and restore to meet the various needs of organizations. Data Insight is instrumental in identifying data that can be safely moved to the cloud to minimize risk and liability. Using AWS cloud storage with NetBackup provides a compelling option for protecting and preserving data from on-premises challenges such as failures, attacks and disasters as well as for off-premises, long-term retention.

References

- NetBackup
  - Product Documentation
    https://www.veritas.com/support/en_US/article.DOC5332
  - NetBackup Deduplication Guide
    https://www.veritas.com/content/support/en_US/doc/25074086-146020141-0/v26653261-146020141
  - NetBackup Cloud Administrator's Guide
    https://www.veritas.com/content/support/en_US/doc/58500769-145530841-0/v58383369-145530841
- Access Appliance
  - Product Documentation
    https://sort.veritas.com/documents/doc_details/AAPP/7.4.2/Veritas%203340/Documentation/
  - White Papers/Data Sheets
    https://www.veritas.com/protection/access-appliance/resources
- AWS Cloud Storage
  - AWS S3 Storage Classes
    https://aws.AWS.com/s3/storage-classes/
  - AWS Snowball
    https://docs.aws.AWS.com/snowball/latest/ug/whatissnowball.html
  - AWS Snowball Edge
    https://docs.aws.AWS.com/snowball/latest/developer-guide/whatisedge.html
- Data Insight
  - Product Documentation
    https://sort.veritas.com/sitemap/document/23

Appendix

This section provides an example of how to configure NetBackup to send deduplicated data to the cloud. This is only a sample deployment, and we recommend readers refer to Veritas product documentation and AWS for definitive and specific installation, administration and configuration details.

The example shown in Figure 18 uses a BYOS NetBackup environment that consists of a primary server and media server configuration. There is a single MSDP storage server with two MSDP disk pools: a local MSDP disk pool for the first copy of the backup image and an MSDP disk pool targeting the AWS S3 Standard bucket for the second copy. A protection plan will be created to do a backup to the MSDP local disk pool and then duplicate to the MSDP cloud disk pool.

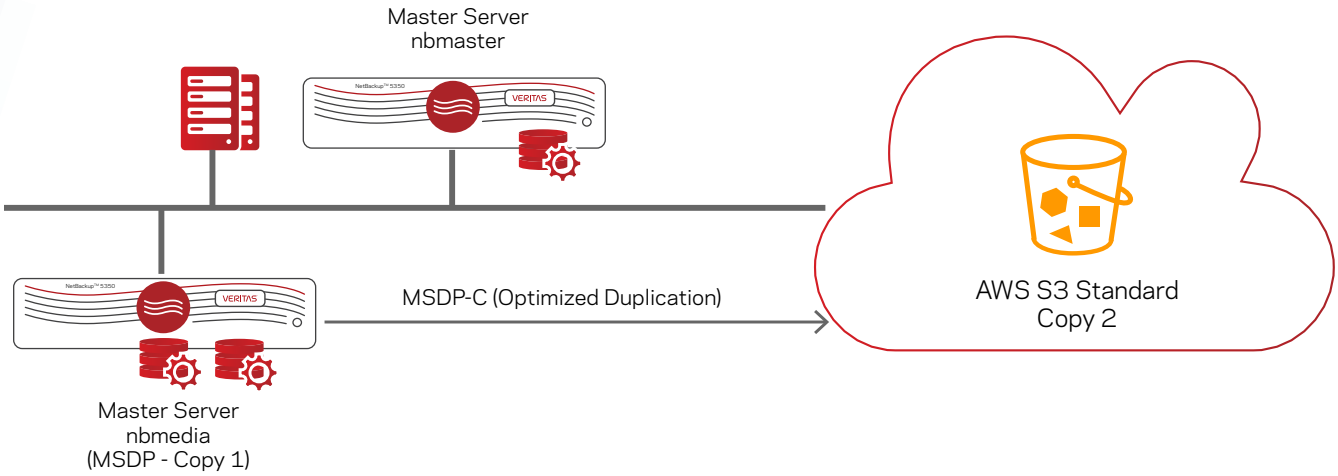
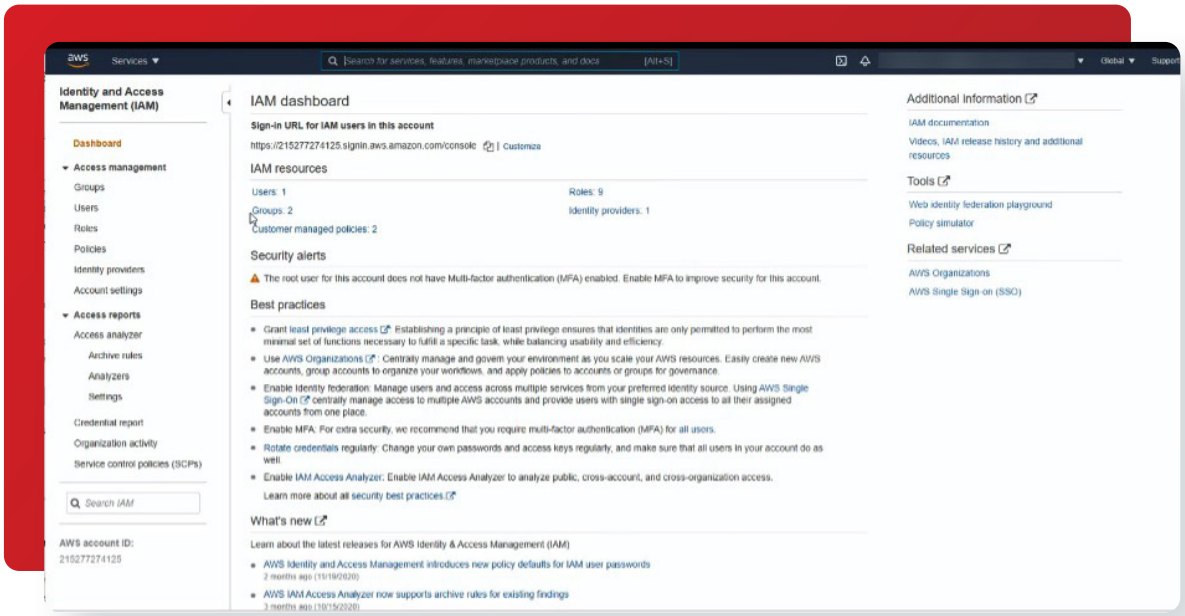


Figure 18. An example NetBackup configuration for sending deduplicated data to the cloud.

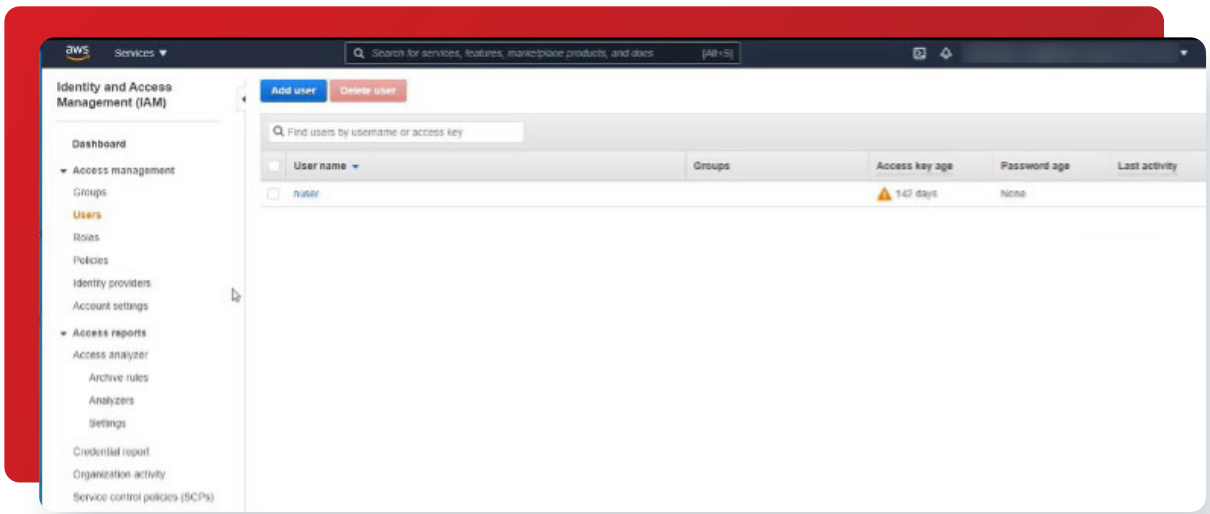
It is assumed the NetBackup components are configured and installed. The example consists of the following main steps:

1. Create a user, access and secret keys on the AWS management console.
2. Create the NetBackup storage server, creating a disk pool for MSDP local storage and creating a disk pool for the S3 Standard storage class as the cloud target.
3. Create a protection plan where backup and deduplication are done on a media server and placed in MSDP local storage, then duplicated to the cloud storage server (S3 Standard) created in the previous step.
4. Attach the protection plan to the VM. Run a manual protection backup.

    1. Log on to the AWS web console GUI. Open the IAM dashboard and click on **Users** under Access management on the left pane.

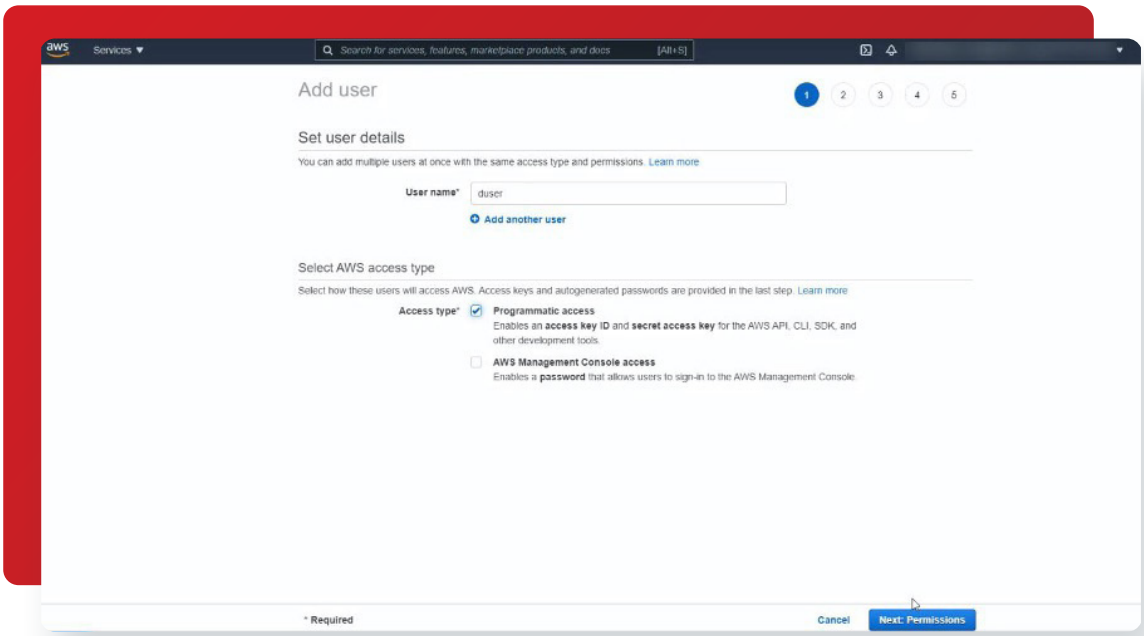


2. Step 2) Click on **Add user**.



5. Verify the backup and duplication.

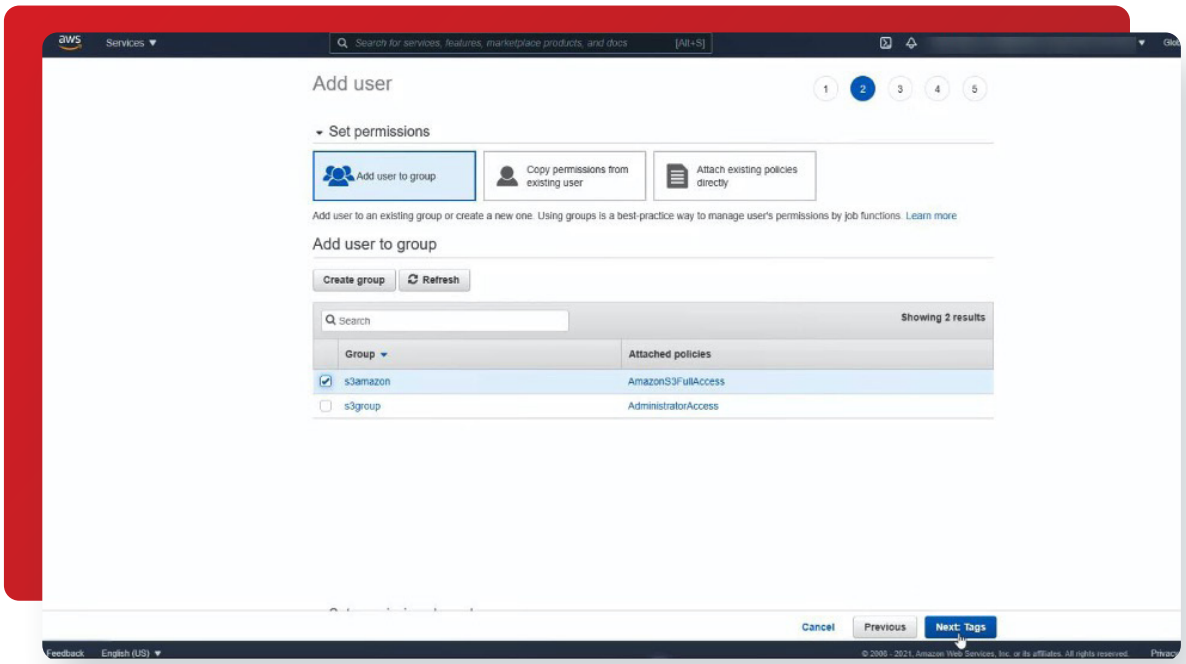
3. Enter the **username** and place a checkmark in the Access type on **Programmatic access**.



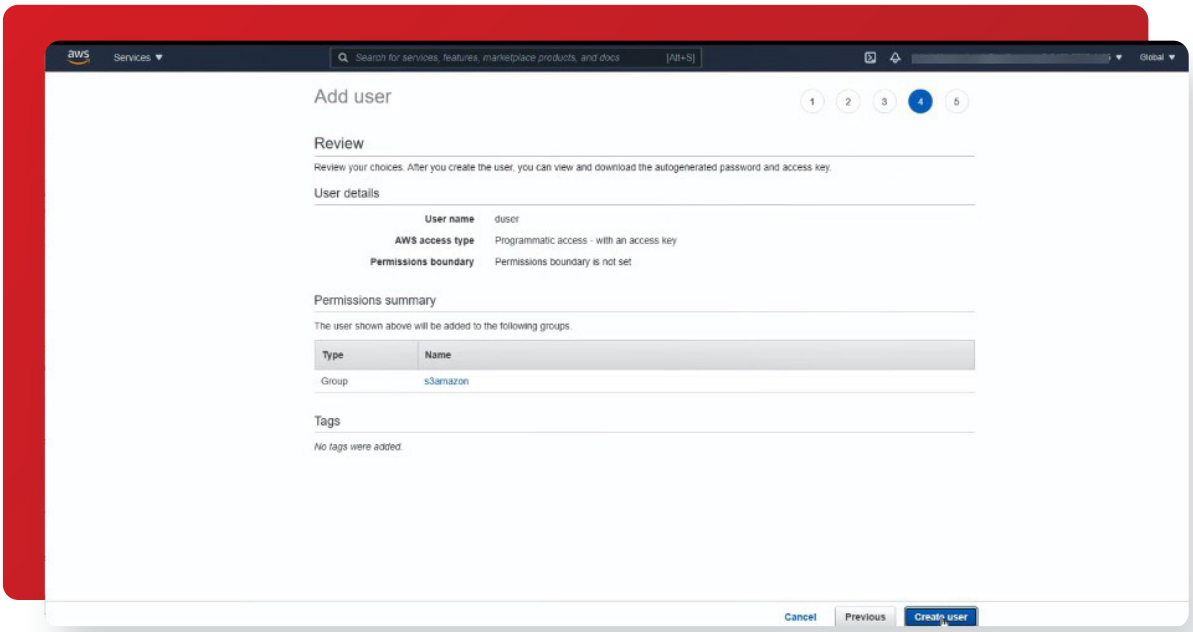
AWS Security and Access Keys

On the AWS management web console, create a user with an associated Access Key ID and Secret Access Key. These keys are used when configuring the AWS cloud storage server within the NetBackup administration console.

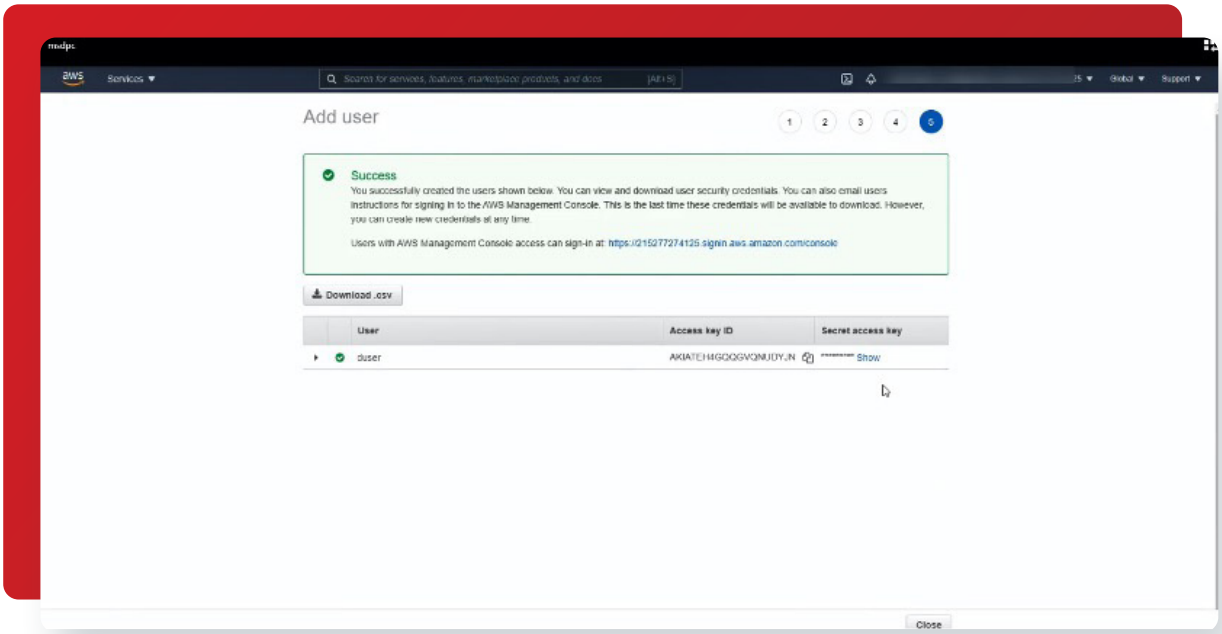
4. Click **Next Tags** and then click **Next: Review**. In this example, the s3AWS group selected has AWSS3FullAccess. For details on the minimum permissions required, refer to the Veritas NetBackup Cloud Administrator's Guide.



5. Click on **Create user**.

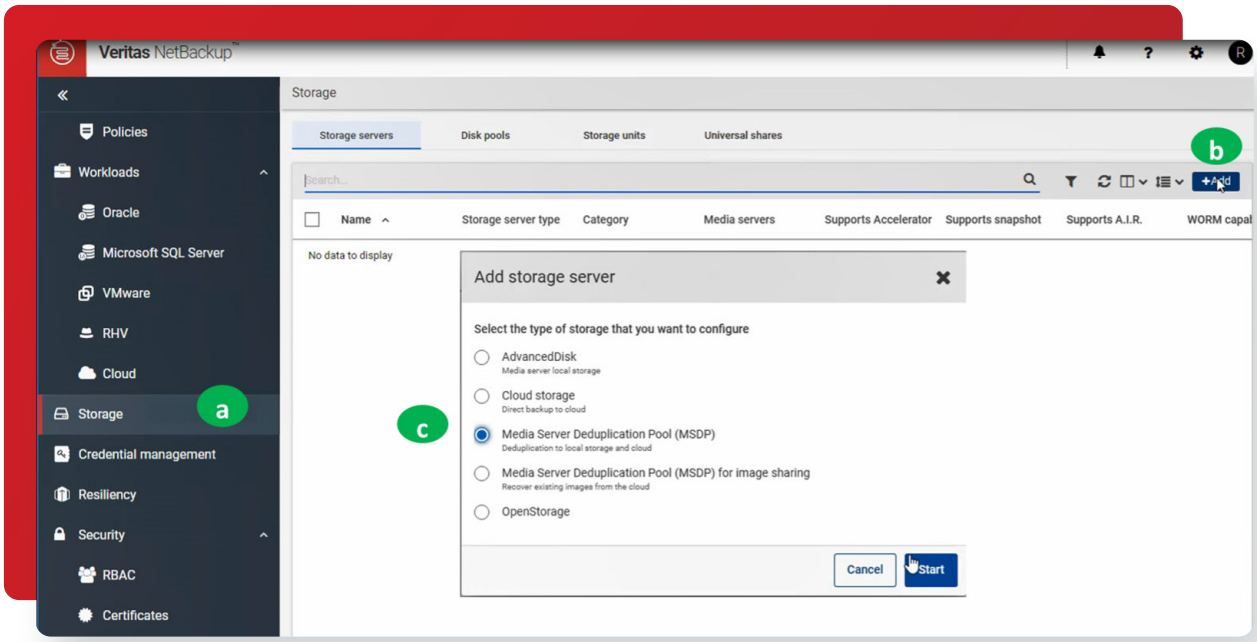


6. Save the **access key and secret key** information in a file.

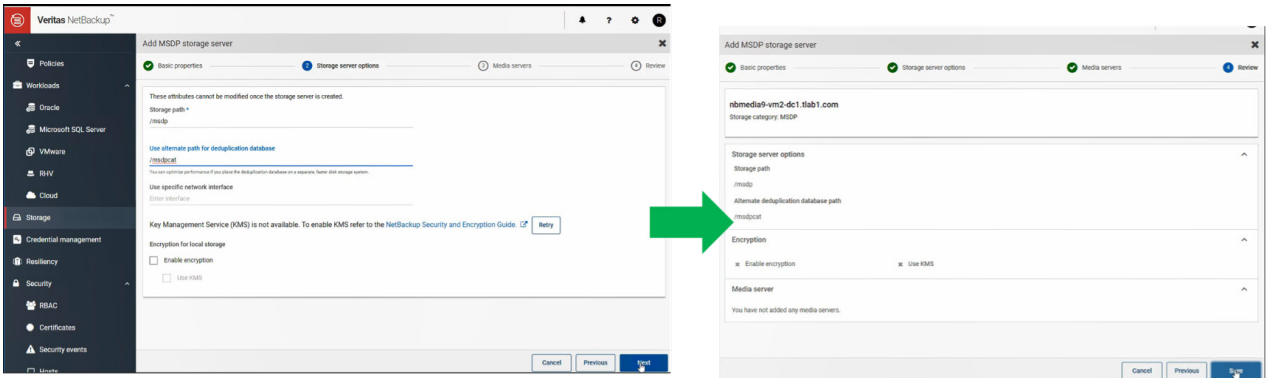


Creation of NetBackup Storage Server, Disk Pools and Storage Units for MSDP Local Storage and MSDP-C Cloud

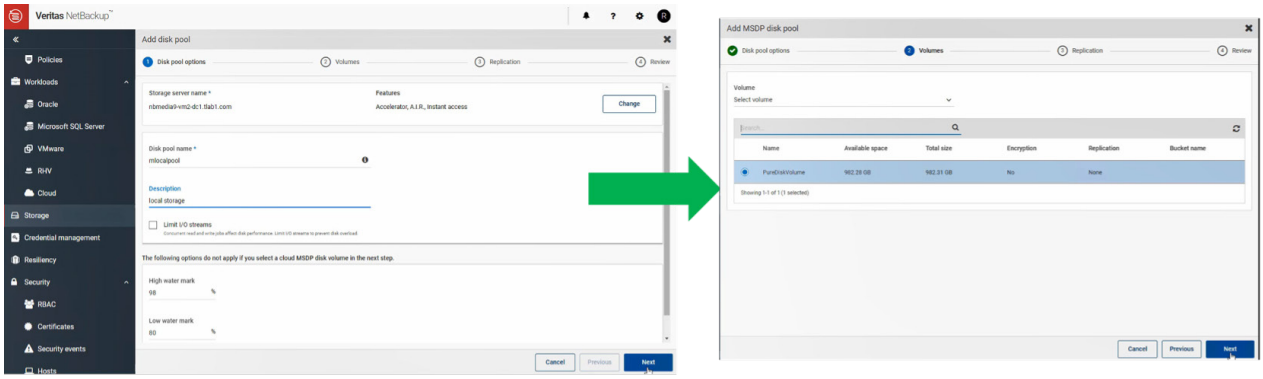
1. Log on to the NetBackup web UI (https://FQDNofMaster/webui) to create the storage servers. Click on **Storage** on the left pane and click **Add**. Select **Media Server Deduplication Pool** to create an MSDP storage server.



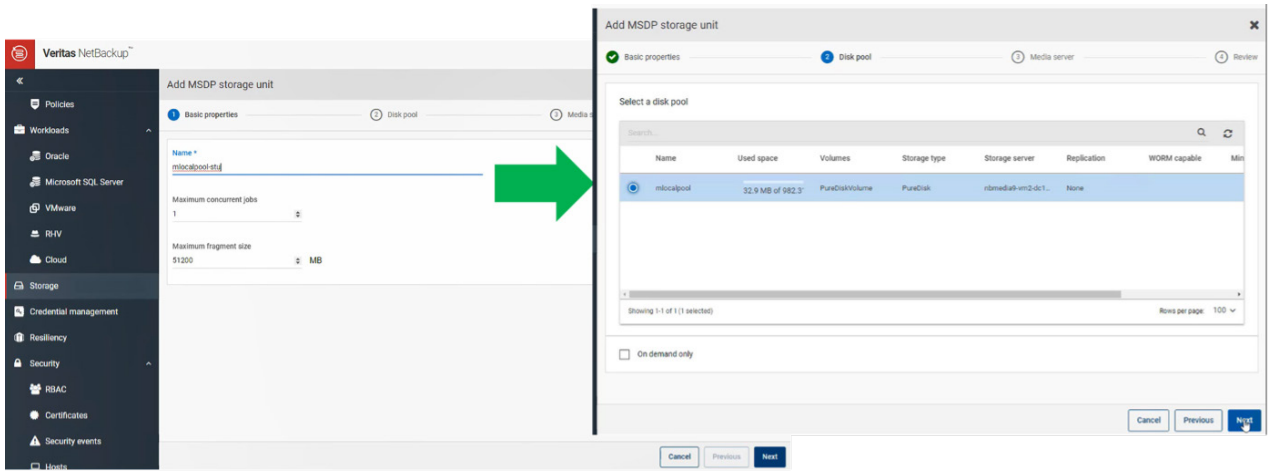
2. Specify the **storage disk path** (for example, /msdp) and optional **alternate deduplication database path** (for example, /msdpcat). Click **Next**, review the entries and click **Save**. NOTE: *The storage disk needs to be a minimum of 1 TB.*



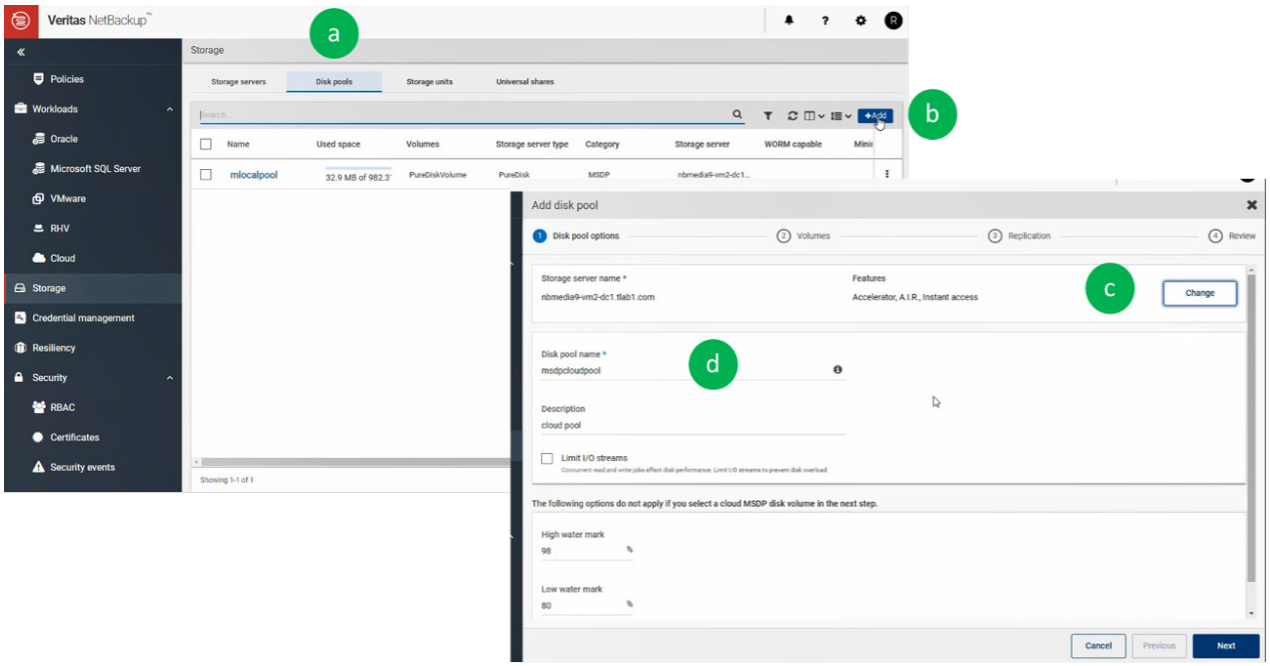
3. Add a local MSDP disk pool. Specify the **disk pool name** and select **PureDiskVolume**, click **Next**, review the entries and click **Finish**.



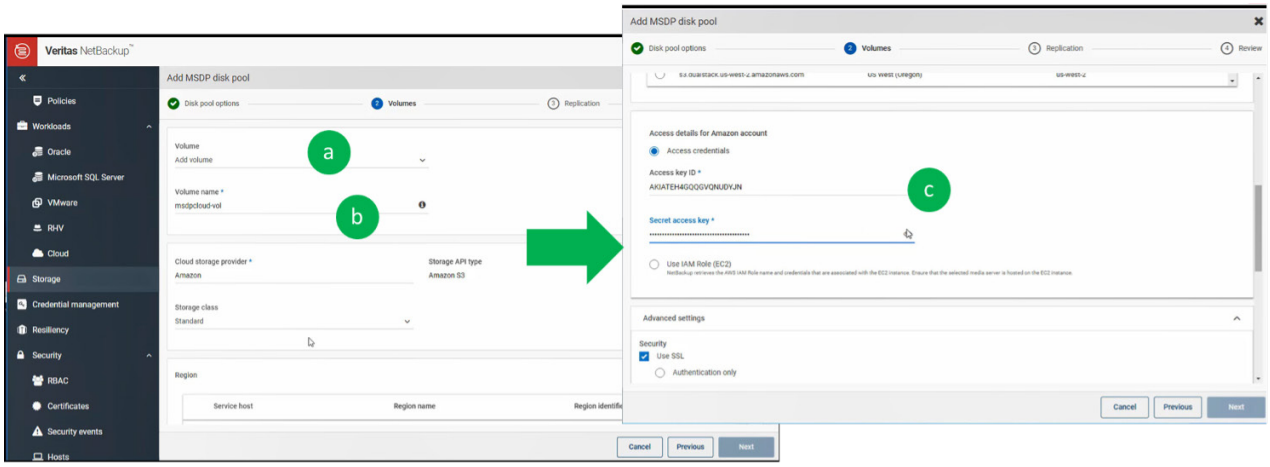
4. Add a storage unit associated with the local MSDP disk pool. Specify the disk pool **name**, click **Next** and in next pane, select the **disk pool** created and click **Next**. Review the entries and then click **Save**.



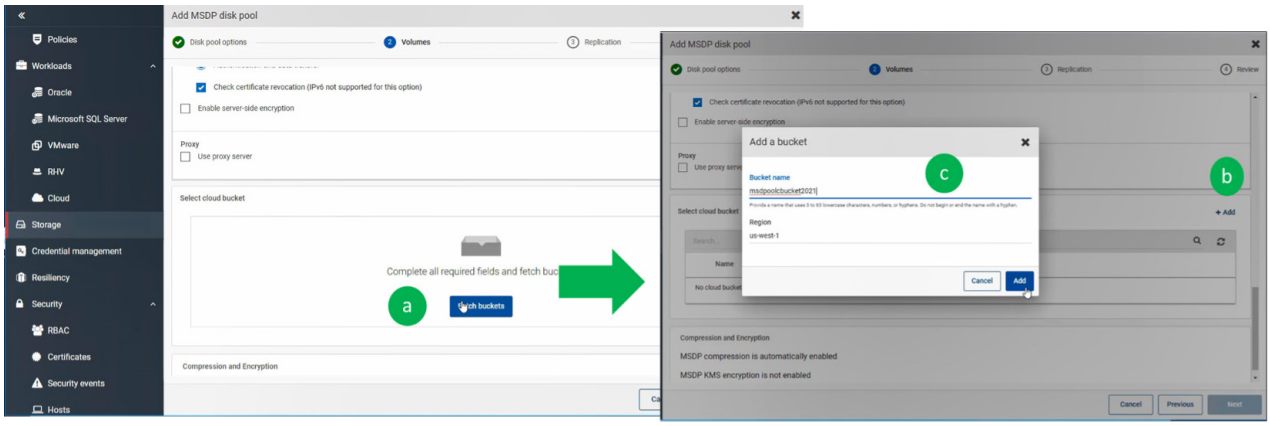
5. Create an MSDP cloud disk pool. Click on the **Disk Pool tab** in the top header row. Click **Add**. On the next pane, click **Change** and select the **storage server name** created in Step 1. Enter the **disk pool name** and click **Next**.



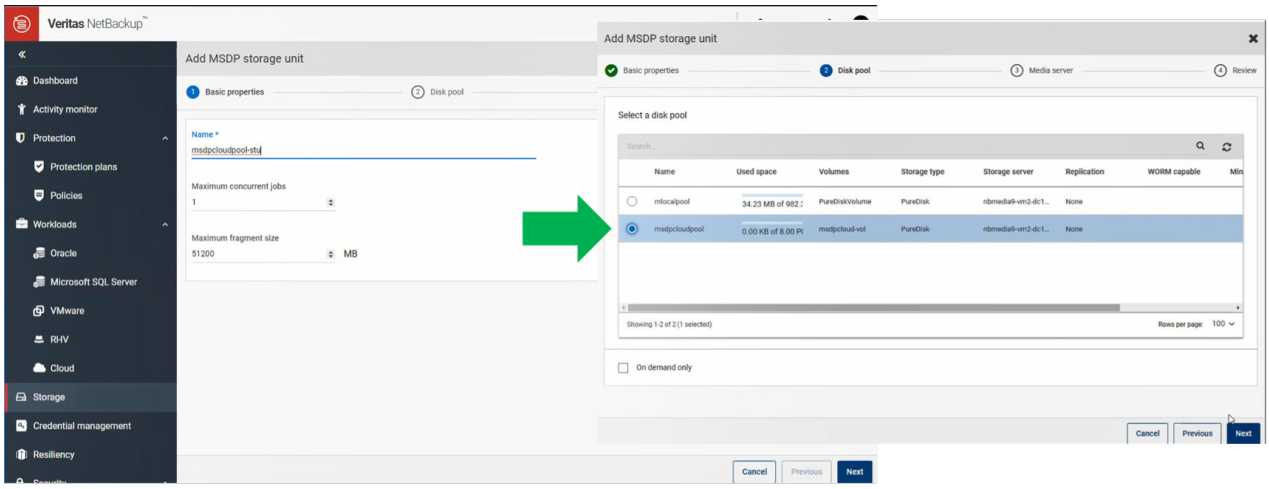
6. Click on the Volume dropdown and select **Add volume**. Specify the **volume name**, the cloud storage provider and the storage class. Scroll down and specify the **region** and enter the **Access and Secret key**.



7. Scroll down further to select the cloud bucket. Click on **Fetch buckets**, then click **Add**. Enter a **unique Bucket name** and click **Add**. Select the **new bucket created** and click **Next**, Next, review the entries and then click **Finish**.

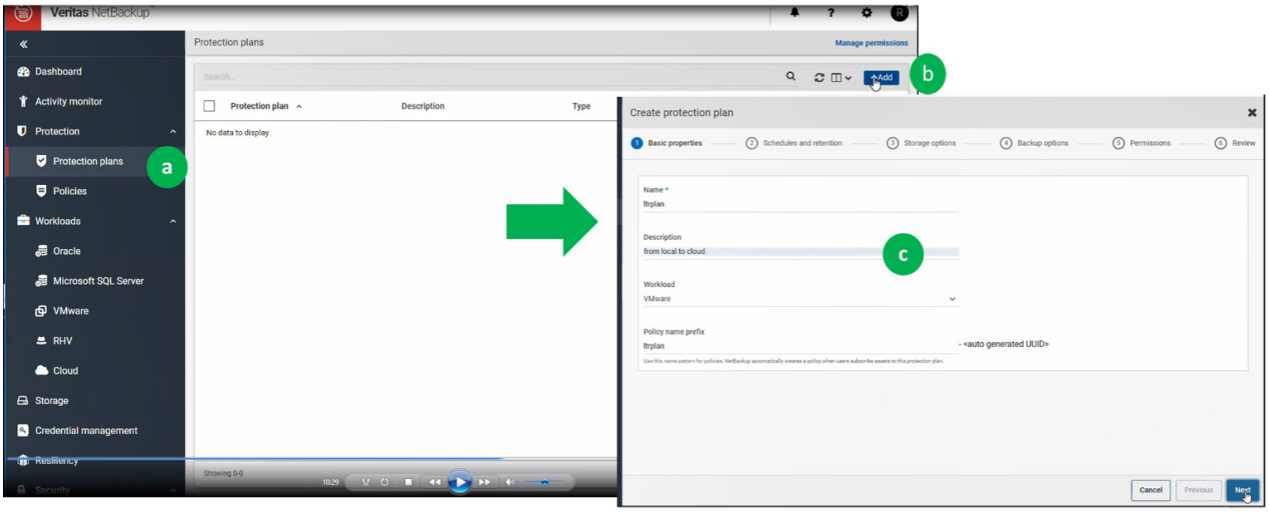


8. Add the storage unit associated with the MSDP cloud disk pool. Enter the **name** of the storage unit, click **Next** and in following pane, select the MSDP cloud disk pool you created in the previous steps. Review the entries and then click **Save**.

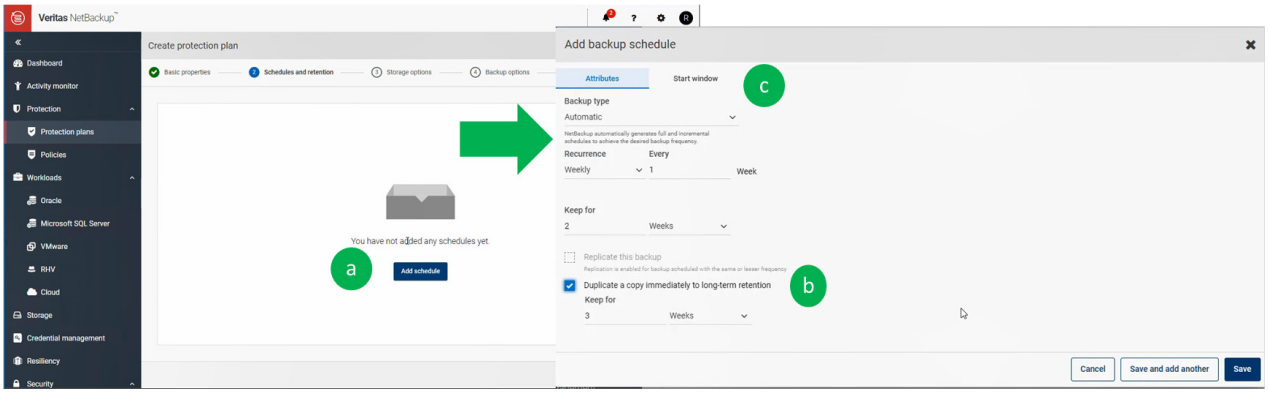


Creation of the Protection Plan

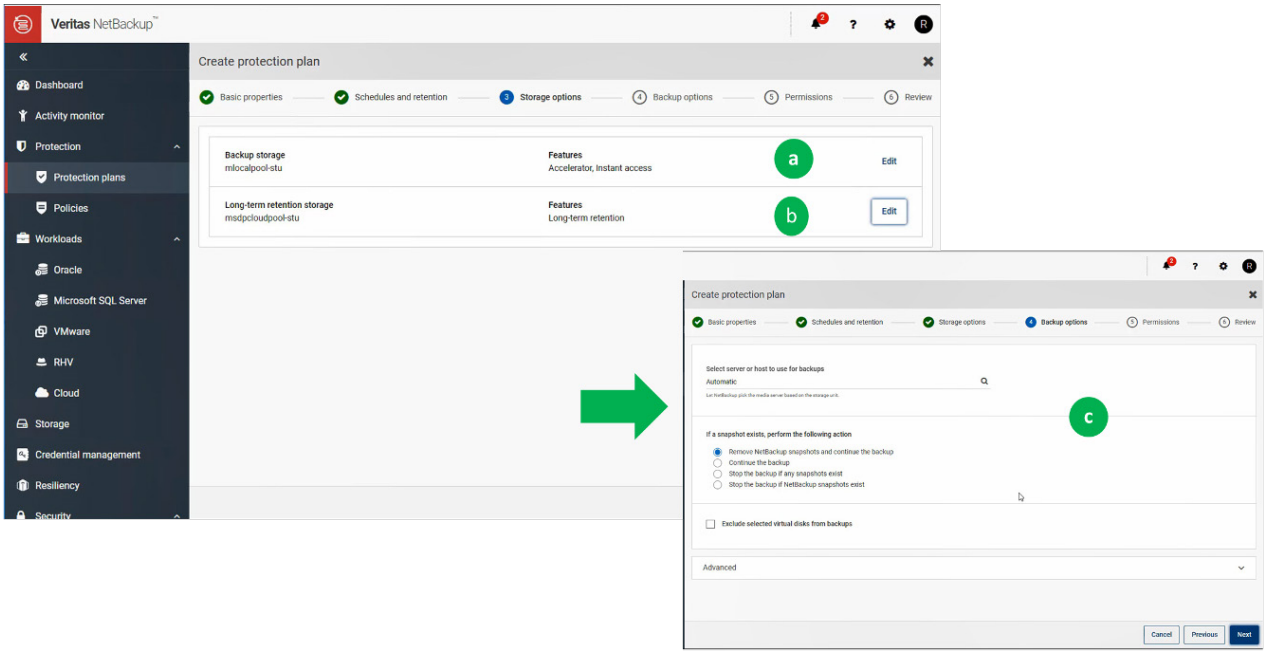
1. Click on **Protection** on the left pane to expand it and then click on **Protection Plans**. Click the **Add** button on the top, specify the **name of plan** and then click **Next**.



2. Click on **Add schedule**. In next pane, select the backup type, recurrence and retention. Click on **Duplicate a copy immediately to long-term retention**. Click on **Start Window tab** and enter the backup windows. When you are done, click **Save**.

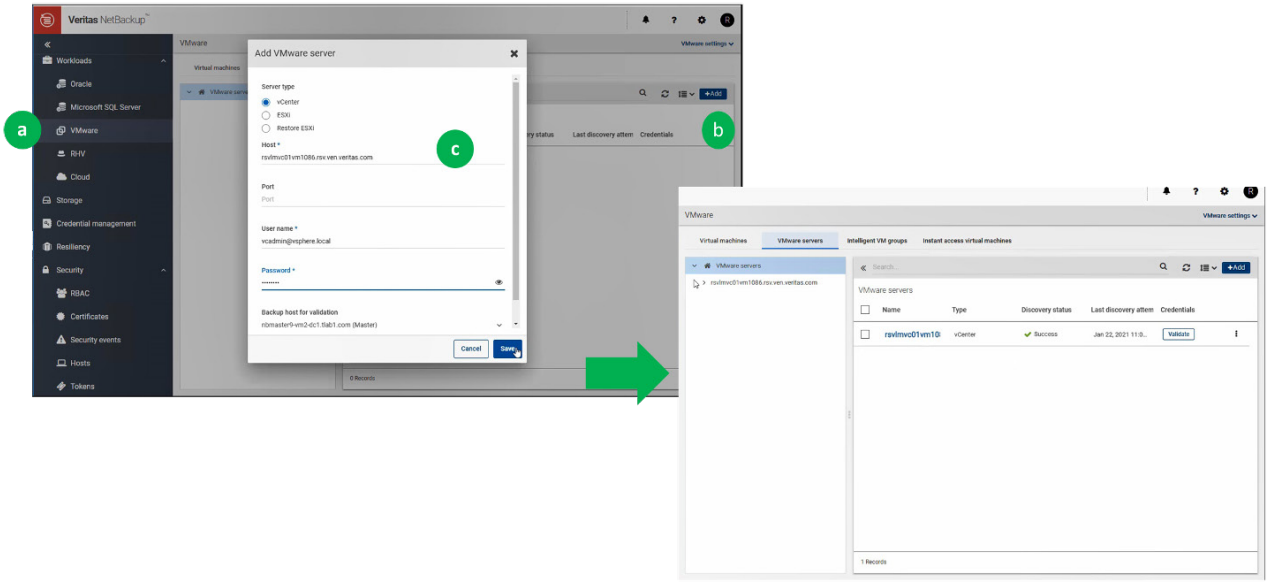


3. In next screen, click **Edit** to select backup storage and then **select MSDP local disk pool**. Click **Edit** to select Long-term retention storage and select **MSDP cloud pool**. Click **Next** and select Backup options. Click **Next** to specify permissions. Review entries and click **Finish**.

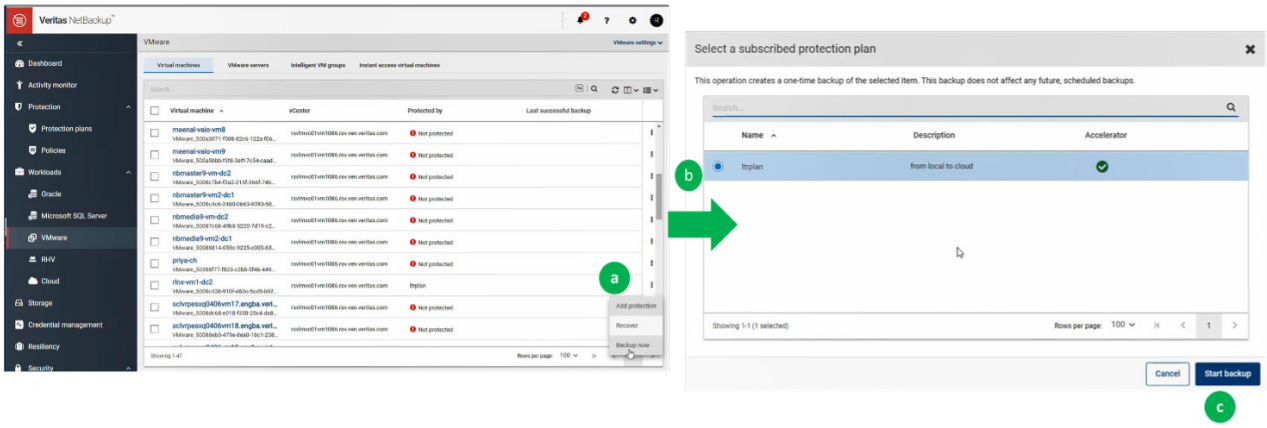


Validation of Setup

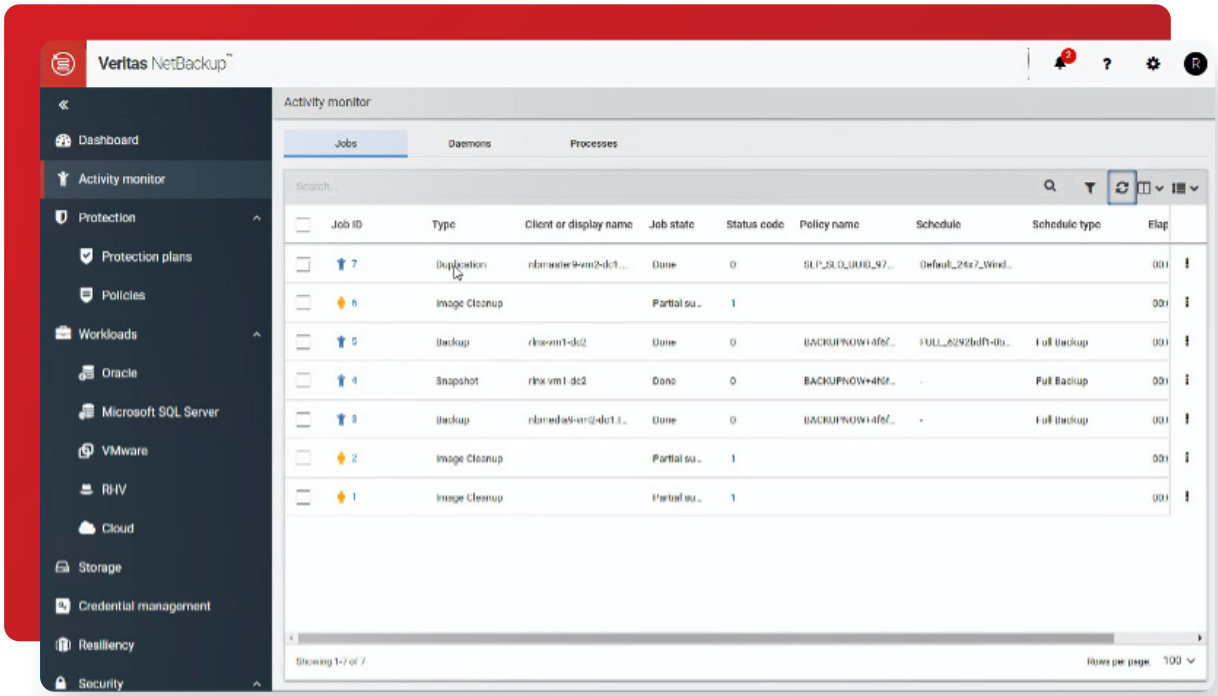
1. Configure a workload to protect. In this example, configure VMware as the workload to protect. Click on **Workloads** on the left pane to expand it and select **VMware**. Enter the VMware host name and credentials. Click **Save**.



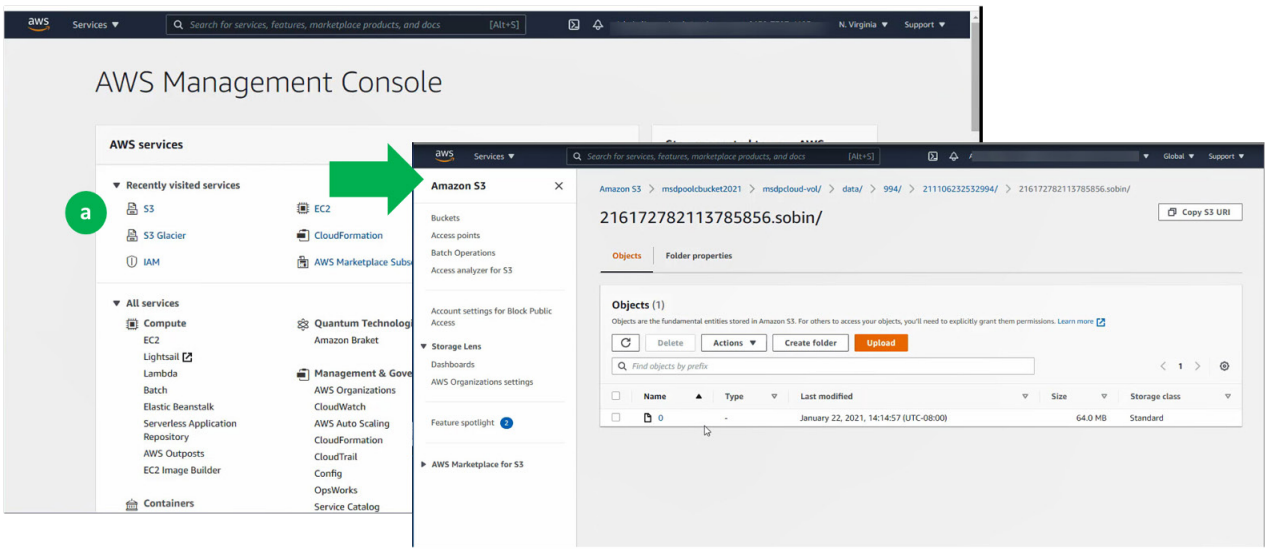
2. Click on the **Virtual machines** tab on the top. Select **VM** to do a manual backup. Click on the three **dots** on the right-hand side of the selected VM and select **Backup Now**. Select the **protection plan** you created in the previous step and click **Start Backup**.



3. Step 3) Click on **Activity monitor** on the left pane and monitor the jobs. As you can see, a backup first occurs and is stored on MSDP local storage, then a duplication job is conducted to duplicate to MSDP cloud storage.



4. Log on to the AWS Management Console. Click on the **S3** service and select the destination bucket and traverse to validate the duplication occurred.



About Veritas

Veritas Technologies is a global leader in data protection and availability. Over 80,000 customers—including 87 percent of the Fortune Global 500—rely on us to abstract IT complexity and simplify data management. The Veritas Enterprise Data Services Platform automates the protection and orchestrates the recovery of data everywhere it lives, ensures 24/7 availability of business-critical applications, and provides enterprises with the insights they need to comply with evolving data regulations. With a reputation for reliability at scale and a deployment model to fit any need, Veritas Enterprise Data Services Platform supports more than 800 different data sources, over 100 different operating systems, more than 1,400 storage targets, and more than 60 different cloud platforms. Learn more at www.veritas.com. Follow us on Twitter at @veritastechllc.

VERITAS

2625 Augustine Drive
Santa Clara, CA 95054
+1 (866) 837 4827
veritas.com

For global contact information visit:
veritas.com/company/contact

Copyright © 2022 Veritas Technologies LLC. All rights reserved. Veritas, the Veritas Logo, and NetBackup are trademarks or registered trademarks of Veritas Technologies LLC or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

V1259 05/22