

AHV: A Virtualization Solution for the Enterprise Cloud

Nutanix White Paper

Copyright

Copyright 2018 Nutanix, Inc.

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

All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws.

Nutanix is a trademark of Nutanix, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

Contents

- 1. Overview of Nutanix and AHV..... 5**
- 2. Core Hypervisor Functions..... 7**
- 3. Manageability..... 10**
 - 3.1. Evaluating Manageability..... 12
- 4. Data Protection..... 14**
 - 4.1. Instant Recovery..... 14
 - 4.2. Backup and Recovery..... 16
 - 4.3. Evaluating Backup and Recovery..... 17
- 5. High Availability..... 18**
 - 5.1. Evaluating High Availability..... 22
- 6. Operational Insight..... 23**
- 7. Networking..... 24**
 - 7.1. Networking Overview..... 24
 - 7.2. Advanced Networking..... 24
- 8. Performance..... 26**
 - 8.1. Evaluating Performance..... 28
- 9. Conclusion..... 30**
- Appendix..... 32**
 - References..... 32
 - About Nutanix..... 32

List of Figures.....33
List of Tables..... 34



1. Overview of Nutanix and AHV

Nutanix delivers a hyperconverged infrastructure (HCI) solution that has pioneered the movement toward a cloud-like consumption model for the enterprise datacenter. The Nutanix solution delivers enterprise-class storage, compute, and virtualization in a single platform, bringing both the performance and the economic benefits of web-scale together in an architecture that is easy to manage and deploy.

When evaluating a virtualization solution and hypervisor, consider the same common set of criteria for the platform that you would for any IT infrastructure investment. The top functionality areas to think about include:

- **Manageability.** How easy is the system to manage? What system management tools are available? A system is only as good as it is simple to manage.
- **Data protection.** To be considered “enterprise ready,” any environment—virtual or physical—must provide features to protect a given solution and make it and its application data recoverable per disaster or business continuity plans.
- **High availability and fault tolerance.** Availability is imperative for business-critical applications. Solutions should have minimal single points of failure, tolerate hardware and software service disruptions, and be able to self-heal with little administrative action. Can the system stay online and active in the event of a planned or an unplanned interruption in system availability? Such a service interruption could relate to issues including the environment (for example, power fluctuations), hardware failure, a hypervisor availability problem, system upgrades, and so on. In HCI, where the hypervisor resides alongside the storage, what if a storage malfunction creates a service concern?
- **Operational insight.** At a glance, can you view information about potential troubles, whether you are nearing a capacity or resource runway limit, or how the cluster is performing? Is there a dashboard that presents predictive analysis to help the administrator? If so, can you customize it?
- **Performance.** Can the hypervisor adapt to changes in performance? Can it predict and analyze the workloads that are currently running? Can it move workloads around to remove hotspots in the environment?

These are just a few of the broad points to consider when determining whether a hyperconverged system meets your business needs. Beyond these higher-level categories, consider details like: how difficult the solution is to deploy, the overall cost (and what cost savings the solution might provide elsewhere), scalability, and support expectations.

Clearly, both the hypervisor and the virtualization industry predate Nutanix, but Nutanix is changing the conversation around these industry trends. As a complete solution, Nutanix has

taken a vendor-agnostic approach to virtualization. Out of the box, Nutanix supports the two industry-standard hypervisors, ESXi and Hyper-V, along with XenServer. However, Nutanix also offers a feature-rich hypervisor of its own—AHV. AHV is an enterprise-ready hypervisor based on proven open source technology.

Nutanix AHV delivers a full-featured experience, providing the most critical features you need to get up and running quickly, that is also uncomplicated. A virtualized environment is meant to abstract the hardware away from the operating system and to help you get the most out of your infrastructure. Realizing the full benefit of your infrastructure shouldn't come at the expense of navigating a labyrinth of little-used features or elaborate deployment options. Nutanix provides simplicity without compromise.

Table 1: Document Version History

Version Number	Published	Notes
1.0	December 2016	Original publication.
2.0	May 2018	Technology updates throughout.
2.1	September 2018	Updated backup vendors with integrated support for AHV.

2. Core Hypervisor Functions

AHV supports all the functions that an enterprise wants and needs when implementing a virtualized infrastructure. You can set up many of these features with only a few clicks.

- Virtual machine (VM) CRUD (create, read, update, delete).

With AHV, you can accomplish these tasks through a series of one-click, fill-in-the-blank processes. To create a VM, for example, use the AHV management interface to input a few details—VM name, description, vCPUs, cores per vCPU, memory, disks, and network adapters—and you're done.

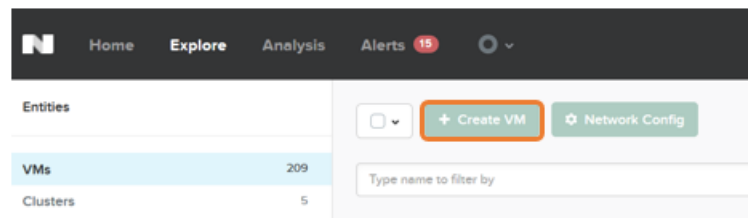


Figure 1: Create, Read, Update, or Delete VMs

- Clone.

Create a new space-efficient VM based on an existing VM. Additionally, with AHV you can create multiple clones through a single modification in the management interface.

General Configuration

NUMBER OF CLONES

PREFIX NAME

STARTING INDEX NUMBER

Example: atvm2-1-1, atvm2-1-2,...

Figure 2: Create Multiple VM Clones

- Power on, off, pause, and resume.

All the options for managing a VM's power state are available in AHV.

- Snapshots.

Create both crash-consistent and application-consistent snapshots.

- Migrate.

Allows a Nutanix AHV solution to be mobile and flexible, moving VMs across different AHV nodes in the cluster. A simple one-click operation initiates the process. The administrator can either select the destination manually or let AHV determine the optimal placement.

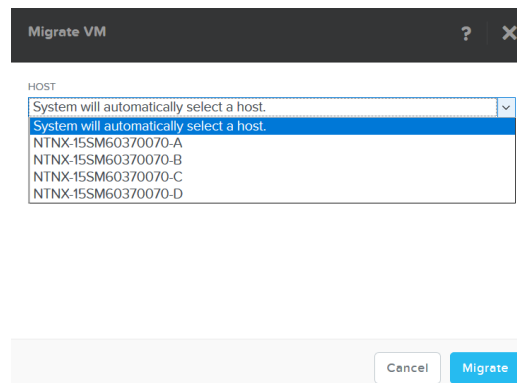


Figure 3: Migrate VMs

- VM guest tools.

The Nutanix Guest Tools (NGT) software package in AHV is similar to other industry hypervisor tool sets. Once enabled, NGT makes a CD-ROM available on the guest to install the guest tools.

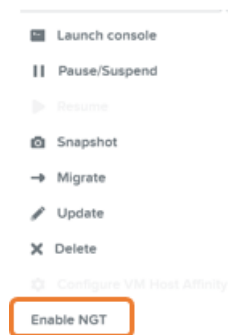


Figure 4: Enable Nutanix Guest Tools (NGT)

The NGT software package consists of the following:

- Nutanix Guest Agent Service: Communicates details about the running machine to the cluster.
- VM mobility drivers: Facilitate relocating a VM from ESXi to AHV and back.
- Self-service file restore: Restores a file within a user VM from a VM snapshot.

- Nutanix VSS agent and provider: Allows you to create application-consistent snapshots on a Windows VM, whether on AHV or ESXi.
- Support for application-consistent snapshots on Linux-based VMs through quiesce scripts.

When you combine the basic features above with the core functions of manageability, data protection, high availability, operational insight, and performance, AHV provides an enterprise-ready solution, capable of competing with the current industry veterans. Much of the Nutanix vision concerns simplifying how enterprises use and consume IT; this idea clearly carries through into AHV.

3. Manageability

The Nutanix AHV virtualization solution—including the tools you need to manage it—ships from the factory installed and ready to go, so you can have the system up and running in minutes once you've racked the cluster and powered it on. As soon as it's up and running, you can maintain the environment through a simple and modern HTML 5 web UI. Prism Element, which is available on each cluster you deploy, integrates this UI with the overall Nutanix solution. You can access Prism Element through each individual Nutanix cluster via the cluster IP or any of the individual Nutanix Controller Virtual Machine (CVM) IP addresses. Prism Element requires no additional software; it is built into every Nutanix cluster and incorporates support for AHV.

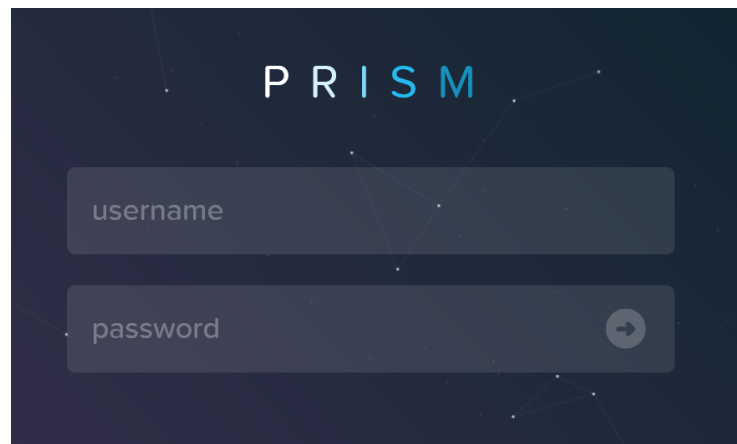


Figure 5: Prism Logon

If you prefer a more centralized mechanism for managing your deployment, Prism Central is available from the [Nutanix portal](#) or can be deployed directly from the Nutanix cluster. Prism Central is a robust optional software appliance VM that can run on ESXi, Hyper-V, or AHV.

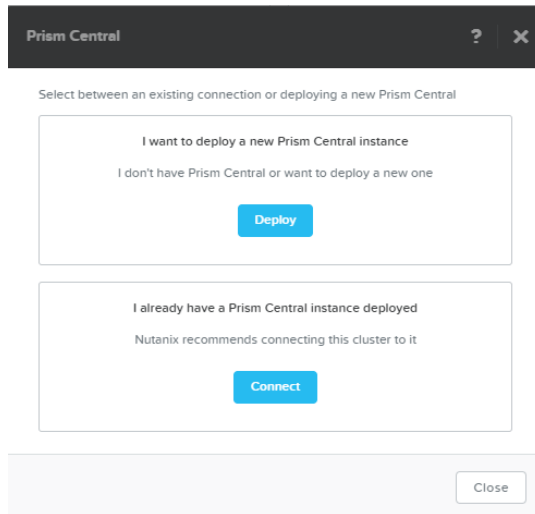


Figure 6: Deploy Prism Central from Cluster

Prism Central is both a platform and a hypervisor-agnostic management interface, providing an aggregate view of your deployed Nutanix clusters. In addition to allowing you to view and manage the cluster, Prism Central provides insight into VMs, hosts, disks, and containers or pooled disks, as shown in the screenshot below.


 Home Explore Analysis	
Entities	
VMs	136
Clusters	5
Hosts	20
Disks	191
Containers	36

Figure 7: Prism Central Management

Prism Central provides a single pane of glass for managing not only multiple Nutanix clusters, but also the native Nutanix hypervisor, AHV. Unlike other hypervisors, AHV requires no additional back-end applications or database to maintain the data rendered in the UI.

Deploying or establishing a connection to Prism Central is very simple:

1. Using a web browser, navigate to and log on to a node in a cluster where you want to establish a connection or deploy a new Prism Central instance.
2. Once logged on, click the appropriate option for your environment: **I want to deploy a new Prism Central instance** or **I already have a Prism Central instance deployed**.
3. Complete the wizard that begins after you select your preferred Prism Central deployment option.

For operational simplicity, backing up Prism Central requires capturing the VMs that comprise the Prism Central Cluster. If you are not a fan of GUIs, we also have a rich set of APIs and commands available via a REST API, PowerShell, or the Acropolis CLI (aCLI).

AHV management scales as the cluster grows. Within a Nutanix cluster, each node has a CVM that runs a service called Prism. The cluster elects a Prism leader to service all Prism Element requests. When a node (or CVM) receives a Prism request and is not the leader, it redirects the request to the current Prism leader. Consequently, should the current leader become unavailable, an election process commences to elect a new leader. This capability results in no single point of failure in the management layer for AHV.

Enterprises don't purchase a product just to spend time designing a management infrastructure for it. Our zero-dependency management stack allows customers to start consuming the Nutanix solution almost immediately. The tools and applications for managing your AHV virtualized environment are either built directly into the Nutanix cluster or, in the case of Prism Central, easily deployed through a management appliance VM available from Nutanix. You can use the same tools you use for Nutanix storage and compute to manage AHV. This capability eliminates the need to learn a different UI or to have a specialized resource to keep the management infrastructure up and running. Running an AHV virtualized datacenter with the built-in Nutanix administration options (Prism Element, REST API, or the aCLI) requires no additional management resources within your deployed Nutanix clusters.

3.1. Evaluating Manageability

When evaluating manageability, ask the following questions:

- How many systems do I need to manage the environment?
- Do I need additional third-party software to complete the management infrastructure?
- What do I need to back up in order to protect the management infrastructure?
- Are the management tools a single point of failure?
- How easy is it to recover from losing the management tools?
- Do we need additional expertise just to manage the management tools? (For example, would a database administrator need to assist with the repository holding the virtualization cluster details?)

Table 2: Evaluating Manageability

Question	Nutanix
How many systems do I need to manage the environment?	One: All management via Prism
Do I need additional third-party software to complete the management Infrastructure?	No: Prism provides management and monitoring
What do I need to back up in order to protect the management infrastructure?	—Prism Element: nothing —Prism Central: single VM
Are the management tools a single point of failure?	No
How easy is it to recover from losing the management tools?	Simple: —Prism Central: restore backup VM or deploy new Prism Central appliance —Prism Element: no action required; Prism service leader moves to other nodes running in the cluster
Do we need additional expertise just to manage the management tools?	No: Prism uses consumer-grade design, presenting a comprehensive yet intuitive interface

4. Data Protection

Enterprises must be able to recover from a loss of data, whether due to human error through accidental deletion, a partial loss of the infrastructure, or, as a worst-case scenario, a disaster causing a site disaster recovery (DR) plan activation. AHV covers these recovery situations with three main data protection categories:

- Instant recovery.

Instant recovery is the ability to expeditiously return a given VM or set of VMs to a previous time and state. AHV delivers this capability via the Prism management interfaces, REST API, or aCLI.

- Backup and recovery.

Backup and recovery generally means shipping your backed-up content to a location that is physically separate from the primary location running the VMs. AHV has resident backup and recovery, providing the flexibility to back up the data and ship it off to another location.

- Archiving.

Archiving typically involves moving the entire backup set (which might entail the “backup and recovery” dataset) to some type of deep cold storage. Enterprises usually manage archiving via third-party guest tools or integrations with third-party applications.

4.1. Instant Recovery

Nutanix provides instant recovery via two built-in methods: on-demand and protection domain recovery. On-demand data protection involves taking a snapshot for backup and recovery via the Prism Element or Prism Central UI.

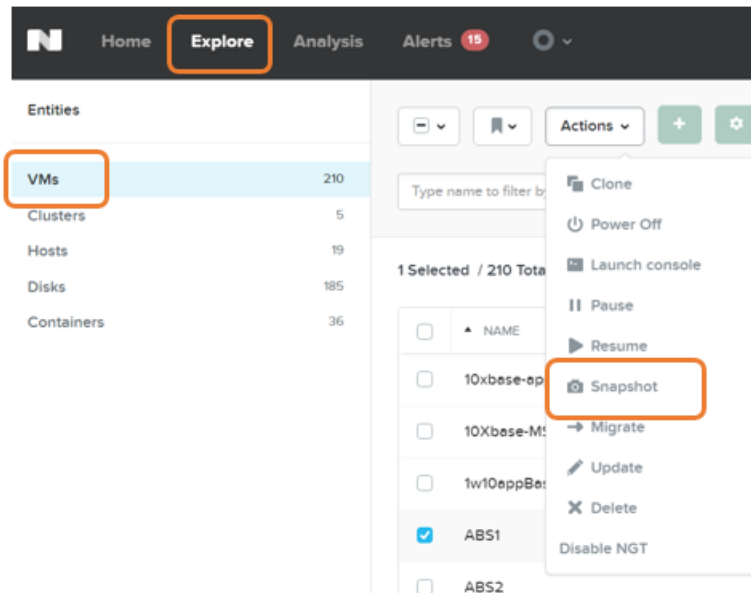


Figure 8: Taking a Snapshot: Prism Central UI

If working from the Prism Element UI, you can take a snapshot from the VM view. Select the VM and click **Take Snapshot**.

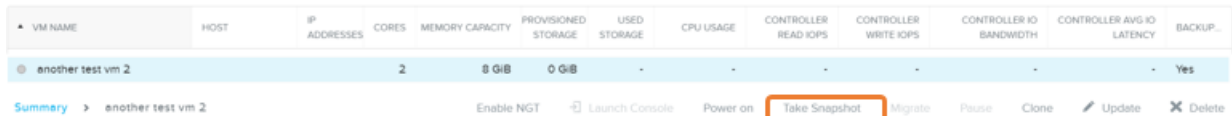


Figure 9: Taking a Snapshot: Prism Element UI

You can create both crash-consistent and application-consistent snapshots with AHV. To take an application-consistent snapshot, you must install NGT on the VM. NGT provides the ability to work with VSS in a Windows OS, as well as preefreeze and post-thaw scripts for Linux that can quiesce the guest prior to snapshot creation.

An administrator can create a protection domain from a VM or set of VMs. This grouping allows you to treat the VM or VMs as a single body with a single RPO. To provide further granularity and organization for VMs within a protection domain, you can create consistency groups. Using consistency groups lets you stipulate that a VM or set of VMs needs to be backed up at the same time. This provision is useful when you need to capture the entire state of a running application at the same time.

Other AHV backup and restore features include:

- Self-service restore, which allows trusted VM administrators to restore their own files on a per-VM basis.
- Recovery of a single VM within a snapshot.
- Clone a VM from a snapshot.
- Rapidly clone a VM using VAAI-like integration with the underlying Nutanix filesystem.

4.2. Backup and Recovery

Snapshots are only part of the picture. Snapshots are local recovery points, residing on the same media as the protected resources. They provide protection in scenarios that don't involve an unrecoverable loss of physical server resources. Traditional RAID architectures could lose an entire backup set if a disk resource failure resulted in unrecoverable data containers (volumes, RAID groups, and so on). A Nutanix solution uses data replication mechanisms instead of RAID, but this difference in data storage mechanics doesn't make snapshots a full solution on their own.

To complete the backup and recovery picture, we must ship data to a physically different location. This separation ensures that you're not storing recovery points in the same place as the protected VMs. AHV can leverage both native solutions and integrations with both cloud and third-party backup applications to provide a complete backup solution.

- Protection domains and instant recovery: Ships the backup data to another remote physical backup target. The backup target can be another Nutanix cluster or a cloud resource.
- Third-party backup application integration: Commvault can integrate with AHV to provide VM-level backup for a hypervisor-like experience, without the need to install the agents within the guest VMs.

With a protection domain and instant recovery, you can replicate a recovery point to a target location. The target can be a separate Nutanix cluster functioning as both a backup target and a DR target. As a recovery target for DR, the cluster can run the backup VMs it receives, restore the VMs to the original cluster, or, if needed, restore the VMs to a brand-new cluster. The target location can also simply serve as a backup target, with no intention of running the backup data it receives.

Nutanix also supports using public cloud resources as the target location for a protection domain. The cloud destination can be an Azure or AWS storage bucket. When pointing to a cloud storage bucket, the setup process creates a single-node Nutanix cluster in the configured cloud vendor. Nutanix replication technology copies the data to the cloud target as it would between two physical clusters. The cloud replication target option is a backup-only solution. You must restore the VMs to a Nutanix cluster in order to power on and use the VMs in a recovery or DR situation.

AHV integration with Commvault provides flexibility similar to that of using protection domains. The backup target can be either a Nutanix cluster holding recovery data, a Nutanix cluster holding recovery data along with running VMs, or a Commvault backup environment. The

integration supports a full life cycle of backup types: full, incremental, or differential. For more in-depth details regarding VM backup and recovery in AHV, see the Nutanix tech note [Data Protection for AHV-Based VMs](#).

The following backup vendors also have integrated support for AHV: HYCU from Comtrade, Rubrik, Veeam, and Veritas. The growing ecosystem around AHV shows that the major backup vendors see its value and are putting work into supporting this enterprise-ready hypervisor.

4.3. Evaluating Backup and Recovery

When evaluating backup and recovery, ask the following questions:

- Does the solution support external backup applications?
- Can you recover individual files?
- Does the solution support snapshots and DR replication?
- Are there APIs for data protection and backup?

Table 3: Evaluating Backup and Recovery

Question	Nutanix
Does the solution support external backup applications?	Yes
Can you recover individual files?	Yes
Does the solution support local snapshots and DR replication?	Yes
Are there APIs for data protection and backup?	Yes

5. High Availability

More apps running in virtualized environments and fewer physical servers mean that we need virtualization implementations to be highly available. In a hyperconverged infrastructure, having a highly available environment requires that we make storage, compute, and virtualization fault-tolerant. With AHV, the environment at the hypervisor level is fully redundant. High availability in virtualization ensures that VMs restart on another hypervisor node within a cluster of virtualization servers.

When a hypervisor node experiences an event that impacts service, running VMs go offline and surviving nodes need to restart those VMs. At creation, a Nutanix cluster defines the number of simultaneous failures it can tolerate before there is an impact on data. AHV uses this information when configuring VM high availability.

Out of the box, AHV enables high availability by default for VMs. The default mechanism is a best-effort recovery model. AHV attempts to recover and power on VMs on other cluster nodes based on the resources available on those nodes. Enabling this enterprise-class feature from the start is very effective, but the best-effort option can result in a situation where a VM does not power on because of a resource limitation.

For a more defined recovery model, AHV can reserve the necessary resources within the cluster to ensure VM high availability. Like the default best-effort feature, reserve mode is built-in and available as a one-click selection. Access the option to configure this capability through the Prism UI.

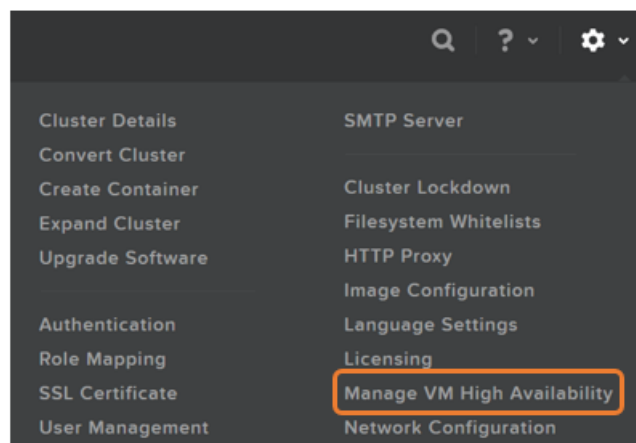


Figure 10: Manage VM High Availability

From this menu selection, **Enable HA** is a simple check box entry.

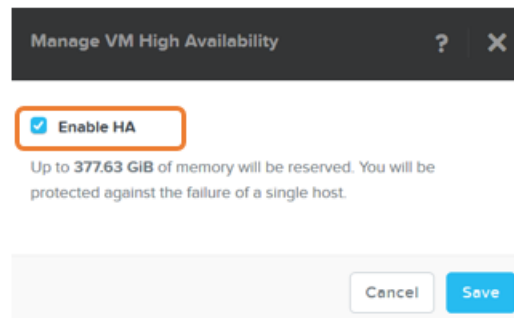


Figure 11: Enable VM High Availability

Once you have made this selection, AHV analyzes the Nutanix cluster and selects the optimal high availability policy for recovery. At this time, the system puts policies in place to ensure that any new VMs are recoverable based on the established high availability policy. AHV bases the policies on the available memory in the cluster and offers the following availability policy:

- **Reserved segments:** AHV spreads the high availability recovery points across the nodes in the cluster rather than dedicating a single node to them. When there is a high availability event, VMs running on a failed host restart on other nodes in the cluster.



Note: From Acropolis 5.0 onward, the default reserved VM high availability policy is reserved segments. The cluster selects this policy no matter how it is configured. Prior to Acropolis 5.0, AHV also offered a reserved host availability policy. For additional details, please consult the [AHV best practices guide](#).

In the reserved configuration, the VM high availability settings reserve resources across the cluster depending on the established Nutanix cluster fault tolerance level. The VM high availability configuration reserves the necessary resources to protect against the following:

1. One AHV host failure, if the Nutanix containers are configured with a replication factor of two.
2. Two AHV host failures, if any Nutanix container is configured with a replication factor of three.

Nutanix simplifies the selection of a high availability recovery method by automatically choosing the recovery option based on the details available from the cluster. Once you have enabled VM high availability beyond the best-effort default, newly created VMs can power on, as long as the policies can ensure that there are sufficient resources to accommodate high availability for the new VM.

Keep in mind that VM high availability configures the recoverability of a running VM only in hypervisor-based high availability events. A hypervisor-based event differs from high availability involving storage or CVM events within the Nutanix solution. Storage or CVM events involve recovery through data replication (based on the defined replication factor) and redirection to remote CVMs in the cluster. From the beginning, Nutanix has supported high availability events

that don't involve the hypervisor portion of the solution. For additional details on best practices for VM high availability, please refer to the Nutanix [AHV best practices guide](#).

In a hyperconverged infrastructure, it's possible for a functional hypervisor to run into an issue with connectivity to the storage layer. Nutanix provides a connection between the AHV host and storage through the Nutanix CVM. Each Nutanix cluster has an elected Acropolis master. The Acropolis master maintains a heartbeat communication with each AHV host in the cluster, sending health check messages once per second. If the Acropolis master does not receive a response from an AHV host over a period of four seconds, it begins evaluating it for possible failure. This polling takes place over a period of several seconds to account for an intermittent network issues that are only temporary and don't result in a complete outage. If there is an interruption in the heartbeat but the Acropolis master receives a response during the next polling interval, then the failure counter resets. This interruption could come in one of the following forms:

- AHV host fails while the Acropolis master is alive. This situation occurs when the Acropolis master can't maintain a network health check connection to the AHV host. If connection isn't reestablished within four seconds, the Acropolis master begins the process of confirming whether recovery steps are necessary. This confirmation process starts a timer. If the Acropolis master receives no further communication from the AHV host before the timer expires, it instructs all the remaining CVMs in the cluster to stop accepting I/O from the unresponsive AHV host. Once all the CVMs have acknowledged the request from the master, the process of restarting the VMs on the remaining available AHV hosts begins.
- AHV host becomes network partitioned while Acropolis master is alive. This situation is a lot like the previous one, but here the AHV host hasn't failed; instead, network connectivity to the Acropolis master has been interrupted. The Acropolis master can't be sure whether the AHV host is available or not. The VMs are still running, but in order to prevent multiple copies of VMs from running in the cluster, the system needs to begin remediation steps. As in the previous example, after a series of timers expire, the Acropolis master instructs all other CVMs to deny access to the isolated AHV host. Once the CVMs have acknowledged the request, the VM restart process begins on the remaining AHV hosts. At this point, the VMs on the partitioned node go into a suspended state, and the partitioned node initiates steps to power off the VMs that are now suspended.
- Acropolis master fails. This situation arises when the CVM experiences service interruptions due to events like: the CVM where the master is running fails, the AHV hypervisor hosting the Acropolis master fails, or the AHV host running the Acropolis master becomes network partitioned. In this situation, the subordinate Acropolis nodes begin an election process after a timer expires. Once a new Acropolis master is elected, the VM restart process begins, as in the other failure scenarios.

In any of the above interruption situations, once the cluster returns to a nondegraded state and the status is healthy, the process of restoring data locality begins. The system attempts to live

migrate the VMs that restarted on the surviving nodes back to the node that previously owned them. This migration occurs to promote data locality, placing the VM with the data it serves.

Taking data locality a step further, AHV communicates with the local CVM in order to interact with storage via iSCSI. AHV uses the CVM to provide data to other VMs. VMs are composed of one or more vDisks presented from AHV. AHV provides a connection to each vDisk via iSCSI, routed through the local CVM for data locality purposes. At the VM level in AHV, a vDisk is a distinct TCP connection to its own target and LUN. The VM sees this connection simply as a disk, much like what it would see if the connection were a physical machine. When the local CVM goes offline (for example, during an upgrade), the local AHV node works to reestablish a connection to another active CVM node in the cluster. AHV needs to be operational even when the CVM is undergoing planned or unplanned service interruptions.

AHV seamlessly stays online through interactions with the Stargate service. Stargate is the Nutanix data I/O manager, responsible for the interface between the hypervisor and data presented from the cluster. When the CVM goes down, AHV initiates new connections to other Stargate services running on other nodes in the cluster, with no impact to the running VMs. Each CVM in the cluster can accept requests for vDisk resources that originally connected through other nodes. This capability ensures that a single CVM doesn't receive all redirected requests for vDisks. The new connection process is efficient, promoting optimal placement to minimize the performance impact to any single cluster node. Once the local CVM node comes back online, AHV ensures that it's stable before moving its connection back over to the local CVM.

Outside of VM high availability, you should also consider the fault tolerance for resources like hypervisor VM networks. Nutanix and AHV provide protection at this level through Open vSwitch implementation. The Nutanix platform achieves networking high availability through the use of bridges (switches) and bonds (port groups) in AHV. These features are similar to what you would find in other hypervisor implementations. You can place the physical interfaces in a Nutanix AHV cluster into different high-performing and fault-tolerant arrangements to provide the needed level of serviceability for your enterprise:

- **Active-backup (default):** Each port in a bond has an active and a backup (idle) state. Each interface connects to separate physical switches for redundancy and fault tolerance. A given VM traverses the active port in the bond. If the active port fails, the system uses the backup.
- **Balance SLB:** Each port in a bond connects to separate physical switches, with the requirement that the switches be interconnected. Unlike in active-backup mode, both ports are active. A hashing algorithm that utilizes the source MAC address routes a given set of VMs through an interface in the bond.
- **Balance TCP:** Each port in a bond connects to separate switches, with the requirement that the ports be grouped together into an aggregation protocol (for example, LACP or similar). This setup allows a single VM to utilize both ports in the bond fully, based on the TCP port traffic flowing across the network.

The high availability and fault tolerance functionality is built into the Nutanix product with no or low configuration requirements. By providing this level of high availability out of the box, Nutanix is continuing along its path toward making the infrastructure as invisible as possible—simplifying without compromising.

5.1. Evaluating High Availability

When evaluating high availability, ask the following questions:

- Does the solution offer a no- or low-configuration high availability option?
- How easy is it to enable VM high availability?
- Can the hypervisor seamlessly recover from losing its connection to storage?
- Can the hypervisor efficiently spread workload across additional nodes upon storage service interruption?

Table 4: Evaluating High Availability

Question	Nutanix
Does the solution offer a no- or low-configuration high availability option?	Yes: out-of-the-box VM high availability
How easy is it to enable VM high availability?	Enabled by default; advanced VM high availability configuration is a one-click operation
Can the hypervisor seamlessly recover from losing its connection to storage?	Yes
Can the hypervisor efficiently spread workload across additional nodes upon storage service interruption?	Yes

6. Operational Insight

The Nutanix architecture can provide a customizable operational dashboard. This dashboard presents administrators with the details they need to quickly assess the overall health of their environment and is available via the Prism UI, both in the native Prism Element and in Prism Central. Using the same Nutanix UIs to monitor all aspects of the environment reduces operational complexity and minimizes the amount of time needed to recover from failures, provide details to management, and perform housekeeping.

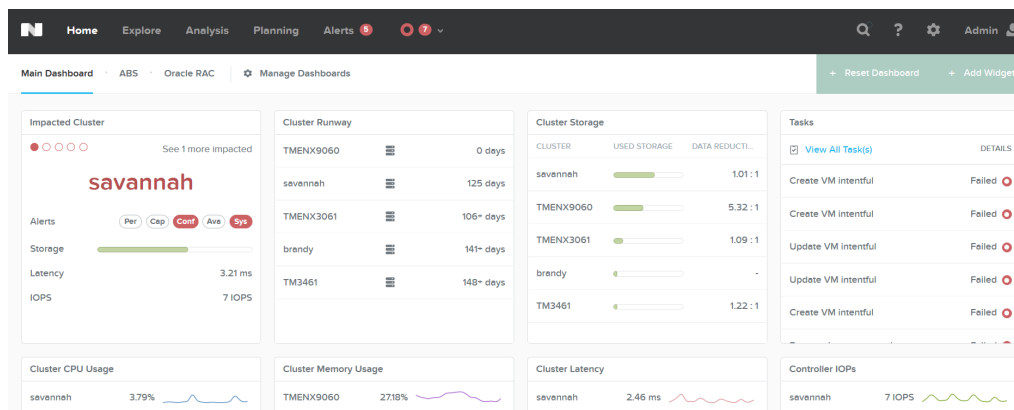


Figure 12: Operational Insights in Prism

The main dashboard view is composed of widgets. You can customize the screen to display the elements that matter most to your enterprise. The dashboard is a native, built-in feature, not only for AHV, but for the entire Nutanix architecture. You don't need to deploy any additional templates or applications in order to get insight into what is going on in the environment. Again, this approach offers simplicity without compromise, allowing the technical resources to provide the people resources with the right information at the right time.

7. Networking

7.1. Networking Overview

Open vSwitch (OVS) is incorporated into AHV to provide virtual networking services. OVS connects the CVM, hypervisor, and guest VMs to each other and to the physical network. Each AHV host in the cluster automatically starts an OVS service that is responsible for providing networking services.

Several components make up the networking stack within AHV. The list below offers a high-level overview of the components that comprise OVS networking. For more details on each component as it relates to AHV, please review the [AHV Networking best practices guide](#).

- **Open vSwitch:** A software switch that functions like a layer-2 learning switch, maintaining a MAC address table. Within a given AHV cluster, each host presents an OVS instance. Together, all of the OVS instances across the cluster present a single logical switch.
- **Bridge:** In this context, a bridge is another term for a virtual switch. A bridge manages the traffic that runs between the physical and virtual interfaces. The default configuration in AHV includes two bridges that are responsible for carrying traffic between the Nutanix CVM, the guest VMs, and the hypervisor. A given bridge can have multiple ports, and each port can have one or more interfaces.
- **Port:** A port is a logical entity within a bridge that allows for connectivity to a given virtual switch.
- **Bond:** A bond joins multiple physical interfaces to form a single entity. From this single entity, you can define three load-balancing modes: active-backup, balance-slb, and balance-tcp.

7.2. Advanced Networking

Securing the datacenter has traditionally focused on separating legitimate inbound traffic from external threats. Administrators typically achieve this security via hardware-based firewalls that contain a series of access control lists and rules with networking details like IP addresses and application port numbers. Once defined, the rules secure north and south activity, allowing or denying access to applications that reside within the datacenter. However, if the external boundary becomes exposed or is breached, the firewall offers little protection for movement from host to host across systems inside the datacenter. To secure this east-west traffic, IT departments often use additional hardware-based firewall devices to isolate portions of their network, repeating the same rules and access control list creation as needed.

Today's datacenters are highly virtualized, and their application environments can be very dynamic. Rapidly starting and stopping workloads places a burden on the traditional static enforcement mechanisms for security, so we need a more modern, scalable, and agile process. Because of this need, the industry is moving toward software-defined networking.

Nutanix AHV offers advanced networking and security services, providing visibility into the virtual network, application-centric protection from network threats, and automation of common networking operations. Fully integrated into Nutanix AHV virtualization and the Nutanix Enterprise Cloud OS, Nutanix features allow organizations to deploy software-defined virtual networking without the complexity of installing and managing additional products that have separate management and independent software maintenance requirements.

Nutanix AHV provides application-centric policies that enable complete visibility and traffic control. This policy model allows administrators to implement fine-grained rules regarding traffic sources and destinations, or microsegmentation. These same policies make it possible to visualize traffic flowing within and between application VMs. This granular level of control is an important part of a defense-in-depth strategy against modern datacenter threats.

Nutanix AHV networking protects against new threats that are designed to spread laterally from one system to another within the same protected datacenter. Because perimeter-based firewalls traditionally only protect the environment from external threats, it can be difficult to repurpose them to protect internal traffic. Nutanix AHV can apply security rules between all applications and VMs within the datacenter, thus adding internal protection behind your perimeter firewall.

8. Performance

This section isn't about the specific numbers of IOPS or whether the workload is random or sequential. All of those factors come into play and are important, but here we are talking about adaptive performance capabilities. In a virtualization environment, you have to take into account what workloads are running on the hypervisor, how those workloads impact the overall system, and what happens as the system grows over time. When building a virtualized infrastructure, the system first needs to be able to place workloads initially on nodes in the environment can best handle the workload. A virtualization environment is no different than a standalone server in that it has a finite number of worker resources, whether they are physical resources like memory and CPU or software-defined resources (threads, queues, and so on). A virtualized environment needs to share these resources across several different running workloads, making initial placement critical to creating a balanced environment from the start.

Over time, however, workloads change. They may change enough that the original placement is no longer optimal, so the hypervisor needs to adjust the environment. These adjustments typically involve moving VMs around to rebalance the environment and remedy hotspots. Nutanix AHV is equipped with the following features to handle these adjustments:

- **Dynamic scheduling (starting with the Acropolis 5.0 release):** An ongoing process monitors resources like CPU, memory, storage, and networking across the nodes in the cluster to ensure that the environment avoids hotspots and that VMs get the resources they need. The scheduler moves VMs to reduce hotspots in a way that does not violate any currently active policies or affinity rules. It makes recommendations based on details such as node CPU, memory, storage CPU (in other words, is the VM doing enough random I/O that it is utilizing the CVM resources more heavily), and the overall amount of local I/O to the node.
- **VM placement:** Balancing workloads up front during initial VM creation, without violating high availability rules. VM placement optimization occurs either when you create and power on a new VM, or when a VM powers off and on for any reason.
- **Affinity rules:** Defining which VMs can run on a particular set of hosts or defining what VMs shouldn't run together on the same host.

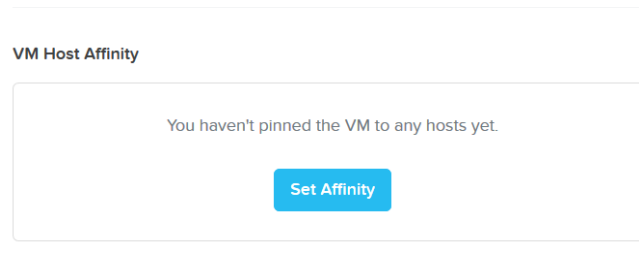


Figure 13: VM Host Affinity

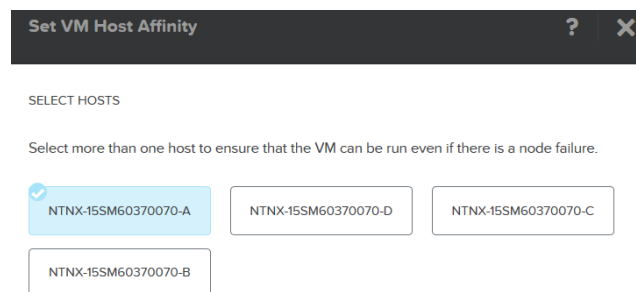


Figure 14: Set VM Host Affinity

In an AHV deployment, placement decisions take into account several metrics available from within the cluster. For initial placement, the AHV scheduler considers CPU and memory usage, looking for the least utilized node. As the system begins to process workloads, the dynamic scheduler determines continued residency for a VM.

The dynamic scheduler evaluates statistics available at both the compute and storage layers. At the compute layer, the environment takes into account CPU utilization, how much memory is used, and contention for resources. The analysis considers whether certain resources have exceeded a defined watermark. The same goes for storage performance. Remember that a hyperconverged infrastructure manages the compute, storage, and hypervisor within the same platform. There may not be any contention or thresholds exceeded for compute, but it is possible that, because of the amount of data a given workload needs, the Nutanix CVMs could be reaching a point of contention. Here is a brief look at some of the details:

- **Compute.** The environment monitors CPU across the nodes. For example, if CPU exceeds a threshold of 85 percent (its default setting here), then the system considers VMs for migration to another node in the cluster.
- **Storage.** The Nutanix storage layer works to ensure that a given VM's data is local to it. This design somewhat negates the need for storage DRS in the typical sense. It doesn't, however, necessarily remove the need to consider moving a given VM's data to another node in the cluster. Dynamic scheduling analyzes each node in the cluster for contention situations. If a given Stargate process exceeds a similar CPU threshold on a given node, then, as in the

compute layer, the system considers it for resource reallocation. Reallocating resources would help balance the storage layer across the cluster.



Note: The above reallocation occurs only when the cluster experiences contention or hotspots. It is not a workload balancing evaluation process to ensure even distribution of resources. Workload balancing says, “I have a four-node cluster. One node is at 50 percent and three nodes are at 10 percent. Rebalance to redistribute the workloads more evenly.” In this scenario, there are no real hotspots, nothing is contending for resources, and the spike on one node may prove to be a temporary situation that resolves itself. Moving VMs can be an expensive operation, so a system should only reallocate them when there is actual contention for a resource. AHV only moves those VMs that are absolutely necessary to mitigate a current hotspot.

The Prism UI logs and reports VM movement. When the system cannot mitigate a hotspot, Prism logs the result of the analysis and generates an alert. Dynamic scheduling aims to resolve all the hotspots the system encounters. However, if moving a VM would violate an affinity rule, the reallocation does not proceed.

AHV continuously evaluates the environment’s performance to determine how best to position VMs. This intelligent placement feature is native to AHV and enabled by default. With this level of optimization, enterprise and IT staff can spend less time focusing on where to place resources and more time building the next generation of their business. When systems become unbalanced and “hot,” the enterprise has to find the imbalance or work around it. With AHV and Nutanix, the goal is to let the enterprise view the environment the same way from one node to the next, with balancing occurring on its own.

8.1. Evaluating Performance

When evaluating performance, ask yourself the following questions:

- Is adaptive performance built into the hypervisor?
- Can I enable adaptive performance with minimal or no configuration changes?
- Does the system preserve affinity and DRS policies?

Table 5: Evaluating Performance

Question	Nutanix
Is adaptive performance built into the hypervisor?	Yes

Question	Nutanix
Can I enable adaptive performance with minimal or no configuration changes?	Yes
Does the system preserve affinity and DRS policies?	Yes

9. Conclusion

Nutanix wants your virtualization solution to streamline moving your enterprise forward. We work to abstract away the complexity in deploying and administering a virtualized solution, while providing the flexibility needed in the modern datacenter. The table below lists common industry-standard hypervisor features and how the AHV feature set fits in.

Table 6: Hypervisor Feature Comparison

Feature	AHV
Management Server / Application	Yes: Prism Element (built in); Prism Central (downloadable VM)
VM vMotion	Yes: Live Migration
Virtual Machine High Availability	Yes: enabled by default as best effort; single-click option to enable resource reservation and ensure VM high availability
Application High Availability	No Note: A similar VMware feature was deprecated and was EOA starting in vSphere 6.0.
Data Protection	Yes: local snapshots and DR asynchronous replication to another Nutanix cluster or cloud repository
Hypervisor Replication	Yes: integration with the Acropolis Distributed Storage Fabric (DSF) and protection domains that can replicate VM snapshots to remote locations
Secure Configurations	Yes: self-healing STIG
Software-Defined Storage Integration	Yes: works with the DSF
VMs (or similar)	Not really applicable. Integration between AHV and the DSF negates the need for this feature.
Hypervisor DRS	Yes: initial placement decision process and ongoing hotspot evaluation

Feature	AHV
Storage-Level DRS	Not really applicable. Integration between AHV and the DSF mitigates the benefits of storage DRS. The Acropolis distributed file system offers intelligent data tiering and with dynamic scheduling from Acropolis 5.0 onward. The system can manage hotspots and remediate resource contention as necessary.
Software Update Management	Yes: native to Prism UI and offered as one-click. Upgrades are also available for Acropolis and firmware. No additional software management tools are necessary.
Storage API / Array Integration	Not really applicable. Integration between AHV and the DSF offloads functions normally associated with this feature from the hypervisor (for example, snapshots, cloning, or vDisk provisioning).
Affinity or Antiaffinity Rules	Yes: starting in Acropolis 5.0
CPU Compatibility (or EVC)	Yes: always on, with nothing to configure
Microsegmentation	Yes
GPU Support	Yes: both passthrough and vGPU

Nutanix has designed AHV to be straightforward to deploy and manage. By providing the most commonly needed features to run your enterprise, AHV turns virtualization decisions into a simple checklist. Even better, AHV delivers these features at no additional hypervisor cost—if you have a Nutanix cluster then you already have AHV, giving you simplicity without complication or compromise.

Appendix

References

1. [Nutanix Data Protection for AHV-Based VMs](#)
2. [Nutanix AHV Best Practices](#)
3. [Nutanix AHV Networking](#)

About Nutanix

Nutanix makes infrastructure invisible, elevating IT to focus on the applications and services that power their business. The Nutanix Enterprise Cloud OS 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](https://twitter.com/nutanix).

List of Figures

Figure 1: Create, Read, Update, or Delete VMs.....	7
Figure 2: Create Multiple VM Clones.....	7
Figure 3: Migrate VMs.....	8
Figure 4: Enable Nutanix Guest Tools (NGT).....	8
Figure 5: Prism Logon.....	10
Figure 6: Deploy Prism Central from Cluster.....	11
Figure 7: Prism Central Management.....	11
Figure 8: Taking a Snapshot: Prism Central UI.....	15
Figure 9: Taking a Snapshot: Prism Element UI.....	15
Figure 10: Manage VM High Availability.....	18
Figure 11: Enable VM High Availability.....	19
Figure 12: Operational Insights in Prism.....	23
Figure 13: VM Host Affinity.....	27
Figure 14: Set VM Host Affinity.....	27

List of Tables

Table 1: Document Version History.....	6
Table 2: Evaluating Manageability.....	13
Table 3: Evaluating Backup and Recovery.....	17
Table 4: Evaluating High Availability.....	22
Table 5: Evaluating Performance.....	28
Table 6: Hypervisor Feature Comparison.....	30