



How Nutanix Works

The Definitive Guide to Hyperconverged Infrastructure

NUTANIXTM

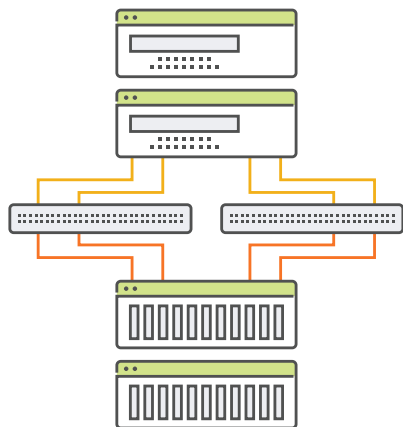
Table of Contents

IT at a crossroads	
Time for a different approach?	7
What is Hyperconverged Infrastructure?	7
The Nutanix Solution	
Nutanix community edition and community edition on-demand	9
The Nutanix Cloud Platform	9
How Nutanix software is deployed	10
Nutanix leads the pack	10
Nutanix Cloud Infrastructure (NCI)	
Infrastructure resilience	13
Tunable redundancy	14
Replication factor versus RAID	14
Data path redundancy	14
Nutanix software upgrades and data path redundancy	15
Integrity checks	15
Availability domains	15
Performance acceleration	
Intelligent tiering	16
Data locality	16
Automatic disk balancing	17
Integrated virtualization	17
Shadow clones	18
Capacity optimization	
Deduplication	20
Compression	20
Pro tip: Compression	21
EC-X	21
Data protection	
What is RTO and RPO?	22
Converged local backups with snapshots and time stream	23
Integrated remote backup and DR using async replication	24
Self-service file restore	24
NearSync	24
Metro availability and sync replication	25
Security	
Data-at-rest encryption	26
Two-factor authentication	27
Secure access	27
Security development lifecycle	27
Security configuration baseline	28
Security automation	28
Freedom to virtualize	
Full-stack industry-standard support	30
AHV	31
Enterprise ready virtualization solution	32
AHV data protection	32
AHV networking	33
Flow network security - network virtualization and security	33
Virtual machine management	34
VM operations	34
Image management	35
AHV VM placement and resource scheduling	36
AOS Dynamic Scheduling	36
Affinity and Anti-affinity	36
Live migration	37
Cross-hypervisor migration	37
Automated high availability	38
High availability out of the box	38
Converged backup and disaster recovery	39
Backup APIs	39
Analytics	39
Performance	41
AHV Turbo	42
vNUMA	42
RDMA	43
GPU support	43
Unify clouds with NCM and Prism	
Prism is highly available by design	44
The Prism approach	45
Software upgrades	47
Pro Tip: Prism Central	48
Nutanix Self-Service and Orchestration	49
About Nutanix	50

IT at a crossroads

IT is increasingly being asked to spend less time on infrastructure and more time (and budget) on application services that add business value. Despite a continuous stream of IT hardware and software enhancements, the infrastructure challenges faced by IT teams continue to rise. The IT infrastructure and virtualization software required to meet the needs of business is complex and expensive, and datacenter management has become painful. Far too much time and effort are focused on just keeping the lights on.

Legacy infrastructure—with separate storage, storage networks, and servers—is not well suited to meet the growing demands of enterprise applications or the fast pace of modern business. The silos created by traditional infrastructure have become a barrier to change and progress, adding complexity to every step, from ordering to deployment to management. New business initiatives require buy-in from multiple teams, and IT needs must be predicted 3-5 years in advance. As most IT teams know, this involves a substantial amount of guesswork and is almost impossible to get right. In addition, vendor lock-in and increasing licensing costs are stretching budgets to the breaking point.



1. **INHERENT COMPLEXITY**
2. **INEFFICIENT SILOS**
3. **FORKLIFT SCALING**
4. **PAINFUL MANAGEMENT**

Figure 1. Challenges of legacy three-tier infrastructure.

Time for a different approach?

Enterprise IT teams today are looking for ways to deliver on-premises IT services with the speed and operational efficiency of public cloud services such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud.

Taking cues from web giants, hyperconverged infrastructure (HCI) combines common datacenter server hardware using locally attached storage devices (spinning disk or flash) with intelligent software to eliminate common pain points associated with legacy infrastructure.

Nutanix delivers a comprehensive hybrid cloud platform that bridges the wide gap between traditional infrastructure and public cloud services. The solution delivers turnkey infrastructure that integrates servers, storage, and virtualization along with end-to-end systems management and operations management capabilities. This allows enterprises to deploy infrastructure in minutes and shift the focus to applications that power the business.

What is hyperconverged infrastructure?

Hyperconverged infrastructure combines common datacenter hardware using locally attached storage resources with intelligent software to create flexible building blocks that replace legacy infrastructure consisting of separate servers, storage networks, and storage arrays.

Hyperconvergence is the fundamental building block for hybrid cloud. This book gives an overview of the Nutanix hyperconverged solution and walks through how different features and functionality provide a fast, highly scalable, and efficient datacenter solution for enterprises of all sizes.

The Nutanix solution

Nutanix converges the entire datacenter stack, including compute, storage, storage networking, and virtualization. Complex and expensive legacy infrastructure is replaced by the Nutanix Cloud Platform (NCP) running on state-of-the-art, industry-standard servers that enable enterprises to start small and scale one node at a time. Each server, also known as a node, includes Intel-powered x86 hardware with flash SSDs and HDDs. Nutanix software running on each server node distributes all operating functions across the cluster for superior performance and resilience.

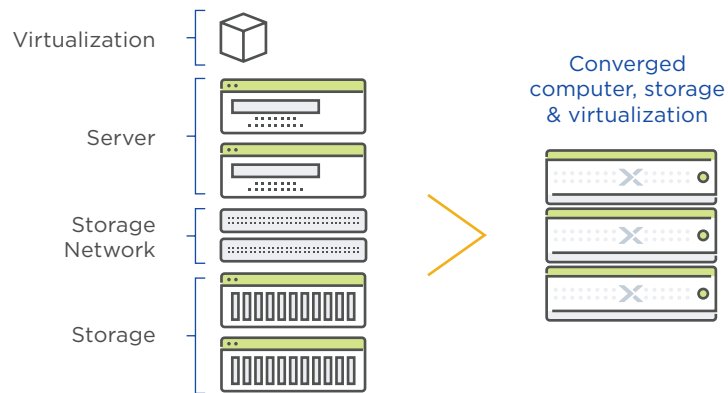


Figure 2. Nutanix converges compute, storage, and virtualization in simple, scalable building blocks.

A single Nutanix cluster can scale as large as the hypervisor cluster it is on. Different hardware platforms are available to address varying workload needs for compute and storage. Nutanix software is hardware agnostic, running on hardware from vendors such as Dell, Lenovo, Fujitsu XF, and HPE ProLiant, and from Nutanix as well.

NUTANIX community edition and community edition on-demand

[Community Edition](#) is a free, 100 percent software solution that lets enterprises easily evaluate the latest Nutanix technology at zero cost on existing hardware or via a [Nutanix Test Drive](#).

Hardware platform configurations are available to fit any workload by independently scaling the various resources (CPU, RAM, or Storage) and can be provisioned with or without GPU for graphics acceleration. All nodes include flash to optimize storage performance, and all-flash nodes are available to deliver maximum I/O throughput with minimum latency for all enterprise applications.

The Nutanix Cloud Platform

Most HCI solutions consist of two fundamental components: a data plane and a management plane. The Nutanix Cloud Infrastructure (NCI) and Nutanix Cloud Manager (NCM) are responsible for data plane and management duties within the NCP.

Nutanix Cloud Infrastructure (NCI) distributed data plane for either VMs or container-based applications that runs across a cluster of nodes delivering enterprise storage and virtualization services.

Nutanix Cloud Manager (NCM) is the distributed management plane that uses advanced data analytics and heuristics to simplify and streamline common workflows, eliminating the need for separate management solutions for servers, storage networks, storage, and virtualization. This suite of features enables application and business requirements to be met within the hyperconverged infrastructure, with no need to rely on external services. In other words, you can build a complete datacenter on HCI within the Nutanix Cloud Platform.

How Nutanix software is deployed

A Nutanix cluster is 100 percent software defined. Each node in a cluster runs a hypervisor (VMware ESXi, Microsoft Hyper-V, or the native Nutanix hypervisor, AHV), and the Nutanix software runs as a virtual machine called the Controller VM (CVM) that runs on every node in the cluster. The CVM includes management and data plane functions.

“ “

My key requirements were to have something that was simple, easy to manage, and ideally a single pane of glass. I wanted a solution that was very powerful and also very versatile. For me, Nutanix ticked all of those boxes.”

PURDIP BAHRA

IT Manager, Joseph Chamberlain College

Nutanix Cloud Infrastructure (NCI)

NCI is the underpinning of an HCI solution that transforms HCI into an enterprise cloud platform. The main components are:

AOS Storage

The AOS Storage fabric simplifies storage and data management for virtual environments. By pooling flash and hard disk drive storage across a Nutanix cluster and exporting it as a data store to the virtualization layer as iSCSI, NFS, and SMB shares, AOS Storage eliminates the need for SAN and NAS solutions.

- Enterprise storage services for applications, eliminating the need for separate solutions from vendors such as NetApp, EMC, and HP
- Includes a comprehensive set of capabilities for performance acceleration, data reduction, data protection, and much more
- Full support for VMware® vSphere, Microsoft® Hyper-V, and Nutanix AHV

Nutanix AHV Virtualization

- Comprehensive virtualization solution included with Nutanix at no additional charge
- Hardened to meet the most stringent enterprise security requirements
- Integrated virtual machine management through Prism
- Intelligent virtual machine (VM) placement, live migration, hypervisor conversion, and cross-hypervisor high availability for maximum flexibility

Nutanix Unified Storage

Nutanix Unified Storage enables you to easily manage and control your enterprise data no matter where it resides — separate application, storage, and geographic silos.

- Files Storage services provides access to Microsoft Windows via SMB 2.1 and to Linux and Unix via the NFS v4 protocol. This solution scales and load balances multiple nodes in the cluster, growing capacity and performance as needed.
- Objects Storage services is a software-defined object storage solution that nondisruptively scales out while lowering overall costs. It supports an industry-standard S3-compatible REST API to handle petabytes of unstructured data.
- Volumes Block Storage services provides iSCSI access to applications that require direct access to block storage. This can be nonvirtualized systems, or virtual machines with specific requirements. Volumes utilizes AOS Storage to scale I/O across the entire cluster and can load balance and accelerate specified Volume Groups.
- Objects Storage provides a simple and scalable S3 compatible storage tier on top of your Nutanix cloud platform.

Nutanix Flow Network Security and Virtual Networking

- Flow Network Security (Microsegmentation) provides a proactive and adaptive approach to VM-to-VM network security
- Flow Virtual Networking provides the ability to create virtual private clouds (VPCs) to simplify network configurations and ensure proper network connectivity is maintained without the time-consuming manual configuration of physical networks, routing, or IP address assignment.
- Nutanix Flow service chains that enable network feature virtualization
- Network visualization and optimization
- Network automation and orchestration

Infrastructure resilience

The Nutanix platform is designed to be fault tolerant, with no single points of failure or performance bottlenecks.

Tunable redundancy

With tunable redundancy, each Nutanix container is configured with a replication factor (RF) of two or three. RF=2 ensures that two copies of data are always maintained, allowing the cluster to survive the failure of a single node or drive. When RF is set to 3 (RF=3), three copies of the data are maintained in a cluster, providing resilience against two simultaneous failures. RF=2 is considered the best practice in most scenarios.

Replication factor versus RAID

RAID has been a popular way of protecting against drive failures while limiting the extra storage capacity required. Rebuilding a multi-TB drive can take days to complete, creating a risk of data loss should further failures occur. RAID has gone from single to double and even triple-parity to try to reduce this risk.

Nutanix Replication Factor (RF) eliminates reliance on RAID, the need for expensive spare drives that sit idle, and the performance penalty that comes with multiple parity calculations.

Data path redundancy

Data path redundancy ensures high availability in the event a Nutanix Controller VM (CVM) becomes unavailable or needs to be brought down for upgrade. If a CVM becomes unavailable for any reason, Nutanix CVM autopathing automatically reroutes requests to a “healthy” CVM on another node. This failover is fully transparent to the hypervisor and applications.

Data path redundancy is possible because every node in a cluster has access to all copies of data—I/O requests can be serviced immediately by any node in the system.

Nutanix software upgrades and data path redundancy

Nutanix software upgrades take advantage of reliable data path redundancy. While the local CVM is unavailable because of software upgrade or a failure, VMs running on the node use data path redundancy to satisfy I/O through a CVM on another node—transparent to users and applications.

Integrity checks

NCI has a variety of features to proactively identify and fix issues related to data consistency and integrity.

- Detection of silent data corruption and repair of data consistency errors
- Automatic data integrity checks during every read
- Automatic isolation and recovery during drive failures

Availability domains

Availability domains offer greater protection from hardware failures by allowing Nutanix clusters to survive the failure of a node or block (multinode chassis) or datacenter rack. Availability domains are created based on the granularity at which failures are likely to occur.

With AOS Storage, data replicas are written to other blocks in the cluster to ensure that in the case of a block failure or planned downtime, the data remains available. This is true for both RF2 and RF3 scenarios, as well as in the case of a block failure. An easy comparison is “node awareness,” where a replica would need to be replicated to another node that will provide protection in the case of a node failure. Block and rack awareness further enhance this by providing data availability assurances in the case of block-or rack-level outages.

Performance acceleration

AOS Storage includes several capabilities that enhance performance:

Intelligent tiering

AOS Storage continually monitors data access patterns and optimizes data placement on either the SSD or HDD tier, achieving the best performance without administrator intervention. Nutanix also supports an all-flash configuration using an all-SSD tier.

The SSD tier provides maximum performance for hot data and random I/O, while the HDD tier provides maximum capacity and economy for cold data and sequential I/O.

Data locality

AOS Storage ensures that as much of a VM's data as possible is stored on the node where the VM is running. This negates the need for read I/O to go through the network. Keeping data local optimizes performance and minimizes network congestion.

Every VM's data is served locally from the CVM and stored preferentially on local storage. When a VM is moved from one node to another using vMotion or live migration (or during an HA event), the migrated VM's data automatically follows the VM in the background based on read patterns.

Automatic disk balancing

Automatic disk balancing ensures that data is distributed uniformly across the entire cluster. Any node in a Nutanix cluster can utilize storage resources across the cluster, without requiring time-consuming and error-prone manual rebalancing.

Automatic disk balancing reacts to changing workloads and allows heterogeneous nodes to be mixed in a single cluster. Once utilization reaches a set threshold, disk balancing keeps it uniform among nodes.

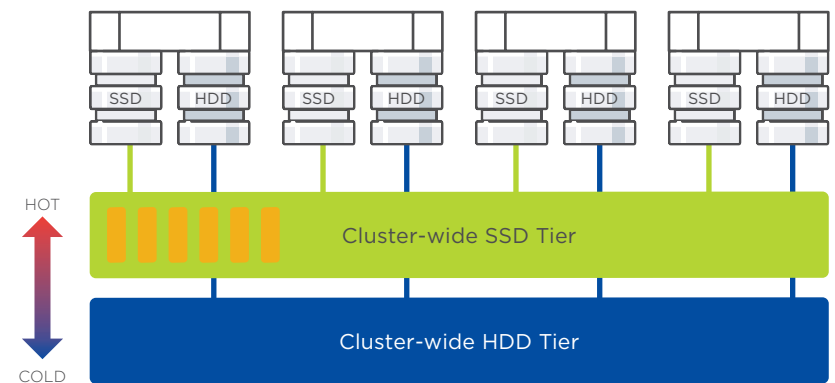


Figure 4. VM Flash Mode allows individual vDisks to be "pinned" in the cluster-wide SSD tier for maximum IOPS and low latency.

Integrated virtualization

Nutanix AHV is a virtualization solution designed to work perfectly with the entire Nutanix platform, which means there is no need for additional configuration or management to get the best application performance. AHV also includes AHV Turbo, which is a hypervisor I/O enhancement built to take advantage of next generation, ultra-low latency storage devices like NVMe. With AHV turbo and data locality, the most demanding applications are guaranteed the performance they need.

Shadow clones

Shadow clones significantly improve performance by caching virtual machine data across a Nutanix cluster. Unique to Nutanix, shadow clones benefit scenarios where there are multiple VMs reading a single source of data, such as deployment servers and repositories. VDI deployments, where many linked clones forward read requests to a central master (such as Citrix MCS Master VM or VMware View replica disks), are an ideal example.

With shadow clones, Nutanix actively monitors vDisk access trends. If there are requests originating from more than two remote Controller VMs (CVMs), as well as the local CVM, and all of the requests are read I/O, the vDisk will be marked as immutable. Once the disk has been marked immutable, the vDisk is then cached locally by each CVM so read operations are now satisfied locally by direct-attached storage resources.

Capacity optimization

AOS Storage incorporates a wide range of storage optimization technologies that work together to make efficient use of the available capacity in a cluster.

Deduplication

Nutanix offers two types of data deduplication to accelerate application performance and optimize storage capacity. Performance-tier deduplication removes duplicate data in-line with the content cache (SSD and memory) to reduce the footprint of an application's working set. In addition, global post-process MapReduce deduplication reduces repetitive data in the capacity tier to increase the effective storage capacity of a cluster. Both forms of deduplication can be easily configured and managed virtual machine and vDisk granularity.

When deduplication is enabled, data is fingerprinted on ingest using a SHA-1 hash. Deduplication operations are software-driven and leverage the hardware-assist capabilities of the Intel chipset for the SHA-1 fingerprint generation. Because SHA-1 is a strong hash, deduplication is performed based on a fingerprint match.

Compression

Data can be compressed inline as it is written to the system, or post-process after the data has been written. Inline and post-process compression is intelligently determined based on sequential or patterns to enable optimal performance. AOS Storage executes post-process compression as a series of distributed MapReduce jobs.

Pro tip: Compression

Use inline compression most of the time; it will not impact random write performance. Inline compression pairs perfectly with erasure coding.

EC-X

Nutanix systems include an innovative implementation of erasure coding technology—Nutanix EC-X—that provides resilience and can increase usable capacity by up to 75 percent. EC-X reduces the capacity cost of replication factor (RF) without taking away any of the resilience benefits and with no impact on write performance.

EC-X encodes a strip of data blocks on different nodes and calculates parity. In the event of a disk or node failure, parity is used to calculate any missing data blocks. AOS Storage uses an extent group as the data block, and each data block in a strip must be on a different node and belong to a different vDisk. The number of data and parity blocks in a strip is configured based on the desired number of failures to withstand.

Data protection

Nutanix offers natively integrated data protection and continuous availability at the VM level. A range of options is available to meet the recovery point objective (RPO) and recovery time objective (RTO) of different applications.

RPO	RTO	Nutanix feature	
Minutes	Minutes	Time Stream	Minor Incidents
Hours	Hours	Cloud Connect	
Minutes	Minutes	Async or Near sync Replication	Major Incidents
Zero	Zero	Sync Replication	

Figure 5. Nutanix data protection options.

What is RTO and RPO?

Recovery Time Objective (RTO) defines how much time you have to recover if an IT failure occurs.

Recovery Point Objective (RPO) defines the maximum amount of data you are willing to lose.

Converged local backups with snapshots

Nutanix can create unlimited metadata-based local snapshots—with VM and application-level consistency—and recover data instantly to meet a wide range of backup and data protection requirements. Metadata based snapshots require minimal disk overhead and ensure high performance recovery.

AOS Storage uses VM-centric snapshots to provide production-level data protection without sacrificing performance. Nutanix utilizes a redirect-on-write algorithm that dramatically improves system efficiency for snapshots.

Many backup vendors combine these capabilities with enterprise storage features from Nutanix. Check our Nutanix Technology Alliance Partners program on www.nutanix.com for a list of our backup partners.

Integrated remote backup and disaster recovery using async replication

Nutanix disaster recovery (DR) and replication capabilities are built on snapshot technology. VM snapshots can be asynchronously replicated or backed up to another datacenter based on a user-defined schedule.

Replication topologies are flexible and bi-directional, enabling one-to-one, one-to-many, and many-to-many deployments. During replication, data is compressed and replicated at the sub-block level for maximum efficiency and lower WAN bandwidth consumption.

Nutanix Cloud Manager interface offers a simplified view of all local and remote snapshots, allowing administrators to restore a VM from a snapshot with a single click. In case of disaster, administrators can also failover a VM to the secondary datacenter with a single click.

Self-service file restore

NCI data protection includes self-service file restore, which allows VM and application owners to recover individual files from VM snapshots without getting an administrator involved.

NearSync

NearSync replication allows for RPOs as low as one minute for protecting your mission-critical applications. By leveraging Nutanix light weight snapshots (LWS), NearSync supports more granular restore capabilities while using the existing DR workflows in NCM (Prism). vSphere and AHV support NearSync and there are no latency or distance restrictions.

Metro availability and sync replication

For critical workloads requiring zero RPO, and near-zero RTO, Nutanix offers metro availability, which ensures continuous data availability across separate sites within a metro. With Nutanix, setting up and managing this feature is simple.

Administrators can set up metro availability bi-directionally between two sites connected over a metro area network. The only network requirement is a round-trip latency of less than five milliseconds. Data is written synchronously to both sites, so it is always available to applications in the event a site fails or needs to undergo maintenance. Virtual machines can be non-disruptively migrated between sites for planned maintenance events or other needs.



We were focused on flexibility and innovation. We were looking for a partner who would be able to understand our business needs. With Nutanix, there was a willingness to listen and propose an innovative solution.”

LAURENT PERRIAULT
Director of Operations, Claranet

Security

NCP is hardened by default, leveraging features from NCI, then managed through NCM. Nutanix utilizes the principle of least privilege, and delivers a true defense-in-depth model with turn-key security policies and controls across features like two-factor authentication and data-at-rest encryption. At the Nutanix core, a security development lifecycle is used across every feature and integrated into product development to help meet the most stringent security requirements.

Nutanix helps achieve deeper alignment to critical frameworks and standards like the zero trust security model or the CISA Cybersecurity Framework to improve Cyber Resilience and are certified across a broad set of evaluation programs to ensure compliance with the strictest standards. Its custom security baseline exceeds the requirements of the U.S. Department of Defense.



Figure 6. Acropolis meets and surpasses some of the most stringent security requirements.

Data-at-rest encryption

Data-at-rest encryption is delivered through self-encrypting drives (SED) that are factory-installed in Nutanix hardware. This provides strong data protection by encrypting user and application data for FIPS 140-2 Level 2 compliance. For SED drives, key management servers are accessed via an interface using the industry-standard Key Management Interface Protocol (KMIP) instead of storing the keys in the cluster.

Nutanix also provides the option to use a native data-at-rest encryption feature that doesn't require specialized hardware from self-encrypting drives (SED). This feature can optionally leverage a built-in, local key management solution, reducing the complexity of deploying a secured environment.

Two-factor authentication

Nutanix solutions support SAML integration and optional two-factor authentication for system administrators in environments requiring additional layers of security. When implemented, administrator logins require a combination of a client certificate and username and password.

Secure access

Nutanix also offers a higher security configuration option called Cluster Lockdown, which restricts access to a Nutanix cluster in security-conscious environments such as government and healthcare datacenters. Cluster lockdown not only disables interactive shell logins automatically but can also enable more restrictive access based on cryptographic keys.

Security development lifecycle

Nutanix uses a unique, well-defined Security Development Lifecycle (SecDL) to incorporate security into every step of the software development process, from design and development to testing and hardening. Threat modeling is used to assess and mitigate customer risk from code changes. SecDL testing is fully automated during development, and all security-related code modifications are timed during minor releases to minimize risk.

Secure configuration baseline

Nutanix provides a security baseline based on the US DISA STIG format in both human readable format and in eXtensible Configuration Checklist Description Format (XCCDF), which allows automated assessment tools, such as Host Based Security System (HBSS), to read it. This provides detailed information on how to assess a Nutanix system to determine compliance with the baseline requirement, cutting down accreditation time from 9-12 months to a matter of minutes.

Security automation

A crucial element of security is the ability to track what configurations have changed and return them back to their desired state. NCI uses built-in automation to self-heal any deviation from the security baseline configuration of the platform.

Freedom to virtualize

NCI supports multiple industry standard virtualization solutions, allowing customers to choose the best solutions for their environments, whether on-prem or in the cloud, or both. The ability to run different virtualization solutions side by side from the same NCM platform prevents vendor lock-in and provides choice for future workloads.

This freedom of choice in hypervisors and clouds allows applications and data to move between runtime environments and includes a broad range of capabilities for migrating between different environments, including:

- Non-Nutanix infrastructure to Nutanix systems
- Between Nutanix systems supporting different hypervisor environments
- Nutanix to a public cloud infrastructure

Full-stack industry-standard support

Nutanix wields [support as a competitive advantage](#) with an industry-leading 90+ Net Promoter Score. Nutanix support covers the entire infrastructure stack—compute, storage, and virtualization.

AOS Storage underpins data services such as VM-centric provisioning, snapshots, clones, data protection, resilience, and availability for all applications. AOS Storage combines with the following technologies to create enterprise-to-public cloud flexibility.

Foundation Central: Standardize and simplify installation of Nutanix software and hypervisor of your choice on a Nutanix cluster across deployments, including remote sites.

Cross-hypervisor backup: Backup application data on remote clusters running different hypervisors and recover quickly with a single click.

Move: A hypervisor and cloud migration tool. Transfer a running VM to AHV with almost no downtime. Automatically turn the machine off in one hypervisor and on in AHV. Also allows you to automatically move and reinstantiate a VM from one location to another.

AHV

The native Nutanix hypervisor, AHV, provides a much simpler and more scalable virtualization solution by leveraging the software intelligence of the hyperconverged architecture. AHV is so easy to use that it allows enterprises to attend to what adds value--the applications--and liberates virtualization from the domain of specialists—making it easily manageable by anyone, from DevOps teams to DBAs.

AHV is performance-tuned for Nutanix HCI, taking advantage of the intelligent storage services provided by AOS Storage. Because AOS Storage is optimized for use with server virtualization, it provides full data resiliency, and data services such as snapshots, clones, and provisioning operations at VM granularity. As a result, AHV is leaner and focused on delivering secure virtual compute services and high availability.

Enterprise ready virtualization solution

Based on Linux KVM/QEMU virtualization, all popular workloads work on AHV, which is hardened to meet the most stringent enterprise security requirements. AHV is included with NCI, so enterprises get full infrastructure and virtualization support from a single vendor with no additional licensing costs.

AHV data protection

Each VM running on AHV is automatically protected according to a designated schedule that can include local snapshots as well as replication to a remote site. AHV has full access to all the data protection capabilities of AOS Storage as described above.

The Nutanix solution's scale-out architecture enables incremental, predictable scaling of capacity and performance in a Nutanix cluster running any hypervisor, including AHV. Administrators can start with as few as three nodes and scale out without limits. The system automatically discovers new nodes and makes them available for use. Expanding clusters is as simple as selecting the discovered nodes you want to add and providing network configuration details. Through NCM, administrators can image or update new nodes to match the AHV version of their preexisting nodes to allow seamless node integration, no matter what version was originally installed.

AHV networking

AHV implements common networking functionality and provides a vSwitch that comes configured on each node. The virtual switch connects the Controller VM, hypervisor, and guest VMs to each other and to physical networks. The switch runs on each AHV node and starts automatically with no direct switch-level configuration required before the cluster is operational. Administrators can change the switch configuration to match network resiliency and redundancy needs, based on the customer's environment. AHV also provides IP Address Management natively, removing the complexity of maintaining a separate IPAM system.

Flow network security - network virtualization and security

Modern networking requirements don't stop with connectivity. Flow Network Security provides application security, visibility, service insertion, and network automation with partner solutions. Security includes east-west firewalls, or microsegmentation, allowing admins to easily manage network isolation and granular VM- and application-level network policy. The extensive visibility makes it easier to see the complex interactions of modern applications, while also improving troubleshooting and simplifying policy creation and maintenance. Service insertion and network automation help extend and enhance networking function through partner or API integrations.

Virtual machine management

VM management on AHV focuses on creation, updates, deletion, data protection, and monitoring of VMs and their resources. These cluster services and features are all available via the NCM interface, a distributed management layer that is available on the CVM on every AHV host.

VM operations

Prism displays a list of all VMs in an AHV cluster along with a wealth of configuration, resource usage, and performance details for each VM. Administrators can create VMs and perform numerous operations on selected VMs, including power on or off, power cycle, reset, shutdown, reboot, snapshots and clones, migration, pause, update, delete, and launch a remote console.

VM	HOST	# VCPUS	MEMORY	STATUS	MEMORY CAPACITY	CPU USAGE	CONTROLLER PROXIES	CONTROLLER WITH DATA	CONTROLLER PROXIES	CONTROLLER AND CLUSTER	BACKUP/PAUSE
server1	brandy3	104584	1	2 GB	3.37%	-	-	-	-	-	Yes
server2	brandy4		1	2 GB	3.65%	-	-	-	-	-	Yes
server3	brandy1	104588	2	2 GB	3.2%	-	-	-	-	-	Yes

Toolbar actions: Launch Console, Power Off/Restart, Take Snapshot, Migrate, Pause, Clone, Update, Delete

Figure 7. VM Operations in Prism.

Image management

The image management service within AHV is a centralized repository that delivers access to virtual media and disk images, as well as the ability to import from external sources. It allows you to store VMs as templates or master images, which you can then use to create new VMs quickly from a known good base image. The image management service can store the virtual disk files that are used to create fully functioning VMs or operating system installation media as an .iso file that you can mount to provide a fresh operating system install experience. Incorporated into NCM (Prism), the image service can import and convert existing virtual disk formats, including .raw, .vhd, .vmdk, .vdi and .qcow2. The previous virtual hardware settings do not constrain an imported virtual disk, allowing administrators the flexibility to fully configure CPU, memory, virtual disks, and network settings at the time of VM provisioning.

Image Configuration ? X

Successfully created image Windows 7 Desktop. X

Manage the images to be used for creating virtual disks.

+ Upload Image

Name	Annotation	Type	State	
Ubuntu Desktop	Linux Desktop Image	DISK	ACTIVE	✎ ✕
Ubuntu Desktop	desktop installer	ISO	ACTIVE	✎ ✕
Windows 2012R2 Ent...	W2012R2 Master Image	DISK	ACTIVE	✎ ✕
Windows 7 Desktop	Win7 Master Image	DISK	ACTIVE	✎ ✕

Close

Figure 8. Image Configuration in Prism.

AHV VM placement and resource scheduling

AHV can place virtual machines intelligently on nodes in a cluster based on deep analytics from virtualization, HCI, and networking inputs. Because it's a single stack, intelligent placement or hot spot avoidance decision are all-inclusive.

AOS Dynamic Scheduling

AOS Dynamic Scheduling (ADS) is an automatic function enabled on every AHV cluster to avoid hot spots within cluster nodes. ADS continually monitors CPU, memory, and storage data points to make migration and initial placement decisions for VMs and volumes. Starting with existing statistical data for the cluster, ADS watches for anomalies, honors affinity controls, and makes move decisions to avoid hot spots. Using machine learning, ADS can adjust move thresholds over time from their initial fixed values to achieve the greatest efficiency without sacrificing performance.

Affinity and Anti-affinity

Affinity controls provide the ability to govern where VMs run. AHV has two types of affinity controls: VM-host affinity and Antiaffinity.

VM-host affinity ties a VM to a host or group of hosts, so the VM only runs on that host or group. Affinity is particularly applicable for use cases that involve software licensing or VM appliances. In such cases, you often need to bind a VM appliance to a single host or limit the number of hosts an application can run on.

Anti-affinity lets you designate VMs that should not run on the same hosts. Antiaffinity gives you a mechanism that allows clustered VMs or VMs that are running a distributed application

to run on different hosts, thereby increasing the application's availability and resiliency. The system overrides this type of rule when a cluster becomes constrained, prioritizing VM availability over VM separation.

Live migration

Live migration allows the system to move VMs from one Acropolis host to another while the VM is powered on, whether the movement is initiated manually or through an automatic process. Live migration can also occur when a host is placed in maintenance mode, which evacuates all VMs.

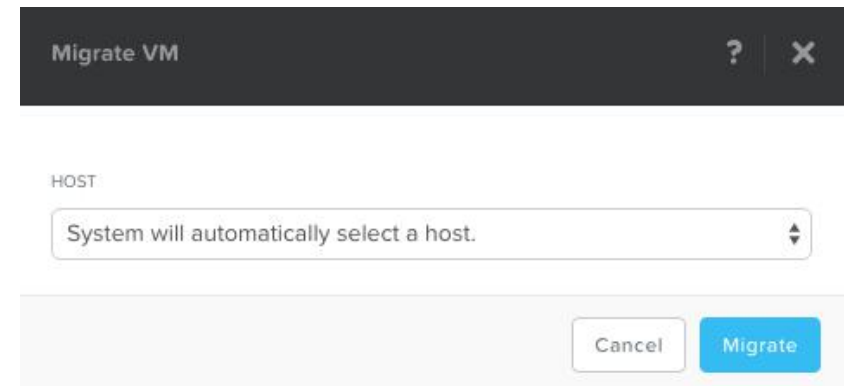


Figure 9. Migrating VMs.

Cross-hypervisor migration

Nutanix simplifies the process of migrating existing VMs between an ESXi cluster and an AHV cluster using built-in data protection capabilities. You can create one or more protection domains on the source cluster and set the AHV cluster as the target remote cluster. Then, snapshot VMs on the source ESXi cluster and replicate them to the AHV cluster, where you can restore them and bring them online as AHV VMs.

Automated high availability

Nutanix offers virtual machine high availability (HA) to ensure VM availability in the event of a host or block outage. If a host fails, the VMs previously running on that host restart on healthy nodes throughout the cluster. There are multiple HA configuration options available to account for different cluster scenarios.

By default, all Nutanix clusters provide best effort HA, even when the cluster is not configured for HA. Best effort HA works without reserving any resources. Admission control is not enforced, so there may not be sufficient capacity available to start all the VMs from the failed host.

You can also configure an Nutanix cluster for HA with resource reservation to guarantee that the resources required to restart VMs are always available. Nutanix offers two modes of resource reservation: host reservations and segment reservations. Clusters with uniform host configurations (for example, RAM on each node) use host reservation, while clusters with heterogeneous configurations use segment reservation.

High availability out of the box

AHV delivers high availability with minimal configuration and no additional software purchase. When a node fails, VMs are automatically restarted on other nodes in the cluster.

Converged backup and disaster recovery

The NCI converged backup and disaster recovery services protect your clusters. Nutanix clusters running any hypervisor have access to these features, which safeguard VMs both locally and remotely for use cases ranging from basic file protection to recovery from a complete site outage. To learn more about the built-in backup and disaster recovery capabilities in the Nutanix platform, read the data protection and disaster recovery technical note.

Backup APIs

To complement the integrated backup that the NCP provides, AHV also publishes a rich set of APIs to support external backup vendors. The AHV backup APIs utilize changed region tracking to allow backup vendors to back up only the data that has changed since the last backup job for each individual VM. Changed region tracking also allows backup jobs to skip reading zeros, further reducing backup times and bandwidth consumed.

Nutanix backup APIs allow backup vendors that build integration to perform full, incremental, and differential backups. Changed region tracking is always on in AHV clusters and does not require you to enable it on each VM. Backups can be either crash-consistent or application-consistent.

Analytics

NCM Intelligent Operations offers in-depth analytics for every element in the infrastructure stack, including hardware, storage, and VMs. Administrators can use element views to monitor these infrastructure stack components, and they can use the analysis view to get an integrated assessment of cluster resources or to drill down to specific metrics on a given element.

NCM Intelligent Operations makes detailed VM data available, grouping it into the following categories:

- VM performance: Multiple charts with CPU and storage-based reports around resource usage and performance.
- Virtual disks: In-depth data points that focus on I/O types, I/O metrics, read source, cache hits, working set size, and latency on a per-virtual disk level.
- VM NICs: vNIC configuration summary for a VM.
- VM snapshots: A list of all snapshots for a VM with the ability to clone or restore from the snapshot or to delete the snapshot.
- VM tasks: A time-based list of all operational actions performed against the selected VM. Details include task summary, percent complete, start time, duration, and status.
- Console: Administrators can open a pop-up console session or an inline console session for a VM.

VM Performance											
Virtual Disks					VM NICs		VM Snapshots		VM Tasks		Console
Default											
Additional Stats											
VIRTUAL DISK	READ LATENCY	WRITE LATENCY	TOTAL IOPS	RANDOM IO	READ SOURCE CACHE	READ SOURCE IOD	READ SOURCE IED	READ WORKING SET IOPS	WRITE WORKING SET IOPS	UNION WORKING SET IOPS	
NFS:860	2.27 ms	4.29 ms	15	74.67%	3.93 KB/s	15.47 KB/s	0 KB/s	682.78 MB/s	289.02 MB/s	953.54 MB/s	

Figure 10. Prism Analytics.

HOST NAME	HOST ID	CPU P	HYPERVISOR	CPU USAGE	CPU CAPACITY	MEMORY USAGE	MEMORY CAPACITY	TOTAL IOPS	IOPS CPU	IOPS IO	DISK LATENCY
brandy-1	10.456.76	10.456.83	AHV	1.8%	94.1% CPU	53.76%	26596 GB	18468 of 177K	0	0 KB/s	0 ms
brandy-2	10.456.79	10.456.83	AHV	1.6%	93.8% CPU	12.30%	26596 GB	3468 of 177K	0	0 KB/s	0 ms
brandy-3	10.456.80	10.456.83	AHV	1.6%	93.6% CPU	18.76%	26596 GB	3468 of 177K	0	0 KB/s	0 ms
brandy-4	10.456.81	10.456.83	AHV	1.2%	93.1% CPU	13.7%	26596 GB	30768 of 177K	0	0 KB/s	0 ms

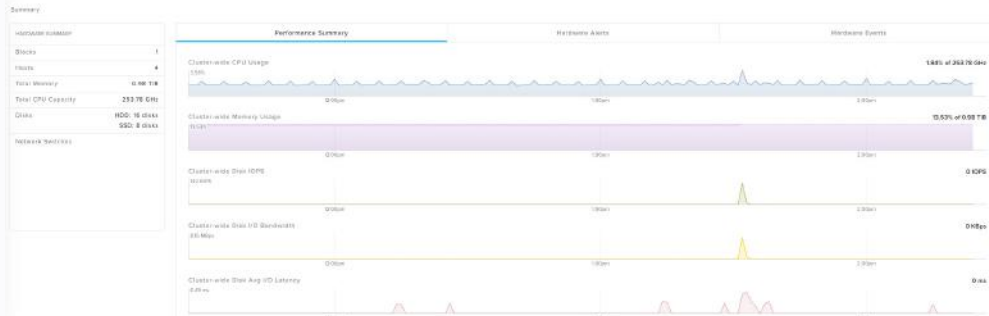


Figure 11. Performance Summary in Prism.

The NCM Intelligent Operations analysis tab gives administrators the tools they need to quickly understand what is going on in their clusters and to identify steps for remediation. You can create custom interactive charts using hundreds of metrics available for elements (such as hosts, disks, storage pools, containers, VMs, protection domains, remote sites, replication links, clusters, and virtual disks) then correlate trends in the charts with alerts and events in the system. You can also choose specific metrics and elements and set a desired time frame when building reports, so you can focus precisely on the data that you're looking for.

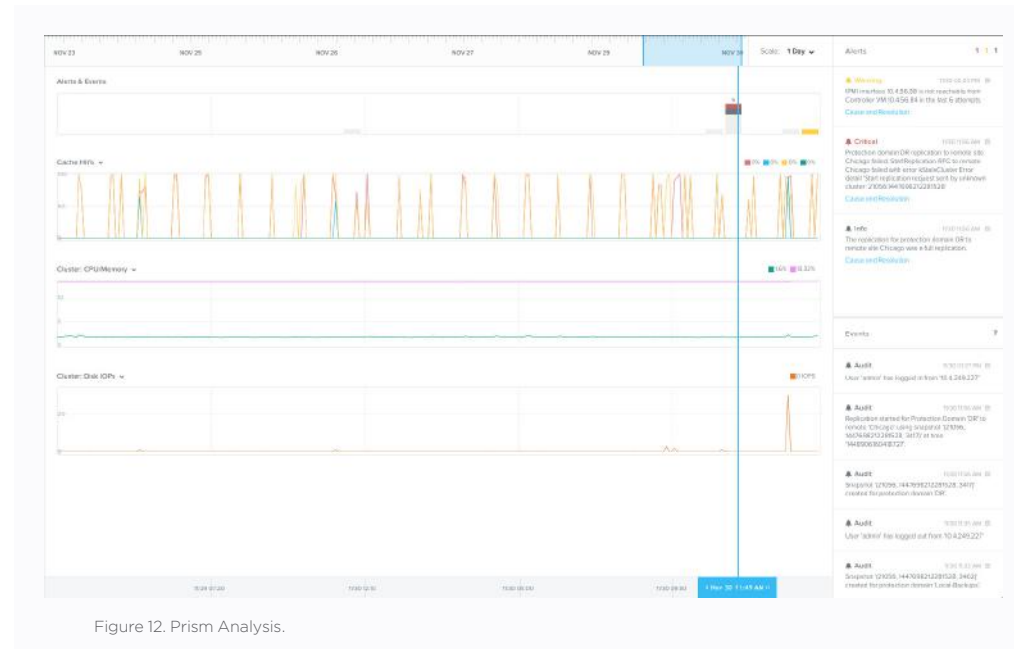


Figure 12. Prism Analysis.

Performance

The Nutanix platform optimizes performance at both the AOS Storage and hypervisor levels. The CVMs that represent the control and data planes contain the AOS optimizations that benefit all supported hypervisors. Built upon a foundation of open source KVM, a significant amount of added innovation makes AHV uniquely a Nutanix offering. The following sections outline a few of the innovations in AHV that are focused on performance.

AHV Turbo

AHV Turbo yields immediate out-of-the-box benefits for the data path, without any configuration.

AHV Turbo provides a new I/O path that bypasses virtualized storage emulation and directly services storage I/O requests. This approach lowers CPU usage and increases the amount of storage I/O available to VMs. AHV Turbo also introduces a multiqueue approach to allow data to flow from a VM to storage, resulting in vastly higher I/O capacity. The storage queues scale out automatically to match the number of vCPUs configured for a given VM, thus making even higher performance possible as the workload scales up.

While these improvements demonstrate immediate benefits, they also prepare AHV for future technologies such as NVMe and persistent memory advances that offer dramatically increased I/O capabilities with lower latencies.

vNUMA

Modern Intel server architectures assign memory banks to specific CPU sockets. In this design, one of the memory banks in a server is local to each CPU, so you see the highest level of performance when accessing memory locally, as opposed to accessing it remotely from a different memory bank. Each CPU and memory pair is a NUMA node. vNUMA is a function that allows a VM's architecture to mirror the NUMA architecture of the underlying physical host. vNUMA is not applicable to most workloads, but it can be very beneficial to very large VMs that are configured with more vCPUs than there are available physical cores in a single CPU socket.

RDMA

Remote Direct Memory Access (RDMA) allows a node to write to the memory of a remote node by allowing a VM running in the user space to directly access a NIC. This approach avoids TCP and kernel overhead, resulting in CPU savings and performance gains. AOS Storage RDMA support is reserved for inter-CVM communications and utilizes the standard RDMA over Converged Ethernet (RoCEv2) protocol on systems configured with RoCE-capable NICs connected to properly configured switches with datacenter bridging (DCB) support.

RDMA support, data locality, and AHV Turbo are not only important performance innovations for current generations, but uniquely position AHV and the Nutanix platform to take full advantage of rapidly advancing flash and memory technologies without requiring network fabric upgrades.

GPU support

A graphics processing unit (GPU) is the hardware or software that displays graphical content to end users. In laptops and desktops, GPUs are either a physical card or built directly into the CPU hardware, while GPU functions in the virtualized world have historically been software-driven and consumed additional CPU cycles. With modern operating systems and applications as well as 3-D tools, more and more organizations find themselves needing hardware GPUs in the virtualized world. You can install physical GPU cards into qualified hosts and present them to guest VMs via passthrough or vGPU mode.

Unify clouds with NCM and Prism

Nutanix Prism is the single pane of glass for NCM and provides an easy way to manage Nutanix environments end to end. It combines multiple management aspects into a single consumer-grade product that lets IT admins manage infrastructure and virtualization, access operational insights, and fix problems, all with a few clicks.

Just as NCI creates a data plane that spans the entire cluster for performance and resiliency, Nutanix Prism creates the same resiliency for management and operational intelligence. It is comprised of two components: Prism Element at the cluster level and Prism Central for multi-cluster management and analytics.

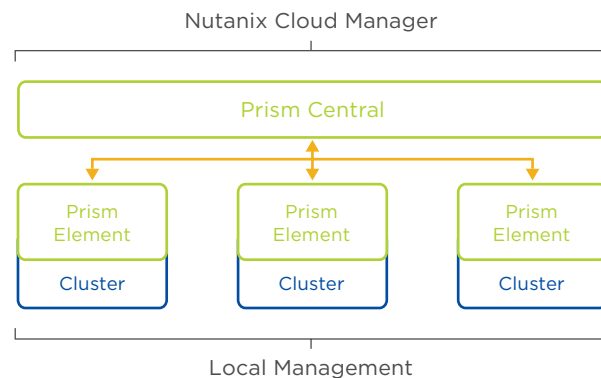


Figure 13. Prism Central.

Prism is highly available by design

Because Prism Element runs on every node in a cluster, it automatically deploys Prism Central as a highly available multi-VM distributed system. There are no external servers or databases to configure.

In addition to the device-friendly HTML5 Prism interfaces, all management capabilities are exposed through comprehensive APIs, PowerShell, and the command line interface (CLI) to facilitate integration and automation.

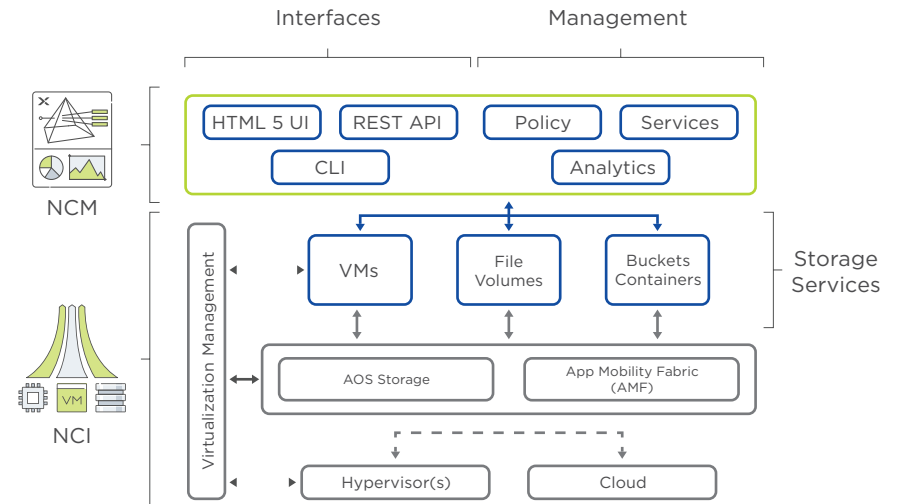


Figure 14. High-Level Prism Architecture.

The Prism approach

Prism offers an uncluttered experience with an intuitive user interface that simplifies and streamlines common enterprise cloud workflows, eliminating the need for different management tools for different tasks. Prism enhances productivity through features such as:

- **Instant search:** Integrated search to query and perform actions quickly.
- **Capacity planning:** The Prism analysis engine forecasts the capacity needs of applications running and allows for detailed “what if?” impact analysis on a Nutanix cluster, giving the IT team the ability to proactively plan for infrastructure needs.
- **Predictive analytics:** Leveraging the rich data sources from the HCI and virtualization stack, Nutanix Xfit machine learning provides predictive analysis and insights into resource utilization and resource demand.

- **Customizable operations dashboard:** Visual dashboard gives an at-a-glance summary of application and infrastructure state.
- **One-click simplicity** for infrastructure management, operational insights, and rapid problem remediation.

The Prism UI is a single comprehensive interface for monitoring and control over all the capabilities provided by Nutanix. There are no additional management VMs or plugins required to take advantage of the full breadth of NCP functionality, which goes well beyond basic HCI management.

Starting with high-level dashboards for access to the most important statistics and alerts at-a-glance, Prism lets you drill down into HCI infrastructures (servers, drives, networking), multi-hypervisor VM management (create, update, consoles), system health, data protection and replication, deep analytics, and alerting and alarming—all from one management plane.

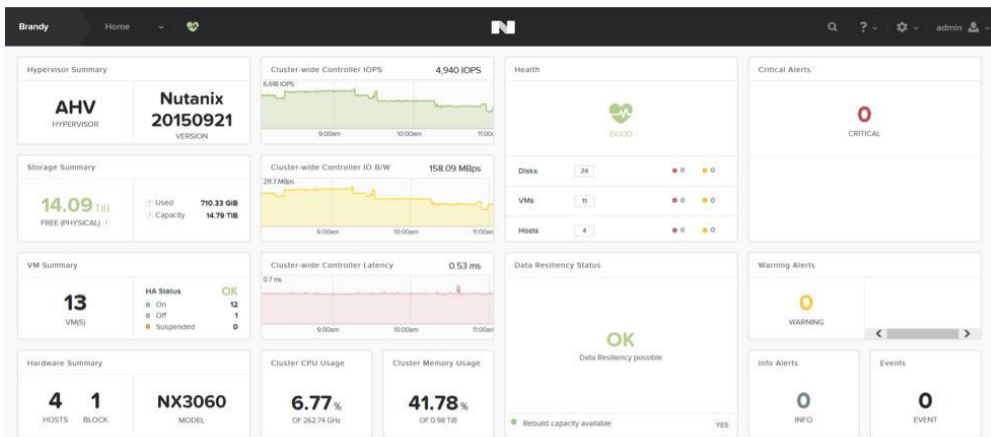


Figure 15. Prism Element — Dashboard.

Software upgrades

Prism provides a single interface for managing the software lifecycle of the entire platform, from HCI infrastructure (hypervisor, BIOS, disk firmware, Nutanix system software) to IT operations. Performing system software upgrades is a simple and non-disruptive process that can occur in the middle of a workday.

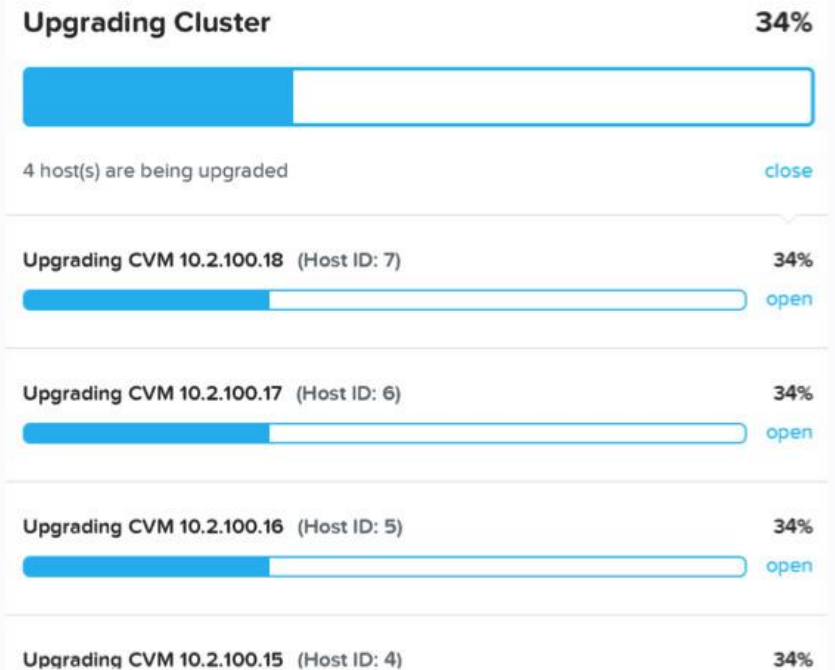


Figure 16.

To conduct an upgrade, just select “Upgrade Software” from the Prism dashboard and download the desired software version from the cloud. Prism automatically orchestrates the software installation across all nodes. That’s it. Three steps, regardless of cluster size.

Pro Tip: Prism Central

Nutanix recommends Prism Central for larger or distributed deployments (more than one Nutanix cluster or multiple sites) to simplify operations and provide a single management UI across all clusters and sites.

With the Prism Central dashboard, administrators can monitor and manage multiple clusters, including consolidated alerts, available storage, performance (bandwidth and IOPS), and more.

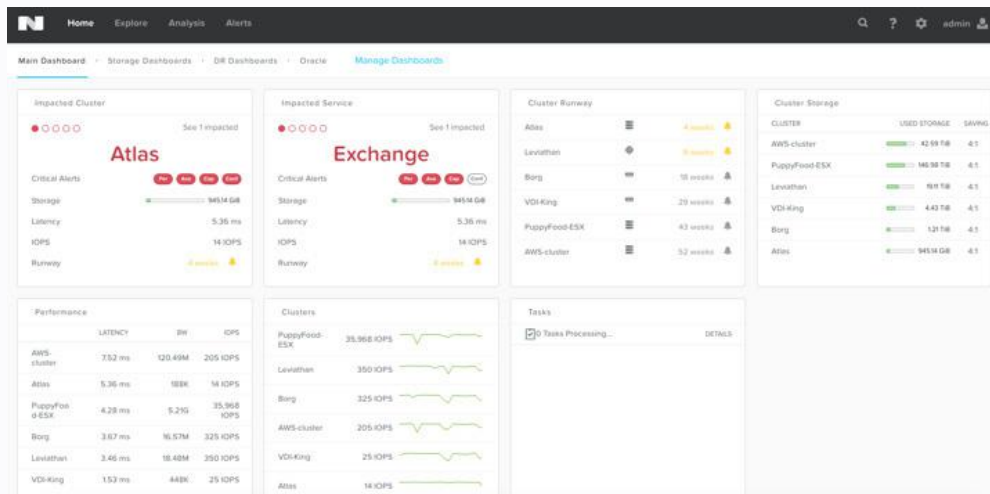


Figure 17. Prism Central — Dashboard.

Nutanix Self-Service and Orchestration

A vital part of NCM, Nutanix Self-Service and Orchestration provides application automation and lifecycle management that spans cloud environments, including Nutanix-powered private clouds and public clouds. Self-Service and Orchestration builds on NCI to make the entire IT infrastructure more agile and application centric.

Automation is built into Nutanix Self-Service and Orchestration to give organizations the ability to run applications on multiple hypervisors and clouds without platform lock-in, empowering them to adjust workloads for business priorities, while also providing the fastest time to market and the lowest operational expense.

Nutanix Self-Service and Orchestration defines applications via simple blueprints that administrators can easily create and instantly deploy. IT managers can either utilize pre-integrated blueprints or create their own, then publish them to the Nutanix Marketplace. IT can empower other teams, such as application developers or lines of business, to set up and manage applications from the Nutanix Marketplace in a self-service manner—while retaining full control of the infrastructure.

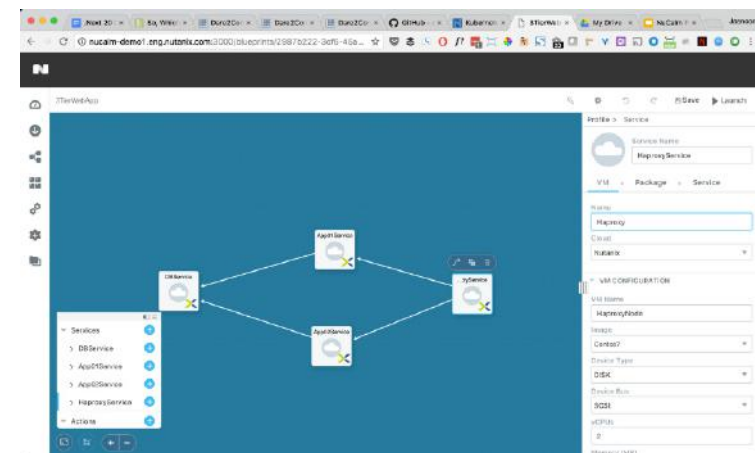


Figure 18. Nutanix Self-Service and Orchestration

Organizations can now fully automate hybrid cloud architecture deployments, scaling both multitiered and distributed applications across different cloud environments.

Ready to learn more about hyperconverged infrastructure and Nutanix? Contact us at info@nutanix.com, follow us on Twitter @nutanix, or send us a request at www.nutanix.com/demo to set up your own customized briefing and demonstration to see how validated and certified solutions from Nutanix can help your organization make the most of its enterprise applications.

Stay engaged with Nutanix experts and customers on the Nutanix Next online community (next.nutanix.com).

About Nutanix

Nutanix elevates IT to focus on the applications and services that power their business. The Nutanix Cloud Platform 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.

NUTANIX™