VERITAS

NetBackup in the Public Cloud

Guidelines for Amazon Web Services (AWS) Deployments.

This technical paper is designed to aid partners and customers looking to protect workloads with Veritas NetBackup™ deployed in the AWS cloud. The guidelines in this paper will assist in designing and implementing data protection solutions based on Veritas products in the public cloud. In addition to these guidelines, partners and customers should also use product documentation, Veritas Educational Services and/or Veritas Consulting Services when necessary.

For the latest in cloud data protection using the NetBackup platform, visit veritas.com/netbackupcloud.

White Paper | September 2022

Contents

Introduction . . . 4

Prerequisites . . . 4

NetBackup Overview . . . 4

- Key Capabilities . . . 4
- Key Features . . . 4

Business Value . . . 5

Why are Customers Using the Cloud? . . . 5

NetBackup and AWS Design Overview . . . 5

- Versatile Deduplication Engine . . . 6
- NetBackup and Cloud Connectivity . . . 6
- NetBackup and Cloud Restore Options . . . 6
- AWS Cloud Versus On-Premises Considerations . . . 6

Use Cases Overview . . . 7

- Standard Backup to AWS S3 Object Storage . . . 7
- Backup to AWS S3 Object Storage with MSDP-C Deduplication . . . 7
- Sending Data Into AWS S3 using AWS Snowball Or Snowball Edge . . . 8
- Sending Data to AWS S3 Using a Third-Party Gateway Appliance . . . 9
- Amazon Storage Gateway Virtual Tape Library . . . 10
- Backup in The Cloud – AWS EC2-Based Architectures . . . 10

Disaster Recovery Using AWS . . . 12

- Auto Image Replication (AIR) to The Cloud – Hybrid Configuration . . . 12
- Leveraging NetBackup Image Sharing for Migration and DR . . . 13
- Recovering Virtual Machines from VMware to AWS Using NetBackup Automated DR . . . 14

NetBackup Intelligent Cloud Policies . . . 14

Immutable Storage in AWS . . . 14

Backup from SnapShot . . . 15

Storage Lifecycle . . . 16

Cost Considerations . . . 16

- NetBackup Cloud Autoscaling . . . 16

Cloud Sizing and Performance . . . 17

- AWS Instance Model . . . 17
- AWS Storage Options . . . 17
- Environment Description and Assumptions for Sizing . . . 18

Contents

NetBackup AWS Instance Sizing . . . 18

NetBackup Management Server . . . 19

Management Server Memory and CPU Requirement . . . 19

Management Server Recommendations – AWS EC2 Sizes . . . 19

NetBackup MSDP Storage . . . 19

Growing the Media Server . . . 20

Media Server Deduplication Pool Recommendations . . . 20

MSDP-C Server – Configuration Walk-Through . . . 22

Additional Architecture Requirements . . . 23

Security of the Information . . . 23

Least Privileged Access . . . 23

Limit Access with Resource Tags . . . 24

About AWS Service Quotas . . . 24

AWS Service Quotas and NetBackup . . . 24

Disaster Recovery Scenarios . . . 25

Cost Overview . . . 25

The Cost of Gets and Puts . . . 25

Storage Costs . . . 25

Compute Costs . . . 25

Deployment Details . . . 25

NetBackup VPC Deployment Configurations . . . 25

Protecting NetBackup Access With EC2 Security Groups . . . 26

Tagging NetBackup Resources . . . 26

Rotating AWS Access Keys for MSDP-C . . . 29

Protecting NetBackup from Faults, Failures and Downtime . . . 30

NetBackup Risk and Audit Management . . . 31

Enabling CloudTrail Logging for NetBackup Resources . . . 31

AWS Scheduled Service Events . . . 32

Summary . . . 32

Appendix A – Additional Information . . . 32

Appendix B – Terminology . . . 33

Disclaimer . . . 33

Introduction

The purpose of this white paper is to provide a technical reference on the capabilities of Veritas NetBackup and Amazon Web Services (AWS). Although this guideline is a stand-alone document, you can find additional information using the links in the Additional Resources section. This document is not a replacement for the NetBackup Cloud Admin Guide, links to which are at the end of this document.

Veritas has partnered with AWS to offer a robust backup and recovery experience to the cloud and in the cloud. Each solution can be tailored to the individual needs of customers.

NOTE: This document contains recommendations that have been shown to work with customer deployments. Because every environment is unique, changes may be required. In addition to these guidelines, you should always consult product documentation and use any additional services (education or consulting) to ensure the best design for unique environments and workloads.

Prerequisites

This document is intended for individuals with a basic understanding of AWS Cloud infrastructure concepts. Users should be familiar with AWS CloudFormation, AWS Identity and Access Management (roles and policies), Amazon S3 (bucket policies and access control lists), Amazon Elastic Compute Cloud (EC2), Amazon VPC, and storage concepts related to enterprise backup and recovery solutions.

NetBackup Overview

As an established market leader in data protection, Veritas provides unparalleled next-generation data protection by minimizing costs and complexity. With NetBackup, a solution that unifies data protection across the entire enterprise, Veritas also ensures greater business continuity.

Key Capabilities

- **Comprehensive**—As a single solution to protect all your data assets, NetBackup provides support for virtually every popular server, storage, hypervisor, database and application platform used in the enterprise today.
- **Scalable**—High performance, elastic automation and centralized management based on a flexible, multi-tier architecture enables NetBackup to adapt to the growing needs of a fast-paced, modern enterprise data center.
- **Integrated**—From purpose-built backup appliances to big data platforms, NetBackup integrates at every point in the technology stack to improve reliability and performance. OpenStorage Technology (OST) provides even tighter integration with third-party storage and snapshot solutions.
- **Innovative**—With hundreds of patents awarded in areas including backup, recovery, virtualization, deduplication and snapshot management, NetBackup continues the long Veritas tradition of bringing advanced technologies to market first.
- **Proven**—For more than a decade, NetBackup has led the industry as the most popular enterprise data protection software by market share and is used by many of the largest enterprises on the planet. When you need your data back, you can trust NetBackup.

Key Features

- One platform, one console unifies virtual and physical global data protection
- Unified global management of snapshots, replicated snapshots, backup and recovery
- Scalable, global deduplication across virtual and physical infrastructures
- Single-pass backup, instant image and single file restore for virtual and physical
- Automated virtual data protection and load-balanced backup performance

Business Value

Many Veritas customers are considering AWS as a supplemental data center—a hybrid of both on-prem and cloud—or as a means of eliminating the traditional data center. These changes in the business model require new strategies to migrate and protect data and workloads. The extensive value of Veritas solutions goes beyond seamlessly protecting data regardless of location to orchestrating the movement of workloads to the cloud.

Whether it's a disaster recovery (DR) requirement or the desire to eliminate physical data center management, customers are thinking cloud more often, and Veritas is there to help every step of the way.

Why Do Customers Use the Cloud?

Customers are using the cloud for several reasons. Smaller customers like not having to maintain a data center and an expensive DR site. Midsize customers enjoy having an off-site copy of their data that is built on highly scalable hardware or uses just-in-time cloud recovery. Large customers with data centers are identifying workloads that can take advantage of cloud availability and cost while freeing up expensive data center space for mission-critical workloads. Sometimes a customer will need a temporary space for a workload and instead of ramping up a new rack of disks in a data center will temporarily use space at a cloud provider to avoid the additional cost of purchasing data center hardware. Cloud subscription models work very well for these sorts of projects with highly scalable and simple-to-use models.

The current megatrend of moving data to the cloud revolves around driving costs down for business. The cloud model is also very agile when it comes to requirements. Organizations can add disks to a server quickly and easily versus having to source the hardware and the rack and stack that comes with it.

The cloud also addresses the issues of hardware maintenance and updates. For example, new firmware for arrays is required on a regular basis, causing risk and downtime to install. Similarly replacing or upgrading hardware impacts the environment by requiring a customer to manage these in the data center. Alternatively, in the cloud these requirements are taken care of by the cloud provider and are invisible to the customer.

Customers will have different reasons to move to cloud-based computing, and Veritas solutions allow customers to run their business seamlessly across physical, virtual or cloud infrastructure.

NetBackup and AWS Design Overview

There are many design cases when it comes to NetBackup and AWS. This section will outline them from a high level. Specific use cases are included in the next section.

A NetBackup server with additional Media Servers and clients is collectively referred to as being part of a NetBackup domain. Veritas has Cloud Formation Templates (CFT) that deploy a NetBackup server instance in minutes from the AWS Marketplace. The CFT provides customization options that include a choice of EC2 instance and disk sizes, NetBackup server type, networking and connectivity options. You can add additional data disks (Elastic Block Storage) to the virtual machine (VM) base to store backup data from the local domain or data duplicated from other NetBackup domains.

Backup data can be stored as large contiguous fragments (AdvancedDisk) or written in a storage-optimized manner using the Veritas Deduplication Engine. The Deduplication Engine is the underlying technology that powers NetBackup storage technologies such as Media Server Deduplication Pool (MSDP) and its cloud tier. This deployment approach is similar to VMs provisioned using other hypervisors. VMs are spun up as needed and managed just like NetBackup running on a physical machine.

Versatile Deduplication Engine

NetBackup MSDP is an intelligent deduplication solution because the data stream is generated by a NetBackup client and NetBackup interprets the metadata for the incoming backup stream to understand what kind of data is being protected. Based on the type of data being backed up, either fixed or variable-length block size is assigned to that particular backup—the block size is optimized based on the type of data being backed up. This approach works well in terms of compressing the data because different file types are compressed to different degrees with different block sizes. And because the block size is already optimized to best compress that particular data type, CPU cycles don't need to be wasted to determine the block boundaries. This approach provides a good balance between performance and resource utilization.

The deduplication cloud tier uses MSDP deduplication technology to upload deduplicated data directly to the cloud. This cloud storage server can be either a NetBackup Appliance or a BYO MSDP Linux Media Server that has had one or more cloud tiers added to it.

NetBackup and Cloud Connectivity

NetBackup can use AWS in several ways, depending on the needs of the customer. As outlined in various places in this document, NetBackup can use Simple Storage Service (S3) object storage, Glacier or Glacier Deep Archive to send data to AWS storage similar to a regular disk pool or use NetBackup MSDP Cloud Tiering (MSDP-C) to send deduplicated data to S3 object storage more efficiently.

If a customer has resources in the cloud, NetBackup can also be installed in the cloud and used to protect these resources in a similar manner to protecting physical resources in a data center. This approach avoids the cost and performance impact of traversing data back to the data center for backups.

NetBackup and Cloud Restore Options

Restores of information in the cloud are as simple as in a local data center. The backup admin has full use of the UI and APIs to recover information. Restore performance is relative to the type of storage (S3, EBS or Glacier). A section at the end of this document

covers the restore process of basic S3 storage and includes screenshots. The restore process is generally the same for other tiers, however the archive tiers will have an automated *warming* phase prior to recovery.

AWS Cloud Versus On-Premises Considerations

Running traditional IT workloads in the cloud can have significant benefits if designed and architected correctly. However, if architected improperly, you could end up paying an unexpected price in terms of cost, workload performance and management headaches.

When protecting workloads in the cloud, consider the following:

- Input/output operations per second (IOPS) available
  - On-premises: In an on-premises environment, you can select the appropriate hardware to meet specific IOPS requirements
  - AWS Cloud: In AWS, you can make your selection of EBS volumes based on IOPS requirements There is a cost associated with guaranteed IOPS
- Peer link limits
  - On-Premises: In an on-premises environment, you can basically have as many peer-to-peer links as required
  - AWS Cloud: In AWS, there are fixed limits on the number of Virtual Private Cloud (VPC) to VPC peer links that are allowed
- Storage targets
  - On-premises: In an on-premises environment, you typically write to block storage, deduplication devices or MSDP
  - AWS Cloud: In AWS, your storage targets are typically EBS or S3 object storage. EBS is generally more expensive and S3 is more durable

Use Cases Overview

There are several different use cases with NetBackup and AWS. A few of them are outlined in this section. This list is not comprehensive. The case will vary with customer requirements; however, the use cases presented here will cover the more popular options.

Standard Backup to AWS S3 Object Storage

With NetBackup, the simplest way to move data to object storage is to use the standard cloud connector interface. This interface allows you to configure a cloud storage target or any number of providers, including Amazon, Amazon GovCloud and supported S3-compatible targets (see Figure 1).

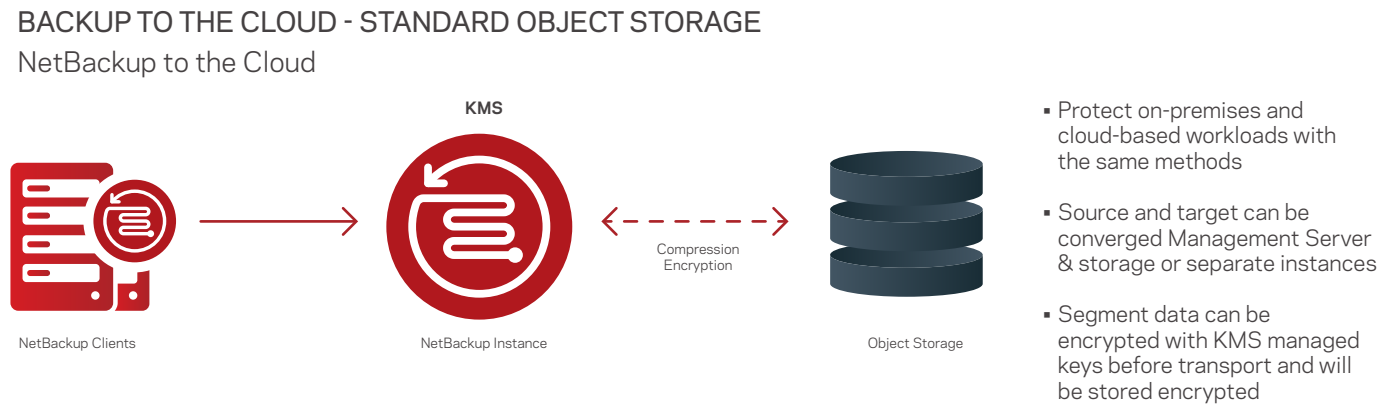


Figure 1. Backing up to the cloud with NetBackup using standard object storage.

This functionality allows you to write to a straightforward and easy-to-implement cloud storage target from any NetBackup server. Standard charges apply based on data ingress and egress charges as documented on the AWS S3 pricing page: https://aws.amazon.com/s3/pricing/

Backup to AWS S3 Object Storage with MSDP-C Deduplication

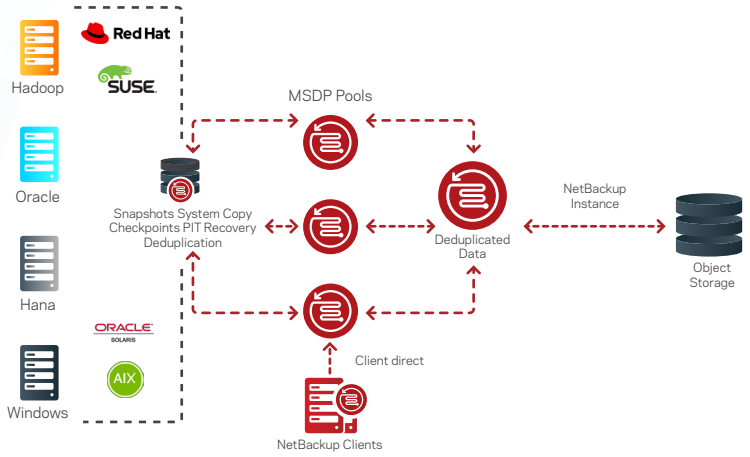
The NetBackup MSDP-C solution combines the performance and flexibility of NetBackup with powerful data deduplication technology to better leverage the cloud for storing backups for DR, cloud data reuse/recovery and long-term data retention. By ensuring backup data remains optimized while in transit to the cloud and while at rest in the cloud, NetBackup MSDP-C greatly reduces cost and increases performance when using cloud storage.

NetBackup MSDP-C can be delivered on a purpose-built appliance, a virtual appliance or as a build-your-own (BYO) software solution. You can add MSDP-C to an existing MSDP pool as of NetBackup 8.3 and higher. In addition, MSDP-C allows customers to send backup data to cloud object storage in deduplicated form. As a storage target, MSDP-C can receive optimized backup images from existing MSDP-compatible sources or directly from a client and transfer the data to an S3 public or private cloud object storage target. A wide variety of S3 object storage has been certified for use with NetBackup MSDP-C.

When using any MSDP-compatible data as the source, the MSDP-C pool does not rehydrate or remove optimization from deduplication. This end-to-end deduplication is a significant difference in how MSDP-C operates compared to other solutions in the market today. The MSDP-C server allows direct recovery of data from the MSDP-C server without first passing through another media server. Using MSDP-C will provide the highest level of functionality and cost savings when using object storage. (See Figure 2.)

BACKUP TO THE CLOUD – DEDUPLICATION TO OBJECT STORAGE

NetBackup Architecture Extended to the Cloud



- Deduplication at the source
- No rehydration of images required between pools
- Automatic protection of all nodes, physical and virtual
- Protect on-premises and cloud-based workloads with the same methods
- Source and target can be converged Management Server & storage or separate instances

Figure 2. Sending data to AWS S3 using NetBackup deduplication.

NetBackup's MSDP-C feature provides a flexible, scalable, high- performing and easy-to-configure solution that lets you use cloud storage more efficiently. Data is stored directly to cloud targets with deduplication. You can configure one MSDP storage server to support multiple storage targets, including one local storage target and zero or more cloud storage targets. You can move data to local and multiple cloud targets simultaneously. The cloud targets can be from the same or different public or private providers and you can add them on demand after the MSDP server is configured and active (see Figure 3).

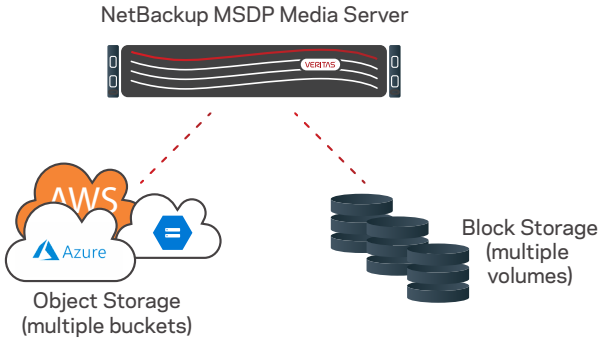


Figure 3. Using MSDP with CloudTier to send data to AWS and other providers.

Multiple cloud targets can coexist in a single cloud bucket or multiple buckets that are distributed in a single cloud provider or different providers. The data and metadata for local storage and multiple cloud targets are isolated to support multi-tenant use. Optimized deduplication is supported within one MSDP server scope so data can be stored to local storage first and then duplicated to cloud targets in the same media server. DR from the cloud targets is enhanced and more straightforward. Finally, the data stored in the target buckets is entirely self-descriptive, allowing one or more recovery servers to be attached to the bucket from cloud or on-premise locations. The Cloud Recovery Server option allows a user to create a NetBackup instance from the AWS Marketplace CFT and attach it to the existing bucket to reuse the backup data in the cloud for testing, recovery or migration. A CRS can also convert VMware images for deployment in EC2 from the NetBackup web UI.

Sending Data into AWS S3 Using AWS Snowball or Snowball Edge

For organizations that must send large amounts of data to Amazon S3 storage classes regularly to initially seed data 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 from NetBackup either via Amazon S3 or the Network File System (NFS) protocols, depending on the Snowball device used (see Figure 4).

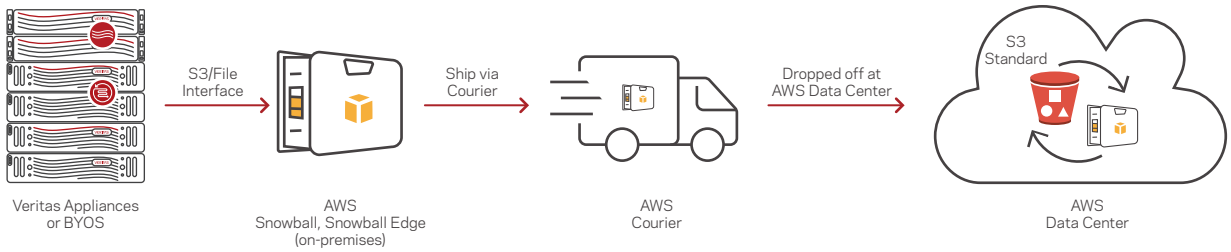


Figure 4. Transferring data to AWS S3 with NetBackup using AWS Snowball or AWS Snowball Edge.

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. 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.

Sending Data to AWS S3 Using a Third-Party Gateway Appliance

This use case is for customers trying to get away from maintaining a tape infrastructure locally in the data center by moving to disk, and would like to take advantage of the conveniences of cloud-based storage. This solution uses a third-party deduplication appliance on-premises that reduces the amount of data sent to the cloud. From a NetBackup standpoint, the dedupe appliance looks like a disk storage unit. Backups are sent to the appliance the same way backups are sent to any disk pool. The appliance will perform the deduplication, so only the changed blocks are forwarded to the cloud via the S3 connector (see Figure 5).

This solution will work with any S3-compatible gateway that presents itself as a disk target to NetBackup, allowing data to be deduplicated for the given environment. Unlike MSDP-C, third-party deduplication appliances are not compatible with MSDP and will require the data to be rehydrated prior to sending it to the device.

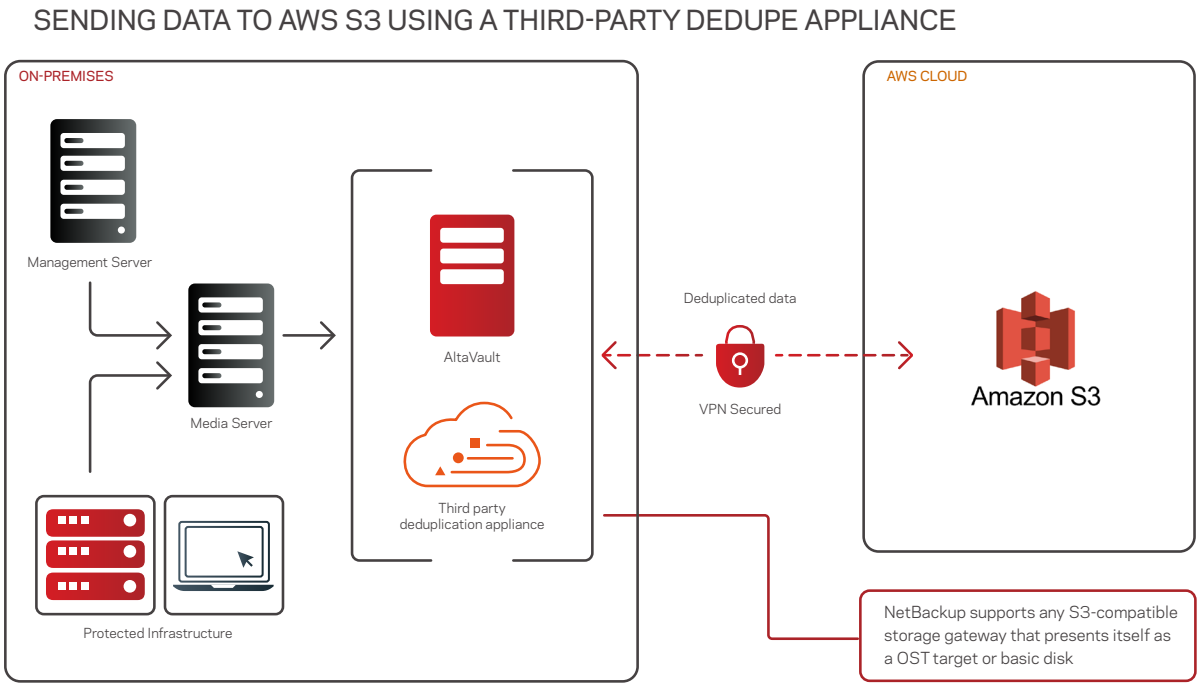


Figure 5. Sending data with NetBackup to AWS S3 using a third-party dedupe appliance.

Amazon Storage Gateway Virtual Tape Library

This use case emphasizes use of the Amazon Storage Gateway. This storage gateway uses iSCSI to provide what looks like a virtual tape library (VTL) to NetBackup. The ability to provide this access requires the downloading and installation of the Tape Gateway Virtual Machine from AWS. This VM will run in ESX or Hyper-V. It will then present an iSCSI connection to NetBackup. Local storage will be required on the VM to allow for data caching prior to the data being sent to AWS.

Once the VM is installed, the NetBackup media server will be able to see the VTL and configure it as it would a physical tape robot, including drives and tape media with associated bar codes. Volumes (tapes) can be configured in various sizes ranging from 100 GB to 2.5 TB. The gateway can have up to 1,500 virtual tapes with a maximum aggregate capacity of 1 PB of storage. Backups are sent to the VTL in the same manner as backups that are sent to a physical tape library. The data is then compressed and sent to AWS across the network connection, then stored in the cloud. NetBackup controls this archive process of moving the tape fromS3 to Glacier (virtual tape shelf) by ejecting a tape. Therefore, restores from Glacier will require additional time because there is a logical performance difference when reading from Glacier (see Figure 6).

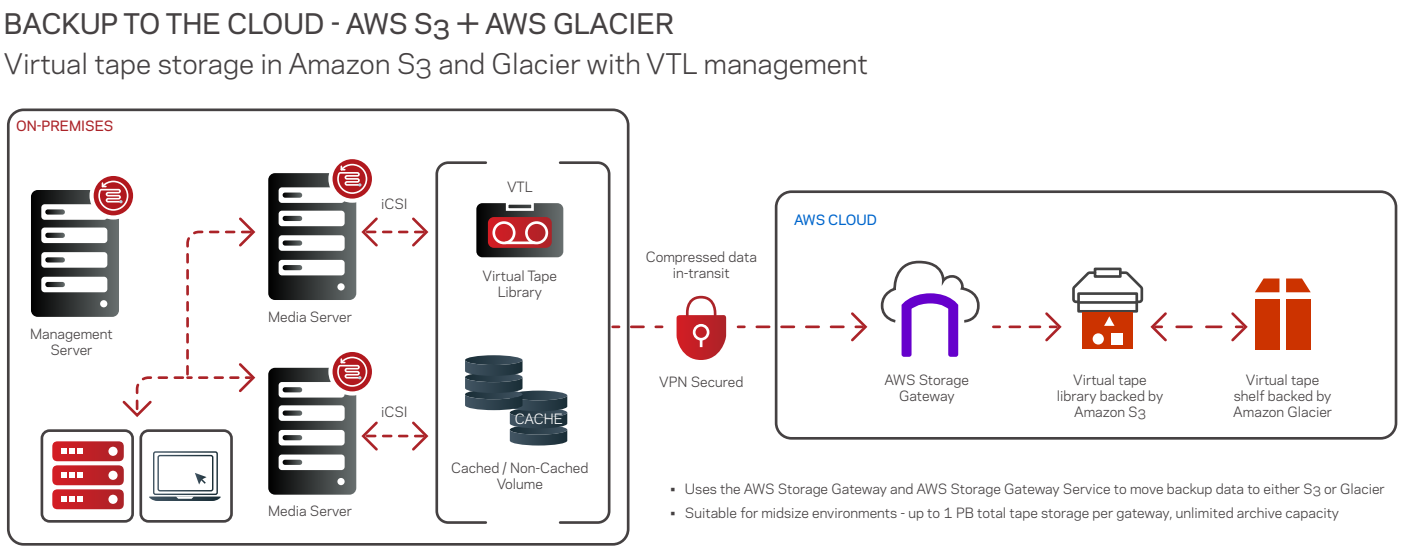


Figure 6. Sending data with NetBackup to an Amazon Storage Gateway VTL

NetBackup will request the data from the VTL, and the VTL will need to pull the data from Glacier back to the VTL and then across the network to NetBackup. This use case is for customers of any size that wish to use what looks like a tape-based tracking solution. Many customers currently use VTL technology and would have to make no changes to their existing infrastructure with this AWS deployment. This approach expands the VTL technology further because the actual storage of the data is in the cloud.

Backup in the Cloud—AWS EC2-Based Architectures

n addition to sending data to the cloud, developing a solution that is completely cloud native is also desirable. This concept is known as infrastructure as a service (or IaaS), and because many customers find that maintaining data centers is not cost-effective, a cloud solution works well for them. This use case involves customer workloads in AWS EC2, which offers the ability to provision VMs similar to any other virtualized environment, with the VM and all storage being in the AWS Cloud environment.

Backups of these workloads are typically still required. EC2 functions similar to a data center that uses a hypervisor environment for VMs. There are built-in safeguards to protect data; however, failures can still occur. NetBackup in EC2 works exactly like NetBackup in a data center. You can provision a NetBackup Management and Media Server from the AWS Marketplace using CFT or manually deploy them in a BYO fashion.

For cloud-native workloads protection, you can launch NetBackup SnapShot Manager from the AWS Marketplace and add it to the NetBackup configuration. NetBackup SnapShot Manager brings cross-cloud functionality and management from the NetBackup UI. NetBackup SnapShot Manager allows automated protection for cloud-native virtual instances, volumes and platform as a service (PaaS) applications from an easy-to-use, central location. (See Figure 7.)

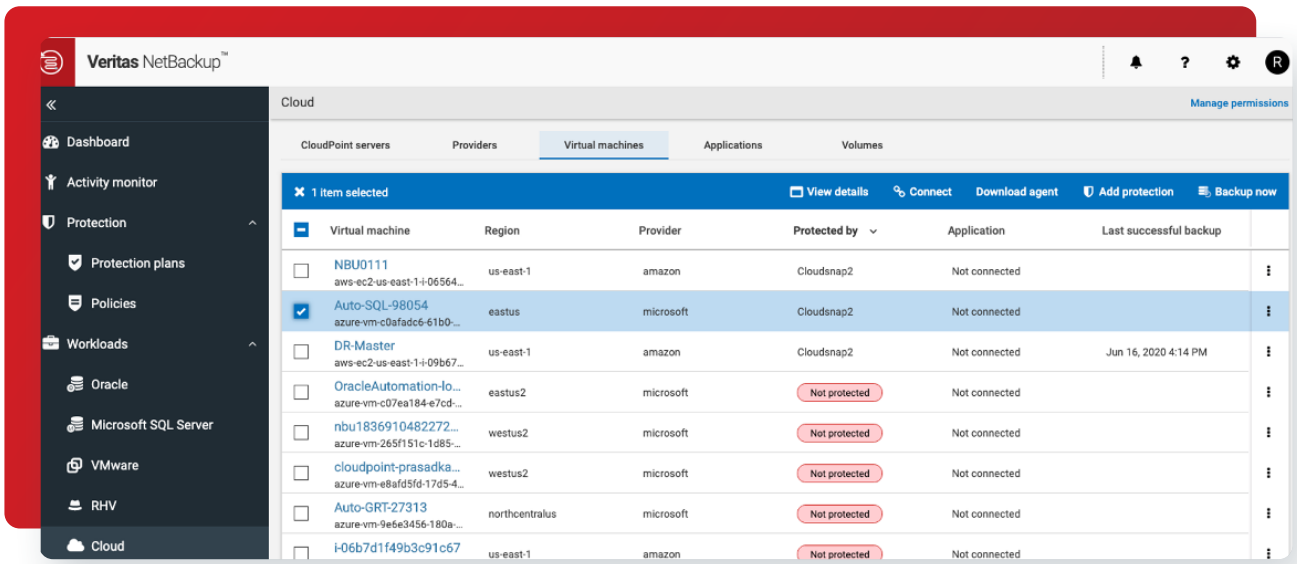


Figure 7. The NetBackup dashboard showing VMs protected by CloudPoint.

Figure 8 shows a Media Server running an MSDP, which means that full or incremental backups will be deduped at the storage (EBS) layer. Alternatively, you can use S3 object storage such as Standard, Standard-IA, Glacier, or Glacier Deep Archive as a storage target for NetBackup running in the cloud. For optimal storage cost savings, use MSDP-C to store duplicated data in S3 object storage. Each option has benefits depending on your specific needs. EC2 instance types and sizing recommendations are covered toward the end of this document.

NetBackup offers an Amazon Machine Image (AMI) and a CFT that is a template for the machine, and makes provisioning of the Management and/or Media Server easy (see Figure 8).

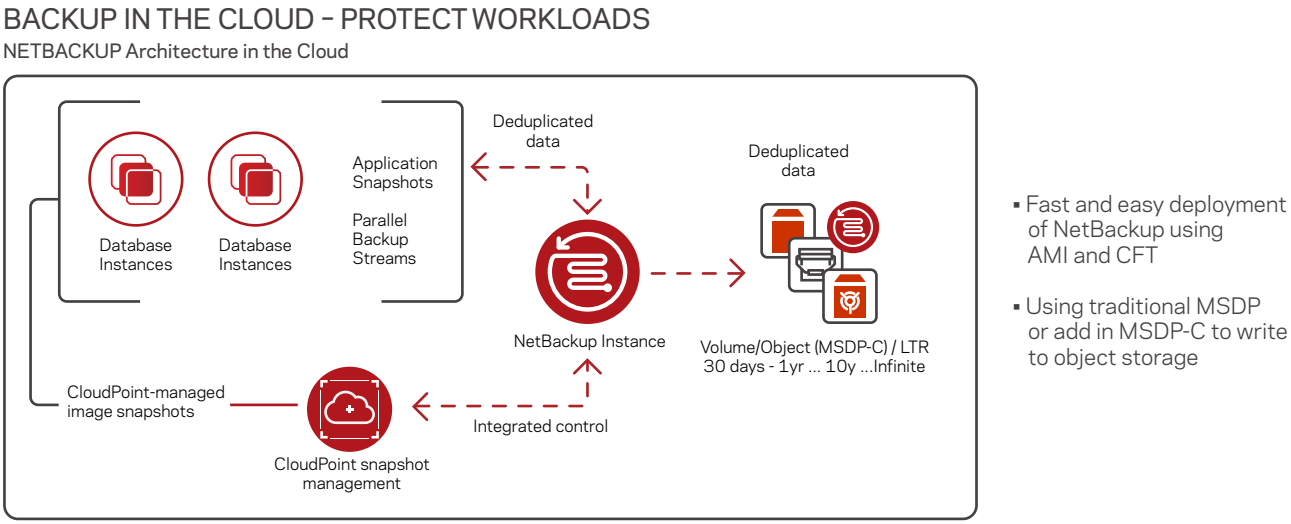


Figure 8. Sending workloads with NetBackup to EC2 storage.

Both NetBackup and NetBackup Snapshot Manager are available in the AWS Marketplace.

NetBackup can work in EC2 with dedupe appliances that also work in the cloud. NetBackup treats such appliances like *basic disks*, the same way it treats them in a data center. The appliance will dedupe the data before sending it on to S3 or to Glacier storage as outlined below.

There are many other configuration options using NetBackup with AWS you can tailor to the customer's needs. These use cases outline a handful of them.

Disaster Recovery Using AWS

Auto Image Replication (AIR) to the Cloud—Hybrid Configuration

Another option to get data into the cloud would be to use a hybrid model where part of the environment is in the data center and a secondary part, for use with DR options, uses the functionality of AIR to get the data to the cloud (see Figure 9).

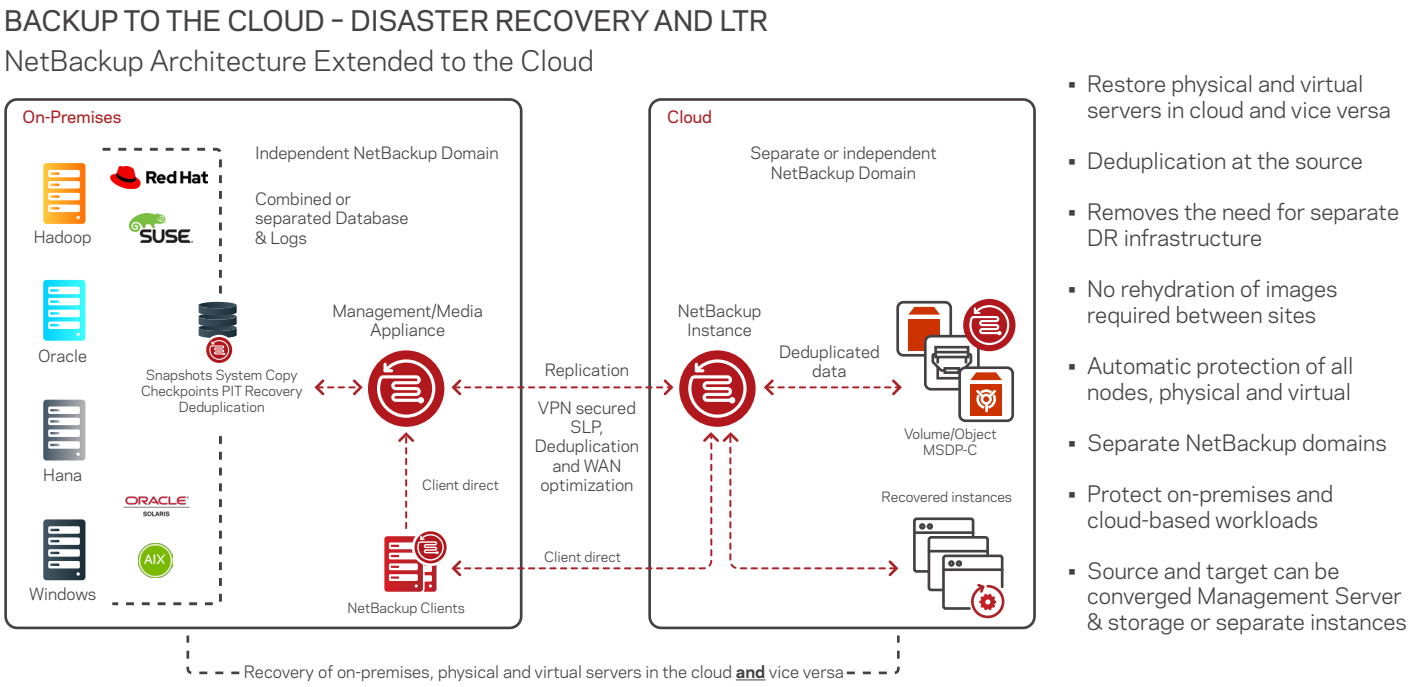


Figure 9. An overview of NetBackup Auto Image Replication (AIR) for cloud DR.

This concept is simple and ties into a number of these use cases. A NetBackup Management and Media Server with MSDP is configured in the data center, and a Management and Media Server with MSDP is configured in AWS. From there, you can use an AIR process to automatically send data from MSDP in data center A to MSDP in AWS. The data can then be imported into the AWS Management Server for use in a DR scenario. This process is the same as using AIR to move data from a data center in San Francisco to a data center in London: The fact that it is in the cloud isn't noted by NetBackup. It's just an AIR target.

This option is ideal for a customer that wants an off-site DR copy of the data at a data center. It's also a good way to migrate to the cloud from a NetBackup perspective.

Veritas offers additional products, such as Veritas Resiliency Platform, that can migrate workloads into the cloud with the help of NetBackup. This method is a perfect blend of creating dual instances of a workload for test/dev/QA while maintaining the original data in the data center. In an existing NetBackup environment, Resiliency Platform can offer orchestration workload recovery. Figure 10 shows how you can use NetBackup and Resiliency Platform to recover workloads into Amazon.

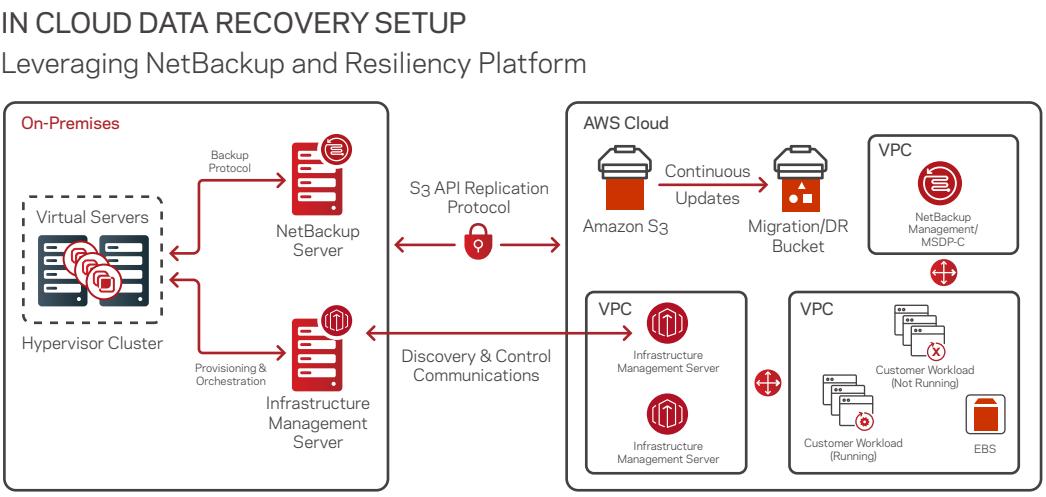


Figure 10. An example of cloud DR using NetBackup and Resiliency Platform.

.Leveraging NetBackup Image Sharing for Migration and DR

Starting with version 8.2, NetBackup has added a new image sharing capability when using MSDP-C to write to AWS S3 object storage. In NetBackup 8.3, the new MSDP cloud tier also includes the same support. This new functionality essentially makes the storage bucket self-descriptive for reuse by an instance other than the one that initially used it. The only data needed to access the data is the bucket name and the necessary authorization.

Using image sharing, you can launch an on-demand NetBackup instance from the AWS Marketplace CFT and attach it to an existing MSDP-C bucket. The new instance will be able to read the bucket data from within the cloud infrastructure and leverage image data to restore workloads in the cloud, or even convert VMware images into an AMI format for host recovery and migration (see Figure 11).

For more advanced migrations of complex environments and their infrastructures, Resiliency Platform integrates with NetBackup to orchestrate recovery and migration operations with push-button simplicity. This capability includes automatic deployment of NetBackup in AWS on-demand instances to leverage MSDP-C data stores in object storage.

.

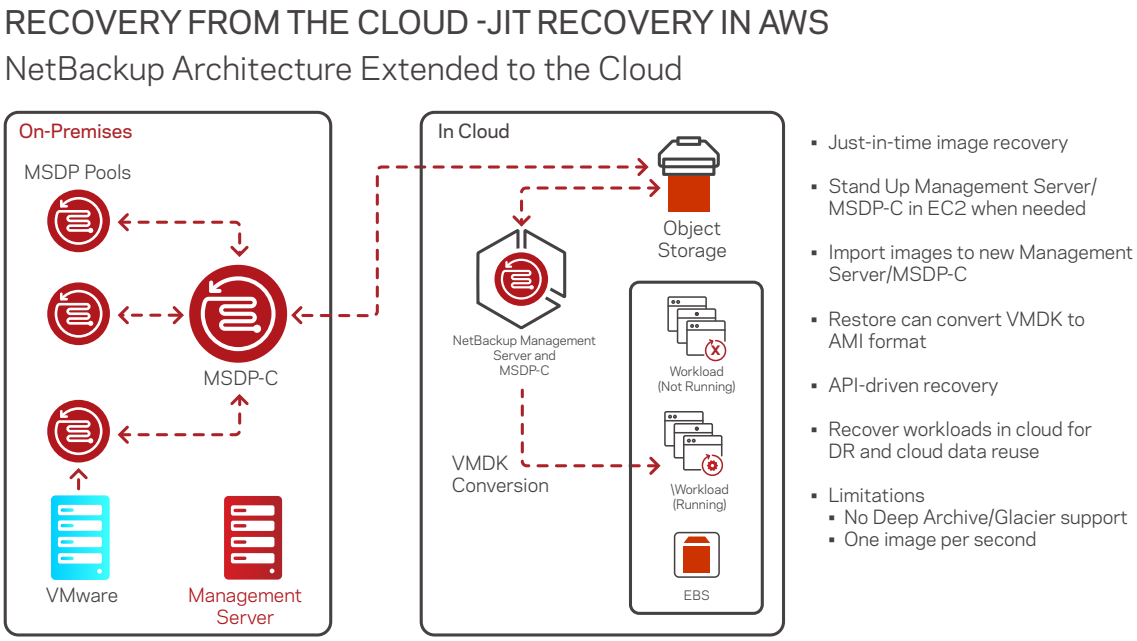


Figure 11. An overview of just-in-time (JIT) recovery in AWS using NetBackup and MSDP-C.

Recovering Virtual Machines from VMware to AWS Using NetBackup Automated DR

The Infrastructure Management Server (IMS) in the on-premises data center discovers the vCenter server and the backup configuration from the NetBackup Management Server. NetBackup MSDP-C with the automated DR feature (available from NetBackup 8.2 onwards) backs up the VM images along with the image metadata from the on-premises data center into the designated S3 bucket.

The recovery using these backup images is achieved using a NetBackup Cloud Recovery Server (CRS) that is deployed in the cloud data center. The image metadata stored in the S3 bucket allows the NetBackup CRS to read the image information and create an Amazon Machine Image (AMI). This AMI is then used to provision cloud instances in the cloud data center during the restore operation.

NetBackup Intelligent Cloud Policies

Most cloud resources utilize one or more Tags in the form of key:value pairs that describe their function, data classification, business owner or purpose. Due to the elastic nature of cloud infrastructure, it is completely reasonable to expect that an array of cloud resources brought online one week, are now no longer available a week later. Therefore, enterprise backup platforms for the cloud must be able to keep up with the dynamic provisioning and deprovisioning of cloud resources.

NetBackup Cloud Intelligent Policies makes life easier on cloud administrators with features such as autodiscovery and SQL-query-based definitions to logically define asset groupings. Creating a Cloud Intelligent Policy (CIP) is easy; use our visual query builder to create a group of cloud assets for protection. No prior SQL expertise is necessary. You can also use NetBackup API to automate the creation of cloud intelligent groups. Schedule protection by associating the Intelligent groups to a Protection plan, or use the Backup Now option to generate an immediate recovery point. As new cloud assets come online with similar tag criteria, they stay protected, and you remain compliant.

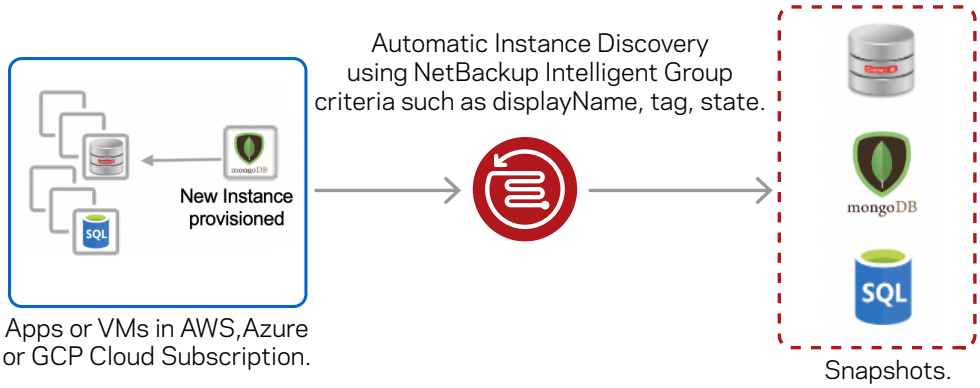


Figure 12. Automated Cloud Asset Discovery.

The automatic discovery of new assets and intelligent filtering drastically reduces the effort required to automate data protection for cloud assets and stay compliant, as enterprises continuously adopt multi-cloud applications and infrastructure. NetBackup protects all your resources across multiple clouds.

Immutable Storage in AWS

NetBackup customers who are looking to take an additional step for security reasons (ransomware, rogue actor) or compliance requirements are leveraging immutable storage. Immutable storage is data storage that will remain completely static or unchangeable for its entire life cycle. Immutable storage allows NetBackup users to designate specific data stored in a form that can never be altered. NetBackup has always supported many storage platforms and continues the tradition of heterogenous support for immutable locks across many different vendor storage platforms providing flexibility, cost savings, and coverage from the edge to the core to the cloud.

NetBackup now supports for AWS S3 Object Lock's using AWS S3 immutable object storage. Amazon S3 Object Lock is an Amazon S3 feature that allows you to store objects using a write once, read many (WORM) model in the Amazon cloud. You can use WORM protection for scenarios where it is imperative that data is not changed or deleted after it has been written, whether your business has a requirement to satisfy compliance regulations, or you want to protect your data from ransomware.

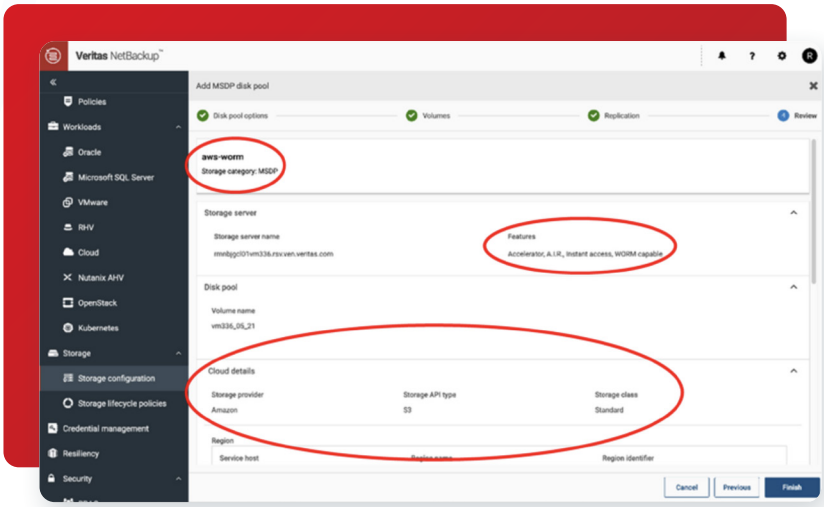


Figure 13. NetBackup media server storage configuration.

The NetBackup administrator can define the retention period for a WORM NetBackup backup policy using a bucket that has been configured to use S3 Object Lock. Individual objects may have multiple references from NetBackup, so an object will remain locked until the longest retention period has passed for that bucket. A NetBackup administrator can configure an object to be copied to another bucket and stored in a locked state on a secondary target. The NetBackup object lock can be applied to any workload data, and a single NBU Media Server can write data in either a locked or non-locked state, offering flexibility and greater use of the Media Server to address locked and un-locked buckets.

Backup from Snapshot

Cloud snapshots provide an excellent option for fast recovery, with rollback restore and the ability to restore to alternate virtual machines. However, it can be cost-prohibitive to use snapshots as long- term retention.

By deduplicating and compressing your data, then storing it on a cheaper storage tier, Veritas can provide storage savings up to 98% compared to snapshots. In addition, as compared to the cost of snapshot storage to AWS Deep Archive tier,

NetBackup can reduce the data stored by up to 50%.

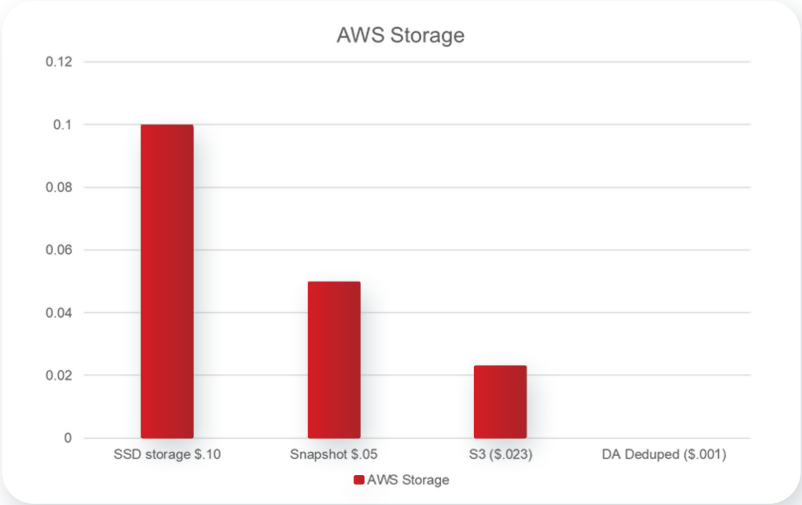


Figure 14. Cost of AWS Storage (SSD/Snapshot/S3/Long Term De-Duped). Source: Veritas

Use Case: Long-Term Retention Point in Time for data copies (compliance, discovery, litigation support, etc.)

As shown below, using Snapshot Storage for the long-term retention of 1TB of data stored is cost prohibitive.

AWS Snapshot Storage

$614/yr

Based on EBS Snaps priced at $0.05/GB/Month

Backup Stored on Archive Tier (with Deduplication) with NetBackup

$12/yr

Based on Deep Archive, priced at $0.001/GB/Month

Figure 15. Cost comparison between AWS Snapshot Storage and NetBackup Backup from Snapshots for long-term retention. Assumptions are based on 1TB (or 1024GB) of annual storage cost. Actual cost savings may vary.

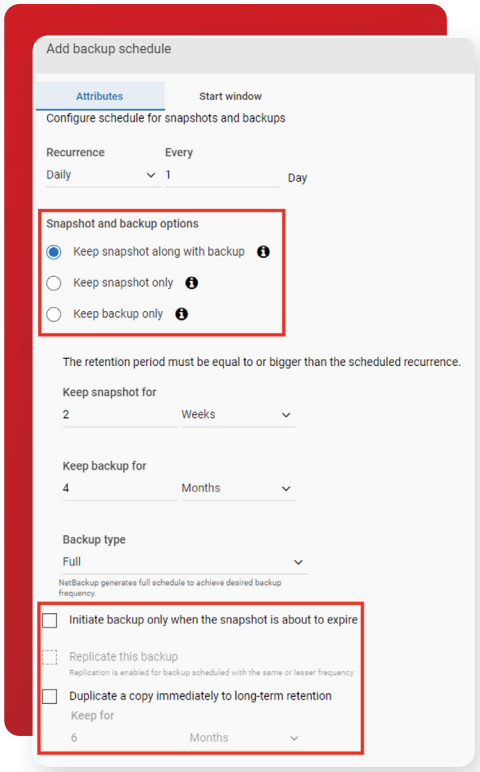


Figure 16. Configuring NetBackup snapshot and backup options.

Storage Lifecycle

NetBackup has long supported the concept of Storage Lifecycle by placing less accessed data on longer-term storage to assist in long-term retention costs. Adopting a similar approach to the cloud, NetBackup supports multiple storage tiers on the cloud and supports storage lifecycle policy (SLP) management of backup copies going from one tier ("hot," for example) to another tier ("Glacier," for example).

Creating backups from snapshot is simple to setup and ready for use when creating a protection plan from within the NetBackup WebUI.

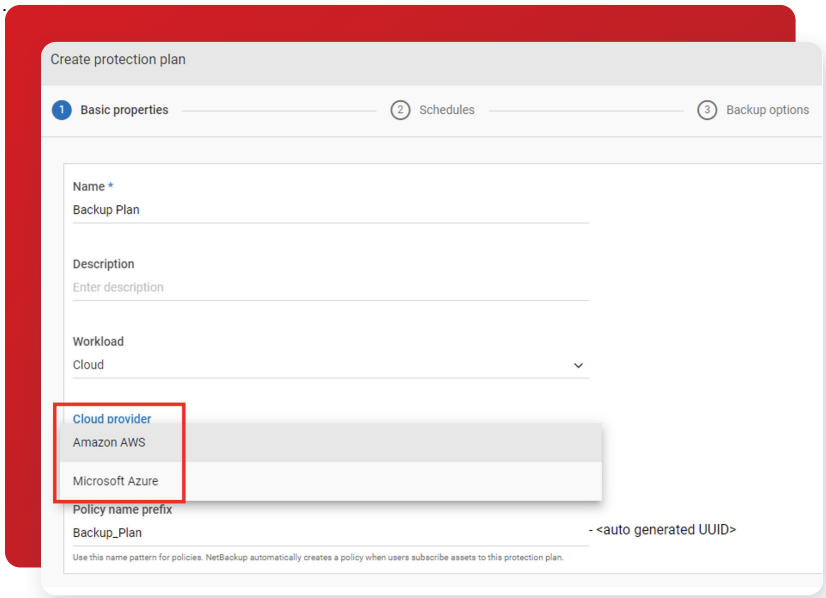


Figure 17. Configuring a NetBackup cloud protection plan.

Available for Amazon AWS, Microsoft Azure and Google GCP, Backup from Snapshot allows backup administrators the ability to retain all the flexibility of NetBackup in the cloud while saving money.

Cost Considerations

The cost of the cloud will vary depending on what a customer needs. Simple backups to AWS can be a very cost-effective solution for a customer that wants to send important data to an off-site location. However, this solution is probably not ideal for a large customer with a large amount of data to send to AWS due to the bandwidth constraints of the network. In addition, AWS Storage is limited in options based on object vs. block-based storage. There is a cost to send the data, store the data and retrieve the data

NetBackup Cloud Autoscaling

Introduced in the latest release of NetBackup, NetBackup Cloud Autoscaling provides a predictable cost envelope when it comes to cloud data protection. For storage cost optimization, thin NetBackup clients read in the snapshot information and move the data to a supported NetBackup storage unit. This includes optimization savings from our deduplication engine and the ability to tier that data to any cloud storage. NetBackup provides this functionally for Azure Stack and Azure resources protection using the NetBackup Snapshot Manager extension.

Figure 16 depicts automatically discovered cloud assets added to an Intelligent Group based on resource attributes or tag criteria, snapshot indexing using the NetBackup Autoscaling client, followed by storing the data onto any configured NetBackup storage.

.

Cloud Sizing and Performance

Sizing and performance of data in the cloud is based on customer need and will vary from customer to customer, which makes providing exact sizing and performance information difficult.

To get data to the cloud, customers can use a simple Internet connection if the data to be transmitted is less than the amount of available bandwidth. (See Figure 18.)

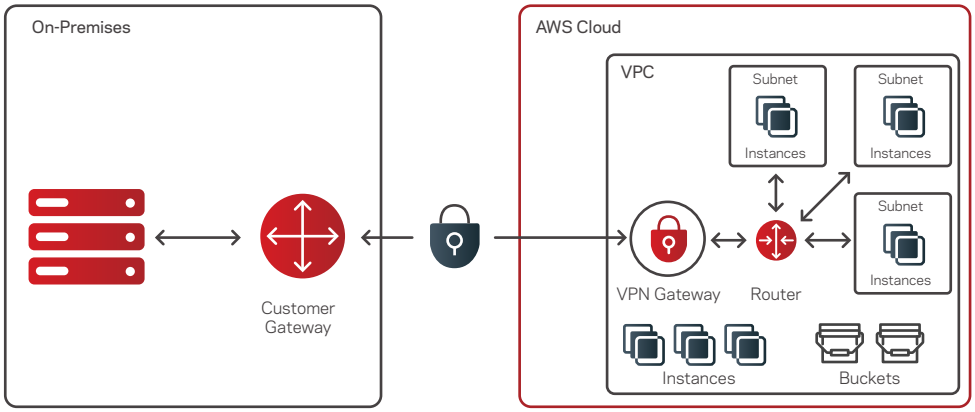


Figure 18. A general view of cloud infrastructure connected with an on-premises data center.

With AWS Direct Connect, customers can get a dedicated link to AWS with performance similar to LAN bandwidth inside a data center. They can compress data at the data center prior to being sent across the network to AWS or use MSDP-C to optimize the data being sent before it is stored in the cloud. They can also *throttle* bandwidth, if desired, to prevent over-saturation of the network pipe.

AWS Instance Model

AWS uses a regional model in which it has configured various Regions across the globe. Each Region has Availability Zones that are similar to data centers within the Region that communicate with each other over high-bandwidth connections. This setup is similar to a customer having multiple physical data centers in a geographical region that are close enough for low-latency connectivity, yet far enough apart to not be impacted by the same natural or artificial disaster.

Data within the Region will typically stay within the Region, but customers have the option to select a geographically dispersed Region that is available for regional DR. Data can be replicated between Availability Zones to provide high availability within the cloud for the customer's data. The loss of a single Availability Zone does not impact the operations of the others. Customers typically choose to operate within the Region that is closest to provide optimized bandwidth for moving data in and out of the AWS Cloud and have the option to select a geographically dispersed Region to provide regional DR.

AWS Storage Options

One of the many benefits of the AWS storage model is the ability to quickly add storage to environments. Customers don't pay for the storage until it is provisioned. This model is much different than a traditional data center where racks of disks may sit idle until needed, thus increasing total cost of ownership (TCO). If the disk is spinning and generating heat, additional cooling and power could be needed to keep the disk spinning even if it is not currently in use. Although next-gen SSD arrays require less cooling and power, idle disks still increase TCO.

Once data is in the cloud, AWS uses various types of storage including object (S3), object with infrequent access (S3-IA) and block (EBS), depending on the type of use case. Other options include EFS for scale-out storage targeted at big data solutions and Glacier for long-term storage of archive data that is rarely accessed. Sizing of the environment is based on the needs of the customer and the workloads placed in the cloud. Pricing is based on the type of storage chosen and is typically priced per GB. For example, standard S3 storage typically runs approximately $0.023/GB per month, whereas Glacier storage currently runs about $0.004/GB per month and even less for Glacier Deep Archive storage. Cost also depends on the Region where the data is stored (see the AWS pricelist for

current costs). Glacier is typically used as long-term archive storage target whereby data is moved there automatically using a variety of methods, and a restore from storage could take hours to access versus seconds for an S3 restore.

Environment Description and Assumptions for Sizing

The following sizing guidelines are based on the assumptions listed and were created using the standard NetBackup Appliance Calculator to determine the storage required for each type of workload. This guideline applies purely to AWS and back up in the cloud workloads only.

The following assumptions were used to size this environment:

- Data assumptions:
  - Data split – 80% FS / 20% DB [no hypervisor level in the cloud]
  - Daily retention 2 weeks / weekly – 4 weeks / monthly 3 months
  - Daily change rate 2%, and YoY growth 10% [sizing for 1 year only]
- Instance Type workload descriptions:
  - Small – FETB <=100 TB <= 100 concurrent jobs
  - Medium – FETB <=500 TB <= 500 concurrent jobs
  - Large – FETB <=1,000 TB <= 1,000 concurrent jobs
  - Extra-Large – FETB > 1 PB >1,000 concurrent jobs

NetBackup AWS Instance Sizing

The architecture is based on a single NetBackup domain consisting of a NetBackup Management Server and multiple MSDP Media Servers in the AWS EC2 cloud.

Typically, backups are written directly to MSDP block stroage for an immediate copy, then opt-duped to a cloud tier to send dedupe data to S3 object storage. However, there is no requirement that backups must go to standard MSDP before MSDP-C. If the solution doesn't require MSDP data to be "local" on block storage, backup data can be sent directly to a cloud tier.

Requirements consist of the following:

- NetBackup Management Server
  - A single NetBackup Management Server can be on any supported operating system.
- NetBackup MSDP Media Server's block storage
  - MSDP Media Servers receive the initial backups from clients and performs deduplication.
- NetBackup MSDP Media Server's Cloud Tier
  - MSDP can have one or more targets on the same storage server that takes the deduplicated backup images from the MSDP Media Server's block storage and stores them in S3.
  - The CloudCatalyst Server is dedicated to performing NetBackup dedupe writes to S3 object storage. It is a dedicated high-end Red Hat server that meets the minimum requirements for CloudCatalyst. It takes the deduplicated backup images from the MSDP Media Servers and stores them in S3.
- Backup Workloads (Clients/Agents)
  - These are the systems or applications that are being protected.

NetBackup Management Server

The NetBackup Management Server should be sized according to the standard Veritas guidelines depending on the load placed on the complete NetBackup domain. Plan accordingly for the initial needs of the environment. AWS does offer the added benefit of being able to scale up the systems as workloads grow. The solution can scale out by adding additional Media Server nodes.

Management Server Memory and CPU Requirement

Table 1. Management Server Memory and CPU Minimum Requirements

Number of Processors	Minimum RAM	Maximum Jobs per Day	Maximum Media Servers per Management Server
4	16 GB	10,000	20
8	32 GB	20,000	50
16	64 GB	30,000	100

The minimum processor (CPU) and server memory requirements referenced in Table 1 are estimates based on the number of Media Servers and the number of jobs the Management Server must support. You may need to increase the amount of RAM and number of processors based on other site-specific factors.

Management Server Recommendations – AWS EC2 Sizes

Small	Medium	Large	Extra Large
32 GiB / 8 vCPU	64 GiB / 8 vCPU	64 GiB / 16 vCPU	122 GiB / 16 vCPU
Install 500 GB EBS Catalog 5 GB EBS	Install 500 GB EBS Catalog 5 GB EBS	Install 500 GB EBS Catalog 10 GB EBS	Install 500 GB EBS Catalog 10 GB EBS
m4.2xlarge	r5.2xlarge	m4.4xlarge	r5.4xlarge

NetBackup MSDP Storage

NetBackup MSDP storage can reside on either a NetBackup Appliance, a Virtual Appliance or a BYO virtual or physical host, including a cloud-based virtual instance. This section will outline MSDP in AWS built on an EC2 instance with EBS storage.

Specifications for MSDP Media Server in EC2

The host computer's CPU and memory constrain how many jobs can run concurrently. The storage server requires sufficient capability for deduplication and for storage management. Processors for deduplication should have a high clock rate and high floating-point performance. Furthermore, high throughput per core is desirable. Each backup stream uses a separate core. (See Table 2.)

Table 2. Recommended Specifications for MSDP Media Servers in AWS

Hardware Component	MSDP Media Server
CPU	▪ Veritas recommends at least a 2.2-GHz clock rate. A 64-bit processor is required ▪ At least 4 cores are required. Veritas recommends 8 cores ▪ For 64 TBs of storage, Intel x86-64 architecture requires 8 cores

Hardware Component	MSDP Media Server
RAM	▪ From 8 TB to 32 TB of storage, Veritas recommends 1 GB of dedicated RAM for 1 TB of block storage consume ▪ However, beyond 32 TB storage, Veritas recommends more than 32 GB of RAM for better and enhanced performance ▪ MSDP-C uses a dynamic spooler cache based on previous and currently running backups and does not leverage the traditional persistent fingerprint pool ▪ MSDP-C also will try and leverage memory as an upload/download cache before falling back on disk. This will be relative to the number of concurrent jobs and each job will use 128 MB of upload cache data. The default max for "CloudUploadCacheSize" is 12 GB, which would allow for roughly 90 concurrent jobs
Storage	▪ MSDP block storage will perform best with storage that is 250 MB/s or faster. Because many volumes/VMs have a 250 MB/s max, it's recommended to use a RAID0/1 stripe ▪ Start out with the expected needed storage based on deduplication rates. Storage can easily be expanded by adding additional volumes to MSDP ▪ MSDP-C does not use a dedicated cache volume. Rather, it will make non-persistent use of free storage on the MSDP server when needed ▪ By default, MSDP-C does require at least 1 TB free space on the MSDP server per cloud tier (configurable in contentrouter.cfg)
Operating System	▪ The operating system must be a supported 64-bit operating system. MSDP-C requires a RHEL/ Centos 7.3 or later server ▪ See the operating system compatibility list at http://www.netbackup.com/compatibility

Growing the Media Server

As the amount of data protected by a server increases, the load requirements on that host will increase. In that scenario, there is a simple solution. You can easily expand any EC2 instance or add EBS volumes to meet higher requirements that may happen over time.

Refer to the following for more information:

- Changing the instance type
- Attaching a volume to an instance

Media Server Deduplication Pool Recommendations

Running traditional MSDP in a cloud environment requires specific resources to be available such as a 10G network and EBS volumes with provisioned IOPS. The recommendations below have been formulated using AWS kits that addresses MSDP pools of different sizes. These are just recommendations and specific customer environments may have different needs. Depending on the AWS footprint, any of the environments below would work based on the sizes.

MSDP Considerations

- Example MSDP storage pool size is up to 96 TB on Linux:
    - Can be a direct backup target, use Fingerprinting Media Servers or a client-side dedupe target
    - MSDP will be storing all data on managed disks
    - The pool will be able to replicate to any Veritas deduplication-compatible target, including MSDP-C

Storage Considerations

Although multiple Media Server deduplication nodes can exist in a NetBackup domain, nodes cannot share servers or storage. Each node manages its own storage. Deduplication within each node is supported; deduplication between nodes is not supported. For a small 32 TB MSDP storage pool performing a single stream read or write operation, storage media 200 MB/sec is recommended for enterprise-level performance. Scaling the disk capacity to 250 TB recommends a performant 500 MB/sec transfer rate. Multiple volumes may be used to provision storage; however, each volume should be able to sustain 250 MB/sec of IO. Greater individual data stream capability or aggregate capability may be required to satisfy your objectives for simultaneous writing to and reading from disk. The suggested layout to use in a cloud infrastructure is a striped RAID0 or RAID1 configuration. More complex RAID configurations are not cost-efficient. For more information, see RAID Configuration on Linux.

Table 3 shows recommended NetBackup Media Server sizing guidelines based on the size of the intended deduplication pool.

Table 3. Recommended Media Server Sizing Guidelines

Dedupe Pool	Instance Type	Storage	Cores	RAM	Networking	IOPS
10TB (Small)	r3.2xlarge	1x160 SSD 1x16 TB EBS-SSD	8	61		
1-20 TB (Small)	c4.8xlarge	1x80 EBS-SSD 1x16 TB EBS SS	36	60	10 GB	EBS Provisioned IOPs (SSD)
32 TB (Medium)	c4.4xlarge	1x80 SSD 2x16 TB EBS-SSD	16	30	10 GB	
	r3.2xlarge	1x160 SSD 2x16 TB EBS-SSD IOPs – 12,000	8	61		12,000
32-64 TB (Large)	m4.10xlarge	1x80 EBS-SSD 2-4x16 TB EBS-SSD	40	160	10 GB	
	c4.8xlarge	1x80 EBS-SSD 2-4x16 TB EBS-SSD	36	60	10 GB	
	r3.4xlarge	1x160 SSD 2x16 TB EBS-SSD IOPs – 12,000	8	61	10 GB	12,000
32-96 TB (xLarge)	m4.10xlarge	1x80 EBS-SSD 2-4x16 TB EBS-SSD	40	160	10 GB	
	r3.4xlarge	2x320 EBS-SSD 2-6x16 TB EBS-SSD IOPs -12,000	32	144	10 GB	12,000

* For instance info at a glance, see the chart "EC2 instances information."

The NetBackup Media Server sizing guidelines referenced in Table 3 are based on the size of the intended deduplication pool.

Table 4. Recommended NetBackup MSDP-C Sizing Guidelines

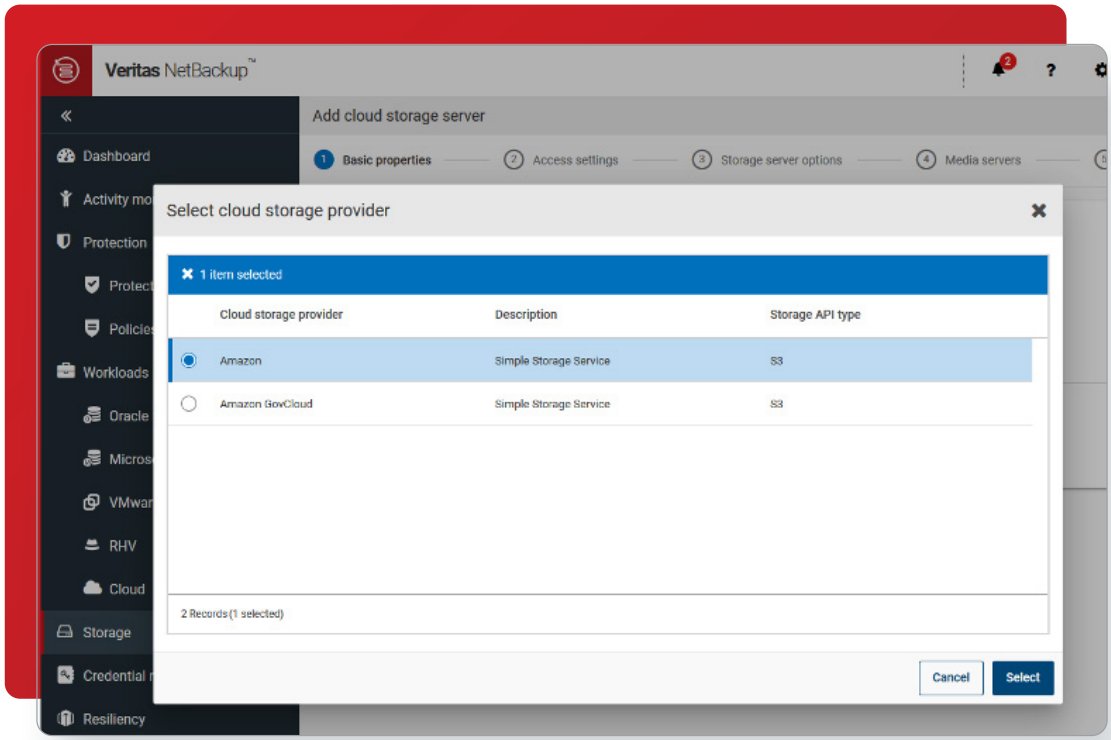
Product	Role	Instance Type	EBS/EFS Storage	CPUs	RAM(GB)
NetBackup	MSDP-C Min	RHEL M5.xlarge	250 GB SSD (gp2) 1+TB	4	16 GB
	MSDP-C Large	RHEL M5.2xlarge	500 GB SSD (gp2) 1+TB	8	32 GB

MSDP-C Server—Configuration Walk-Through

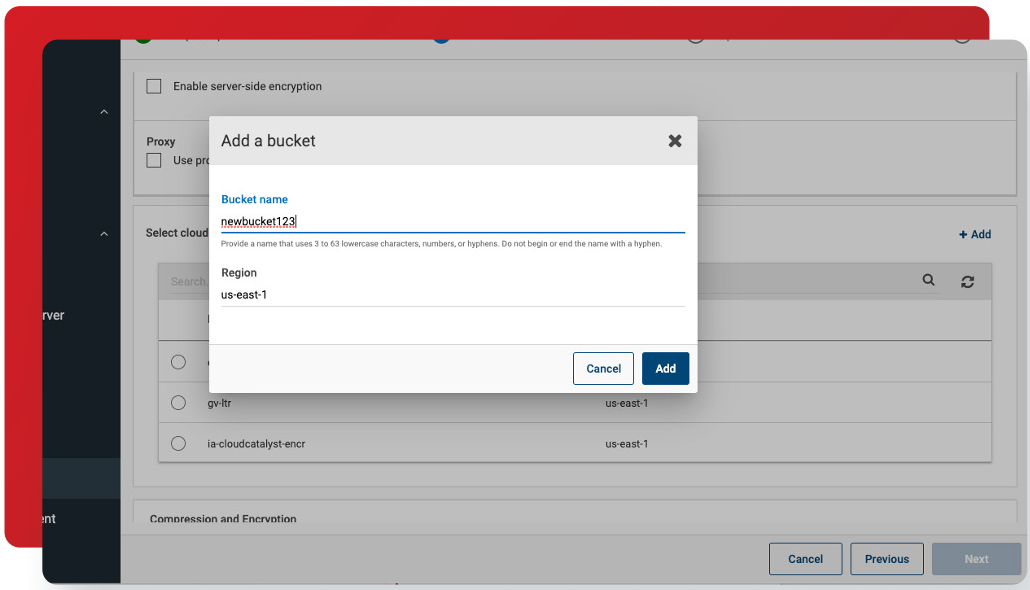
For NetBackup MSDP-C servers, Table 4 lists a minimum configuration and a large configuration along with the AWS VM instance type and the storage configuration. Customers should start with the larger instance recommendation unless they are using MSDP-C for basic functionality testing. The SSD disk listed is for the operating system and NetBackup installation files. The 1 TB volume represents the local cache volume and mount location required for MSDP-C deployments.

Adding a NetBackup MSDP Cloud Tier is easy.

Once an MSDP server is configured to add a cloud tier, select Disk Pools -> Add. After picking the MSDP storage server, select *Add Volume* and select Amazon as the cloud storage provider from the list. If the cloud service end point for your vendor doesn't appear, you must add the cloud storage instance first.



Next, select the Storage Class (such as S3 or Glacier) and the Region and enter the credentials for the AWS account you want to use. The final step is to select or create a new bucket to use.



After reviewing the inputs, click Next and Finish to complete the setup. Follow the remaining wizards to compete the storage unit configuration. For a more detailed walk-through, check out this video: Adding a cloud tier to an existing MSDP storage server

Additional Architecture Requirements

In addition to the use case architectures noted in this document, there are several other topics customers will need to consider when looking at a move or partial move to the cloud.

Security of the Information

In-Flight

Starting with NetBackup 8.1, data security has been heightened because more data is now going to the cloud and out of the ownership of the data center. With NetBackup, the use of SSL and certificates guarantees that the servers and clients being protected and the data being received are from authenticated endpoints.

At-Rest

NetBackup MSDP can deliver source-side encryption from the client end that encrypts data in transit and at rest. In addition, any data from that client that is sent to another pool will maintain that encryption, even when going to the cloud via MSDP-C.

Data coming from NetBackup moving into the cloud can use encryption before the data is sent to the AWS environment from the Media Server. This encryption can use the Key Management Service (KMS) from the NetBackup software or an external KMS to handle the keys. The data in the cloud at rest will be encrypted. The only drawback to this option is that during a restore, the KMS server must be available to have the keys available to decrypt the data. In most cases, this would not be an issue unless the original Management is not available.

Least Privileged Access

For security concerns, it's always important to practice a "least privileged" approach. This approach means that users are limited only to access that they absolutely need. When running NetBackup on AWS or in the cloud, the minimum permissions needed for NetBackup to write and retrieve data from S3 are:

- S3:CreateBucket
- S3:ListAllMyBuckets
- S3:ListBucket
- S3:GetBucketLocation
- S3:GetObject
- S3:PutObject
- S3:DeleteObject

For Amazon Glacier, the following additional permissions are necessary:

- S3:PutLifecycleConfiguration
- S3:GetLifecycleConfiguration
- S3:PutObjectTagging

In order to use the NetBackup Web UI to covert VMware images to Amazon Machine Images:

- ec2:DescribeImages

For more details, refer to the Knowledge Base article How to configure Amazon Web Services when using I AM Role with NetBackup.

These permissions should not be assigned to a user directly. Instead, they should be applied in a limited context using resource tags and prefixes of S3 bucket names.

Limit Access with Resource Tags

Resource tags are a recommended way to manage access to resources. When working with S3 buckets, NetBackup can be limited to the above permissions for specific buckets using a defined resource tag that only grants the above permissions when working with the specified buckets.

For example, in the following policy, the permissions indicated above are only granted to the user when accessing buckets that have names beginning with the string *nbu* (for NetBackup):

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
                "s3:CreateBucket",
                "s3:ListBucket",
                "s3:DeleteObject",
                "s3:GetBucketLocation",
                "s3:ListAllMyBuckets"
            ],
            "Resource": [
                "arn:aws:s3:::nbu*"
            ]
        }
    ]
}
```

For more information, see the details from Amazon on S3 access tags, user specific access, folder access and object tagging.

You can also use another comprehensive AWS pricing calculator to determine basic cost models.

About AWS Service Quotas

Every AWS account has default service quotas, which used to be referred to as service limits. These limits implement constraints on the number of RDS database clusters, the number of a specific instance type that can be running at once or how many elastic IPs the user is allowed.

For more information on service quotas, review the Amazon documentation.

AWS Service Quotas and NetBackup

Because NetBackup is running on dedicated EC2 instances and VPC that have predefined architectures, it's not likely to be impacted directly by service quotas. A specific limit NetBackup could experience as an issue is the maximum number of EBS snapshots (default 100,000). If this limit were hit, some snapshot-based backup operations would fail with a snap create failure. To avoid this problem, the user will need to request a quota increase in the AWS Service Quotas console.

Disaster Recovery Scenarios

Disasters can strike anywhere and at any time, even in cloud infrastructures. That's why it's always important to have a plan for how to recover from failures. Part of an effective plan is to have proper recovery testing. With NetBackup, there are two levels of DR: One is to recover from an EC2-based snapshot of the NetBackup Server, and the preferred level is to perform DR with NetBackup's built-in DR capabilities, which is preferred.

To test recovery from a snapshot, do the following:

- Stop NetBackup services to ensure consistency or the image will only be crash consistent
- Perform EC2 snapshots using the EC2 console, AWS cli or APIs
- Recover to a new instance in an isolated, dedicated VPC from the snapshot
- The VPC should prevent connectivity to production resources during testing
- Test functionality to validate that the recovery was successful
- Discard the test instance(s)

The preferred method of recovery is to use NetBackup's built-in recovery capabilities. Doing so involves a Catalog Backup Policy, an isolated DR backup location such as S3 or an EFS share and a recovery wizard to recover in the following scenarios:

- Corruption of the database
- Loss of important configuration information
- Rollback after an upgrade failure or other significant change
- Recovery from complete loss of the system

The recovery wizard will walk you through the recovery process. Depending on the location of the catalog backup, you may need to do some configuration to access the DR data, such as mounting the EFS share. Complete details on DR and NetBackup is available in the NetBackup Troubleshooting Guide's Disaster Recovery section. If you're using MSDP or MSDP-C, be sure to enable the MSDP DR policy (see the details here).

Cost Overview

The cost of the cloud will vary depending on what a customer needs. Simple backups to S3 storage can be a cost-effective solution for a customer that wants to send important data to an off-site location. However, this solution is probably not ideal for a large customer with a large amount of data to send to S3 due to the bandwidth constraints of the network. In addition, S3 storage is limited in options based on object versus block-based storage. There is a cost to send the data, store the data and retrieve the data.

The Cost of Gets and Puts

When writing data to Amazon S3 object storage, there is a cost associated for each time you or an application uploads/updates a file or object (PUT) and every time you retrieve an object (GET) from Amazon S3. To optimize data transfer to S3, NetBackup breaks data down to 64 MB of deduplicated data before sending it to the configured S3 bucket. Each 64-MB chunk write or read will incur a GET or PUT request. For more information on costs, see Amazon S3 pricing.

Storage Costs

Just as in the above example of costs associated with putting objects into Amazon S3, there is also a cost for using S3 storage. The costs and available S3 storage types will vary based on region, so be sure to check prices in your intended region when calculating costs (see Amazon S3 pricing).

Compute Costs

EC2 environments whereby machines are configured in the cloud using EBS disk will incur additional cost based on the number of processors needed, RAM usage and amount of disk provisioned. Backups in this environment will also incur costs based on moving the data from the client to the NetBackup environment. This cost is dependent on the location of the NetBackup Media Server in relation to the source client or data. Most options in cloud-based computing come "à la carte," where customers pay only for what they use. In some cases, these costs can be less than maintaining a data center and in some cases, the cost is more. That said, the peace of mind that comes from knowing the data is highly available and at a relatively secure off-site data center might be worth the extra expense of the cloud environment.

To better understand the cost structure, AWS has created a number of calculators you can use to determine approximate cost options. For example, the TCO calculator allows the customer to enter information about a planned deployment and receive cost estimates. Another comprehensive Pricing Calculator to determine basic cost models is also available from AWS.

When you deploy an Amazon EC2 instance, the cost of the instance is determined by the hardware type (CPU, memory), EBS volume usage and utilization. For example, running a small NetBackup instance to protect less than 100 TB for front-end data would cost about $301 (compute of $250 and EBS storage of $50.50) per month, with NetBackup running 16 hours per day. If NetBackup doesn't need to run for 16 hours per day, then the total cost decreases.

As noted, cost should not always be the determining factor when it comes to a cloud-based solution. There is more to a data center than hard costs. The AWS infrastructure can provide additional options that might be more expensive up front; however, the flexibility provided by on-demand storage, uptime guarantees and staffing may merit a move to AWS. And remember that NetBackup will be there to protect the data that is in the cloud.

Deployment Details

NetBackup VPC Deployment Configurations

NetBackup uses a three-tier architecture that consists of a single Management Server, multiple Media Servers and clients, which allows flexibility when deploying NetBackup in AWS. By default, NetBackup supports single-AZ (Availability Zone) and multi-AZ as well as multi-Region environments.

NetBackup Deployment Options

To optimize data movement as well as costs, we recommend that you place the NetBackup Management Server in the Region where management of all EC2 instances takes place as well as in a subnet where it can communicate with the deployed Media Servers. Media Servers should be placed as close to the target EC2 instances as possible; however, as long as network connectivity exists, a Media Server in one AZ can protect clients in different AZs (see Figure 19).

There are two deployment options when customers require multi-Region protection: a NetBackup domain in each Region or a single NetBackup domain with Media Servers in each Region. By using a NetBackup domain

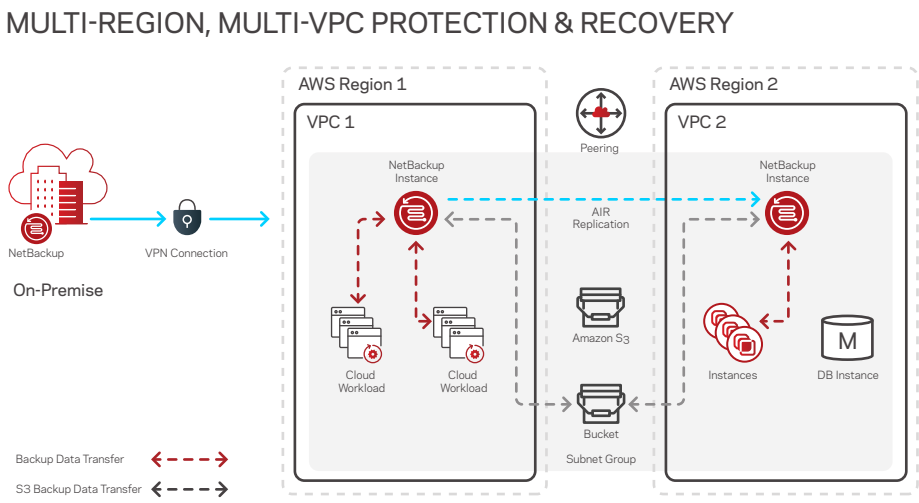


Figure 19. An example of a NetBackup multi-Region deployment in AWS.

for each Region, traffic will be isolated in each Region and NetBackup AIR can be used to offer DR in case of a Region failure. The downside to this approach is that each NetBackup domain will need to be managed independently.

Using a single NetBackup domain and deploying Media Servers into each Region eliminates the need to manage multiple NetBackup domains independently; if there is an outage or failure, however, all backup operations are impacted for all Regions.

Protecting NetBackup Access with EC2 Security Groups

Security and data protection are still a requirement as data is migrated to the cloud. When deploying NetBackup into AWS, there are two ways to limit network connectivity and access to the NetBackup instance: security groups or VPC access control lists (ACLs).

An EC2 security group acts as a virtual firewall for your instance to control inbound and outbound traffic. These security group firewall settings are attached to the EC2 network interface to only allow connections to NetBackup and out.

To limit the inbound connectivity to NetBackup, create a security group to control access to the EC2 instance or a network ACL to control access to the subnet with the custom TCP settings shown in Figure 20.

Inbound rules | Outbound rules | Tags

Inbound rules

Type	Protocol	Port range	Source	Description
Custom TCP	TCP	10102	65.52.193.0/32	spad
Custom TCP	TCP	1556	65.52.193.0/32	pbx
SSH	TCP	22	65.52.193.0/32	SSH for Adm
HTTPS	TCP	443	65.52.193.0/32	HTTPS
Custom TCP	TCP	10082	65.52.193.0/32	spoold

Figure 20. Using NetBackup to control EC2 access with a security group of a network ACL.

These inbound security group settings will allow all NetBackup Management Servers, Media Servers and clients to communicate with each other on the network. For NetBackup Management Servers, the pbx (1556) and https (443) ports need to be accessible. When using MSDP, the spad (10102) and spoold (10082) ports are needed for clients and other MSDP servers to communicate with the storage pool. As a best practice, access should be restricted to only the VPC, hosts or CIDR blocks that need to work with the NetBackup environment.

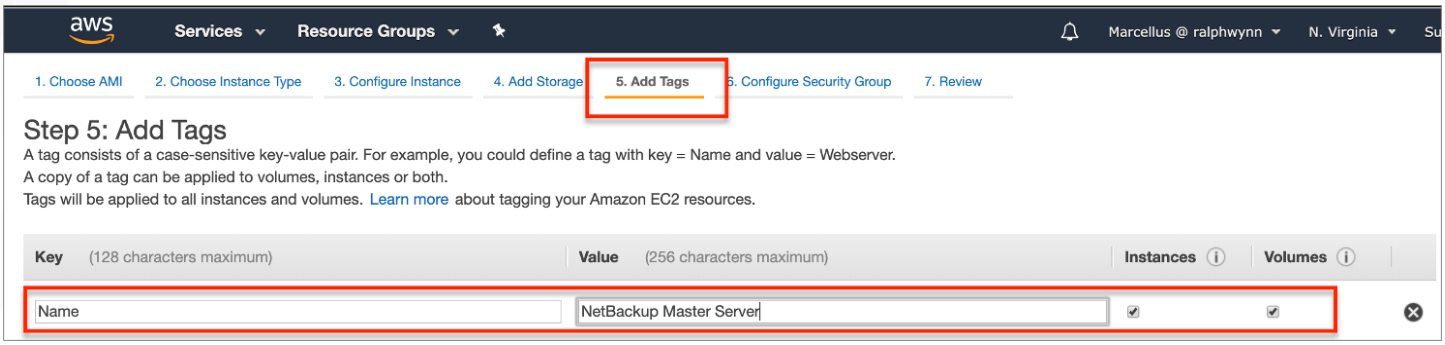
Tagging NetBackup Resources

Tagging your AWS resources is an important process because it allows for more precise filtering, searching and reporting. In relation to data protection, tagging all NetBackup resources will help identify all the resources needed to provide NetBackup data protection to the targeted AWS infrastructure.

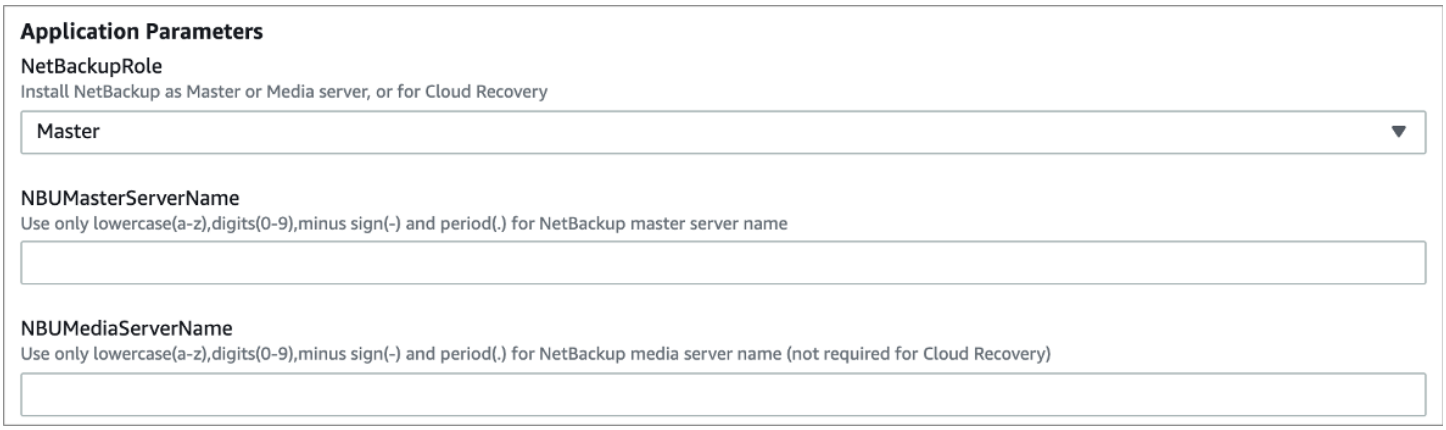
There are two ways to tag NetBackup resources: when launching an EC2 instance or post deployment.

Tagging During Launch

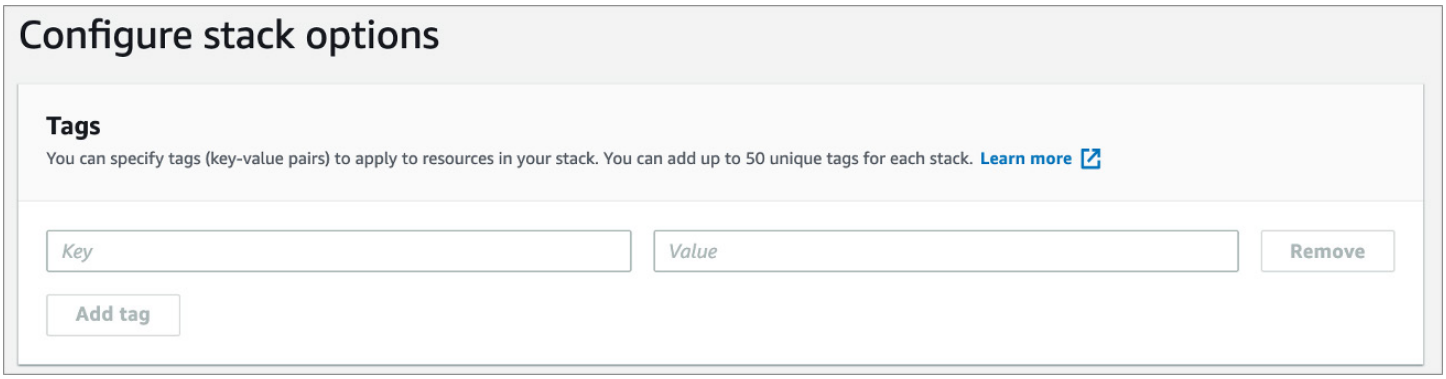
1. Using the EC2 Launch Instance Wizard, under **Add Tags**, specify the Key of the **Name** and the value of the desired NetBackup server name. Make sure **Instances** and **Volumes** are checked.



2. If you are deploying NetBackup using the CFT, the template will automatically create a tag for the instance name.



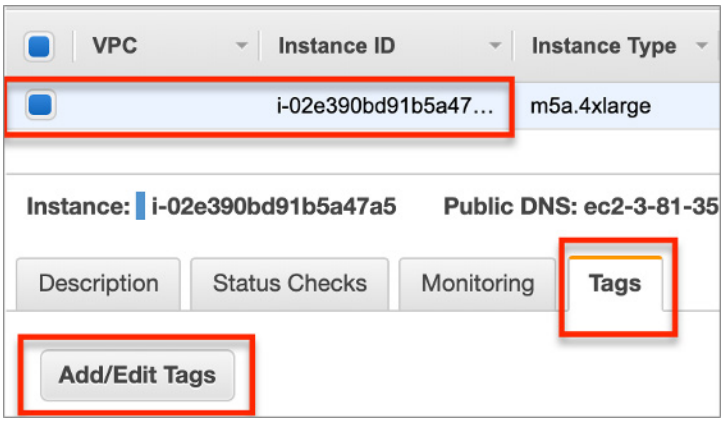
3. If you need additional tags, you can create them under **Configure stack options**.



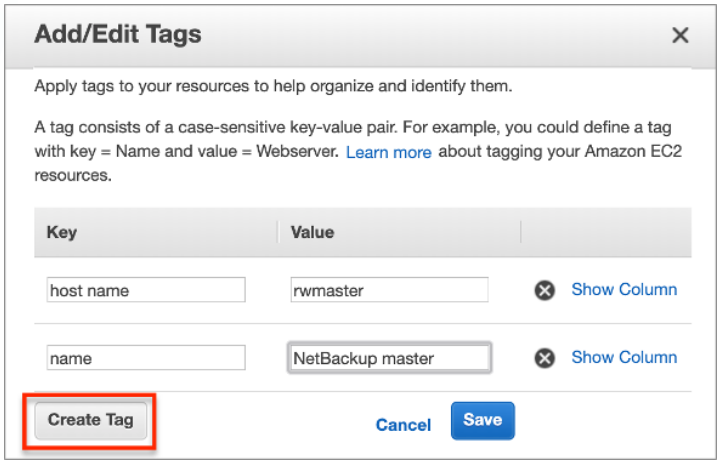
Tagging Post Deployment

You can add a tag to NetBackup resources already in use at any time.

1. Select the NetBackup EC2 resource that requires a tag.
2. Select the **Tags** tab and select **Add/Edit Tags**.



3. Specify a **Name** (such as name) and **Value** for this tag. Click **Create Tag**. You can add up to 50 tags per resource. You can also hide or show the column for each tag. When you are finished adding tags, click **Save**.



4. Once saved, the Name column will now display the name you provided as well as any additional tags you created.

Rotating AWS Access Keys for MSDP-C

NetBackup uses a MSDP-C server to send deduplicated data to AWS S3 buckets. MSDP-C uses IAM roles and access keys to make API calls to the AWS S3 services. If AWS access keys are configured, they are stored encrypted with the MSDP-C server. If security policies dictate that access keys must be changed, you will need to update each MSDP-C server configured to use the access key.

To update the server with a new key requires the use of the tpconfig command:

1. Connect to the MSDP-C server using ssh.
2. Under `/usr/openv/volmgr/bin, run` **`tpconfig -dsh -all_hosts`** to show the current access key user and the storage server type for Amazon.
3. Run the following command:

   `tpconfig -update -storage_server server name -stype server_type -sts_user_id AWS Access Key -password`

   ```
   Storage Server:          amazon.com
   User Id:                 AKIAT3NJE7RRGNTO4XZG
   ```

   ```
   Storage Server Type:     PureDisk_amazon_rawd
   ```

   `AWS Secret Key`

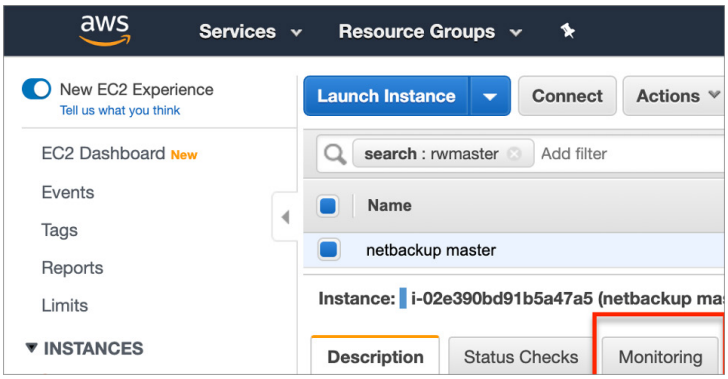
4. Once completed, the output of **`tpconfig -dsh -all_hosts`** will reflect the new access key. Once this change is made, NetBackup will use the new access and secret key to write to the configured S3 bucket.

Protecting NetBackup from Faults, Failures and Downtime

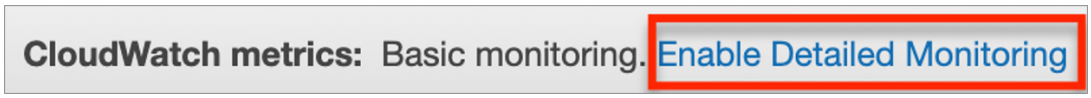
NetBackup uses the AWS EC2 service to provide data protection in an AWS VPC. This setup means that if there is an Availability Zone or Region failure impacting the EC2 service, NetBackup will be affected as well. Although those types of failures are uncommon, network connectivity can be interrupted, EC2 instances can go offline and EBS volumes can become corrupted, so choosing the right deployment model that optimizes DR (see the section on disaster recovery) is important to plan for potential failure.

Using alarms and monitoring for the NetBackup environment, you can detect and avoid the long-term risk of downtime or failure from an Availability Zone, instance or application fault. You can also use CloudWatch alarms to alert users when data written to an S3 bucket exceeds a defined threshold. To do so, create CloudWatch alarms for System Status and Instance Status Check failures for NetBackup Server EC2 instances by following these steps:

1. Select the NetBackup EC2 Instance and then select the **Monitoring** tab.

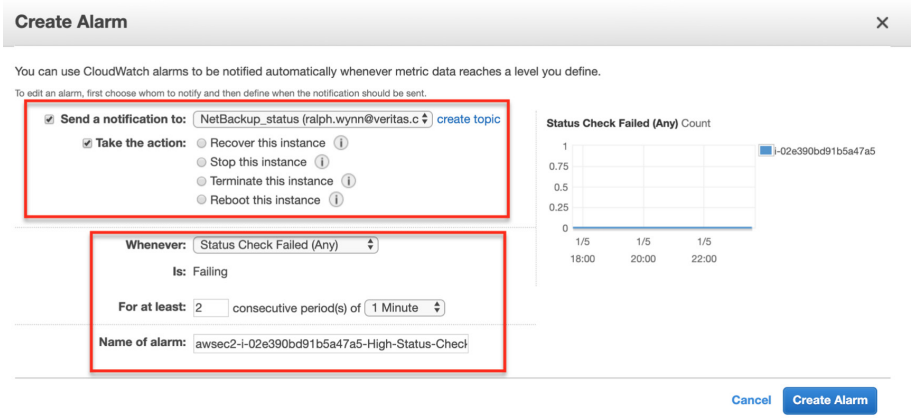


2. Click **Enable Detailed Monitoring** so metrics will be captured every 1 minute.



3. On the monitoring table, click **Create Alarm**.

4. For the alarm, specify (or create a new topic) to receive alerts of status check failures.
5. Select **Take the action** and then select **Reboot this instance** to enable this action whenever any status checks fail.



6. If the NetBackup EC2 instance has a hardware (AWS infrastructure) or a software failure (OS, memory), AWS sends a notification and performs the configured action.
7. You can use additional alarms to monitor network packets in and out as well as write operations to attached EBS volumes to detect possible network or storage failures and provide remediation steps.
8. Creating a CloudWatch alarm for the S3 Metric of BucketSizeBytes will send a notification when data written to an S3 bucket exceeds a threshold.

NetBackup Risk and Audit Management

NetBackup has built-in role-based access controls (RBAC) that can limit and control what actions users can take when performing backups and restores as well as adding clients.

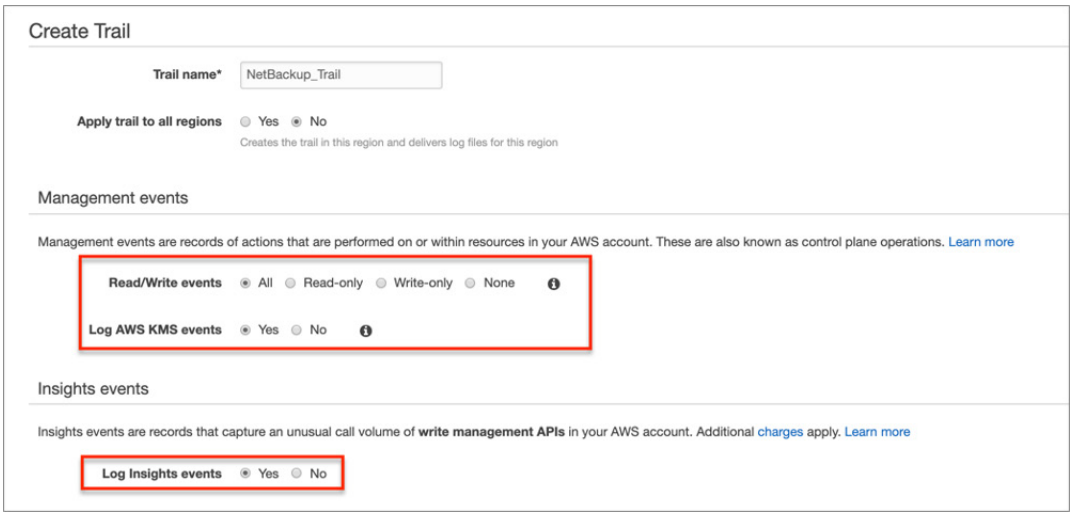
For more information on using NetBackup RBAC, see the NetBackup Web UI Security Administrator's Guide.

If you need additional info about what API calls NetBackup used to communicate with AWS S3 or what API calls NetBackup is making in the target VPC, CloudTrail can provide the right level of detail. You can also use CloudTrail to find out which account deleted an S3 bucket.

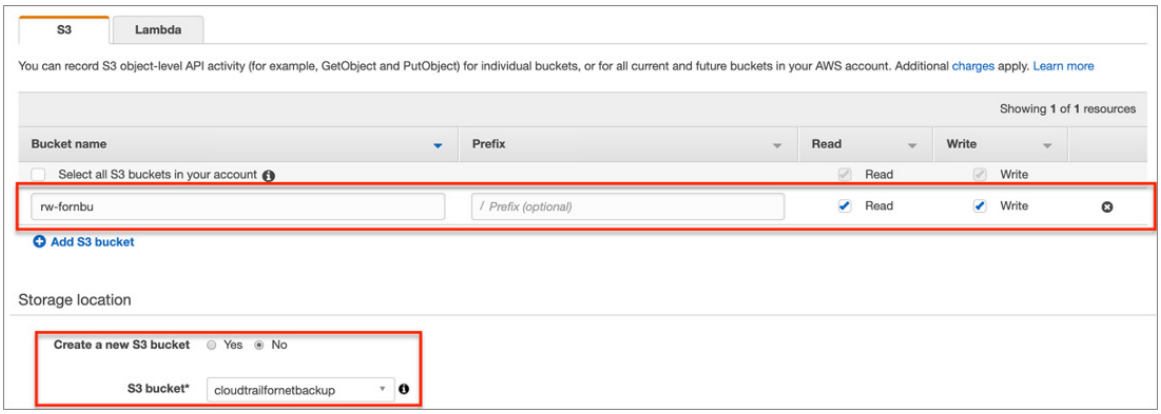
Enabling CloudTrail Logging for NetBackup Resources

To enable CloudTrail logging for NetBackup resources:

1. Create a CloudTrail that captures all create and delete API calls and provides Log Insights.



2. Specify either all buckets or the buckets used by NetBackup and the target S3 bucket for collections.



9. 3. Once created, CloudTrail will log all API calls to the target bucket. AWS recommends using AWS Athena to query the CloudTrail logs for actions by the NetBackup IAM user or Access key. For more information on using Athena with CloudTrail, see Using AWS CloudTrail to identify Amazon S3 requests.

AWS Scheduled Service Events

In rare cases, Amazon can schedule service events for instances. These events can include things like restarting or even retiring an instance. In the event of this scheduled activity, Amazon will notify the controlling user via email ahead of the event with additional details. It is important to make sure that the email address in your account is up to date to avoid an unexpected interruption.

When dealing with AWS scheduled service events, follow the instructions from Amazon under Scheduled Events for your Instances.

Summary

Customers are moving to the cloud and a number of cloud providers are moving to the forefront of the cloud megatrend. Amazon Web Services (AWS) and Veritas have partnered to create a usable, scalable solution for customers who want to create a cloud presence.

There are multiple paths to the cloud, which means proper planning and research is required to ensure the path you take will yield the expected outcome. In this document, we've highlighted the most common cloud use cases customers are deploying. By following the guidelines for these use cases, your cloud journey is more likely to be successful.

Appendix A – Additional Information

- Veritas Information
- Veritas NetBackup Cloud Administrator's Guide
- Veritas NetBackup Deduplication Guide
- AWS Marketplace
- NetBackup Resources
- Adding a cloud tier to an existing MSDP storage server
- NetBackup Security and Encryption Guide
- How to configure Amazon Web Services when using IAM Role with NetBackup
- AWS Service Quotas
- Protecting Amazon S3 objects using NetBackup with S3fs – white paper: v8.2
- How to resize an EC2 instance
- How to modify an EBS volume
- AWS RAID configuration guidance
- NetBackup Deduplication Guide
- AWS Regions and Availability Zones
- AWS EC2 Additional Information
- AWS S3 Additional Information
- AWS Glacier Storage Additional Information

Appendix B – Terminology

Amazon Elastic Block Store (Amazon EBS): A storage file system that provides persistent block storage volumes that can be used with Amazon EC2 instances in the AWS Cloud.

Amazon Elastic Compute Cloud (Amazon EC2): A web service that provides secure, resizable compute capacity in the cloud.

Amazon Machine Image (AMI): A file that provides the information necessary to launch an instance (or virtual server) in the AWS Cloud.

Amazon CloudFormation Template (CFT): A CFT allows for automation of deployment of services in AWS.

Disclaimer

THIS PUBLICATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID. VERITAS TECHNOLOGIES LLC SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS PUBLICATION. THE INFORMATION CONTAINED HEREIN IS SUBJECT TO CHANGE WITHOUT NOTICE.

No part of the contents of this book may be reproduced or transmitted in any form or by any means without the written permission of the publisher.

About Veritas

Veritas Technologies is a leader in multi-cloud data management. Over 80,000 customers—including 95 percent of the Fortune 100—rely on Veritas to help ensure the protection, recoverability, and compliance of their data. Veritas has a reputation for reliability at scale, which delivers the resilience its customers need against the disruptions threatened by cyberattacks, like ransomware. No other vendor is able to match the ability of Veritas to execute, with support for 800 + data sources, 100+ operating systems, 1,400+ storage targets, and 60+ clouds through a single, unified approach. Powered by Cloud Scale Technology, Veritas is delivering today on its strategy for Autonomous Data Management that reduces operational overhead while delivering greater value. 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.

V1072 09/22