VERITAS

NetBackup in the Public Cloud

Guidelines for Azure deployments.

This technical paper is designed to aid partners and customers looking to protect workloads with Veritas NetBackup™ deployed in the Azure 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

Contents

Introduction

The purpose of this white paper is to provide a technical reference on the capabilities of Veritas NetBackup and Microsoft Azure. 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, a link to which is provided at the end of this document.

Veritas has partnered with Microsoft Azure to offer a robust backup and recovery experience both 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 might 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 Microsoft Azure cloud infrastructure concepts. Users should be familiar with Azure Resource Manager Access Control using Azure Identity and Access Management (IAM), IAM Role Assignments and Compute, Networking and Storage Accounts concepts (Access Tiers and Blob types) 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 and ensuring greater business continuity with NetBackup, a solution that unifies data protection across the entire enterprise.

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 enable 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 the Azure Cloud as a supplemental data center—a hybrid of on-premises 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 are Customers Leveraging the Cloud?

Customers are using the cloud for several reasons. Smaller customers like not having to maintain a data center and an expensive disaster recovery (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 (JIT) cloud recovery. Large customers with data centers are identifying workloads that can take advantage of cloud availability and cost while freeing 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 types 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 very agile when it comes to requirements. Organizations can add disks to a server quickly and easily versus having to source 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 regularly, causing risk and downtime to install. Similarly, replacing or upgrading hardware impacts the environment by requiring a customer to manage them in the data center. In the cloud, however, 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. Veritas solutions allow customers to run their business seamlessly across physical, virtual or cloud infrastructure.

NetBackup and Azure Design Overview

There are many design cases when it comes to NetBackup and Azure. This section will outline them from a high level.

A NetBackup server with additional Media Servers and clients are collectively referred to as being part of a NetBackup domain. Veritas has created a solution template based on Azure Resource Manager (ARM) that deploys a NetBackup server instance in minutes from the Azure Marketplace. You can add additional data disks to the virtual machine (VM) based on the required capacity. You can configure these data disks to store backup data or backup images 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), NetBackup MSDP Cloud Tiering (MSDP-C) and MSDP with 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.

For more information regarding the automated deployment of NetBackup and Cloud Point, please refer to the following guide: veritas.com/content/support/en_US/doc/Marketplace_azure.

A Versatile Deduplication Engine

NetBackup MSDP is an intelligent, plug-in-based solution because precious CPU cycles don't need to be wasted to determine the file or block boundaries in the data stream sent by the NetBackup client. The deduplication plug-in interprets the metadata for the incoming backup stream to understand what kind of data is being protected and uses fixed or variable-length block size to segment the backup stream into unique blocks. The block size is also optimized based on the type of data being protected. This approach also works well in terms of compressing the data because different file types are split and compressed to different degrees with different block sizes based on the metadata in the backup stream. This approach provides a good balance between performance and resource utilization.

NetBackup's Deduplication Cloud Tier uses MSDP deduplication technology to upload deduplicated data directly to the cloud. This cloud storage server can be either a Veritas 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 Azure in different ways, depending on the needs of the customer. As outlined in various places in this document, NetBackup can use Block Blob Storage, Hot, Cool or Archive Access tiers like a regular disk pool or employ NetBackup MSDP-C to efficiently send deduplicated data to Block Blob or Archive Storage.

If a customer has resources in the cloud, NetBackup installed in the cloud protects these resources similarly 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 data in the cloud are as simple as in a local data center. The backup administrator has full use of the UI and APIs to recover data. Restore performance is relative to the type of storage (Hot, Cool or Archive), with restore from a Hot tier within the cloud being like that of on-prem recovery in a data center.

Azure 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—You can select the appropriate hardware to meet specific IOPS requirements.
- Azure Cloud—Select the appropriate disk type based on IOPS requirements when using a particular VM instance size. Cost varies based on the selected instance size.

Peer link limits

- On-Premises—You can have as many peer-to-peer links as required.
- Azure Cloud—There are fixed limits on the number of one-to-many VNET peering links allowed.

Storage targets

- On-premises—Systems typically write to block storage, deduplication devices or MSDP pools.
- Azure Cloud—Storage targets are typically object storage (Block Blob) or block storage disks (Page Blob) for storing data without dedupe optimization.

Use Cases Overview

There are different use cases with NetBackup and Azure. Because use cases vary with customer requirements, this section outlines a few of the more popular options.

Backup in the Cloud—Azure VM-Based Architectures

In addition to sending data to the cloud for DR or tape elimination, developing a solution that is completely cloud-native is also desirable. This concept is known as infrastructure as a service (IaaS), and many customers are finding that running workloads entirely in the cloud is more cost-effective and offers the ability to provision VMs, with the VM and all storage being in Azure. When protecting Azure-based workloads, it is important from a cost perspective to minimize data movement to on-prem by running NetBackup in the cloud as well.

Backups of these workloads are still required to protect them from corruption and malicious activity such as ransomware. Azure VMs function similar to a data center, using a hypervisor environment so there are built-in safeguards to improve data availability; however, failures and corruption can still occur. NetBackup in Azure IaaS works exactly like NetBackup in a data center. You can provision a NetBackup Management and Media Server from the Azure Marketplace using Azure Resource Manager (ARM) or manually deploy them in a BYO fashion.

For cloud-native workload protection, you can launch NetBackup CloudPoint from the Azure Marketplace and add it to the NetBackup configuration. NetBackup CloudPoint brings cross-cloud functionality and management from the NetBackup UI. CloudPoint 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 1).

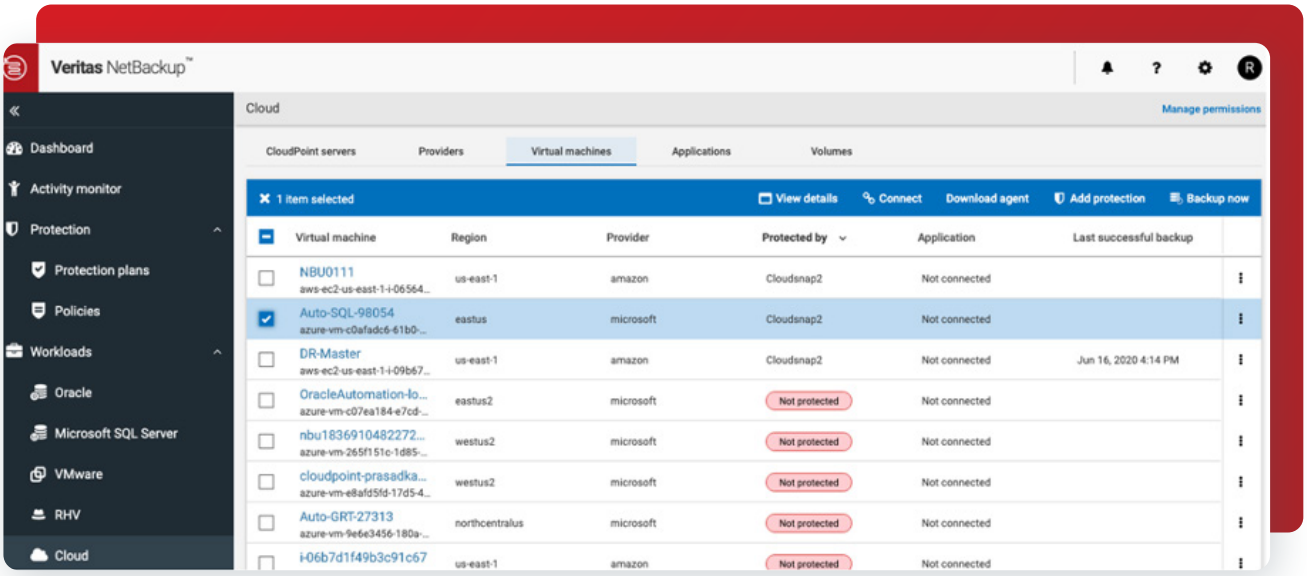


Figure 1. The NetBackup dashboard showing VMs protected by CloudPoint.

Figure 2 shows a Media Server running an MSDP so full and incremental backups will be deduped onto local fixed disks attached to the VM (Page Blob/Data Disk). The storage target could also use Block Blob (or object) storage that you provision using a storage account with one of the Hot, Cool, or Archive access tiers. For optimal storage cost savings, use MSDP-C or MSDP with cloud tier to store duplicated data in Block Blob storage. Each option has benefits, depending on the organization's needs and the data type. Azure VM instance types and sizing recommendations are covered later in this document.

NetBackup offers an ARM-based solution template in the Azure Marketplace to simplify and automate NetBackup server provisioning.

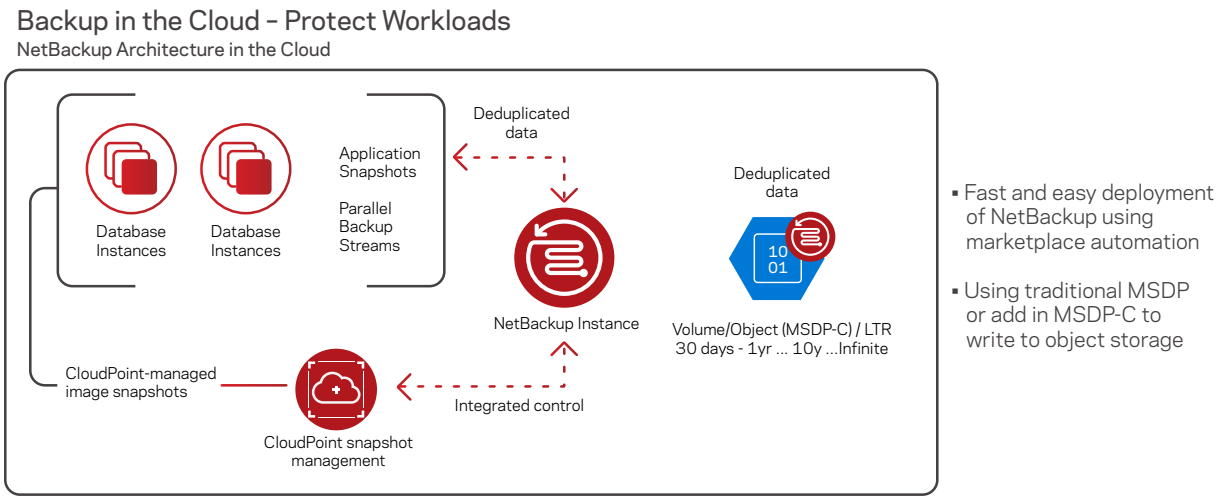


Figure 2. Sending workloads with NetBackup to Azure VMs.

Similar to on-premises deployments, NetBackup in the cloud can also send data to Azure using OST-compatible third-party dedupe solutions. NetBackup treats these solutions as basic disk storage in a data center. The third-party solution performs the work to dedupe the data before sending it to Azure Block Blob storage. These third-party deduplication solutions are not compatible with MSDP and will require the data to be rehydrated from its proprietary format before sending to MSDP. There are many other configuration options using NetBackup with Azure that you can tailor to the customer's needs. These use cases outline a handful of them.

NetBackup Intelligent Cloud Policies

Most cloud resources use 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 last week, is 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 make life easier for cloud administrators with features like auto-discovery and SQL-query-based definitions to logically define asset groupings (see Figure 3). Creating a Cloud Intelligent Policy 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 the 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.

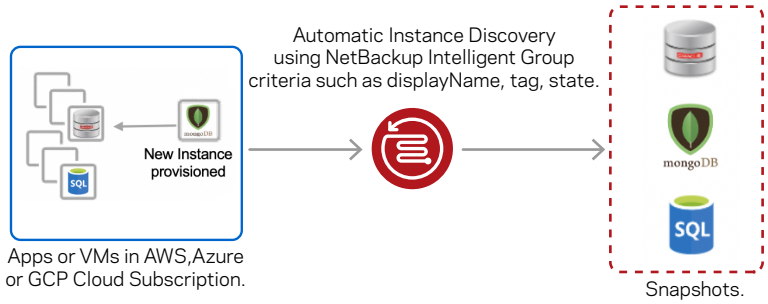


Figure 3. Automated cloud asset discovery.

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.

Backup to the Cloud—LTR Solution Using Microsoft Azure Archive

NetBackup helps you move data to Azure Archives. When writing to an Azure Archive, you must use a storage account with the Hot access tier. Image data and metadata are written to a Hot access tier and object data comprising image fragments are reclassified to the Archive tier (see Figure 5). In addition, with NetBackup MSDP-C now supports Azure Archive as a target, allowing for deduplication to low-cost cloud storage and dramatically lowering the cost of long-term retention (LTR) in the cloud.

Restoring data from an Azure Archive tier requires a retrieval operation where data is recalled from the Archive tier to a Hot tier. Once the image fragments are available in the Hot tier, they're read by the NetBackup Media Server and data is sent to the client for restore purposes. Once the restore is complete, the image fragments in the Hot tier are moved back to the Archive tier. (See Figure 4.)

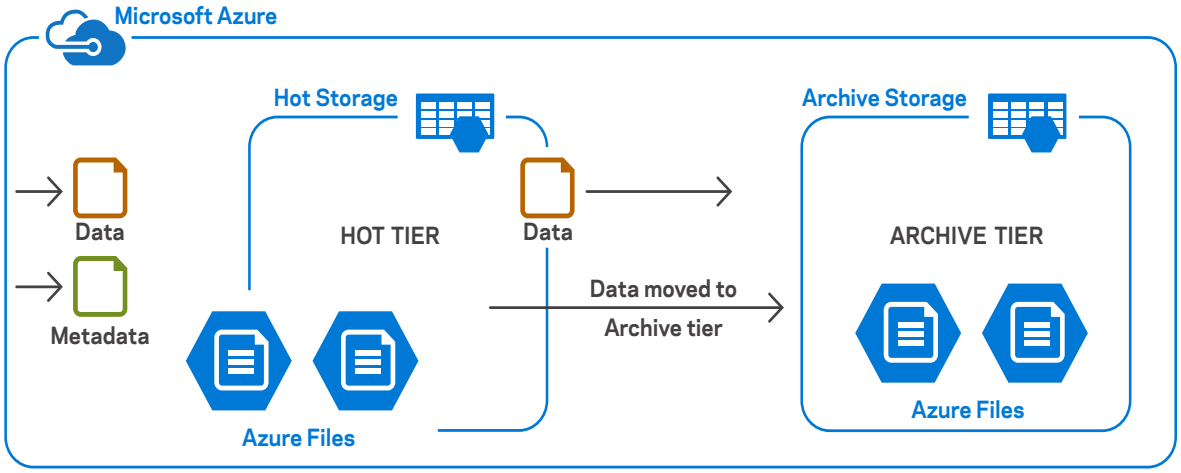


Figure 4. Backing up to an Azure Archive using a Hot tier storage account type.

Standard Backup from On-Premises to Azure Block Blob Storage

With NetBackup, the simplest way to move data to Azure object storage is to use the built-in Azure cloud connector. This interface allows you to configure a cloud object storage target such as those available in Microsoft Azure or Microsoft Azure GovCloud (see Figure 5).

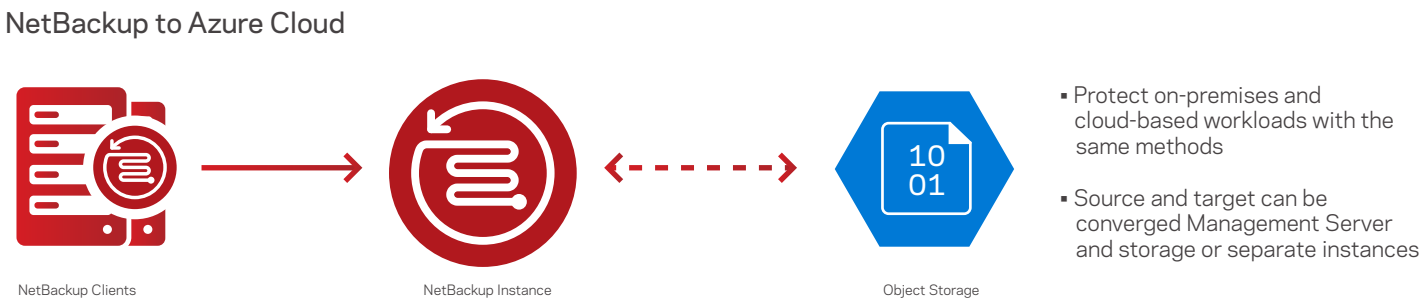


Figure 5. Backing up to the cloud with NetBackup using standard object storage.

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 Azure Storage Overview pricing page.

Optimized Backup to Block and Object Storage with Multi-Cloud Tiering

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 or 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 as 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 process optimized backup images from existing MSDP-compatible sources or directly from a client for transfer to an Azure Block Blob storage target. The Hot, Cool and Archive tiers of Azure Block Blob storage have been certified for use with NetBackup MSDP-C as well.

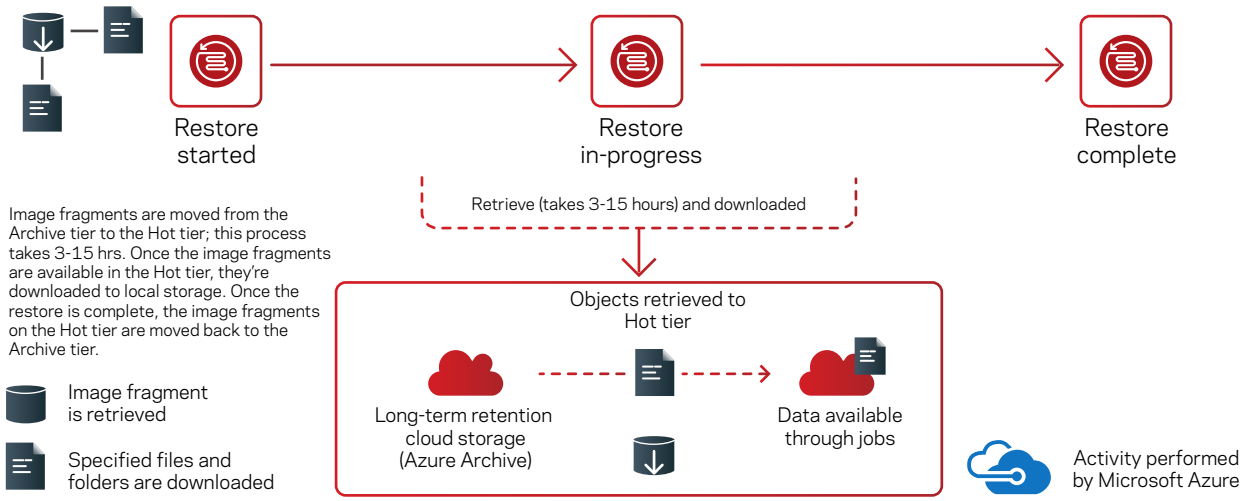


Figure 6. Restoring data from an Azure Archive with NetBackup.

When using any MSDP-compatible data as the source, the MSDP-C pool server 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 7).

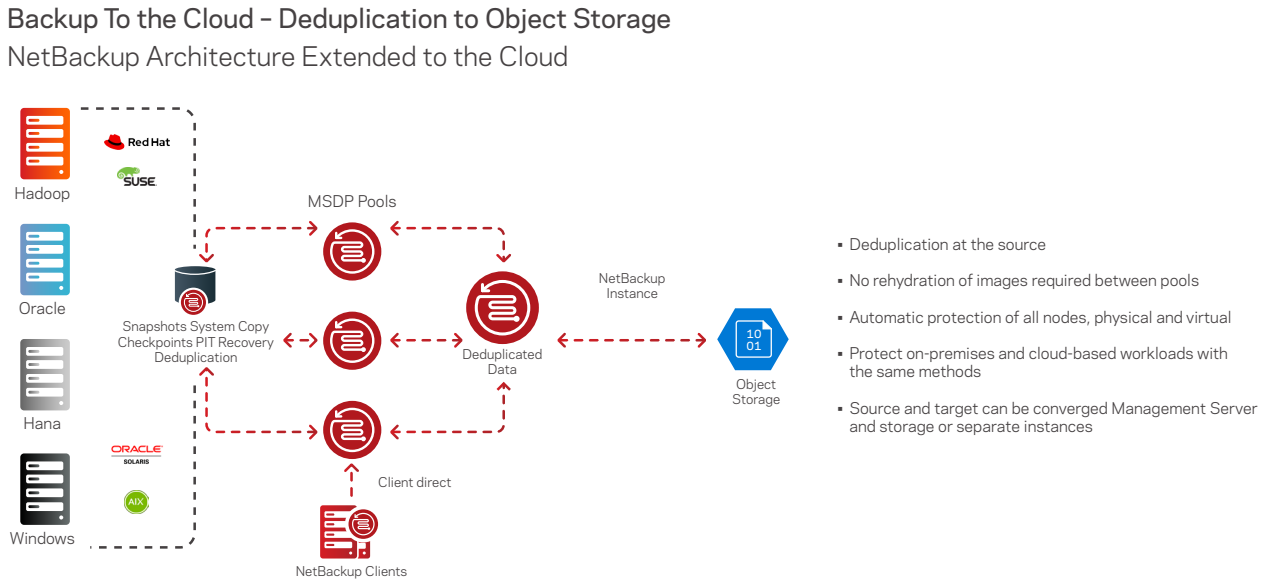


Figure 7. Protecting data sent to Azure Blob Storage using NetBackup MSDP-C.

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 (see Figure 8).

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.

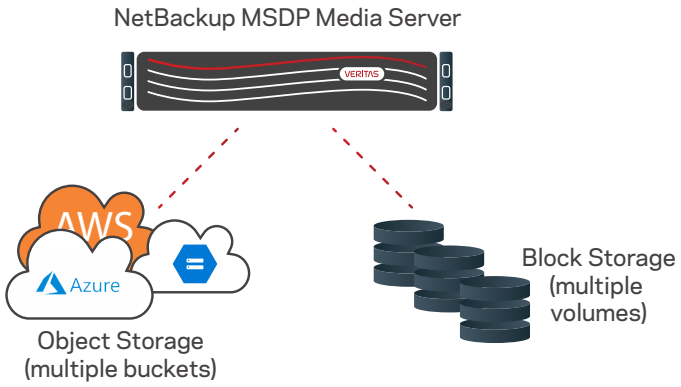


Figure 8. Using MSDP with cloud tier to send data to Azure and other providers.

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 multiple tenant usage. Optimized deduplication is supported within one MSDP server scope so that 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 the cloud or on-premises locations. You can use this approach to recover an on-prem system into a cloud infrastructure for DR, testing, or other cloud data reuse scenarios such as running cloud-based analytics against a database.

Sending Data from On-Premises to Azure Using a Third-Party Gateway Appliance

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 (see Figure 9).

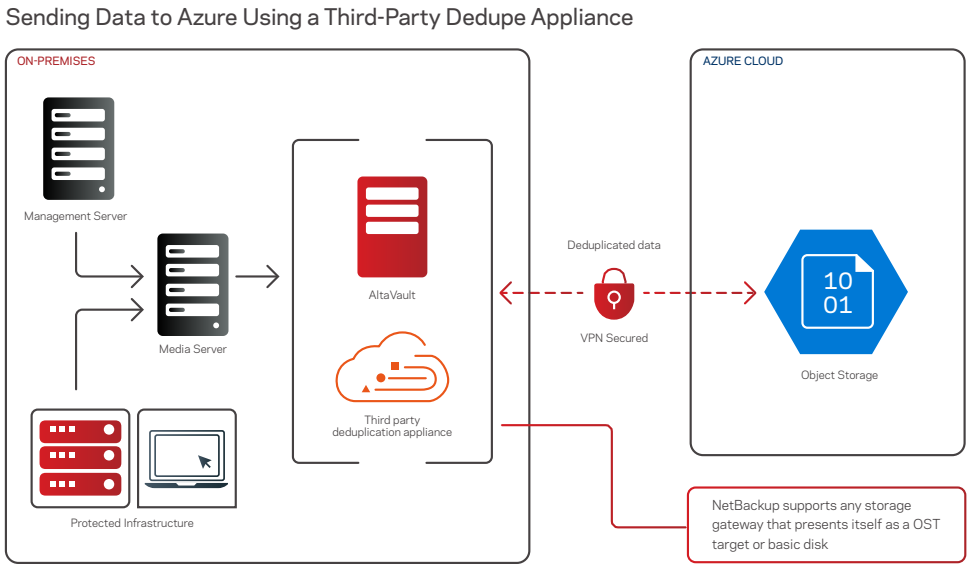


Figure 9. Sending data with NetBackup to Azure using a third-party dedupe appliance.

The appliance performs the deduplication before the changed blocks are forwarded on to the cloud via Azure Storage APIs. This solution will work with any Azure-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, which requires the data to be rehydrated before sending it to the gateway device.

Disaster Recovery Using Azure

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

One 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 second part is running in the cloud. Such a setup would allow you to duplicate and send a storage-optimized data stream from the data center to the cloud using AIR technology (see Figure 10).

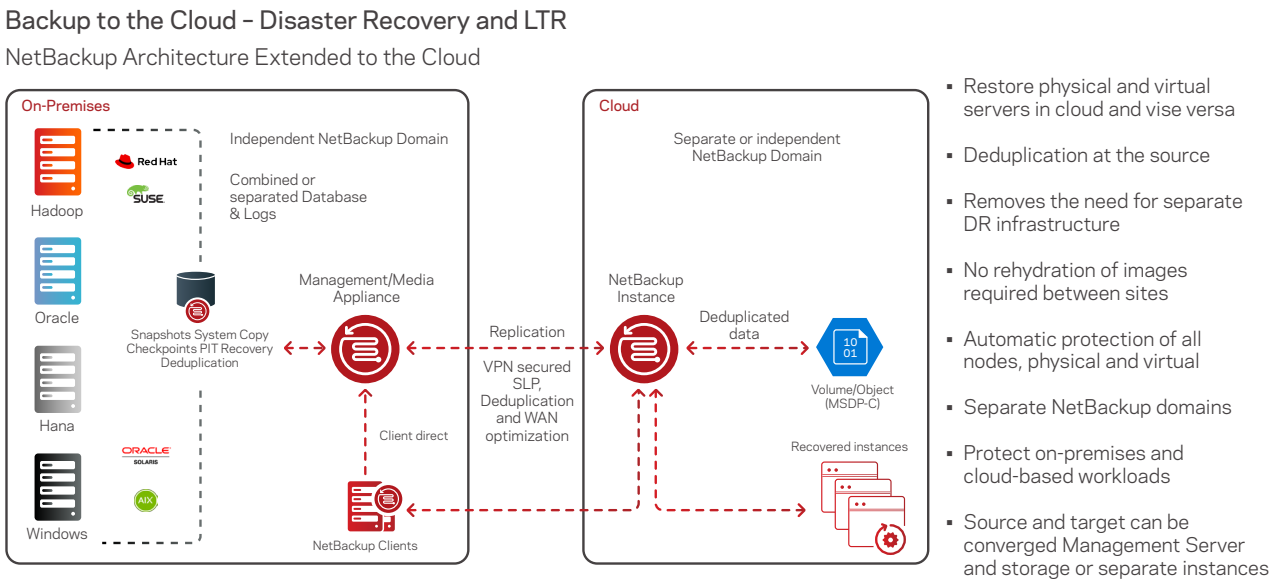


Figure 10. An overview of NetBackup Auto Image Replication (AIR) for cloud DR.

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 Azure. From there, you can use an AIR process to automatically send data from MSDP in the data center to MSDP or MSDP-C in Azure. The transferred data's metadata is encapsulated in the transfer so the import into the Azure NetBackup domain is nearly instantaneous after the data is copied. Customers have been using this model for global DR protection—to move data from a data center in San Francisco to a data center in London, for example—for a while. Leveraging this technology for a cloud target is no different for NetBackup. It's just another AIR target.

This option is ideal for a customer that wants an off-site DR copy of the data, and it's also a good way to migrate to the cloud from a NetBackup perspective. Veritas offers additional solutions such as the Veritas Resiliency Platform (VRP) that automate workload migration to the cloud and can integrate with 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. Resiliency Platform can also orchestrate workload recovery. Figure 11 shows how you can use NetBackup and Resiliency Platform to recover workloads into Azure.

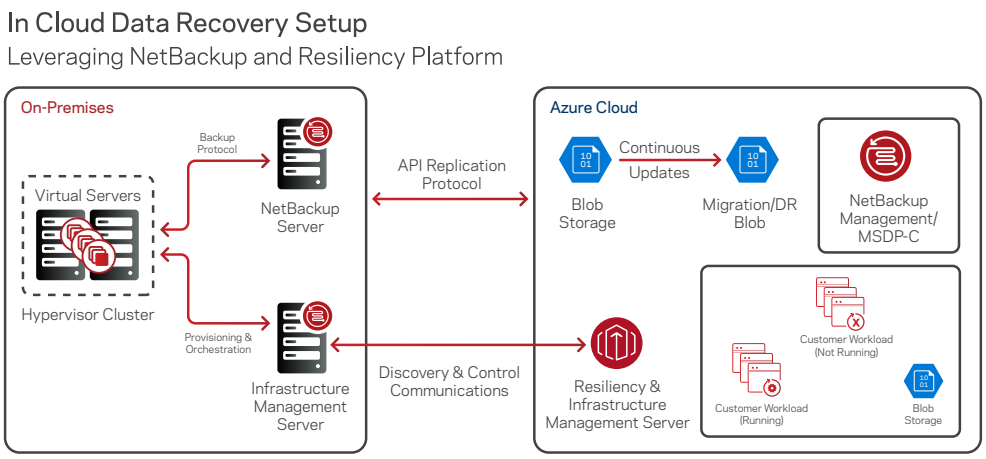


Figure 11. An example of cloud DR using NetBackup and Resiliency Platform.

Leveraging NetBackup Image Sharing for Migration and DR

NetBackup extends its image-sharing capability when using MSDP-C to write to Azure Blob storage, which is in addition to the cloud tier support for MSDP (see Figure 12).

This functionality makes the data in the storage container self-descriptive for reuse by an instance other than the one that initially wrote it. The only data needed to access the data is the storage account name and the necessary authorization.

Using image sharing, you can launch an on-demand NetBackup instance from the ARM Solution Template from the Azure Marketplace 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.

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 Azure on-demand instances to leverage MSDP-C data stores in object storage.

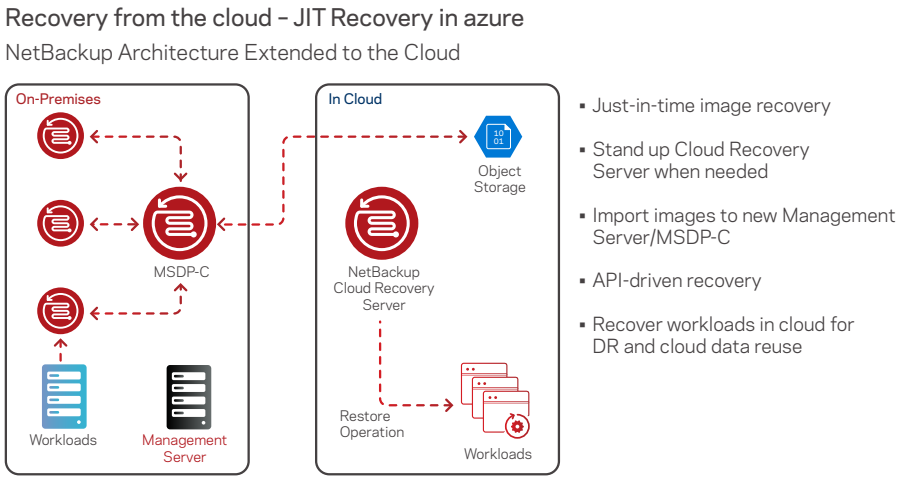


Figure 12. An overview of just-in-time recovery in Azure using NetBackup and MSDP-C.

Using NetBackup to Covert VMware VMDK Images to Native Azure VHD

With the latest release of NetBackup, 9.1 users can now convert VMware backup images into Azure VHD (only available for Azure) using an image-sharing server (Cloud Recovery Server), which is uploaded into an Azure blob. Image sharing in the cloud is a self-describing storage solution over MSDP-C and uploads the NetBackup catalog along with backup images. This approach allows users to restore data from the cloud without having an on-premises NetBackup server.

VMware VMDK to native Azure conversion is activated by default on the image-sharing server and does not need user intervention to be enabled.

Accelerator for Azure and Azure Stack

The Accelerator for NetBackup has been available since NetBackup 7.5 and can now be used with Azure and Azure Stack to increase the speed of full backups. To learn more visit: https://www.veritas.com/content/support/en_US/doc/18716246-142505646-0/v64540182-142505646.

The accelerator combines the benefits of the O/S file system change journal, an Accelerator Track Log, deduplication, and optimized backups to shorten backup windows while optimizing cost. After an initial full backup is performed, subsequent backups will check the cloud snapshot(s), and only backup changed blocks. By reading only the changed data from the cloud provider, you can minimize egress/ingress costs.

The Accelerator for NetBackup is enabled by default and supports the following NetBackup storage server types:

1. MSDP
2. MSDP-C
3. Open Storage
4. Cloud Storage

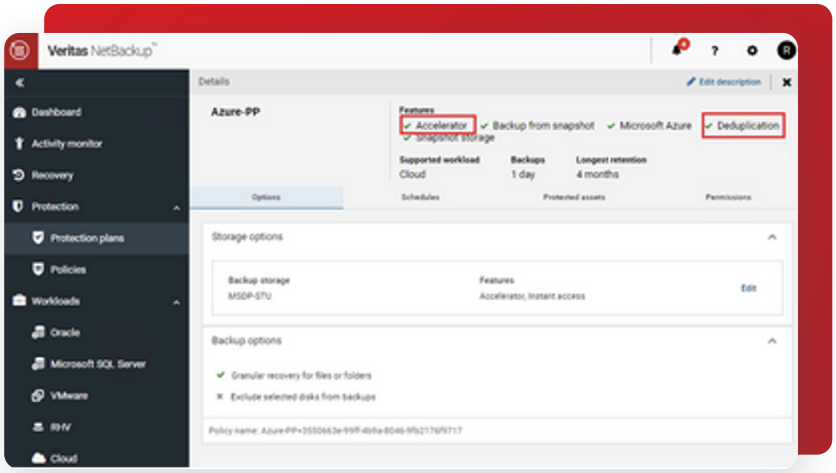


Figure 13: Protection Plans can utilize the Accelerator for NetBackup as of NetBackup 10.0.

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 less expensive storage tier, Veritas can provide storage savings up to 98 percent compared to snapshots. In addition, as compared to the cost of snapshot storage by cloud providers such as Azure, NetBackup can reduce the data stored by up to 50 percent.

Use Case

- Long Term Retention of point in time data copies (compliance, discovery, litigation support)
- Snapshot storage is cost prohibitive
- Example: *1TB annual storage cost:
- Snapshot: $614
- Backup stored on Archive tier (w/dedupe): $12

Retention: Daily snaps for two weeks, monthly for one year.

- EBS Snaps priced at $.05/gb/mo
- Deep Archive priced at $.001/gb/mo

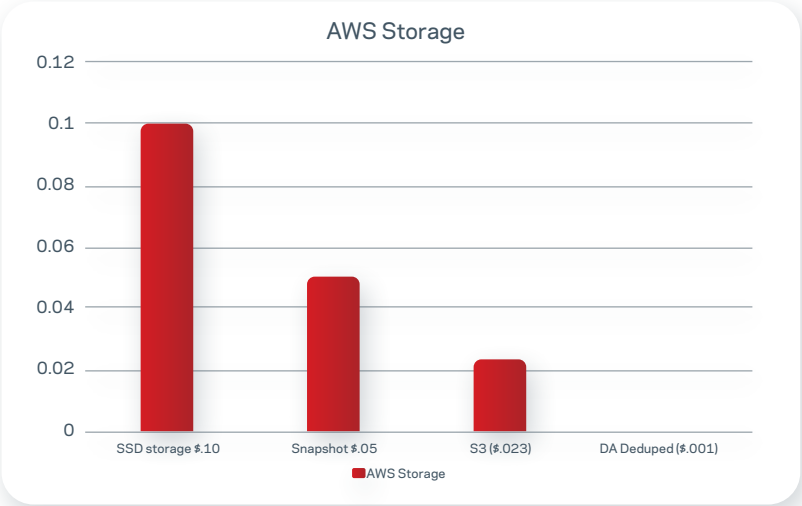


Figure 14: Cost of AWS Storage (SSD/Snapshot/S3/Long Term De-Duped).

Math:

Snapshot: .05 x 1024 = 51.2 x 12 months = $614 per year.
Long Term De-Duped: .001 x 1024 = 0.124 x 12 months = $12 per year.

Actual savings may vary

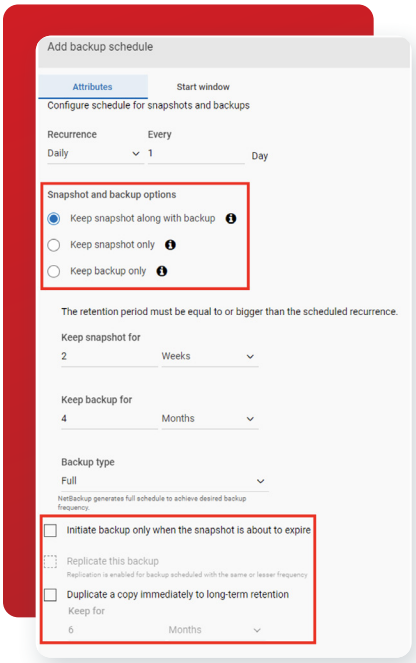


Figure 15. Configuring NetBackup snapshot and backup options.

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 cloud, NetBackup supports multiple storage tiers on 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.

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.

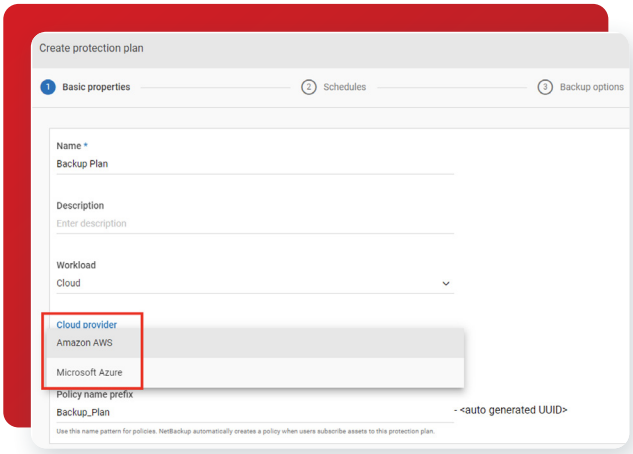


Figure 16. Configuring a NetBackup cloud protection plan.

Cost Considerations

The cost of the cloud will vary depending on what a customer needs. Simple backups to Azure Storage 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 Azure due to the bandwidth constraints of the network. In addition, Azure Storage is limited in options based on object versus block-based storage. There is a cost to send, store and retrieve the data.

Storage Costs

Just as there are costs associated with putting objects into Azure Storage, there is also a cost for using Azure Storage. The costs and available Azure Storage types will vary based on region, so be sure to check prices in your intended region when calculating costs (see Block Blob pricing). This is where MSDP comes in as a powerful way to cut storage costs with deduplication.

The Cost of Gets and Puts

When writing data to Azure Storage, there is a cost associated each time you or an application uploads/updates a file or object (PUT) and every time you retrieve an object (GET) from Azure. To optimize data transfer to Azure NetBackup breaks data down to 64 MB of deduplicated data before sending it to the configured Azure Storage account. Each 64 MB chunk written or read will incur a GET or PUT request.

NetBackup Storage Lifecycle Policies can help alleviate these costs by automating the movement of data to the cloud so it occurs only after a time the customer defines as beyond the likely restore period or by automating the removal of the on-site copy to occur only after that period. For example, if data is rarely requested from backups after two weeks, the customer can choose to replicate the data immediately for DR but maintain an on-site copy for two weeks to address any restore requests. NetBackup automation will manage the backup, copy, and expiration of images as part of the overall backup operation. This model optimizes on-site storage utilization and minimizes GET cost impacts from frequent restore requests.

Compute Costs

Azure environments where VMs are configured in the cloud using disk storage will incur an additional cost based on the number of processors needed, RAM usage, and the amount of disk provisioned. Backups in this environment will also incur costs based on copying the data from the VM 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.

The cost of running these workloads in Azure is part of the customer's business analysis to determine if moving a workload to the cloud provides a cost benefit. In some cases, these costs can be less than maintaining a data center. In other cases, the cost is more, which will drive the rate of cloud adoption. That said, the peace of mind that comes from knowing the data is highly available and protected can weigh the decision toward the cloud. NetBackup's ability to run natively in the cloud alleviates the PUT/GET penalty of protecting cloud workloads with an on-prem-only solution.

When you deploy an Azure VM, the cost of the instance is determined by the hardware type (CPU, memory), disk, and utilization.

As noted, cost should not always be the sole determining factor when it comes to a cloud-based solution. There is more to maintaining a data center than hard costs. The Azure infrastructure can provide additional benefits that might seem more expensive upfront; however, the flexibility provided by on-demand storage, uptime guarantees, and staffing may make a move to Azure make sense. Azure is a modern and efficient infrastructure that can support many business needs, and NetBackup will be there to protect the data cloud.

To better understand the cost structure, Azure has created a Cloud Adoption Framework you can use to determine cost models, business cases, and other planning considerations. You can input information about a planned deployment and review cost estimates right in the framework.

You can also use the comprehensive Azure pricing calculator to determine basic cost models.

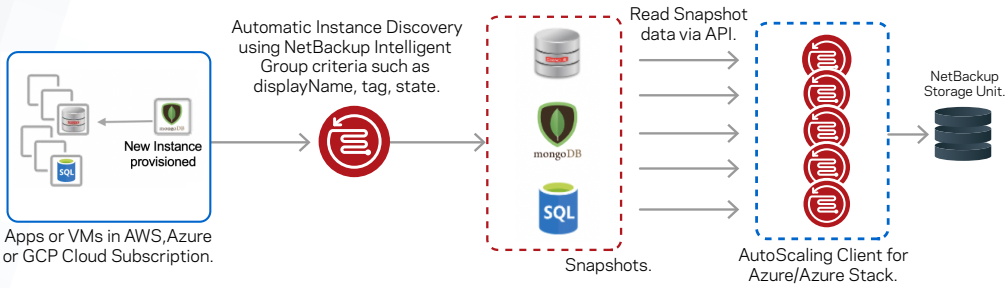


Figure 16. An example of NetBackup Cloud Autoscaling.

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 today for Azure Stack and Azure resources protection using the NetBackup CloudPoint 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.

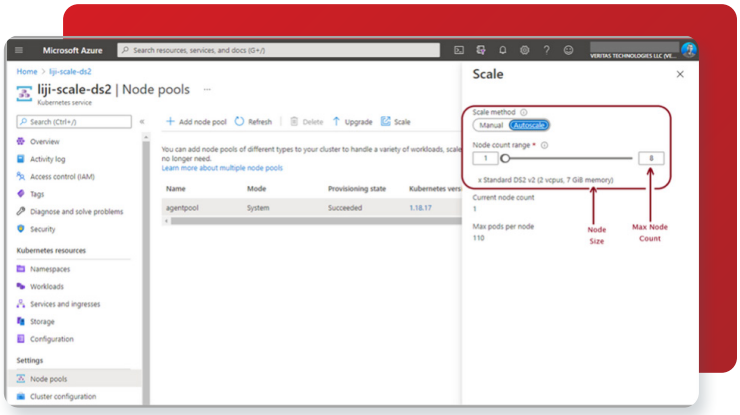


Figure 17. Estimate cloud data protection costs using Azure node pool count, node size and duration of your backup window.

For compute optimization, the NetBackup thin client is implemented using an Azure Kubernetes (K8s) cluster where the node pool Scale Method is set to Autoscale. Figure 17 illustrates the Scale settings for both node size and node count range for an Azure K8s node pool. The administrator can define the maximum node count and node size for each node in the pool. Doing so helps to establish a predictable cost for compute resources in the cloud. A cost-conscious approach is used here to ensure additional nodes are spun up only after the K8s cluster has determined the current node's capacity has been reached.

For Azure Stack Hub, an on-premises host or VM compatible with NetBackup CloudPoint is sufficient for the thin client containers. In this scenario, additional thin clients are launched on the host VM to complete the data protection workflow.

Organizations can stop pre-provisioning resources and running up cloud costs. Sustainability is a core value at Veritas, and when it comes to protecting cloud resources using NetBackup Cloud Autoscaling, organizations can be confident they only have to pay for the cloud resources they use.

Cloud Sizing and Performance

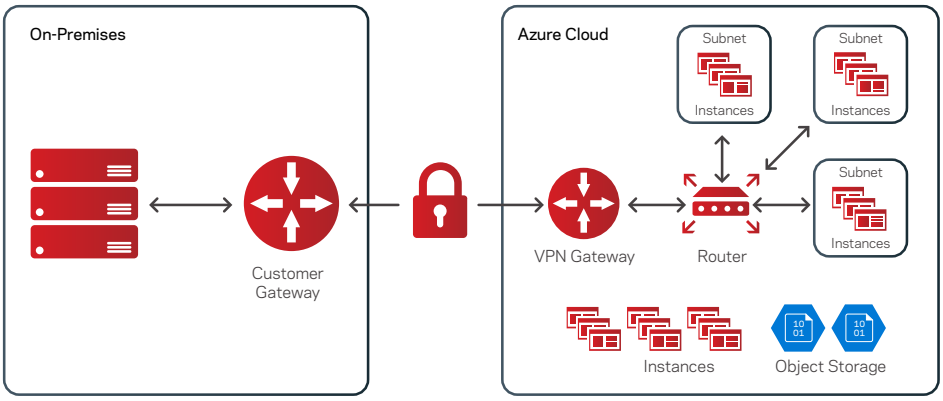


Figure 18. A general view of cloud infrastructure connected with an on-premises data center.

Sizing and performance in the cloud is based on customer needs and will vary from customer to customer. Therefore, use this information as a guideline and modify actual deployment as needed.

To get data to the cloud, customers can use any Internet connection if the available bandwidth allows all the data to be transmitted in the target timeframe (see Figure 18).

With Azure Express Route, customers can get a dedicated link to Azure with dedicated high-speed WAN connectivity in the data center. They can compress data at the data center prior sending it across the network to Azure or use MSDP-C to optimize the data being sent before it is transmitted to the cloud. They can also "throttle" bandwidth, if desired, to prevent over-saturation of the network pipe.

Azure Instance Model

Azure uses a regional model in which it has configured various regions across the globe. Many regions have Availability Zones, which are multiple data centers within the region that communicate with each other over high-bandwidth, low-latency 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 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 Azure Cloud and have the option to select a geographically dispersed region to provide regional DR.

Azure Region Pairs

Each Azure region is paired with another region within the same geography, making a regional pair. The exception is Brazil South, which is paired with a region outside its geography. Azure serializes platform updates (planned maintenance) across the regional pairs, so that only one paired region is updated at a time. In the event of an outage affecting multiple regions, at least one region in each pair will be prioritized for recovery. For more information, see the Azure documentation on business continuity and disaster recovery for paired regions.

Azure allows the selection of a storage account to be Geo-Redundant Storage (GRS) to provide an enhanced level of durability and protect against a region-wide disaster. For more information, see the Azure documentation on storage redundancy.

Azure Storage Options

One of the many benefits of the Azure 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, Azure uses various types of storage including object (Block Blob across Premium, Hot, Cool and Archive tiers) and block (Page Blob/Managed Disks), depending on the type of use case. Azure Block Blob Storage can offer a hierarchical name space (HNS) that provides access to the data via HDFS APIs for big data and analytics workloads. For more information, refer to the Azure documentation.

Other options include Azure Files or Azure NetApp Files for high-performance file system targets. Sizing of the environment is based on the customer's needs and the workloads placed in the cloud. Pricing is based on the type of storage chosen and is priced per GB, per transaction and on data egress. Hot storage has the highest cost per GB but the lowest transaction cost. Conversely, Archive storage has the lowest cost per GB but has a high cost for transactions and data retrieval. The latter is best suited for long-term storage, where there is low probability of data retrieval, and is often used as a tape archive replacement.

You can find details about pricing on the Azure Pricing Overview page.

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 Azure 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 Azure Instance Sizing

This example architecture is based on a single NetBackup domain consisting of a NetBackup Management Server, several MSDP Media Servers and a single NetBackup MSDP-C Server for Azure Storage.

Typically, backups are written directly to each Media Server's MSDP storage for an immediate copy, then duplicated via MSDP-C to Azure Storage. Alternatively, 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 Servers' block storage
  - MSDP Media Servers receive the initial backups from clients and perform deduplication.
- Multiple MDSP pools also help distribute copies across different regions and infrastructures.
- 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 Servers' block storage and stores them in Azure Blob storage.
- 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 and expected growth. Azure does offer the added benefits 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 details the minimum processor and memory requirements for the various environment sizes.

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

These estimates are based on the number of Media Servers and the total 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 – Azure Instance 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
Standard_DS4_v2	Standard_DS13_v2	Standard_DS5_v2	Standard_DS14_v2

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 Azure built on an Azure VM with Azure Data Disks (Page Blob) storage.

Specifications for MSDP Media Server in Azure

The host computer's CPU and memory constrain how many jobs can run concurrently. The storage server requires sufficient capability for deduplication and 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 Azure

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
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

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 Azure VM instance to meet higher requirements that may happen over time. For more information, refer to the following:

- Resizing an Azure VM
- Adding disks to an Azure VM

Media Server Deduplication Pool Recommendations

Running traditional MSDP in a cloud environment requires specific resources to be available such as a 10G network and managed disks with required performance. The recommendations below have been formulated using Azure kits that address MSDP pools of different sizes. These are just recommendations. Specific customer environments may have different needs. Depending on the Azure 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.

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

	Storage	Cores	RAM	Networking	IOPS
10TB (Small)	1x160 SSD 1x16 TB EBS-SSD	8	61		
1-20 TB (Small)	1x80 EBS-SSD 1x16 TB EBS SS	36	60	10 GB	EBS Provisioned IOPs (SSD)
32 TB (Medium)	1x80 SSD 2x16 TB EBS-SSD	16	30	10 GB	
	1x160 SSD 2x16 TB EBS-SSD IOPs – 12,000	8	61		12,000

32-64 TB (Large)	1x80 EBS-SSD 2-4x16 TB EBS-SSD	40	160	10 GB	
	1x80 EBS-SSD 2-4x16 TB EBS-SSD	36	60	10 GB	
	1x160 SSD 2x16 TB EBS-SSD IOPs – 12,000	8	61	10 GB	12,000
32-96 TB (xLarge)	1x80 EBS-SSD 2-4x16 TB EBS-SSD	40	160	10 GB	
	2x320 EBS-SSD 2-6x16 TB EBS-SSD IOPs -12,000	32	144	10 GB	12,000

Table 4: Recommended NetBackup MSDP-C Sizing Guidelines

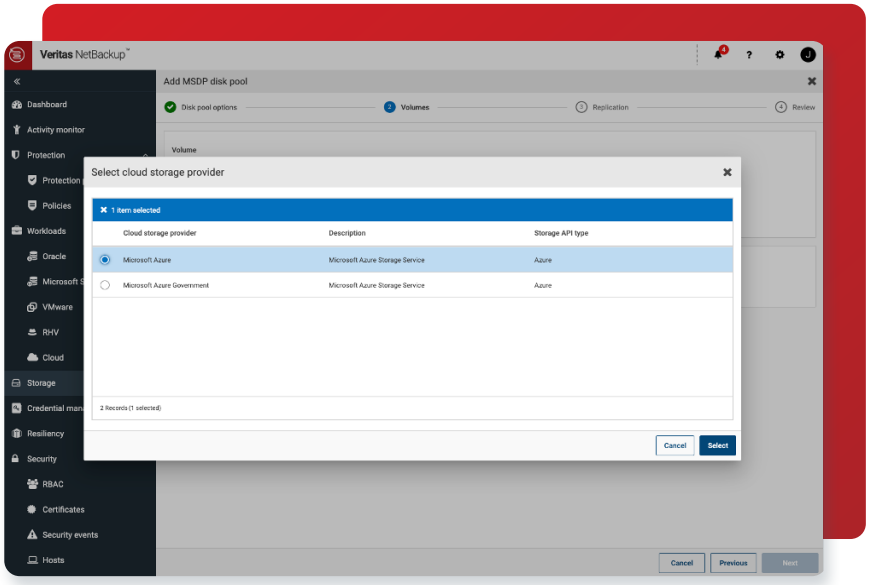
Product	Role	Instance Type	EBS/EFS Storage	CPUs	RAM(GB)
NetBackup	MSDP-C Min	RHEL M5.xlarge	250 GB SSD (gp2) 1+TB	4	16 GB
	MSDP-C Large	RHEL M5.2xlarge	500 GB SSD (gp2) 1+TB	8	32 GB

For NetBackup MSDP-C servers, Table 4 lists a minimum configuration and a large configuration along with the Azure VM instance type and the storage configuration. Customers should start with the larger instance recommendation unless they are using CloudCatalyst 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.

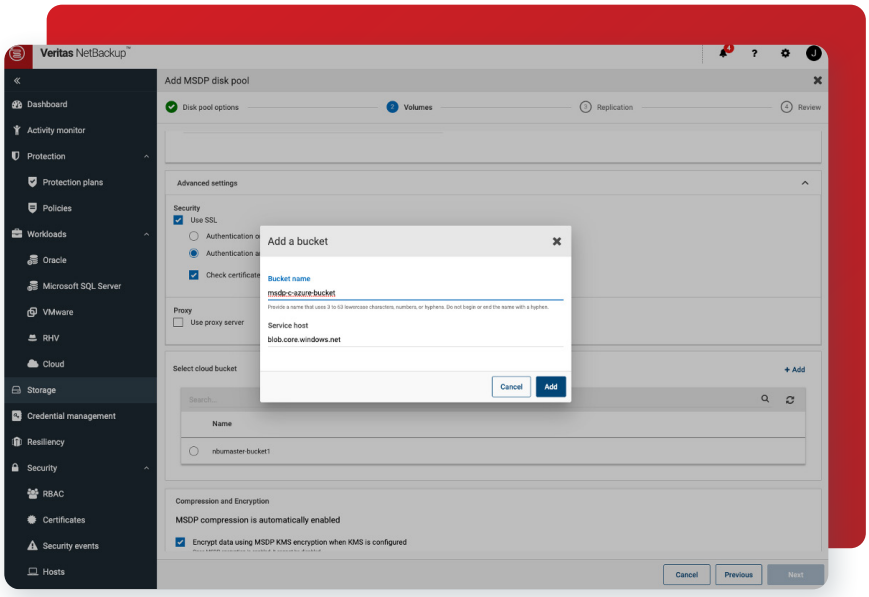
Adding a NetBackup MSDP Cloud Tier is easy.

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 Microsoft 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 Azure Archive), the region and enter the credentials for the storage account you want to use.



The final step is to select or create a new bucket for use.



After reviewing the inputs, click Next and Finish to complete the setup. Follow the remaining wizards to complete 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 starting from the client, in transit and at rest. In addition, any data that is sent to another MSDP 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 Azure 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.

Summary

Customers are moving to the cloud and many cloud providers are moving to the forefront of the cloud megatrend. Microsoft 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 should be successful.

Appendix A – Additional Information

Description	Link
Veritas Information	http://www.veritas.com
Veritas NetBackup Cloud Administrators Guide	https://www.veritas.com/content/support/en_US/doc/58500769-150013608-0/v79187862-150013608
Veritas NetBackup Deduplication Guide	https://www.veritas.com/content/support/en_US/doc/25074086-149019166-0/v95646212-149019166
Azure Marketplace	https://azuremarketplace.microsoft.com/en-us/marketplace/apps?search=veritas
NetBackup Security and Encryption Guide	https://www.veritas.com/content/support/en_US/doc/21733320-146139160-0/v141139708-146139160
NetBackup Resources	https://www.veritas.com/protection/netbackup/resources
Adding a cloud tier to an existing MSDP storage server	https://youtu.be/Q9wrBlYmCgo
How to Resize an Azure VM	https://docs.microsoft.com/en-us/azure/virtual-machines/windows/resize-vm
Azure Regions and Region Pairs	https://docs.microsoft.com/en-us/azure/best-practices-availability-paired-regions
Azure Availability Zones	https://docs.microsoft.com/en-us/azure/availability-zones/az-overview
Azure VMs Additional Information	https://azure.microsoft.com/services/virtual-machines/
Azure Blob Storage additional information	https://azure.microsoft.com/en-us/services/storage/blobs/
Azure Archive Storage additional information	https://azure.microsoft.com/services/storage/archive/

Disclaimer

About Veritas

Veritas Technologies is a global leader in data protection and availability. Over 80,000 customers—including 95 percent of the Fortune 100—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

V1073 09/22