v5.3 | March 2021 | BP-2005

BEST PRACTICES

Data Protection and Disaster Recovery

NUTANIX™
YOUR ENTERPRISE CLOUD

Copyright

Nutanix, Inc.
1740 Technology Drive, Suite 150
San Jose, CA 95110

Contents

1. Executive Summary

The Nutanix enterprise cloud software is a hyperconverged infrastructure system delivering storage, compute, and virtualization services for any application. Designed for supporting multiple virtualized environments, including Nutanix AHV, VMware ESXi, and Microsoft Hyper-V, Nutanix invisible infrastructure is exceptionally robust and provides many ways to achieve your required recovery point objectives (RPOs).

Enterprises are increasingly vulnerable to data loss and downtime during disasters as they rely on virtualized applications and infrastructure that their legacy data protection and disaster recovery solutions can no longer adequately support. This best practices guide discusses the optimal configuration for achieving data protection using the native Nutanix disaster recovery capabilities and the Leap disaster recovery orchestration features available both on-premises, in Xi, and in public cloud providers like Amazon Web Services (AWS). Whatever your use case, you can protect your applications with drag-and-drop functionality. The Nutanix Prism UI facilitates seamless management to configure the shortest recovery time objectives (RTOs) possible, so customers can build complex disaster recovery workflows at a moment’s notice. With Leap built in, Prism Central allows you to apply protection policies across all your managed clusters. Once the business has decided on the required RPO, you can activate recovery plans to validate, test, migrate, and fail over in a seamless fashion. Recovery plans can protect availability zones both on-premises and hosted in Xi.

As application requirements change and grow, Nutanix can easily adapt to business needs. Nutanix is uniquely positioned to protect and operate in environments with minimal administrative effort because of its web-scale architecture and commitment to enterprise cloud operations.

2. Introduction

Audience

This best practices guide is part of the Nutanix Solutions Library. We wrote it for IT administrators and architects who want more information about the data protection and disaster recovery features built into the Nutanix enterprise cloud software. Readers should have basic familiarity with Nutanix and AOS.

Purpose

This document provides best practice guidance for data protection solutions implementation on Nutanix servers running AOS 5.19. We present the following concepts:

- Scalable metadata.
- Backup.
- Crash-consistent versus application-consistent snapshots.
- Protection domains.
- Protection policies.
- Recovery plans.
- Scheduling snapshots and asynchronous replication.
- Sizing disk space for local snapshots and replication.
- Scheduling lightweight snapshots (LWS) and NearSync replication.
- Sizing disk space for LWS and NearSync replication.
- Determining bandwidth requirements.
- File-level restore.

Table 1: Document Version History

Version Number	Published	Notes
1.0	December 2014	Original publication.
2.0	March 2016	Updated recommendations for current best practices throughout.
2.1	June 2016	Updated Backup and Disaster Recovery on Remote Sites section.
2.2	July 2016	Updated bandwidth sizing information.
2.3	December 2016	Updated for AOS 5.0.
2.4	May 2017	Updated information on sizing SSD space on a remote cluster.
3.0	December 2017	Updated for AOS 5.5.
3.1	September 2018	Updated overview and Remote Site Setup section.
4.0	December 2018	Updated for AOS 5.10 and Xi Leap.
4.1	February 2019	Updated Sizing Space section and Leap product details.
4.2	August 2019	Updated for AOS 5.11.
4.3	October 2019	Updated for AOS 5.11.1 and Nutanix Files support for NearSync replication.
5.0	May 2020	Updated for AOS 5.17.
5.1	September 2020	Updated for AOS 5.18.
5.2	December 2020	Updated for AOS 5.19 and Leap multisite replication.
5.3	March 2021	Refreshed content.

3. Nutanix Enterprise Cloud Overview

Nutanix delivers a web-scale, hyperconverged infrastructure solution purpose-built for virtualization and both containerized and private cloud environments. This solution brings the scale, resilience, and economic benefits of web-scale architecture to the enterprise through the Nutanix enterprise cloud platform, which combines the core HCI product families—Nutanix AOS and Nutanix Prism management—along with other software products that automate, secure, and back up cost-optimized infrastructure.

Available attributes of the Nutanix enterprise cloud OS stack include:

- Optimized for storage and compute resources.
- Machine learning to plan for and adapt to changing conditions automatically.
- Intrinsic security features and functions for data protection and cyberthreat defense.
- Self-healing to tolerate and adjust to component failures.
- API-based automation and rich analytics.
- Simplified one-click upgrades and software life cycle management.
- Native file services for user and application data.
- Native backup and disaster recovery solutions.
- Powerful and feature-rich virtualization.
- Flexible virtual networking for visualization, automation, and security.
- Cloud automation and life cycle management.

Nutanix provides services and can be broken down into three main components: an HCI-based distributed storage fabric, management and operational intelligence from Prism, and AHV virtualization. Nutanix Prism furnishes one-click infrastructure management for virtual environments running on AOS. AOS is hypervisor agnostic, supporting two third-party hypervisors

—VMware ESXi and Microsoft Hyper-V—in addition to the native Nutanix hypervisor, AHV.

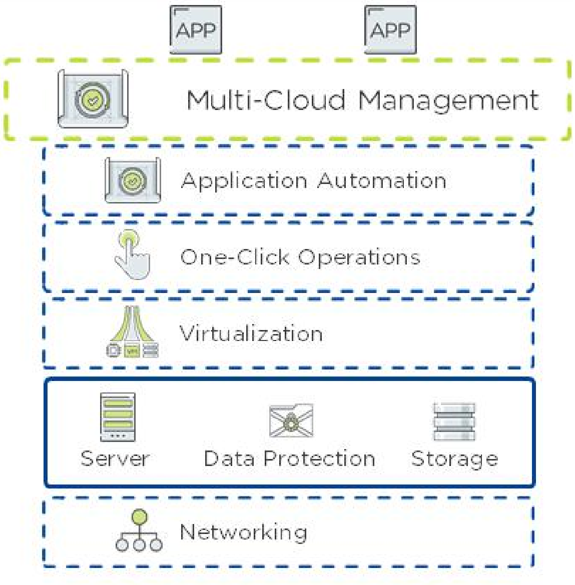


Figure 1: Nutanix Enterprise Cloud OS Stack

Nutanix HCI Architecture

Nutanix does not rely on traditional SAN or network-attached storage (NAS) or expensive storage network interconnects. It combines highly dense storage and server compute (CPU and RAM) into a single platform building block. Each building block delivers a unified, scale-out, shared-nothing architecture with no single points of failure.

The Nutanix solution requires no SAN constructs, such as LUNs, RAID groups, or expensive storage switches. All storage management is VM-centric, and I/O is optimized at the VM virtual disk level. The software solution runs on nodes from a variety of manufacturers that are either entirely solid-state storage with NVMe for optimal performance or a hybrid combination of SSD and HDD storage that provides a combination of performance and additional capacity. The storage fabric automatically tiers data across the cluster to different classes of storage devices using intelligent data placement algorithms. For best performance, algorithms make sure the most frequently used data is available in memory or in flash on the node local to the VM.

To learn more about Nutanix enterprise cloud software, visit the Nutanix Bible and Nutanix.com.

4. Web-Scale Data Protection

One of the key architectural differentiators for Nutanix is the ability to scale. Nutanix isn't bound by the same limitations as dual-controller architectures or federations relying on special hardware like NVRAM or custom ASICs for performance. When it comes to snapshots and disaster recovery, scaling metadata is a key part of delivering performance while ensuring availability and reliability. Each Nutanix node is responsible for a subset of the overall platform's metadata. All nodes in the cluster serve and manipulate metadata entirely through software, eliminating traditional bottlenecks.

Because each node has its own virtual storage controller and access to local metadata, replication scales with the system. Every node participates in replication to reduce hotspots throughout the cluster.

Nutanix uses two different forms of snapshots: full snapshots for asynchronous replication (when the RPO is 60 minutes or greater) and lightweight snapshots (LWS) for NearSync replication (when the RPO is between 15 minutes and 1 minute). Full snapshots keep system resource usage low when you use many snapshots over an extended period. The LWS feature reduces metadata management overhead and increases storage performance by decreasing the high number of storage I/O operations that long snapshot chains can cause.

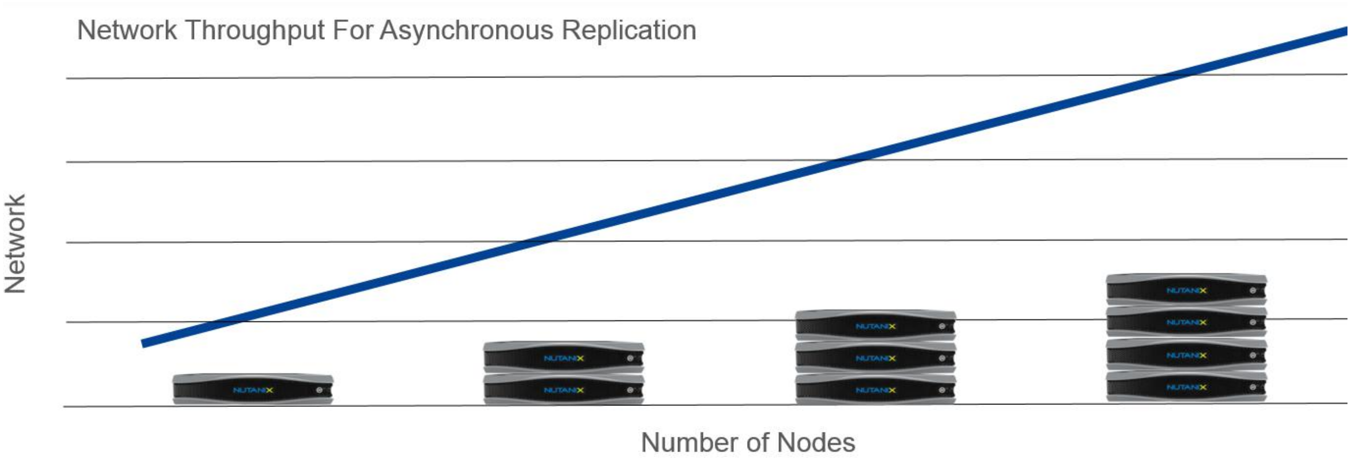


Figure 2: Scalable Replication

In asynchronous replication, every node can replicate four files, up to an aggregate of 100 MBps at one time. Thus, in a four-node configuration, the cluster can replicate 400 MBps or 3.2 Gbps. As you grow the cluster, the virtual storage controllers keep replication traffic distributed. In many-to-one deployments, as when remote branch offices communicate with a main datacenter, the main datacenter can use all its available resources to handle the increased replication load from the branch offices. When the main site is scalable and reliable, administrators don't have multiple replication targets to maintain, monitor, and manage. You can protect both VMs and volume groups with asynchronous replication.

NearSync offers unbound throughput. Because all writes go to SSD, we want to make sure that the performance tier doesn't fill up. To this end, AOS automatically allocates 7 percent of the performance tier to be used by NearSync. NearSync, which covers both VMs and volume groups, is supported for bidirectional replication between two clusters.

Nutanix also provides cross-hypervisor disaster recovery natively with asynchronous replication. Existing vSphere clusters can target AHV-based clusters as their disaster recovery and backup targets. Thanks to true VM mobility, Nutanix customers can place their workloads on the platform that best meets their needs.

5. Deployment Overview

Nutanix meets real-world requirements with native backup and replication infrastructure and management features that support a wide variety of enterprise topologies.

Native Nutanix Snapshots

Per-VM or per–volume group snapshots enable instant recovery. Depending on the workload and associated SLAs, customers can tune the snapshot schedule and retention periods to meet the appropriate RPOs. With the intuitive UI snapshot browser, you can perform restore and cloning operations instantly on the local cluster.

Two-Way Mirroring

The ability to mirror VM and volume group replication between multiple sites is necessary in environments where all sites must support active traffic. Consider a two-site example. Site B is the data protection target for selected workloads running on site A. At the same time, site A serves as the target for designated workloads running on site B. While asynchronous replication is supported for all workflows listed in this section, two-way mirroring is the only supported topology for NearSync.

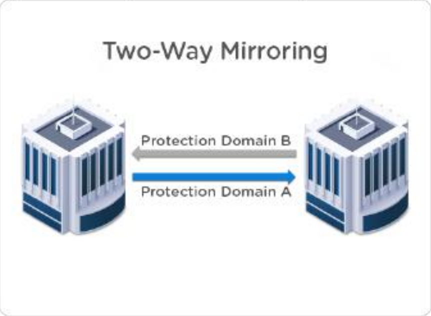


Figure 3: Two-Way Mirroring

Many-to-One

In a many-to-one or hub-and-spoke architecture, you can replicate workloads running on sites A and B to a central site C. Centralizing replication to a single site may improve operational efficiency for geographically dispersed environments. Remote and branch offices (ROBO) are a classic many-to-one topology use case.

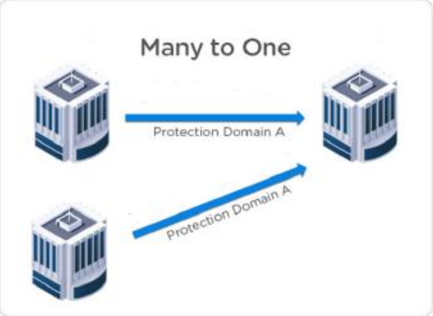


Figure 4: Many-to-One Architecture

To the Cloud

With Cloud Connect, customers can now use the public cloud as a destination for backing up their on-cluster VMs and volume groups. At this time, Nutanix supports Amazon Web Services (AWS) as the cloud destination. This option is particularly suitable for customers who don’t have an offsite location for their backups or who are currently relying on tapes for storing their backups offsite. Cloud Connect provides customers with backup options for both Hyper-V and ESXi using the same Prism UI. Customers can also deploy a Nutanix cluster in a public cloud provider like AWS, setting up protection domains and employing the same workflows they use with on-prem clusters.

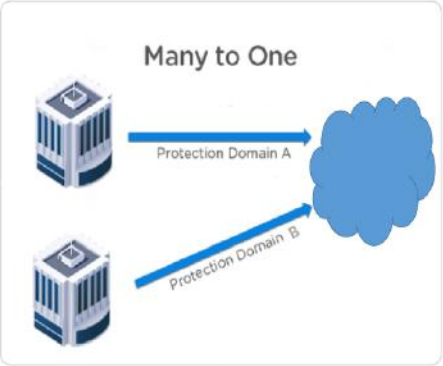


Figure 5: Public Cloud as a Backup Destination

Single-Node Backup

Nutanix has added support for single-node backup as a cost-efficient solution for providing full native backups to branch offices and SMBs. Using the same underlying disaster recovery technology, the single node can be either on-site or remote. Nutanix protects data on the node from single drive failure and provides native backup end to end. Single-node systems can also run VMs for testing or remote branch offices.

Leap: Disaster Recovery Orchestration

Prism Central provides a single web console for monitoring and managing multiple clusters. AOS versions 5.11 and later provide protection policies and recovery plans in Prism Central, offering an easy way to orchestrate operations around migrations and unplanned failures. Now you can apply orchestration policies for ESXi and AHV from a central location, ensuring consistency across all your sites and clusters.

To help manage these protection policies and recovery plans, Nutanix uses a construct called availability zones. On-premises, an availability zone includes all the Nutanix clusters managed by one Prism Central. An availability zone can also represent a region in Nutanix Xi Cloud Services. For disaster recovery, availability zones exist in pairs—either on-prem to on-prem or on-prem to Xi. Once you have paired your on-prem environment to a Xi-based availability zone, you can take advantage of Xi Leap, which is Nutanix disaster recovery as a service. There are multiple Xi Leap subscription plans available so you don't have to pay the full cost associated with buying a secondary cluster up front and you save the time it takes to manage and operate the infrastructure.

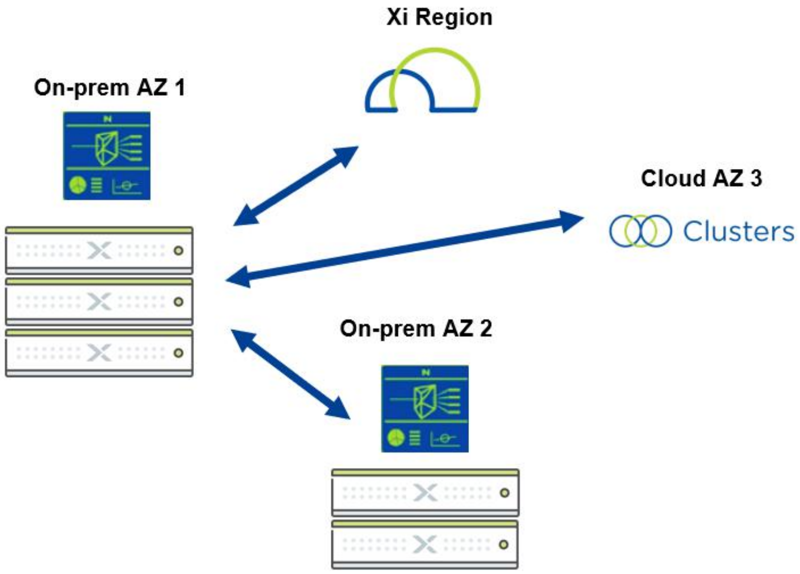


Figure 6: On-Prem, Cloud, and Xi-Based Availability Zones

6. Local Backup with Snapshots

Native Snapshots

Nutanix native snapshots provide production-level data protection without sacrificing performance. Nutanix uses a redirect-on-write algorithm to dramatically improve system efficiency for snapshots. Native snapshots operate at the VM level, and our crash-consistent snapshot implementation is the same across hypervisors. Implementation varies for application-consistent snapshots because of differences in the hypervisor layer. Nutanix can create local backups and recover data instantly to meet a wide range of data protection requirements.

Best practices:

- All VM files should sit on Nutanix storage. If you make non-Nutanix storage available to store files (VMDKs), the storage should have the same file path on both the source and destination clusters.
- Remove all external devices, including ISOs and floppy devices.

Crash-Consistent vs. Application-Consistent Snapshots

VM snapshots are by default crash-consistent, which means that the vDisks captured are consistent with a single point in time. The snapshot represents the on-disk data as if the VM crashed or the power cord was pulled from the server —it doesn’t include anything that was in memory when the snapshot was taken. Today, most applications can recover well using crash-consistent snapshots.

Application-consistent snapshots capture the same data as crash-consistent snapshots, with the addition of all data in memory and all transactions in process. Because of their extra content, application-consistent snapshots are the most involved and take the longest to perform.

While most organizations find crash-consistent snapshots to be sufficient, Nutanix also supports application-consistent snapshots. The Nutanix application-consistent snapshot uses the Nutanix Volume Shadow Copy Service (VSS) to quiesce the file system for ESXi and AHV prior to taking the snapshot. You can configure which type of snapshot each protection domain should maintain.

VSS Support with Nutanix Guest Tools

The Nutanix Guest Tools (NGT) software package for VMs plays an important role in application-consistent snapshots. ESXi and AHV-based snapshots call the Nutanix VSS provider from the Nutanix Guest Agent, which is one component of NGT. VMware Tools talk to the guest VM's VSS writers. Application-consistent snapshots quiesce all I/O, complete all open transactions, and flush the caches so everything is at the same point. VSS freezes write I/O while the native Nutanix snapshot takes place, so all data and metadata is written in a consistent manner. Once the Nutanix snapshot takes place, VSS thaws the system and allows queued writes to occur. Application-consistent snapshots don't snapshot the OS memory during this process.

Requirements for Nutanix VSS snapshots:

- Configure an external cluster IP address.
- Guest VMs should be able to reach the external cluster IP on port 2074.
- Guest VMs should have an empty IDE CD-ROM for attaching NGT.
- Only available for ESXi and AHV.
- Virtual disks must use the SCSI bus type.
- For NearSync support, you must have NGT version 1.3 installed.
- VSS isn't supported with NearSync or if the VM has any delta disks (hypervisor snapshots).

- Nutanix VSS snapshots are only available for these supported versions:
  - Windows 7
  - Windows Server 2008 R2 and later
  - CentOS 6.5 and 7.0
  - Red Had Enterprise Linux (RHEL) 6.5 and 7.0
  - Oracle Linux 6.5 and 7.0
  - SUSE Linux Enterprise Server (SLES) 11 SP4 and 12
- VSS must be running on the guest VM. Check the PowerShell Scripts section of the appendix for a script that verifies whether the service is running.
- The guest VM must support the use of VSS writers. Check the PowerShell Scripts section of the appendix for a script that ensures VSS writer stability (Windows only).
- VSS isn't supported for volume groups.
- You can't include volume groups in a protection domain configured for Metro Availability.
- You can't include volume groups in a protected VStore.
- You can't use Nutanix native snapshots to protect VMs that have VMware fault tolerance enabled.

For ESXi, if you haven't installed NGT, the process fails back using VMware Tools. Because the VMware Tools method creates and deletes an ESXi-based snapshot whenever it creates a native Nutanix snapshot, it generates more I/O stress. To eliminate this stress, we strongly recommend installing NGT.

Best practices:

- Schedule application-consistent snapshots during off-peak hours. NGT takes less time to quiesce applications than VMware Tools, but application-consistent snapshots still take longer than crash-consistent snapshots.
- Increase cluster heartbeat settings when using Windows Server Failover Cluster (WSFC).

- To avoid accidental cluster failover when you perform a vMotion, follow VMware best practices to increase heartbeat probes:
    - Change the tolerance of missed heartbeats from the default of 5 to 10.
    - Increase the number to 20 if your servers are on different subnets.
    - If you run Windows Server 2012, adjust the RouteHistoryLength to double the CrossSubnetThreshold value.

Table 2: MS Failover Settings Adjusted for Using VSS

(get-cluster).SameSubnetThreshold = 10
(get-cluster).CrossSubnetThreshold = 20
(get-cluster).RouteHistoryLength = 40

VSS on Hyper-V

Hyper-V on Nutanix supports VSS only through third-party backup applications, not snapshots. Because the Microsoft VSS framework requires a full share backup for every virtual disk in the share, Nutanix recommends limiting the number of VMs on any container using VSS backup.

Best practices:

- Create different containers for VMs that need VSS backup support. Don’t exceed 50 VMs on each container.
- Create a separate large container for crash-consistent VMs.

7. Protection Domains

A protection domain is a group of VMs or volume groups that you can either snapshot locally or replicate to one or more clusters when you have a remote site configured. Prism Element uses protection domains when replicating between remote sites.

Best practices:

- Protection domain names must be unique across sites.
- No more than 200 VMs per protection domain.
  - VMware Site Recovery Manager and Metro Availability protection domains are limited to 3,200 files.
  - No more than 10 VMs per protection domain with LWS.
- Group VMs with similar RPO requirements.
  - NearSync can only have one schedule, so place NearSync VMs in their own protection domain.

Consistency Groups

Administrators can create a consistency group for VMs and volume groups that are part of a protection domain and need to be snapshotted in a crash-consistent manner.

Best practices:

- Keep consistency groups as small as possible. Collect dependent applications or service VMs into a consistency group to ensure that they're recovered in a consistent state (for example, put a web server and database in the same consistency group).
- For all hypervisors, try to limit consistency groups to fewer than 10 VMs following the above best practices. Although we have tested consistency

groups with up to 50 VMs, it’s more efficient to have smaller consistency groups.

- Each consistency group using application-consistent snapshots can contain only one VM.
- When you provide disaster recovery for VDI using VMware View Composer or Machine Creation Services (MCS), place each protected VM in its own consistency group (including the gold image) inside a single protection domain.

8. Backup and Disaster Recovery on Remote Sites

Nutanix allows administrators to set up remote sites and select whether they use those remote sites for simple backup or for both backup and disaster recovery.

Remote sites are a logical construct. You must configure any AOS cluster—either physical or based in the cloud—as a remote site from the perspective of the source cluster before you use it as the destination for storing snapshots. Similarly, on this secondary cluster, you must configure the primary cluster as a remote site before snapshots from the secondary cluster start replicating to it.

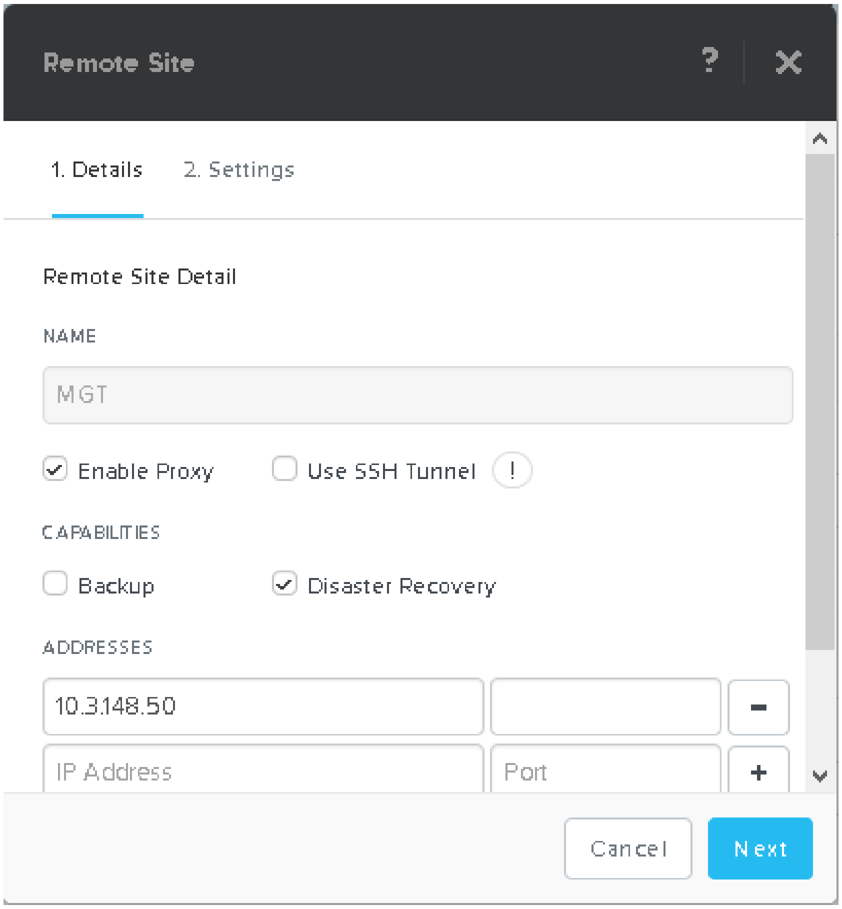
Configuring the backup option on Nutanix allows an organization to use its remote site as a replication target. This option means you can back up data to this site and retrieve snapshots from it to restore locally, but failover protection (that is, running failover VMs directly from the remote site) isn't enabled. Backup supports using multiple hypervisors; as an example, an enterprise might have ESXi in the main datacenter but use Hyper-V at a remote location. With the backup option configured, the Hyper-V cluster could use storage on the ESXi cluster for backup. Using this method, Nutanix can also back up to AWS from Hyper-V or ESXi.

Configuring the disaster recovery option allows you to use the remote site both as a backup target and as a source for dynamic recovery. In this arrangement, failover VMs can run directly from the remote site. Nutanix provides cross-hypervisor disaster recovery between ESXi and AHV clusters. Currently, Hyper-V clusters can only provide disaster recovery to other Hyper-V-based clusters.

For data replication to succeed, configure forward (DNS A) and reverse (DNS PTR) DNS entries for each ESXi management host on the DNS servers used by the Nutanix cluster.

Remote Site Setup

You can customize several options when you set up a remote site. Protection domains inherit all remote site properties during replication.



Remote Site Detail

NAME

MGT

☑ Enable Proxy □ Use SSH Tunnel

CAPABILITIES

□ Backup ☑ Disaster Recovery

ADDRESSES

10.3.148.50

IP Address Port

Figure 7: Setup Options for a Remote Site

Address

Use the external cluster IP as the address for the remote site. The external cluster IP is highly available, as it creates a virtual IP address for all the virtual storage controllers. You can configure the external cluster IP in the Prism UI under cluster details.

Other recommendations include:

- Try to keep both sites at the same AOS version. If both sites require compression, both must have the compression feature licensed and enabled.
- Open the following ports between both sides: 2009 TCP, 2020 TCP, 9440 TCP, and 53 UDP. If you use the SSH tunnel, also open 22. Use the external cluster IP address for source and destination. Cloud Connect uses a port between 3000–3099 but that setup occurs automatically. You must allow all CVM IPs to pass replication traffic between sites with the ports detailed above. To simplify firewall rules, you can use the proxy described in the following section.

Enable Proxy

The enable proxy option redirects all egress remote replication traffic through one node. This remote site proxy is different from the Prism proxy. When you enable the proxy, replication traffic goes to the remote site proxy, which then forwards it to other nodes in the cluster. This arrangement significantly reduces the number of firewall rules you need to set up and maintain.

Best practice:

- Use the proxy in conjunction with the external address.

SSH Tunnel

An SSH tunnel is a point-to-point connection—one node in the primary cluster connects to a node in the remote cluster. By enabling proxy, we force replication traffic to go over this node pair. You can use the SSH tunnel between Cloud Connect and physical Nutanix clusters when you can't set up a virtual private network (VPN) between the two clusters. We recommend using an SSH tunnel as a fail-back option in lieu of a VPN.

Best practices:

- To use an SSH tunnel, enable the proxy.
- Open port 22 between external cluster IPs.
- Only use SSH tunnel for testing—not production. Use a VPN between remote sites or a Virtual Private Cloud (VPC) with AWS.

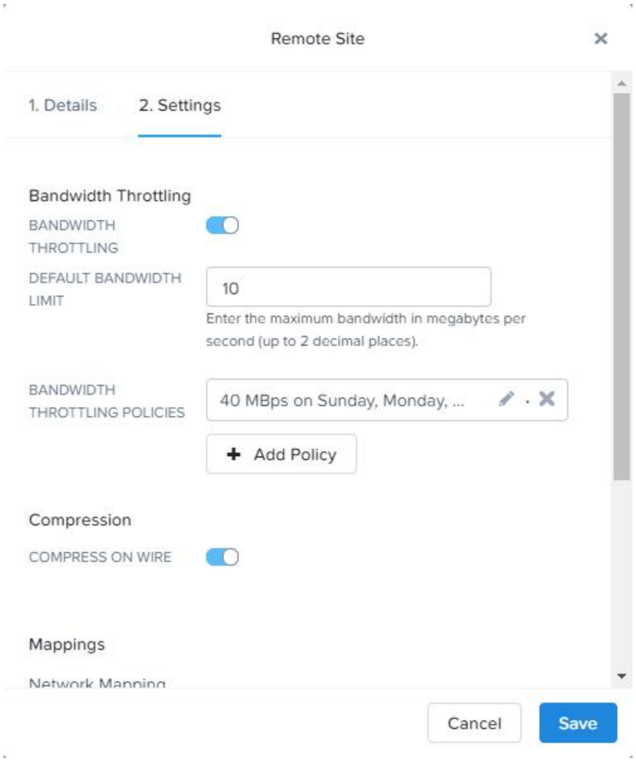
Capabilities

The disaster recovery option requires that both sites either support cross-hypervisor disaster recovery or have the same hypervisor. Today, Nutanix supports only ESXi and AHV for cross-hypervisor disaster recovery with full snapshots. When using the backup option, the sites can use different hypervisors, but you can't restore VMs on the remote side. You also use the backup option when you back up to AWS and Azure.

Bandwidth Throttling

Max bandwidth is set to throttle traffic between sites when no network device can limit replication traffic. The max bandwidth option allows for different settings throughout the day, so you can assign a max bandwidth policy when your sites are busy with production data and disable the policy when they aren't as busy. Max bandwidth doesn't imply a maximum observed throughput.

When you talk with your networking teams, note that this setting is in megabytes per second, not megabits per second. NearSync doesn't currently honor maximum bandwidth thresholds.



Remote Site

1. Details 2. Settings

Bandwidth Throttling

BANDWIDTH THROTTLING

DEFAULT BANDWIDTH LIMIT 10

Enter the maximum bandwidth in megabytes per second (up to 2 decimal places).

BANDWIDTH THROTTLING POLICIES 40 MBps on Sunday, Monday, ...

Compression

COMPRESS ON WIRE

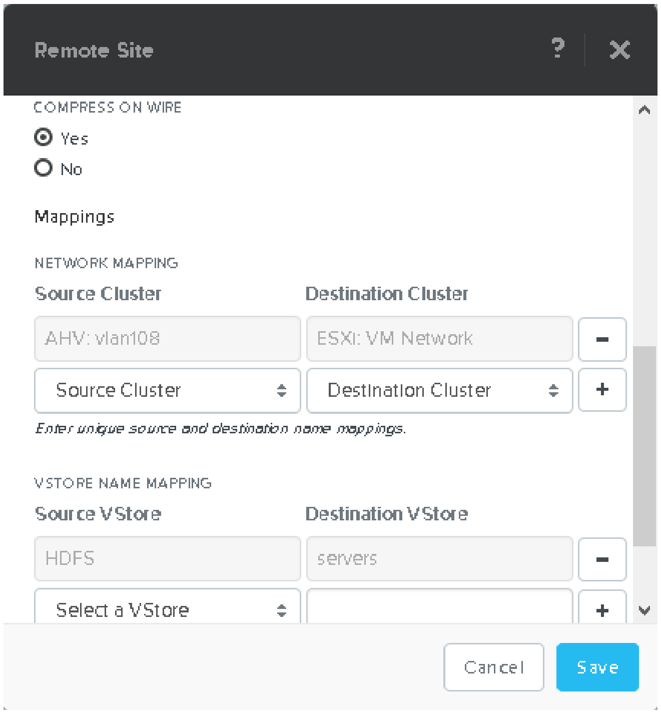
Mappings

Network Mapping

Figure 8: Max Bandwidth Settings

Remote Container

VStore name mapping identifies the container on the remote cluster used as the replication target. When you establish the VStore mapping, we recommend that you create a new, separate remote container with no VMs running on it on the remote side. This configuration allows the hypervisor administrator to distinguish failed-over VMs quickly and apply polices on the remote side easily in case of a failover.



COMPRESS ON WIRE

● Yes
○ No

Mappings

NETWORK MAPPING

Source Cluster	Destination Cluster
AHV: vlan108	ESXi: VM Network

VSTORE NAME MAPPING

Source VStore	Destination VStore
HDFS	servers

Figure 9: VStore Container Mappings for Replication

Best practices:

- Create a new remote container as the target for the VStore mapping.
- If you're backing up many clusters to one destination cluster, use only one destination container if the source containers have similar advanced settings.
- Enable MapReduce compression if licensing permits.
- If you use vCenter Server to manage both the primary and remote sites, don't have storage containers with the same name on both sites.

If the aggregate incoming bandwidth required to maintain the current change rate is less than 500 Mbps, we recommend skipping the performance tier. This setting saves your flash for other workloads while also saving on SSD write endurance. To skip the performance tier, use the following command from the nCLI:

```
ncli ctr edit sequential-io-priority-order=DAS-SATA,SSD-SATA,SSD-PCIe
 name=<container-name>
```

You can reverse this command at any time.

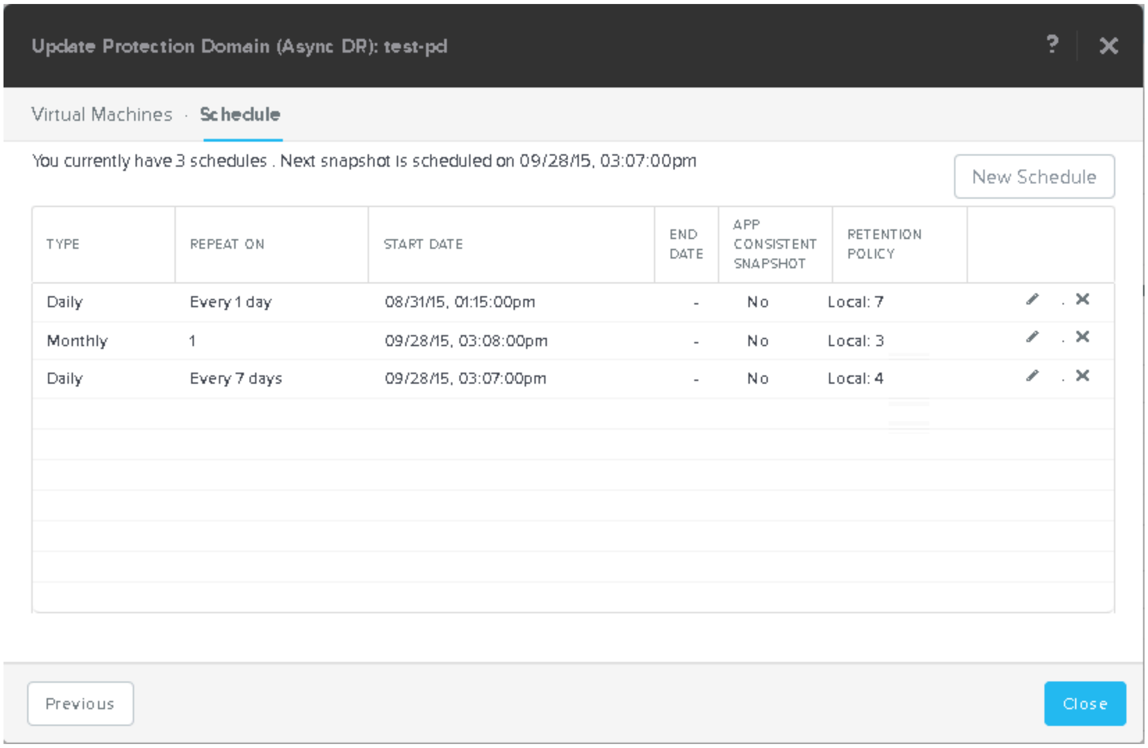
Network Mapping

AOS supports network mapping for disaster recovery migrations moving to and from AHV. Whenever you delete or change the network attached to a VM specified in the network map, modify the network map accordingly.

Scheduling Full Snapshots and Asynchronous Replication

The snapshot schedule should be equal to your desired RPO. In practical terms, the RPO determines how much data you can afford to lose in the event of a failure. The failure could be due to a hardware, human, or environmental issue. Taking a snapshot every 60 minutes for a server that changes infrequently or when you don't need a low RPO uses resources that could benefit more critical services.

The RPO is set from the local site. If you set a schedule to take a snapshot every hour, bandwidth and available space at the remote site determine if you can achieve the RPO. In constrained environments, limited bandwidth may cause the replication to take longer than the one-hour RPO, increasing the RPO. We list guidelines for sizing bandwidth and capacity to avoid this scenario later in this document.



You currently have 3 schedules . Next snapshot is scheduled on 09/28/15, 03:07:00pm

TYPE	REPEAT ON	START DATE	END DATE	APP CONSISTENT SNAPSHOT	RETENTION POLICY	
Daily	Every 1 day	08/31/15, 01:15:00pm	-	No	Local: 7	✎ · ✕
Monthly	1	09/28/15, 03:08:00pm	-	No	Local: 3	✎ · ✕
Daily	Every 7 days	09/28/15, 03:07:00pm	-	No	Local: 4	✎ · ✕

Figure 10: Multiple Schedules for a Production Domain

You can create multiple schedules for a protection domain (PD) using full snapshots, and you can have multiple PDs. The previous figure shows seven daily snapshots, four weekly snapshots, and three monthly snapshots to cover a three-month retention policy. This policy manages metadata on the cluster more efficiently than a daily snapshot with a 180-day retention policy.

Best practices:

- Stagger replication schedules across PDs. If you have a PD starting at the top of the hour, stagger the PDs by half of the most common RPO. The goal is to spread out replication impact on performance and bandwidth.
- Configure snapshot schedules to retain the lowest number of snapshots while still meeting the retention policy, as shown in the previous figure.

Remote snapshots implicitly expire based on how many snapshots there are and how frequently they are taken. For example, if you take daily snapshots and keep a maximum of five, on the sixth day the first snapshot expires. At that point, you can't recover from the first snapshot because the system deletes it automatically.

In case of a prolonged network outage, Nutanix always retains the last snapshot to ensure that you don't ever lose all the snapshots. You can modify the retention schedule from the nCLI by changing the min-snap-retention-count. This value ensures that you retain at least the specified number of snapshots, even if all the snapshots have reached the expiry time. This setting works at the PD level.

Scheduling LWS and NearSync

Nutanix offers NearSync replication with a telescopic schedule (time-based retention). When you set the RPO to at most 15 minutes and at least 1 minute, you can save your snapshots for X number of weeks or months. Once you select NearSync, you can't add any more schedules.

The following table presents the default telescopic schedule to save recovery points for one month.

Table 3: Default Telescopic Schedule for One Month

Type	Frequency	Retention
Minute increments	Every minute	15 minutes
Hourly	Every hour	6 hours
Daily	Every 24 hours	7 days
Weekly	Every week	4 weeks
Monthly	Every month	1 month

As NearSync continues to improve, please refer to the latest set of requirements and limitations in the Prism Web Console Guide.

Limit the number of VMs to 10 or fewer per protection domain. If you can, maintain one VM per protection domain to help you transition back to NearSync if you run out of LWS reserve storage.

Cross-Hypervisor Disaster Recovery

Nutanix provides cross-hypervisor disaster recovery for migrating between ESXi and AHV clusters. The migration works with one click and uses the Prism data protection workflow. Once you have installed the mobility drivers through NGT, VMs can move freely between the hypervisors.

Best practices:

- Configure the CVM external IP address.
- Obtain the mobility driver from NGT.
- Don’t migrate VMs:
  - With delta disks (hypervisor-based snapshots).
  - Using SATA disks.
- Ensure that protected VMs have an empty IDE CD-ROM attached.
- Ensure that network mapping is complete.

Single-Node Backup Target

Nutanix offers the ability to use an NX-1155 or NX-1175 appliance as a single-node backup target for an existing Nutanix cluster. Because this target has different resources than the original cluster, its primary use case is to provide backup for a small set of VMs. This utility gives SMB and ROBO customers a fully integrated backup option.

Best practices:

- Combined, all protection domains should have fewer than 30 VMs.
- Limit backup retention to a three-month policy.
  - We recommend a policy that includes seven daily, four weekly, and three monthly backups.
- Only map an NX-1155 or NX-1175 to one physical cluster.
- Snapshot schedule should be at least six hours.

- Turn off deduplication.

Cloud Connect

The CVM running in AWS and Azure has limited SSD space, so we recommend following these best practices when you size:

- Try to limit each protection domain to one VM to speed up restores. This approach also saves money, as it limits the amount of data going across the WAN.
- The RPO shouldn’t be lower than four hours.
- Turn off deduplication.
- Try to use Cloud Connect to protect workloads that have an average change rate of less than 0.5 percent.

9. Leap: Disaster Recovery Orchestration

The following best practices for disaster recovery orchestration cover both on-premises environments and Xi Leap. We have noted any differences between the two.

Required Infrastructure for Disaster Recovery Orchestration

- Deploy Prism Central.
  - For on-prem environments, deploy two Prism Central instances or one Prism Central instance at a separate site for high availability.
  - Leap for ROBO sites only requires one Prism Central, at the main site. When you replicate between two clusters managed from the same Prism Central instance, the minimum RPO is one hour or more and cross-hypervisor disaster recovery isn't supported.
- Deploy Prism Central on a subnet that doesn't fail over.
- Place the CVM and hypervisor IPs on a subnet different from the subnets used by VMs.
- The test network for on-prem disaster recovery orchestration requires a nonroutable VLAN.

Availability Zones

Paired availability zones synchronize the following disaster recovery configuration entities:

- Protection policies.
- Recovery plans.
- Categories used in protection polices and recovery plans.

Issues such as network connectivity loss between paired availability zones or user actions such as unpairing availability zones followed by pairing those availability zones again can affect entity synchronization. Pairing previously unpaired availability zones triggers an automatic synchronization event. If you don't update entities before you resolve a connectivity issue or pair the availability zones again, the synchronization behavior resumes. If you update entities in either or both availability zones before you resolve such issues or pair unpaired availability zones again, you can't synchronize the entities. In such a scenario, you can force the entities in one availability zone to synchronize with the paired availability zone. This forced synchronization overwrites entities at the paired availability zone.

Observe the following recommendations to avoid inconsistencies and the resulting synchronization issues:

- During network connectivity issues, don't update an entity at both availability zones in a pair. You can safely make updates at any one location. After the connectivity issue is resolved, force synchronization from the availability zone you made the updates in. Failure to adhere to this recommendation results in synchronization failures.
- You can safely create new entities in either or both availability zones as long as you don't assign the same name to entities in both availability zones. After you resolve the connectivity issue, force synchronization from the availability zone where you created the entities.
- If one of the availability zones becomes unavailable or if a service in the paired availability zone is down, perform a forced sync from the paired availability zone after you resolve the issue.

Protection Policies

A protection policy automates the creation and replication of snapshots. When you configure a protection policy for creating local snapshots, simply specify the RPO, retention policy, and the entities you want to protect. If you want to automate snapshot replication to a remote location, you can also specify the remote location. Leap uses entity-centric disaster recovery. There are a few requirements for using protection policies:

- A VM can only belong to a PD or a protection policy, not both.
- If you don't use Nutanix AHV IPAM and need to retain your IP addresses, you must install NGT on the VMs you want protected.
- You can associate a VM with a maximum of two protection policies.
- When you protect a VM with more than one protection policy, only one of those policies can use NearSync or have Xi Leap as the target.

Best practices for using protection policies:

- Apply protection polices using categories.
- For VMs that haven't been replicated before, create a container with the same name on both sides. If the VM was in a PD before, it continues to use the same container.
- All protection policies that use NearSync must have fewer than 200 VMs.

Cross-Hypervisor Disaster Recovery

Beginning with AOS 5.11, cross-hypervisor disaster recovery is supported for Xi Leap for ESXi on-premises. (Because Xi Leap already runs AHV, AHV on-premises just works.) Cross-hypervisor disaster recovery preserves these elements of the ESXi configuration on failback when your on-prem environment is once again healthy:

- Port group (based on network mapping).
- Virtual hardware version.
- Network adapter types.
- Disk adapter type.
- VMware Tools state.
- Number of vCPUs.
- Number of cores per vCPU.
- Static MAC address.
- Disk provisioning (thick or thin).
- OS type.

Cross-hypervisor disaster recovery has the following limitations:

- Doesn’t preserve SCSI controllers. Nutanix uses LSI logic on failback.
- No support for UEFI boot.
- Doesn’t preserve HA and Distributed Resource Scheduler (DRS) settings.
- No support for hypervisor-based snapshots, VMware VSS, or linked clones.

Recovery Plans

A recovery plan orchestrates restoring protected VMs at a backup location. Recovery plans can either recover all specified VMs at once or, using what is essentially a runbook functionality, use power-on sequences with optionally configurable interstage delays to recover applications gracefully and in the required order. Recovery plans that restore applications in Xi Leap can also create the required networks during failover and can assign public-facing IP addresses to VMs.

Requirements for recovery plans:

- To enable disaster recovery orchestration, you must have set up a Prism Element external data services IP.
- Prism Central must run on a Nutanix cluster with the external data services IP.

Xi Leap doesn’t allow you to create a recovery plan if:

- The recovery network that you specify exists with the same name on multiple clusters.
- The networks have the same name but different IP address spaces.

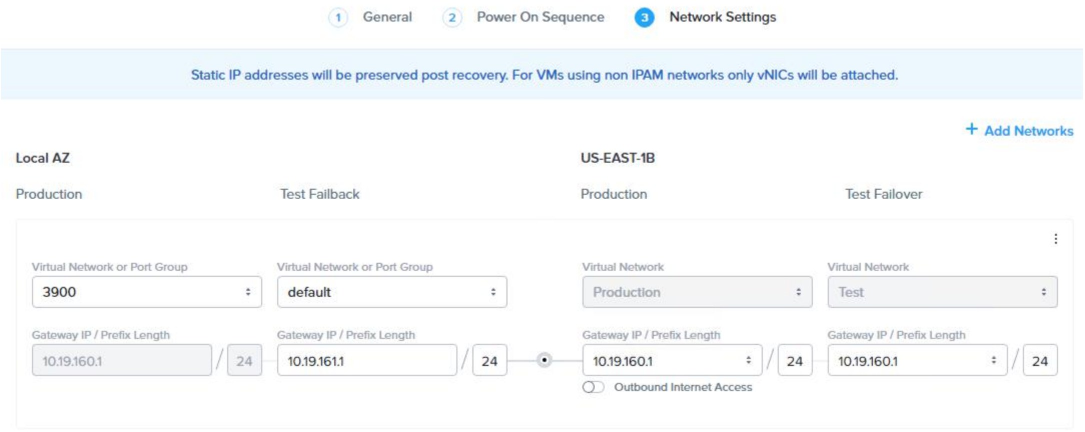
> Note: If you add a VM to multiple recovery plans and perform failover simultaneously on those recovery plans, each recovery plan creates an instance of the VM at the recovery location. You must manually clean up the additional instances.

Best practices:

- For on-premises availability zones, create a nonroutable network for testing failovers.

- Run the Validate workflow after you make changes to recovery plans.
- After you run the Test workflow, run the Clean-Up workflow instead of manually deleting VMs.
- A recovery plan should cover a maximum of 200 VMs.
- Maximum of 50 categories in a recovery plan.
- Maximum of 20 stages in a recovery plan.
- Maximum of 15 categories per stage in a recovery plan.
- Only run a maximum of 5 recovery plans in parallel.
- Keep asynchronous or NearSync replication VMs in separate recovery plans from synchronous replication VMs because synchronous recovery plans don't currently support planned failover.

Network Mapping



Static IP addresses will be preserved post recovery. For VMs using non IPAM networks only vNICs will be attached.

Local AZ		US-EAST-1B	
Production	Test Failback	Production	Test Failover
Virtual Network or Port Group: 3900	Virtual Network or Port Group: default	Virtual Network: Production	Virtual Network: Test
Gateway IP / Prefix Length: 10.19.160.1 / 24	Gateway IP / Prefix Length: 10.19.161.1 / 24	Gateway IP / Prefix Length: 10.19.160.1 / 24 ○ Outbound Internet Access	Gateway IP / Prefix Length: 10.19.160.1 / 24

Figure 11: Network Mapping for Xi Leap

Virtual Networks in On-Premises Clusters

Virtual networks in on-premises Nutanix clusters are virtual subnets bound to a single VLAN. At physical locations, including the recovery location, administrators must create these virtual subnets manually, with separate virtual subnets created for production and test purposes. You must create these virtual subnets before you configure recovery plans.

When you configure a recovery plan, map the virtual subnets at the source location to the virtual subnets at the recovery location.

Virtual Networks in Xi Cloud

Xi Cloud features two built-in virtual networks: Production and Test. These virtual networks aren't analogous to the virtual subnets used in on-premises Nutanix clusters; they only provide two separate IP address spaces for production and testing so that activities performed in one do not affect the other. These separate Production and Test IP address spaces contain virtual subnets that are analogous to the virtual subnets in on-premises Nutanix clusters.

The Production virtual network contains subnets used for production workloads. These production workloads can be either workloads created and maintained in Xi Cloud or workloads that have failed over from a paired physical location.

The Test virtual network contains the subnets you want to recover VMs to when you test failover from the virtual subnets at a paired physical location. When you configure a recovery plan, map the virtual subnets on your on-premises clusters to virtual subnets in the Production and Test networks in Xi Cloud.

Best practices:

- Set up administrative distances on VLANs for subnets that completely fail over. If you don't set up administrative distances, shut down the VLAN on the source side after failover if the VPN connection is maintained between the two sites. If you are failing over to a new subnet, set up the subnet beforehand so you can test the routing.
- The prefix length for network mappings at the source and the destination must be the same.
- If you don't use Nutanix IPAM, you must install NGT to maintain a static address.
- To maintain a static address for Linux VMs that don't use Nutanix IPAM, the VMs must have the NetworkManager command-line tool (nmcli) version 0.9.10.0 or later installed. Additionally, you must use NetworkManager to

manage the network for the Linux VMs. To enable NetworkManager on a Linux VM, set the value of the NM_CONTROLLED field to yes in the interface configuration file (for example, in CentOS, the file is /etc/sysconfig/network-scripts/ifcfg-eth0). After you set this field, restart the network service on the VM.

For networking best practices specific to Xi Leap, refer to the Xi Connectivity Tech Note.

Xi Leap Hypervisor Support

- Xi Leap supports clusters running AHV and ESXi as the source.

Xi Leap VM Configuration Limitations

VMs with the following configurations can't start:

- VMs configured with a GPU resource.
- VMs configured with four vNUMA sockets.

10. Sizing Space

Full Local Snapshots

To size space for local snapshots, you need to account for the rate of change in your environment and how long you plan to keep your snapshots on the cluster. Reduced snapshot frequency may increase the rate of change because there's a greater chance of common blocks changing before the next snapshot.

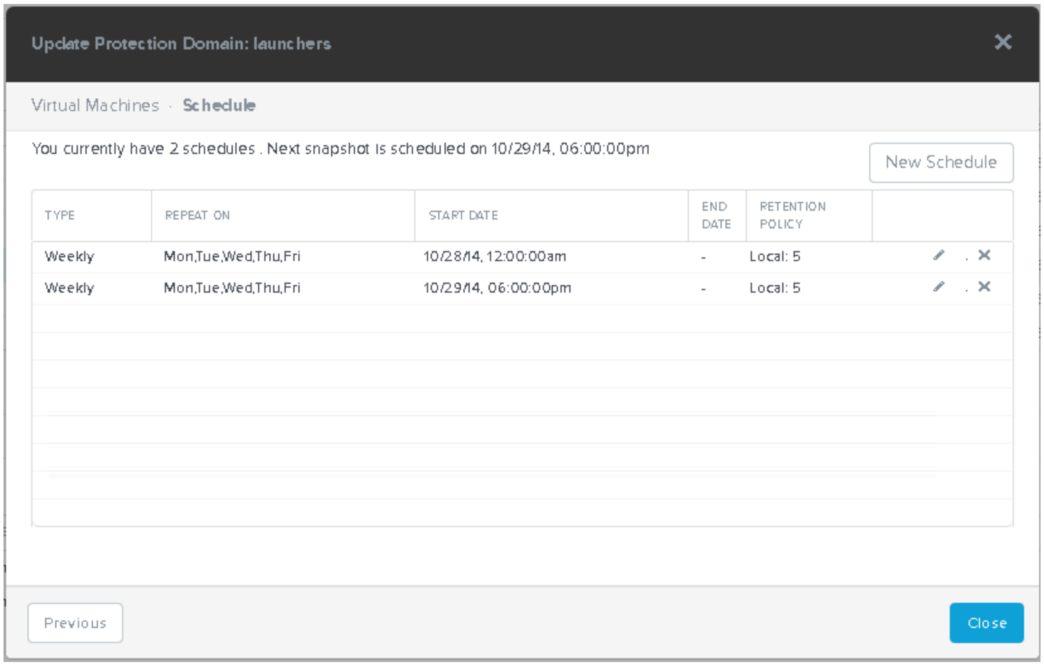
To find the space needed to meet your RPO, you can use the following formula. As you lower the RPO for asynchronous replication, you may need to account for an increased rate of transformed garbage. Transformed garbage is space that was allocated for I/O optimization or space that was assigned but to which the metadata no longer refers.

Table 4: Local Full Snapshot Reserve Formula

snapshot reserve = (frequency of snapshots × change rate per frequency) + (change rate per frequency × # of snapshots in a full curator scan × 0.1)

Note: A full curator scan runs every six hours.

You can look at your backups and compare the incremental differences between them to find the change rate. You could also take a conservative approach and start with a low snapshot frequency and a short expiry policy, then gauge the size difference between backups before consuming too much space.



Update Protection Domain: launchers

Virtual Machines · **Schedule**

You currently have 2 schedules . Next snapshot is scheduled on 10/29/14, 06:00:00pm

TYPE	REPEAT ON	START DATE	END DATE	RETENTION POLICY
Weekly	Mon,Tue,Wed,Thu,Fri	10/28/14, 12:00:00am	-	Local: 5
Weekly	Mon,Tue,Wed,Thu,Fri	10/29/14, 06:00:00pm	-	Local: 5

Figure 12: Example Snapshot Schedule: Snapshot at Noon and 6 PM

Using the local snapshot reserve formula presented above and assuming for demonstration purposes that the change rate is 35 GB of data every six hours and that we keep ten snapshots:

```
snapshot reserve = (frequency of snapshots × change rate per frequency) +
(change rate per frequency × # of snapshots in a full curator scan × 0.1)
= (10 × 35,980 MiB) + (35,980 MiB × 1 × 0.1)
= 359,800 + (35,980 × 1 × 0.1)
= 359,800 + 3,598
= 363,398 MiB
= 363 GiB
```

Asynchronous Replication

Asynchronous replication uses the same process, but you must include the first full copy of the PD plus delta changes based on the set schedule.

Table 5: Remote Snapshot Reserve Formula

snapshot reserve = (frequency of snapshots × change rate per frequency) + (change rate per frequency × # of snapshots in a full curator scan × 0.2) + total size of the source PD

For the minimum amount of space needed at the remote side, 130 percent of the PD is a good average to work from.

If the remote target also runs a workload, incoming replication uses the performance tier. If you use a hybrid cluster, size for the additional hot data. You can also skip the performance tier by creating a separate container for incoming replication and following the steps provided in the Remote Container section.

Lightweight Snapshots (LWS)

Sizing for LWS is very similar to sizing for full snapshots in that you must account for change rate and retention time. During the LWS retention time, you don’t have to size for transformed garbage space, but LWS does use additional SSD space that you have to size for. By default, 7 percent of each node’s SSD space forms the LWS reserve, which is an additional factor.

Duplicating the workload from the full snapshot example, let’s change the RPO to one minute. LWS data exists for 75 minutes plus one extra snapshot based on the RPO value. Because the frequency rate in this scenario is high, it’s very important to account for overwrites in the change rate. Because all data goes through the oplog, it’s compressed; the type and amount of compression varies by workload. Using inline compression, we typically see a 2:1 compression rate.

Example Using 35 GB Change Rate Over 6 Hours

If your change rate is 35 GB of data every six hours, add 5 GB to account for overwrites. 40 GB every six hours has a change rate of approximately 114 MB per minute. If the cluster is running with replication factor 2, you must account for the total physical space.

Table 6: LWS Reserve

LWS reserve	= (frequency of snapshots × change rate per frequency) × oplog compression × replication factor
	= (75 × 114 MB per minute) × 0.50 × 2
	= 8,550 MB
	= 8.6 GB

If you run a 3460 hybrid system with two 1.9 TB SSDs, your LWS reserve is shown with the following calculation.

Table 7: Cluster LWS Reserve

LWS cluster reserve	= ((total SSD capacity per node - (CVM + metadata + oplog + cache overhead)) × number of nodes) × LWS reserve percentage
	= (3,539 GiB - (120 GiB + 30 GiB + 200 GiB + 40 GiB)) × 4 × 0.07
	= 3,149 GB × 4 × 0.07
	= 881.72 GiB
	= ~945 GB

Apply the 7 percent of the SSD space used by the extent store after you account for the rest of the system. Because the LWS cluster reserve is 945 GB, this system has lots of room for the workload as well as for additional business-critical applications.

As we discussed in the Scheduling LWS and NearSync section, a telescopic schedule has a total of six hourly, seven daily, four weekly, and one monthly snapshots. You only need to account for additional garbage space for the dailies because of the higher frequency. Using the change rate above, you can calculate each separate schedule and add them all together. Because full snapshots occur less often than LWS, you don't have to be concerned with overwrites and can use the original 35 GB change rate for every six hours.

Example of a Telescopic NearSync Snapshot Schedule

```
Hourly change rate:      5,833 MB
```

```
Daily change rate:     140,000 MB

Weekly change rate:    980,000 MB

Monthly change rate: 3,920,000 MB

snapshot overall capacity reserve = (hourly schedule + daily schedule + weekly
 schedule + monthly schedule)

For the hourly schedule:

(frequency of snapshots × change rate per frequency) + (change rate per frequency
 × # of snapshots in a full curator scan × 0.1)

The remaining schedules:

= (frequency of snapshots × change rate per frequency)

snapshot overall capacity reserve = ((6 × 5,833)+(5,833 × 6 × 0.1)) + (7 ×
 140,000) + (4 × 980,000) + (2 × 3,920,000)

= 34,998 + 3,500 + 980,000 + 3,920,000 + 8,257,538

= 13,196,036 MB

= ~13 TB
```

For most small and medium businesses, a daily change of 140 GB is considered high. This NearSync example highlights the difference between keeping a lot of snapshots around and keeping only the 10 snapshots in the full snapshot example.

NearSync

NearSync uses the same process as the snapshot process outlined above, but you must include the first full copy of the protection domain as well as delta changes based on the set schedule.

For the minimum amount of space needed at the remote site, start with 130 percent of the PD plus the required LWS reserve space.

Best practices:

- Use drives with the same or greater capacity at the remote site compared to those in your primary cluster to ensure that there’s enough LWS reserve space.
- Don’t enable deduplication on the source container for NearSync.

11. Bandwidth

You must have enough available bandwidth to keep up with the replication schedule. If you're still replicating when the next snapshot is scheduled, the current replication job finishes first. The newest outstanding snapshot then starts to get the newest data to the remote side first. To help replication run faster when you have limited bandwidth, you can seed data on a secondary cluster at the primary site before shipping that cluster to the remote site.

Seeding

To seed data for a new site:

- Set up a secondary cluster with local IPs at the primary site.
- Enable compression on the remote site in the production domain.
- Set the initial retention time to 3 months.
- Once setup completes, reconfigure the secondary cluster with remote IPs.
- Shut down the secondary cluster and ship it to the remote site.
- Power on the remote cluster and update the remote site on the primary cluster to the new IP.

If you can't seed the protection domain at the local site, you can create the remote cluster as a normal install and turn on compression over the wire. Manually create a one-time replication with retention time set to 3 months. We recommend this retention time setting because of the extra time it takes to replicate the first data set across the wire.

To figure out the needed throughput, you must know your RPO. If you set the RPO to one hour, you must be able to replicate the changes within that time.

Assuming you know your change rate based on incremental backups or local snapshots, you can calculate the bandwidth needed. The next example uses a change rate of 15 GB and an RPO of one hour. We don't use deduplication in the calculation, partly so the dedupe savings can serve as a buffer in the overall calculation and partly because the one-time cost for deduped data going over the wire has less impact once the data is present at the remote site. We assume an average of 30 percent bandwidth savings for compression on the wire.

Table 8: Bandwidth Sizing

Bandwidth needed = (RPO change rate × (1 - compression on wire savings %)) / RPO
Example:
(15 GB × (1 - 0.3)) / 3,600s
(15 GB × 0.7) / 3,600s
10.5 GB / 3,600s
(10.5 × 1,000 MB) / 3,600s (changing to MBps)
(10,500 MB) / 3,600s
10,500 MB / 3,600 = 2.92 MBps
Bandwidth needed = 23.33 Mbps

You can easily perform the calculation online using http://www.wolframalpha.com/input/?i=(15+GB+*+(1-30%25))%2F(1+hour).

If you keep missing your replication schedule, either increase your bandwidth or your RPO. To allow more RPO flexibility, you can run different schedules on the same production domain; for example, have one daily replication schedule and create a separate schedule to take local snapshots every two hours.

12. Failover: Migrate vs. Activate

Nutanix offers two options for handling VMs and volume groups during failover: migrate and activate.

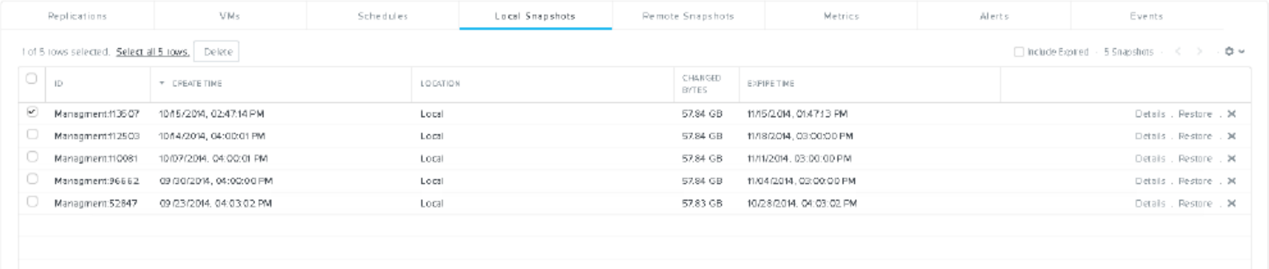
Use migrate when you know that both the production site and the remote site are still healthy. This option is only available on the active production site. Migrate shuts down the VM or volume group, takes another snapshot, and replicates the VM or volume group to the selected remote site.

The activate option for restoring a VM is only available on the remote site. When you select activate, the system uses the last snapshot on the remote side to bring up the VM regardless of whether the active production site is healthy. You can't sync any outstanding changes to the VM from the production site.

If you activate a PD because the primary site is down but the primary site comes back up after the failover, you can have a split-brain scenario. To resolve this situation, deactivate the PD on the former primary site. The following command is hidden from the nCLI because it deletes the VMs, but it resolves the split while keeping the existing snapshots:

```
ncli> pd deactivate_and_destroy_vms name=<protection_Domain_Name>
```

You can test a VM at the remote site without breaking replication using the restore or clone functionality.



1 of 5 rows selected. Select all 5 rows. Delete

□ Include Expired · 5 Snapshots

□	ID	▾ CREATE TIME	LOCATION	CHANGED BYTES	EXPIRE TIME	
☑	Managment:113607	10/15/2014, 02:47:14 PM	Local	57.84 GB	11/15/2014, 01:47:13 PM	Details · Restore · ✕
□	Managment:112503	10/14/2014, 04:00:01 PM	Local	57.84 GB	11/18/2014, 03:00:00 PM	Details · Restore · ✕
□	Managment:110081	10/07/2014, 04:00:01 PM	Local	57.84 GB	11/11/2014, 03:00:00 PM	Details · Restore · ✕
□	Managment:96662	09/30/2014, 04:00:00 PM	Local	57.84 GB	11/04/2014, 03:00:00 PM	Details · Restore · ✕
□	Managment:52847	09/23/2014, 04:03:02 PM	Local	57.83 GB	10/28/2014, 04:03:02 PM	Details · Restore · ✕

Figure 13: Cloning a VM for Testing Using the Remote Local Snapshot Browser

Using the local snapshot browser on the inactive production domain at the remote site, choose the restore option to clone a VM to the datastore. Add a prefix to the VM’s path.

Best practices:

- When you activate PDs on the remote site, use intelligent placement for Hyper-V and DRS for ESXi clusters. Intelligent placement evenly spreads out the VMs on boot during a failover. AOS powers on VMs uniformly at boot time.
- Install NGT on machines using volume groups.
- Configure the data services cluster IP on the remote cluster.

Protection Domain Cleanup

Because inactive protection domains still consume space in existing snapshots, you should remove any unused PDs to reclaim that space.

To remove a PD from a cluster, follow these steps:

1. Remove existing schedules.
2. Remove both local and remote snapshots.
3. Remove the VMs from the active PD.
4. Delete the PD.

13. Self-Service File Restore

Self-service file restore is available for ESXi and AHV. The goal is to offer self-service file-level restore with minimal support from infrastructure administrators. Once you've installed NGT, you can open a web browser and go to http://localhost:5000 to browse snapshots sorted by today, last week, last month, and user-defined criteria. Self-service works with both protection domains and Leap.

- Guest VMs should be:
  - Windows 2008, Windows 7, or a later version
  - Centos 6.5+ and 7.0+
  - Red Hat 6.5+ and 7.0+
  - OEL 6.5+
  - SLES 11+
- Install VMware Tools.
- Remove JRE 1.8 or later if installed with a previous release.
- Configure a cluster external IP address.
- Add the VM to a protection domain.
- Use the default disk.EnableUUID = true for the VM in advanced settings for ESXi.
- The VM must have an IDE-based CD-ROM configured for installation. On newer versions of ESXi, a SATA-based CD-ROM is added by default.
- Detach the mounted disk after you restore your files.

Limitations:

- For Linux VMs, logical volumes spanning multiple disks aren't supported.

14. Third-Party Backup Products

Nutanix provides its own hypervisor-agnostic changed-region tracking API that vendors can access using a REST API. Currently, Cohesity, Commvault, HYCU, and Rubrik provide backup support specifically for AHV. Nutanix systems can integrate with any backup vendor that supports vStorage APIs for Data Protection for ESXi. Nutanix also offers a VSS provider for Hyper-V backup software vendors. Visit the Nutanix Elevate Technology Alliance Partner Program page for additional details on certified third-party solutions.

Backups with Replication

> Note: Hypervisor-based snapshots left on the VM when you take a hardware-based snapshot may cause the recovery to fail. You must manually turn on the VM. Refer to VMware KB 1025279 for more information.

Backup software interacting with AHV uses scoped production domains, also known as backup snapshots. When the backup vendor requests a snapshot of a VM that isn't protected under any user-created production domain, temporarily protect the VM by placing it in a scoped production domain. Issue a snapshot of this production domain, then remove the VM from the production domain. Backup snapshots also allow you to back up VMs and use Leap disaster recovery orchestration. The only possible issue is that while the snapshot operation on the scoped production domain is in progress, the user-initiated VM protection may fail in the regular production domain.

Best practices:

- Avoid using hypervisor-based snapshots.
- Try to schedule hypervisor snapshots or user-created hardware-based snapshots at different times than the backup window.

Nutanix has published recommendations for optimizing and scaling third-party backup solutions in best practices guides focused on an enhanced disk-to-disk backup architecture. Find these guides on the Nutanix Support Portal.

Note: Nutanix Files deployments aren’t certified for backup with Cohesity because of the current Cohesity architectural implementation. Issues related to Cohesity backups and configuration for Nutanix Files aren’t supported by Nutanix.

15. Conclusion

Nutanix offers granular data protection based on the required recovery point objectives across many different deployment models. As your application requirements change and your cluster grows, you have the flexibility to add and remove VMs from protection domains. Sizing capacity and bandwidth are key to achieving optimal data protection and maintaining cluster health. Snapshot frequency and the daily change rate affect the capacity and bandwidth needed between sites to meet the needs of the business.

With data protection features that are purpose-built, fully integrated, and 100 percent software defined, Nutanix provides the ultimate in adaptability and flexibility for meeting your enterprise’s backup and recovery needs.

Appendix

Best Practices Checklist

Following are high-level best practices for backup and disaster recovery on Nutanix.

General

- All VM files should sit on Nutanix storage. If non-Nutanix storage stores files externally, it should have the same file path on both sides.
- Remove all external devices, including ISOs or floppy devices.

Nutanix Native VSS Snapshots

- Configure an external cluster IP address.
- Guest VMs should be able to reach the external cluster IP on port 2074.
- Guest VMs should have an empty IDE CD-ROM for attaching NGT.
- Only available for ESXi and AHV.
- Virtual disks must use the SCSI bus type.
- VSS must be running in the guest VM. The appendix has a Check for VSS Service PowerShell script that verifies whether the service is running.
- The guest VM needs to support the use of VSS writers. The appendix has a Check VSS Writers PowerShell script that makes sure the VSS writers are stable.
- Schedule application-consistent snapshots during off-peak hours or ensure that additional I/O performance is available. If you take a VSS snapshot during peak usage, the delta disk from the hypervisor-based snapshot could become large. When you delete the hypervisor-based snapshot, collapsing it takes additional I/O; account for this additional I/O to avoid affecting performance.

Hyper-V VSS Provider

- VSS support is only for backup.
- Create different containers for VMs that need VSS backup support. Limit the number of VMs on each container to 50.
- Create a separate large container for crash-consistent VMs.

Protection Domains

- Protection domain names must be unique across sites.
- Group VMs with similar RPO requirements.
- Maximum of 200 VMs per protection domain for full snapshots.
- Maximum of 10 VMs per protection domain for LWS.
- VMware Site Recovery Manager and Metro Availability protection domains are limited to 50 VMs.
- Linked clone VMs (typically nonpersistent View desktops) aren't supported with NearSync.
- Remove unused protection domains to reclaim space.
- If you must activate a protection domain rather than migrate it, deactivate the old primary protection domain when the site comes back up.

Consistency Groups

- Keep consistency groups as small as possible. Keep dependent applications or service VMs in one consistency group to ensure that they are recovered in a consistent state (for example, App and DB).
- Each consistency group using application-consistent snapshots can contain only one VM.

Disaster Recovery and Backup

- Ensure that you configure forward (DNS A) and reverse (DNS PTR) DNS entries for each ESXi management host on the DNS servers used by the Nutanix cluster.

Remote Sites

- Use the external cluster IP as the address for the remote site.
- Use disaster recovery proxy to limit firewall rules.
- Use max bandwidth to limit replication traffic.
- When activating protection domains, use intelligent placement for Hyper-V and DRS for ESXi clusters on the remote site. Intelligent placement evenly spreads out the VMs on boot during a failover. AOS powers on VMs uniformly at boot time.
- If you use vCenter Server to manage both the primary and remote sites, don’t use storage containers with the same name on both sites.

Remote Containers

- Create a new remote container as the target for the VStore mapping.
- When you back up many clusters to one destination cluster, use only one destination container if the source containers have similar advanced settings.
- Enable MapReduce compression if licensing permits.
- If the aggregate incoming bandwidth required to maintain the current change rate is less than 500 Mbps, we recommend skipping the performance tier to save flash capacity and increase device longevity.

Network Mapping

- Whenever you delete or change the network attached to a VM specified in the network map, modify the network map accordingly.

Scheduling

- To spread out replication impact on performance and bandwidth, stagger replication schedules across protection domains. If you have a PD starting at the top of the hour, stagger the PDs by half of the most common RPO.
- Configure snapshot schedules to retain the smallest number of snapshots while still meeting the retention policy.
- Metro and SRM-protected containers aren’t supported.

- Deduplication on the source container isn't currently supported for NearSync.

Cross-Hypervisor Disaster Recovery

- Configure CVM external IP address.
- Obtain the mobility driver from NGT.
- Don't migrate VMs with delta disks (hypervisor-based snapshots), using SATA disks, or using volume groups.
- Ensure that protected VMs have an empty IDE CD-ROM attached.
- Ensure that network mapping is complete.

Disaster Recovery Orchestration

- Deploy Prism Central to each on-prem site.
- Deploy Prism Central on a subnet that doesn't fail over.
- Place CVM and hypervisor IPs on a different subnet than the subnets used by VMs.
- On-prem disaster recovery orchestration requires a nonroutable VLAN for the test network.

Availability Zones

- If one of the availability zones becomes unavailable or if a service in the paired availability zone is down, perform a forced sync from the paired availability zone after the issue is resolved.

Protection Polices

- A VM can only belong to either a protection domain or a protection policy.
- If you don't use Nutanix AHV IPAM and need to retain your IP addresses, install NGT on the VMs to be protected.
- Use categories to apply protection policies.
- For on-premises Leap, create the same container name on both sides. If the container name doesn't match on both sides, data replicates by default to the SelfServiceContainer.

Cross-Hypervisor Support with ESXi for Protection Polices

- No support for UEFI boot.
- Doesn’t preserve HA and DRS settings.
- No support for hypervisor-based snapshots, VMware VSS, or linked clones.

Recovery Plans

- For on-premises availability zones, create a nonroutable network for testing failovers.
- Run the Validate workflow after making changes to recovery plans.
- After you run the Test workflow, run the Clean-Up workflow instead of manually deleting VMs.
- A recovery plan should cover a maximum of 200 VMs.
- Maximum of 50 categories in a recovery plan.
- Maximum of 20 stages in a recovery plan.
- Maximum of 15 categories per stage in a recovery plan.
- You can run a maximum of 5 recovery plans in parallel.

Network Mapping

- Set up administrative distances on VLANs for subnets that completely fail over. If you don’t set up administrative distances, shut down the VLAN on the source side after failover if the VPN connection is maintained between the two sites. If you’re failing over to a new subnet, set up the subnet beforehand so you can test the routing.
- The prefix length for network mappings at the source and the destination must be the same.
- If you don’t use Nutanix IPAM, you must install the NGT software package to maintain a static address.
- To maintain a static address for Linux VMs that don’t use Nutanix IPAM, the VMs must have the NetworkManager command-line tool (nmcli) version 0.9.10.0 or later installed. Additionally, you must use NetworkManager to manage the network for the Linux VMs. To enable NetworkManager on a

Linux VM, set the value of the NM_CONTROLLED field to yes in the interface configuration file (for example, in CentOS, the file is /etc/sysconfig/network-scripts/ifcfg-eth0). After you set the field, restart the network service on the VM.

Xi Leap Hypervisor Support

- Xi Leap only supports clusters running AHV.

Xi Leap Virtual Machine Configuration Restrictions

- Can't power on VMs configured with a GPU resource.
- Can't power on VMs configured with four vNUMA sockets.

Single-Node Backup

- Combined, all protections domains should be under 30 VMs.
- Limit backup retention to a three-month policy. We recommended seven daily, four weekly, and three monthly backups.
- Only map an NX-1155 to one physical cluster.
- Snapshot schedule should be at least six hours.
- Turn off deduplication.

Cloud Connect

- Try to limit each protection domain to one VM to speed up restores. This approach also saves money, as it limits the amount of data going across the WAN.
- The RPO shouldn't be lower than four hours.
- Turn off deduplication.
- Try to use Cloud Connect to protect workloads that have an average change rate of less than 0.5 percent.

Sizing

- Size local and remote snapshot usage using the application's change rate.

- Remember to either size the performance tier for hybrid clusters to accommodate incoming data or bypass the performance tier and write directly to disk.

Bandwidth

- Seed locally for replication if WAN bandwidth is limited.
- Set a high initial retention time for the first replication when you seed.

File-Level Restore

- Guest VMs should be Windows 2008, Windows 7, or a later version.
- Install VMware tools.
- Install JRE 1.8 or later.
- Configure a cluster external IP address.
- Add the VM to a protection domain.
- Use the default disk.EnableUUID = true for the VM in advanced settings.
- The VM must have a CD-ROM configured.
- Detach the mounted disk after restoring your files.

PowerShell Scripts

Check for VSS Service

This PowerShell script checks whether the VSS service is running as required for application-consistent snapshots.

```
#Connect to the Nutanix cluster of your choice, try to use the external address.

Connect-NutanixCluster -AcceptInvalidSSLCerts -server External_cluster_IP -
UserName admin

#load Nutanix CMDlets, make sure your local version matches the cluster version

Add-PSSnapin NutanixCmdletsPSSnapin

#Get a list of all Consistency Groups

$pdvss = Get-NTNXProtectionDomainConsistencyGroup

#array of all the appConsistentVMs
```

```
$appConsistentVM = @()
Foreach ($vssVM in $pdvss)
  {
    if ($vssVM.appConsistentSnapshots)
  {
     $appConsistentVM += $vssVM.consistencyGroupName
    }
  }
 get-service -name VSS -computername $appConsistentVM | format-table -property
 MachineName, Status, Name, DisplayName -auto
```

Check VSS Writers

This PowerShell script checks whether VSS writers are stable for VMs that run application-consistent snapshots.

```
function Get-VssWriters {
<#
 .Synopsis
 Function to get information about VSS Writers on one or more computers
 .Description
 Function will parse information from VSSAdmin tool and return object containing
 WriterName, StateID, StateDesc, and LastError
 Function will display a progress bar while it retrives information from
 different
 computers.
 .Parameter ComputerName
 This is the name (not IP address) of the computer.
 If absent, localhost is assumed.
 .Example
 Get-VssWriters
 This example will return a list of VSS Writers on localhost
 .Example
 # Get VSS Writers on localhost, sort list by WriterName
 $VssWriters = Get-VssWriters | Sort "WriterName"
 $VssWriters | FT -AutoSize # Displays it on screen
 $VssWriters | Out-GridView # Displays it in GridView

 $VssWriters | Export-CSV ".\myReport.csv" -NoTypeInformation # Exports it to CSV

 .Example

 # Get VSS Writers on the list of $Computers, sort list by ComputerName

 $Computers = "xHost11","notThere","xHost12"

 $VssWriters = Get-VssWriters -ComputerName $Computers -Verbose | Sort
 "ComputerName"

 $VssWriters | Out-GridView # Displays it in GridView

 $VssWriters | Export-CSV ".\myReport.csv" -NoTypeInformation # Exports it to CSV

 .Example

 # Reports any errors on VSS Writers on the computers listed in
 MyComputerList.txt, sorts list by ComputerName

 $Computers = Get-Content ".\MyComputerList.txt"

 $VssWriters = Get-VssWriters $Computers -Verbose |

  Where { $_.StateDesc -ne 'Stable' } | Sort "ComputerName"

 $VssWriters | Out-GridView # Displays it in GridView

 $VssWriters | Export-CSV ".\myReport.csv" -NoTypeInformation # Exports it to
 CSV

 .Example

 # Get VSS Writers on all computers in current AD domain, sort list by
 ComputerName

 $Computers = (Get-ADComputer -Filter *).Name

 $VssWriters = Get-VssWriters $Computers -Verbose | Sort "ComputerName"

 $VssWriters | Out-GridView # Displays it in GridView

 $VssWriters | Export-CSV ".\myReport.csv" -NoTypeInformation # Exports it to CSV

 .Example

 # Get VSS Writers on all Hyper-V hosts in current AD domain, sort list by
 ComputerName

 $FilteredComputerList = $null

 $Computers = (Get-ADComputer -Filter *).Name

Foreach ($Computer in $Computers) {

   if (Get-WindowsFeature -ComputerName $Computer -ErrorAction SilentlyContinue
 |

    where { $_.Name -eq "Hyper-V" -and $_.InstallState -eq "Installed"}) {

     $FilteredComputerList += $Computer

   }
 }
 $VssWriters = Get-VssWriters $FilteredComputerList -Verbose | Sort
 "ComputerName"
 $VssWriters | Out-GridView # Displays it in GridView
 $VssWriters | Export-CSV ".\myReport.csv" -NoTypeInformation # Exports it to CSV
 .OUTPUT
 Scripts returns a PS Object with the following properties:
  ComputerName

  WriterName
  StateID

  StateDesc

  LastError

 .Link
 https://superwidgets.wordpress.com/category/powershell/
 .Notes
 Function by Sam Boutros
 v1.0 - 09/17/2014
#>
[CmdletBinding(SupportsShouldProcess=$true,ConfirmImpact='Low')]
  Param(
    [Parameter(Mandatory=$false,
         ValueFromPipeLine=$true,
         ValueFromPipeLineByPropertyName=$true,
         Position=0)]
      [ValidateNotNullorEmpty()]
      [String[]]$ComputerName = $env:COMPUTERNAME
  )
  $Writers = @()
  $k = 0
  foreach ($Computer in $ComputerName) {
    try {
      Write-Verbose "Getting VssWriter information from computer $Computer"

      $k++

      $Progress = "{0:N0}" -f ($k*100/$ComputerName.count)

      Write-Progress -Activity "Processing computer $Computer ... $k out of
 $($ComputerName.count) computers" `

        -PercentComplete $Progress -Status "Please wait" -CurrentOperation
 "$Progress% complete"

      $RawWriters = Invoke-Command -ComputerName $Computer -ErrorAction Stop -
ScriptBlock {

        return (VssAdmin List Writers)

      }

      for ($i=0; $i -lt ($RawWriters.Count-3)/6; $i++) {

        $Writer = New-Object -TypeName psobject

        $Writer| Add-Member "ComputerName" $Computer

        $Writer| Add-Member "WriterName" $RawWriters[($i*6)+3].Split("'")[1]

        $Writer| Add-Member "StateID" $RawWriters[($i*6)+6].SubString(11,1)

        $Writer| Add-Member "StateDesc" $RawWriters[($i*6)+6].SubString(14,
$RawWriters[($i*6)+6].Length - 14)

        $Writer| Add-Member "LastError" $RawWriters[($i*6)+7].SubString(15,
$RawWriters[($i*6)+7].Length - 15)

        $Writers += $Writer

      }

      Write-Debug "Done"

    } catch {

      Write-Warning "Computer $Computer is offline, does not exist, or cannot be
 contacted"

    }

  }

  return $Writers

}

#Connect to the Nutanix cluster of your choice, try to use the external address.

Connect-NutanixCluster -AcceptInvalidSSLCerts -server External_cluster_IP -
UserName admin

#load Nutanix CMDlets, make sure your local version matches the cluster version

Add-PSSnapin NutanixCmdletsPSSnapin

#Get a list of all Consistency Groups
```

```
$pdvss = Get-NTNXProtectionDomainConsistencyGroup
#array of all the appConsistentVMs
$appConsistentVM = @()
Foreach ($vssVM in $pdvss)
  {
    if ($vssVM.appConsistentSnapshots)
  {
     $appConsistentVM += $vssVM.consistencyGroupName
    }
  }
$VssWriters = Get-VssWriters $appConsistentVM -Verbose |
  Where { $_.StateDesc -ne ‘Stable’ } | Sort “ComputerName”
$VssWriters | Out-GridView # Displays it in GridView
$VssWriters | Export-CSV “.\vssWriterReport.csv” -NoTypeInformation # Exports it to CSV
```

About the Author

Dwayne Lessner is a Principal Technical Marketing Engineer on the Product and Technical Marketing team at Nutanix. Follow Dwayne on Twitter at @dlink7.

Mike Umphreys is a Sr. Technical Marketing Engineer on the Product and Technical Marketing team at Nutanix.

About Nutanix

Nutanix makes infrastructure invisible, elevating IT to focus on the applications and services that power their business. The Nutanix enterprise cloud software leverages web-scale engineering and consumer-grade design to natively converge compute, virtualization, and storage into a resilient, software-defined solution with rich machine intelligence. The result is predictable performance, cloud-like infrastructure consumption, robust security, and seamless application mobility for a broad range of enterprise applications. Learn more at www.nutanix.com or follow us on Twitter @nutanix.

List of Figures

List of Tables