simpliVity™

Technical White Paper

# SimpliVity Backup Hub

## Table Of Contents

## Executive Summary

This paper documents technical information on SimpliVity's Backup Hub offering along with the benefits it provides in terms of longer retention of data, improved RPO's and basic operational efficiency. Backup methodologies, use cases and best practices are also detailed in this paper.

### Purpose

The purpose of this paper is to introduce the reader to SimpliVity's Backup Hub offering. It discusses deployment options and use cases for the Backup Hub within a SimpliVity environment. Best practices and recommendations for backing up data is also covered in this paper.

### Audience

The audience for this paper are customers, IT professionals who want to learn about the Backup Hub product offering and want to deploy it within their data centers.

## Introduction

Acquiring, implementing and managing affordable backup storage for growing VM volumes, onsite and off-site, continues to be a challenge for enterprise data centers today.

A well-established data center strategy for reducing enterprise recovery point objectives (RPO) is to backup data more frequently. But even in highly virtualized environments, storing larger quantities of operational backups can be costly, depending on the size of the VMs, how much they change, and how often they are backed up. In order to meet increasingly stringent RPOs, enterprises that back up their data more frequently are seeking cost effective storage solutions for their growing volume of operational VM backups.

In addition, global enterprises today face a growing and ever-changing landscape of country-level and industry specific regulations associated with data retention and protection. Keeping data for longer periods helps organizations comply with a wider range of government regulations, but requires more backup storage, which again, can be costly. Enterprises today are looking for ways to cost-effectively migrate and store their operational backups at both local and offsite storage locations, to support longer data retention and backup aging policies that help them comply with expanded regulations for data retention and protection

### Why SimpliVity?

SimpliVity's hyperconverged infrastructure solution transforms the data center by virtualizing data and incorporating all IT infrastructure and services below the hypervisor into commodity x86 building blocks. With 3X total cost of ownership (TCO) reduction, SimpliVity provides software-defined hyperconverged infrastructure delivers the best of both worlds: the enterprise-class performance, protection and resiliency that today's organizations require, with the cloud economics businesses demand.

Designed to work with any hypervisor or industry-standard x86 server platform, the SimpliVity solution provides a single, shared resource pool across the entire IT stack, eliminating point products and inefficient siloed IT architectures. The solution is distinguished from other converged infrastructure solutions by three unique attributes:  accelerated data efficiency, built-in data protection functionality and global unified management capabilities.

- Accelerated Data Efficiency: Performs inline data deduplication, compression and optimization on all data at inception across all phases of the data lifecycle, all handled with fine data granularity of just 4KB-8KB. On average, SimpliVity customers achieve 40:1 data efficiency while simultaneously increasing application performance.
- Built-In Data Protection:  Includes native data protection functionality, enabling business continuity and disaster recovery for critical applications and data, while eliminating the need for special-purpose backup and recovery hardware or software. OmniStack's inherent data efficiencies minimize I/O and WAN traffic, reducing backup and restore times from hours to minutes.
- Global Unified Management:  A VM-centric approach to management that eliminates manually intensive, error-prone administrative tasks. System administrators are no longer required to manage LUNs and volumes; instead, they can manage all resources and workloads centrally, using familiar interfaces such as VMware vCenter and VMware vRealize Automation.

The SimpliVity solution includes its OmniStack software and related technologies, packaged on popular x86 platforms—either on 2U servers marketed as SimpliVity OmniCube, or with partner systems from Cisco or Lenovo, marketed as OmniStack with Cisco UCS and OmniStack with Lenovo System x, respectively.

An individual OmniStack node includes:

- A compact hardware platform - a 2U industry-standard virtualized x86 platform containing compute, memory, performance-optimized SSDs and capacity-optimized HDDs protected in hardware RAID configurations, and 10GbE network interfaces.
- A hypervisor such as VMware vSphere/ESXi.
- OmniStack virtual controller software running on the hypervisor.
- An OmniStack Accelerator Card – a special-purpose PCIe card with an FPGA, flash, and DRAM, protected with super capacitors; the accelerator card offloads CPU-intensive functions such as data compression, deduplication and optimization from the x86 processors.

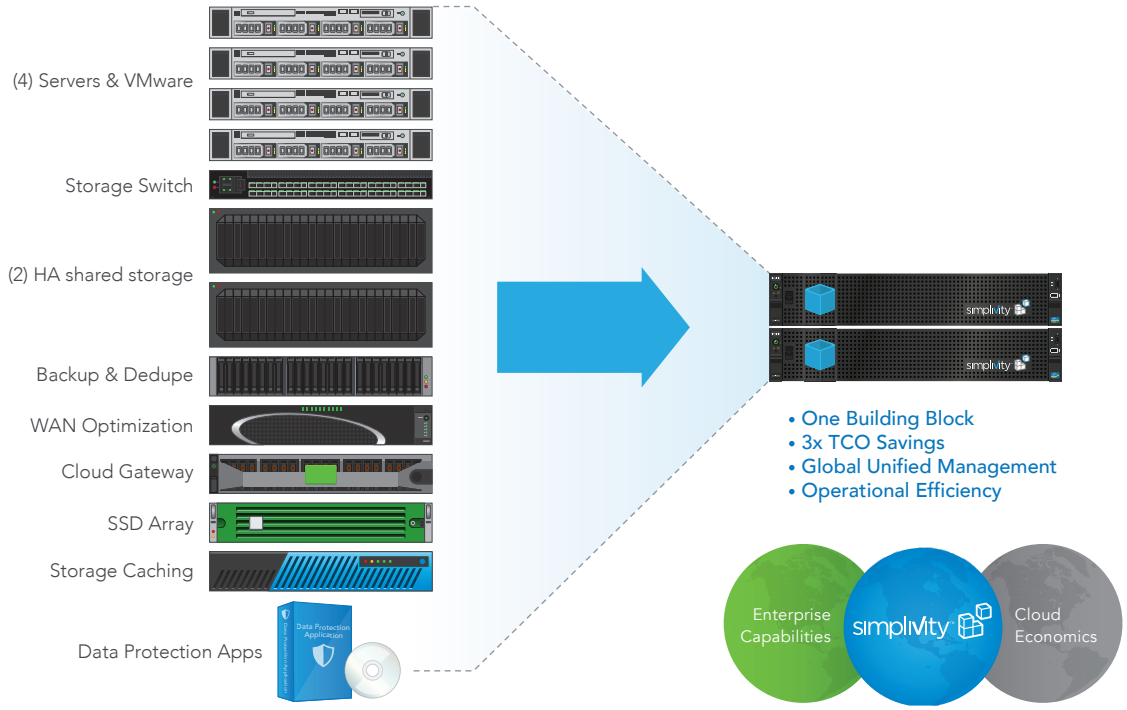


Figure 1 – Legacy Comparison

## SimpliVity Technology Overview and Key Features

OmniStack was specifically designed to meet the stringent price-performance, scalability, agility and resiliency demands of today's data-intensive, highly virtualized IT environments. Key benefits and advantages include:

- Superior economics: OmniStack eliminates infrastructure cost and complexity by consolidating a variety of IT functions (compute, storage, network switching, replication, backup, etc.) onto commodity virtualized x86 hardware, with global unified management. The solution contains CAPEX by eliminating IT silos, converging technology stacks, and optimizing storage capacity; and it reduces OPEX by containing power, cooling, rack space and system administration expenses.
- Linear scalability: The SimpliVity solution features a scale-out architecture that minimizes upfront investments and provides a high degree of flexibility and extensibility. OmniStack nodes are installed in an incremental fashion to accom-

modate growth, enable new applications or extend system availability. Two or more OmniStack nodes can be federated to create a massively scalable pool of shared resources that is administered as a cohesive system, with a single administrative interface.

- VM-centric design: OmniStack was designed from the ground up with virtualization in mind. The solution abstracts data from the underlying hardware; virtual machine files are mapped directly to blocks on storage. All data storage, management, and protection functions are inherently optimized for virtualization. And all administrative tasks—including managing data protection policies, analyzing performance and troubleshooting problems—are performed at the VM level. From an administrative perspective, a datastore is simply a logical construct, decoupled from the underlying physical infrastructure. Concepts like LUNs, volumes, shares, and disk groups simply don't apply with SimpliVity.
- Accelerated IT service agility: OmniStack's inherent data efficiencies and VM-centric management capabilities dramatically simplify operations and boost IT service agility. With OmniStack, system administrators can spin up IT services and clone VMs in just seconds with two or three mouse clicks.
- High resiliency: The SimpliVity solution is designed to be highly resilient, with no single point of failure. The solution supports both RAID (redundant array of independent disks) for disk-level resiliency and RAIN (redundant array of independent nodes) for node-level resiliency. In a high availability RAIN implementation, the complete set of data associated with a VM is simultaneously written to two distinct nodes, protecting data in the event of disk or node failures.

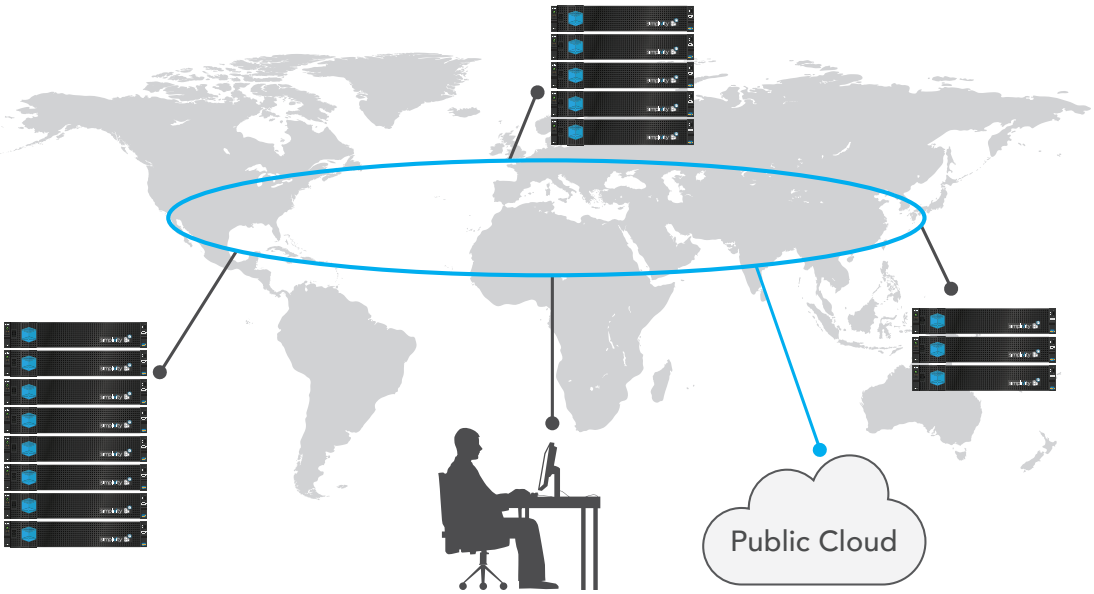


Figure 2 – An OmniStack Federation

## Solution Overview

Designed and optimized for SimpliVity hyperconverged infrastructure, the Backup Hub provides extra-large operational storage capacity that's affordable, efficient, and easy to implement and manage.

SimpliVity's Backup Hub is specifically configured and affordably priced for operational VM backup storage. Its highly efficient OmniStack Data Virtualization Platform maximizes the number of VM backups that can be safely stored in a data center rack. The Backup Hub also supports most third-party tape backup solutions, for archiving VM backups to cold storage.

SimpliVity's Backup Hub is automatically detected, recognized and registered when connected to a SimpliVity Federation. It can be managed with the same VM-centric tools that are used to manage your overall SimpliVity hyperconverged infrastructure.

### Backup and Restore technology

Data protection is fully integrated into SimpliVity's hyperconverged offering and architected into its technological foundation, eliminating the need for purpose-built backup/recovery and disaster recovery applications and appliances. Management of backup policies is performed at the VM-level, with just a few clicks, and fully managed within VMware vCenter.

All backups and restores are full logical copies of VMs with negligible impact on running applications. SimpliVity reduces downtime with its ability to restore Terabyte (TB) VMs in seconds, with backups as frequently as every 10 minutes.

For SimpliVity enterprise customers that want affordable storage for their growing operational VM backup requirements, the Backup Hub delivers extremely efficient, easily implemented and highly resilient large capacity storage to support more frequent backups for improved RPOs and longer data retention.

Some key benefits are:

- Cost effectively and efficiently support increased backup frequencies that improve your RPOs
- Quickly extend OmniStack's manageability to your backup storage
- Get enterprise-class resiliency and protection for your backup data

## Deployment Options

This section covers the deployment options for the Backup Hub within a SimpliVity environment. There are two scenarios to deploy to the Backup Hub.

Add the Backup Hub to an existing SimpliVity deployment

Add the Backup Hub to a new SimpliVity deployment.

**Note:** It is recommended to use the same Omnistack version across the SimpliVity Federation.

### Add Backup Hub to an existing SimpliVity Deployment

A Backup Hub can be added to any existing SimpliVity environment. The Backup Hub can be deployed locally in the same site as the production servers or at a remote site to provide data availability in case of an outage at the production site. A Backup Hub can also be added to an existing DR solution. Note, that the Backup Hub and the DR nodes should be two separate solutions.

Perform the following steps to ensure the most optimal usage of the Backup Hub and increase operational efficiency.

1. Move existing backups from production servers to the Backup Hub using one of the following methods.
   a. Use the copy backup feature to copy a backup from one SimpliVity datacenter to another. If desired, delete the backup after the copy completes.
   b. Create a script using SVTCLI or REST API to sequentially copy multiple backups from one datacenter to another and delete the backup from the production system.
2. Create backup policies and rules to take local backups on the production SimpliVity nodes. Set a retention policy based on business needs.
3. For long term retention, create backup policies to take remote backups on the Backup Hub. Set the retention policy based on the compliance requirements of the organization.
4. If DR is required, create a separate backup policy that backs up data to separate nodes dedicated for DR (and not the Backup Hub) based on the RPO objectives of the organization.

### Add Backup Hub to a new SimpliVity deployment

A Backup Hub can be added to a greenfield SimpliVity deployment. The Backup Hub can be added at either the production site or at a remote site (central site in a ROBO use case). Backup Hub is not meant to provide a DR solution. It can be deployed alongside a DR solution, but only as an independent solution used exclusively for backup storage.

1. Perform the following steps after deploying a new SimpliVity federation with a Backup Hub.
2. Create a local backup policy. Set the frequency and retention based on RPO objectives and business needs.
3. For longer retention of backups, create a remote backup policy to backup VMs to the Backup Hub. Set the frequency and retention of the backups based on organization compliance needs.
4. If DR is required, create a separate backup policy that backs up data to separate dedicated DR nodes (and not the Backup Hub) based on the RPO objectives of the organization.

## Use Cases for Backup Hub

This section discusses the use cases for the Backup Hub within a SimpliVity environment. There are two primary Backup Hub use cases that are illustrated here.

1. Backup and restore
2. Moving backups to a tape library.

### Use Case 1 – Backup and Restore

One of the main goals of the Backup Hub is to act as a target location for backups from either one or multiple sites. Organizations that need to have longer retention of data can utilize the Backup Hub without filling up their production environments.

Backups are not useful if you cannot restore them when required. The Backup Hub enables organizations to restore previous backups to production sites in case of data corruption or data loss. Backups can also be restored from the Backup Hub to development and test environments.

### Architecture

The diagram below depicts a ROBO environment where data is being backed up and restored from the branch office to the Backup Hub located at the central or remote site. Note that the Backup Hub does not necessarily need to be deployed in a ROBO environment, it can also be deployed in a non-ROBO environment where the production nodes and the Backup Hub live in the same site or separate sites.

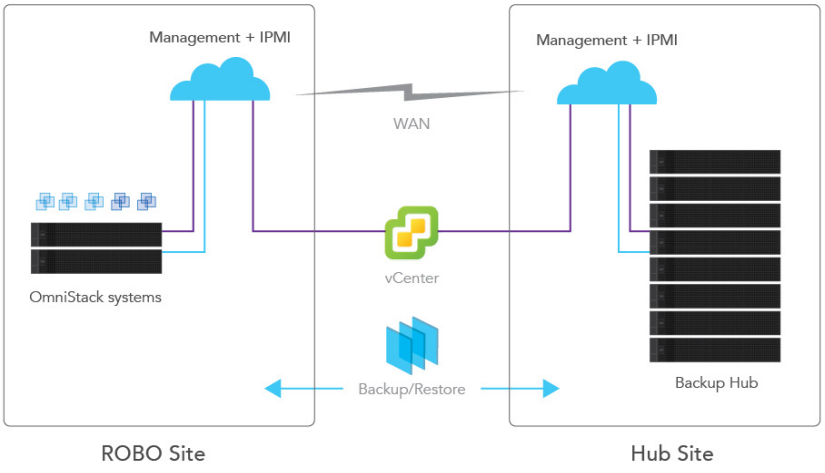


Figure 3

## How to: Backup and Restore Data

### Backing up a VM

Backups on a SimpliVity OmniStack system can be either policy-driven or manual.

### Policy-based backup

Policy driven backups allow you to schedule backup operations for virtual machines running on OmniStack systems. A backup policy contains certain rules, which include

- Backup schedule
- Destination Datacenter for the backup
- Retention period for the backup

Backup policies can be applied at a datastore level or at a VM level. A VM backup policy takes precedence over a datastore level backup policy.

### Manual Backup

A VM running application data can also be backed up on-demand manually. Perform the following steps to manually backup a VM.

1. In the vSphere client right click the VM.
2. Hover over all SimpliVity Actions
3. Click Backup Virtual Machine

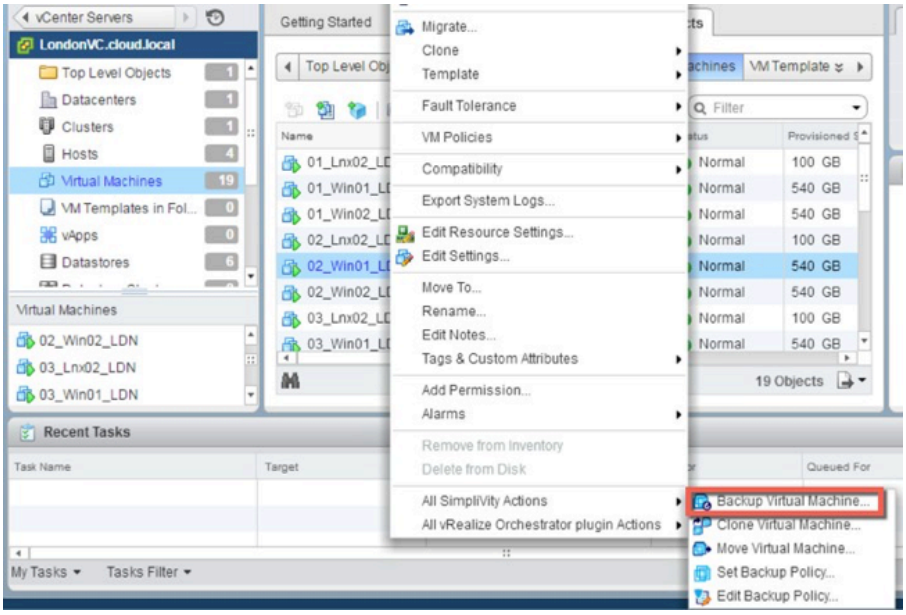


Figure 4

4. Select the datacenter to backup the VM to and optionally provide a name for the backup in the next dialog box. Advanced settings for the backup in terms of retention period can be configured as well.

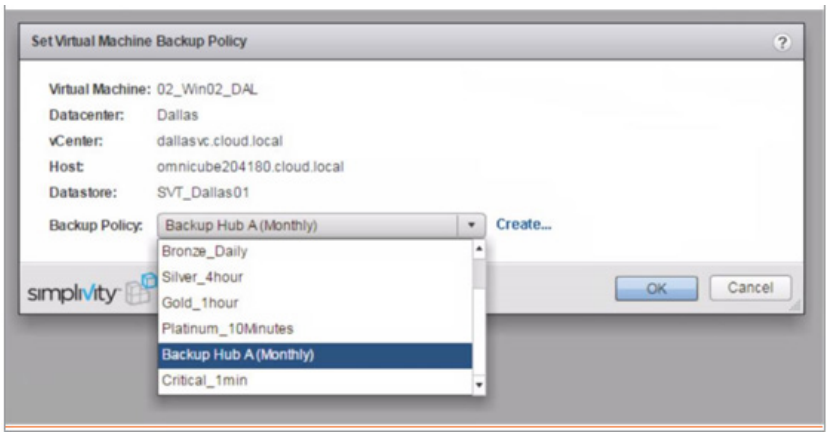


Figure 5

## Restoring a VM

Restoring a VM from a backup on a SimpliVity OmniStack system is extremely efficient and significantly faster than other backup-restore technologies available. Either the original VM can be restored from the backup or a new VM can be created from it.

To restore a VM from the Backup Hub perform the following steps:

1. Find the backup to be restored
2. Right click the backup and select restore.
3. From the restore VM dialog box.
   a. Confirm the "Create New VM" option is chosen.
   b. Optionally provide a specific name for the restored VM
   c. Select a remote datacenter to restore the VM to
   d. Click OK.

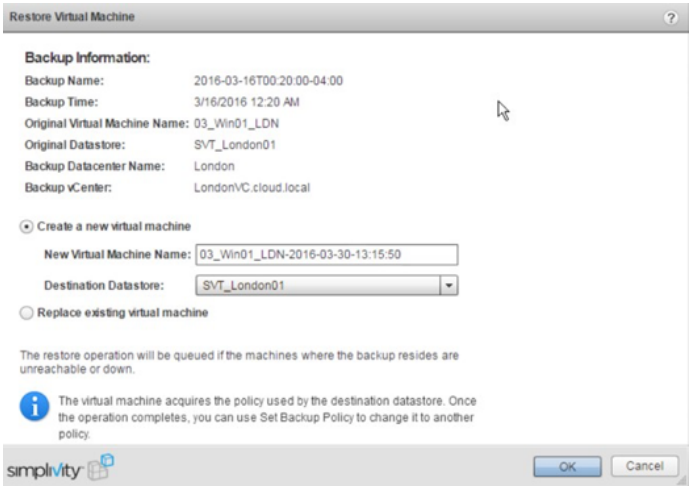


Figure 6

After the restore operation completes a new VM will appear under the datacenter selected in step 3.c

### Use Case 2 – Moving Backups to Tape Library

Most organizations have compliance regulations that mandates them to retain data on tape drives for several years. In this use case we discuss moving the backups from the SimpliVity Backup Hub to a Tape Library.

SimpliVity Backup Hub supports moving data to a tape library by restoring the backup as a powered off VM on the Backup Hub and leveraging existing tape solution to move the VM on to the tape library.

This approach enables customers to archive data directly from the Backup Hub instead of using up resources from production systems for the operation.

### Architecture

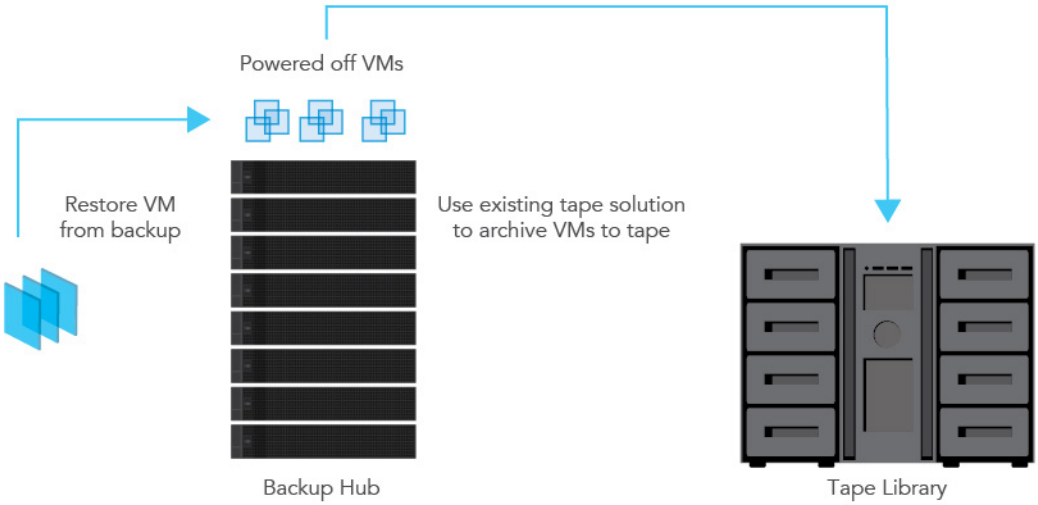


Figure 7

### How to: Archive VMs to Tape

To move VMs to tape:

1. Identify VM backup/backups to archive
2. Restore the backup from the Backup Hub using SimpliVity restore functionality on to the Backup Hub. The restored VMs are in a powered off state.
3. Leverage existing tape-out solution to move the restored VMs to the tape library.
   **Note:** The powered off VMs can be exported as .OVF through vCenter to be stored eventually on the Tape Library

## Planning and Sizing

This section provides planning and sizing information for the Backup Hub and general practices around backups within a SimpliVity environment. Several factors that affect the speed and efficiency of backups are discussed here.

### Backup Considerations

#### Backup Frequencies

Backup frequency within a backup policy dictates how often a VM is backed up. The frequency is generally determined by the Recovery Point Objective (RPO) of the organization. Higher the frequency of backups, more often data is sent over the WAN in a remote backup scenario. 10-minute backup frequency is the lowest time interval you can set with the current version of the OmniStack software (3.5.1).

### Backup Retention

The retention period of a backup policy decides when the backup and associated data is deleted from the system. For example, if a backup policy is created to take backups every day and retain it for 7 days, then on the 8th day the oldest backup will be deleted and a new backup will be taken. The longer the retention period, the more likely it is to fill up the system with backup data.

### Change Rate

Change rate is defined as the rate at which new data is written or created against an existing data object within a given time period. Change rate is one of the most significant factors that determine the size of a backup. For example, if the change rate for a 100GB VM is 10% per day and backups are taken daily then each backup will be around 10GB. Along with this if the retention period for a backup policy is 7 days, then there will be 70GB of backup data associated with the policy on the SimpliVity platform. Note that data efficiency savings on SimpliVity platforms through deduplication and compression may require less data to be stored.

### VM Size

VM size is also a critical factor that decides the size of the backup and the amount of data that is transferred over the WAN in a remote backup scenario. As mentioned above, the size of the VM or data object along with the change rate typically dictates the size of the backup. A very large VM with a small change rate can generate the same amount of backup data as a small VM with a high change rate. For example, a 100GB VM with a 10% daily change rate will produce the same size daily backup as a 1000GB VM with a 1% change rate.

### Backup Scheduling

If there are multiple backup policies sending data over the WAN at the same time, there is a chance that the network might get congested and the RPO of the organization will be missed. Therefore, backups must be scheduled in such a way that not all data is being transferred over the network at one time but instead is distributed over time.

### WAN and Storage Latency

Latency and existing load on the production and Backup Hub system can affect the RPO of a backup environment. If the latency and load is constantly high, then backup policies may take more time to complete due to contention of resources.

### Examples

A few backup calculation examples are provided below as reference. For simplicity, the size of the VM is kept constant when calculating change rate in the examples below. In the examples below we are assuming no deduplication and compression. The user may see additional efficiency in terms of their deduplication and compression ratios.

### Example 1:

The following example shows how to calculate the size of a remote backup.

**Details**

- VM Size - 200GB
- Change rate – 5%
- Backup frequency – daily
- Retention Period – 1 week

$$\text{Remote Backup Size} = \frac{(\text{VM Size} \times \text{Change Rate})}{(\text{Backup Frequency})}$$

$$= \frac{(200\text{GB} \times 0.05)}{1}$$

$$= 10\text{GB per day}$$

Example 2:

Remote backup size using the same numbers as Example 1 but with a different frequency. In this example we have changed the frequency from daily to hourly. In this case we need convert the daily change rate to an hourly rate. It should be noted that organizations typically have 8 hours as a work day when user activity is the busiest. If there is not much fluctuation in user activity during a full day then the work day should be used as a 24 hour period. For this example, we are assuming 8 hours as the typical work day.

Details

- VM Size – 200GB
- Change rate – 5%
- Backup frequency – hourly
- Retention Period – 1 week

$$\text{Remote Backup Size} = \frac{(\text{VM Size} \times \text{Change Rate})}{(\text{Backup Frequency})}$$

$$= \frac{(200\text{GB} \times 0.05)}{8}$$

$$= 1.25\text{GB per Hour}$$

Example 3:

In this example we will work out WAN bandwidth considerations. It is assumed that the data being backed up is pre-calculated per site along with the average change rate.

Details

| Site | Data Size | Daily Change Rate % | Backup Frequency | Size of Backup |
|---|---|---|---|---|
| 1 | 200GB | 1% | Daily | 5GB/day |
| 2 | 200GB | 2% | Daily | 4GB/day |
| 3 | 500GB | 1% | Daily | 5GB/day |
| 4 | 100GB | 4% | Hourly | 0.5GB/hour |
| 5 | 200GB | 5% | Hourly | 1.25GB/hour |

Based on the data above, assuming the backups are not staggered and all data is unique and not de-duplicated at the Hub site, the maximum data that would be transferred over the WAN at a given hour would be:

10GB+4GB+10GB+0.5GB+1.25GB=15.75GB

The backups should complete in an hour since that is the smallest RPO amongst the 5 sites. Therefore, minimum WAN bandwidth required is:

$$= \frac{(15.75\ \text{GB})}{(3600\ \text{seconds})}$$

$$= \frac{(15.75 \times 1024\ \text{MB})}{(3600\ \text{seconds})}$$

$$= 4.48\ \text{MB/second}$$

$$= 35.84\ \text{Mb/second}$$

Based on the calculations above a minimum of 35.84 Mbps is required..

More information on WAN speeds can be found here:

http://www.comptechdoc.org/independent/networking/cert/netwan.html

**Note:** As mentioned above, example assumes a worst case with no deduplication at the remote site and that all the data generated is new and unique. SimpliVity's global inline deduplication and compression reduces WAN traffic and speeds up the backup process significantly.

Example 4:

In this example we will use the same details as example 3. However, in this example we will consider staggering the backup policies to reduce the WAN bandwidth requirement by avoiding all 5 sites backup at the same time.

There are 3 daily backups and 2 hourly backups.

The daily backups can be scheduled after an 8-hour work day (9am to 5pm) to run at 6PM, 1AM and 8AM respectively. 7 hours between the backups should provide enough time for the daily backups to meet their RPO's.

5GB daily change is the maximum data that needs to be backed up.

At any given hour during the workday (9am to 5pm) the maximum backup data for the hourly backup would be:

Site 4 + Site5
= 1.25+0.5

= 1.75GB

To transfer 1.75GB and achieve an RPO of 1 hour, the WAN bandwidth required is:

$$= \frac{(1.75 \times 1024\ \text{MB})}{(3600\ \text{seconds})}$$

=0.49 MB/second

=3.98 Mb/second

Therefore, scheduling backup policies can significantly reduce WAN bandwidth requirements.

Example 5:

### Sizing Backup Hub

There are several factors that need to be considered for determining the size of the Backup Hub needed in an environment. Since the first backup is always a full copy, minimum space required will be equal to the sum of the size of all VMs being backed up. Subsequent backups are incremental and will depend on the rate of change of the VM and the frequency at which backups are taken. Lastly, the retention period of the backups will also increase the space requirements as the backups won't be deleted until the retention policy for the backup.

Therefore, in general space required on the backup hub would be dependent on number of VMs being backed up, size of VMs, frequency of each backup, change rate of each VM and retention period of each backup.

- VM size – 100GB
- Change rate per day – 10%
- Frequency of backup – daily
- Retention – 1 week

Minimum size required at the backup hub for this backup = 100 + [100 (0.1) x 7] = 170GB.

Therefore a minimum of 170GB storage space is required at the backup hub for the VM described above. If there are 10 VMs with the same profile being backed up to the Backup Hub, then approximately you would need 1700GB of space at the backup hub (not considering any efficiency savings).

**Note:** Deduplication and compression will provide additional space savings and will vary from one environment to another.

## Best Practices and Recommendations

This section provides general guidance and recommendations for Backup Hub

### Manage VM size

Backup VM sizes should be managed to ensure that individual hub nodes are not saturated, causing other backups to begin on a different Hub node. It is recommended to have VMs no more than 5TB in size.

### Pre-seed Backup Hub

Pre-seed Backup Hub nodes with common data so that less data needs to be transferred over the WAN and to speed up the backup process.

### Schedule backups

Stagger backup policies so that backups are initiated at different times throughout the day and RPO objectives are easily met. If the customer triggers multiple concurrent backups from many remote sites (e.g., Backup every site daily at mid-night), backups could queue at the Hub and fail, resulting in missed customer RPOs.

### Monitor System Performance

Monitor performance on the Backup Hub and verify RPO's are being met after adding a new backup policy that backups to the hub. Verify that the system has periods with low activity so that new backup policies can be scheduled during that time. Periodically, monitor the free space available within the datacenter and ensure that space is available for existing backup policies and for any growth.

### Estimate Backup Size and WAN bandwidth

Estimate backup load using rate of change of data, frequency of backups and VM size and ensure sufficient network bandwidth is available to avoid network congestion at the hub site. **Note:** Refer to the examples provided above.

### Add Backup Policies Incrementally

Protect data locally by creating short term retention local backup policies and protect data at the hub by creating long term retention backup policies. Most of the time, errors and corrections are identified shortly after they occur. Restoring/ Recovering data from a local backup will be faster instead of transferring data across the WAN.

### Manage retention of backups

If the objective of the organization is to maintain a copy of the VM data and doesn't require the delta between backups, then policies should be created to delete small RPO backups when a longer RPO backup is taken. For example, if you are backing up a VM every day for a week, then you will have 7 backups in the week, if the delta of the changes is not required for the 7 days, then a single weekly backup should be taken via a different policy to replace the other 7 daily backups.

### Deploy DR as an independent solution

The Backup Hub is not meant for DR purposes. If DR is required, then implement a separate DR solution so that any outages of the Backup Hub does not compromise DR strategy.

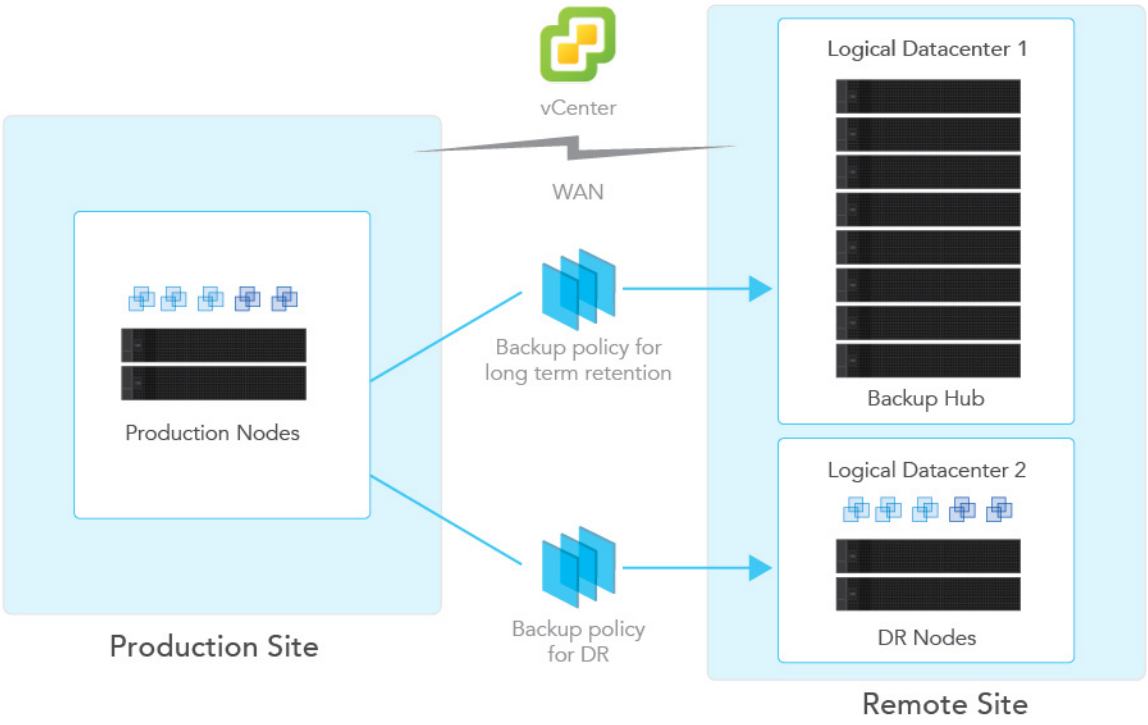


Figure 8

The above diagram shows independent solutions for long term retention of backups and for DR. In this case there would be two backup policies – one policy backing up to the Backup Hub and retaining backups based on compliance requirements and another policy backing up VMs to the DR node based on RPO of the organization.

## Summary/Conclusion

SimpliVity's Backup Hub provides a solution to multiple challenges faced by a customer today. The Backup Hub provides an optimum solution for longer retention of backups, it can serve as a single highly available backup target in a ROBO environment and since the Backup Hub is running the OmniStack Software, backups and restores from and to production sites are still as fast and provide significant space savings. The Backup Hub also supports archiving VMs to tape to enable organizations achieve compliance as demanded by their specific business industry.

Affordability, data efficiency, ease of integration and management along with resiliency and high availability make the Backup Hub the perfect complement to a complete well-rounded data protection solution.

## For more information, visit:

www.simplivity.com



J1001_Backup Hub-WP-EN-0816