

# Section 1 – Nutanix Enterprise Cloud Concepts

## Define and differentiate features and technologies present in Acropolis, Prism and Calm

### Acropolis

Acropolis is a distributed multi-resource manager, orchestration platform and data plane. It is broken down into the following three main components:

### Distributed Storage Fabric (DSF)

- This is at the core and birth of the Nutanix platform and expands upon the Nutanix Distributed Filesystem (NDFS). NDFS has now evolved from a distributed system pooling storage resources into a much larger and capable storage platform.
- Distributed multi-resource manager, orchestration platform, data plane
- All nodes form DSF
- Appears to Hypervisor as storage array
- All I/O's are handled locally for highest performance

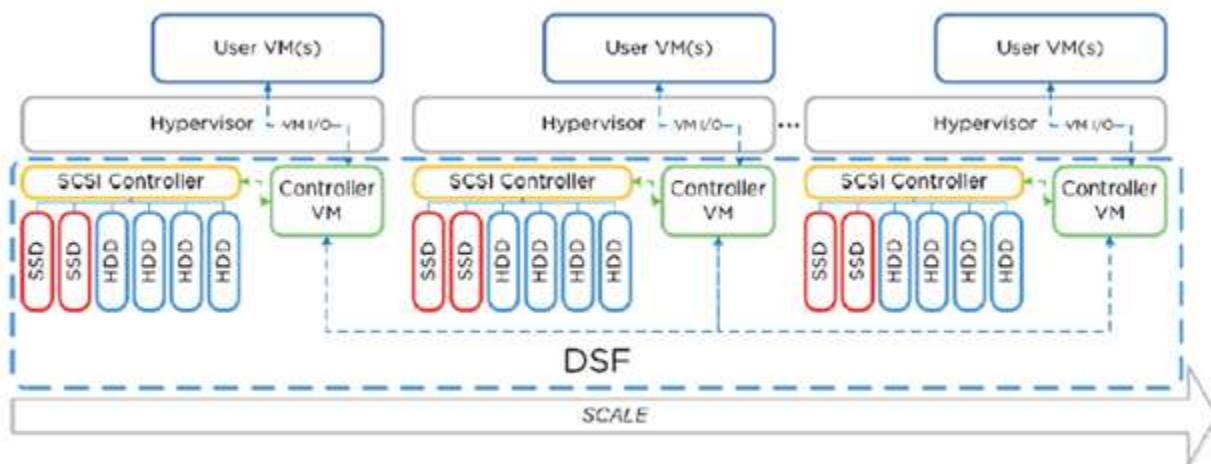


Image credit: <https://nutanixbible.com>

# App Mobility Fabric (AMF)

- Hypervisors abstracted the OS from hardware, and the AMF abstracts workloads (VMs, Storage, Containers, etc.) from the hypervisor. This will provide the ability to dynamically move the workloads between hypervisors, clouds, as well as provide the ability for Nutanix nodes to change hypervisors.

## Hypervisor

- A multi-purpose hypervisor based upon the CentOS KVM hypervisor.

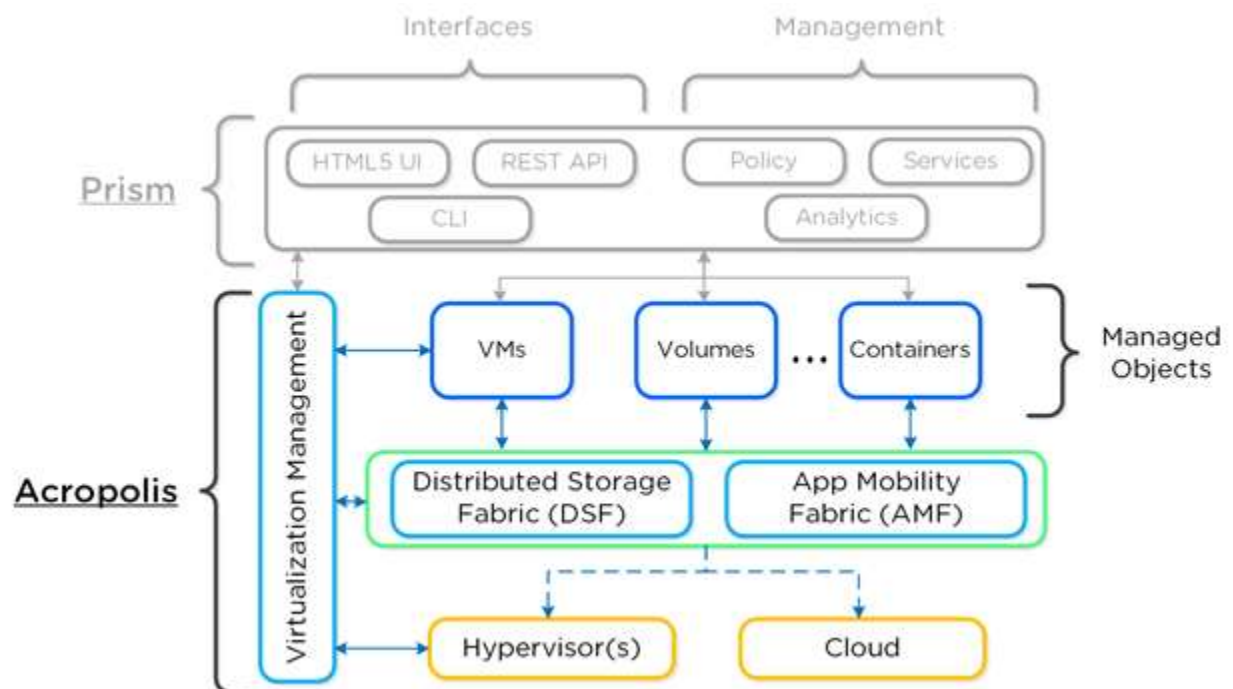


Image credit: <https://nutanixbible.com>

- Acropolis service allows for workloads/resource management/provisioning/operations
- Seamlessly move workloads between hypervisors/clouds

## Architecture

- CVM runs as a VM and disks are presented using PCI Pass-through
- Allows for full PCI controller to be passed to CVM and bypass hypervisor
- Based on CentOS

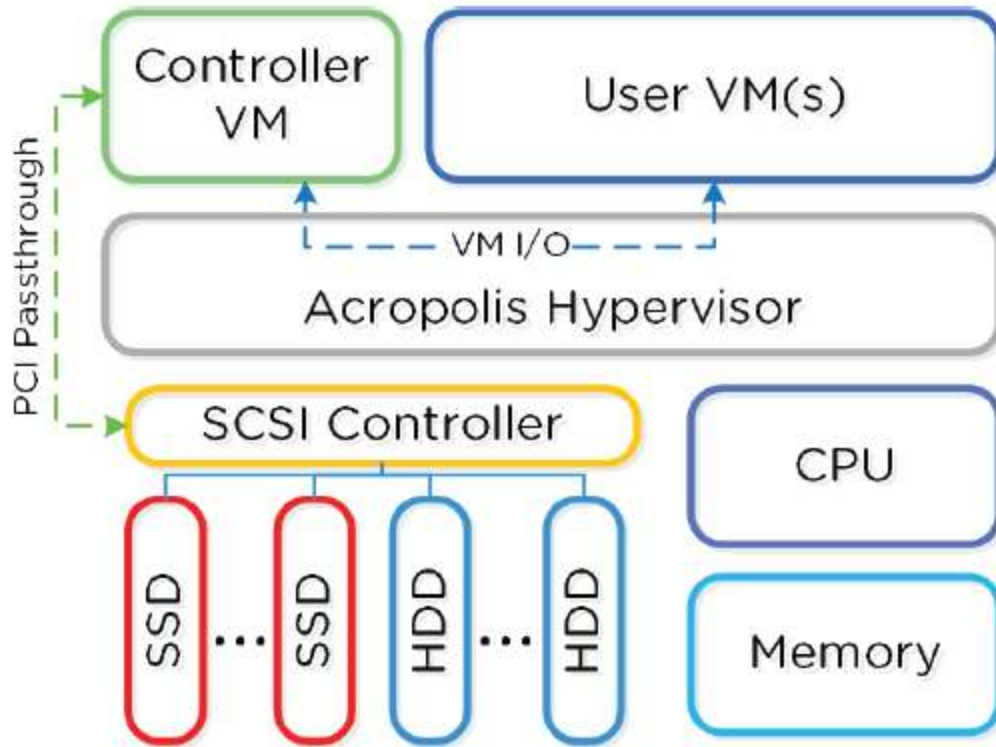


Figure 4.1.1. AHV Node

Image credit: <https://nutanixbible.com>

- Includes HA, live migration, etc.
- KVM-kmod – kernel module
- Libvirtd – API, daemon, and management tool for managing KVM/QEMU.
- Communication between Acropolis/KVM/QEMU occurs through libvirtd.
- Qemu-kvm – machine emulator/virtualizer runs in userspace for every VM.
- Used for hardware assisted virtualization

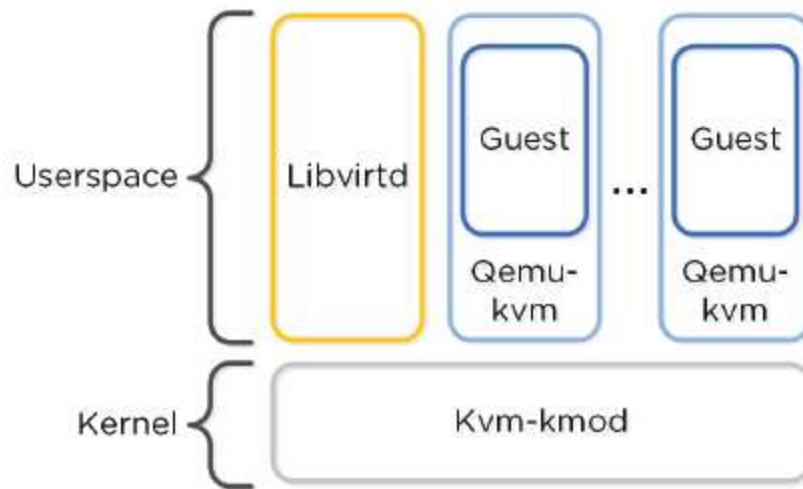


Figure 4.1.2. KVM Component Relationship

Image credit: <https://nutanixbible.com>

- AHV has an EVC-like feature.
- Determines lowest processor generation in cluster and constrains all QEMU domains to that level

## Scalability

- Max cluster size: N/A (same as Nutanix cluster)
- Max vCPU/VM: Number of physical cores/host
- Max mem/VM: 2TB
- Max VM/host: N/A (limited by mem)
- Max VM/cluster: N/A (limited by mem)

## Networking

- Leverages Open vSwitch (OVS)
- Each VM NIC connected to TAP interface

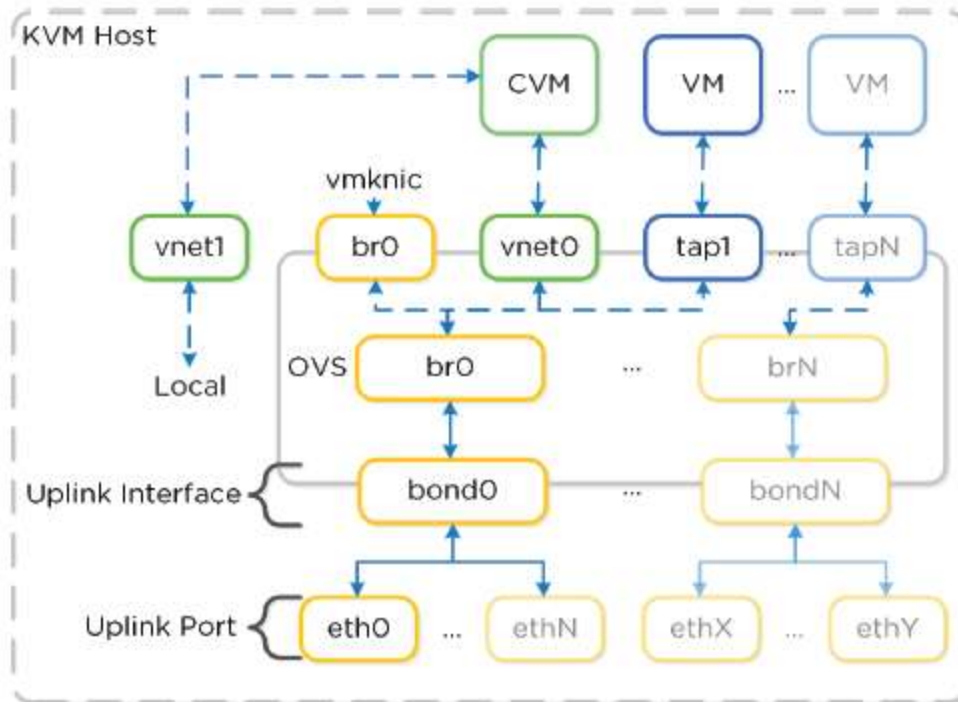


Figure 4.1.3. Open vSwitch Network Overview

Image credit: <https://nutanixbible.com>

- Supports Access and Trunk ports

## iSCSI Multipathing

- Each host has a iSCSI redirector daemon running which checks Stargate health with NOP OUT commands
- QEMU configured with iSCSI redirector as iSCSI target portal  
Redirector will redirect to healthy Stargate

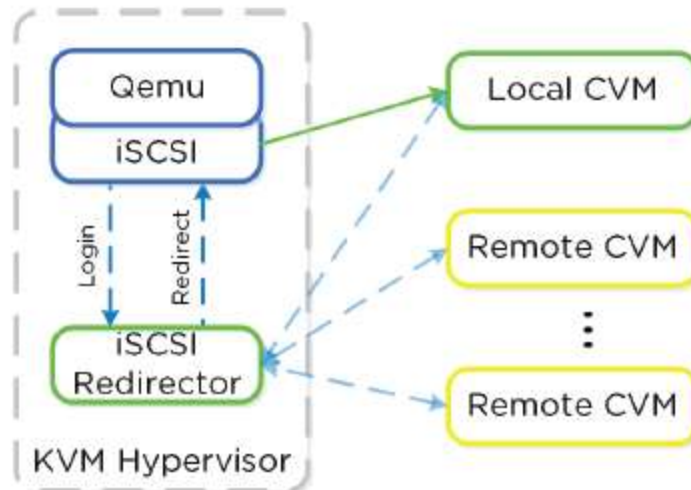


Figure 4.2.1. iSCSI Multi-pathing - Normal State

Image credit: <https://nutanixbible.com>

- If Stargate goes down (stops responding to NOP OUT commands), iSCSI redirector marks local Stargate as unhealthy.
- Redirector will redirect to another healthy Stargate

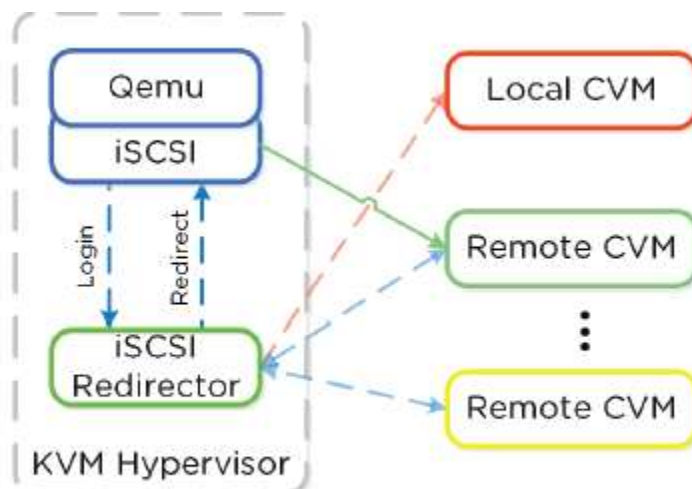


Figure 4.2.2. iSCSI Multi-pathing - Local CVM Down

Image credit: <https://nutanixbible.com>

- Once local comes back, redirector will perform TCP kill and redirect back to local Stargate.

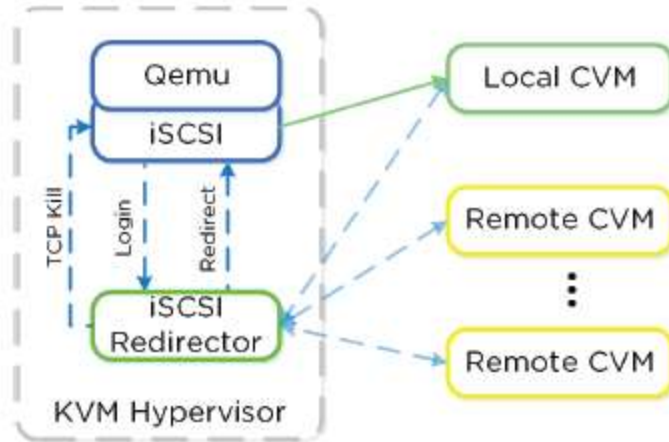


Figure 4.2.3. iSCSI Multi-pathing - Local CVM Back Up

Image credit: <https://nutanixbible.com>

## Troubleshooting iSCSI

- Check iSCSI Redirector logs: `cat /var/log/iscsi_redirector`
  - In the `iscsi_redirector` log (located in `/var/log/` on the AHV host), you can see each Stargate's health:

```
2017-08-18 19:25:21,733 - INFO - Portal 192.168.5.254:3261 is up
2017-08-18 19:25:25,735 - INFO - Portal 10.3.140.158:3261 is up
2017-08-18 19:25:26,737 - INFO - Portal 10.3.140.153:3261 is up
```

- NOTE: The local Stargate is shown via its 192.168.5.254 internal address
  - In the following you can see the `iscsi_redirector` is listening on 127.0.0.1:3261:

```
[root@NTNX-BEAST-1 ~]# netstat -tnlp | egrep tcp.3261
Proto ... Local Address Foreign Address State PID/Program nametcp ...
127.0.0.1:3261 0.0.0.0: LISTEN 8044/python
```

## Frodo I/O Path (aka AHV Turbo Mode)

As storage technologies continue to evolve and become more efficient, so must we. Given the fact that we fully control AHV and the Nutanix stack this was an area of opportunity.

In short Frodo is a heavily optimized I/O path for AHV that allows for higher throughput, lower latency and less CPU overhead.

## Controller VM (CVM)

- All I/O for host

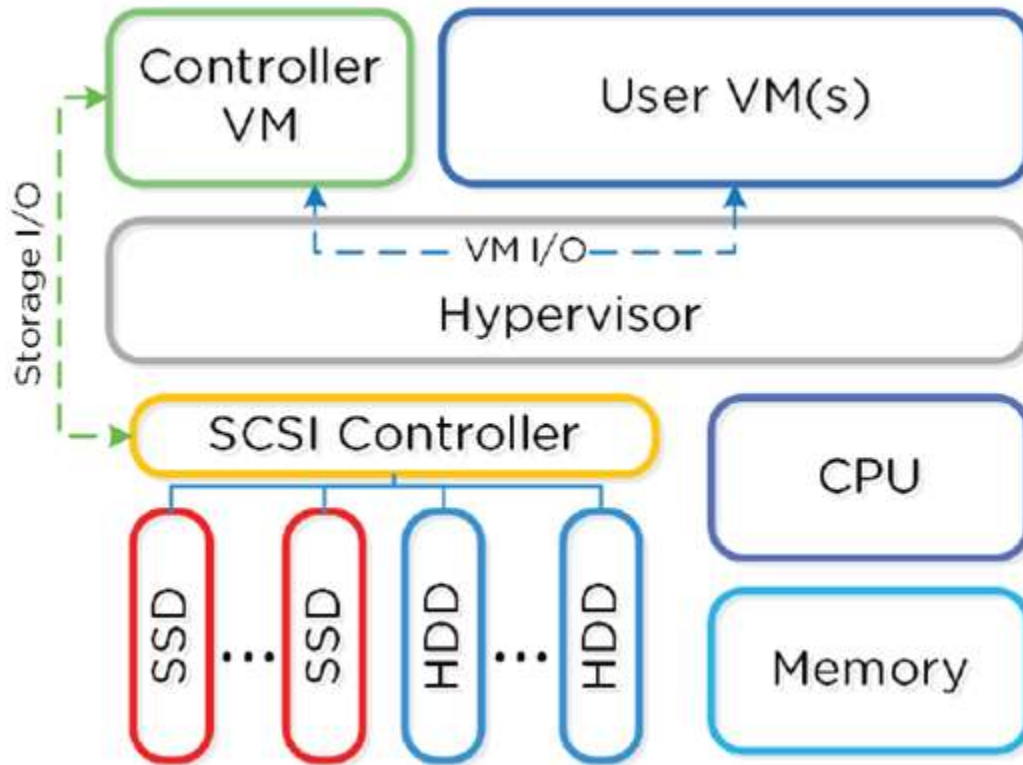


Image credit: <https://nutanixbible.com>

- Doesn't rely on hardware offloads or constructs for extensibility
- Deploy customer features via software
- Allows newer generation features for older hardware



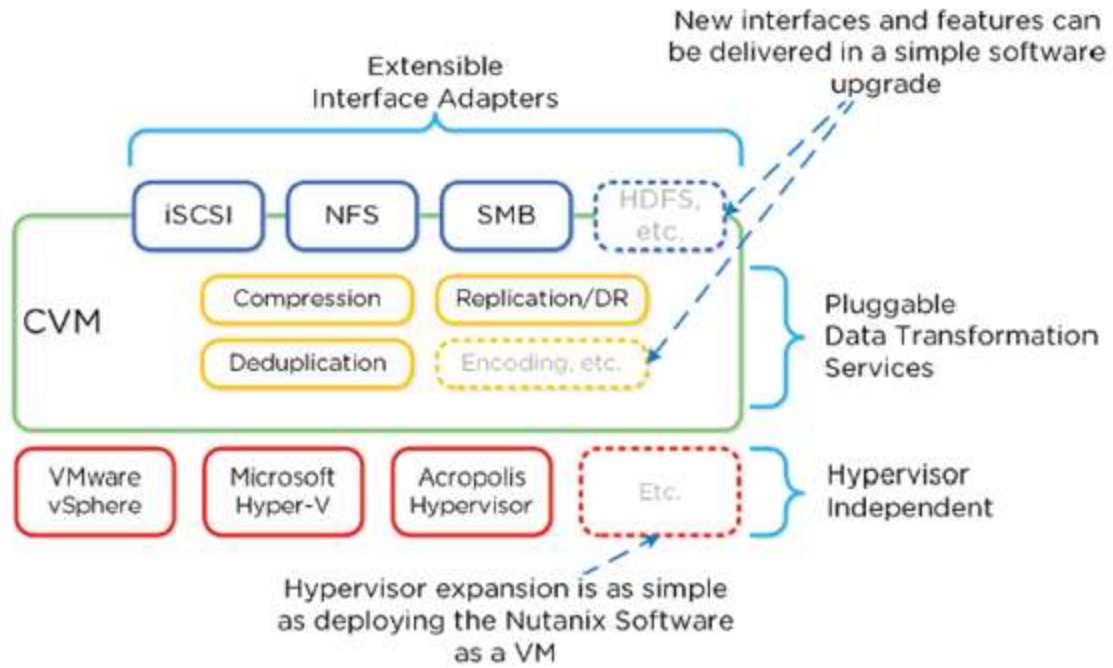


Image credit: <https://nutanixbible.com>

## I/O Path and Cache

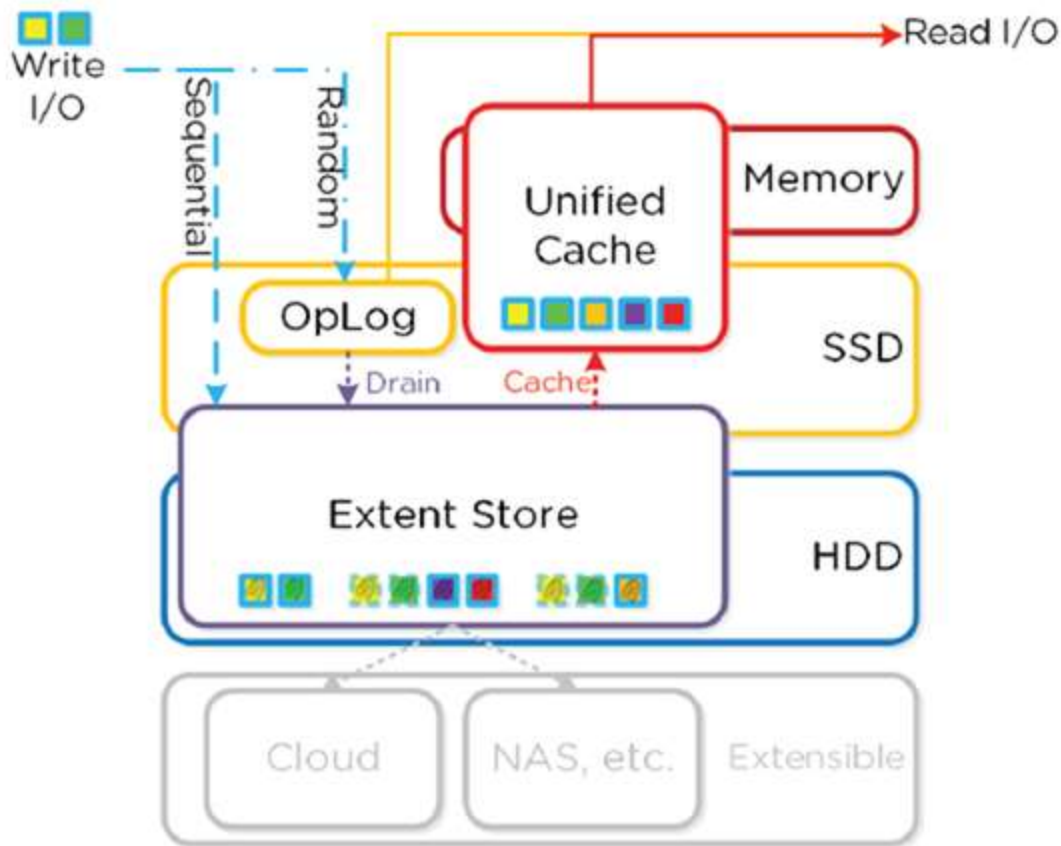


Image credit: <https://nutanixbible.com>

## OpLog

- Staging area for bursts of random writes, stored on SSD Tier
- Writes are coalesced, and sequentially drains to extent store
- Synchronously replicates to other CVM's OpLog's before ack'ing write
- All CVM's partake in replication
- Dynamically chosen based on load
- For **sequential workloads, OpLog is bypassed** (writes go directly to extent store)
- If data is in OpLog and not drained, all read requests fulfilled from OpLog
- When Dedupe is enabled, write I/O's will be fingerprinted to allow them to be deduplicated based on fingerprints.

## Extent Store

- Persistent bulk storage of DSF and spans SSD/HDD.

- Data is either :
  - Being drained from OpLog
  - Sequential in nature
- ILM determines tier placement
- Sequential = more than 1.5MB of outstanding write I/O

## Unified Cache

- Deduplicated read cache spanning CVM memory + SSD
- Data not in cache = single-touch pool (completely in memory)
- Subsequent requests move to multi-touch pool (memory + SSD)
- Evicted downwards

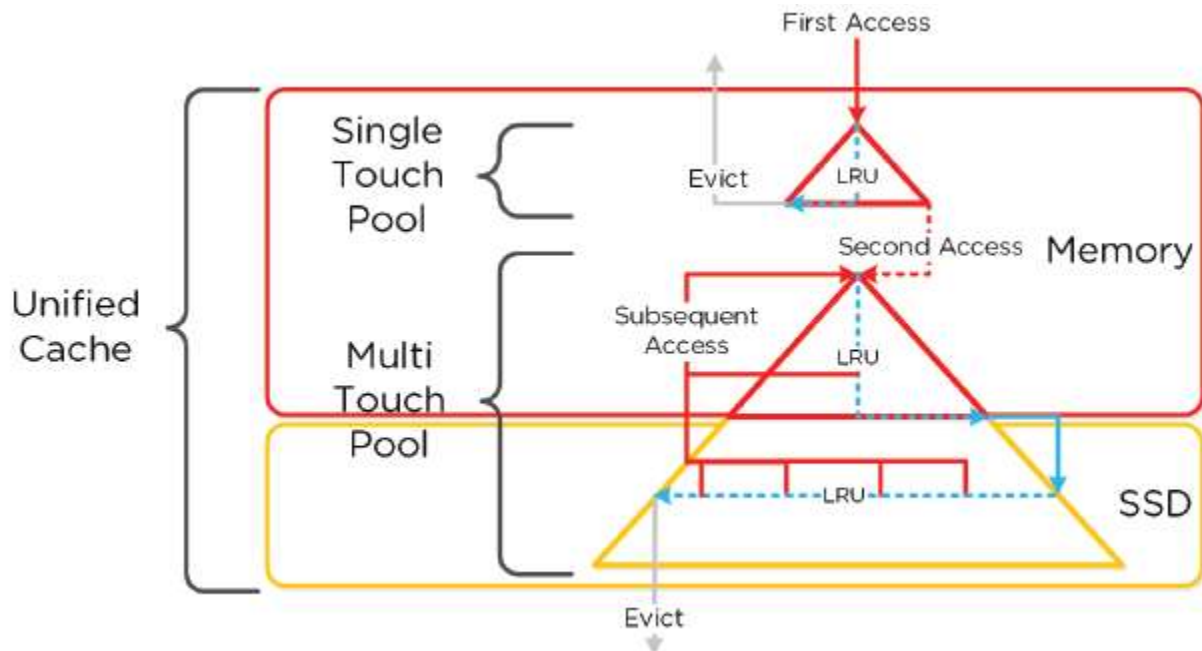


Image credit: <https://nutanixbible.com>

## Extent Cache

- In-memory read cache completely in CVM memory.
- Stores non-fingerprinted extents for containers where fingerprinting/dedupe are disabled

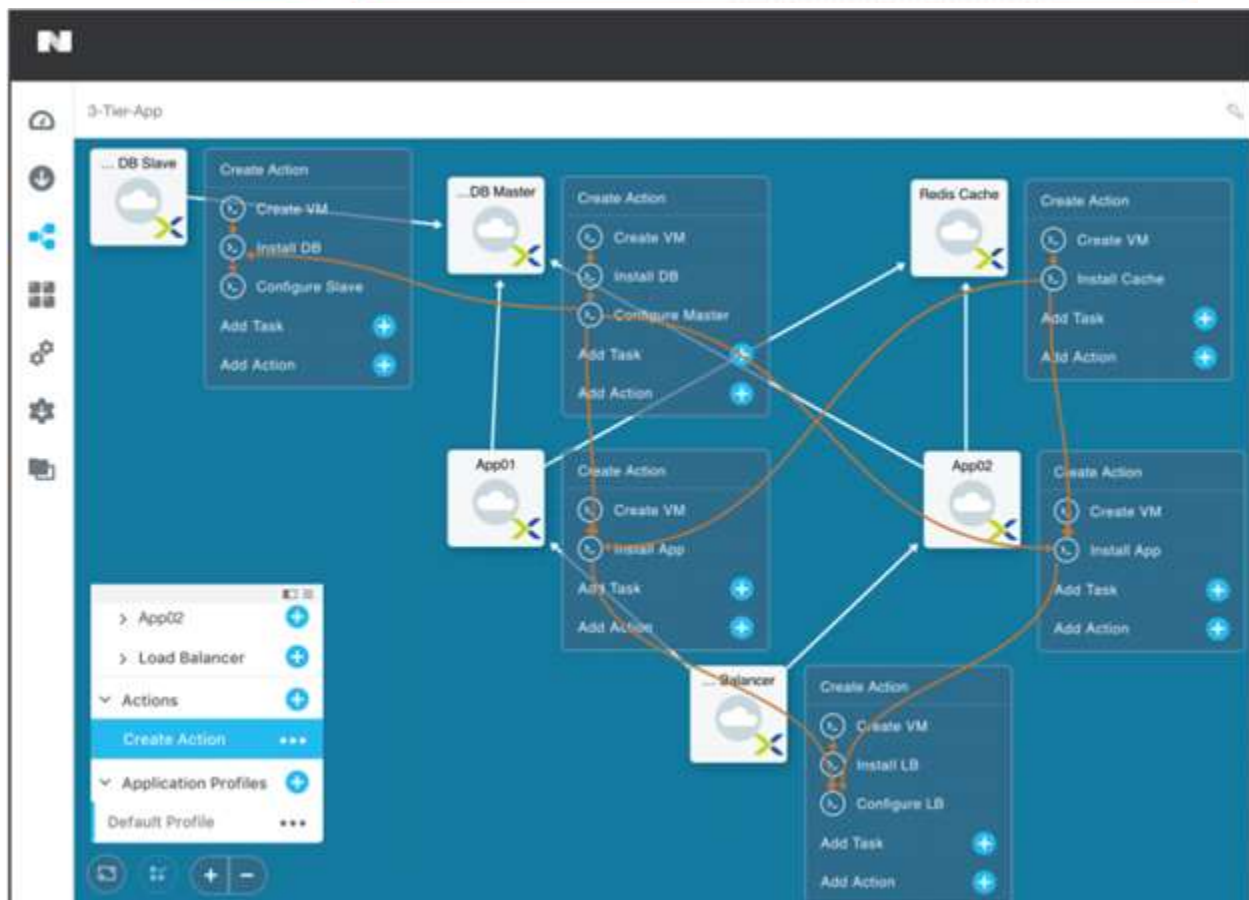
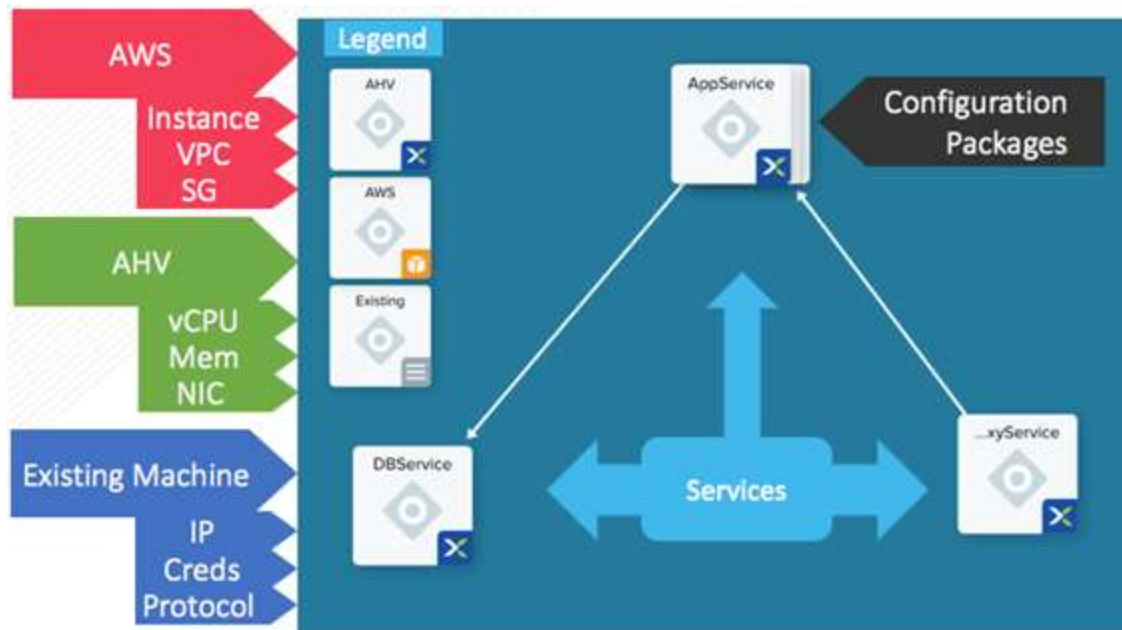
## Prism

Nutanix Prism is the proprietary management software used in Nutanix Acropolis hyper-converged appliances. Nutanix Prism provides an automated control plane that uses machine learning to support predictive analytics and automated data movement.

## Calm

Calm is a multi-cloud application management framework delivered by Nutanix. Calm provides application automation and lifecycle management natively integrated into the Nutanix Platform. With Calm, applications are defined via simple blueprints that can be easily created using industry standard skills and control all aspects of the application's lifecycle, such as provisioning, scaling, and cleanup. Once created, a blueprint can be easily published to end users through the Nutanix Marketplace, instantly transforming a complex provisioning ticket into a simple one-click request.

- **Application Lifecycle Management:** Automates the provision and deletion of both traditional multi-tiered applications and modern distributed services by using pre-integrated blueprints that make management of applications simple in both private (AHV) and public cloud (AWS).
- **Customizable Blueprints:** Simplifies the setup and management of custom enterprise applications by incorporating the elements of each app, including relevant VMs, configurations and related binaries into an easy-to-use blueprint that can be managed by the infrastructure team.
- **Nutanix Marketplace:** Publishes the application blueprints directly to the end users through Marketplace.
- **Governance:** Maintains control with role-based governance thereby limiting the user operations that are based on the permissions.
- **Hybrid Cloud Management:** Automates the provisioning of a hybrid cloud architecture, scaling both multi-tiered and distributed applications across cloud environments, including AWS.



**Explain the relationship between nodes, blocks and clusters**

## Node

- A physical server contained in a Nutanix block
- Model Comparison:
  - 2 Node (6000/7000)
  - 4 Node (1000/2000/3000/3050)
- Each node runs a hypervisor (ESXi/HV/KVM) + Nutanix CVM (Controller VM)

## Block

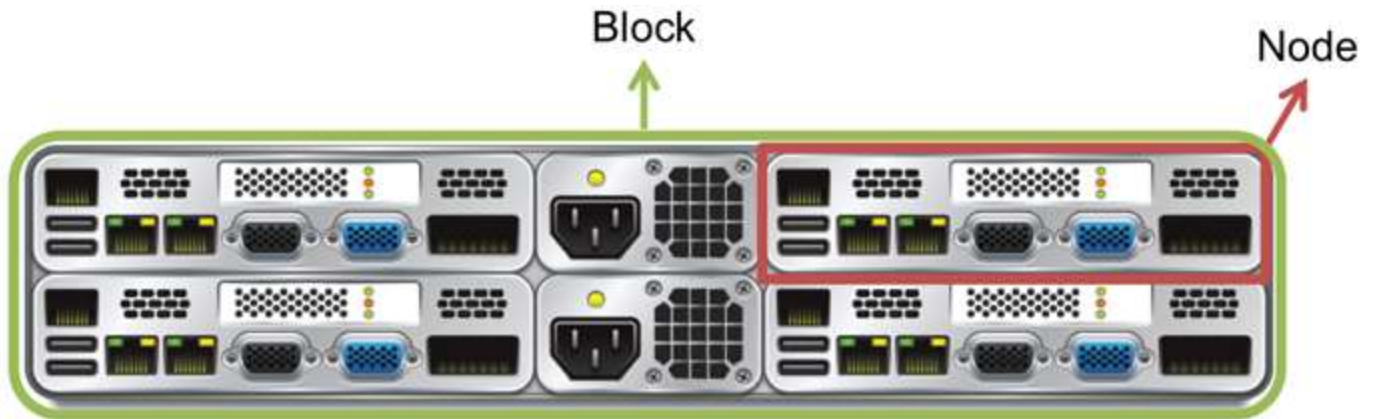
- A set of Nutanix nodes contained in a single enclosure
- Uniform populated blocks recommended
  - Prevents storage skew
  - Ensures if block fails/maintenance needed, system can run without interruption

## Cluster

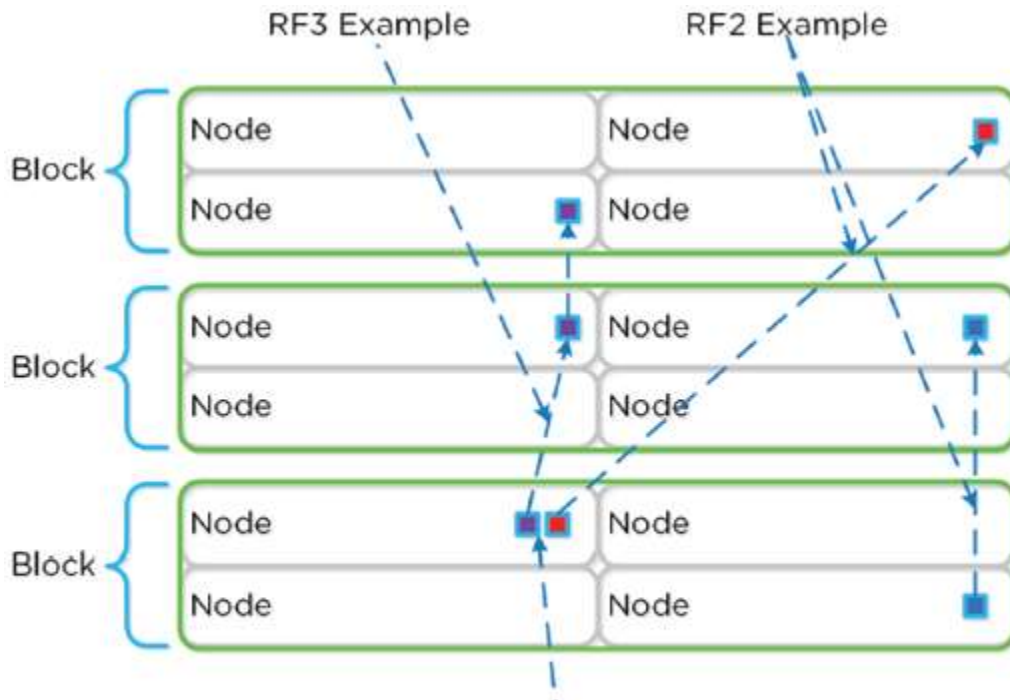
A group of nodes contained in one or more Nutanix blocks

## Node/Block/Rack Awareness

- Determines component/data placement
- DSF is node/block aware
- **Node awareness** is default
  - Replicas are replicated to another node to protect against node failure
- Min. 3 blocks needed for **block awareness**
  - Replicas are written to other blocks in cluster
  - Provides data availability in case of block outage
  - As of AOS 4.5, block awareness is best effort
- Increases to rack aware as cluster scales



- Within block, PSU + fans are the only shared components

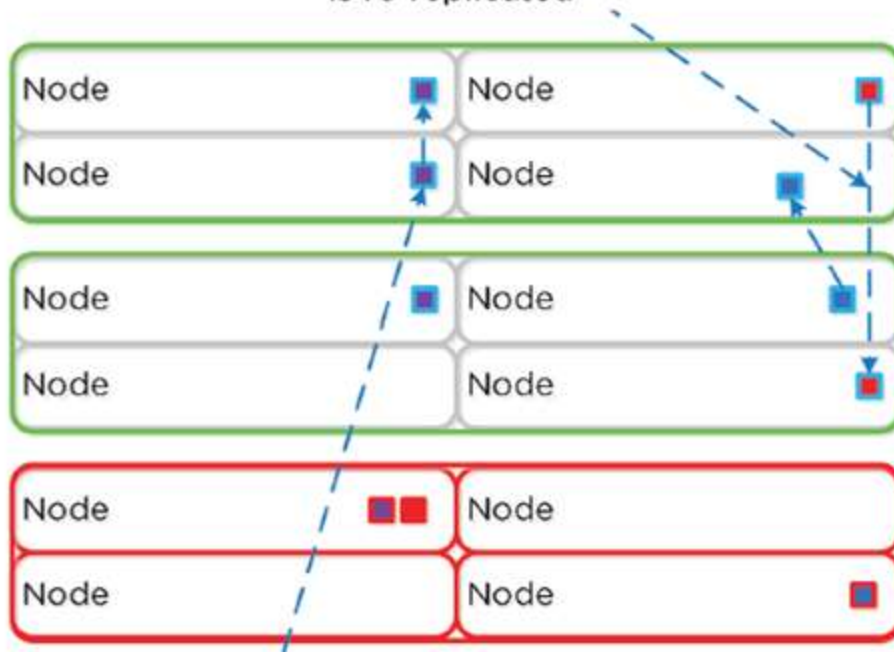


Block awareness ensures that the replicas will be replicated to another block in the cluster

Image credit: <https://nutanixbible.com>

- In event of block failure, block awareness maintained and re-replicated blocks are replicated to other blocks in cluster:

Block awareness is maintained even in the case of a block failure where data is re-replicated



In the event where full block awareness cannot be fulfilled node awareness will still be fulfilled to maintain the RF

Image credit: <https://nutanixbible.com>

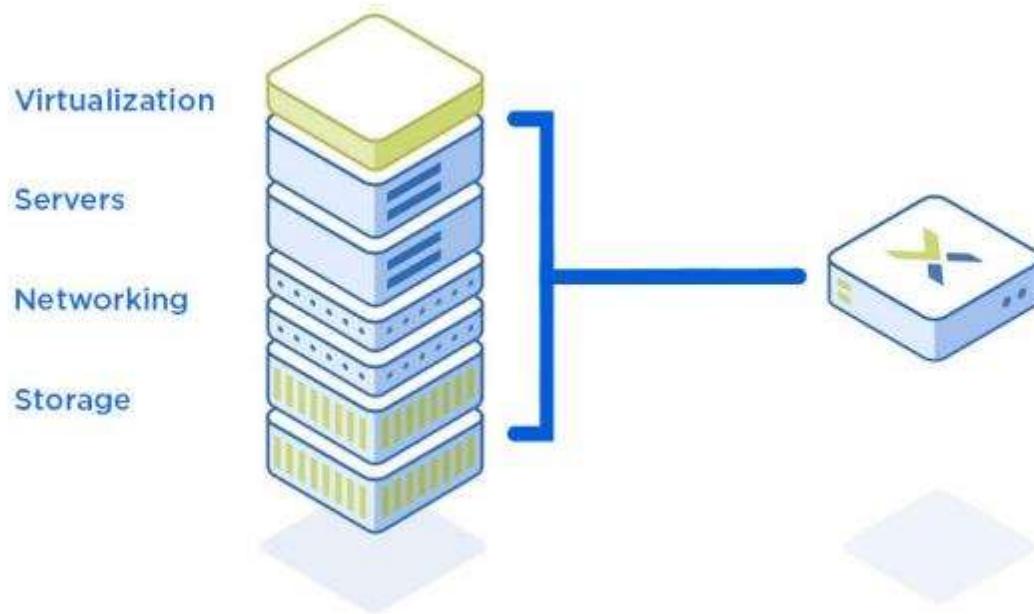
Number of Blocks	Awareness Type	Simultaneous Failure Tolerance	
		Cluster FT1	Cluster FT2
<3	NODE	SINGLE NODE	DUAL NODE
3-5	NODE+BLOCK	SINGLE BLOCK (up to 4 nodes)	SINGLE BLOCK (up to 4 nodes)
5+	NODE+BLOCK	SINGLE BLOCK (up to 4 nodes)	DUAL BLOCK (up to 8 nodes)



Desired Awareness Type	FT Level	EC Enabled?	Min. Units	Simultaneous failure tolerance
Node	1	No	3 Nodes	1 Node
Node	1	Yes	4 Nodes	1 Node
Node	2	No	5 Nodes	2 Node
Node	2	Yes	6 Nodes	2 Nodes
Block	1	No	3 Blocks	1 Block
Block	1	Yes	4 Blocks	1 Block
Block	2	No	5 Blocks	2 Blocks
Block	2	Yes	6 Blocks	2 Blocks
Rack	1	No	3 Racks	1 Rack
Rack	1	Yes	4 Racks	1 Rack
Rack	2	No	5 Racks	2 Racks
Rack	2	Yes	6 Racks	2 Racks

## Recognize the benefits of a Nutanix hyperconverged infrastructure solution

Hyperconverged Infrastructure converges the entire datacenter stack, including compute, storage, storage networking, and virtualization. Complex and expensive legacy infrastructure is replaced by a platform running on turnkey, industry-standard servers that enable enterprises to start small and scale one node at a time. Software running on each server node distributes all operating functions across the cluster for superior performance and resilience.



Source: <https://www.nutanix.com>

## Components of a Hyperconverged Infrastructure

### Distributed Plane

Runs across a cluster of nodes delivering storage, virtualization, and networking services for guest applications—whether they're VMs or container-based apps.

### Management plane

Lets you easily administer HCI resources from one place and one view and eliminates the need for separate management solutions for servers, storage networks, storage, and virtualization.

Nearly all modern hyperconverged infrastructure solutions are 100% software-defined, with no dependency on proprietary hardware. Each HCI node in a cluster runs a hypervisor (Nutanix AHV, VMware ESXi, or Microsoft Hyper-V), and the HCI control features run as a separate virtual machine on every node, forming a fully distributed fabric that can scale resources with the addition of new nodes.

# Differentiate between physical and logical cluster components

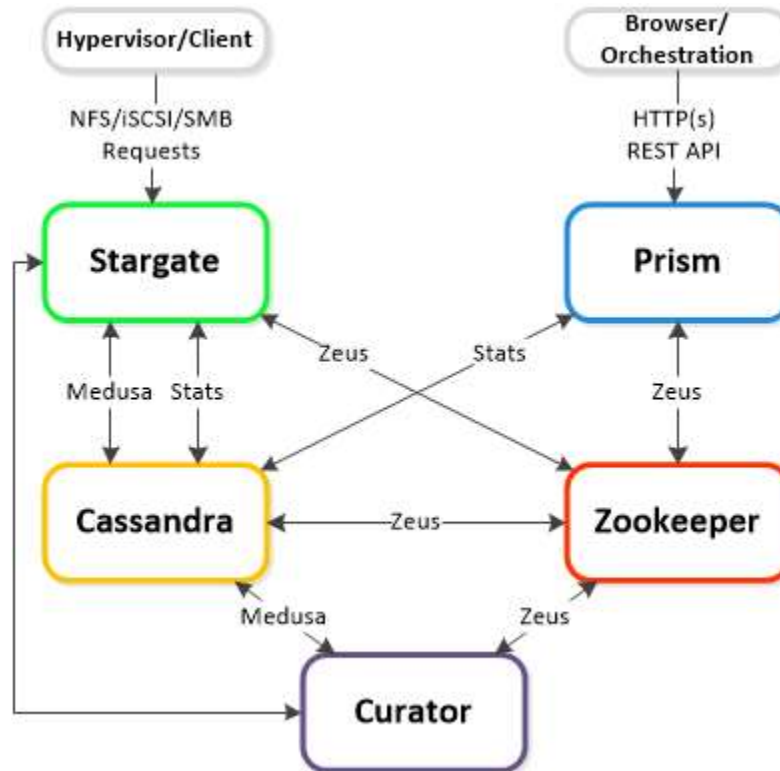


Image credit: <https://nutanixbible.com>

## Cassandra

- Distributed metadata store, based on heavily modified Apache Cassandra
- Stores/manages all metadata in ring-like manner
- Metadata describes where and how data is stored on a filesystem
- Let's system know which node/disk and in what form the data resides
- Runs on every node in the cluster
- Accessed via Medusa
- RF used to maintain availability/redundancy
- Upon write/update, row is written to a node in the ring and then replicated to n number of peers (n depends on cluster size)
- Majority must agree before commit

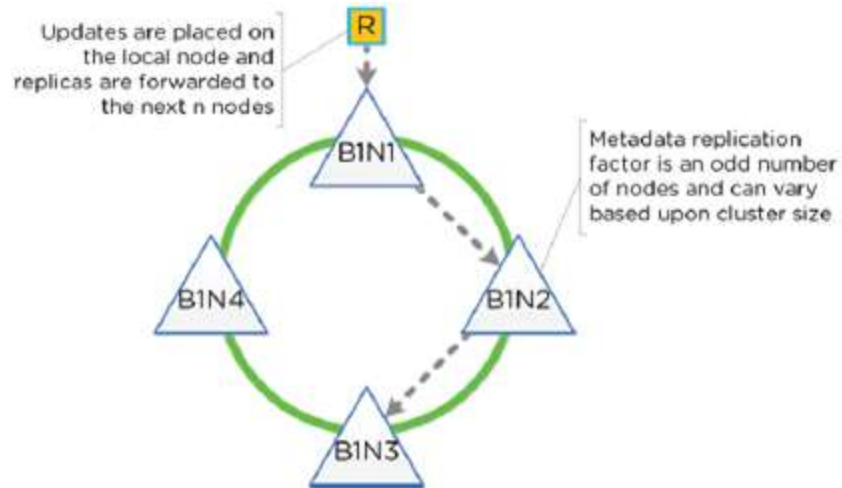


Image credit: <https://nutanixbible.com>

- Each node is responsible for a subset of overall platform metadata
- Eliminates traditional bottlenecks (served from all nodes vs dual controllers)
- When cluster scales, nodes are inserted for “block awareness” and reliability

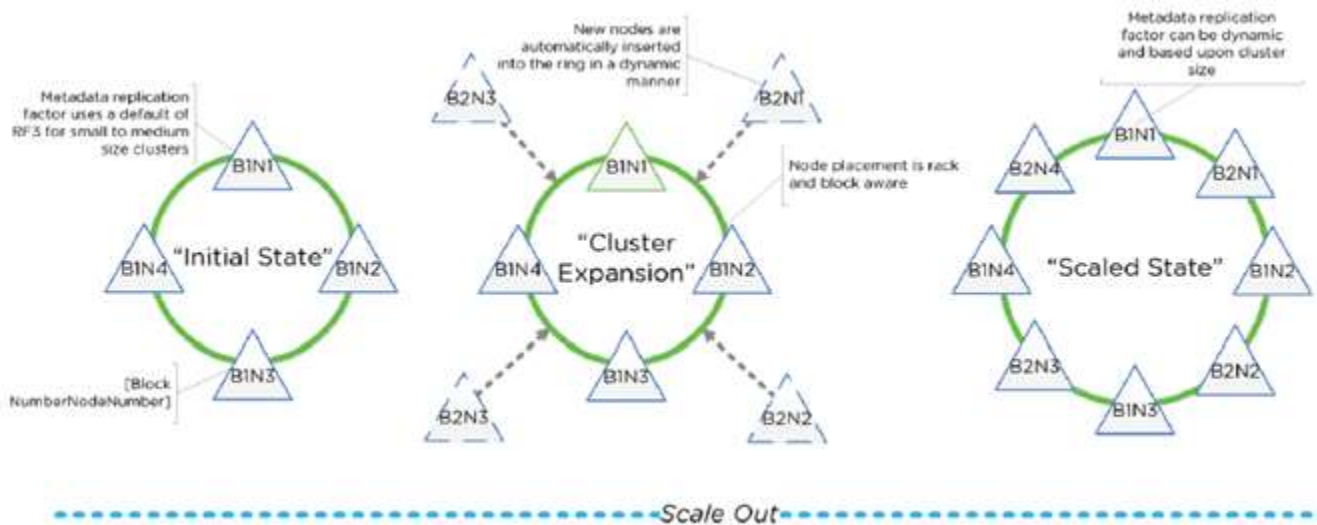


Image credit: <https://nutanixbible.com>

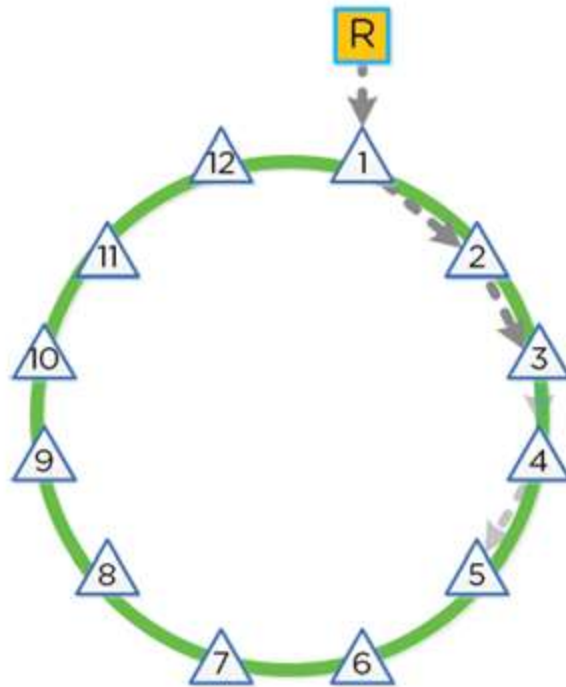


Image credit: <https://nutanixbible.com>

- Peer replication iterates through nodes in clockwise manner
- With block awareness, peers are distributed among blocks to ensure no two peers are in same block

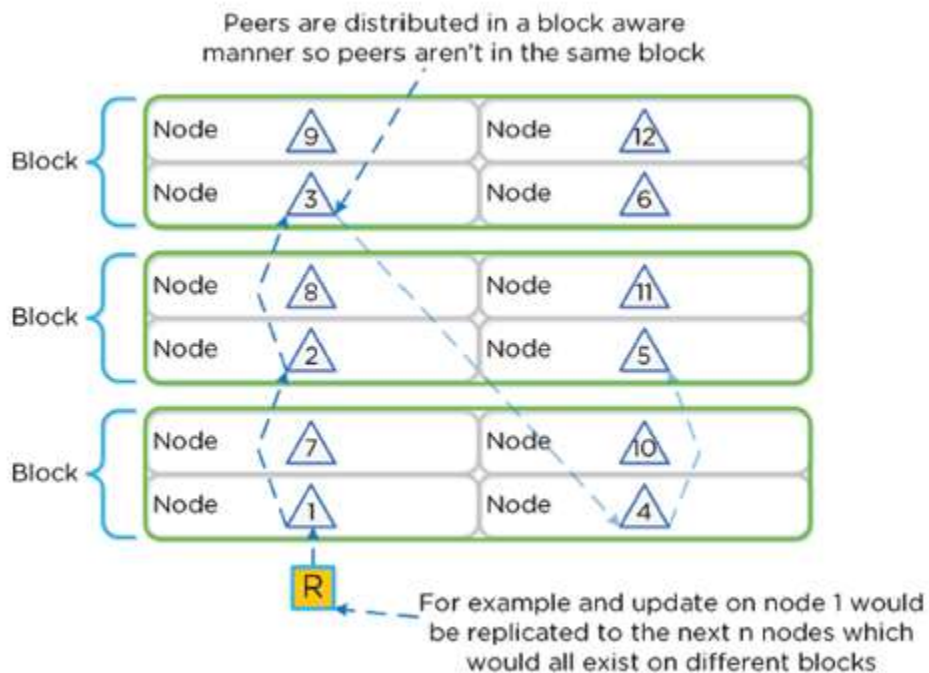


Image credit: <https://nutanixbible.com>

- If a block failure occurs, there are at least two copies of data available

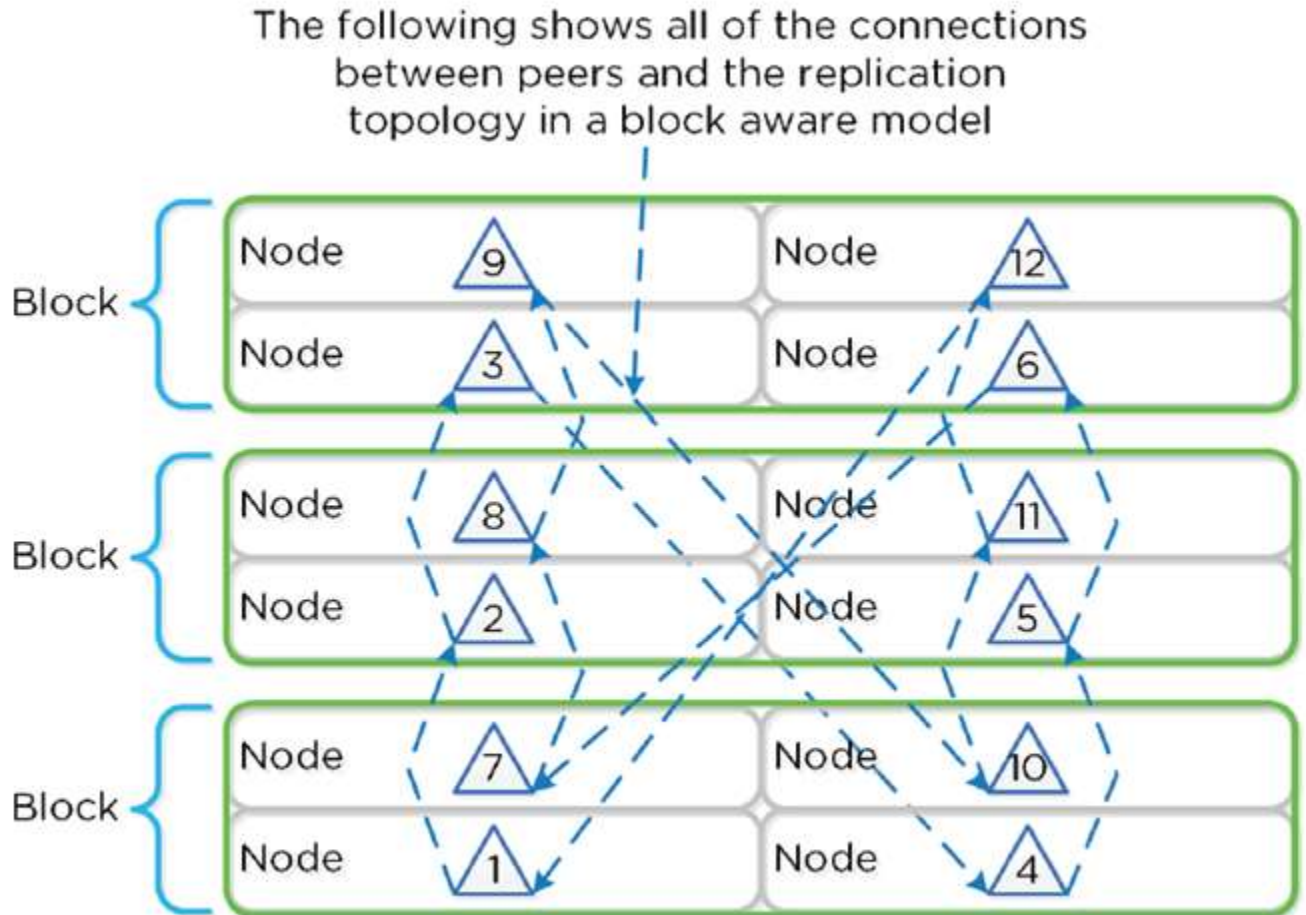


Figure 3.2-14. Full Cassandra Node Block Aware Placement

Image credit: <https://nutanixbible.com>

## Metadata Awareness Conditions

### Fault Tolerance Level 1 (FT1):

Data is Replication Factor 2 (RF2) / Metadata is Replication Factor 3 (RF3)

### Fault Tolerance Level 2 (FT2):

Data is Replication Factor 3 (RF3) / Metadata is Replication Factor 5 (RF5)

# Zookeeper

- Cluster configuration manager, based on Apache Zookeeper
- Configuration including hosts/IPs/state/etc
- 3 Nodes in cluster; 1 leader
- Accessed via Zeus
- Distributed in block-aware manner
- Ensures availability in event of block failure

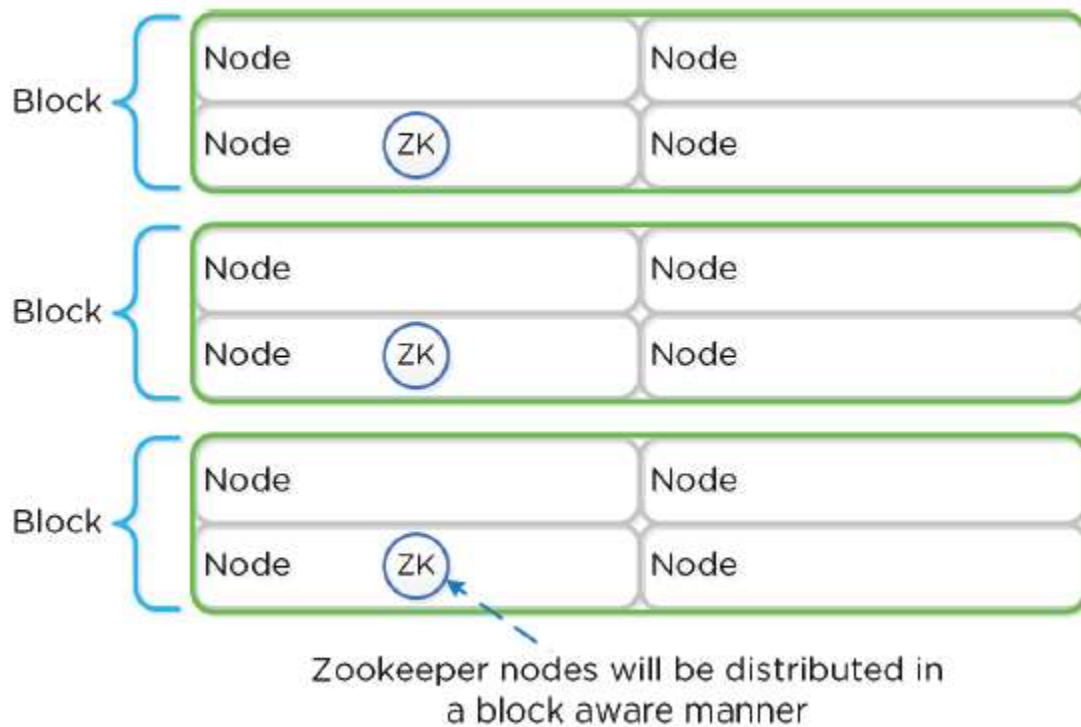


Image credit: <https://nutanixbible.com>

- In the event of a Zookeeper Outage, the ZK role transferred to another node:

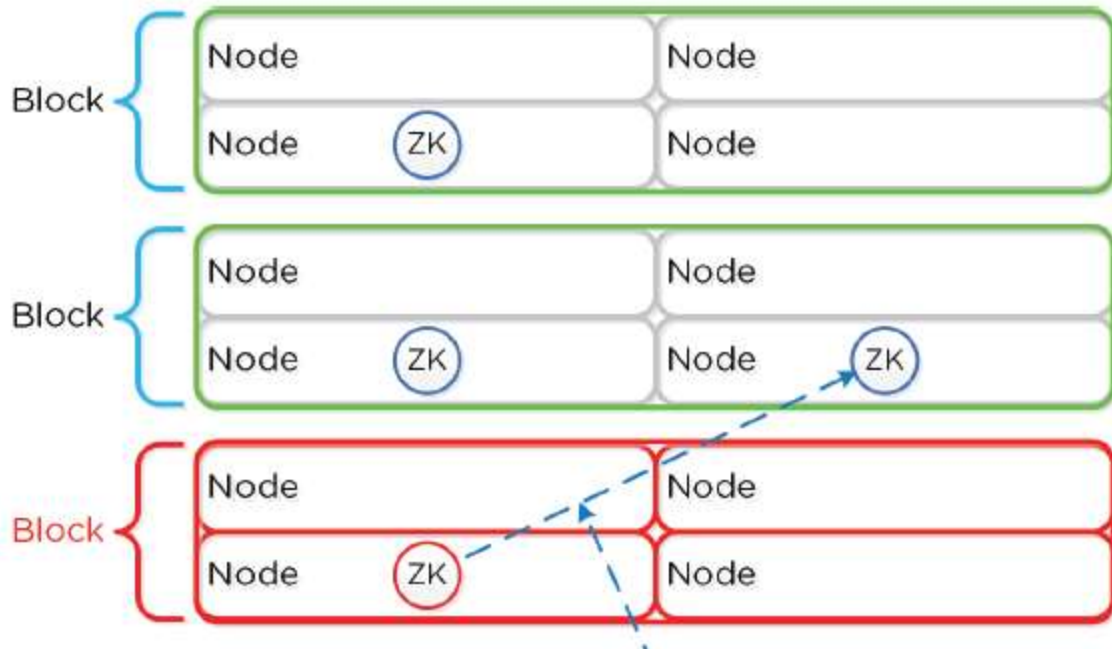


Image credit: <https://nutanixbible.com>

## Stargate

- Data I/O Manager
- Responsible for all data management and I/O
- Main interface to hypervisors (via iSCSI, NFS, SMB)
- Runs on every node in the cluster

## Curator

- Map reduce cluster management/cleanup
- Managing/distributing tasks throughout cluster
- Disk balancing/Proactive Scrubbing
- Runs on every node in the cluster; 1 leader

## Prism

- UI/API
- Runs on every node in the cluster; 1 leader

## Genesis

- Cluster Component/Service Manager



- Runs on each node, responsible for service interactions (start/stop/etc)
- Runs independently of cluster
- Requires Zookeeper

## **Chronos**

- Job/task Scheduler
- Takes jobs/tasks from Curator scan and schedules/throttle amongst nodes
- Runs on every node in the cluster; 1 leader

## **Cerebro**

- Replication/DR Manager
- Scheduling snapshots/replication to remote sites/migrations, failovers
- Runs on every node in the cluster; 1 leader
- All nodes participate in replication

## **Pithos**

- vDisk Configuration Manager
- vDisk (DSF) configuration data
- Runs on every node in the cluster

## **Foundation**

- Bootstrap, Imaging, and deployment tool for Nutanix Clusters
- Imaging process will install the desired version of the AOS software as well as the hypervisor of choice
- AHV pre-installed by default
- Foundation must be used to leverage a different hypervisor
- NOTE: Some OEMs will ship directly from the factory with the desired hypervisor

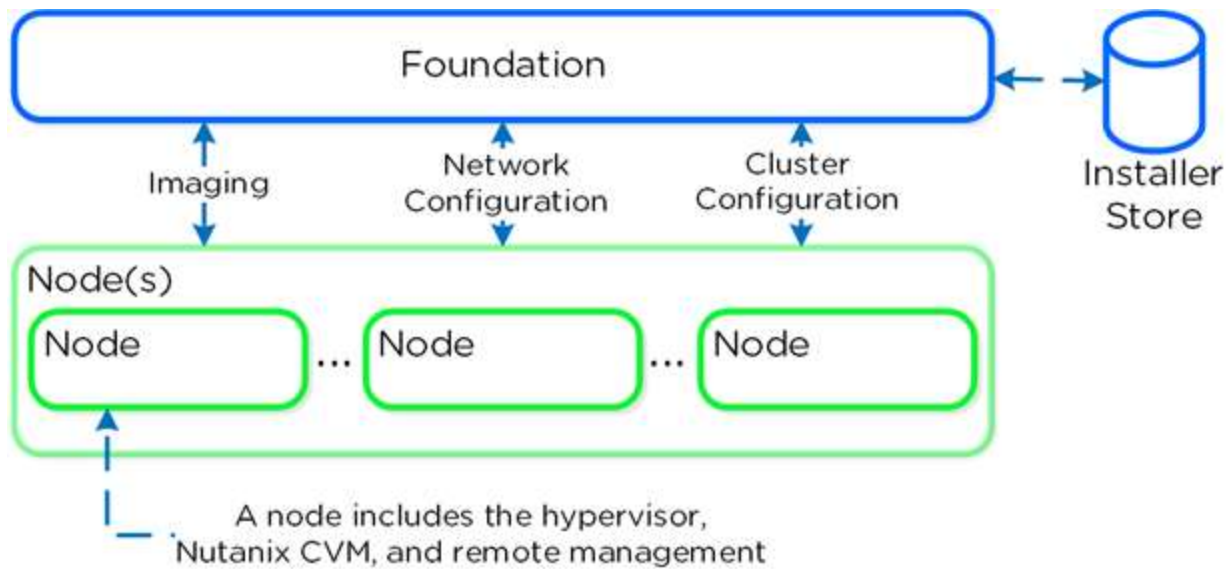


Image credit: <https://nutanixbible.com>

## Disk Partitioning

### SSD Devices

- Nutanix Home (CVM core)
- Cassandra (metadata storage)
- OpLog (persistent write buffer)
- Unified Cache (SSD cache portion)
- Extent Store (persistent storage)

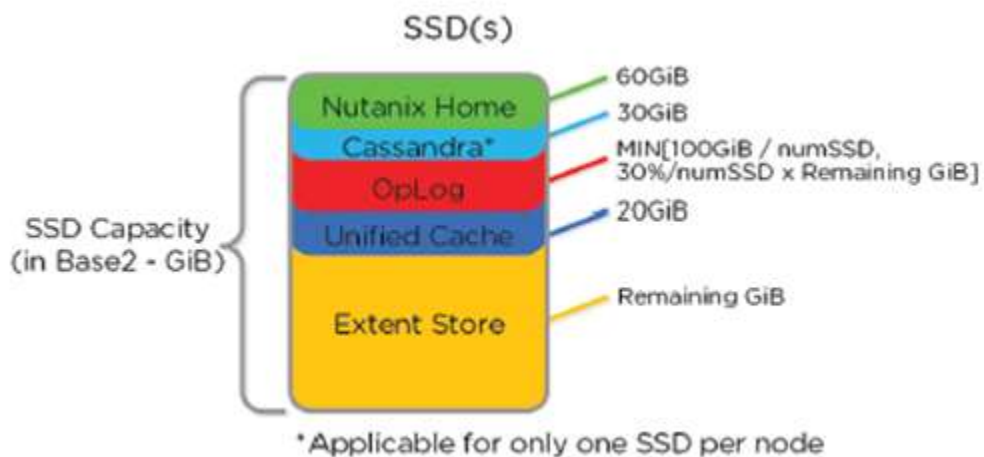


Image credit: <https://nutanixbible.com>

- Home is mirrored across first two SSD's
- Cassandra is on the first SSD
- If SSD fails, the CVM will be restarted and Casandra storage will be on second SSD

## HDD Devices

- Curator Reservation (Curator Storage)
- Extent Store (Persistent Storage)

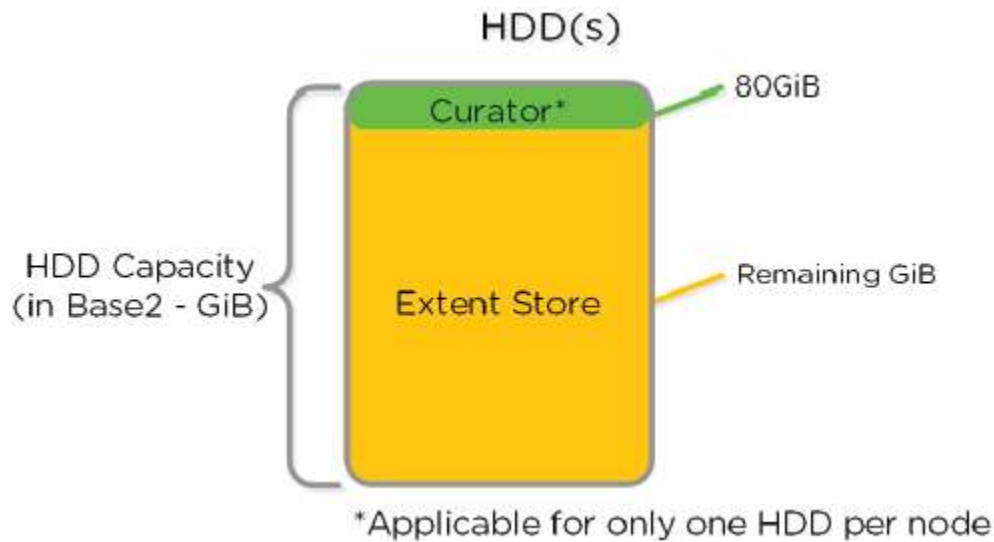


Image credit: <https://nutanixbible.com>

## Describe some of the primary AOS services running on the CVM

An Acropolis Slave runs on every CVM with an elected Acropolis Master which is responsible for task scheduling, execution, IPAM, etc. Similar to other components which have a Master, if the Acropolis Master fails, a new one will be elected.

The role breakdown for each can be seen below:

### Acropolis Master

- Task scheduling & execution
- Stat collection / publishing

- Network Controller (for hypervisor)
- VNC proxy (for hypervisor)
- HA (for hypervisor)

## Acropolis Slave

- Stat collection / publishing
- VNC proxy (for hypervisor)

An Acropolis Master is elected per cluster and is responsible for task scheduling, HA, VNC proxy, etc.

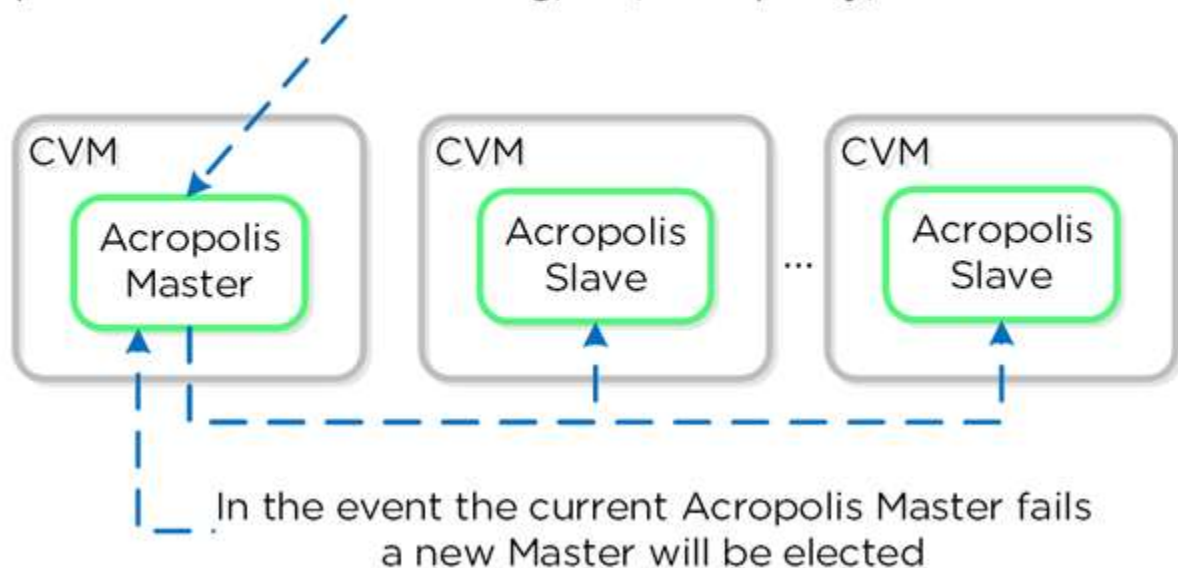


Image credit: <https://nutanixbible.com>

## Dynamic Scheduler

Efficient scheduling of resources is critical to ensure resources are effectively consumed. The Acropolis Dynamic Scheduler extends the traditional means of scheduling that relies upon compute utilization (CPU/MEM) to make placement decisions. It leverages compute, as well as storage and others to drive VM and volume (ABS) placement decisions. This ensures that resources are effectively consumed and end-user performance is optimal.

Resource scheduling can be broken down into two key areas:

- Initial placement
  - Where an item is scheduled at power-on

- Runtime Optimization
  - Movement of workloads based upon runtime metrics

The original Acropolis Scheduler had taken care of the initial placement decisions since its release. With its release in AOS 5.0, the Acropolis Dynamic Scheduler expands upon this to provide runtime resources optimization.

The figure shows a high-level view of the scheduler architecture:

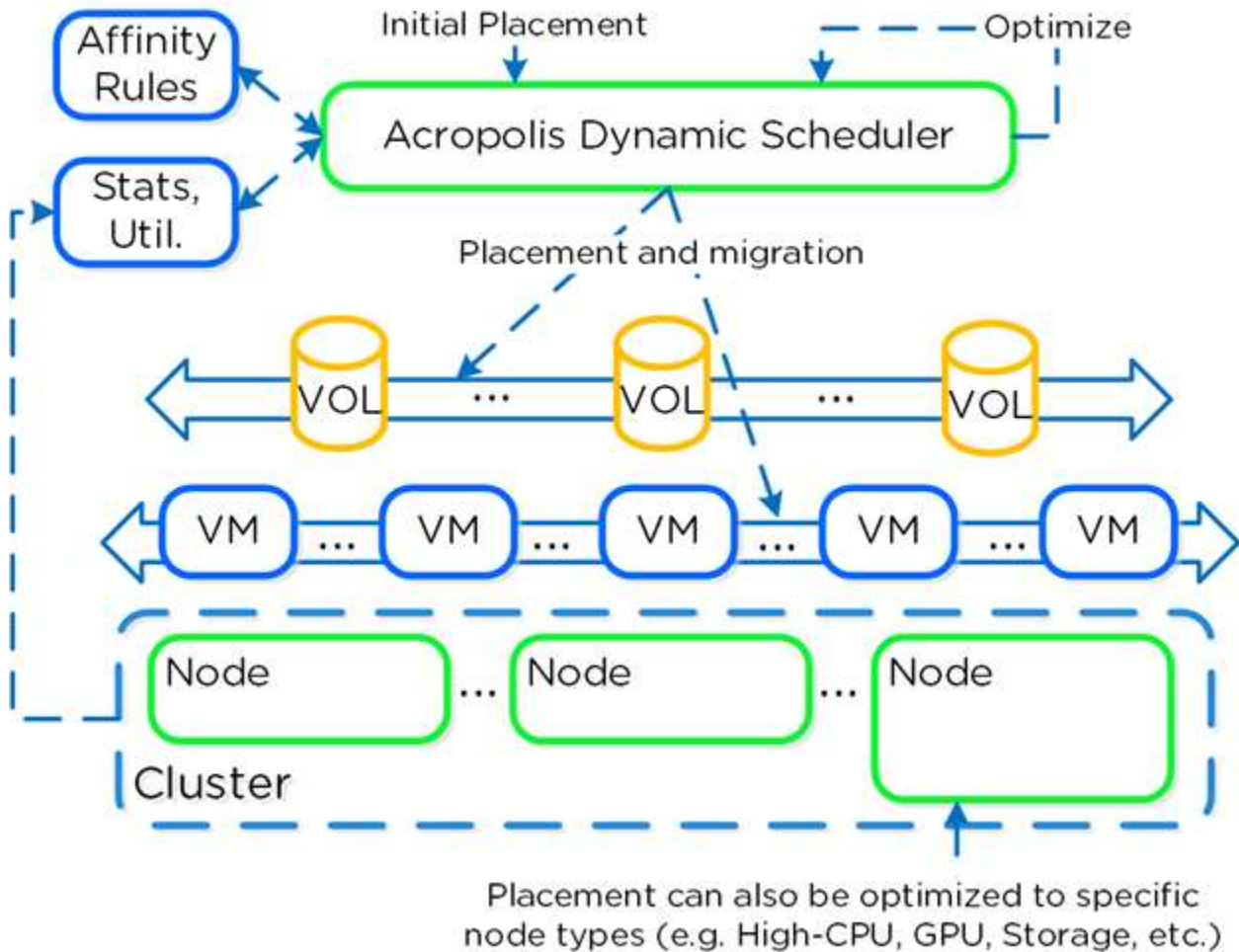


Image credit: <https://nutanixbible.com>

The dynamic scheduler runs consistently throughout the day to optimize placement (**currently every 15 minutes** | Gflag: `lazan_anomaly_detection_period_secs`). Estimated demand is calculated using historical utilization values and fed into a smoothing algorithm. This estimated demand is what is used to determine movement, which ensures a sudden spike will not skew decisions.

# Section 2 – Managing a Nutanix Cluster

## Identify methods for managing a Nutanix Enterprise Cloud

### Interfaces

HTML5 UI, REST API, CLI, PowerShell CMDlets, etc.

### Management

Policy definition and compliance, service design and status, analytics and monitoring

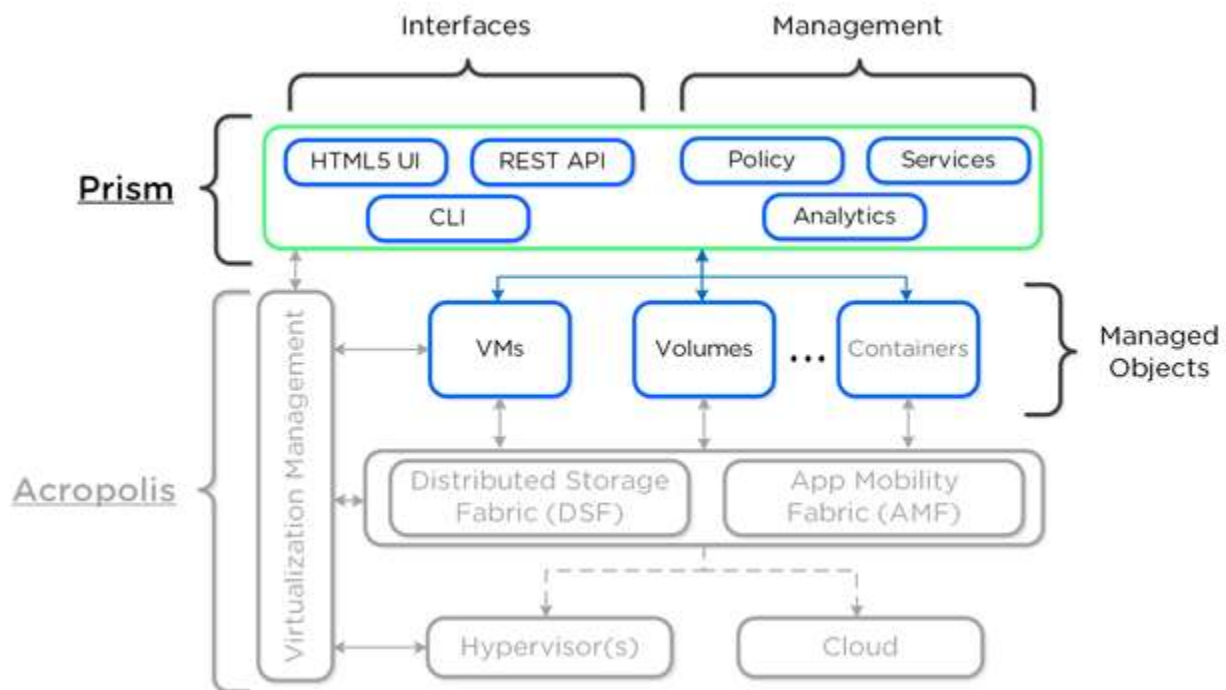


Image credit: <https://nutanixbible.com>

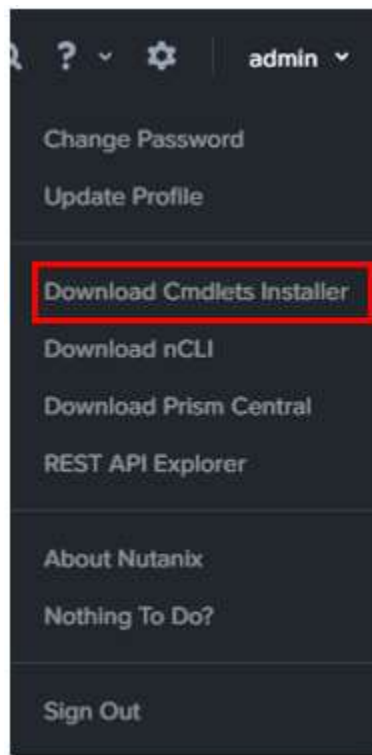
## Identify how to download and configure tools and applications like Prism Central, Cmdlets, and REST API

# Prism Central

Coming Soon

## PowerShell cmdlets

1. Sign in to the Nutanix web console.
2. Click the user icon in the upper-right corner of the web console and select Download Cmdlets Installer.



After the installer completes downloading, double-click the installer and follow the prompts.

The cmdlets are installed and a desktop shortcut NutanixCmdlets is created. Double-click the shortcut to start a PowerShell window that has the Nutanix cmdlets loaded.

## Load Nutanix Snapin

```
if ( (Get-PSSnapin -Name NutanixCmdletsPSSnapin -  
ErrorAction SilentlyContinue) -eq $null )
```

```
{  
  
    Add-PsSnapin NutanixCmdletsPSSnapin  
  
}
```

## List Nutanix CMDlets

```
Get-Command | Where-Object{$_ .PSSnapin.Name -eq "NutanixCmdletsPSSnapin"}
```

## Connect to an Acropolis Cluster

```
Connect-NutanixCluster -Server $server -UserName "myuser" -Password (Read-  
Host "Password: " -AsSecureString) -AcceptInvalidSSLCerts
```

## Get Nutanix VMs matching a certain search string

```
$searchString = "myVM"  
  
$vms = Get-NTNXVM | where {$_.vmName -match $searchString}
```

## Interactive

```
Get-NTNXVM | where {$_.vmName -match "myString"}
```

## Interactive and formatted

```
Get-NTNXVM | where {$_.vmName -match "myString"} | ft
```

## Get Nutanix vDisks

```
$vdisk = Get-NTNXVDisk
```

## Interactive

```
Get-NTNXVDisk
```

## Interactive and formatted



```
Get-NTNXVDisk | ft
```

## **Get Nutanix Containers**

```
$containers = Get-NTNXContainer
```

## **Interactive**

```
Get-NTNXContainer
```

## **Interactive and formatted**

```
Get-NTNXContainer | ft
```

## **Get Nutanix Protection Domains**

```
$pds = Get-NTNXProtectionDomain
```

## **Interactive**

```
Get-NTNXProtectionDomain
```

## **Interactive and formatted**

```
Get-NTNXProtectionDomain | ft
```

## **Get Nutanix Consistency Groups**

```
$cgs = Get-NTNXProtectionDomainConsistencyGroup
```

## **Interactive**

```
Get-NTNXProtectionDomainConsistencyGroup
```

## **Interactive and formatted**

## Utilize Prism Element to configure and monitor a cluster

- The web console is a graphical user interface (GUI) that allows you to monitor cluster operations and perform a variety of configuration tasks.
- Nutanix employs a license-based system to enable your entitled Nutanix features, and you can install or regenerate a license through the web console.
- You can upgrade a cluster when a new AOS release is available through the web console. You can also update other components such as disk firmware and hypervisor software.
- If you have multiple clusters, you can manage them all through a single web interface.

## Describe, differentiate and utilize nCLI/aCLI to configure and monitor a cluster

### Acropolis CLI (ACLI)

The Acropolis CLI (aCLI) is used for managing the Acropolis portion of the Nutanix

List hosts:

```
host.list
```

Create network:

```
net.create [TYPE].[ID].[VSWITCH] ip_config=[A.B.C.D]/[NN]
#Example
net.create vlan.133 ip_config=10.1.1.1/24
```

List network(s):

```
net.list
```

Create DHCP scope:

```
net.add_dhcp_pool [NET NAME] start=[START IP A.B.C.D] end=[END IP W.X.Y.Z]
```

Note: .254 is reserved and used by the Acropolis DHCP server if an address for the Acropolis DHCP server wasn't set during network creation

### Get an existing network's details:

```
net.list_vms [NET NAME]
```

```
#Example  
net.list_vms vlan.133
```

### Configure DHCP DNS servers for network:

```
net.update_dhcp_dns [NET NAME] servers=[COMMA SEPARATED DNS IPs]  
domains=[COMMA SEPARATED DOMAINS]
```

```
#Example net.set_dhcp_dns vlan.100 servers=10.1.1.1,10.1.1.2  
domains=splab.com
```

### Create Virtual Machine:

```
vm.create [COMMA SEPARATED VM NAMES] memory=[NUM MEM MB] num_vcpus=[NUM VCPU]  
num_cores_per_vcpu=[NUM CORES] ha_priority=[PRIORITY INT]
```

### Bulk Create Virtual Machine:

```
vm.create [CLONE PREFIX][STARTING INT..[END INT] memory=[NUM MEM MB]  
num_vcpus=[NUM VCPU] num_cores_per_vcpu=[NUM CORES] ha_priority=[PRIORITY  
INT]
```

```
#Example  
vm.create testVM[000..999] memory=2G num_vcpus=2
```

### Clone VM from existing:

```
vm.clone [CLONE NAME(S)] clone_from_vm=[SOURCE VM NAME]
```

```
#Example  
vm.clone testClone clone_from_vm=MYBASEVM
```

### Bulk Clone VM from existing:

```
vm.clone [CLONE PREFIX][[STARTING INT]..[END INT]] clone_from_vm=[SOURCE VM  
NAME]
```

```
#Example  
vm.clone testClone[001..999] clone_from_vm=MYBASEVM
```

### Create disk and add to VM:

```
vm.disk_create [VM NAME] create_size=[Size and qualifier, e.g. 500G]
container=[CONTAINER NAME]
```

```
#Example
vm.disk_create testVM create_size=500G container=default
```

### Add NIC to VM:

```
vm.nic_create [VM NAME] network=[NETWORK NAME] model=[MODEL]
```

```
#Example
vm.nic_create testVM network=vlan.100
```

### Set to boot from specific disk id:

```
vm.update_boot_device [VM NAME] disk_addr=[DISK BUS]
```

```
#Example
vm.update_boot_device testVM disk_addr=scsi.0
```

### Set VM's boot device to CD-ROM:

```
vm.update_boot_device [VM NAME] disk_addr=[CD-ROM BUS]
```

```
#Example
vm.update_boot_device testVM disk_addr=ide.0
```

### Create CD-ROM with ISO:

```
vm.disk_create [VM NAME] clone_nfs_file=[PATH TO ISO] CD-ROM=true
```

```
#Example
vm.disk_create testVM clone_nfs_file=/default/ISOs/myfile.iso CD-ROM=true
```

### If a CD-ROM is already created just mount it:

```
vm.disk_update [VM NAME] [CD-ROM BUS] clone_nfs_file[PATH TO ISO]
```

```
#Example
vm.disk_update atestVM1 ide.0 clone_nfs_file=/default/ISOs/myfile.iso
```

### Detach ISO from CD-ROM:

```
vm.disk_update [VM NAME] [CD-ROM BUS] empty=true
```

### Power on VM(s):

```
vm.on
```

```
#Example
vm.on testVM
```

**Power on all VMs:**

```
vm.on *
```

**Power on all VMs matching a prefix:**

```
vm.on testVM*
```

**Power on range of VMs:**

```
vm.on testVM[0-9][0-9]
```

## Nutanix command-line interface (nCLI)

The Nutanix command-line interface (nCLI) allows you to run system administration commands against the Nutanix cluster from any of the following machines:

- **Your local machine (preferred)**
- Any Controller VM in the cluster

**Add subnet to NFS whitelist:**

```
ncli cluster add-to-nfs-whitelist ip-subnet-masks=10.2.0.0/255.255.0.0
```

**Display cluster version**

```
ncli cluster version
```

**Display hidden NCLI options:**

```
ncli helpsys listall hidden=true [detailed=false|true]
```

**List Storage Pools:**

```
ncli sp ls
```

**List containers:**

```
ncli ctr ls
```

**Create container:**

```
ncli ctr create name=[NAME] sp-name=[SP NAME]
```

## List VMs:

```
ncli vm ls
```

## List public keys:

```
ncli cluster list-public-keys
```

## Add public key to cluster:

```
ncli cluster add-public-key name=myPK file-path=~/.mykey.pub
```

## Remove public key:

```
ncli cluster remove-public-keys name=myPK
```

## Create protection domain:

```
ncli pd create name=[NAME]
```

## Create remote site:

```
ncli remote-site create name=[NAME] address-list=[Remote Cluster IP]
```

## Create protection domain for all VMs in container:

```
ncli pd protect name=[PD NAME] ctr-id=[Container ID] cg-name=[NAME]
```

## Create protection domain with specified VMs:

```
ncli pd protect name=[PD NAME] vm-names=[VM Name(s)] cg-name=[NAME]
```

## Create protection domain for DSF files (aka vDisk):

```
ncli pd protect name=[NAME] files=[File Name(s)] cg-name=[NAME]
```

## Create snapshot of protection domain:

```
ncli pd add-one-time-snapshot name=[PD NAME] retention-time=[seconds]
```

## Create snapshot and replication schedule to remote site:

```
ncli pd set-schedule name=[PD NAME] interval=[seconds] retention-policy=[POLICY] remote-sites=[REMOTE SITE NAME]
```

## List replication status:

```
ncli pd list-replication-status
```

**Migrate protection domain to remote site:**

```
ncli pd migrate name=[PD NAME] remote-site=[REMOTE SITE NAME]
```

**Activate protection domain:**

```
ncli pd activate name= [PD NAME]
```

**Enable DSF shadow clones:**

```
ncli cluster edit-params enable-shadow-clones=true
```

**Enable dedup for vDisk:**

```
ncli vdisk edit name=[VDISK NAME] fingerprint-on-write=[true/false] on-disk-dedup=[true/false]
```

**Check cluster resiliency status:**

```
# Node status
ncli cluster get-domain-fault-tolerance-status type=node

# Block status
ncli cluster get-domain-fault-tolerance-status type=rackable_unit
```

## Differentiate between Pulse and Alert technologies

### Pulse

After you have completed initial setup, created a cluster, and opened ports 80 or 8443 in your firewall, each cluster sends a **Pulse message** once every 24 hours to a Nutanix Support server by default. Each message includes cluster configuration and health status that can be used by Nutanix Support to address any cluster operation issues. Pulse HD can also send automatic alert email notifications.

### Access Requirements

In order to successfully send Pulse and alert messages from a cluster to the Nutanix support servers, Pulse requires the following access:

- Messages are sent from the Zeus (cluster configuration manager) leader, so the firewall must allow the Zeus leader IP address. Because the Zeus leader can change, it is recommended that the IP addresses for all Controller VMs in the cluster be open in the firewall.
- Required firewall rules configured to allow traffic from the cluster to Nutanix support servers.

Pulse data is collected through the following methods:

- By default, emails are sent daily through port 8443 via nsc01.nutanix.net or nsc01.nutanix.net to nos-asups@nutanix.com.
  - If a SMTP server is configured on a cluster, emails will be sent via the SMTP server to nos-asups@nutanix.com.
- Throughout the day, each Controller VM in the cluster sends Pulse data to insights.nutanix.com over port 443 using the HTTPS REST endpoint.

Pulse

Pulse is the Nutanix automated support system. Configure it below for this cluster.

Enable

EMAIL RECIPIENTS

Nutanix Support (nos-asups@nutanix.com)

Comma separated email list.

VERBOSITY

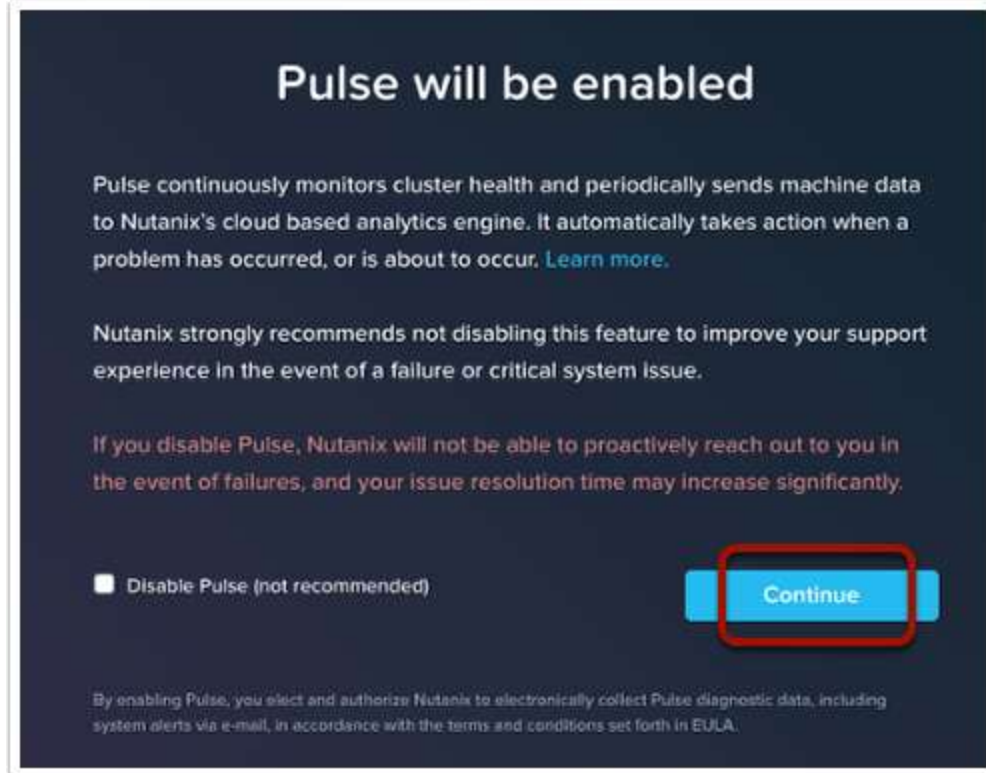
Basic CoreDump

CONNECTION STATUS

Mode	Default Nutanix Tunnel
Http Proxy	10.1.59.187
Status	<span style="color: green;">●</span> SUCCESS
Service Center	nsc02.nutanix.net
Established Since	06/21/14, 06:28:02pm

Cancel Save





Source: <https://www.virtualdennis.com/deploying-and-registering-prism-central-on-nutanix-aos-5-5/>

## Alerts

Alternately, Pulse can use your SMTP server to send messages and **alert notifications** if you have configured an SMTP server (see Configuring an SMTP Server). In this case the cluster only needs access to your SMTP server (not the Internet) to send the messages. If you do not use an SMTP server, another option is to implement an HTTP proxy as part of your overall support scheme.

Support Service	Default Setting	Options
Pulse	Enabled and ready to send status messages over customer-opened ports 80 or 8443	<ol style="list-style-type: none"> <li>1. Disabled.</li> <li>2. Enabled through customer SMTP server over customer-opened ports 80 or 8443.</li> <li>3. Enabled by implementing an HTTP proxy over customer-opened ports 80 or 8443.</li> </ol>

Alert email notifications	Enabled and ready to send alert notifications through customer-opened ports 80 or 8443	<ol style="list-style-type: none"> <li>1. Disabled.</li> <li>2. Enabled through your SMTP server over customer-opened ports 80 or 8443.</li> <li>3. Enabled by implementing an HTTP proxy over customer-opened ports 80 or 8443. You can add email accounts to also receive these notifications.</li> </ol>
Remote Support Services	Disabled	<ol style="list-style-type: none"> <li>1. Disabled.</li> <li>2. Enable a temporarily or permanently established SSH tunnel through customer-opened ports 80 or 8443.</li> <li>3. Enabled by implementing an HTTP proxy over customer-opened ports 80 or 8443.</li> </ol>

## Use the REST API Explorer to retrieve and/or make changes to a cluster

### REST API Methods

The HTTP verbs comprise a major portion of our “uniform interface” constraint and provide us the action counterpart to the noun-based resource. The primary or most-commonly-used HTTP verbs (or methods, as they are properly called) are POST, GET, PUT, PATCH, and DELETE. These correspond to create, read, update, and delete (or CRUD) operations, respectively. There are a number of other verbs, too, but are utilized less frequently. Of those less-frequent methods, OPTIONS and HEAD are used more often than others.

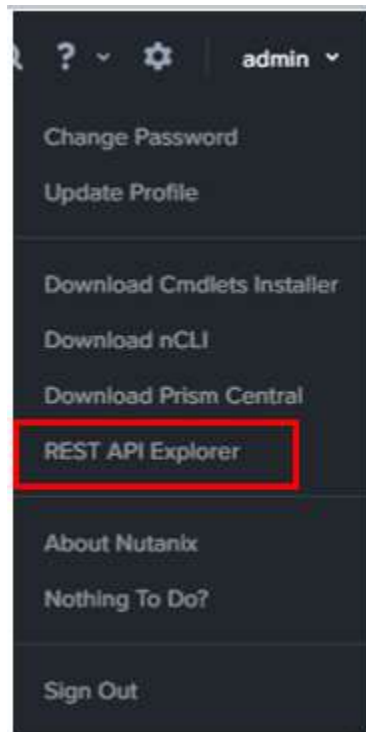
Verb	CRUD
POST	Create
GET	Read
PUT	Update/Replace
PATCH	Update/Modify
DELETE	Delete

The Nutanix REST APIs allow you to create scripts that run system administration commands against the Nutanix cluster. The API enables the

use of HTTP requests to get information about the cluster as well as make changes to the configuration. Output from the commands are returned in JSON format.

There are two versions of the Nutanix REST API.

- **v1:** The original Nutanix REST API.
- **v2:** An update of the v1 API. Users of the v1 API are encouraged to migrate to v2.



The REST API Explorer displays a list of the cluster objects that can be managed by the API. Each line has four options:

- **Show/Hide:** Expand or reduce the detail shown for the object
- **List Operations:** Show all operations that can be run on this object
- **Expand Operations:** Show the detailed view of the operations that can be run on this object

**Nutanix API V2**

[For additional information, code samples, please refer to the Nutanix Developer Portal](#)

API Version 2

**alerts**

Show/Hide | List Operations | Expand Operations

- GET /alerts/ Get the list of Alerts.
- POST /alerts/acknowledge Acknowledge Alerts.
- GET /alerts/configuration Get the Alert configuration.
- PATCH /alerts/configuration Modify the Alert configuration.
- PUT /alerts/configuration Update the Alert configuration.
- GET /alerts/policies Get all User Defined Alert Policies.
- POST /alerts/policies Create a new User Defined Alert Policy.
- PUT /alerts/policies Update a User Defined Alert Policy.
- DELETE /alerts/policies/{policy\_id} Delete the specified User Defined Alert Policy.

**/storage\_pools**

Show/Hide | List Operations | Expand Operations | Raw

GET /storage\_pools/ Get the list of Storage Pools.

**Implementation Notes**

Get the list of Storage Pools configured in the cluster.

**Parameters**

Parameter	Value	Description	Data Type
count	<input type="text"/>	Number of Storage Pools to retrieve	int
filter	<input type="text"/>	Filter criteria	string
sort	<input type="text"/>	Sort criteria	string

**Error Status Codes**

HTTP Status Code	Reason
500	Any internal exception while performing this operation

[Try it out!](#)

Try it out! [Hide Response](#)

**Request URL**

```
https://10.3.177.187:9440/PrismGateway/services/rest/v1/storage_pools/
```

**Response Body**

```
{
  "metadata": {
    "totalEntities": 1,
    "filterCriteria": "",
    "sortCriteria": "",
    "nextCursor": null,
    "previousCursor": null
  },
}
```

**Response Code**

```
200
```

**Response Headers**

```
Date: Wed, 24 Jul 2013 20:08:07 GMT
Server: Apache-Coyote/1.1
Transfer-Encoding: chunked
Content-Type: application/json
```

## REST API Status Codes

Status Code	Definition
200	The API request was successful and received a response.
201	The API request was successful and created an object.
400	The API request was malformed and could not be processed.
401	You have no access and/or are not authorized.
403	You are authorized but do not have the privileges for this API.
404	The URL was not found.
405	The called method is not allowed or is not supported.
408	The request timed out (20 seconds maximum).
500	The API request was received but there was a server error.

503	Service unavailable at this time or too early to process.
505	HTTP other than 1.1 not supported.

## Section 3 – Securing a Nutanix Cluster

### Describe how Nutanix provides cluster security

1. **User accounts** control access, and the web console allows you to set the authentication method.
2. Nutanix uses **SSL to secure communication** with a cluster, and the web console allows you to install SSL certificates.
3. Nutanix supports **key-based SSH access** to a cluster, but you have the option to disable such access.
4. Nutanix provides an option to configure the cluster for enhanced **data-at-rest security** through the use of self-encrypting drives.

### Security Policies

**Nutanix Flow** includes a policy-driven security framework that inspects traffic within the data center. The framework works as follows:

- Security policies inspect traffic that originates and terminates within a data center and help eliminate the need for additional firewalls within the data center.
- The framework uses a workload-centric approach instead of a network-centric approach. Therefore, it can scrutinize traffic to and from VMs no matter how their network configurations change and where they reside in the data center. The workload-centric, network-agnostic approach also enables the virtualization team to implement these security policies without having to rely on network security teams.
- Security policies are applied to categories (a logical grouping of VMs) and not to the VMs themselves. Therefore, it does not matter how many VMs are started up in a given category. Traffic associated with the VMs in a category is secured without administrative intervention, at any scale.
- Prism Central offers a visualization-based approach to configuring policies and monitoring the traffic to which a given policy applies.

# Types of Policies

Policy Type	Use Case
Application Security Policy	<p>Use an application security policy when you want to secure an application by specifying allowed traffic sources and destinations. This method of securing an application is typically called application ring fencing.</p> <p>For example, use an application security policy when you want to allow only those VMs in the categories department: engineering and department: customersupport (the whitelisted sources) to communicate with an issue tracking tool in the category AppType: IssueTracker (the secured application), and you want the issue tracking tool to be able to send traffic only to an integrated customer relationship management application in the category AppType: CRM.</p> <p>The secured application itself can be divided into tiers by the use of categories (the built-in AppTier category). For example, you can divide the issue tracking tool into web, application, and database tiers and configure tier-to-tier rules.</p>
Isolation Environment Policy	<p>Use an isolation environment policy when you want to block all traffic, regardless of direction, between two groups of VMs identified by their category. VMs within a group can communicate with each other.</p> <p>For example, use an isolation environment policy when you want to block all traffic between VMs in the category Environment: sandbox and VMs in the category Environment: production, and you want to allow all the VMs within each of those categories to communicate with each other.</p>
Quarantine Policy	<p>Use a quarantine policy when you want to isolate a compromised or infected VM and optionally want to subject it to forensics.</p>

## Security Policy Model

### Application-centricity

The security policy model uses an application-centric policy language instead of the more complex, traditional network-centric policy language. Configuring an application security policy involves specifying which VMs belong to the application you want to protect and then identifying the entities or networks, in the inbound and outbound directions, with which you want to allow communication.

All the entities in an application security policy are identified by the categories to which they belong and not by their IP address, VLAN, or other network attributes. After a VM is associated with a category and the category is specified in a security policy, traffic associated with the VM is monitored even if it migrates to another network or changes its IP address.

The default options for allowing traffic on the inbound and outbound directions are also inherently application centric. For application security policies, the default option for inbound traffic is a whitelist, which means that a whitelist is usually the recommended option for inbound traffic. The default option can be changed to allow all traffic. The default option in the outbound direction allows the application to send traffic to all destinations, but you can configure a destination whitelist if desired.

For forensic quarantine policies, the default option in both directions is a whitelist, but you can allow all traffic in both directions. For strict quarantine policies, no traffic is allowed in either direction.

All the VMs within a category can communicate with each other. For example, in a tiered application, regardless of how you configure tier-to-tier rules, the VMs within a given tier can communicate with each other.

## **Whitelist-Based Policy Expression**

An application security policy is expressed in terms of the categories and subnets with which you want the application to communicate and therefore, by extension, the traffic you want to allow. A more granular policy expression can be achieved by specifying which protocols and ports can be used for communication.

Any category or subnet that is not in the allowed list (the whitelist) is blocked. You cannot specify the categories and subnets you want to block because the number of such entities are typically much larger and grow at a much higher rate than the categories and subnets with which an application should be allowed to communicate. Expressing a policy in terms of allowed traffic results in a smaller, tighter policy configuration that can be modified, monitored, and controlled more easily.

## **Enforcement Modes**



All policies, whether associated with securing an application, isolating environments, or quarantining VMs, can be run in the following modes: Apply Mode Blocks all traffic that is not allowed by the policy. Monitor Mode Allows all traffic, including traffic that is not allowed by the policy. This mode enables you to visualize both allowed and disallowed traffic and fine-tune the policy before applying it.

You can switch a policy between these two modes as many times as you want.

## **Automated Enforcement**

A policy uses categories to identify the VMs to which it must apply. This model allows the automatic enforcement of a policy to VMs regardless of their number and network attributes. Connectivity between Prism Central and a registered AHV cluster is required only when creating and modifying policies, or when changing the mode of operation (applied or monitoring) of a policy. Policies are applied to the VMs in a cluster even if the cluster temporarily loses network connectivity with the Prism Central instance with which it is registered. New policies and changes are applied to the cluster when connectivity is restored.

## **Priorities Between Policies**

Prism Central does not provide a way for you to specify priorities between policies of a single type. For example, you cannot prioritize one security policy over another. There is no limit to the number of inbound and outbound rules that you can add to a security policy, allowing you to define all of an application's security requirements in a single policy. This makes priorities between policies unnecessary.

However, priorities exist between the different policy types. Quarantine policies have the highest priority followed by isolation environment policies and application security policies, in that order.

Isolation environment rules take precedence over application security rules, so make sure that isolation environment policies and application security policies are not in conflict. An isolation environment rule and an application security rule are said to be in conflict if they apply to the same traffic (a scenario that is encountered when VMs in one of the categories in the isolation environment send traffic to an application in the other category, and

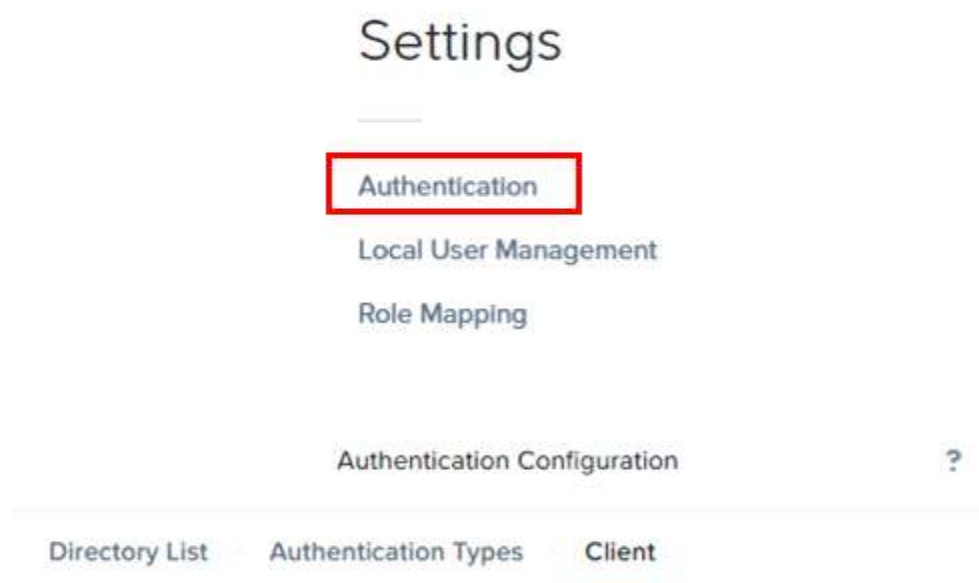
some or all of that traffic is either whitelisted or disallowed by the application security policy). The effect that an isolation environment policy has on a conflicting application security policy depends on the mode in which the isolation environment policy is deployed, and is as follows:

- If the isolation environment policy is in the applied mode, it blocks all traffic to the application, including the traffic that is whitelisted by the application security policy.
- If the isolation environment policy is in the monitoring mode, it allows all traffic to the application, including any traffic that is disallowed by the application security policy.

## Explain security concepts such as two-factor authentication, key management and cluster lockdown

### Two Factor Authentication

You can enable two-factor authentication for users through a combination of a client certificate and/or username/password to address stringent security needs.



**Configure Client Chain Certificate**

Force the authentication of all clients of the REST API (including the Prism UI and NCLI). Requires a client chain certificate.

CLIENT CHAIN CERTIFICATE

UNIVERSAL\_CA\_CHAIN.CER

Enable Client Authentication

**Configure Service Account** [Edit](#)

Configure to enable swipe access authentication instead of two-step token authentication.

DIRECTORY

AD1

SERVICE USERNAME

SERVICE PASSWORD

**Note:** Enabling CAC Authentication will also enable Client Authentication.

Enable CAC Authentication

## Key Management

Nutanix supports key-based SSH access to a cluster. Adding a key through the Prism web console provides key-based access to the cluster, Controller

VM, and hypervisor host. Each node employs a public/private key pair, and the cluster is made secure by distributing and using these keys.

You can create a key pair (or multiple key pairs) and add the public keys to enable key-based SSH access. However, when site security requirements do not allow such access, you can remove all public keys to prevent SSH access.

## Cluster Lockdown

You can easily lock down access to Nutanix clusters if your environment mandates heightened security requirements. Cluster Shield restricts access to a Nutanix cluster by disabling interactive shell logins.

Cluster lockdown is the ability to disable password based CVM access and/or only allow key based access.

### Settings


Cluster Lockdown

Data at Rest Encryption

Filesystem Whitelists

SSL Certificate

Cluster lockdown disables the ability to login to the cluster via SSH by password-challenge. However, you can add your public ssh-key via Prism and continue to login via SSH by using your ssh key. This adds a layer of non-repudiation to the connection, since the key used to access the emergency account via SSH is logged. All CVMs have a set of SSH keys that are generated at installation, so all CVMs in a cluster can still communicate with each other using keys, and you can SSH between them using keys once you gain access. ✕

 Cluster is not locked down.

Cluster lockdown makes your connection to the cluster more secure. To lock down the cluster, delete all keys in the cluster and disable remote login with password.

Enable Remote Login with Password

+ New Public Key

NAME	KEY
------	-----

No Keys Found.

## Explain Data-at-Rest Encryption (DARE) functionality

The data-at-rest encryption feature is being released with NOS 4.1 and allow Nutanix customers to encrypt storage using strong encryption algorithm and only allow access to this data (decrypt) when presented with the correct credentials, and is compliant with regulatory requirements for data at rest encryption.

Nutanix data-at-rest encryption leverages **FIPS 140-2 Level-2** validated self-encrypting drives, making it a future proof since it uses open standard protocols **KMIP** and **TCG**.

#### Encryption Type

Encrypt the cluster using:

Drive-based encryption

Use the cluster's self-encrypting drives (SEDs) to encrypt all cluster data.

Software-based encryption



Nutanix provides data-at-rest encryption via **three main options**:

- Native software-based encryption (FIPS-140-2 Level-1) \*released in 5.5
- Using self-encrypting drives (SED) (FIPS-140-2 Level-2)
- Software + hardware encryption

This encryption is configured at either the cluster or the container level, and is dependent on the hypervisor type:

- Cluster level encryption:
  - AHV, ESXi, Hyper-V
- Container level encryption:
  - ESXi, Hyper-V

## Data-at-Rest Encryption

Encrypting your cluster will help keep your information safe.



Manage Keys

**Encryption State of Cluster:** Software encryption is enabled.

Edit Configuration

Close

## Data-at-Rest Encryption

Encrypting your cluster will help keep your information safe.



Manage Keys

**Encryption State of Cluster:** Encrypt data by creating encrypted storage containers.

### Encrypted Storage Containers

alert\_test

alert\_test2

alert\_test3

All cluster data will be encrypted with software.

### Select Key Management Server (KMS)

The KMS manages the encryption keys used to encrypt data.

Cluster's local KMS

Keep your keys safe with the cluster's local KMS. Prerequisites: The local KMS can only be used via software encryption and the cluster must contain at least three nodes.

An external KMS

Configure Key Management Servers and upload SVM certificates. You will manually download and upload certificates to validate the KMS. Nutanix recommends having two or more Key Management Servers for redundancy.

Save KMS Type



### Certificate Signing Request Information

Enter the following information to generate the Certificate Signing Requests for the cluster.

EMAIL	ORGANIZATION	ORGANIZATIONAL UNIT
COUNTRY CODE	CITY	STATE

Save CSR Info

Download CSRs

### Key Management Server

Configure Key Management Servers and upload SVM certificates. Nutanix recommends having two or more Key Management Servers for redundancy.

Add New Key Management Server

### KMS CA Certificates

Configure Certificate Authorities used to validate Key Management Server authenticity. At least one certificate authority is required for encryption.

Add New Certificate Authority

< Back

Enable Encryption

## Configure user authentication

Prism currently supports integrations with the following authentication providers:

### Prism Element (PE)

- Local

- Active Directory
- LDAP

## Prism Central (PC)

- Local
- Active Directory
- LDAP
- SAML Authn (IDP)

### Settings

---

Authentication

Local User Management





Role Mapping

### Authentication Configuration

Configure one or more authentication directories to be used by the Nutanix software. You can also select the authentication types as well as configure client authentication.

Directory List · Authentication Types · Client **select**

**+ New Directory** **add** **edit**

Name	Domain	URL	
vamsi	testad.nutanix.com	ldaps://10.3.56.7 9:636	 
qa	qa.nutanix.com	ldap://10.1.59.115: 3268	 

**delete** **Close**

Directory List · **Authentication Types**

Enabled	Auth. Type
<input checked="" type="checkbox"/>	Local
<input type="checkbox"/>	Directory Service

**Directory List** · Authentication Types

Name

Domain

Directory URL

Directory Type

Connection Type

## Install an SSL certificate

Nutanix supports SSL certificate-based authentication for console access. To install a self-signed or custom SSL certificate, do the following:

### Settings

**SSL Certificate**

## SSL Certificate



**Note:** Any change made to a certificate will restart the prism session and the user will be logged out automatically.

### Installed Certificate

Organization	Nutanix Inc.
Organization Unit	Manageability
Common Name	*.nutanix.local
Location	San Jose
State	CA
Country	US
Key Type	rsa2048
Signing Algorithm	SHA256withRSA
Expiry Date	Fri Sep 07 12:00:38 UTC 2029

[Replace Certificate](#)

## SSL Certificate



---

**Note:** Any change made to a certificate will restart the prism session and the user will be logged out automatically.

- Regenerate Self Signed Certificate
- Import Key and Certificate

---

Cancel

Next

## SSL Certificate



**Note:** Any change made to a certificate will restart the prism session and the user will be logged out automatically.

**Guidelines for RSA 2048:** Please use a SHA-256 or SHA-384 signature algorithm with 2048 bit certificate for security and performance.

PRIVATE KEY TYPE

RSA 2048 bit

PRIVATE KEY 

Choose File

No file chosen

PUBLIC CERTIFICATE

Choose File

No file chosen

CA CERTIFICATE/CHAIN

Choose File

No file chosen

Cancel

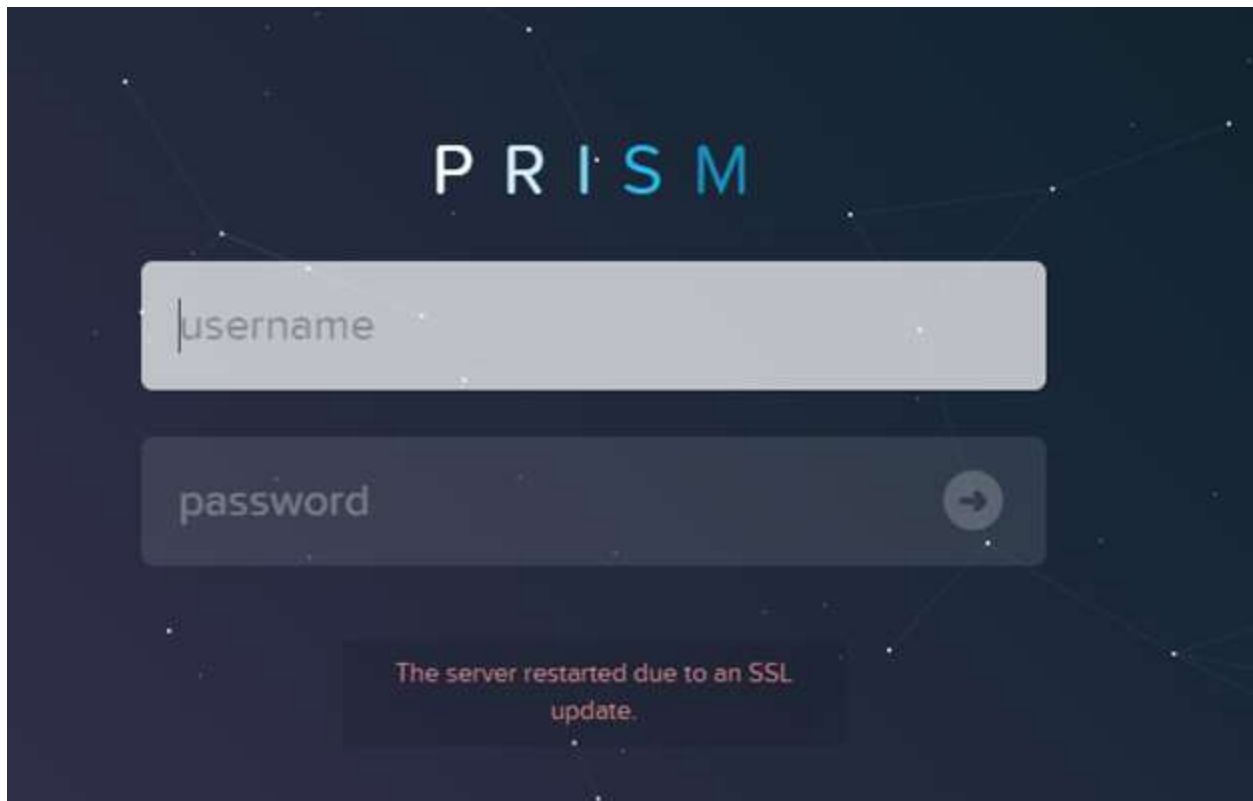
Import Files

## SSL Certificate



The certificate has been successfully updated and the server is in the process of restarting. This page will automatically refresh shortly.





## Recommended Key Configurations

Key Type	Size/Curve	Signature Algorithm
RSA	2048	SHA256-with-RSAEncryption
EC DSA 256	prime256v1	ecdsa-with-sha256
EC DSA 384	secp384r1	ecdsa-with-sha384
EC DSA 521	secp521r1	ecdsa-with-sha512

## Section 4 – Networking

### Differentiate AHV managed and unmanaged networks

A virtual network can have an IPv4 configuration, but it is not required. A virtual network with an IPv4 configuration is a **managed network**; one without an IPv4 configuration is an **unmanaged network**. A VLAN can have at most one managed network defined. If a virtual network is managed, every NIC must be assigned an IPv4 address at creation time.

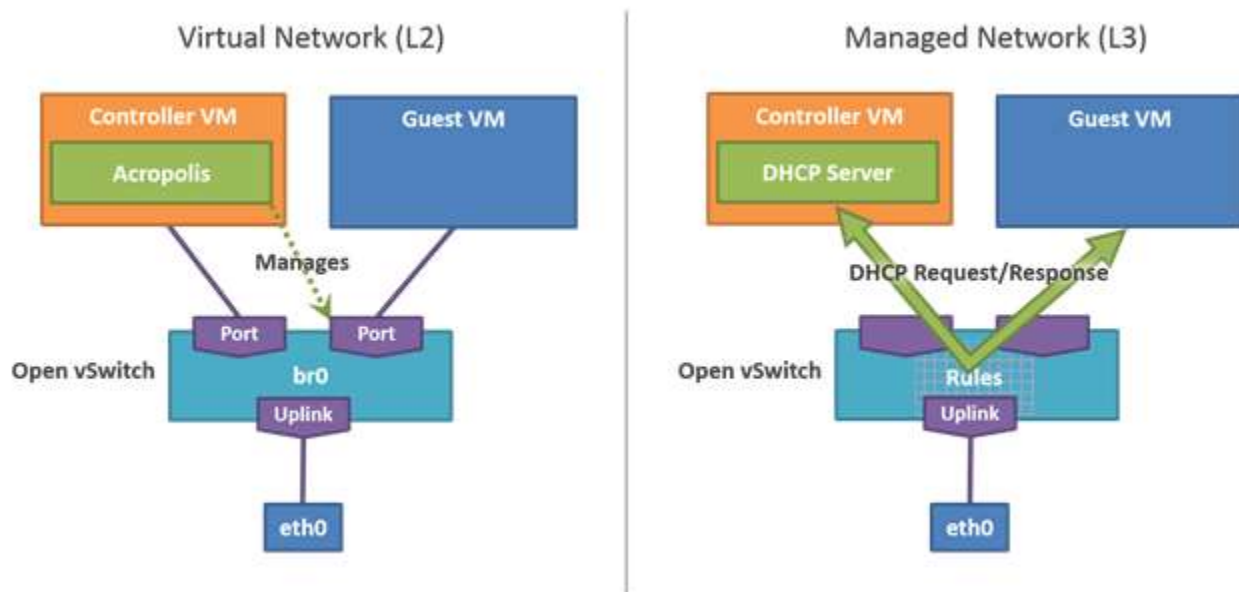


A managed network can optionally have one or more non-overlapping DHCP pools. Each pool must be entirely contained within the network's managed subnet.

If the managed network has a DHCP pool, the NIC automatically gets assigned an IPv4 address from one of the pools at creation time, provided at least one address is available. Addresses in the DHCP pool are not reserved. That is, you can manually specify an address belonging to the pool when creating a virtual adapter. If the network has no DHCP pool, you must specify the IPv4 address manually.

All DHCP traffic on the network is rerouted to an internal DHCP server, which allocates IPv4 addresses. DHCP traffic on the virtual network (that is, between the guest VMs and the Controller VM) does not reach the physical network, and vice versa.

**A network must be configured as managed or unmanaged when it is created.** It is not possible to convert one to the other.



## Describe AHV networking components and configuration settings

- No backplane for internode communication
- All I/O's handled by hypervisor on private network
- I/O is forwarded from hypervisor to CVM

- CVM replicates with other nodes with external IP over public 10GB network
- Read requests are served locally
- Typically, the only traffic on the 10G public is replication
- Occasionally CVM will forward requests in event CVM is down or data is remote or for cluster tasks such as disk balancing

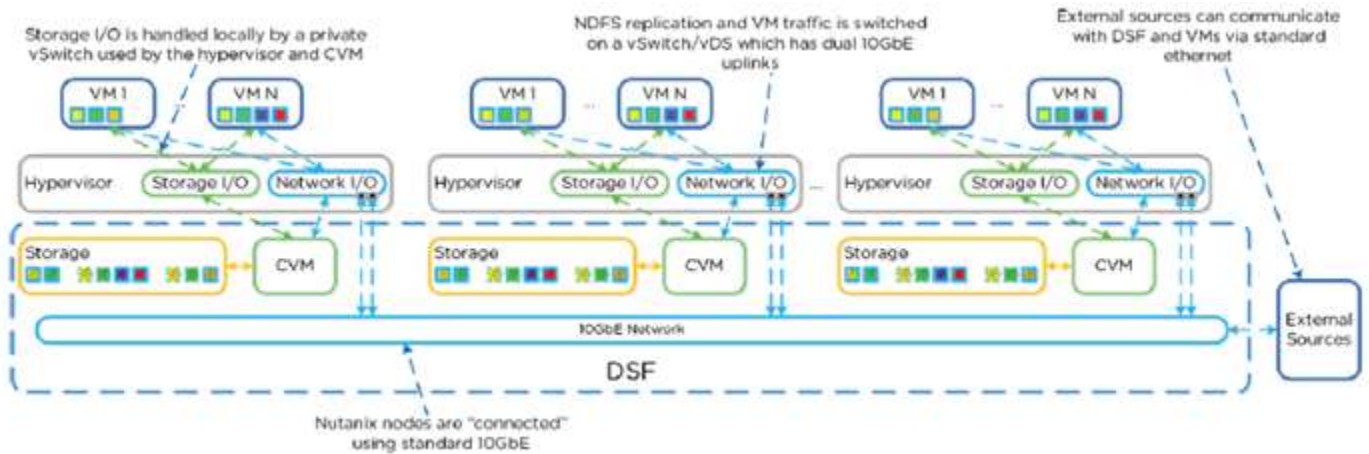
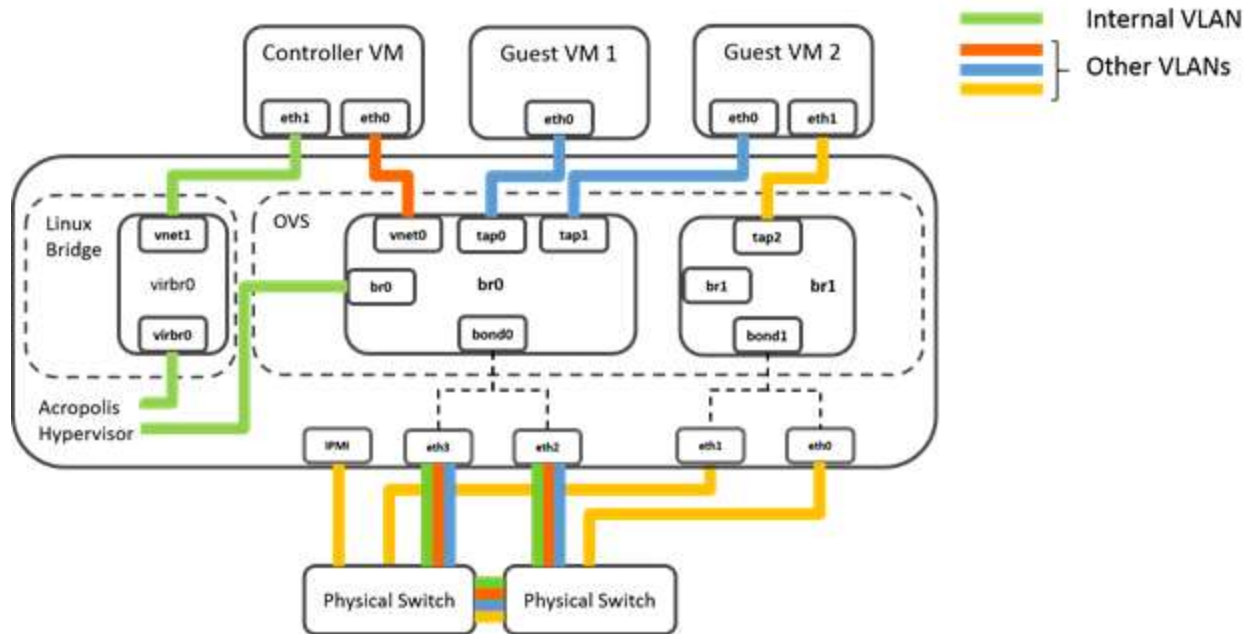


Image credit: <https://nutanixbible.com>

## Recommended Network Configuration

Network Component	Recommendations
Open vSwitch	Do not modify the OpenFlow tables that are associated with the default OVS bridge br0.
VLANs	Add the Controller VM and AHV to the same VLAN. By default, the Controller VM and the hypervisor are assigned to VLAN 0, which effectively places them on the native VLAN configured on the upstream physical switch. Do not add any other device, including guest VMs, to the VLAN to which the Controller VM and hypervisor host are assigned. Isolate guest VMs on one or more separate VLANs.
Virtual bridges	Do not delete or rename OVS bridge br0. Do not modify the native Linux bridge virbr0.
OVS bonded port (bond0)	Aggregate the 10 GbE interfaces on the physical host to an OVS bond on the default OVS bridge br0 and trunk these interfaces on the physical switch. By default, the 10 GbE interfaces in the OVS bond operate in the recommended active-backup mode.

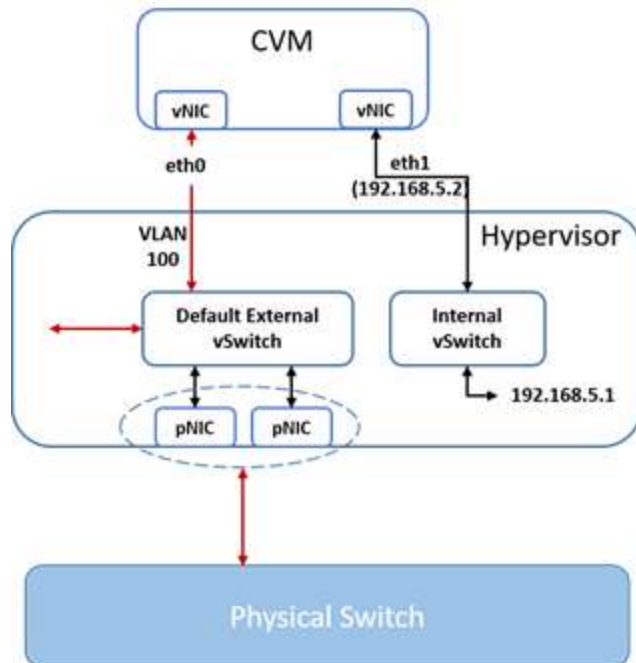
1 GbE and 10 GbE interfaces (physical host)	<p>If you want to use the 10 GbE interfaces for guest VM traffic, make sure that the guest VMs do not use the VLAN over which the Controller VM and hypervisor communicate.</p> <p>If you want to use the 1 GbE interfaces for guest VM connectivity, follow the hypervisor manufacturer's switch port and networking configuration guidelines. Do not include the 1 GbE interfaces in the same bond as the 10 GbE interfaces. Also, to avoid loops, do not add the 1 GbE interfaces to bridge br0, either individually or in a second bond. Use them on other bridges.</p>
IPMI port on the hypervisor host	Do not trunk switch ports that connect to the IPMI interface. Configure the switch ports as access ports for management simplicity.
Upstream physical switch	<p>Nutanix does not recommend the use of Fabric Extenders (FEX) or similar technologies for production use cases. While initial, low-load implementations might run smoothly with such technologies, poor performance, VM lockups, and other issues might occur as implementations scale upward (see Knowledge Base article <a href="#">KB1612</a>). Nutanix recommends the use of 10Gbps, line-rate, non-blocking switches with larger buffers for production workloads.</p> <p>Use an 802.3-2012 standards-compliant switch that has a low-latency, cut-through design and provides predictable, consistent traffic latency regardless of packet size, traffic pattern, or the features enabled on the 10 GbE interfaces. Port-to-port latency should be no higher than 2 microseconds.</p> <p>Use fast-convergence technologies (such as Cisco PortFast) on switch ports that are connected to the hypervisor host.</p> <p>Avoid using shared buffers for the 10 GbE ports. Use a dedicated buffer for each port.</p>
Physical Network Layout	<p>Use redundant top-of-rack switches in a traditional leaf-spine architecture. This simple, flat network design is well suited for a highly distributed, shared-nothing compute and storage architecture.</p> <p>Add all the nodes that belong to a given cluster to the same Layer-2 network segment.</p> <p>Other network layouts are supported as long as all other Nutanix recommendations are followed.</p>
Controller VM	Do not remove the Controller VM from either the OVS bridge br0 or the native Linux bridge virbr0.



## Explain and implement network segmentation

### Unsegmented Network

In the default, unsegmented network in a Nutanix cluster, the Controller VM has two virtual network interfaces—eth0 and eth1. Interface eth0 is connected to the built-in external virtual switch, which is in turn connected to the external network through a bond or NIC team that contains the host’s physical uplinks. Interface eth1 is connected to an internal network that enables the CVM to communicate with the hypervisor. In this network, all traffic, whether backplane traffic or management traffic, uses interface eth0. These interfaces are on the default VLAN on the virtual switch.



	Internal vSwitch	External vSwitch
AHV	virbr0	br0
ESXi	vSwitchNutanix	vSwitch0
Hyper-V	InternalSwitch	ExternalSwitch

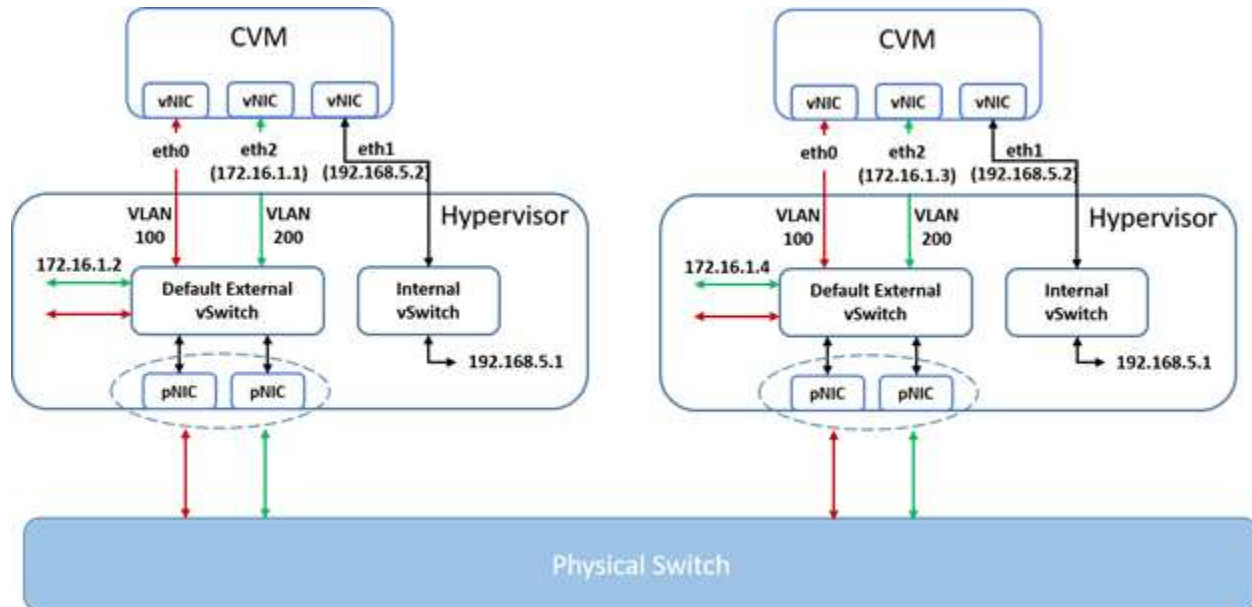
Network segmentation is not supported in the following configurations:

- Clusters on which the CVMs have a manually created eth2 interface.
- Clusters on which the eth2 interface on one or more CVMs have been assigned an IP address manually. During an upgrade to an AOS release that supports network segmentation, an eth2 interface is created on each CVM in the cluster. Even though the cluster does not use these interfaces until you configure network segmentation, you must not manually configure these interfaces in any way.
- ESXi clusters in which the CVM is connected to a VMware distributed virtual switch.
- Clusters that have two (or more) vSwitches or bridges for CVM traffic isolation. In this release, the CVM management network (eth0), which carries user VM traffic, and the CVM backplane network (eth2) must reside on a single vSwitch or bridge. These CVM networks cannot be placed on separate vSwitches or bridges.

## Segmented Network

In a segmented network, management traffic uses interface eth0 and the backplane traffic uses interface eth2. The backplane network uses either the default VLAN or, optionally, a separate VLAN that you specify when segmenting the network.

See [Create a Backplane Network](#)



Network segmentation is supported in the following environment:

- The hypervisor must be one of the following:
  - AHV
  - ESXi
  - Hyper-V
- The AOS version must be 5.5 or later.
- RDMA requirements:
  - Network segmentation is supported with RDMA for AHV and ESXi hypervisors only.
  - For the NX-9030-G5 platform, each node must have two Mellanox CX-3 Pro network cards.
  - For G6 platforms, each node must have two Mellanox CX-4 network cards. (For this reason, RDMA is not supported on platforms that have only one NIC per node.)
- The Controller VM interfaces eth0 and eth2 must be configured as follows:
  - On AHV, both interfaces must be on bridge br0.
  - On ESXi, both interfaces must be on vSwitch0 and must have the same physical network adapters (vmnic#) as uplinks.
- In all cases, the network adapters must be 10 GbE adapters.

You can segment the network on an existing cluster by using the Prism web console. The network segmentation process creates a separate network for backplane communications on the existing default virtual switch and places

the eth2 interfaces (that are created on the CVMs during upgrade) and the host interfaces on the newly created network. From the specified subnet, IP addresses are assigned to each new interface. Two IP addresses are therefore required per node. If you specify the optional VLAN ID, the newly created interfaces are placed on the VLAN. A separate VLAN is highly recommended for the backplane network to achieve true segmentation.

## Explain how to separate 1GbE and 10GbE interfaces

To avoid running any traffic on the 1 Gbps NICs there are two simple commands available which will remove all the 1 Gbps NICs from the bond.

```
nutanix@cvm$ allssh manage_ovs -interfaces 10g update_uplinks
nutanix@NTNX- -C-CVM:192.168. :~$ manage_ovs --interfaces 10g update_uplinks
2017-02-03 12:55:07 INFO manage_ovs:281 Bridge not specified. Using br0 by default.
2017-02-03 12:55:07 INFO manage_ovs:346 Deleting OVS ports: br0-up
2017-02-03 12:55:07 INFO manage_ovs:359 Adding bonded OVS ports: eth3 eth2
2017-02-03 12:55:08 INFO manage_ovs:416 Sending gratuitous ARPs for 192.168.
```

Replace interfaces with one of the following values:

- A comma-separated list of the interfaces that you want to include in the bond. For example, eth0,eth1.
- A keyword that indicates which interfaces you want to include. Possible keywords:
  - 10g. Include all available 10 GbE interfaces
  - 1g. Include all available 1 GbE interfaces
  - all. Include all available interfaces

## Identify the default AHV network configuration

The default AHV configuration includes an OVS bridge called **br0** and a native Linux bridge called **virbr0**.

The **virbr0** Linux bridge carries management and storage communication between the CVM and AHV host.

All other storage, host, and VM network traffic flows through the **br0** OVS bridge. The AHV host, VMs, and physical interfaces use **ports** for connectivity to the bridge.

```
nutanix@cvm$ manage_ovs show_interfaces
name mode link speed
eth0 1000 True 1000
eth1 1000 True 1000
eth2 10000 True 10000
eth3 10000 True 10000
```

Replace **bridge** with the name of the bridge for which you want to view uplink information. Omit the `-bridge_name` parameter if you want to view uplink information for the default OVS bridge `br0`.

```
nutanix@cvm$ manage_ovs -bridge_name bridge show_uplinks
Bridge: br0
Bond: br0-up
bond_mode: active-backup
interfaces: eth3 eth2 eth1 eth0
lacp: off
lacp-fallback: false
lacp_speed: slow

root@ahv# ovs-appctl bond/show bond0
-- bond0 --
bond_mode: active-backup
bond may use recirculation: no, Recirc-ID : -1
bond-hash-basis: 0
updelay: 0 ms
downdelay: 0 ms
lacp_status: off
active slave mac: 0c:c4:7a:48:b2:68 (eth0)
slave eth0: enabled
active slave
may_enable: true
slave eth1: disabled
may_enable: false
```

Replace `bond_name` with a name for the bond. The default value of `-bond_name` is `bond0`.

## Explain IP Address Management (IPAM)

An **unmanaged network** does not perform IPAM functions and gives VMs direct access to an external Ethernet network. Therefore, the procedure for configuring the PXE environment for AHV VMs is the same as for a physical machine or a VM that is running on any other hypervisor. VMs obtain boot file information from the DHCP or PXE server on the external network.



A **managed network** intercepts DHCP requests from AHV VMs and performs IP address management (IPAM) functions for the VMs. Therefore, you must add a TFTP server and the required boot file information to the configuration of the managed network. VMs obtain boot file information from this configuration.

## IP Address Management

- IPAM establishes DHCP scope to assigns IP's
- Leverages VXLAN and Open Flow to intercept DHCP requests

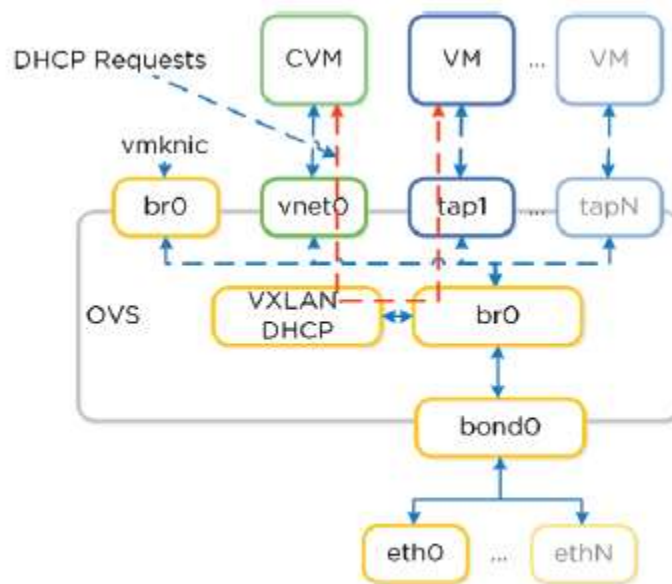


Figure 4.2.4. IPAM - Local Acropolis Master

Image credit: <https://nutanixbible.com>

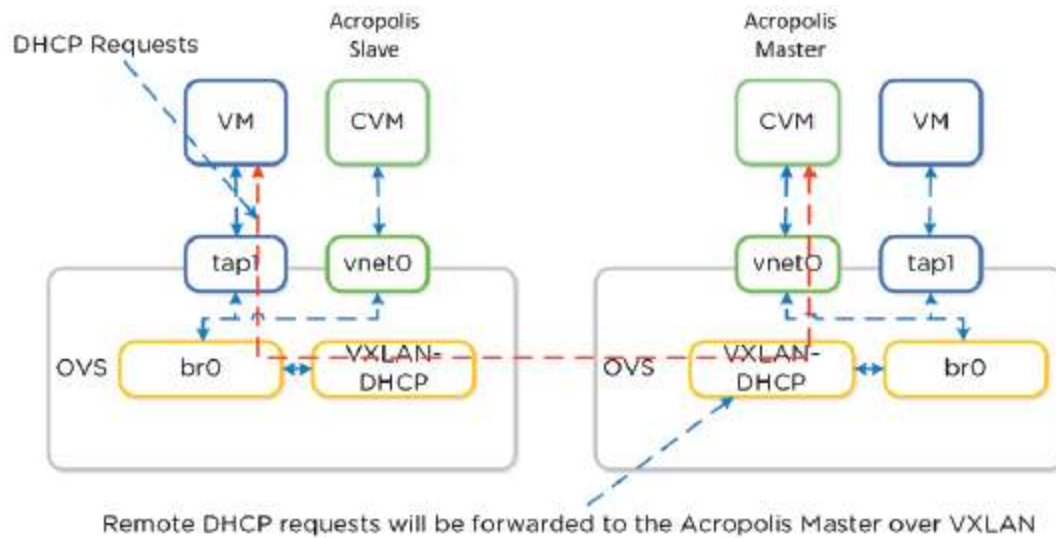


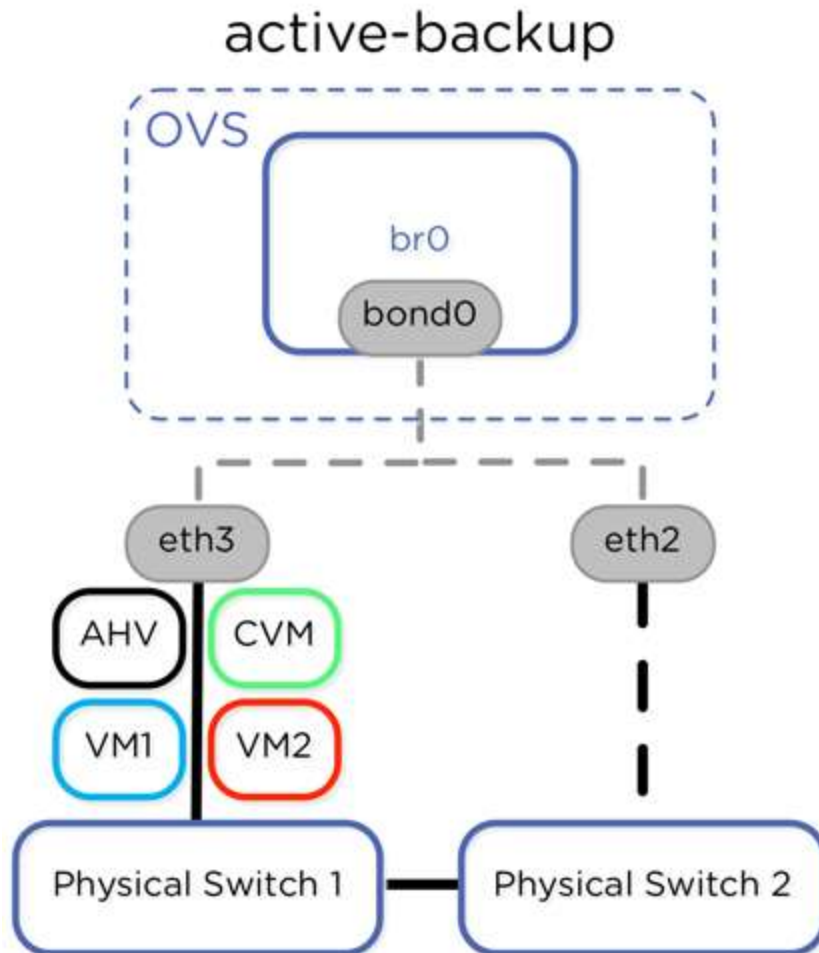
Figure 4.2.5. IPAM - Remote Acropolis Master

Image credit: <https://nutanixbible.com>

## Define and differentiate AHV Bond Modes

### Active-backup

The Active-Backup bond mode is the simplest, easily allowing connections to multiple upstream switches without any additional switch configuration. The downside is that traffic from all VMs use only the single active link within the bond. All backup links remain unused. In a system with dual 10 gigabit Ethernet adapters, the maximum throughput of all VMs running on a Nutanix node is limited to 10 Gbps.



Active-backup mode is **enabled by default**, but can be configured with the following AHV command:

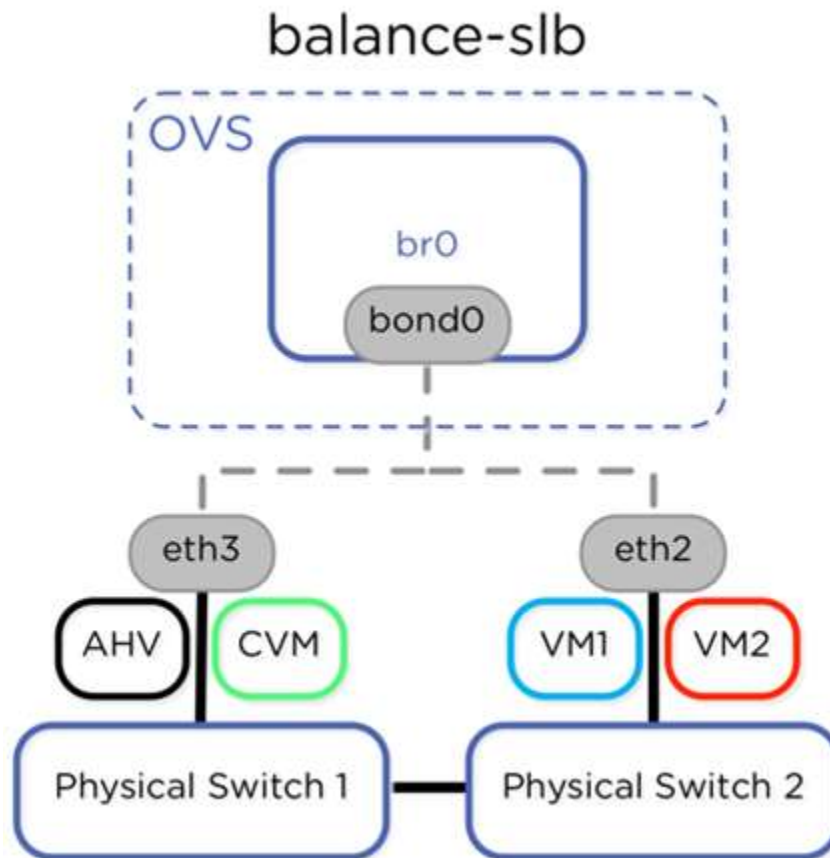
```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0 bond_mode=active-backup"
```

## Balance-slb

The Balance-slb bond mode takes advantage of the bandwidth provided by multiple upstream switch links, we **recommend** configuring the bond mode as balance-slb. The balance-slb bond mode in OVS takes advantage of all links in a bond and uses measured traffic load to rebalance VM traffic from highly used to less used interfaces. When the configurable bond-rebalance-interval expires, OVS uses the measured load for each interface and the load for each source MAC hash to spread traffic evenly among links in the bond.

Traffic from source MAC hashes may be moved to a less active link to more

evenly balance bond member utilization. Perfectly even balancing is not always possible. Each individual virtual machine NIC uses only a single bond member interface, but traffic from multiple virtual machine NICs (multiple source MAC addresses) is distributed across bond member interfaces according to the hashing algorithm. As a result, it is possible for a Nutanix AHV node with two 10 gigabit interfaces to use up to 20 gigabits of network throughput, while individual VMs have a maximum throughput of 10 gigabits per second.



The default rebalance interval is 10 seconds, but we recommend setting this to 60 seconds to avoid excessive movement of source MAC address hashes between upstream switches. We've tested this configuration using two separate upstream switches with the Acropolis hypervisor. No additional configuration (such as link aggregation) is required on the switch side, as long as the upstream switches are interconnected.

The balance-slb algorithm is configured for each bond on all AHV nodes in the Nutanix cluster with the following commands:

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0
bond_mode=balance-slb"

nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0
other_config:bond-rebalance-interval=60000"
```

Verify the proper bond mode with the following commands:

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-appctl bond/show bond0"
-- bond0 --
bond_mode: balance-slb
bond-hash-basis: 0
updelay: 0 ms
downdelay: 0 ms
next rebalance: 59108 ms
lacp_status: off

slave eth2: enabled
may_enable: true
hash 120: 138065 kB load
hash 182: 20 kB load

slave eth3: enabled
active slave
may_enable: true
hash 27: 0 kB load
hash 31: 20 kB load
hash 104: 1802 kB load
hash 206: 20 kB load
```

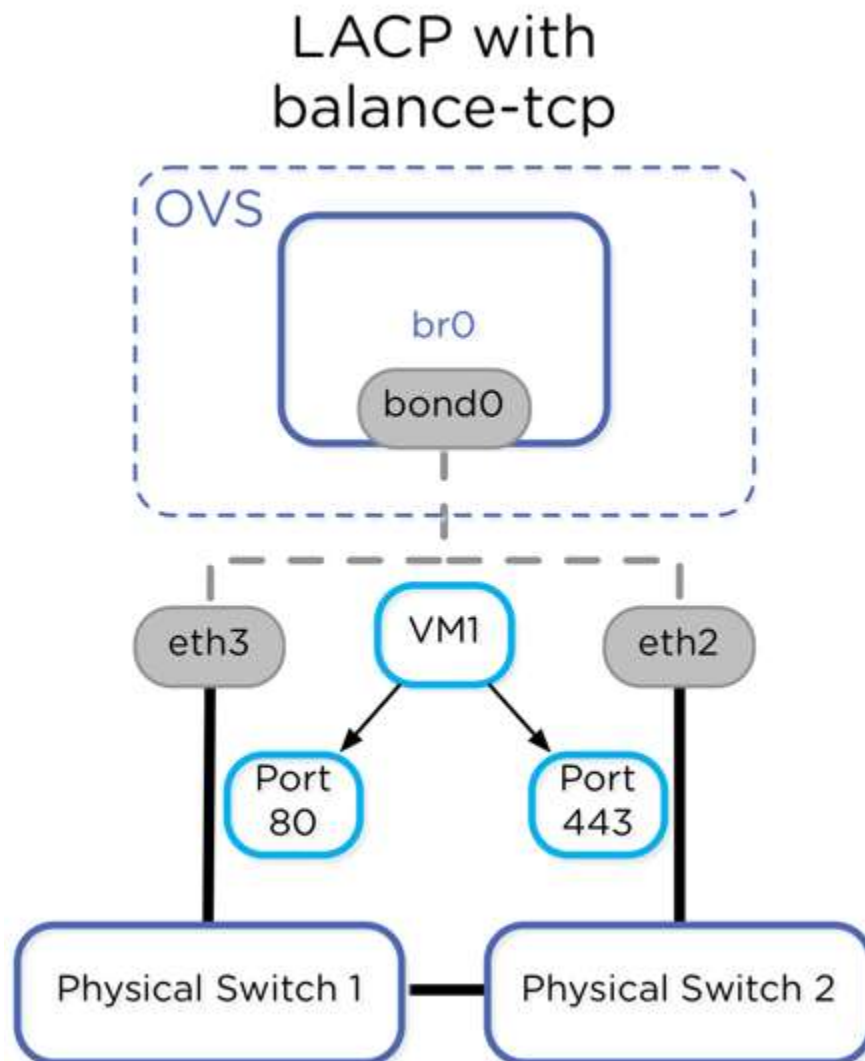
## LACP and Link Aggregation

Because LACP and balance-tcp require upstream switch configuration, and because network connectivity may be disabled if cables from AHV nodes are moved to incorrectly configured switches, Nutanix does not recommend using link aggregation or LACP.

However, to take full advantage of the bandwidth provided by multiple links to upstream switches from a single VM, link aggregation in OVS using Link Aggregation Control Protocol (LACP) and balance-tcp is required. Note that appropriate configuration of the upstream switches is also required. With LACP, multiple links to separate physical switches appear as a single Layer-2 link. Traffic can be split between multiple links in an active-active fashion based on a traffic-hashing algorithm.

Traffic can be balanced among members in the link without any regard for switch MAC address tables, because the uplinks appear as a single L2 link. We recommend using balance-tcp when LACP is configured, since multiple

Layer-4 streams from a single VM could potentially use all available uplink bandwidth in this configuration. With link aggregation, LACP, and balance-tcp, a single user VM with multiple TCP streams could potentially use up to 20 Gbps of bandwidth in an AHV node with two 10Gbps adapters.



Configure LACP and balance-tcp with the following commands. Upstream switch configuration of LACP is required.

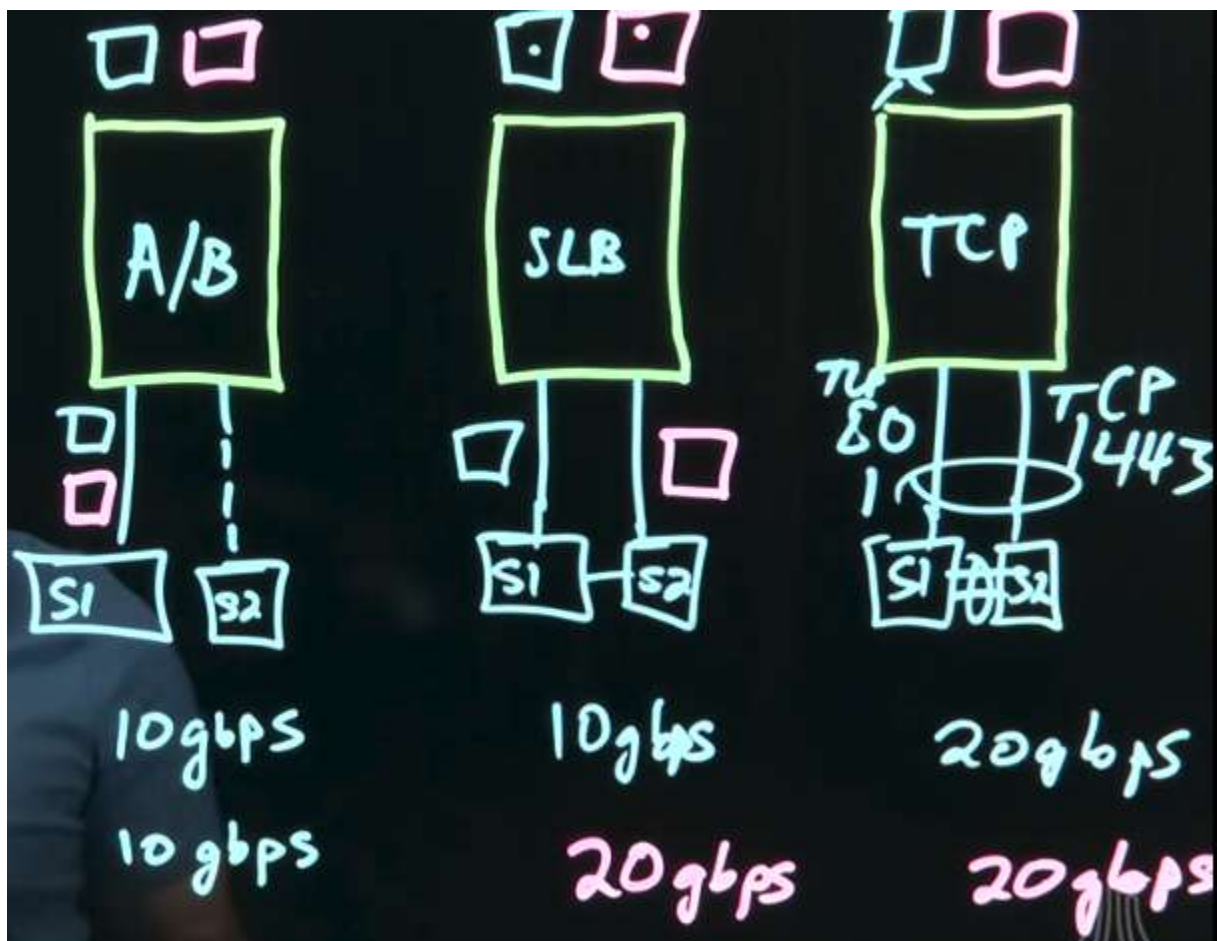
```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0 lacp=active"  
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0  
bond_mode=balance-tcp"
```

If upstream LACP negotiation fails, the default configuration is to disable the bond, which would block all traffic. The following command allows **fallback to active-backup** bond mode in the event of LACP negotiation failure.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port bond0  
other_config:lacp-fallback-ab=true"
```

## Finding the right balance

Use your virtualization requirements to choose the bond mode that's right for you! The following methods are arranged from least complex to most complex configuration. For simple and reliable failover with up to 10Gbps of host throughput with minimal switch configuration, choose active-backup. For instances where more than 10Gbps of throughput is required from the AHV host, use balance-slb. Where more than 10Gbps of throughput is required from a single VM, use LACP and balance-tcp.



# Create a Backplane Network

Backplane traffic is **intra-cluster traffic** that is necessary for the cluster to function, and comprises traffic between CVMs, traffic between CVMs and hosts, storage traffic, and so on. (For nodes that have RDMA-enabled NICs, the CVMs use a separate RDMA LAN for Stargate-to-Stargate communications.)

## Settings

- Network Configuration**
- Network Switch
- NTP Servers
- SNMP

Network Configuration ?

---

Virtual Networks    **Internal Interfaces**

NAME	INTERFACE	
Backplane LAN ⓘ	eth2	<b>Configure</b>
Hypervisor LAN	eth1	
Management LAN ⓘ	eth0	



## Create Interface



NAME	Backplane LAN
INTERFACE	eth2

SUBNET IP

NETMASK

VLAN ID

Enable VLAN ID on Physical Switch as well.

Cancel

Verify and Save

## Create a User VM Network

## Network Configuration

?

Virtual Networks

Internal Interfaces

**create**

+ Create Network

NAME	VLAN ID	
vlan0	vlan.0	 
vlan1016	vlan.1016	 

**delete**

**update**

## Create Network

?

Name

VLAN ID 

Enable IP address management

This gives AHV control of IP address assignments within the network.

Network IP Address / Prefix Length

Gateway IP Address

Configure Domain Settings

Domain Name Servers (Comma Separated)

Domain Search (Comma Separated)

Domain Name

TFTP Server Name

Boot File Name

---

IP Address Pools [?](#)

[+ Create Pool](#)

None defined.

---

Override DHCP server [?](#)

[Cancel](#) [Save](#)

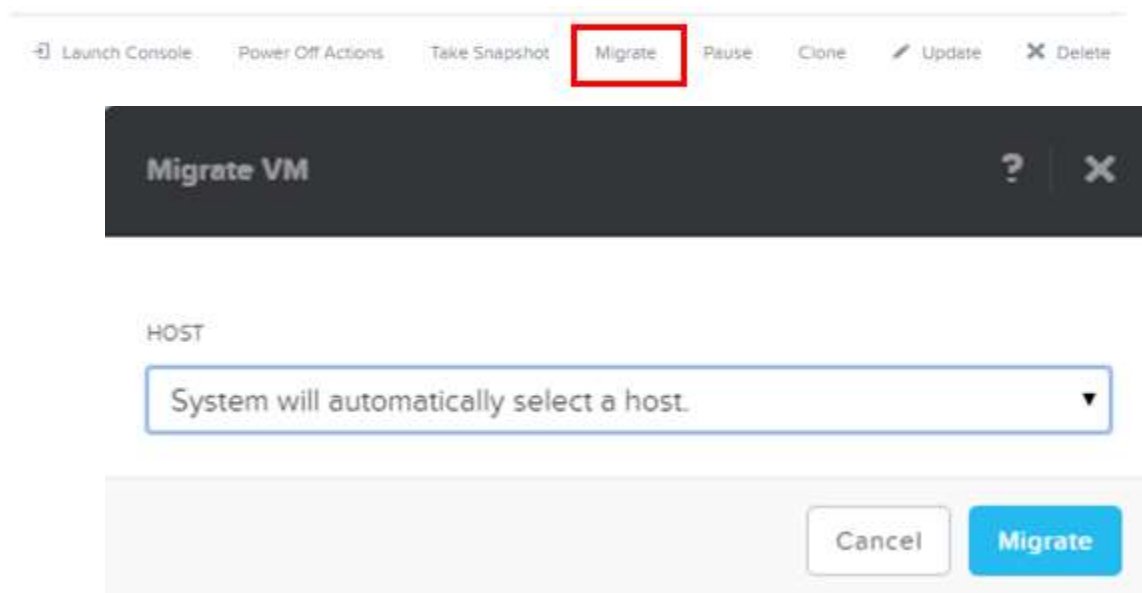
## Section 5 – VM Creation and Management

### Explain Live Migration

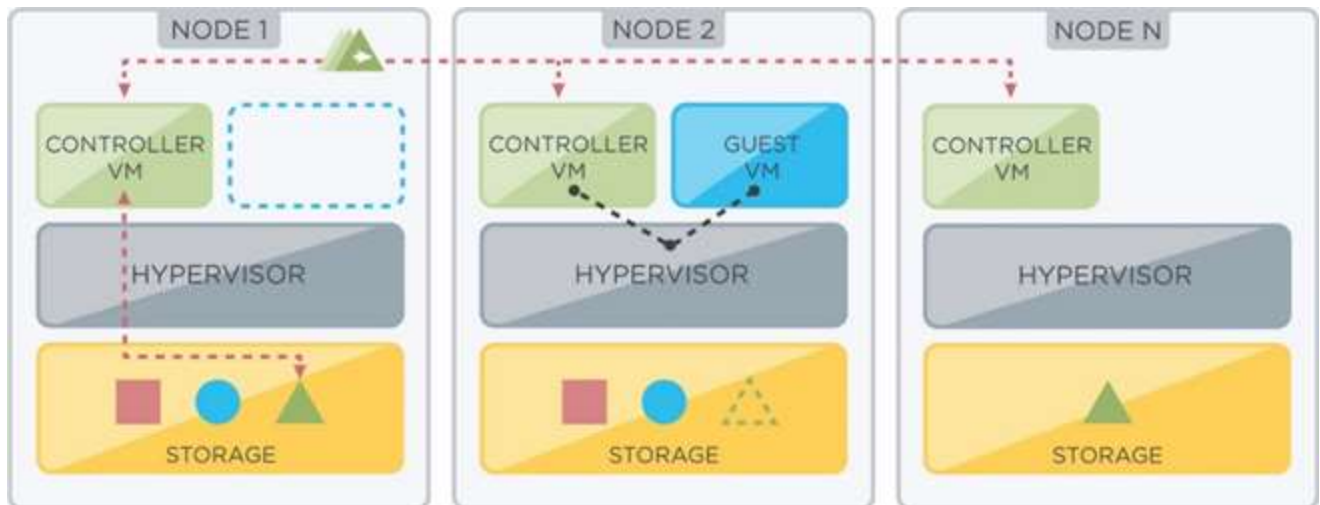
Live migration lets you move a user VM from one Acropolis host to another while the VM is powered on. This feature follows similar resource rules as VMHA to determine if migration can occur—as long as enough RAM and CPU cycles are available on the target host, live migration will initiate.

Live migration can be started with any of the following methods:

- Put the Acropolis host in maintenance mode. (VM evacuation)
- PRISM UI (VM Page)
- aCLI (automatic, targeted, or maintenance mode)
- REST API (automatic, targeted, or maintenance mode)



- Storage vMotion is not included in the Live Migrate option
- Acropolis selects a target host automatically, but you can specify a target if required.



## Describe VM High Availability functionality

### VM High Availability Modes

**Default:** This does not require any configuration and is included by default when an Acropolis Hypervisor-based Nutanix cluster is installed. When an AHV host becomes unavailable, the failed VMs that were running on the failed AHV host restart on the remaining hosts, depending on the available resources. Not all of the failed VMs will restart if the remaining hosts do not have sufficient resources.

**Guarantee:** This non-default configuration reserves space to guarantee that all failed VMs will restart on other hosts in the AHV cluster during a host failure.

### Settings

Data Resiliency

Configure Witness

Degraded Node Settings

Manage VM High Availability

Redundancy State

Enable HA Reservation

High Availability ensures that VMs can be migrated and restarted on another node in the case of a single-host failure.

Disabled: Resources are not reserved to handle high availability of VMs. In the case of a failure, VMs will be revived on a best-effort basis if resources are available.

Save

- Acropolis Master responsible for restarting VMs
- Master monitors host health by monitoring all connections to libvirt

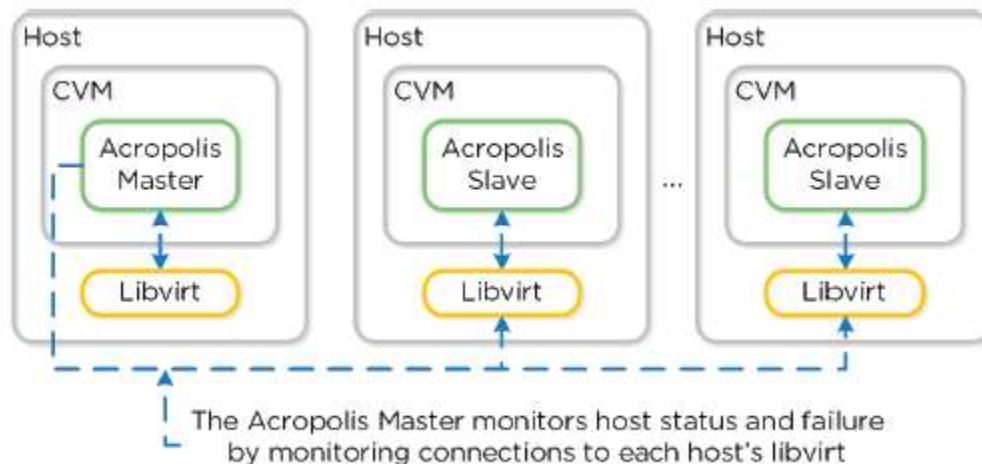


Figure 4.2.6. HA - Host Monitoring

Image credit: <https://nutanixbible.com>

- If Master is Isolated/Partitions/Failed, new Master elected.
- If cluster becomes partitioned, side with quorum will remain up and VMs restarted on those hosts.

# Restart Policy

- By default, AHV will do its best to restart VMs
- Best effort = ability to restart dependent on AHV resources
- With 5.0 and later we now only support a **segment based reservation** which is **automatically** implemented when the **Guarantee HA mode is selected**.

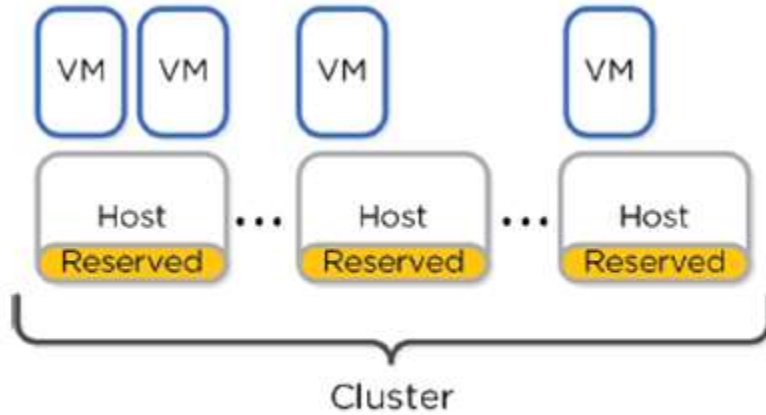


Figure 4.2.9. HA - Reserved Segment

Image credit: <https://nutanixbible.com>

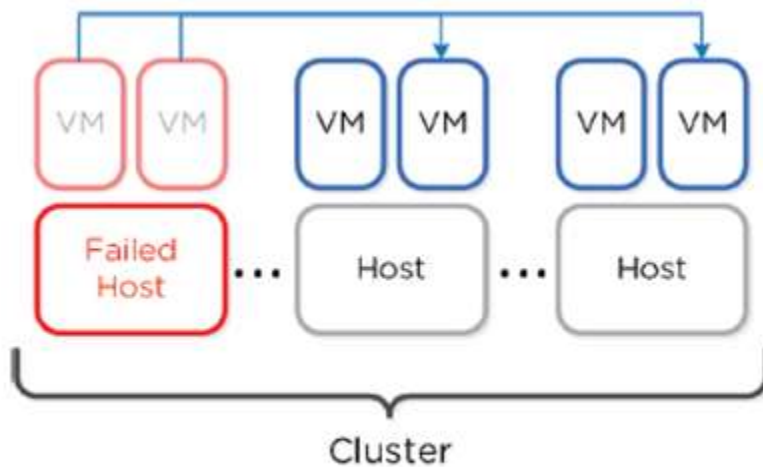


Figure 4.2.10. HA - Reserved Segment - Fail Over

Image credit: <https://nutanixbible.com>

# Reserve Segments Calculation

- Automatically calculates reserved segments per host reservation

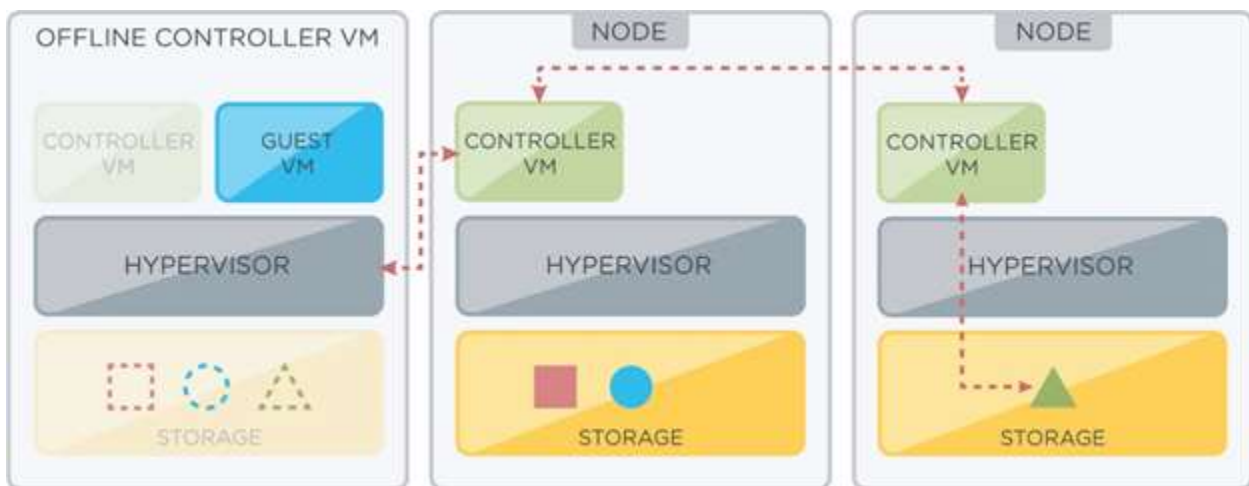
- Fixed size segments are used
- Size corresponds to largest VM in system
- Unlike VMware, AHV can pack multiple smaller VMs in single fixed segment
- Reduces fragmentation

## Describe VM Data Path Redundancy

The Nutanix cluster automatically selects the optimal path between a hypervisor host and its guest VM data. The Controller VM has multiple redundant paths available, which makes the cluster more resilient to failures.

When available, the optimal path is through the local Controller VM to local storage devices. In some situations, the data is not available on local storage, such as when a guest VM was recently migrated to another host. In those cases, the Controller VM directs the read request across the network to storage on another host through the Controller VM of that host.

Data Path Redundancy also responds when a local Controller VM is unavailable. To maintain the storage path, the cluster automatically redirects the host to another Controller VM. When the local Controller VM comes back online, the data path is returned to this VM.



## Perform guest customization on a Virtual Machine

### Cloud-init



Cloud-init is a utility that is used to customize Linux VMs during first-boot initialization. The utility must be pre-installed in the operating system image used to create VMs. Cloud-init runs early in the boot process and configures the operating system on the basis of data that you provide (user data). You can use Cloud-init to automate tasks such as setting a host name and locale, creating users and groups, generating and adding SSH keys so that users can log in, installing packages, copying files, and bootstrapping other configuration management tools such as Chef, Puppet, and Salt. For more information about Cloud-init, see <https://cloudinit.readthedocs.org/>.

## Sysprep

Sysprep is a utility that prepares a Windows installation for duplication (imaging) across multiple systems. Sysprep is most often used to generalize a Windows installation. During generalization, Sysprep removes system-specific information and settings such as the security identifier (SID) and leaves installed applications untouched. You can capture an image of the generalized installation and use the image with an answer file to customize the installation of Windows on other systems. The answer file contains the information that Sysprep needs to complete an unattended installation.

To customize a **Linux VM by using Cloud-init**, do the following:

1. Log in to the web console by using the Nutanix credentials.
2. In the VM dashboard (see VM Dashboard), do one of the following:
  1. To create a VM, click Create VM.
  2. To clone a VM, click the VM that you want to clone, and then click Clone.
3. In the Create VM or Clone VM dialog box, specify a name for the VM and allocate resources such as vCPUs, memory, and storage. Select the Custom Script check box and specify how you want to customize the VM.

For information about creating a VM and specifying customization options, see [Creating a VM \(AHV\)](#). For information about cloning a VM, see [Managing a VM \(AHV\)](#).

- In the VM dashboard, select the VM, and then click Power On. The VM is powered on and initialized based on the directives in the user data file. To create a reference image from the VM, use Image Service. See [Image](#)

Service in the VM Management chapter of the Acropolis App Mobility Fabric Guide.

CloudInit must be installed on Linux VM

## Input Formatting

Begin with #!

```
#!/bin/bash
touch /tmp/fooTest
mkdir /tmp/barFolder

#include
http://s3.amazonaws.com/path/to/script/1
http://s3.amazonaws.com/path/to/script/2

#cloud-config

# Set hostname
hostname: foobar

# Add user(s) users:
- name: nutanix
sudo: ['ALL=(ALL) NOPASSWD:ALL']
ssh-authorized-keys:
- ssh-rsa: <PUB KEY>
lock-passwd: false
passwd: <PASSWORD>

# Automatically update all of the packages
package_upgrade: true
package_reboot_if_required: true

# Install the LAMP stack packages:
- httpd
- mariadb-server
- php
- php-pear
- php-mysql

# Run Commands after execution runcmd:
- systemctl enable httpd
```

To customize a **Windows VM by using Sysprep**, you need to perform the following tasks:

- Create a reference image by using Sysprep.
- Create a VM from the reference image.

- You can also customize a VM when performing a fresh installation of Windows with an ISO file.

The Customization Process in a Nutanix Cluster: You can use Cloud-init or Sysprep both when creating and when cloning VMs in a Nutanix cluster. For unattended provisioning, you can specify a user data file for Cloud-init and an answer file for Sysprep. All Cloud-init user-data formats are supported. For example, you can use the Cloud Config format, which is written in YAML, or you can provide a multi-part archive. To enable Cloud-init or Sysprep to access the script, AOS creates a temporary ISO image that includes the script and attaches the ISO image to the VM when you power on the VM.

Note: The ISO image is mounted on bus IDE 3, so ensure that no other device is mounted on that bus.

You can also specify source paths to the files or directories that you want to copy to the VM, and you can specify the target directories for those files. This is particularly useful if you need to copy software that is needed at start time, such as software libraries and device drivers. For Linux VMs, AOS can copy files to the VM. For Windows VMs, AOS can copy files to the ISO image that it creates for the answer file.

After customizing a VM, you can copy the VDisk of the VM to Image Service for backup and duplication.

## **Perform a Self-Service Restore of a VM**

The Nutanix administrator should deploy NGT on the VM and then enable this feature. For more information on enabling and mounting NGT, see the [Enabling and Mounting Nutanix Guest Tools](#). After the feature is enabled and a disk is attached, the guest VM administrator can recover files within the guest operating system. If the guest VM administrator fails to detach the disk, it gets automatically detached from the VM after 24 hours.

Note:

- The Nutanix administrator can enable this feature for a VM only through nCLI, and in-guest actions can be performed only by using NGT.
- Only Async-DR workflow is supported for the self-service restore feature.

Verify that NGT is enabled and mounted on the VM. For more information, see Enabling and Mounting Nutanix Guest Tools.

1. To enable self-service restore, click Manage Guest Tools.
2. To enable self-service restore feature, click Self Service Restore (SSR) check box. The Self-Service Restore feature is enabled on the VM. The guest VM administrator can restore the desired file or files from the VM.
3. Click Submit.

The screenshot displays the 'Snapshots' view in vCenter, showing a list of 11 snapshots for a VM. The selected snapshot is 3290413, taken on 09-08-2016 at 10:51 AM. Below the snapshot list, the 'Disk Action' menu is open, showing a table of disks for the selected snapshot.

	DISK	ORIGINAL DRIVE LETTERS	MOUNTED DRIVE LETTERS
<input checked="" type="checkbox"/>	Disk 0	C:	

### Mounted Snapshots

Here you manage snapshot drives that are currently mounted to your virtual machine. We have correlated the original snapshot drive letters with their current drive letters.

Select All

Snapshot ID: 3293562 09-09-2016 03:06 PM

<input checked="" type="checkbox"/>	DISK	ORIGINAL DRIVE LETTERS	MOUNTED DRIVE LETTERS
<input checked="" type="checkbox"/>	Disk 0	C:	H:

Snapshot ID: 3294330 09-09-2016 04:08 PM

<input type="checkbox"/>	DISK	ORIGINAL DRIVE LETTERS	MOUNTED DRIVE LETTERS
<input type="checkbox"/>	Disk 0	C:	G:

## Nutanix Guest Tools

Enables advanced VM management

- **Guest Agent Service**
  - Self-Service Restore (SSR) aka File-Level Restore (FLR) CLI
  - VM Mobility drivers (VirtIO for AHV)
    - VirtIO contains **vNIC, iSCSI, and Balloon Driver**
  - VSS Agent and HW Provider for Windows
  - App Consistent snapshot support for Linux (via scripts to quiece)
- **Gateway Tools Service:** gateway from Acropolis and Nutanix services and Guest Agent
  - Distributed across all CVMs with elected NGT Master (runs on Prism leader)
- **Guest Agent:** deployed on VM OS.
  - Handles local functions

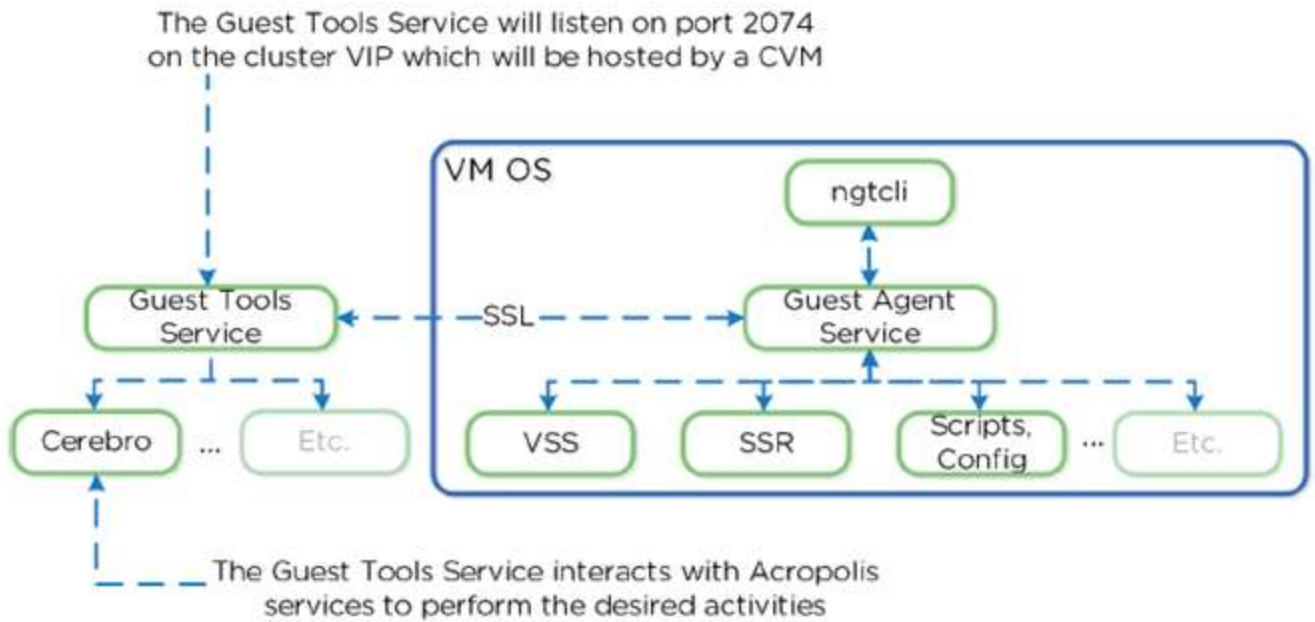


Image credit: <https://nutanixbible.com>

- **Guest Tools Service**

- NGT Master: handles requests from NGT proxy
- Interfaces with Acropolis
- Dynamically elected per cluster
- Listens internally on 2073
- NGT Proxy: runs on every CVM
- Forwards request to NGT Master
- Listens externally on 2074

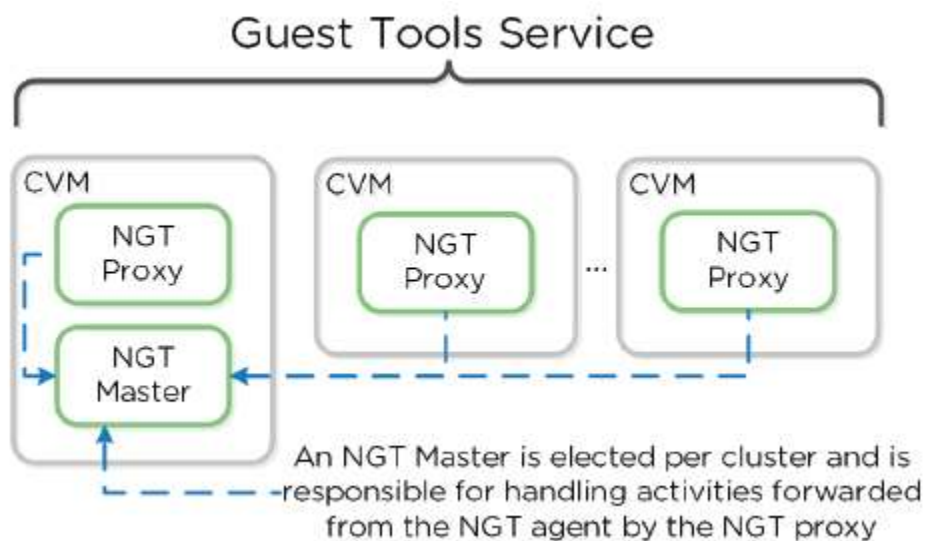


Image credit: <https://nutanixbible.com>

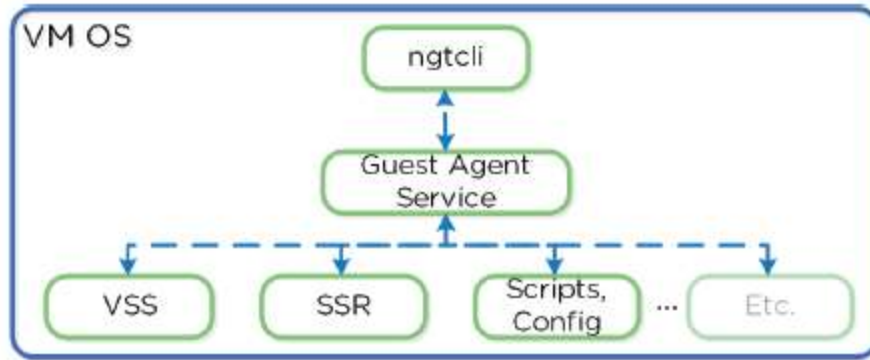


Image credit: <https://nutanixbible.com>

- Communicates via Nutanix Cluster IP via SSL
- Acts as CA (Certificate Authority)
- Responsible for generating certificate pairs for each NGT enabled UVM
- Installs similar to VMTools
- Can deploy in base image

## Use the Image Service to deploy a VM

Select the image that you have created by using the image service feature.

This field appears only when Clone from Image Service is selected. It specifies the image to copy.

## Creating the Image

### Settings

General

Cluster Details

Configure CVM

Convert Cluster

Expand Cluster

Image Configuration

## Image Configuration



Manage the images to be used for creating virtual disks.

[+ Upload Image](#)

NAME	ANNOTATION	TYPE	STATE	SIZE	
tinycore_loadrun...		DISK	ACTIVE	307.34 MiB	
tinyhrapp		DISK	ACTIVE	307.34 MiB	

## Create Image



Name

Annotation

Image Type

Not Selected



Storage Container

SelfServiceContainer



Image Source

From URL

Upload a file

No file chosen

[← Back](#)



## Creating the VM

Select **Clone from Image Service** to copy an image that you have imported by using image service feature onto the disk.

The screenshot shows the 'Update Disk' dialog box. At the top, there is a title bar with the text 'Update Disk' and icons for help and close. Below the title bar is a blue message box that says 'The CDROM is empty.' with a close icon. The main area contains four dropdown menus: 'TYPE' set to 'CDROM', 'OPERATION' set to 'CLONE FROM IMAGE SERVICE' (highlighted with a red border), 'BUS TYPE' set to 'IDE', and 'IMAGE' set to 'Windows2012R2'. At the bottom right are 'Cancel' and 'Update' buttons.

## Section 6 – Health Monitoring and Alerts

**Identify dashboards and monitoring tools that can be used to resolve cluster issues**

### Advanced Pages

CVMIP/DNS:Port

- 2009: Stargate page to monitor backend storage
  - Check QoS Queue and OpLog QoS queue (admitted/outstanding IO's)
  - Cache hit rates should be 80-90%+ if workload is ready heavy for best possible performance
  - Check Avg Latency, Avg Op Size, Avg. Outstanding

Start time	20150618-11:12:52-GMT-0700
Build version	el6-release-master-038d4c7d75cbc6ed21e64d357c94303350159807
Build last commit date	2015-05-30 14:14:03 -0700
Stargate handle	10.3.140.151:2009
iSCSI handle	10.3.140.151:3261
SVM id	7
Incarnation id	30986558
Highest allocated opid	38138394
Highest contiguous completed opid	36132415
Content cache total hits(NonDedup)	86.17%
Content cache flash pagin pct(NonDedup)	0
Content cache total hits(Dedup)	0%
Content cache flash pagin pct(Dedup)	0%
Content cache memory usage	3220 MB
Content cache physical memory usage	3224 MB
Content cache flash usage	0 MB
QoS Queue (size/admitted)	0/72
Oplog QoS queue (size/admitted)	0/0
NFS Flush Queue (size/admitted)	0/0
NFS cache usage	0 MB

Image credit: <https://nutanixbible.com>

QoS Queue (size/admitted)	0/26
Oplog QoS queue (size/admitted)	0/0

Image credit: <https://nutanixbible.com>

Content cache total hits(NonDedup)	86.17%
Content cache flash pagin pct(NonDedup)	0
Content cache total hits(Dedup)	0%
Content cache flash pagin pct(Dedup)	0%
Content cache memory usage	3220 MB
Content cache physical memory usage	3224 MB
Content cache flash usage	0 MB

Image credit: <https://nutanixbible.com>

SVM Id	IP:port	Incarnation	SSD-PCIe	SSD-SATA		DAS-SATA			
7	<a href="https://10.3.140.151:2009">10.3.140.151:2009</a>	30986558		154 (188/209)	153 (188/209)	152 (477/862)	151 (477/862)	150 (477/862)	149 (438/782)
8	<a href="https://10.3.140.152:2009">10.3.140.152:2009</a>	30174288		146 (190/209)	145 (190/209)	144 (474/862)	143 (476/862)	142 (474/862)	141 (434/782)
9	<a href="https://10.3.140.153:2009">10.3.140.153:2009</a>	30972235		50 (188/209)	49 (188/208)	48 (487/862)	47 (488/862)	44 (486/862)	43 (440/782)
10	<a href="https://10.3.140.154:2009">10.3.140.154:2009</a>	30989925		139 (189/209)	138 (189/209)	137 (483/862)	136 (482/862)	135 (484/862)	134 (432/782)
11	<a href="https://10.3.140.155:2009">10.3.140.155:2009</a>	30332545		90 (186/209)	89 (190/209)	88 (594/862)	87 (591/862)	86 (591/862)	85 (533/782)
13	<a href="https://10.3.140.157:2009">10.3.140.157:2009</a>	30813522		123 (165/209)	122 (165/209)	121 (481/862)	120 (480/862)	119 (481/862)	117 (429/782)
14	<a href="https://10.3.140.158:2009">10.3.140.158:2009</a>	30460780		78 (189/209)	77 (189/208)	76 (482/862)	75 (477/862)	74 (477/862)	73 (436/782)

Image credit: <https://nutanixbible.com>

vDisk Name	Unstable data			Outstanding ops			Ops/s			KB/s		Avg latency (usec)		Avg op size	Avg outstanding	% busy
	KB	Ops/s	KB/s	Read	Write	Read	Write	Error	Read	Write	Read	Write	Read	Write		
NFS:31181822 (55b04a56-8e98-40d4-8c0f-617a57cb0450)	0	0	0	0	6	2248	907	0	8992	3628	178	2740	4096	5	85	
NFS:31181826 (ede19589-df09-40a8-9640-6cad4e2d48bd)	0	0	0	1	5	2228	922	0	8912	3688	172	2756	4096	6	85	
NFS:31182359 (0f1ae5e0-be91-4068-9ac0-f5a3227f5c480)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
NFS:31181823 (1a8ef41e-ac3f-4293-a46e-49d7e6c1e907)	0	0	0	0	6	2192	986	0	8768	3944	104	2198	4096	1	82	
NFS:31181821 (813b97ae-99c1-4198-a3aa-791bd915b959)	0	0	0	0	5	2254	936	0	9016	3744	173	2761	4096	3	86	
NFS:35678286 (bdc1e686-fb82-4157-9657-b8b038ab3e00)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
NFS:31181824 (3d9becd64-93de-4891-9351-500e589db25b)	0	0	0	0	3	2258	921	0	9032	3684	185	3243	4096	3	86	
NFS:31181825 (4a3fc350-73a3-4ead-9104-4d5823143f48)	0	0	0	1	1	2289	956	0	9156	3824	133	2575	4096	3	84	

Image credit: <https://nutanixbible.com>

- 2009/latency: Stargate page backend latency
- 2009/vdisk\_stats: Stargate page histograms of I/O, latency, writes
  - If high read latency, check read source for vDisk and look where I/Os are served from (more than likely reads coming from HDD)
  - Random/smaller IO (<64K) written directly to OpLog. Larger/sequential IO bypass
  - OpLog and written to extent store (Estore)
  - Ops and Randomness section shows if IO is random or sequential
  - Working Set Size shows last 2 minutes and 1 hour
  - Read source shows where IO's are being served from
  - Write Destination shows where IO's are going
  - Extent Group Up-Migration shows data up-migrated in last 300, 3600, and 86400 seconds

#### Hosted vDisks

The dedup usage is only periodically computed by the curator and may be stale  
[Oplog / vdisk statistics](#)  
[Stats for all vDisks](#)

vDisk Id	vDisk Name	Usage (GB)	Dedup (GB)	Oplog			Outstanding ops			Ops/s				KB/s		Avg latency (usec)	Avg op size	Avg qlen	% busy	
				KB	Fragments	Ops/s	KB/s	Read	Write	Estore	Read	Write	Error	Random	Read					Write
<a href="#">14437295</a>	NFS:20581239 (a549adb6-3809-423e-a89e-df2f02d66536)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
<a href="#">31181823</a>	NFS:31181823 (1a8ef41e-ac3f-4293-a46e-49d7e6c1e907)	2	0	198372	49593	990	3960	0	0	0	2192	1	0	2193	8768	320	46	4243	0	10
<a href="#">31181821</a>	NFS:31181821 (813b97ae-99c1-4198-a3aa-791bd915b959)	2	0	196920	49250	939	3756	0	0	0	2253	0	0	2253	9012	0	38	4096	0	11

Image credit: <https://nutanixbible.com>

- 2009/h/traces: Stargate page monitor activity traces for operations
- 2009/h/vars: Stargate page monitor various counters
- 2010: Curator page for Curator runs
  - Need to be on the Curator master
  - Partial scan every 60 minutes
  - Full scan every 6 hours
  - Also triggered by:

- Periodic (normal state)
- Disk/Node/Block failure
- ILM Imbalance
- Disk/Tier Imbalance
- Partial scans have a single MapReduce job
- Full scans have four

### Curator Jobs

Job id	Execution id	Job name	Status	Reasons	Background tasks			Start time	End time	Total time (secs)	Scan timeout (secs)
					Submitted	Canceled	Generated				
1	<a href="#">21063</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 13:59:14	Jul 13 14:05:12	358	43200
1	<a href="#">21061</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 12:53:07	Jul 13 12:59:13	366	43200
1	<a href="#">21059</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 11:46:33	Jul 13 11:53:06	393	43200
0	<a href="#">21054</a>	Full Scan	Succeeded	Periodic	34	0	34	Jul 13 10:29:30	Jul 13 10:46:32	1022	43200
1	<a href="#">21052</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 09:55:01	Jul 13 10:01:12	371	43200
1	<a href="#">21050</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 08:48:44	Jul 13 08:55:00	376	43200
1	<a href="#">21048</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 07:42:04	Jul 13 07:48:43	399	43200
1	<a href="#">21046</a>	Partial Scan	Succeeded	Periodic	0	0	0	Jul 13 06:36:08	Jul 13 06:42:03	355	43200
1	<a href="#">21044</a>	Partial Scan	Succeeded	Periodic	6	0	6	Jul 13 05:29:30	Jul 13 05:36:07	397	43200
0	<a href="#">21039</a>	Full Scan	Succeeded	Periodic	37	0	37	Jul 13 04:12:12	Jul 13 04:29:29	1037	43200

Image credit: <https://nutanixbible.com>

### MapReduce Jobs

Job id	Job name	Status	Map tasks done	Reduce tasks done	Fg tasks	Bg tasks	Errors	Start time	End time	Total time (secs)
<a href="#">21064</a>	PartialScan MapReduce	Succeeded	35/35	35/35	2493	0	0	Jul 13 14:00:14	Jul 13 14:05:12	298
<a href="#">21062</a>	PartialScan MapReduce	Succeeded	35/35	35/35	2492	0	0	Jul 13 12:54:07	Jul 13 12:59:12	305
<a href="#">21060</a>	PartialScan MapReduce	Succeeded	35/35	35/35	2493	0	0	Jul 13 11:47:34	Jul 13 11:53:05	331
<a href="#">21058</a>	FullScan MapReduce #4	Succeeded	14/14	28/28	2621	34	0	Jul 13 10:41:13	Jul 13 10:46:30	317
<a href="#">21057</a>	FullScan MapReduce #3	Succeeded	7/7	7/7	0	0	0	Jul 13 10:38:26	Jul 13 10:41:13	167
<a href="#">21056</a>	FullScan MapReduce #2	Succeeded	7/7	7/7	0	0	0	Jul 13 10:33:47	Jul 13 10:38:26	279
<a href="#">21055</a>	FullScan MapReduce #1	Succeeded	15/15	14/14	33	0	0	Jul 13 10:33:30	Jul 13 10:33:47	17
<a href="#">21053</a>	PartialScan MapReduce	Succeeded	35/35	35/35	2493	0	0	Jul 13 09:56:01	Jul 13 10:01:11	310
<a href="#">21051</a>	PartialScan MapReduce	Succeeded	35/35	35/35	2492	0	0	Jul 13 08:49:45	Jul 13 08:54:59	314
<a href="#">21049</a>	PartialScan MapReduce	Succeeded	35/35	35/35	2492	0	0	Jul 13 07:43:04	Jul 13 07:48:42	338

Image credit: <https://nutanixbible.com>

Activity Full	Scan	Partial Scan
ILM	X	X
Disk Balancing	X	X
Compression	X	X
Deduplication	X	
Erasur Coding	X	
Garbage Cleanup	X	

- 2010/master/control: Curator page for manually starting jobs
- 2011: Chronos page monitors jobs/tasks scheduled by Curator
- 2020: Cerebro page monitors PDs/rep status/DR
- 2020/h/traces: Cerebro page to monitor activity traces for PD/rep
- 2030: Main Acropolis page for details about hosts/tasks/networking
- 2030/sched: Acropolis page with info about VM/resource scheduling/placement decisions
- 2030/tasks: Acropolis page shows info about Acropolis tasks/state
- 2030/vms: Information about Acropolis VMs

## Cluster Commands

### Cluster Status:

```
cluster status
```

### Local CVM Status:

```
genesis status
```

### Check Upgrade Status:

```
upgrade_status
```

### Stop Cluster Service:

```
cluster stop [Service Name]
```

### Start Stopped Cluster Services:

```
cluster start #NOTE: This will start all stopped services
```

### Start Single Service:

```
cluster state [Service Name]
```

### Restart Local Service:

```
genesis stop [Service Name]
```

### Find Cluster ID:

```
zeus_config_printer | grep cluster_id
```

### Find AOS Version:

```
allssh "cat /etc/nutanix/release_version"
```

Find CVM Version:

```
allssh "cat /etc/nutanix/svm-version"
```

Run NCC Health Checks:

```
ncc health_checks run_all
```

## Logs

All cluster logs:

```
allssh "cat ~/data/logs/Acropolis.log"
```

Errors logs:

```
allssh "cat ~/data/logs/[COMPONENT].ERROR"
```

Fatal logs:

```
allssh "cat ~/data/logs/[COMPONENT].FATAL"
```

## Storage Layers and Monitoring

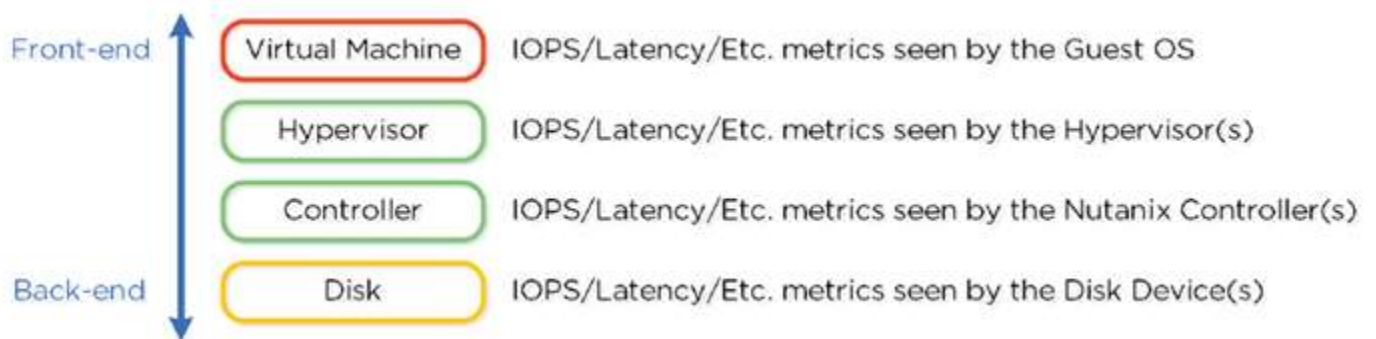


Image credit: <https://nutanixbible.com>

## VM Layer

- Metrics reported by hypervisor for VM
- Represent performance VM is seeing
- Indicative of I/O
- Usage: troubleshooting VM level

## Hypervisor Layer

- Metrics reported by hypervisor
- Represent performance Hypervisor is seeing
- Usage: detailed/valuable metrics

## Controller Layer

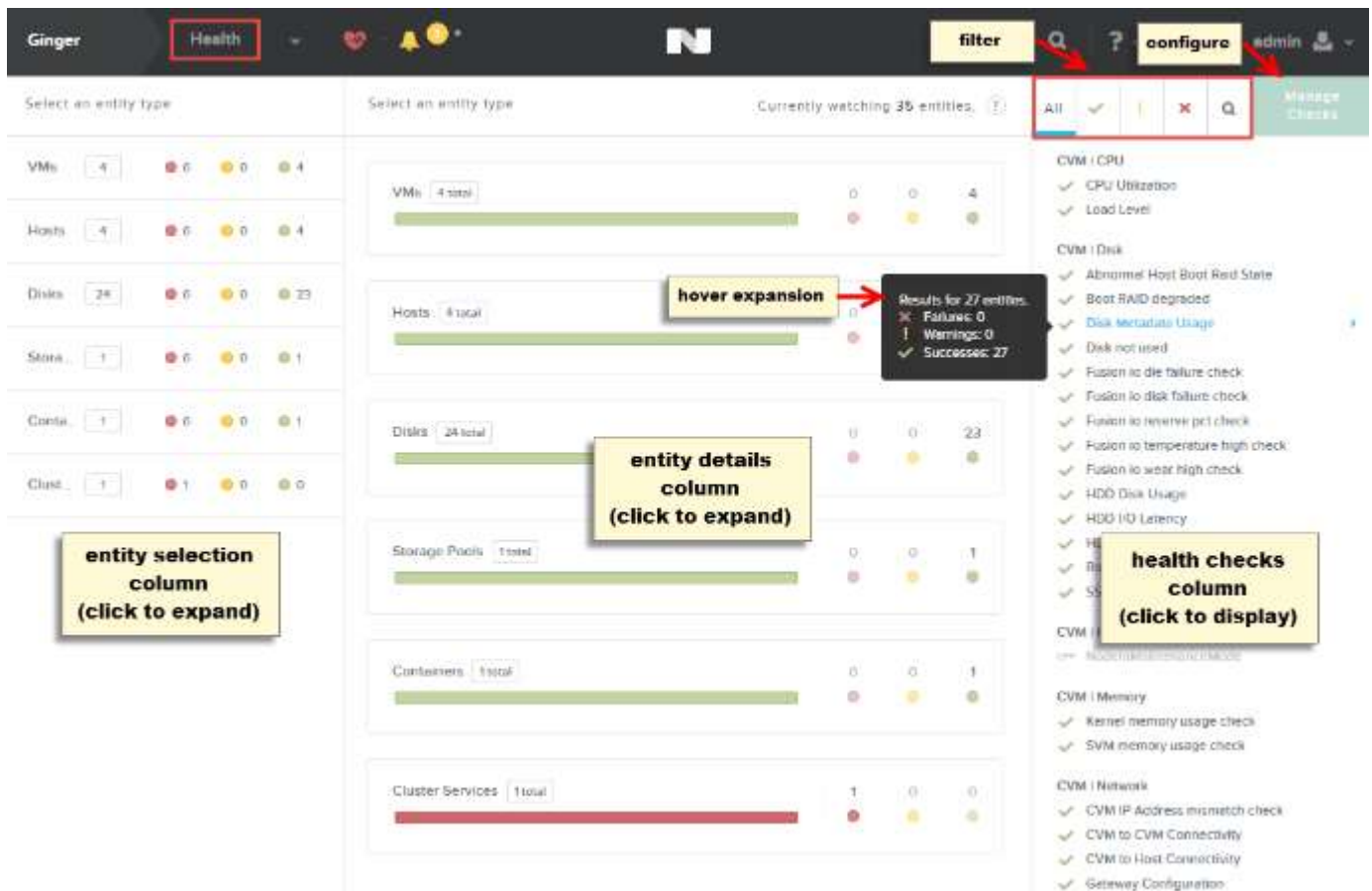
- Metrics from Nutanix Controllers
- From CVM's
- What front-end is seeing from NFS/SMB/iSCSI or backend operations (balancing, ILM).
- Should normally match hypervisor layer
- Usage: show backend operations

## Disk Layer

- Metrics reported by disk devices
- Pulled from physical disks.
- Data hitting OpLog or Extent Store.
- Usage: See how many ops served from cache or disk

## Utilize the Health dashboard and its major components

The Health dashboard displays dynamically updated health information about VMs, hosts, and disks in the cluster. To view the Health dashboard, select **Health** from the pull-down list on the left of the main menu.



## Screen Details

The Health dashboard is divided into three columns:

- The left column displays tabs for each entity type (VMs, hosts, disks, storage pools, containers, cluster services, and [when configured] protection domains and remote sites). Each tab displays the entity total for the cluster (such as the total number of disks) and the number in each health state. Clicking a tab expands the displayed information (see following section).
- The middle column displays more detailed information about whatever is selected in the left column.
- The right column displays a list of the available health checks along with the current status of each check (success, warning, failure, or disabled).
  - Hovering the cursor over an entry displays more information about that health check.
  - You can filter the list by selecting one of the buttons at the top of the column.

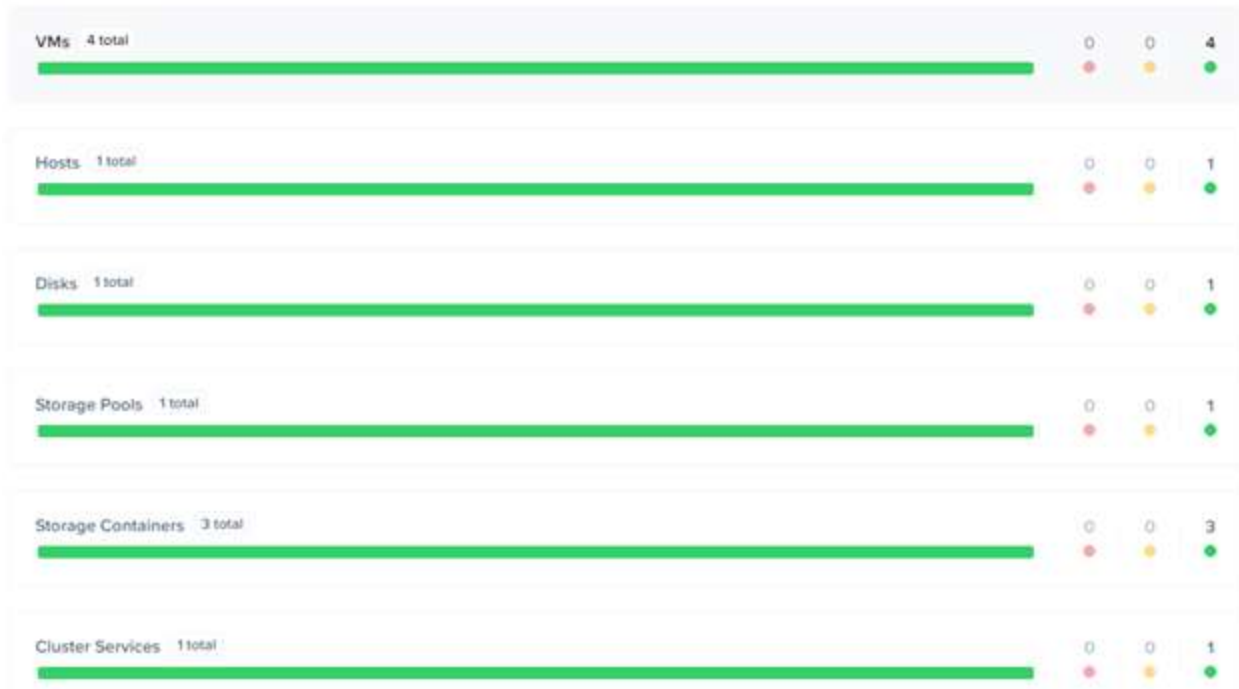


- You can learn more about (or reconfigure) the health checks by clicking one of the health checks in the list or the **Manage Checks** button.

Select an entity type				
VMs	4	<span style="color: red;">●</span> 0	<span style="color: yellow;">●</span> 0	<span style="color: green;">●</span> 4
Hosts	1	<span style="color: red;">●</span> 0	<span style="color: yellow;">●</span> 0	<span style="color: green;">●</span> 1
Disks	1	<span style="color: red;">●</span> 0	<span style="color: yellow;">●</span> 0	<span style="color: green;">●</span> 1
Stora...	1	<span style="color: red;">●</span> 0	<span style="color: yellow;">●</span> 0	<span style="color: green;">●</span> 1
Stora...	3	<span style="color: red;">●</span> 0	<span style="color: yellow;">●</span> 0	<span style="color: green;">●</span> 3
Cluste...	1	<span style="color: red;">●</span> 0	<span style="color: yellow;">●</span> 0	<span style="color: green;">●</span> 1

Select an entity type

Currently watching 11 entities. 



**Summary**    Checks    Actions 

 ALL CHECKS    646

BY CHECK STATUS

 Passed    615

 Failed    0

 **Warning**    **0**

 Error    16

 Off    15

BY CHECK TYPE

 Scheduled    228

 Not Scheduled    118

 Event Triggered    300

colossus05 Health admin

All > Disks Currently watching 18 / 18 total disks.

**grouping selection**

GROUP DISKS BY

- Storage Tier** 2 Groups
- Disk Usage 5 Groups
- Disk Capacity 4 Groups
- Health 3 Groups

FILTER BY HEALTH

- Critical 0
- Warning 0
- Good 18

DISK MODE

- Online 18
- Offline 0

DISK STATUS

- Normal 18
- Data Migration Initiated 0
- Marked for removal 0
- Detachable 0

**grouping filters**

**view options**

**select list**

Sort by Health

**order list**

You are grouping Disks by 'Storage Tier'.

DAS-SATA 15

0 0 15

**hover expansion**

Disk Serial #: 9XG04GYT  
Disk Id: 39  
Storage Tier: DAS-SATA  
Used (Physical): 0 GB  
Capacity (Physical): 92578 GB

SSD-PCIe 3

0 0 3

**grouping details**

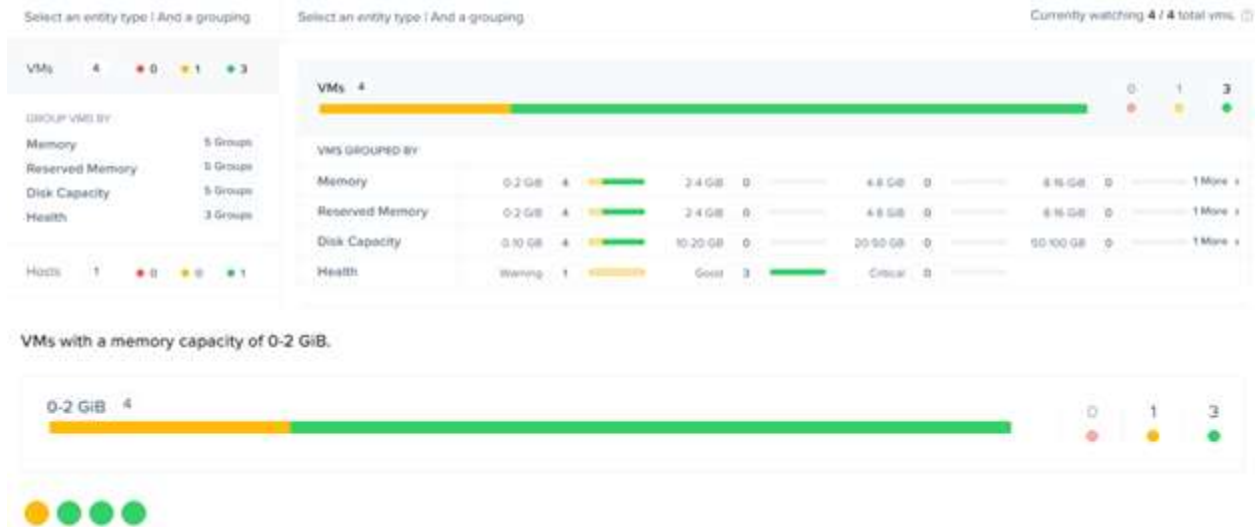
1 VMs 4 0 1 3

Group VMs by memory All Groups Currently watching 4 / 4 total vms.

You are grouping VMs by 'memory'.

0-2 GiB 4 2-4 GiB 0 4-8 GiB 0 8-16 GiB 0

See all in this group None in this group None in this group None in this group



## Configure Alert e-mail settings for a cluster

### Configuring Alert Emails

Prism Central allows you to configure the alert messages sent by Prism Central. To configure alert settings, reporting rules, and message templates, do the following:

#### Settings

Email and Alerts

**Alert Email Configuration**

Alert Policies

SMTP Server

Settings · Rules · **Template**

### Email Delivery

Configure the frequency of alert email notifications.

**Individual Alert**

An email is sent for every alert.

**Daily Digest**

A single summary email is sent once per day.

**Skip empty digest messages**

Only send a daily digest email if there are one or more alerts.

### Email Recipients

**Nutanix Support (nos-alerts@nutanix.com)**

Additional email recipients

#### TUNNEL CONNECTION

Mode **Default Nutanix Tunnel**

Status **● DISABLED**

Settings · **Rules** · Template

Configure alert rules which control the conditions for triggering alerts and the recipients for receiving notifications.

+ New Rule

Settings Rules **Template**

Preview

Subject Data Disk Space Usage High

Body Disk space usage for {mount\_path} on {entity} [{ip\_address}] has exceeded {disk\_usage\_threshold}%.

Prepend Content To The Email Subject

Append Content To The Email Body

## Configuring Alert Policies

The system monitors a variety of conditions and sends an alert whenever one of the alert conditions is detected (when alerting is enabled). There are default policies for these alerts, but you have the option to modify the default policies and add new policies.

## Alert Policies

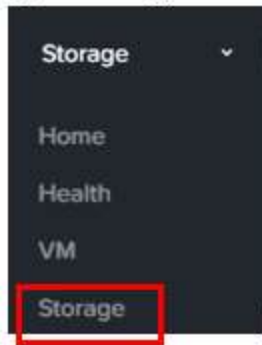
ID	Title	Impact Type	Entity Type	Rule	Enabled	Auto Resolve	Last Update Time	Action
A1001	CVM Connectivity Failure	System Indicator	Cluster	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A1003	Protection Domain Replication Expired	System Indicator	Protection Domain	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A1005	Data Disk Space Usage High	Capacity	Host	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A1006	Host IP Not Reachable	Availability	Host	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A1008	IPMI IP Address Mismatch	Configuration	Host	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A1009	CVM IP Address Mismatch	Configuration	Host	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A1010	Protected VM Not Found	System Indicator	Protection Domain	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎
A101047	Cluster can not tolerate node failure and guarantee available rebuild capacity	System Indicator	Storage Pool	Default	Enabled	Enabled	09/10/19, 10:33:23 AM	✎

## Section 7 – Distributed Storage Fabric

### Identify methods for creating a Storage Container

A storage container is a defined subset of available storage within a storage pool. Storage Containers allow you to apply rules or transformations such as compression to a data set.

- Logical segmentation of SP containing groups of VM's/files (vDisks).
- Typically have 1:1 mapping with datastore



## VMware

Storage presented via NFS (default) or iSCSI

**Create Storage Container** ? X

Enter a name for your storage container and select a storage pool for it. You can provision the storage container for all hosts or select individual hosts.

NAME

STORAGE POOL  
 +

MAX CAPACITY  
**12.95 TiB** (Physical) Based on storage pool free unreserved capacity

NFS DATASTORE  
 Mount on all ESXi hosts  
 Mount on the following ESXi hosts

## AHV

Storage presented via iSCSI



### Create Storage Container ? ×

Enter a name for your storage container and select a storage pool for it. You can provision the storage container for all hosts or select individual hosts.

NAME

STORAGE POOL  
 ▼ +

MAX CAPACITY  
**13.77 TiB** (Physical) Based on storage pool free unreserved capacity

⚙️ Advanced Settings Cancel Save

## Hyper-V

Storage presented via SMB

## Create Storage Container



Enter a name for your storage container and select a storage pool for it. You can provision the storage container for all hosts or select individual hosts.

NAME

STORAGE POOL

MAX CAPACITY

**10.27 TiB** (Physical) Based on storage pool free unreserved capacity

MAKE THIS STORAGE CONTAINER THE DEFAULT STORE FOR VMS ON HYPER V HOSTS

- None
- Make default on all Hyper V hosts
- Make default on particular Hyper V hosts

 Advanced Settings

Cancel

Save

## ADVANCED SETTINGS

REPLICATION FACTOR [?](#)

2

RESERVED CAPACITY (GiB)

0

ADVERTISED CAPACITY (GiB)

Total GiB

COMPRESSION

DEDUPLICATION

CACHE [?](#)

Perform inline deduplication of read caches to optimize performance.

CAPACITY [?](#)

Perform post-process deduplication of persistent data.

ERASURE CODING ?

Erasure coding enables capacity savings across solid-state drives and hard disk drives.

#### FILESYSTEM WHITELISTS

Enter comma separated entries

Use this format for entries: **nnn.nnn.nnn.nnn/xxx.xxx.xxx.xxx**. Also, note that setting a storage container whitelist will override any global whitelists for this storage container.

## Storage Pool

- Group of physical devices (PCIe SSD, SSD, HDD) that can span multiple nodes.
- Typically only a single pool.

## vDisk

- Any file over 512KB on DSF

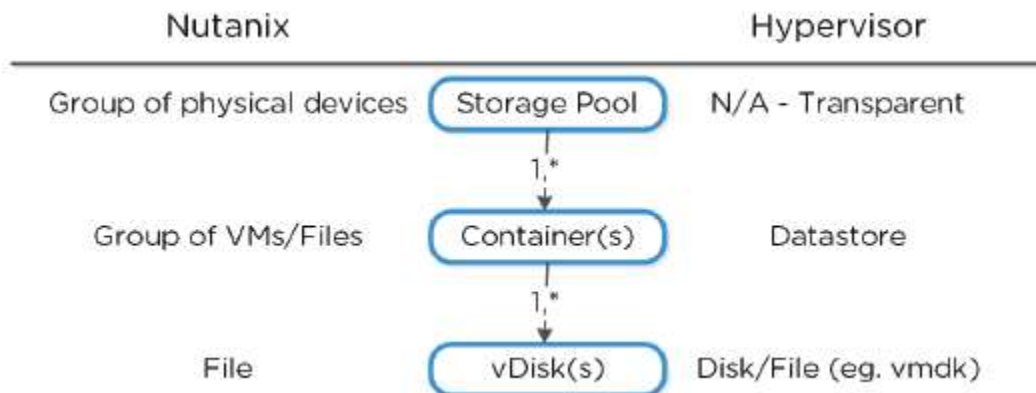


Image credit: <https://nutanixbible.com>

- Composed of extents which are groups/stored on disk as extent group

## Extent

- 1MB piece of logically contiguous data which consists of n number of contiguous blocks
- Written/Read/Modified on sub-extent basis (aka slice)
- May be trimmed when moving into the cache

## Extent Group

- 1MB or 4MB piece of physically contiguous stored data
- Stored as a file and owned by CVM
- Dynamically distributed among groups to provide striping across nodes/disks

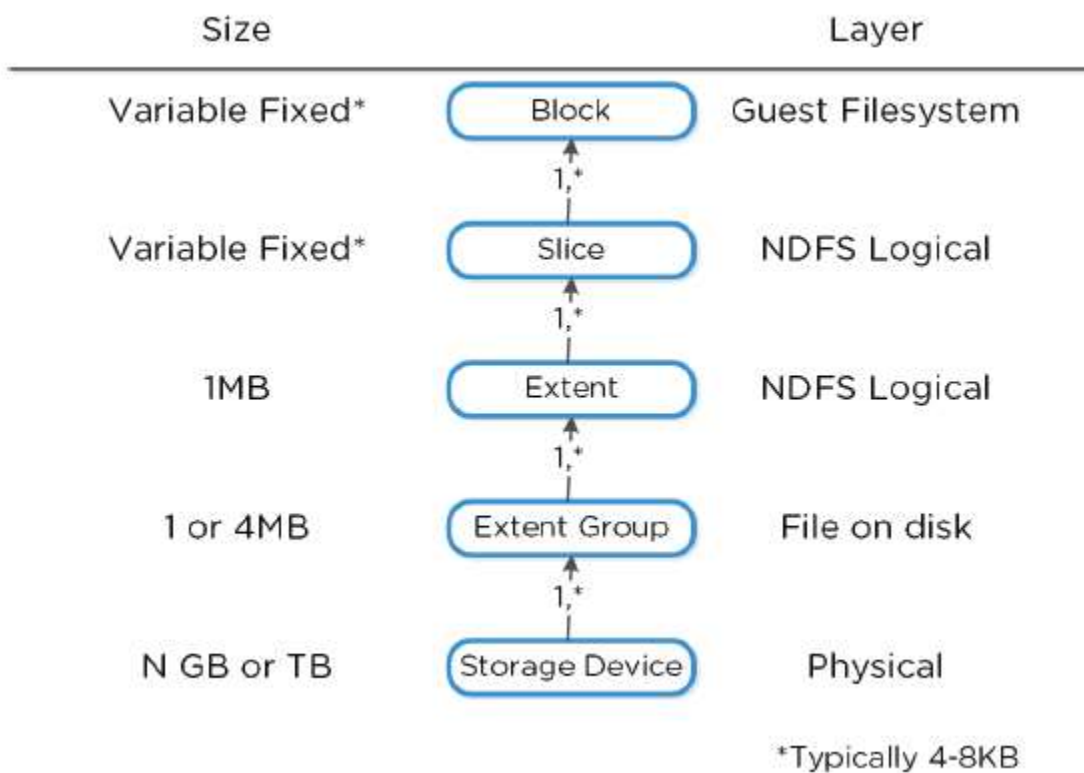


Image credit: <https://nutanixbible.com>

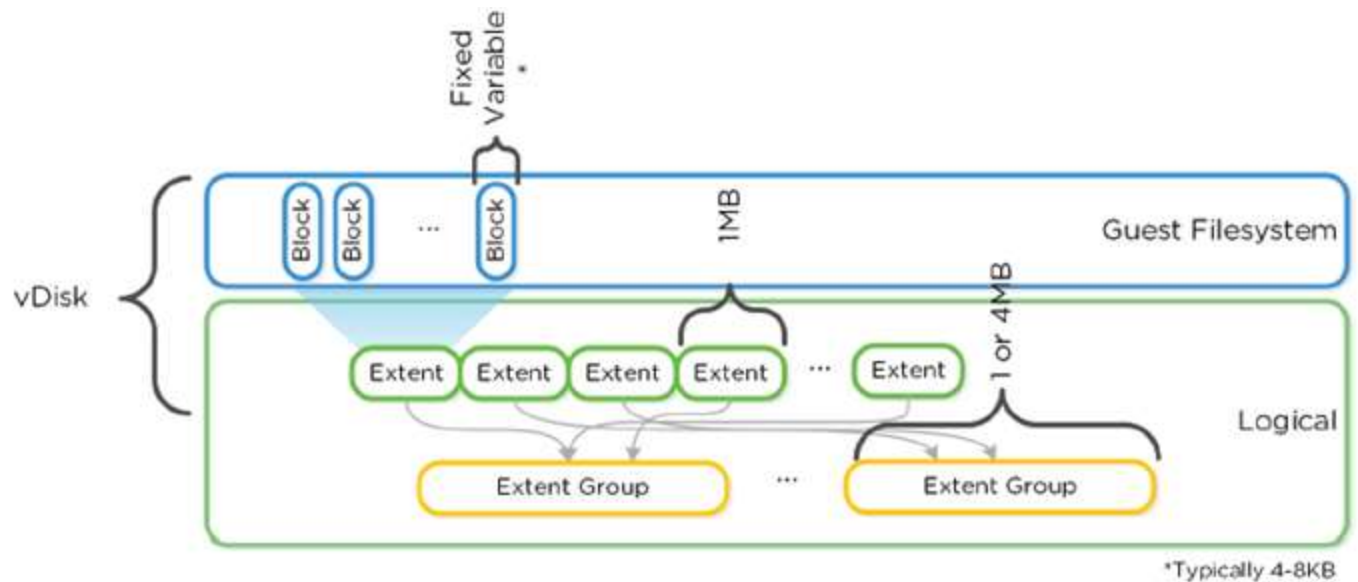


Image credit: <https://nutanixbible.com>

## Determine what capacity optimization method(s) should be used based on a given workload

Data Transform	Application[s]	Comments
Erasure Coding	All	Provides higher availability with reduced overheads than traditional RF. No impact to normal write or read I/O performance. Does have some read overhead in the case of a disk / node / block failure where data must be decoded.
Inline Compression	All	No impact to random I/O, helps increase storage tier utilization. Benefits large or sequential I/O performance by reducing data to replicate and read from disk.
Offline Compression	None	Given inline compression will compress only large or sequential writes inline and do random or small I/Os post-process, that should be used instead.
Perf Tier Dedup	P2V/V2V,Hyper-V (ODX),Cross-container clones	Greater cache efficiency for data which wasn't cloned or created using efficient Acropolis clones.
Capacity Tier Dedup	Same as perf tier dedup	Benefits of above with reduced overhead on disk.

# Erasure Coding

- Similar to RAID where parity is calculated, EC encodes a strip of data blocks on different nodes to calculate parity
- In event of failure, parity used to calculate missing data blocks (decoding)
- Data block is an extent group, and each block is on a different node belonging to a different vDisk
- Configurable based on failures to tolerate data blocks/parity blocks

## EC Strip Size:

- Ex. RF2 = N+1
  - 3 or 4 data blocks + 1 parity strip = 3/1 or 4/1
- Ex. RF3 = N+2
  - 3 or 4 data blocks + 2 parity strips = 3/2 or 4/2

## Overhead:

	FT1 (RF2 equiv.)		FT2 (RF3 equiv.)	
Cluster Size [nodes]	EC Strip Size [data/parity blocks]	EC Overhead [vs. 2X of RF2]	EC Strip Size [data/parity]	EC Overhead [vs. 3X of RF3]
4	2/1	1.5X	N/A	N/A
5	3/1	1.33X	N/A	N/A
6	4/1	1.25X	N/A	N/A
7+	4/1	1.25X	4/2	1.5X

- Recommended to have cluster size which is at least 1 more node than combined strip size (data + parity)
- Allows for rebuilding in event of failure
- Ex. 4/1 strip would have 6 nodes
- Encoding is done post-process leveraging Curator MapReduce framework

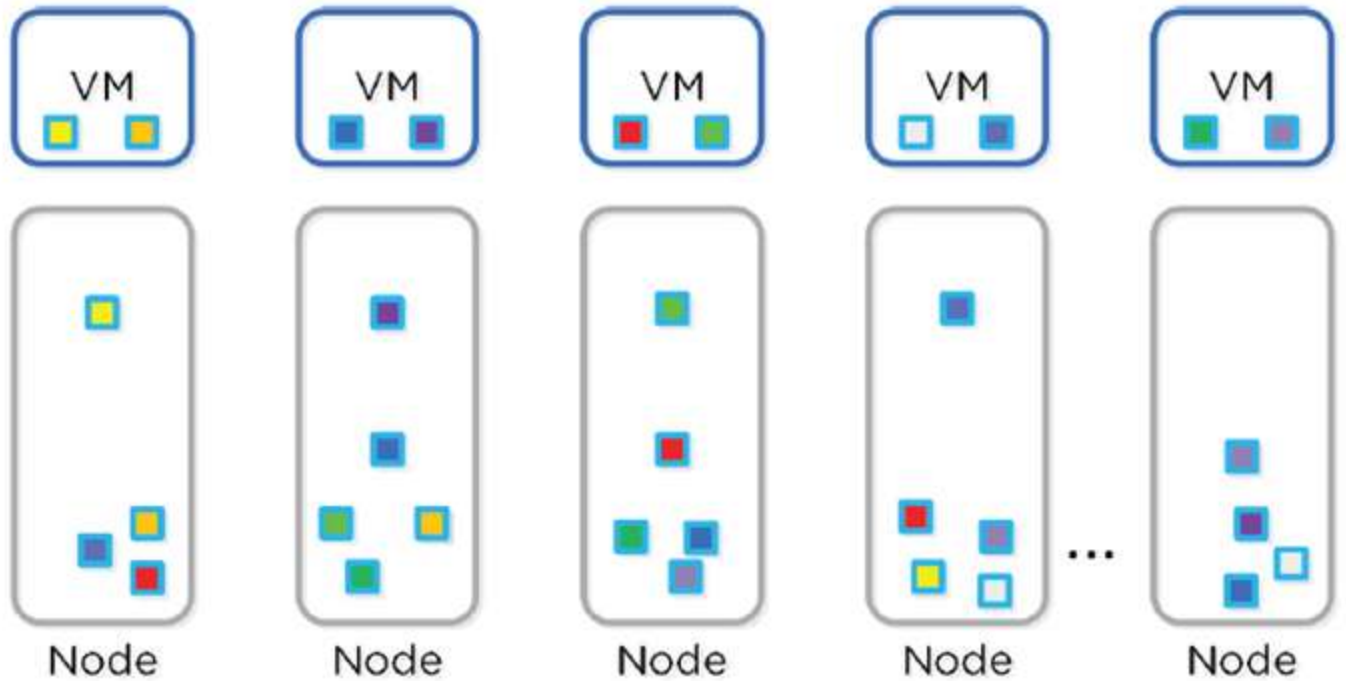


Image credit: <https://nutanixbible.com>

- When Curator scan runs, it finds eligible extent groups to be encoded.
- Must be “write-cold” = haven’t been written to > 1 hour
- Tasks are distributed/throttled via Chronos



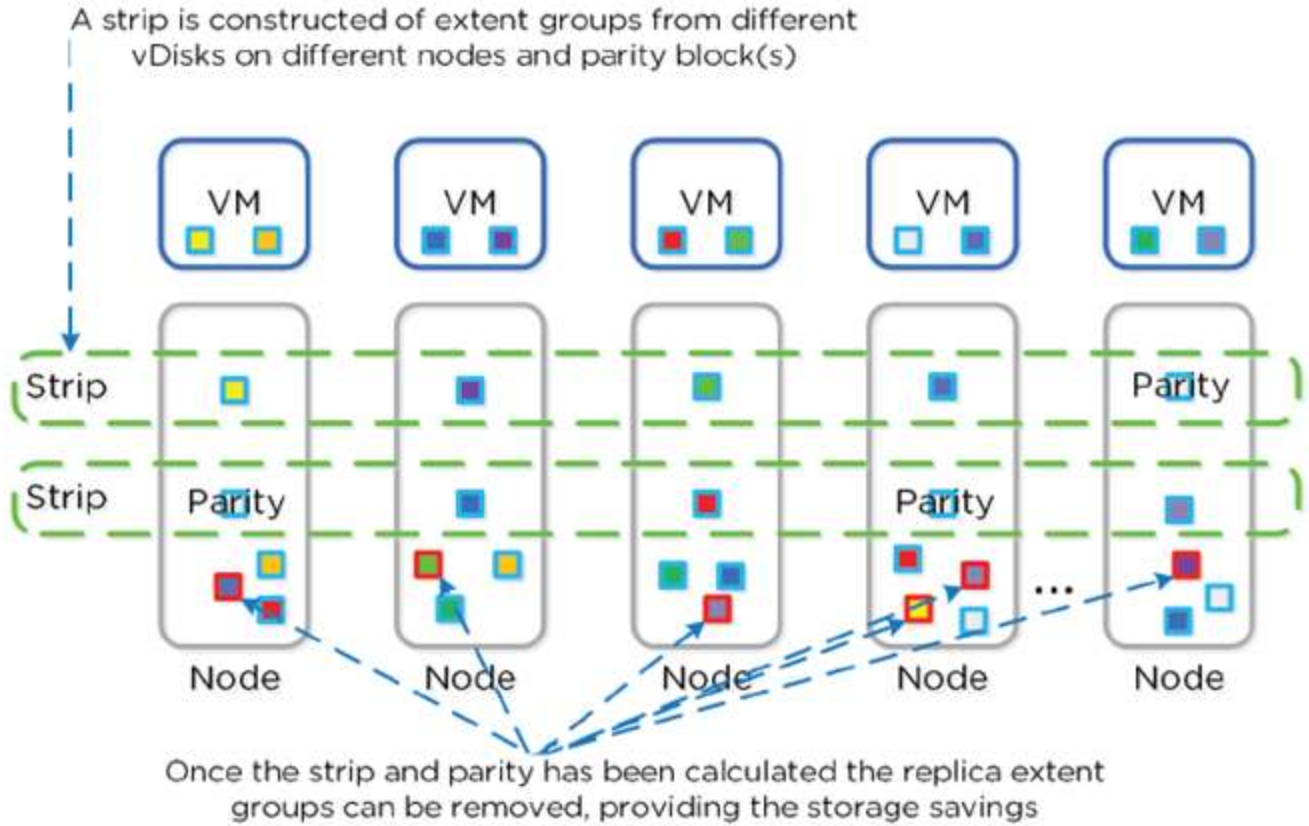


Figure 3.2-18. DSF Encoded Strip - Pre-savings

Image credit: <https://nutanixbible.com>

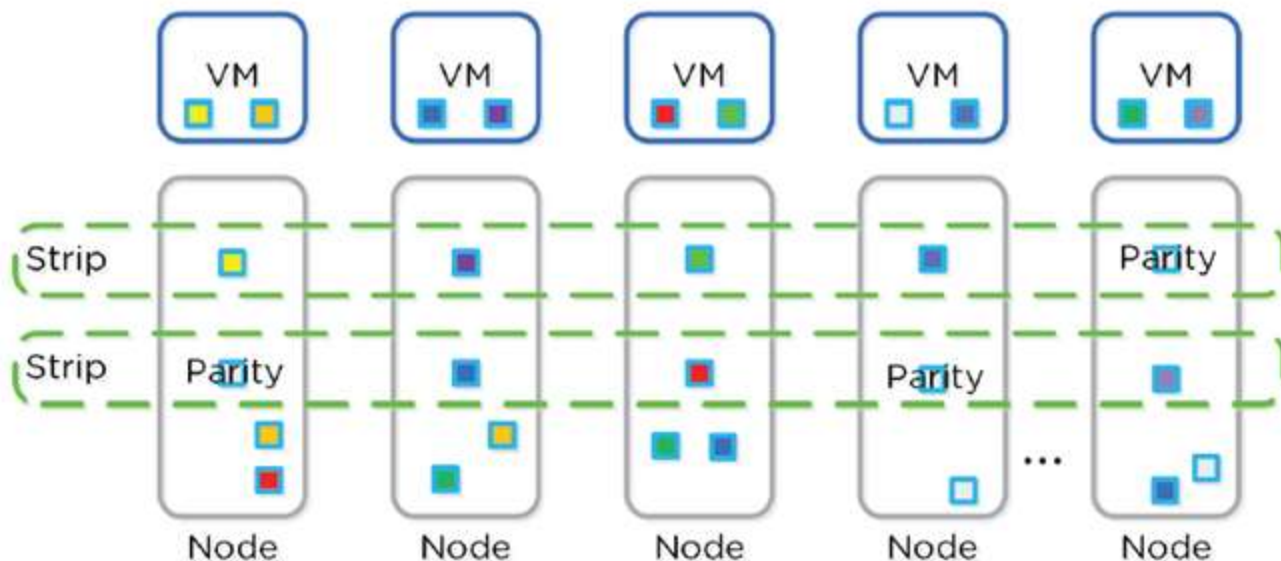


Figure 3.2-19. DSF Encoded Strip - Post-savings

Image credit: <https://nutanixbible.com>

- EC pairs well with Inline Compression

## Compression

Capacity Optimization Engineer (COE) performs data transformations to increase data efficiency on disk

## Inline

- Sequential streams of data or large I/O in memory before written to disk
- Random I/O's are written uncompressed to OpLog, coalesced, and then compressed in memory before being written to Extent Store
- Leverages Google Snappy compression library
- For Inline Compression, **set the Compression Delay to "0" in minutes.**

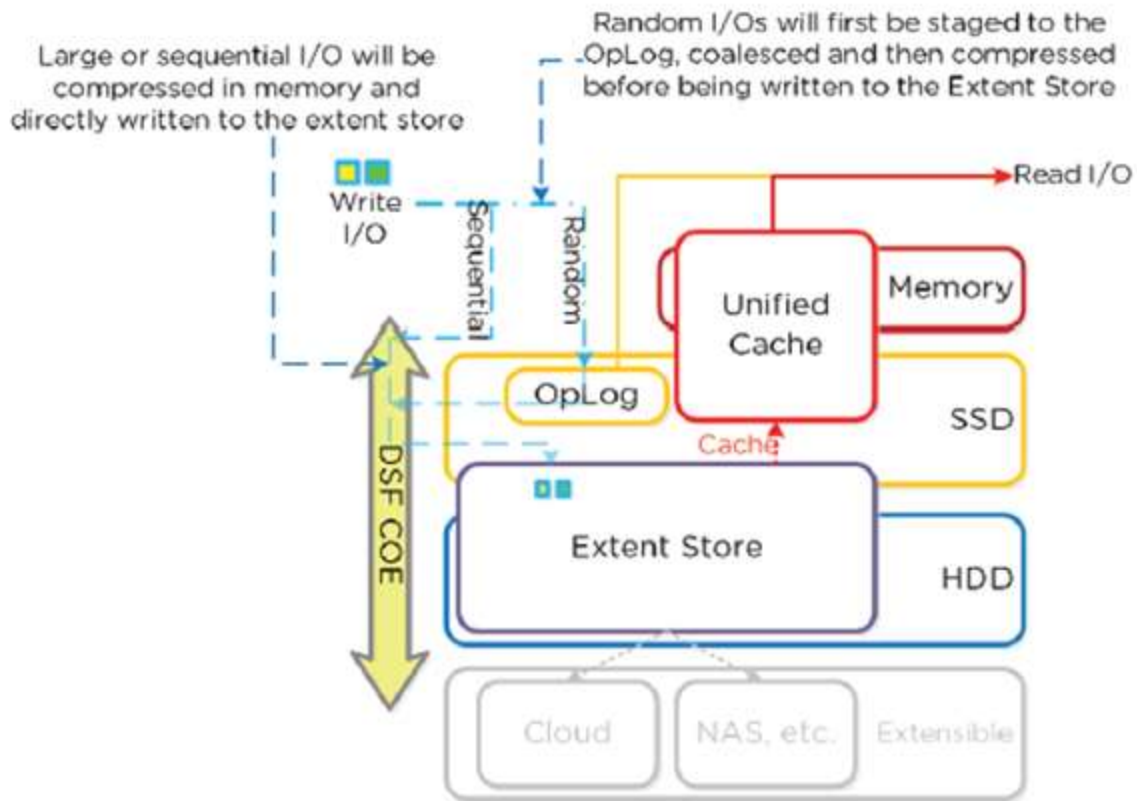


Figure 3.2-20. Inline Compression I/O Path

Image credit: <https://nutanixbible.com>

## Offline

- New write I/O written in uncompressed state following normal I/O path
- After compression delay is met, data = cold (migrated down to HDD tier via ILM) data can be compressed
- Leverages Curator MapReduce framework
- All nodes perform compression task
- Throttled by Chronos

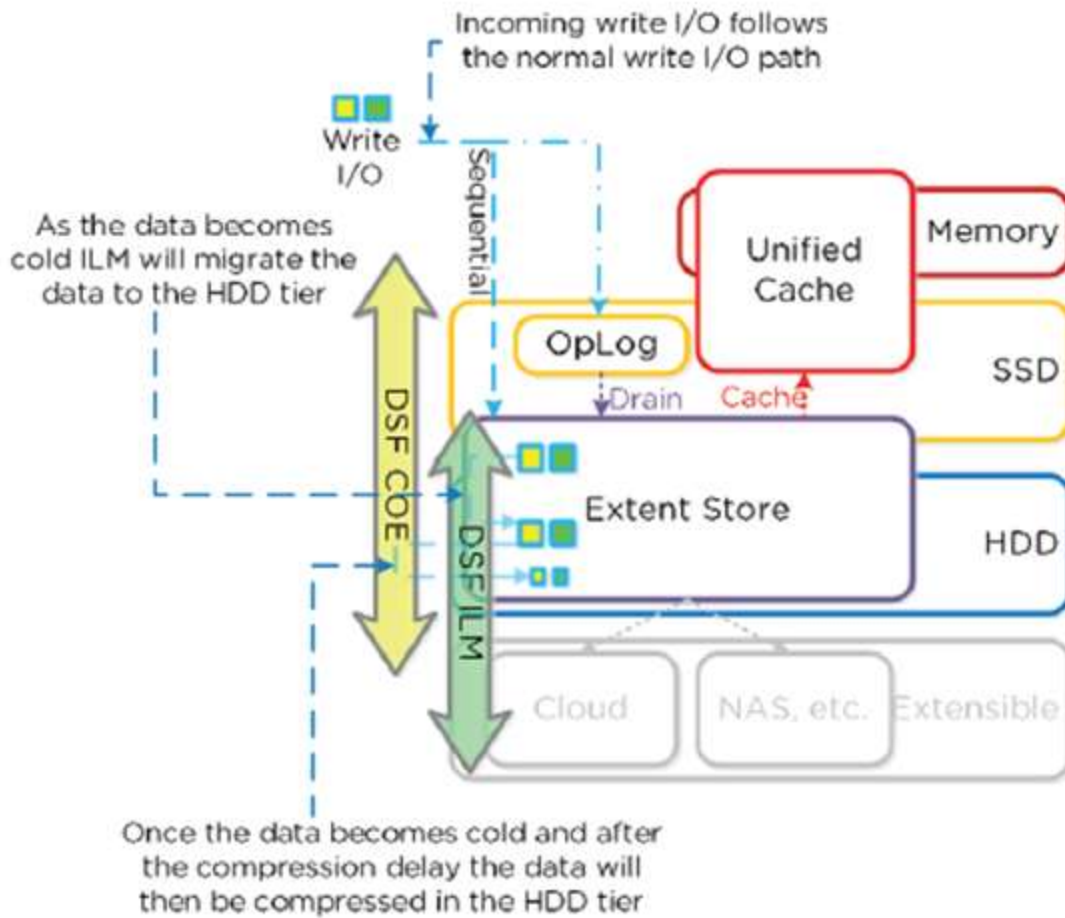


Figure 3.2-21. Offline Compression I/O Path

Image credit: <https://nutanixbible.com>

- For Read/IO, data is decompressed in memory and then I/O is served.
- Heavily accessed data is decompressed in HDD tier and leverages ILM to move up to SSD and/or cache

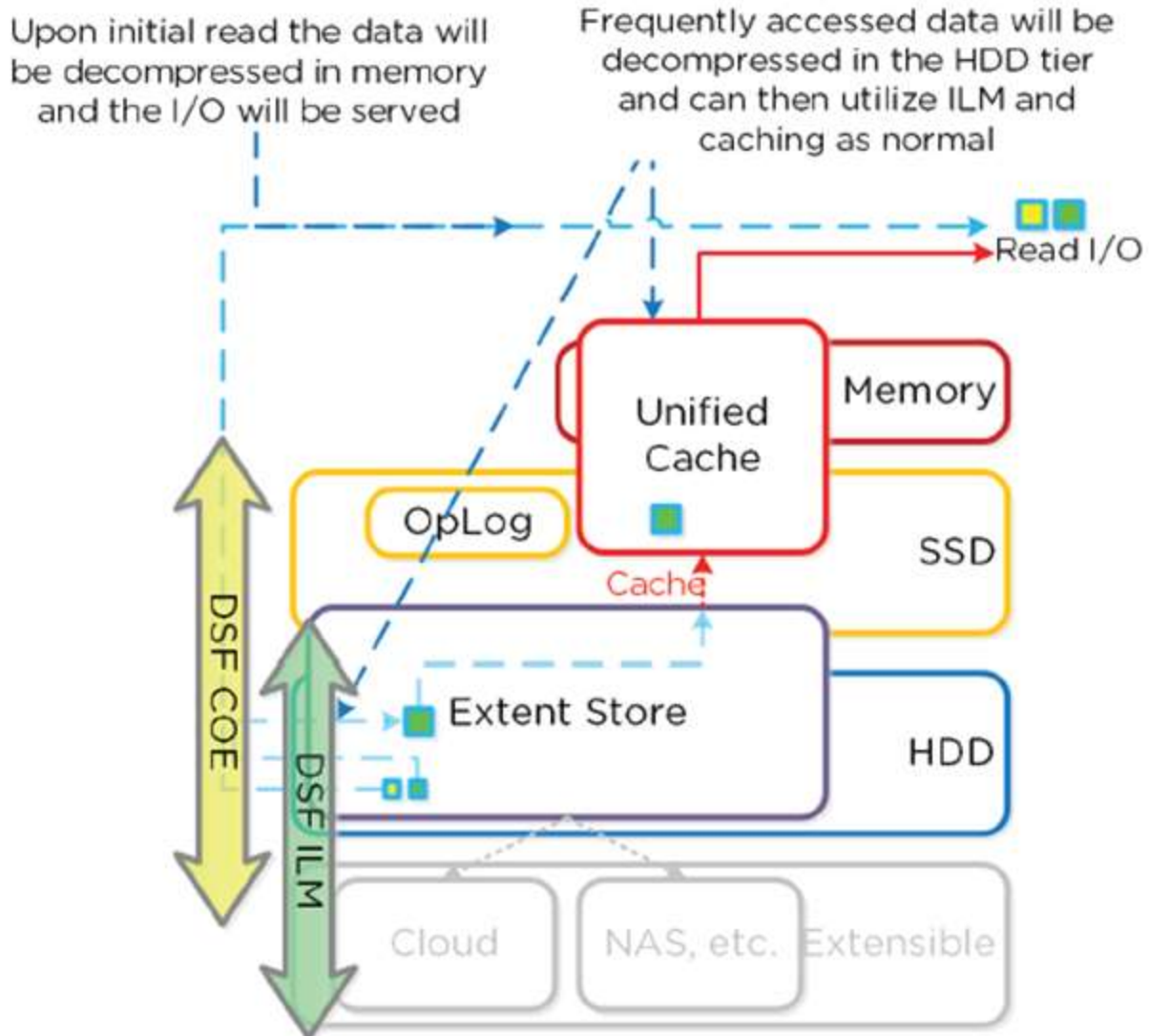


Figure 3.2-22. Decompression I/O Path

Image credit: <https://nutanixbible.com>

## Elastic Dedupe Engine

- Allows for dedupe in capacity and performance tiers
- Streams of data are fingerprinted during ingest using SHA1 hash at 16k
- Stored persistently as part of blocks' metadata
- Duplicate data that can be deduplicated isn't scanned or re-read; dupe copies are just removed.
- Fingerprint reccounts are monitored to track dedupability

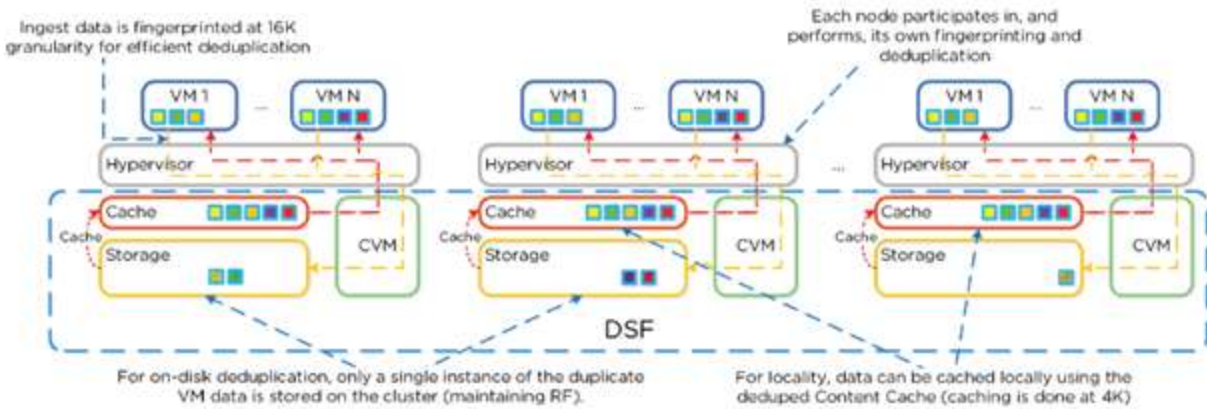


Image credit: <https://nutanixbible.com>

- Intel acceleration is leveraged for SHA1
- When not done on ingest, fingerprinting done as background process
- Where duplicates are found, background process removed data with DSF Map Reduce Framework (Curator)

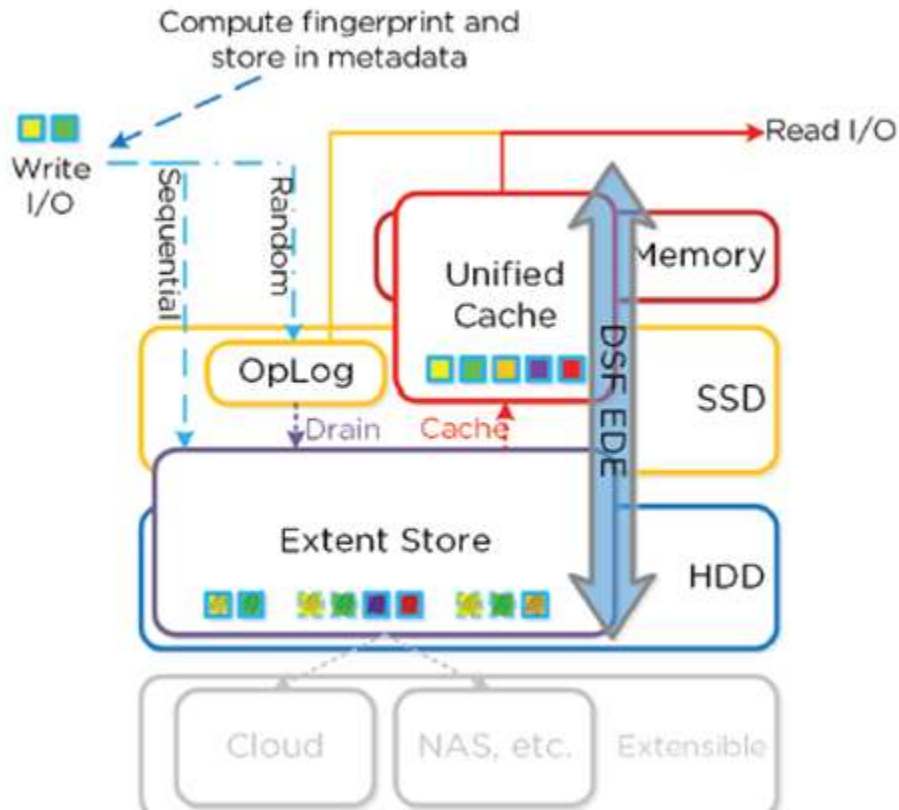


Figure 3.2-24. EDE I/O Path

Image credit: <https://nutanixbible.com>

# Global Deduplication

- DSF can dedupe by just updating metadata pointers
- Same concept in DR/Replication
- Before sending data over the wire, DSF queries remote site to check fingerprint(s) on target
- If nothing, data is compressed/sent to target
- If data exists, no data sent/metadata updated

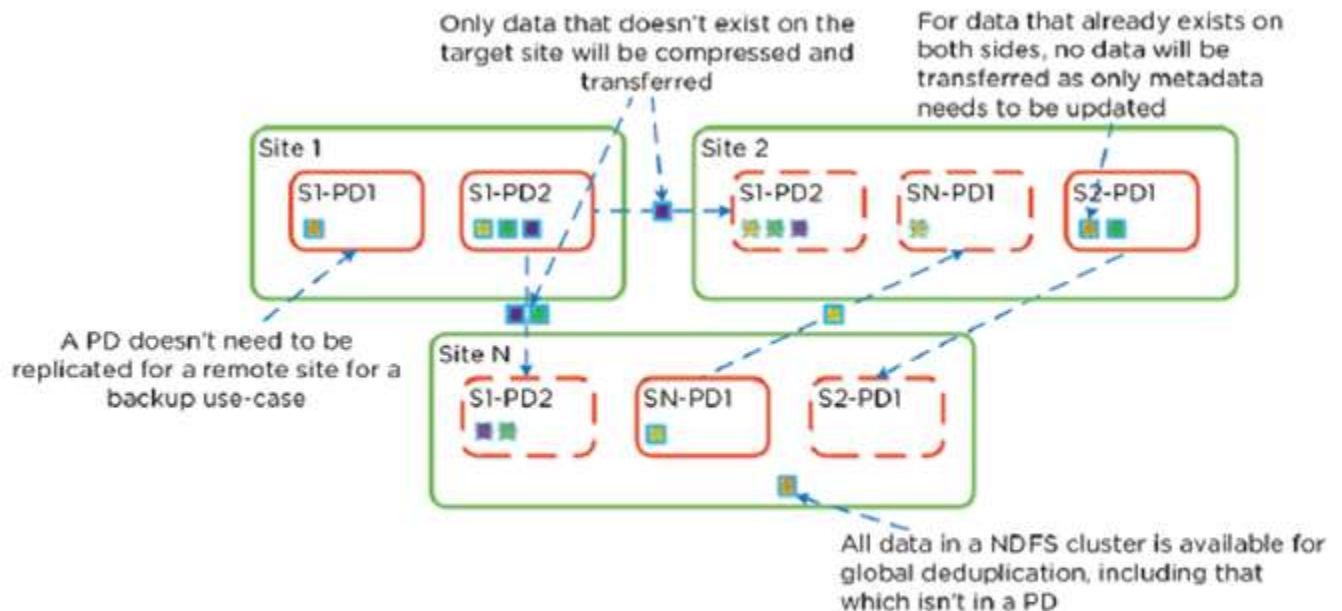
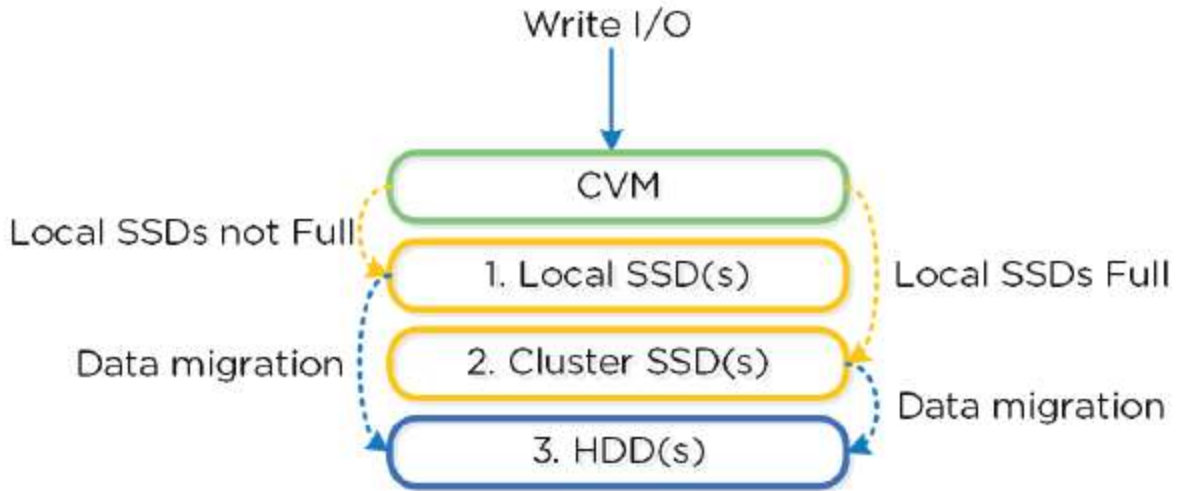


Figure 3.4.22. Replication Deduplication

Image credit: <https://nutanixbible.com>

# Storage Tiering + Prioritization

- ILM responsible for triggering data movement events
  - Keeps hot data local DSF
  - ILM constantly monitors I/O patterns and down/up migrates as necessary
- Local node SSD = highest priority tier for all I/O
- When local SSD utilization is high, disk balancing kicks in to move coldest data on local SSD's to other SSD's in cluster
- All CVM's + SSD's are used for remote I/O to eliminate bottlenecks



\*NOTE: Sequential IO can be configured to bypass SSD and be directly written to the HDD tier.

Image credit: <https://nutanixbible.com>

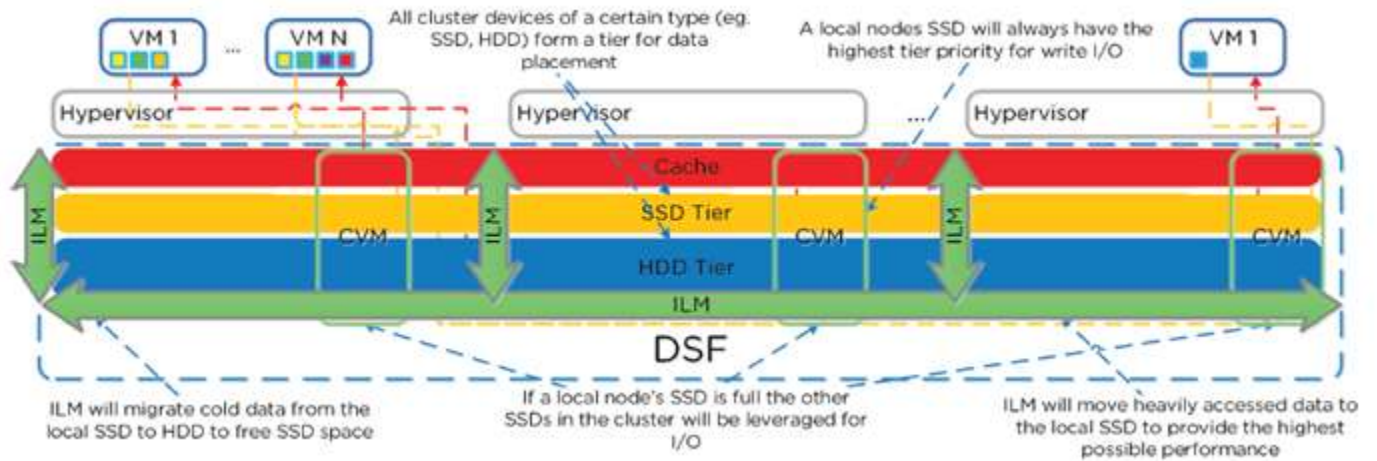


Image credit: <https://nutanixbible.com>



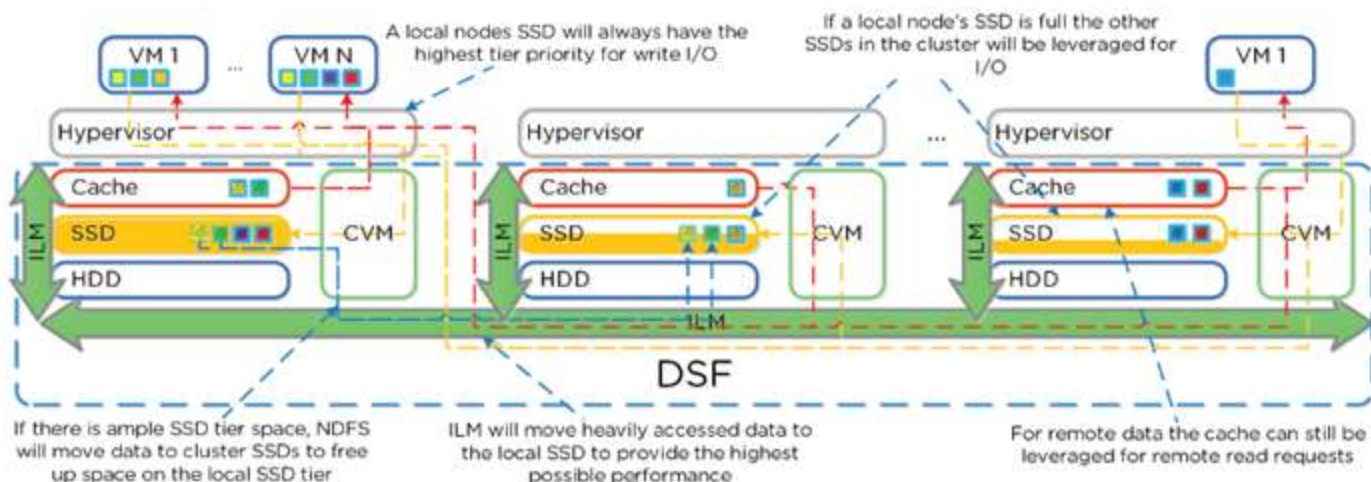


Image credit: <https://nutanixbible.com>

## Data Locality

- VM data is served locally from CVM on local disks under CVM's control
- When reading old data (after HA event for instance) I/O will be forwarded by local CVM to remote CVM
- DSF will migrate data locally in the background
- Cache Locality: vDisk data stored in Unified Cache. Extents may be remote.
- Extent Locality: vDisk extents are on same node as VM.

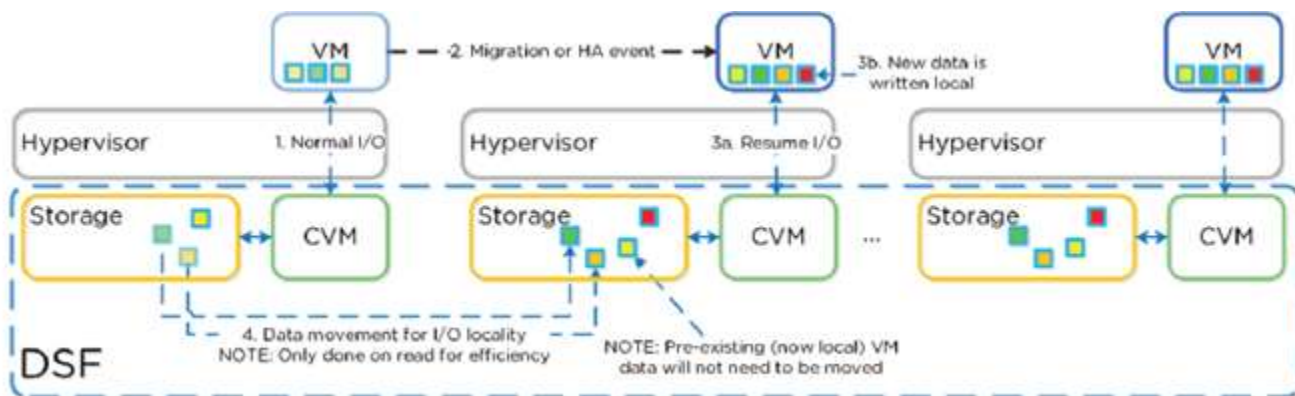


Image credit: <https://nutanixbible.com>

- Cache locality determined by vDisk ownership

## Disk Balancing

- Works on nodes utilization of local storage
- Integrated with DSF ILM

- Leverages Curator
- Scheduled process

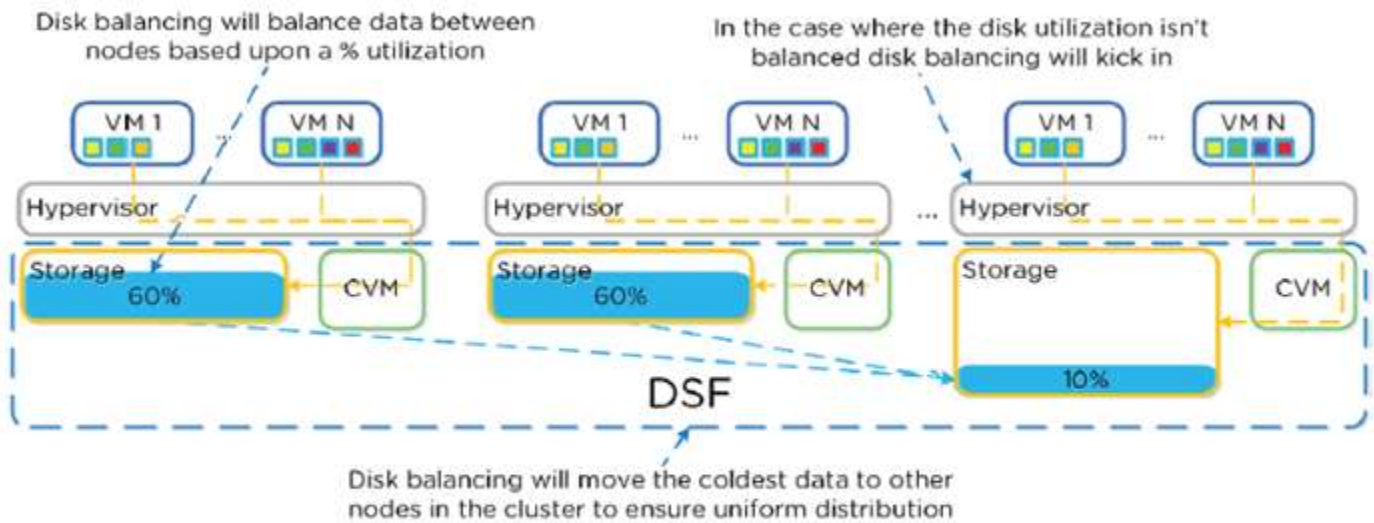


Figure 3.2-29. Disk Balancing - Unbalanced State

Image credit: <https://nutanixbible.com>

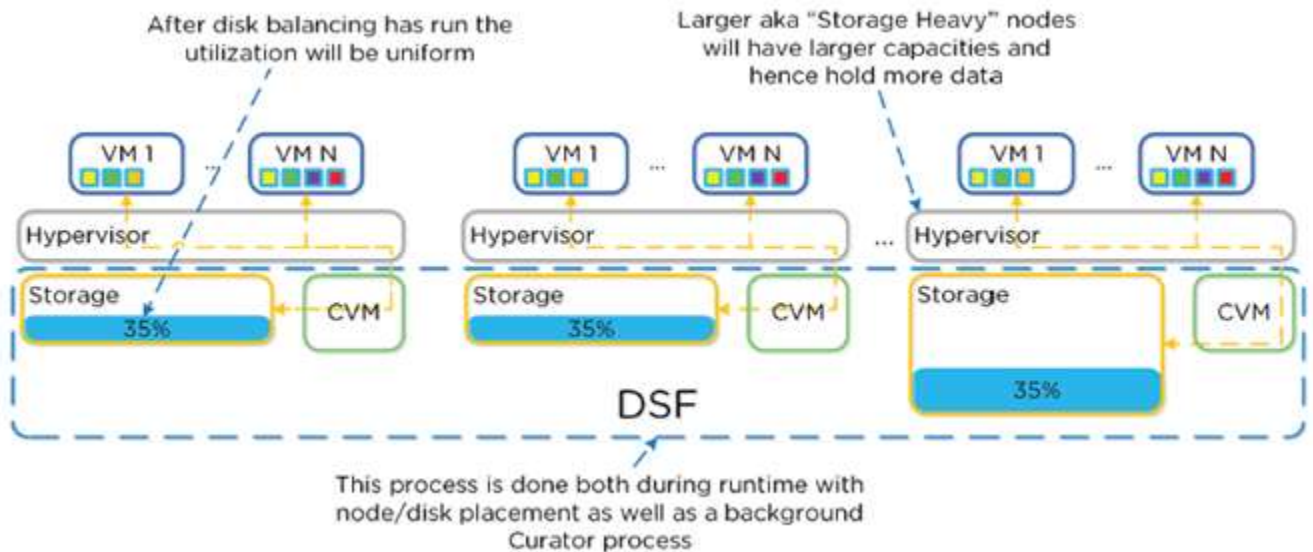


Image credit: <https://nutanixbible.com>

- With "storage only" node, CVM can use nodes full memory to CVM for much larger read cache

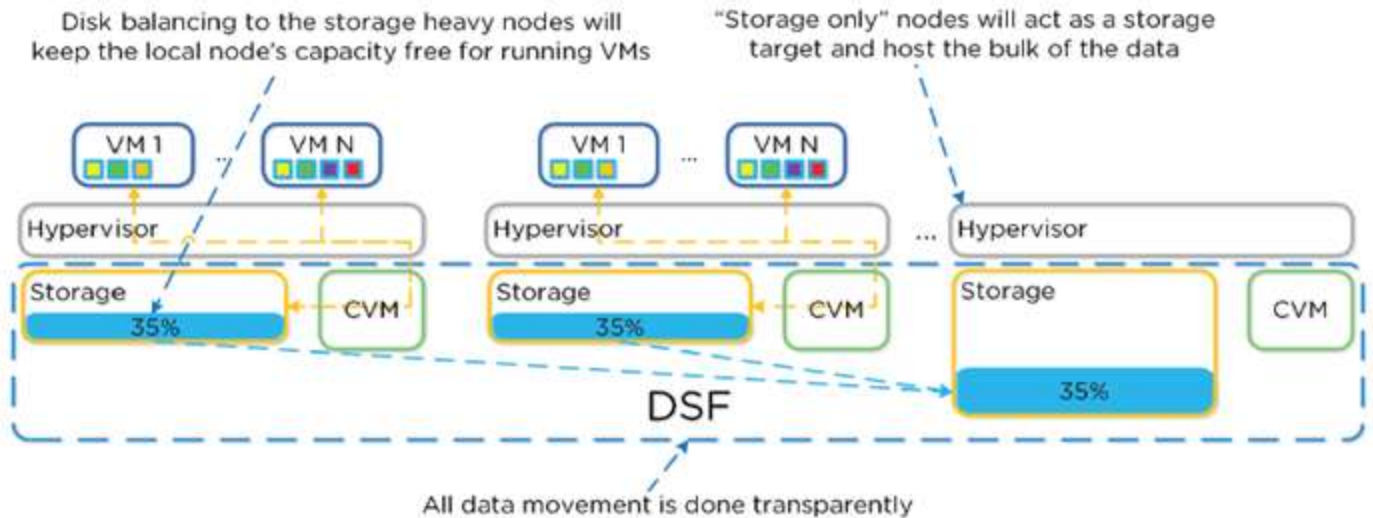


Figure 3.2-31. Disk Balancing - Storage Only Node

Image credit: <https://nutanixbible.com>

## Describe and differentiate technologies used in conjunction with a Distributed Storage Fabric, including snapshots, clones, high availability and disaster recovery

### Snapshots + Clones

- DSF provides native support (VAAI, ODX, etc)
- Leverage **redirect-on-write** algorithm
- VM data consists of files (vmdk/vhdx) which are vDisks
- Snapshot taken = vDisk marked immutable
  - New vDisk created as read/write
  - Both vDisks have same block map
  - Metadata mapping to corresponding extents
  - Since each vDisk is its own block map, it **eliminates need for snapshot chain**

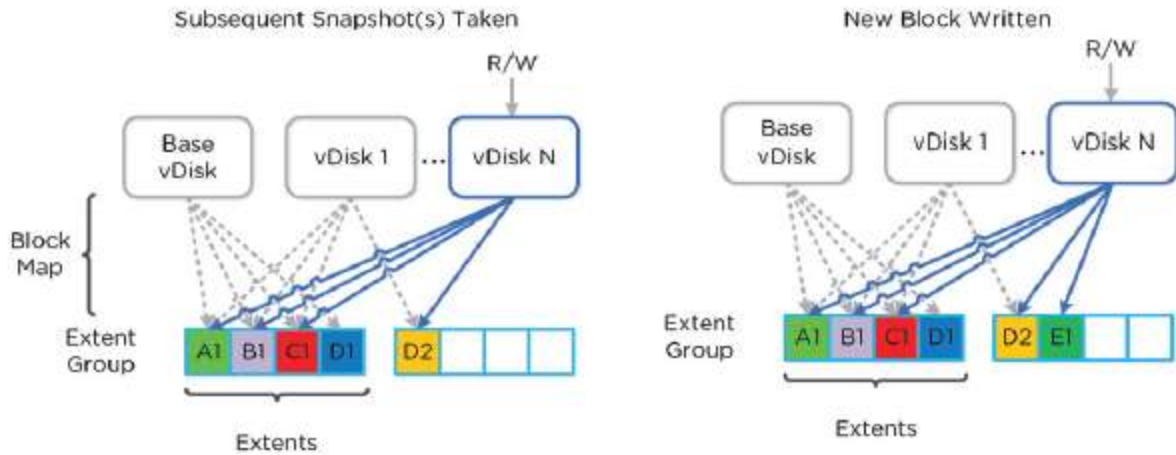


Figure 3.2-33. Multi-snap Block Map and New Write

Image credit: <https://nutanixbible.com>

- When VM/vDisk is cloned, current block mapped is locked
- Clones created
- Previously cloned VM acts as “base disk”

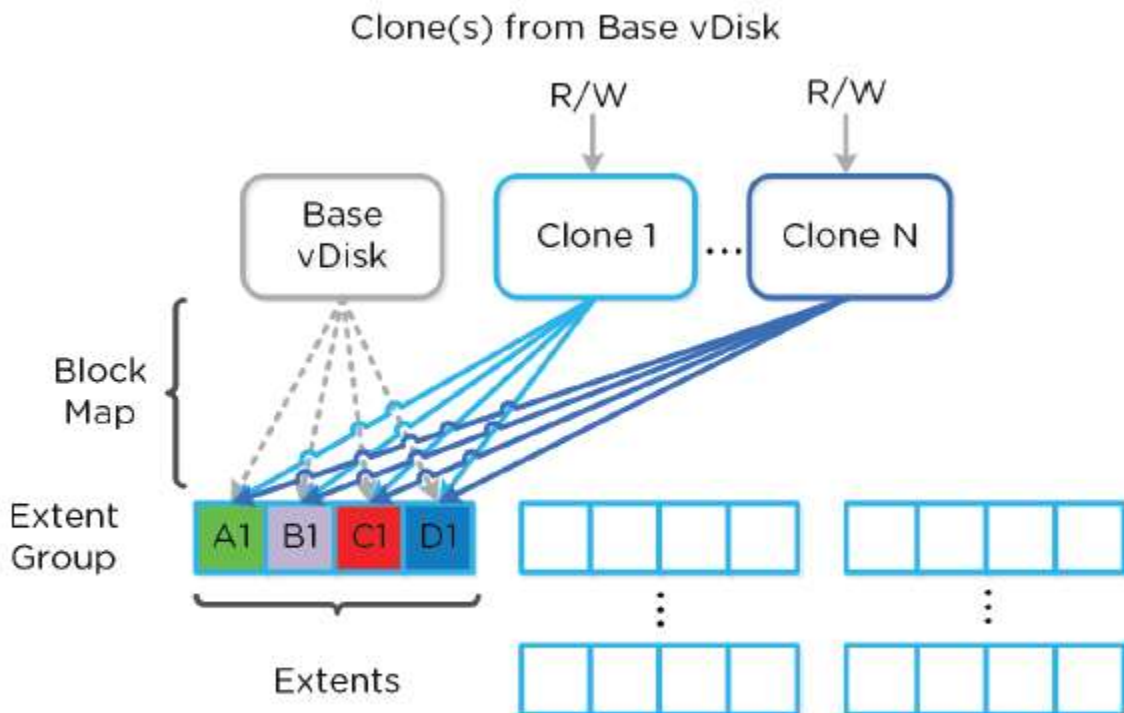


Figure 3.2-34. Multi-Clone Block Maps

Image credit: <https://nutanixbible.com>

- Clones from base VM have their own block map

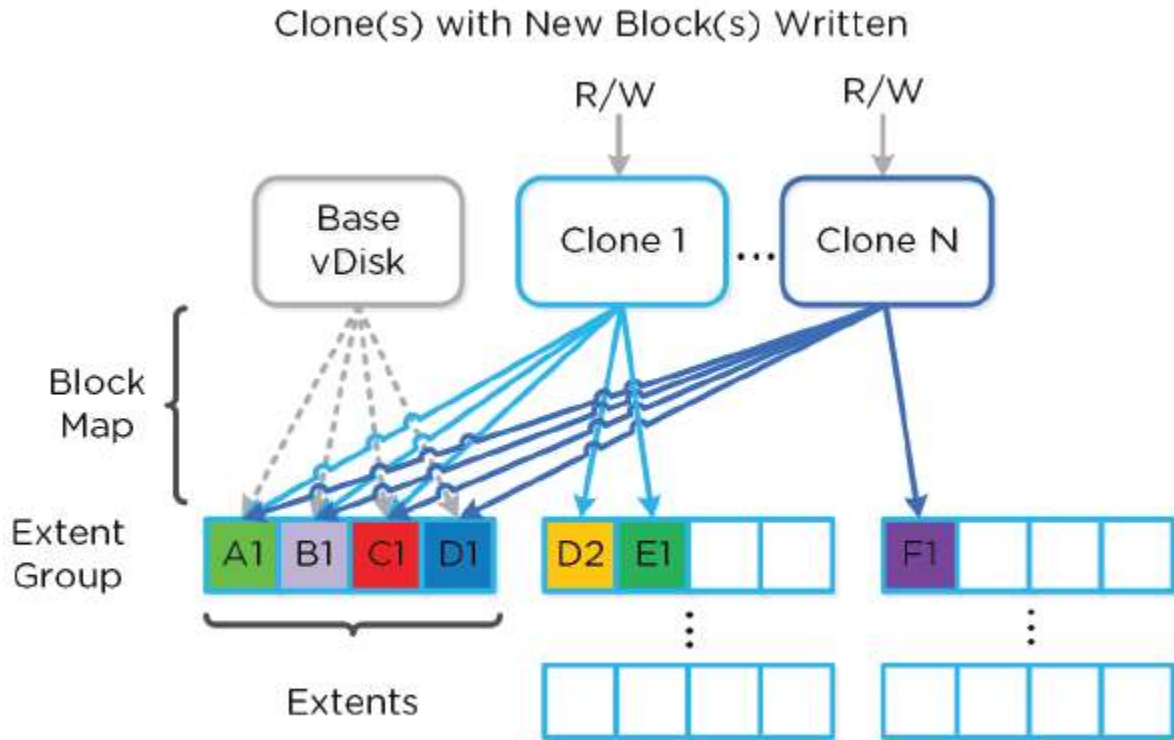


Figure 3.2-35. Clone Block Maps - New Write

Image credit: <https://nutanixbible.com>

- New writes/updates occur there

## App Consistent Snapshots

- Native VSS for queiscing included
- VmQuiesced Snapshot Service (Windows + Linux)

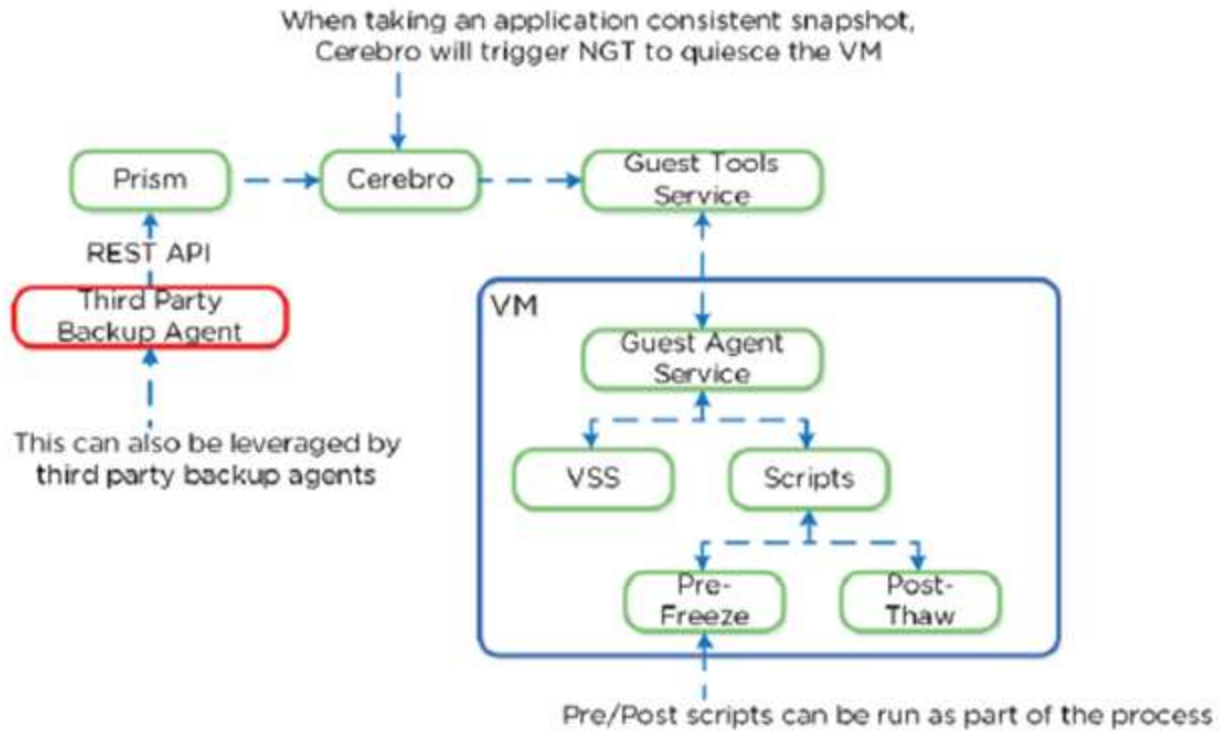


Image credit: <https://nutanixbible.com>

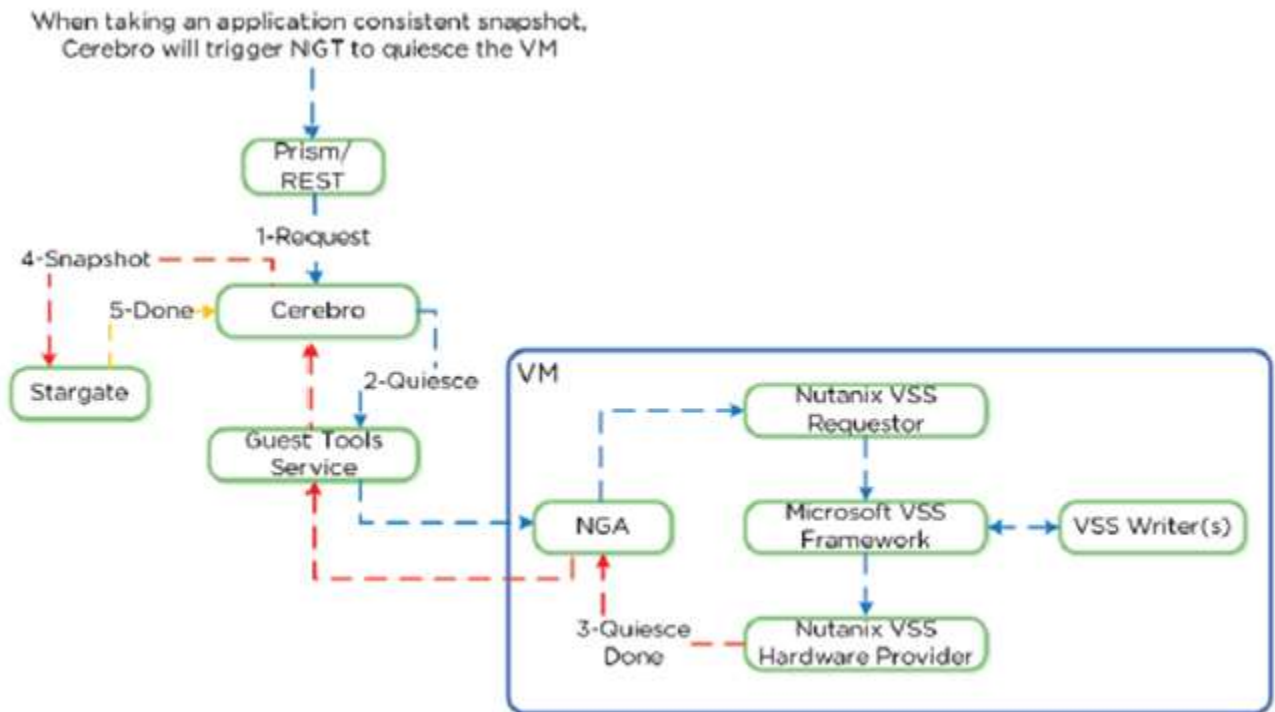


Figure 3.4.15. Nutanix VSS - Windows Architecture

Image credit: <https://nutanixbible.com>

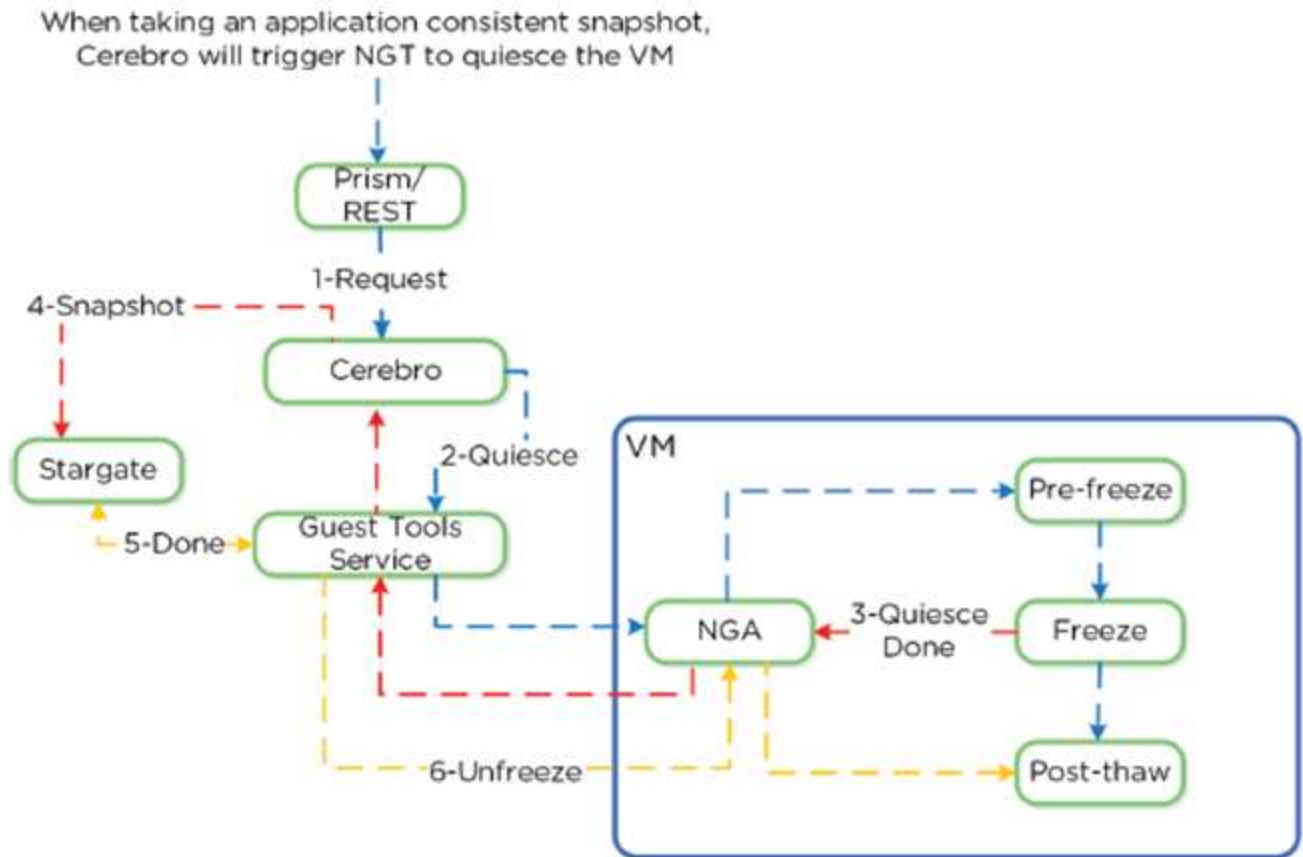


Figure 3.4.17. Nutanix VSS - Linux Architecture

Image credit: <https://nutanixbible.com>

- Post-free: /sbin/pre\_freeze
- Post-thaw: /sbin/post\_thaw

## Eliminating ESXi Stun:

- When delta disks are created, ESXi “stuns” the VM in order to remap disks to new deltas.
- Also occurs when snapshots are deleted.
- During this process, OS cannot execute operations (stuck)
- Duration depends on number of VMDK’s, speed of datastore, etc.
- Nutanix VSS bypasses VMware snapshot/stun process = more efficient

## Shadow Clones

- Distributed caching of vDisks/VM data in multi-reader scenario

- Ex. VDI deployment with many linked clones
- Read requests forwarded to “base VM”
- Read requests occur from more than two remote CVM’s (all read I/O) = vDisk marked immutable
- vDisk then cached locally by each VM
- Allows for each node to have a “base VM”

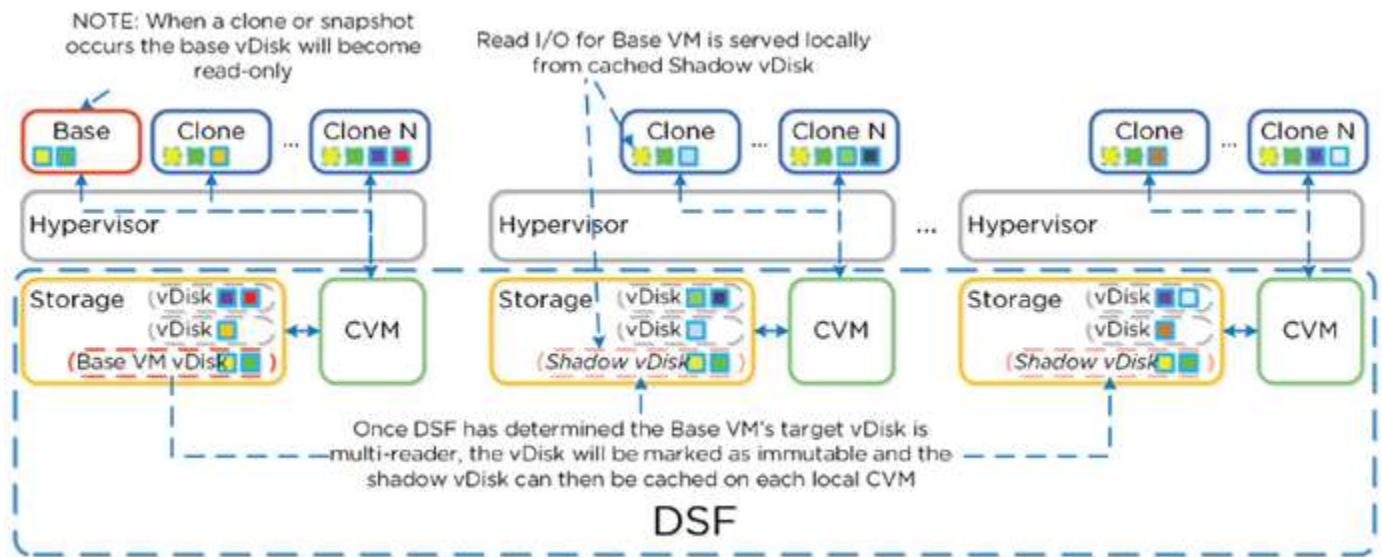


Image credit: <https://nutanixbible.com>

## Configure Deduplication, Compression, and Erasure Coding on Nutanix containers

### Compression

Select the check box to enable compression. A Delay (In Minutes) field appears after checking the box. Enter a zero to enable inline compression or a value (number of minutes) to enable post-write compression, which can begin (up to) that number of minutes after the initial write.



Compression

Perform post-process compression of all persistent data. For inline compression, set the delay to 0.

DELAY (IN MINUTES)

All data in the storage container is compressed when this box is checked. See Compression for guidelines about using compression.

## Deduplication

Select the **CACHE** check box to perform **inline deduplication of read caches** to optimize performance. If you enable this option, the Controller VMs must be configured to have at least 24 GB of RAM. This feature is primarily recommended for full-clone, persistent desktops, and physical to virtual migration use cases. Turning deduplication on for VAAI clone or linked clone environments is not recommended.

DEDUPLICATION

Cache ?

Perform inline deduplication of read caches to optimize performance.

Capacity ?

Perform post-process deduplication of persistent data.

Select the **CAPACITY** check box to perform **post-process deduplication of persistent data**. This option is recommended primarily for full clone, persistent desktops, and physical to virtual migration use cases that need storage capacity savings (not just performance savings from deduplication). It is further recommended that the Controller VMs have at least 32GB of RAM and 300GB SSDs for the metadata disk to use this option.

## Erasure Coding

Select the check box to enable erasure coding. Erasure coding increases the effective or usable capacity on a cluster.

ERASURE CODING ⓘ

Enable

Erasure coding enables capacity savings across solid-state drives and hard disk drives.

## > Data Efficiency (Math)

Node avoided by Data or Parity

Node used for Parity

Node used for Data

NX-8150 (20 TB raw per node)

3 nodes Not recommended

	Cluster Size	Raw	Usable	After Erasure	
	3 nodes	60 TB	30 TB	→ 40 TB	$60/1.5 = 40$
	4 nodes	80 TB	40 TB	→ ~53 TB	$80/1.5 = \sim 53$
	5 nodes	100 TB	50 TB	→ 75 TB	$100/1.33 = 75$
	6 nodes	120 TB	60 TB	→ 96 TB	$120/1.25 = 96$
	7 nodes	140 TB	70 TB	→ 112 TB	$140/1.25 = 112$

10.  
10

## Section 8 – AHV Workload Migration

**Describe the steps needed to perform an ESXi to AHV workload migration from preparation through completion**

# Supported Source Environments

You can migrate VMs from the following source hypervisors to AHV:

- VMware ESXi
- Microsoft Hyper-V

## Unified Extensible Firmware Interface (UEFI) Support Information

In the current AOS and AHV releases, the UEFI implementation is limited. The following table describes the level of support available for various usage scenarios:

Scenario	Support Level
Generation 2 VM (UEFI) migrated from Hyper-V to AHV	Limited support
UEFI VM migrated from ESXi to AHV	Not supported

## Configure a Filesystem Whitelist

A whitelist is a set of addresses that are allowed access to the cluster. Whitelists are used to allow appropriate traffic when unauthorized access from other sources is denied.

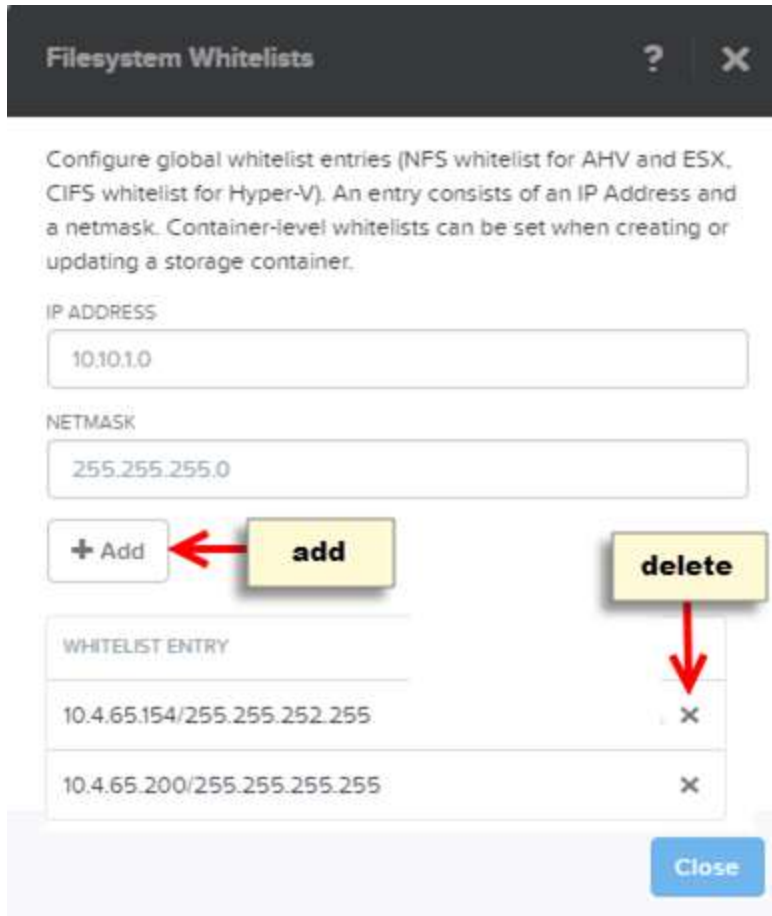
### Settings

Cluster Lockdown

Data at Rest Encryption

**Filesystem Whitelists**

SSL Certificate



## Configure Images

In AHV clusters, you can import and configure operating system ISO and disk image files through the web console. You can also convert previously imported files to the format that AHV uses.

The image service feature allows you to build a store of imported files that you can use to create a CD-ROM from an ISO image or an operating system Disk from a disk image when creating a VM. The image service supports raw, vhd, vhdx, vmdk, vdi, iso, and qcow2 disk formats.

Image create, update, and delete (CUD) behavior depends on whether a cluster (also known as Prism Element) is registered to Prism Central.

- If the Prism Element (PE) cluster is newly created and has never been registered with a Prism Central instance (that is, never managed by Prism Central), all image CUD operations will be allowed on the PE cluster.

- If the PE cluster registered with a Prism Central instance (that is, managed by Prism Central) CUD operations will be blocked on the PE cluster. Ownership of images is migrated from PE to Prism Central. CUD operations must be performed through Prism Central.
- If the PE cluster is unregistered from Prism Central, new images can be created, updated and deleted on the PE clusters. Update operations are blocked on PE cluster for images that previously migrated to Prism Central.
- In the case of a local image local file upload, with more than one PE cluster managed by Prism Central, the image state is active on that PE cluster. All other PE clusters will show the image as inactive. If you create a VM from that image, the image bits are copied to the other PE clusters. The image then appears in an active state on all managed PE clusters.

### Create Image ? ×

NAME

ANNOTATION

IMAGE TYPE

CONTAINER

IMAGE SOURCE

From URL

Upload a file  No file chosen

# Migrate a VM from an ESXi cluster to an AHV cluster

## Windows VM Migration Prerequisites

1. On the source hypervisor, power off all the VMs that you want to migrate.
2. Ensure that the source VMs do not have any hypervisor snapshots associated with them.
3. (Optional) Clone any VMs that you want to preserve.
4. (Optional) Create a storage container on the AHV cluster.
5. (For ESXi source environments) Windows VMs using Unified Extensible Firmware Interface (UEFI) are not supported for migration in AOS or AHV.
6. By default, CD-ROM drives are supported and configured only on the IDE bus.

## VirtIO

1. Install Nutanix VirtIO on all the VMs that need to be migrated. You can download Nutanix VirtIO from the Nutanix Support Portal (see Nutanix VirtIO for Windows).
2. Verify that you have completed all pre-migration tasks. See also Windows VM Migration Prerequisites.
3. From the following scenarios, choose the scenario that applies to you and migrate the Windows VMs to the AHV cluster.

## Scenario 1: The AHV Cluster Can Access the Source Virtual Disk Files over NFS or HTTP

1. Provide the target AHV cluster with read access to the virtual disk files on the NFS or HTTP server. See Providing Read Access to the Nutanix Cluster.
2. Import the virtual disks by using Image Service. See Configuring Images. This step must be performed for each virtual disk on the source VM. Note the following when using Image Service:
3. For a source Nutanix ESXi or Hyper-V cluster, specify the IP address of any of the Controller VMs in the source cluster.

4. For a source non-Nutanix NFS server, specify the IP address the NFS server.
5. Use the Prism web console to create a VM from the imported image. Use the Clone from Image Service option. Also assign a vNIC to the VM. See [Creating a Windows VM on AHV after Migration](#).

## Scenario 2: The Source Virtual Disk Files Are Not Accessible to the AHV Cluster

1. On the target AHV cluster, provide access to the source hypervisor hosts by adding the host IP addresses to the AHV cluster's filesystem whitelist. See [Configuring a Filesystem Whitelist](#).
2. Adding a source hypervisor host's IP address to the target AHV cluster's filesystem whitelist enables the host to mount the target cluster's container as a temporary NFS datastore or SMB share.
3. Mount the target AHV cluster's container on the source hypervisor host as an NFS datastore or SMB share. See [Mounting an AHV Container on a Source Hypervisor Host](#).
4. From the source hypervisor host, copy the virtual disk files from their original location to the temporary NFS datastore or SMB share mounted from the target AHV cluster. See [Migrating VM Disks to AHV Storage](#).
5. In the Prism web console, use Image Service to convert the virtual disk files to the raw format that AHV can use. See [Configuring Images](#).
6. Use the Prism web console to create a VM from the imported image. Use the Clone from Image Service option. Also assign a vNIC to the VM. See [Creating a Windows VM on AHV after Migration](#).

## Scenario 3: The Source Virtual Disk Files Have Been Exported

1. Use an SFTP client to connect to the target AHV cluster. Connect to port 2222 on any Controller VM IP address and log in as the Prism admin user.
2. Optionally, create a subdirectory in the target AHV container to use as a staging area for the virtual disk files. Make sure that the container you use is the container in which the virtual disk files will eventually reside.
3. By using the SFTP client, copy the virtual disk files (\*.flat.vmdk files) to the target AHV container.
4. In the Prism web console, use Image Service to convert the virtual disk files to the raw format that AHV can use. See [Configuring Images](#).

5. Use the Prism web console to create a VM from the imported image. Use the Clone from Image Service option. Also assign a vNIC to the VM. See [Creating a Windows VM on AHV after Migration](#).

## Creating a Windows VM on AHV after Migration

Create a disk from the disk image by clicking **Add New Disk** and completing the indicated fields.

1. **TYPE:** DISK
2. **OPERATION:** CLONE FROM IMAGE
3. **BUS TYPE:** SCSI
4. **CLONE FROM IMAGE SERVICE:** Select the image you created previously from the drop-down menu.
5. Click **Add** to add the disk drive.

The **Path** field is displayed when **Clone from ADSF file** is selected from the **Operation** field. For example, you can specify the image path to copy as `nfs://127.0.0.1/container_name/vm_name/vm_name.vmdk` or `nfs://127.0.0.1/container_name/vm_name/vm_name-flat.vmdk`.

## Post Migration Tasks, Windows

Complete the post-migration tasks in the Prism web console.

1. After the VM is created, the *Received operation to create VM* dialog box appears. Click **View the task details** and then select the VM. The Summary line (middle of screen) displays the VM name with a set of relevant action links on the right.
2. (For Generation 2 VMs migrated from a Hyper-V host) Before you power on a UEFI guest VM, configure the VM with the aCLI option `uefi_boot=True`. For example:

```
acli vm.update vm_id uefi_boot=True
```

- **Note:** Only Generation 2 (Gen 2) Windows VMs that use UEFI to start (boot) are supported for migration. However, support is limited.
3. Click **Power on** to start the Windows VM.
  4. After the VM is started, the *Received operation to power on the VM* dialog box appears. Click **View the task details** and then select the VM. The Summary line (middle of screen) displays the VM name with a



set of relevant action links on the right. Click **Launch Console** to log on to the VM through the console.

5. Configure an IP address for the VM. Follow the prompts in the console to configure an IP address.
6. (For VMs migrated from ESXi) Open the **Control Panel** in the Windows server VM and remove the VMware Tools and other VMware related software.
7. Restart the VM.

## Linux VM Migration Prerequisites

1. Prepare the VM for migration.
  1. Install Nutanix VM Mobility by enabling and mounting the Nutanix Guest Tools on the Linux VM. See Nutanix Guest Tools in the Prism Web Console Guide.
  2. Check that the virtIO drivers are installed.
2. In the Prism web console, add the source hypervisor host IP address to the target AHV cluster's filesystem whitelist. See Configuring a Filesystem Whitelist.
3. Use vSphere Storage vMotion to migrate the VM disk image to the AHV storage container. Mount the Acropolis storage container as a temporary NFS datastore or SMB share.
4. Create a VM and attached the imported disk image.
5. Power on the VM and log in to the VM's console. Optionally, you can uninstall VMWare tools, if installed.

## Linux VM Migration Requirements

- The SUSE/Ubuntu Linux kernel must include appropriate virtIO drivers to migrate Linux servers.
- Ensure virtIO modules are loaded on the VM to be migrated.
- vSphere files on a Acropolis storage container mounted as a temporary NFS Datastore or SMB share.

Minimum AOS version	AOS 4.6.1.1
Minimum AHV version	AHV-20160217.2
Minimum vSphere version	vSphere 5.0 U2
Minimum Ubuntu version	For SCSI bus:Ubuntu 12.04.3 and later Ubuntu 14.04.x

	SUSE 2.6.1 and later For PCI bus: Ubuntu 12.04.2 and earlier
Connectivity type between clusters	AHV network connected Acropolis storage container mounted as a temporary NFS Datastore or SMB share

## Checking VirtIO Module Status

Verify that your kernel has the correct virtIO modules built in.

```
$ grep -i virtio /boot/config-3.0.101-63-default
```

Check the output to verify if the drivers are installed.

```
CONFIG_NET_9P_VIRTIO=m
CONFIG_VIRTIO_BLK=m
CONFIG_SCSI_VIRTIO=m
CONFIG_VIRTIO_NET=m
CONFIG_VIRTIO_CONSOLE=m
CONFIG_HW_RANDOM_VIRTIO=m
CONFIG_VIRTIO=m
#Virtio drivers
CONFIG_VIRTIO_PCI=m
CONFIG_VIRTIO_BALLOON=m
#CONFIG_VIRTIO_MMIO is not set
```

- For CONFIG\_VIRTIO\_PCI and CONFIG\_SCSI\_VIRTIO, the =m output means the VirtIO SCSI driver is built directly into the kernel and is a loadable kernel module.

## Prerequisites for Migrating Ubuntu

Check the Ubuntu version and confirm the installed virtIO drivers on the Ubuntu VM.

1. On vSphere, log into the Ubuntu VM and open a terminal window.
2. Verify that the minimum Ubuntu version is at least 12.04.

```
$ cat /etc/lsb-release
```

3. Output might look similar to the following.

```
DISTRIB_ID=Ubuntu DISTRIB_RELEASE=14.04
DISTRIB_CODENAME=name DISTRIB_DESCRIPTION="Ubuntu 14.04 LTS"
```

4. Check that the virtIO drivers are installed.

```
$ grep -i virtio /boot/config-`uname -r`
```

Check the output to verify if the drivers are installed.

- For CONFIG\_VIRTIO\_PCI, the =y output means the VirtIO PCI driver is built directly into the kernel.
  - For CONFIG\_SCSI\_VIRTIO, the =m output means the VirtIO SCSI driver is built directly into the kernel and is a loadable kernel module.
5. Confirm that the virtio\_scsi module is built into the initramfs image.
    - a. Copy the initramfs image to a temporary location.

```
nutanix@ubuntu12045:~$ cp -p /boot/initrd.img-`uname -r`  
/tmp/initrd.img-`uname -r`.gz
```

- b. Check the virtIO SCSI module is built in

```
nutanix@ubuntu12045:~$ zcat /tmp/initrd.img-`uname -r`.gz | cpio  
-it | grep virtio
```

## Migrating VM Disks to AHV Storage

Migrate the virtual disk files from the source hypervisor to the temporarily mounted container. It is recommended to create one or more sub-directories in the container on the target AHV cluster. Copying the virtual disk files to sub-directories helps organize migrated virtual disk files. The virtual disk files are only required until converted by Image Service and can be easily identified and deleted when no longer required.

- If the source hypervisor is ESXi, do one of the following:
  - Use Storage vMotion to move the virtual disk files of running VMs from their original location to the temporarily mounted NFS datastore on the target AHV cluster. when specifying a datastore for Storage vMotion, you can enter the file path `nfs://127.0.0.1/container_name/vm_name/vm_name.vmdk`.
  - Replace container\_name with the name of the storage container where the image is placed and replace vm\_name with the name of the VM where the image is placed.
  - After Storage vMotion moves the virtual disk files, use Image Service to convert the files to the raw format that AHV can use. When Storage vMotion is complete, shut down the source VM.

- Use Vmkfstools to copy the virtual disk files from the datastore on the source hypervisor host. The following commands create a subdirectory in the container on the target AHV cluster and then move the virtual disk files to the subdirectory.

```
~ # mkdir /vmfs/volumes/container_name/subdirectory

~ # vmkfstools -i /vmfs/volumes/original_datastore/win7-
vm/virt_disk_file.vmdk

/vmfs/volumes/container_name/subdirectory/win7-vm.vmdk
```

- Replace container\_name with the name of the container on the target AHV cluster, subdirectory with a name for the subdirectory, and virt\_disk\_file with the name of the virtual disk file.

## Creating a Linux VM on AHV after Migration

Create a disk from the disk image by clicking the **+ New Disk** button and completing the indicated fields.

1. **TYPE:** DISK
2. **OPERATION:**
  - For Ubuntu VMs, select **CLONE FROM ADS FILE**
  - For SUSE VMs, select **CLONE FROM NDFS FILE**
3. **BUS TYPE:** SCSI
4. **PATH:** From the drop down list, choose the path name. Type a forward slash and pick the file name, */container\_name/vm\_name/flat\_vmdk\_file*.
5. Replace container\_name with the name of the storage container.
  - Replace vm\_name with the name of the VM you migrated.
  - Replace flat\_vmdk\_file For example, a file path might look similar to */default-container-32395/Ubuntu12345VMW/Ubuntu12345VMW-flat.vmdk*.
6. Click **Add** to add the disk driver.

## Post Migration Tasks, Linux

1. Log in to the Prism Web Console and navigate to **Home > VM**. Choose the **Table** view.
2. Select the VM you created in the table and click **Power on** from the action links.

3. After the VM powers on, click **Launch Console** and log into the VM. This opens a Virtual Network Computing (VNC) client and displays the console in a new tab or window. This option is available only when the VM is powered on.
4. If they are installed, remove the VMware tools from the VM.
  - From a tar installation:

```
$ sudo /usr/bin/vmware-uninstall-tools.pl
```

- From an rpm installation:

```
rpm -e VMwareTools
```

## Section 9 – Acropolis Services

### Define and differentiate Acropolis Block Services (ABS) and Acropolis File Services (AFS)

#### Acropolis Block Services (ABS)

- Exposes backend DSF to external consumers via iSCSI
- Use Cases:
  - Oracle RAC
  - MSCS
  - Containers
  - Bare-metal
  - Exchange on vSphere
- Constructs:
  - Data Services IP: Cluster Wide VIP for iSCSI logins
  - Volume Group: iSCSI target/group of disk devices
  - Disks: Devices in Volume Group
  - Attachment: Permissions for IQN access
  - Backend = VG's disk is just a vDisk on DSF

#### Path High-Availability

- Data Services VIP leveraged for discovery
- Single address without knowing individual CVM IP's

- Assigned to iSCSI master
- In event of failure, new master elected with address

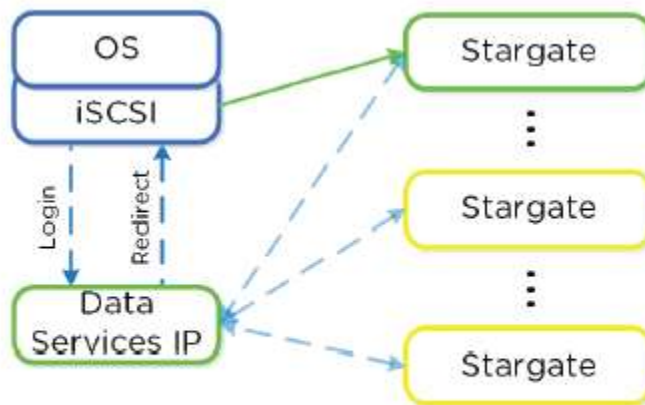


Figure 3.3.17. Block Services - Login Redirect

Image credit: <https://nutanixbible.com>

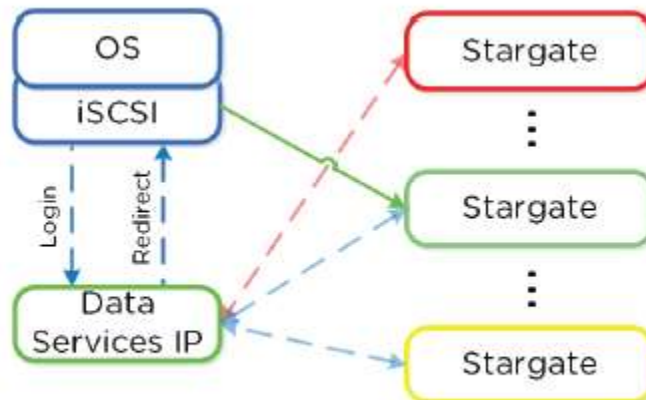


Figure 3.3.18. Block Services - Failure Handling

Image credit: <https://nutanixbible.com>

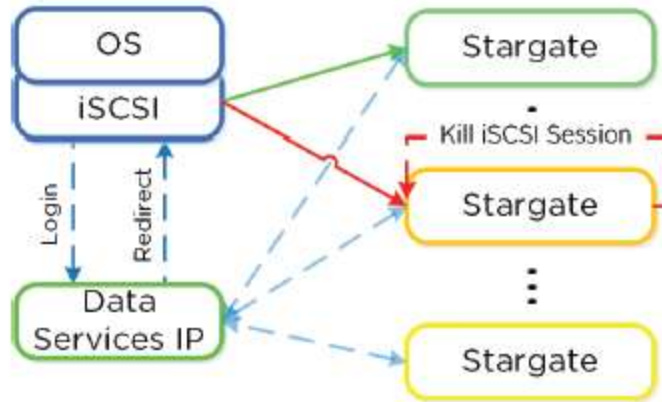


Figure 3.3.19. Block Services - Failback

Image credit: <https://nutanixbible.com>

- With this, client side MPIO is not needed
- No need to check “enable multi-path” in Hyper-V

## Multi-Pathing

- iSCSI protocol mandates single iSCSI session/target
- 1:1 relationship from Stargate to target
- Virtual targets automatically created per attached initiator and assigned to VG
- Provides iSCSI target per disk device
- Allows each disk device to have its own iSCSI session, hosted across multiple Stargates

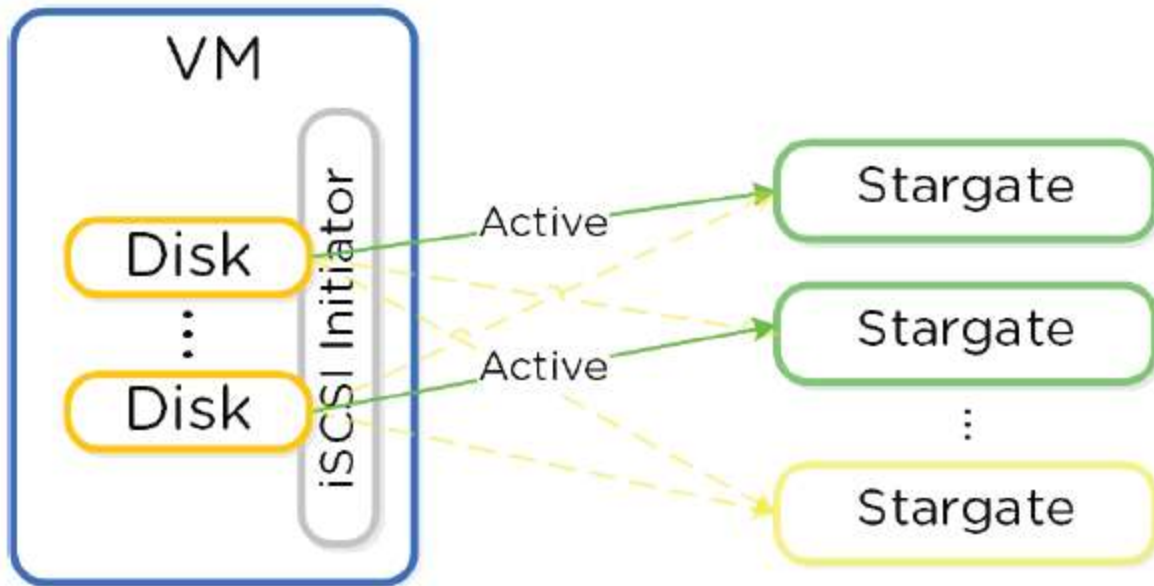


Figure 3.3.22. Block Services - Multi-Path

Image credit: <https://nutanixbible.com>

- Load balancing occurs during iSCSI session establishment per target
- Performed via hash
- TRIM supported (UNMAP)

## Acropolis File Services (AFS)

- SMB is only supported protocol
- File Services VMs run as agent VM's
  - **Minimum of 3 VM's** deployed by default for scale
- Transparently deployed
- Integrated into AD/DNS
- During install, the following are created:
  - AD Computer Account
  - AD SPN for file server and each FSVM
  - DNS for FS pointing to all FSVM(s)
  - DNS for each FSVM
  - Each FSVM uses Acropolis Volumes API accessed via ingest iSCSI
  - Allows FSVM to connect to any iSCSI target in event of outage



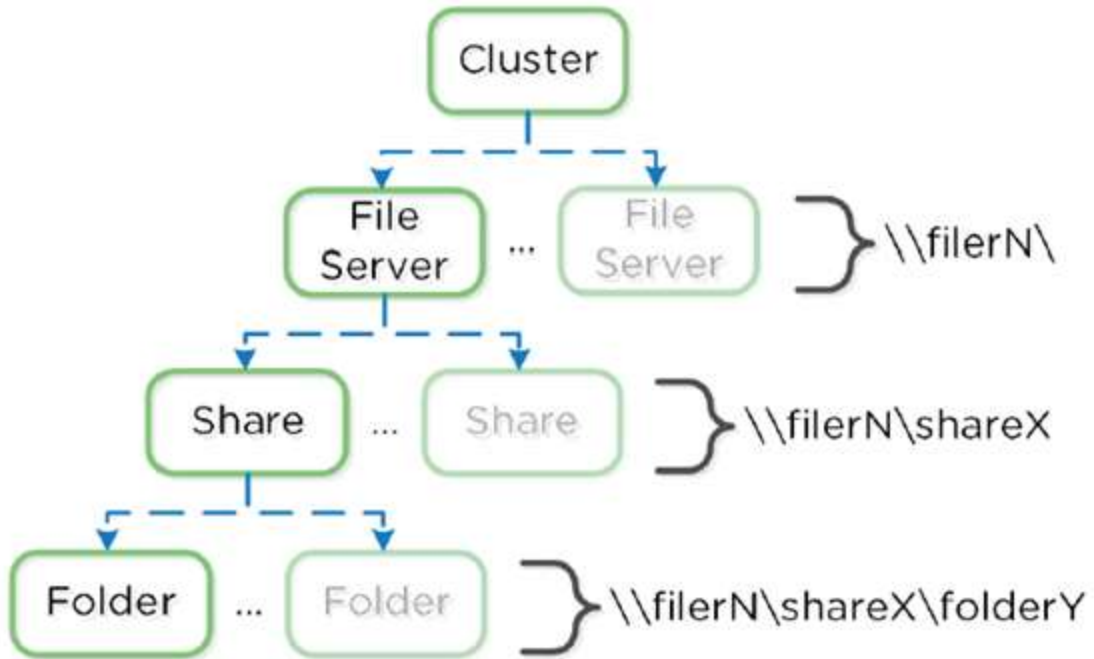


Figure 3.3.22. File Services Mapping

Image credit: <https://nutanixbible.com>

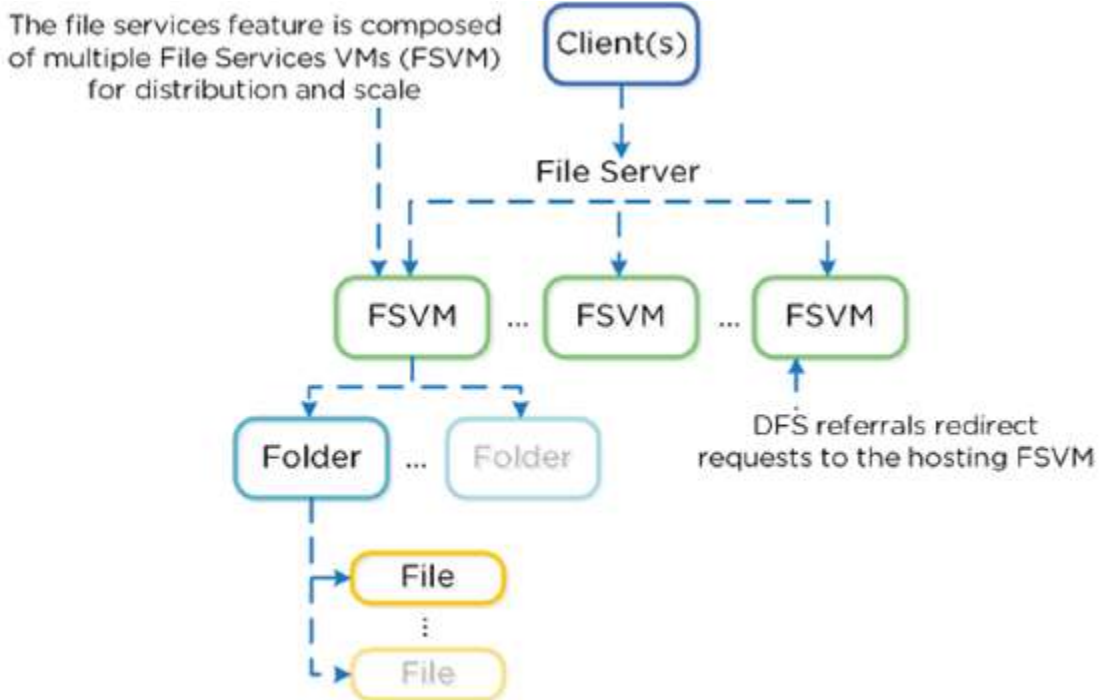


Figure 3.3.23. File Services Detail

Image credit: <https://nutanixbible.com>

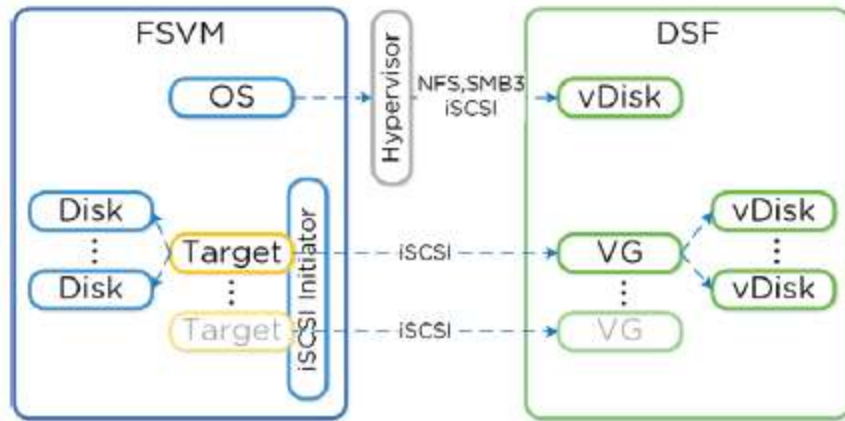


Figure 3.3.25. FSVM Storage

Image credit: <https://nutanixbible.com>

- For availability in Linux, DM-MPIO (Linux Multipathing) is leveraged within FSVM with active path set to Local CVM.

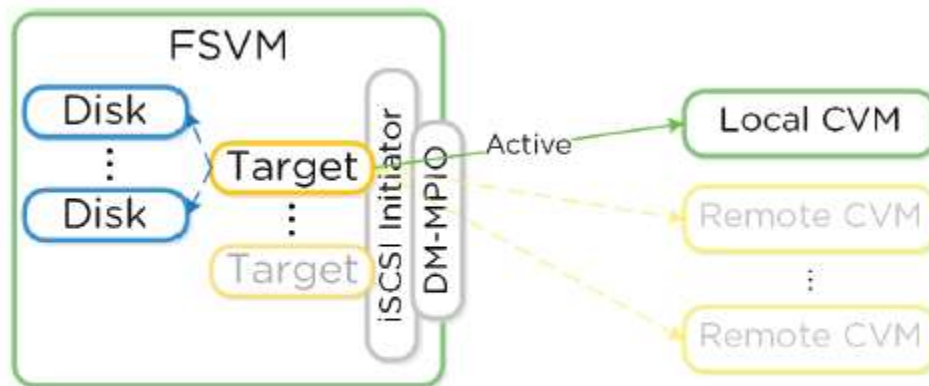


Figure 3.3.26. FSVM MPIO

Image credit: <https://nutanixbible.com>

- In event of failure, DM-MPIO will active failover path on remote CVM.

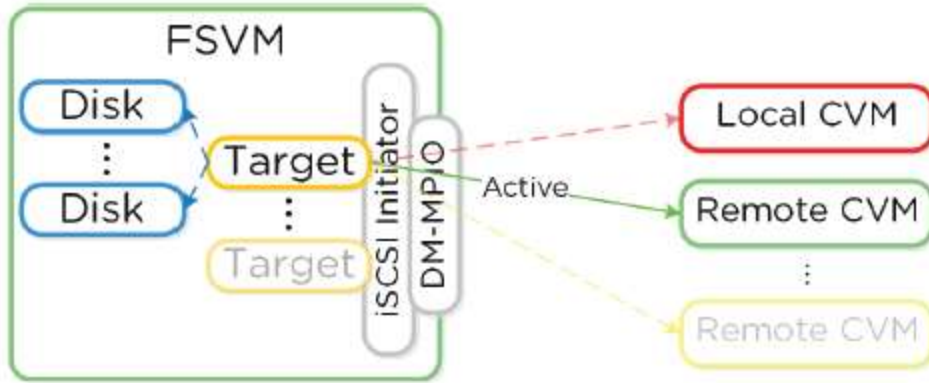


Figure 3.3.27. FSVM MPIO Failover

Image credit: <https://nutanixbible.com>

- Each FSVM will have IP which clients use to communicate to FSVM as part of DFS.

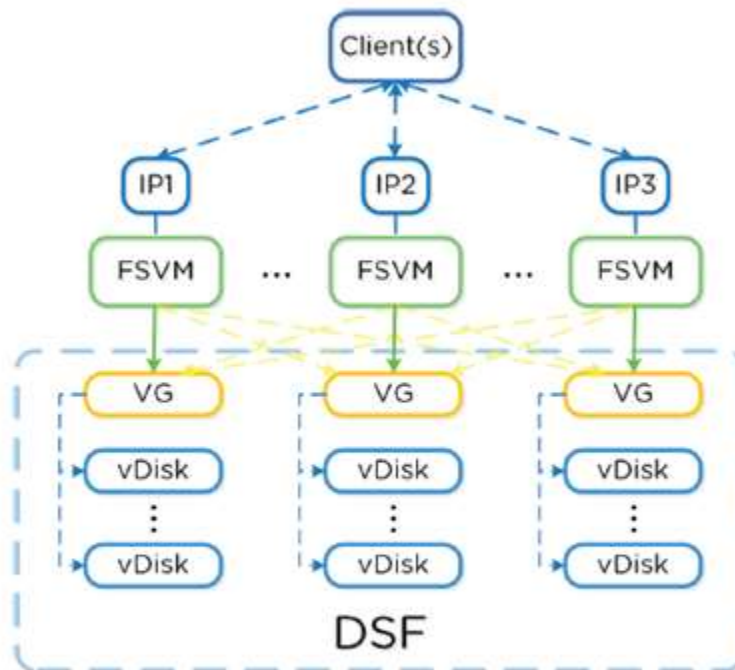


Figure 3.3.28. FSVM Normal Operation

Image credit: <https://nutanixbible.com>

- In event of failure, VG and IP of failed FSVM will be taken over by another FSVM:

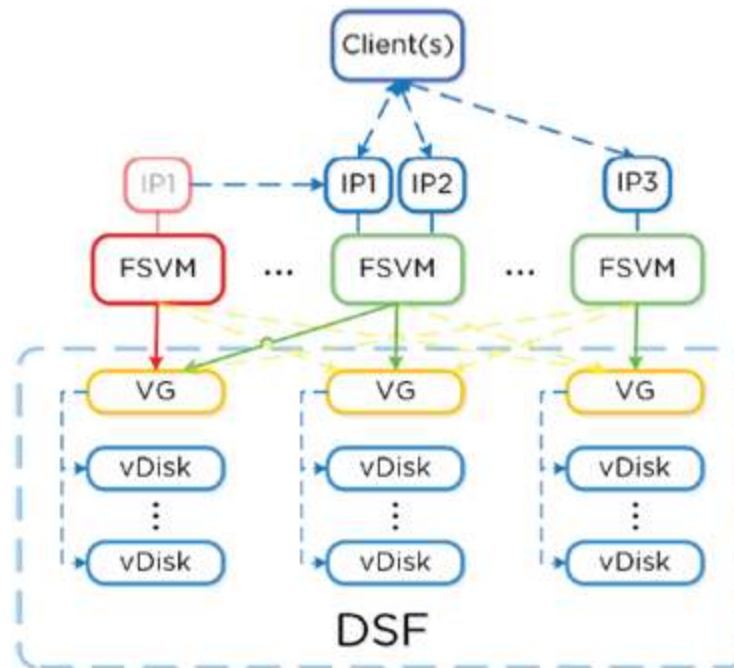


Figure 3.3.29. FSVM Failure Scenario  
 Image credit: <https://nutanixbible.com>

- When failed FSVM comes back and is stable, it will re-take its IP and VG.

## Configure Acropolis Block Services (ABS)

### Requirements and Limitations

- Ensure that ports 3260 and 3205 are open on any clients accessing the cluster where Acropolis Block Services is enabled.
- You must configure an external data services IP address in **Cluster Details** available from the Prism web console.
- Synchronous Replication or Metro Availability are not currently supported for volume groups.
- Linux guest VM clustering is not supported for solutions other than Oracle RAC with Oracle Clusterware and Microsoft Windows Failover Clusters.

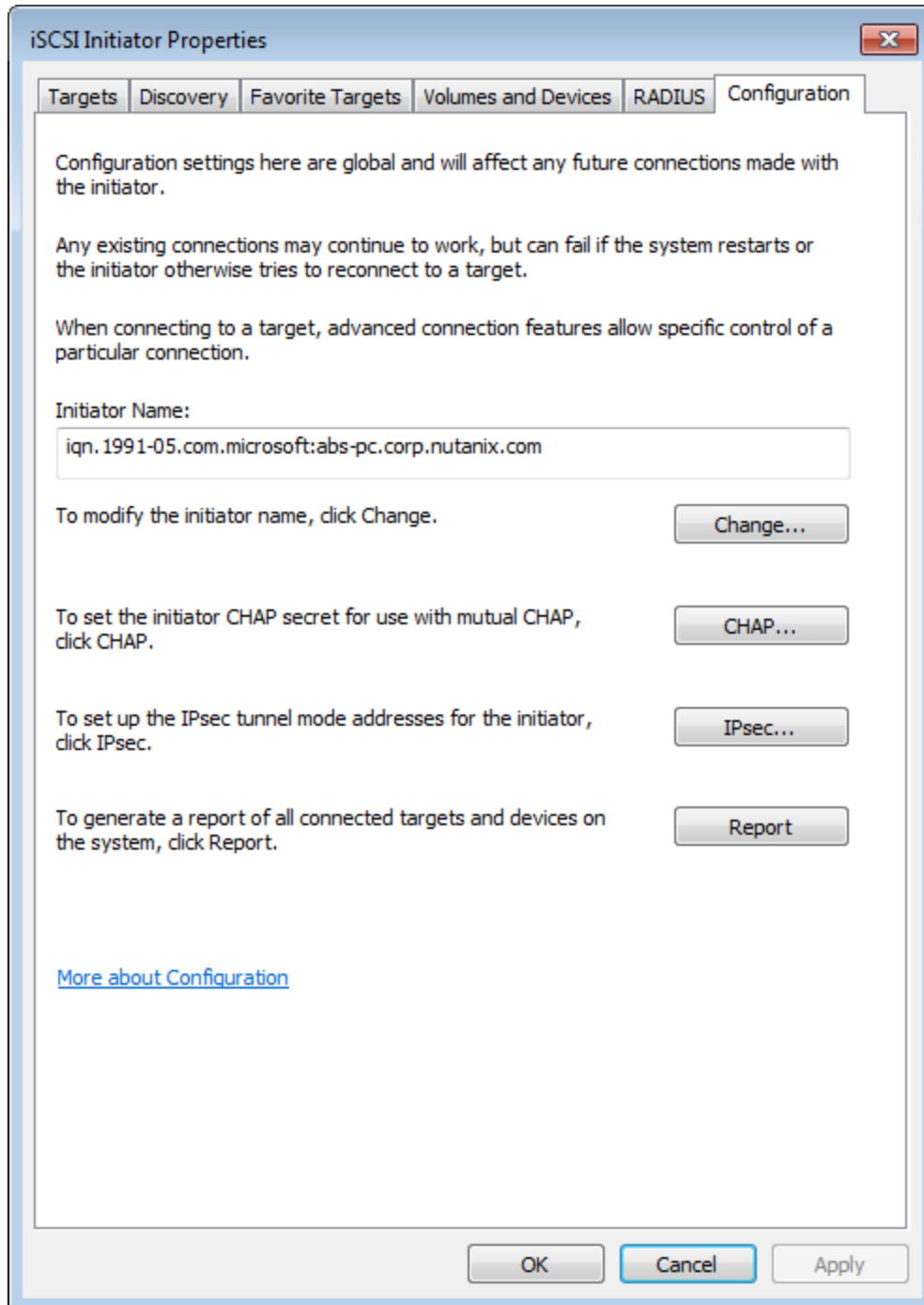
### Support Operating Systems

- Microsoft Windows Server 2012 R2
- Microsoft Windows Server 2008 R2
- Red Hat Enterprise Linux 6.7

- Oracle Linux 6.x
- Oracle Linux 7.x

## Obtaining the Windows iSCSI Initiator Name

1. Open the *iSCSI Initiator Properties* window.
  1. If *iSCSI Initiator* is not available from Administrative Tools, you can open it by clicking **Start**, typing iSCSI in the search box, and clicking **iSCSI Initiator** under *Programs*.
2. In the *iSCSI Initiator Properties* window, click the **Configuration** tab.
  1. The **Initiator Name** field contains the initiator IQN name. Copy this name for use with the procedures in this section.



## Obtaining the Linux iSCSI Initiator Name

On the Linux client, open a terminal window and type:

```
$ sudo cat /etc/iscsi/initiatorname.iscsi
```

For example, the command displays:

InitiatorName=iqn.1991-05.com.redhat:8ef967b5b8f

## Creating a Volume Group

The screenshot shows a 'Create Volume Group' dialog box with the following sections:

- General Configuration**
  - NAME**: A text input field containing 'Name'.
  - ISCSI TARGET NAME**: A text input field containing 'iSCSI target name'.
  - DESCRIPTION**: A text input field containing 'Description'.
- DISKS**: A button labeled '+ Add new disk'.
- Share across multiple iSCSI Initiators or multiple VMs**: A checkbox that is currently unchecked.
- INITIATORS**
  - IQN**: A text input field with an empty space and a '+ Add' button to its right.
- VMs**: A button labeled '+ Attach to a VM' with an information icon (i) to its right.

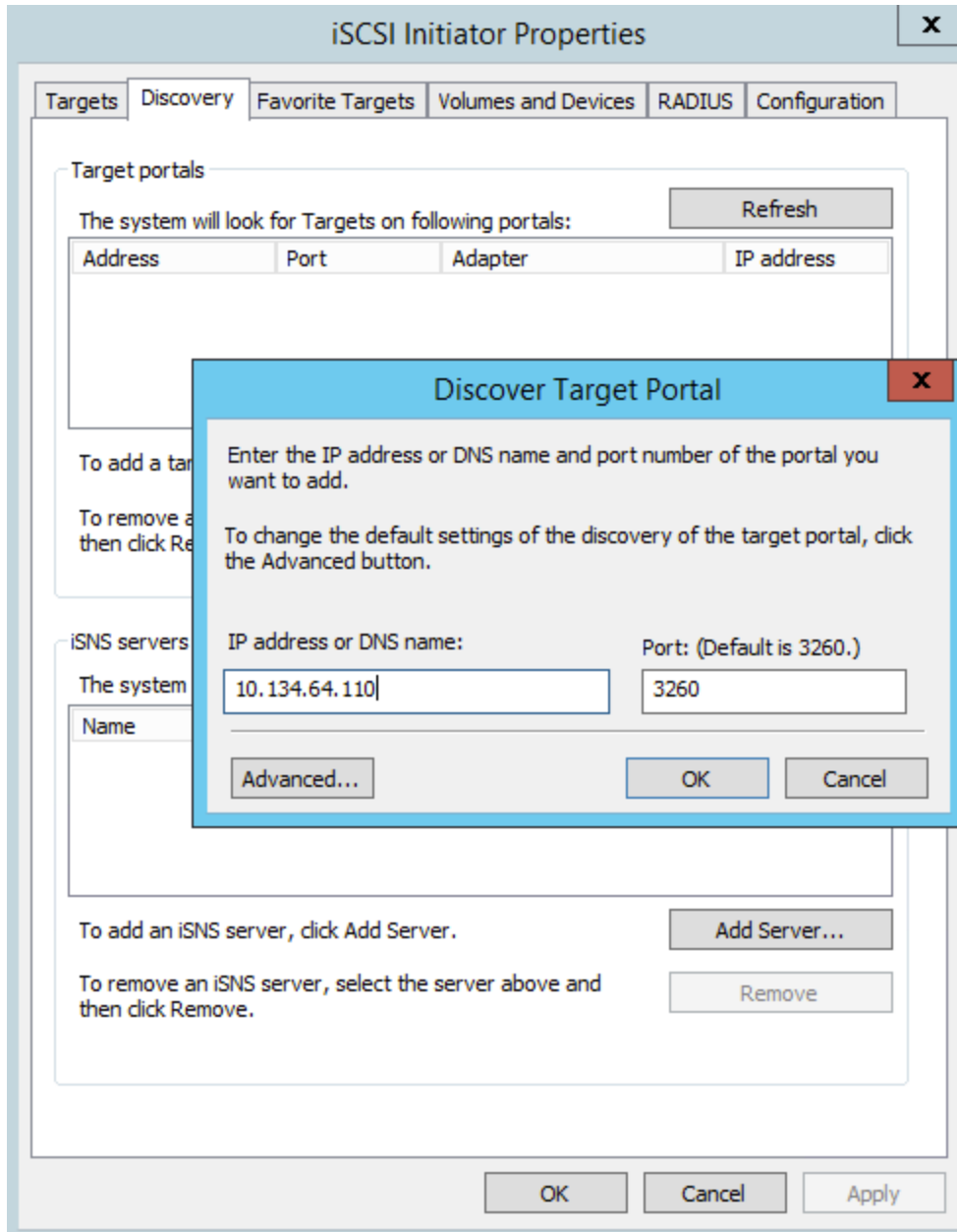
At the bottom right of the dialog, there are two buttons: 'Cancel' and 'Save'.

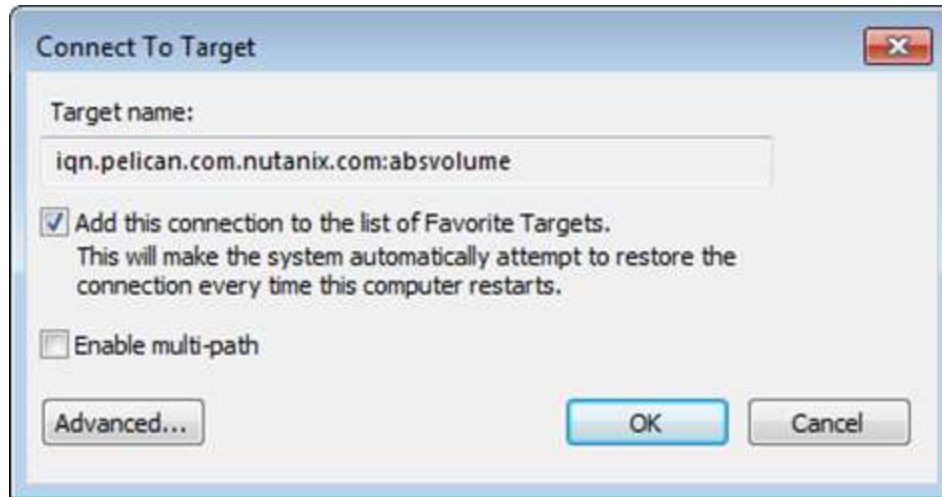
1. To add one or more disks to the volume group, do the following:
  - a. In the **Disks** section, click **Add New Disk**.

- b. In the **Add Disk** dialog box, in **Operation**, specify **Allocate on Container**.
  - c. In **Container**, select the storage container to use from the pull-down list. The list includes all containers created on this cluster.
  - d. In **Size**, enter the disk size in GiBs.
  - e. Click **Add**.
  - f. Repeat these steps to add another disk for this volume group.
2. To enable multiple initiators to access the volume group, select **Share across multiple iSCSI initiators or multiple VMs**.
3. To whitelist the initiators that must access the volume group, in the **Initiators** section, do the following:
  - a. In **IQN**, enter the Initiator iSCSI Qualified Name (IQN) of the VM.
  - b. Click **Add**.

## Discovering and Connecting to ABS from Windows







## Discovering and Connecting to ABS from Linux

Ensure that the iSCSI service is started.

For Red Hat Enterprise Linux 6.0:

```
$ sudo /etc/init.d/iscsi status
```

For Red Hat Enterprise Linux 6.7.

```
$ sudo service iscsid status
```

Discover the ABS target by specifying the external data services IP address on the default port 3260.

```
$ sudo /sbin/iscsiadm -mode discovery -type sendtargets \ -portal  
external_data_services_IP_address:3260
```

The command output will appear similar to `external_data_services_IP_address:3260, 1 iqn_name`, where `iqn_name` is the ABS target IQN.

Connect the ABS target by specifying `iqn_name` from the previous command.

```
$ sudo /sbin/iscsiadm -mode node -targetname iqn_name \ -portal  
external_data_services_IP_address:3260,1 -login
```

## Configure Acropolis File Services (AFS)

# Preparations

- Ensure each cluster has a minimum configuration of 4 vCPUs and 12 GiB of memory available on each host.
  - Ensure you have configured or defined internal and external networks.
  - An Active Directory, Domain Name Server, and a Network Time Protocol Server.
  - You need Active Directory administrator credentials, enterprise administrator credentials, and at least domain administrator credentials to complete this step.
1. Log in to the web console.
  2. Navigate to the file server by clicking **Home > File Server** in the left corner.
  3. Click **+ File Server**.

**New File Server** ? X

1. File Server Tech Preview · 2. **Define File Server** · 3. Configure Networks · 4. Join Active Domain

**File Server Details**

**NAME**

demo

**Capacity**  
The default configuration can be edited based on your needs.

**FILE SERVER SIZE**

1024 GiB

**NUMBER OF FILE SERVER VMS (MIN 3)**

3

**NUMBER OF VCPUS PER FILE SERVER VM**

4

**MEMORY PER FILE SERVER VM**

12 GiB

Back Cancel Next

## New File Server

1. File Server Tech Preview · 2. Define File Server · 3. **Configure Networks** · 4. Join Active Domain

### Internal & External Network Config (VLANs)

Select an Internal network which communicates between File Server VMs and CVMS.

vlan.36

Select an external network for clients to communicate with File Server VMs.

vlan.37

DOMAIN NAME SERVERS (COMMA SEPARATED)

XX.X.XX.XX

NTP SERVERS (COMMA SEPARATED)

pool.ntp.org

Back Cancel Next

## New File Server

1. File Server Tech Preview · 2. Define File Server · 3. Configure Networks · 4. **Join Active Domain**

### Active Directory Details

This is the user account location to join.

DOMAIN NAME

like: dom.companyname.com

### Credentials

Username should have admin privileges in the domain to join. [Learn more](#)

USERNAME

user\_name

PASSWORD

.....

Back Cancel Create

# Managing a File Server

After creating a file server, use the web console to update every file server in the cluster, expand the file server by adding more file server VMs, modify vCPU or RAM, leave or join an active directory domain and delete a file server. To accomplish one or more of these tasks, do the following.



# Expand a File Server

## Expand File Server



NUMBER OF FILE SERVER VMS



Adding a VM will require AD credentials for the following domain.

REALM NAME

CREDENTIAL

The user should be the name of the account with admin privileges in the domain you are joining. [Learn more](#)

USERNAME

PASSWORD

Cancel

Save

### Update File Server

NUMBER OF VCPUS PER FILE SERVER VM

MEMORY PER FILE SERVER VM

 GiB

Internal network settings to be used for the File Server VM created.

INTERNAL NETWORK (VLANS)

External network settings to be used for the File Server VM created.

EXTERNAL NETWORK (VLANS)

DOMAIN NAME SERVERS (COMMA SEPARATED)

NTP SERVERS (COMMA SEPARATED)

Cancel Save

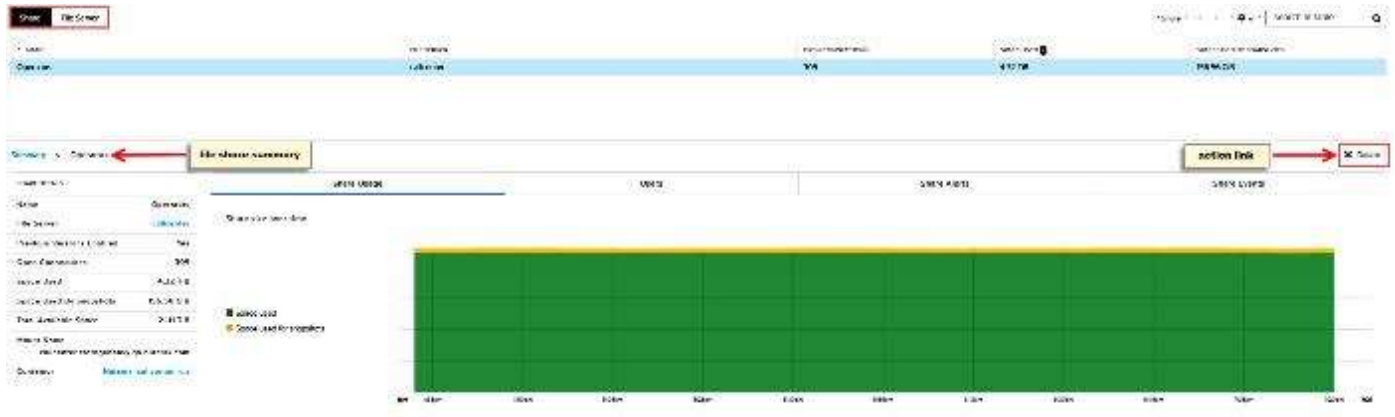
## Creating a File Share

**Create Share**
?
X

**NAME**

**FILE SERVER**

**ENABLE PREVIOUS VERSION**



## Determine and implement storage services based on a given workload

### Acropolis Block Services

Acropolis Block Services can support use cases including but not limited to:

- iSCSI for Microsoft Exchange Server. ABS enables Microsoft Exchange Server environments to use iSCSI as the primary storage protocol.
- Shared storage for Windows Server Failover Clustering (WSFC). ABS supports SCSI-3 persistent reservations for shared storage-based Windows clusters, commonly used with Microsoft SQL Server and clustered file servers.



- Bare-metal environments. ABS enables existing server hardware separate from a Nutanix cluster to consume the Acropolis Distributed Storage Fabric (DSF) resources. Workloads not targeted for virtualization can also use the DSF.

## Acropolis File Services

- Enable you to share files among user work stations or VMs in a centralized protected location to help eliminate the requirement for a third-party file server.
- Acropolis File Services uses a scale-out architecture that provides Server Message Block (SMB) file shares to Windows clients for home directory and user profiles. Acropolis File Services consist of three or more file server VMs (FSVM). There is one file server maximum per cluster image. A set of file server VMs is also known as a Acropolis File Services cluster. Multiple file server clusters can be created on a Nutanix cluster.

## Section 10 – Data Resiliency

### Describe the concept of the Redundancy Factor and related requirements

### What is the different between Redundancy Factor and Replication Factor?

**Redundancy Factor** (aka FT – Fault Tolerance) in the simplest terms, is the number of components that a Nutanix cluster can withstand at any time +1. These components include disks, NIC's, and nodes. For example, in a two block environment and the **default Redundancy Factor of 2**, you can lose a disk, NIC or node and still maintain no data loss.

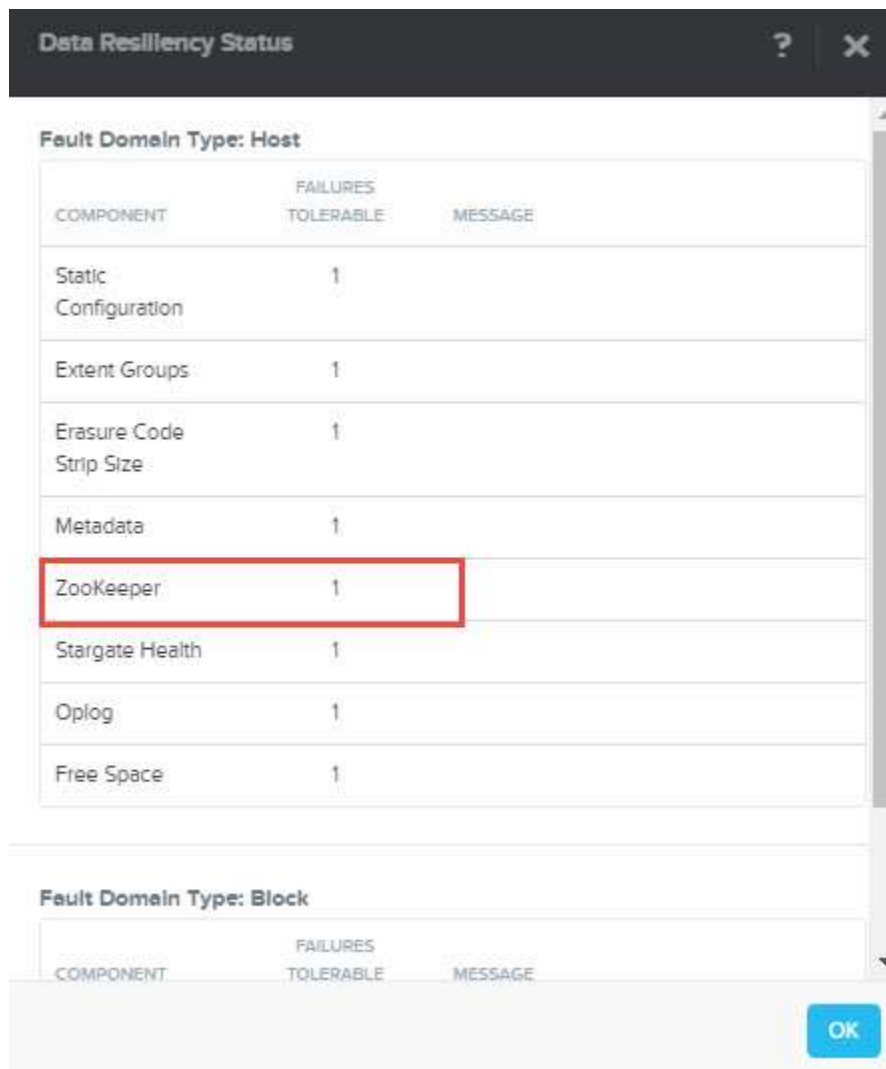
This is not to be confused with **Replication Factor**, which is the number of times that any piece of data is duplicated on the system. This is directly tied to the Redundancy Factor setting. For example, if you have Redundancy Factor of 2 on the cluster, the only option for Replication Factor on Storage Containers is 2. A Redundancy Factor of 3 on the cluster will allow for a Replication Factor of 2 or 3 for Storage Containers.

**NOTE #1:** You can have a container with RF=2 and another container with RF=3 configured on the same Nutanix cluster.

**NOTE #2:** Changing container replication Factor from RF=2 to RF=3, consumes more storage space on Nutanix cluster (because systems have to keep 3 copies of the VM data which are on containers with RF=3).

## Redundancy Factor (FT)

By default, Nutanix clusters have redundancy factor 2, which means they can tolerate the **failure of a single drive, NIC or node**. The larger the cluster, the more likely it is to experience multiple failures. Without redundancy factor 3, multiple failures cause cluster unavailability until the failures are repaired.



The screenshot shows a 'Data Resiliency Status' dialog box. It is divided into two sections: 'Fault Domain Type: Host' and 'Fault Domain Type: Block'. Each section contains a table with columns for 'COMPONENT', 'FAILURES TOLERABLE', and 'MESSAGE'. In the 'Host' section, the 'ZooKeeper' row is highlighted with a red border. The 'Block' section is partially visible at the bottom.

COMPONENT	FAILURES TOLERABLE	MESSAGE
Static Configuration	1	
Extent Groups	1	
Erasure Code Strip Size	1	
Metadata	1	
ZooKeeper	1	
Stargate Health	1	
Oplog	1	
Free Space	1	

COMPONENT	FAILURES TOLERABLE	MESSAGE
-----------	--------------------	---------

Default Redundancy Factor 2

# Redundancy Factor 3 (FT3)

Redundancy factor 3 has the following requirements:

- A cluster must have at least five nodes for redundancy factor 3 to be enabled.
- For guest VMs to tolerate the simultaneous failure of two nodes or drives in different blocks, the data must be stored on storage containers with replication factor 3.
- Controller VM must be configured with enough memory to support redundancy factor 3.

## *Changing the Redundancy Factor from 2 to 3*

**NOTE:** Setting the Redundancy Factor can only be configured before the cluster is created.

1. To increase the cluster from redundancy factor 2 to redundancy factor 3, log on to any Controller VM in the cluster through SSH and start the nCLI.
2. To view the cluster redundancy factor state:

```
ncli> cluster get-redundancy-state
```

3. Output similar to the following shows redundancy factor 2.

```
Current Redundancy Factor : 2  
Desired Redundancy Factor : 2  
Redundancy Factor Status :  
kCassandraPrepareDone=true;kZookeeperPrepareDone=true
```

4. Set the cluster to redundancy factor 3.

```
ncli> cluster set-redundancy-state desired-redundancy-factor=3
```

5. Output similar to the following is displayed.

```
Current Redundancy Factor : 2  
Desired Redundancy Factor : 3  
Redundancy Factor Status : -
```

6. The nCLI output might take several minutes to update the redundancy factor.
7. Verify that the redundancy factor is now 3.

```
ncli> cluster get-redundancy-state
```

8. Output similar to the following shows redundancy factor 3.

```
Current Redundancy Factor : 3  
Desired Redundancy Factor : 3  
Redundancy Factor Status :  
kCassandraPrepareDone=true;kZookeeperPrepareDone=true
```

**Data Resiliency Status** ? X

**Fault Domain Type: Host**

COMPONENT	FAILURES TOLERABLE	MESSAGE
Static Configuration	2	
ZooKeeper	2	
Stargate Health	2	
Oplog	2	
Erasure Code Strip Size	2	
Extent Groups	1	Based on placement of extent group replicas the cluster can tolerate a maximum of 1 node failure(s)
Metadata	2	
Free Space	2	

**Fault Domain Type: Block**

OK

Post Configuration to Redundancy Factor 3

# Identify Data Resiliency requirements and policies related to a Nutanix Cluster

## Data Resiliency Levels

The following table shows the level of data resiliency (simultaneous failure) provided for the following combinations of replication factor, minimum number of nodes, and minimum number of blocks.

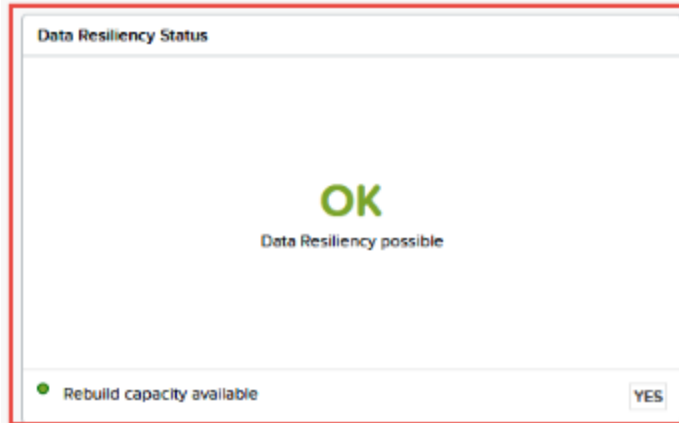
Replication Factor	Minimum Number of Nodes	Minimum Number of Blocks	Data Resiliency
2	3	1	1 node or 1 disk failure
2	3	3 (minimum 1 node each)	1 block or 1 node or 1 disk failure
3	5	2	2 nodes or 2 disk failures
3	5	5 (minimum 1 node each)	2 blocks or 2 nodes or 2 disks
3	6	3 (minimum 2 nodes each)	1 block or 2 nodes or 2 disks
Metro Cluster	3 nodes at each site	2	1 cluster failure

Once block fault tolerance conditions are met, the cluster can tolerate a specific number of block failures:

- A **replication factor two** or replication factor three cluster with three or more blocks can tolerate a **maximum failure of one block**.
- A **replication factor three** cluster with five or more blocks can tolerate a **maximum failure of two blocks**.

Block fault tolerance is one part of a resiliency strategy. It does not remove other constraints such as the availability of disk space and CPU/memory resources in situations where a significant proportion of the infrastructure is unavailable.

The state of block fault tolerance is available for viewing through the Prism Web Console and Nutanix CLI.



## Replication Factor

Replication Factor is the number of times that any piece of data is duplicated on the system. This is directly tied to the Redundancy Factor setting. For example, if you have Redundancy Factor of 2 on the cluster, the only option for Replication Factor on Storage Containers is 2. A Redundancy Factor of 3 on the cluster will allow for a Replication Factor of 2 or 3 for Storage Containers.

Replication Factor	Requirement	Example
Replication factor 2	There must be at least 3 blocks populated with a specific number of nodes to maintain block fault tolerance. To calculate the number of nodes required to maintain block fault tolerance when the cluster RF=2, you need twice the number of nodes as there are in the block with the most or maximum number of nodes.	There must be at least 5 blocks populated with a specific number of nodes to maintain block fault tolerance. To calculate the number of nodes required to maintain block fault tolerance when the cluster replication factor 3 you need four times the number of nodes as there are in the block with the most or maximum number of nodes
Replication factor 3	If a block contains 4 nodes, you need 16 nodes distributed across the remaining (non-failing) blocks to maintain block fault tolerance for that cluster. X = number of nodes in the block with the most nodes. In this case, 4 nodes in a block. 4X = 16 nodes in the remaining blocks	If a block contains 4 nodes, you need 16 nodes distributed across the remaining (non-failing) blocks to maintain block fault tolerance for that cluster. X = number of nodes in the block with the most nodes. In this case, 4 nodes in a block. 4X = 16 nodes in the remaining blocks.

- Replication Factor (RF) + checksum to ensure redundancy/availability
  - RF3 = minimum of 5 nodes (metadata will be RF5)

- RF configured via Prism and at Container level
- All nodes participate in OpLog replication = linear performance
- When data is written, checksum is computed and stored as part of its metadata.
- Data then drained to extent store where RF is maintained.
- In case of failure, data is re-replicated amongst all nodes to maintain RF.
- Checksum is computed to ensure validity on every read.
- In case of no match, replica of data will be read and replace non-valid copy.
- Data consistently monitored to ensure integrity
- Stargate's scrubber operation scans through extent groups to perform checksum validation when disks aren't heavily utilized

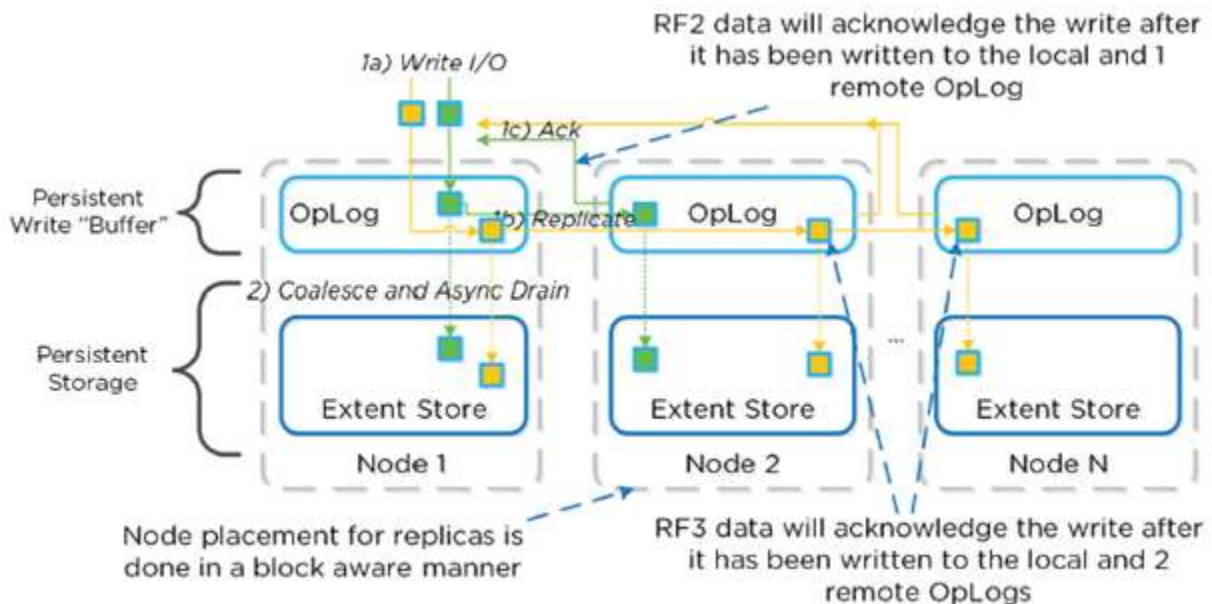


Image credit: <https://nutanixbible.com>

- Nutanix built with idea that hardware will eventually fail
- Systems designed to handle failures in elegant/non-disruptive manner

## Describe and differentiate component, service, and CVM failover processes such as Disk Failure, CVM Failure, and Node Failure

## Disk Failure

- Monitored via SMART data
- Hades responsible for monitoring
- VM impact:
  - HA Event: NO
  - Failed I/O: NO
  - Latency: NO
- In event of failure, Curator scan occurs immediately
- Scans metadata to find data previously hosted on failed disk
- Re-replicates (distributes) to nodes throughout cluster
- All CVM's participate

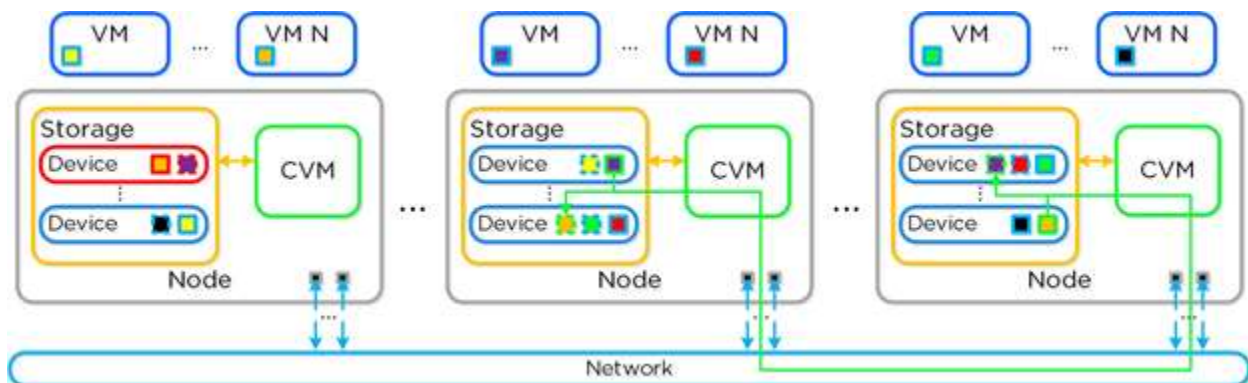


Image credit: <https://nutanixbible.com>

## CVM Failure

- Failure = I/O's redirected to other CVM's in cluster
- VM impact:
  - HA Event: NO
  - Failed I/O: NO
  - Latency: Potentially higher given I/O's are over network (not local)
- ESXi/Hyper-V handle via CVM Autopathing = leverages HA.py (happy) where routes are modified to forward traffic from internal address (192.168.5.2) to external IP of another CVM.
- Keeps datastore intact
- Once local CVM is back online, route is removed and local CVM takes back I/O
- KVM = iSCSI multipathing leveraged
  - Primary path = local CVM, other two paths = remote CVM



# Node Failure

- VM impact:
  - HA Event: Yes
  - Failed I/O: NO
  - Latency: NO
- VM HA event will occur, restarting VMs on other nodes
- Curator will find data previously hosted on node and replicate
- In event node is down for prolonged period of time, downed CVM will be removed from metadata ring.
- Will re-join after up and stable

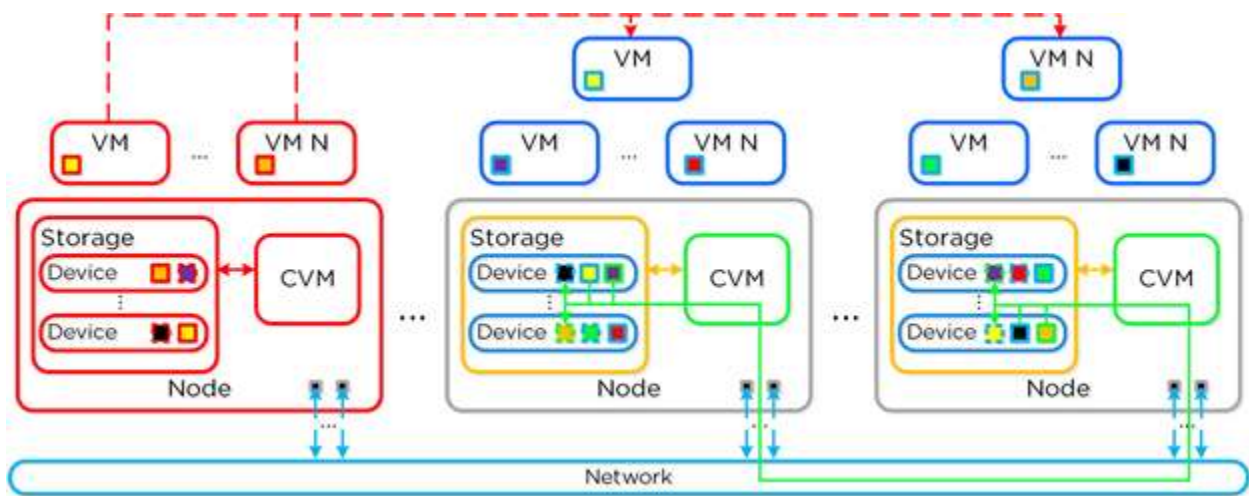


Image credit: <https://nutanixbible.com>

## Section 11 – Data Protection

**Describe and differentiate Nutanix data protection technologies such as NearSync, Cloud Connect, and Protection Domains**

### NearSync

Building upon the traditional asynchronous (async) replication capabilities mentioned previously; Nutanix has introduced support for near synchronous replication (NearSync).

NearSync provides the best of both worlds: zero impact to primary I/O latency (like async replication) in addition to a very low RPO (like sync replication (metro)). This allows users have a very low RPO without having the overhead of requiring synchronous replication for writes.

This capability uses a new snapshot technology called light-weight snapshot (LWS). Unlike the traditional vDisk based snapshots used by async, this leverages markers and is completely OpLog based (vs. vDisk snapshots which are done in the Extent Store).

**Mesos** is a new service added to manage the snapshot layer and abstract the complexities of the full/incremental snapshots. Cerebro continues to manage the high-level constructs and policies (e.g. consistency groups, etc.) whereas Mesos is responsible for interacting with Stargate and controlling the LWS lifecycle.

The following figure shows an example of the communication between the NearSync components:

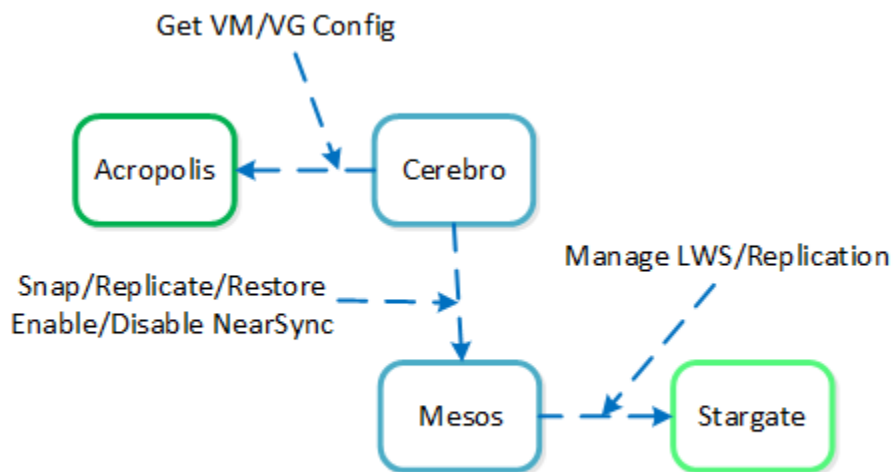


Figure. NearSync Component Interaction

Image credit: <https://nutanixbible.com>

When a user configures a snapshot frequency  **$\leq 15$  minutes, NearSync is automatically leveraged**. Upon this, an initial seed snapshot is taken then replicated to the remote site(s). Once this completes in  $< 60$  minutes (can be the first or n later), another seed snapshot is immediately taken and replicated in addition to LWS snapshot replication starting. Once the second seed

snapshot finishes replication, all already replicated LWS snapshots become valid and the system is in stable NearSync.

In the event NearSync falls out of sync (e.g. network outage, WAN latency, etc.) causing the LWS replication to take > 60 minutes, the system will automatically switch back to vDisk based snapshots. When one of these completes in < 60 minutes, the system will take another snapshot immediately as well as start replicating LWS. Once the full snapshot completes, the LWS snapshots become valid and the system is in stable NearSync. This process is similar to the initial enabling of NearSync.

Some of the advantages of NearSync are as follows.

- Protection for the mission-critical applications. Securing your data with minimal data loss in case of a disaster, and providing you with more granular control during the restore process.
- No latency or distance requirements that are associated with fully synchronous replication feature.
- Allows resolution to a disaster event in minutes.
- To implement the NearSync feature, Nutanix has introduced a technology called Lightweight Snapshots (LWS) to take snapshots that continuously replicates incoming data generated by workloads running on the active cluster. The LWS snapshots are created at the metadata level only. These snapshots are stored in the LWS store, which is allocated on the SSD tier. LWS store is automatically allocated when you configure NearSync for a protection domain.

## Cloud Connect

- Extends to cloud providers
  - Only AWS, Azure
- Cloud remote site is spun up based on Acropolis
  - All native capabilities available
- Storage performed with “cloud disk”. Logical disk backed by S3 or BlobStore.

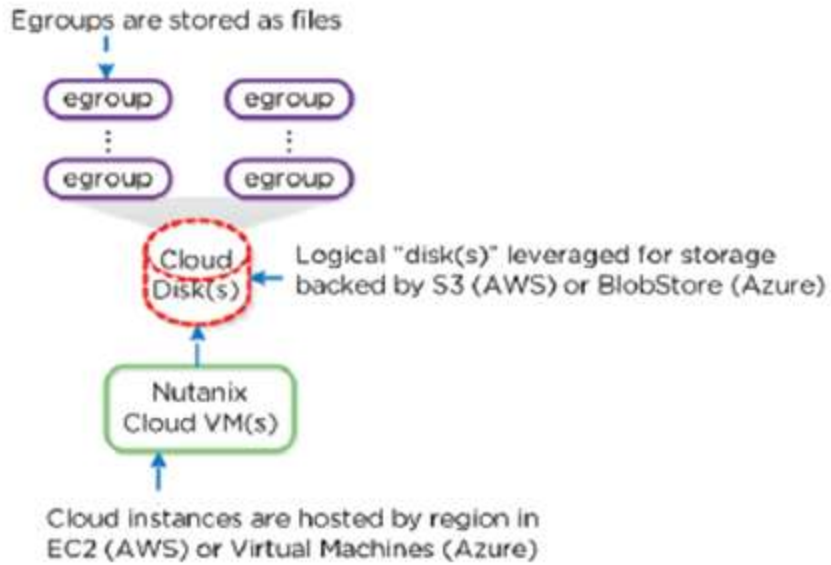


Figure 3.4.23. Cloud Connect Region

Image credit: <https://nutanixbible.com>

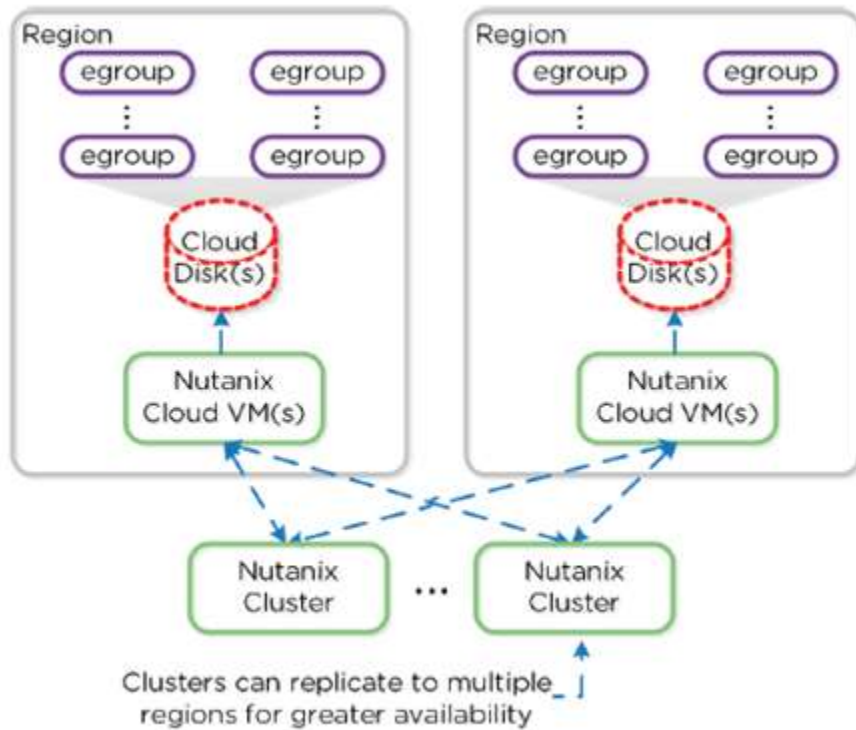


Figure 3.4.24. Cloud Connect Multi-region

Image credit: <https://nutanixbible.com>

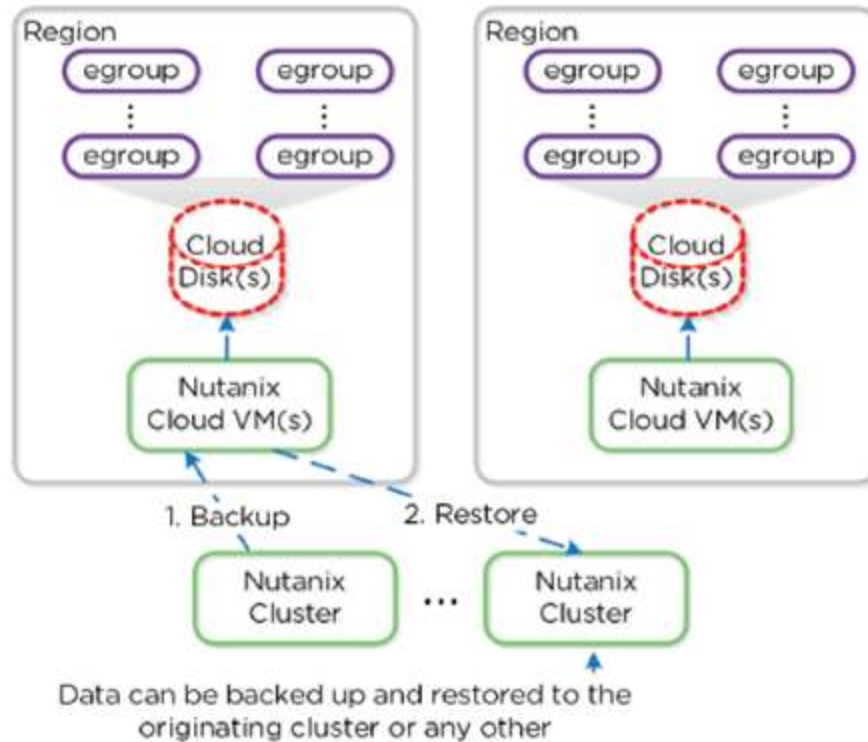


Figure 3.4.25. Cloud Connect - Restore

Image credit: <https://nutanixbible.com>

## Metro Availability

- Compute spans two locations with access to shared pool of storage (stretch clustering)
- Synchronous replication
- Acknowledge writes
- If link failure, clusters will operate independently
- Once link is re-established, sites are resynchronized (deltas only)

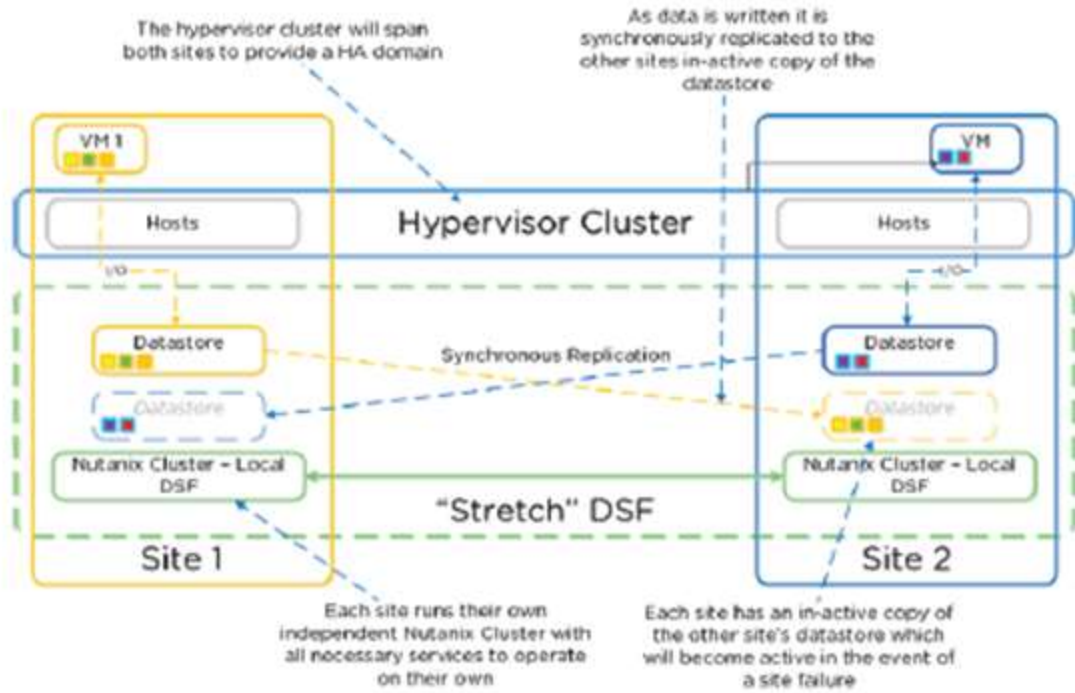


Figure 3.4.26. Metro Availability - Normal State

Image credit: <https://nutanixbible.com>

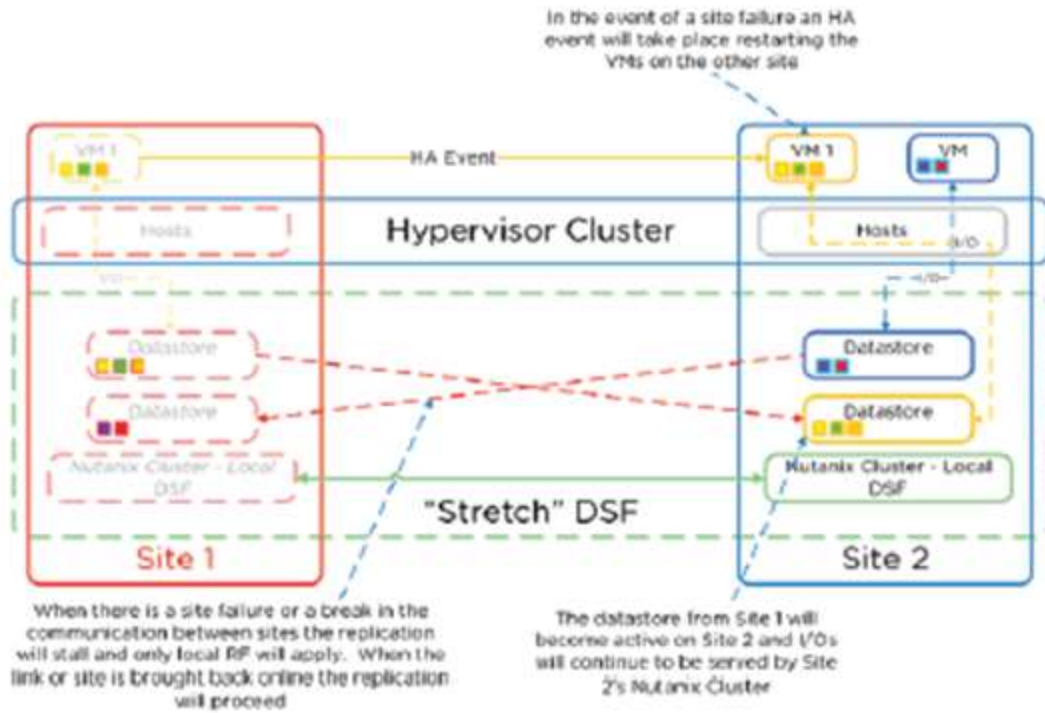


Figure 3.4.27. Metro Availability - Site Failure

Image credit: <https://nutanixbible.com>

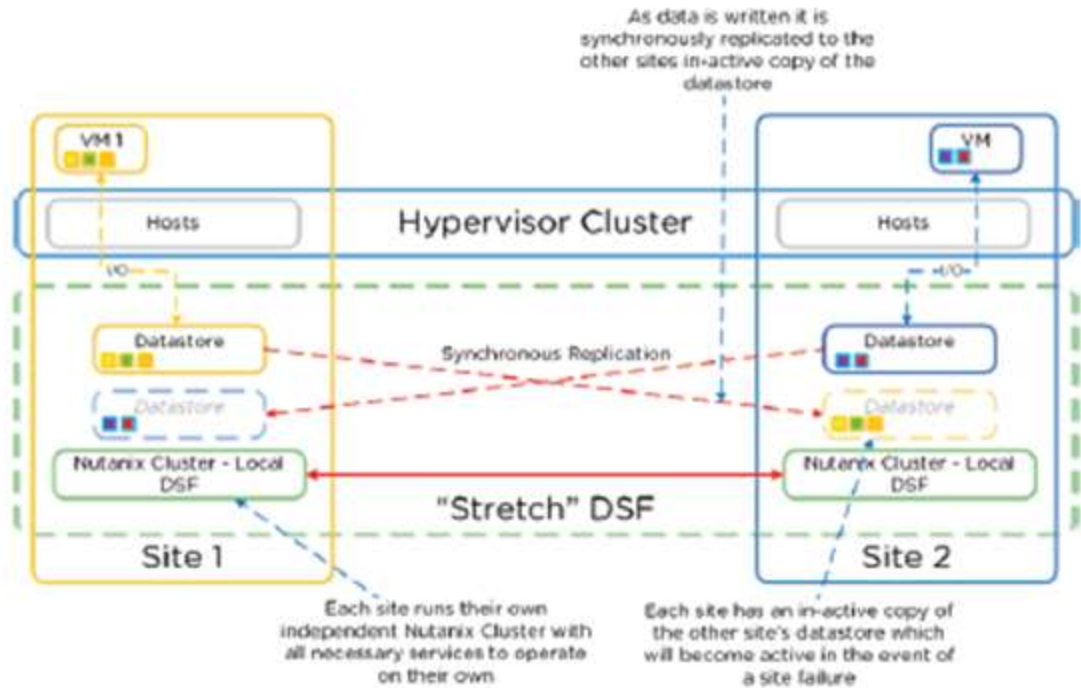


Figure 3.4.28. Metro Availability - Link Failure

Image credit: <https://nutanixbible.com>

## Protection Domains

- Macro group of VM's/files to protect
- Replicated together on schedule
- Protect full container or individual VMs/files
- Tip: create multiple PDs for various service tiers driven by RTO/RPO.

**Consistency Group (CG):** subset of VMs/files in PD crash-consistent

- VMs/files part of PD which need to be snapshotted in crash-consistent manner
- Allows for multiple CG's

**Snapshot Schedule:** snap/replication schedule for VMs in PD/CG

**Retention Policy:** number of local/remote snaps to retain.

- Remote site must be configured for remote retention/replication



**Remote Site:** remote cluster as target for backup/DR

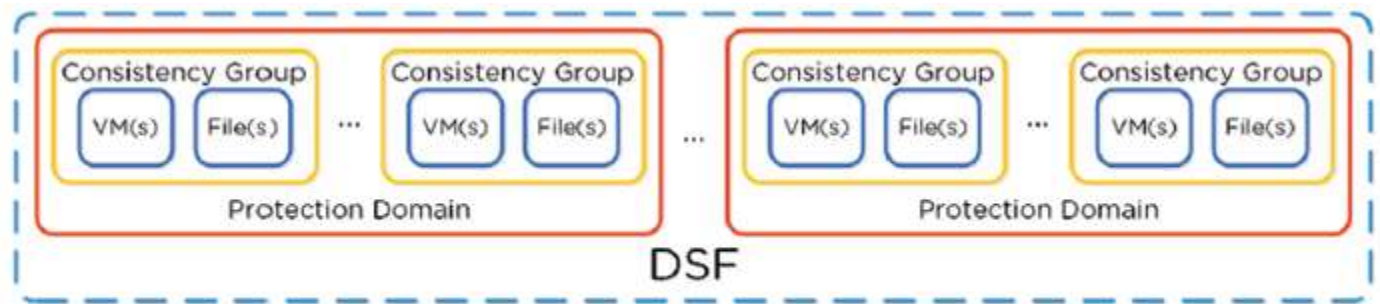


Figure 3.4.1. DR Construct Mapping

Image credit: <https://nutanixbible.com>

## Replication and Disaster Recovery

- Cerebro manages DR/Replication
- Runs on every node with one master
- Accessed via CVM:2020

## Replication Topologies

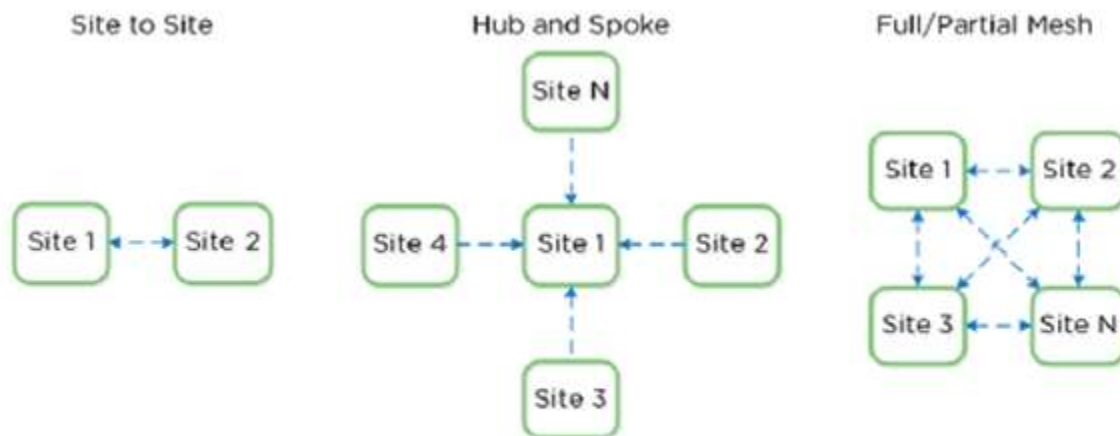


Figure 3.4.18. Example Replication Topologies

Image credit: <https://nutanixbible.com>

## Replication Lifecycle

- Leverages Cerebro
- Manages task delegation to local Cerebro slaves/coordinates with Cerebro master
- Master determines data to be replicated
- Tasks delegated to slaves
- Slaves tell Stargate what to replicate/where
- Extent reads on source are check summed to ensure consistency
- New extents check summed at target

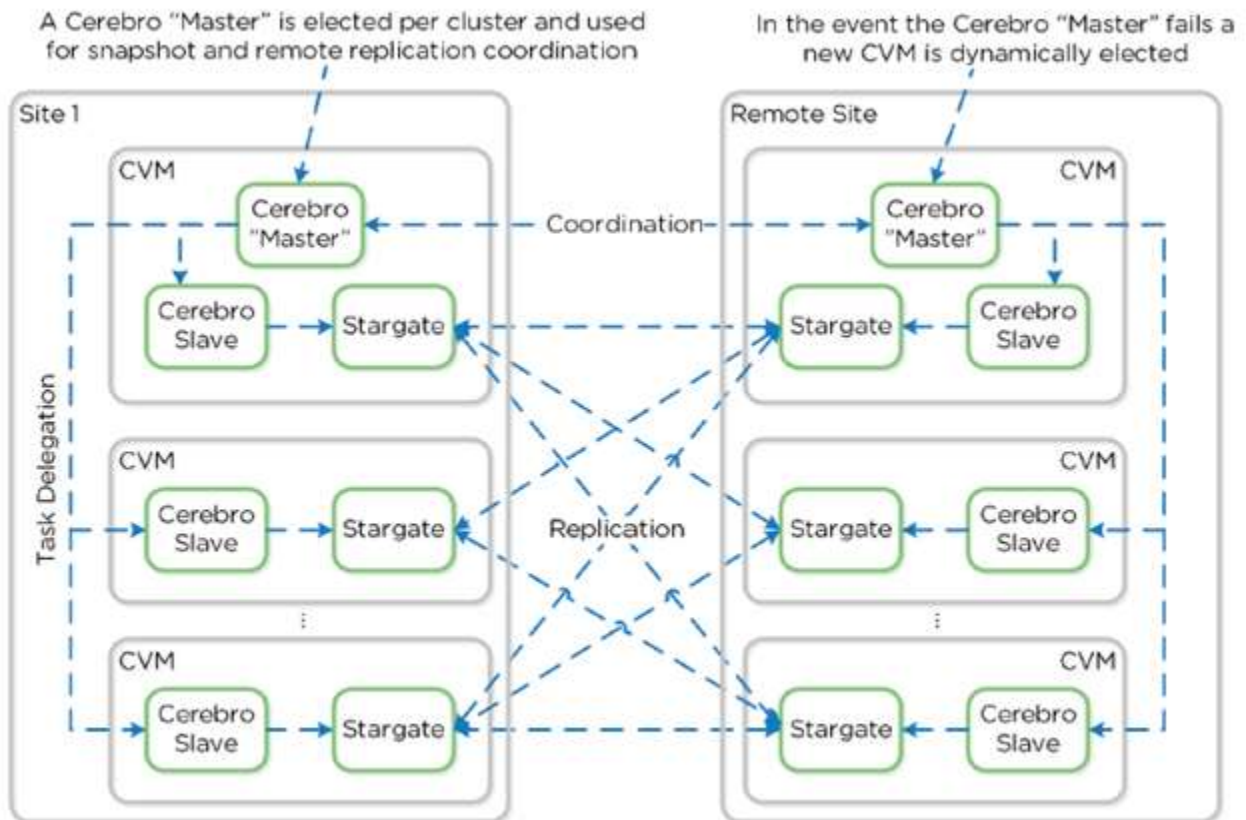


Figure 3.4.19. Replication Architecture

Image credit: <https://nutanixbible.com>

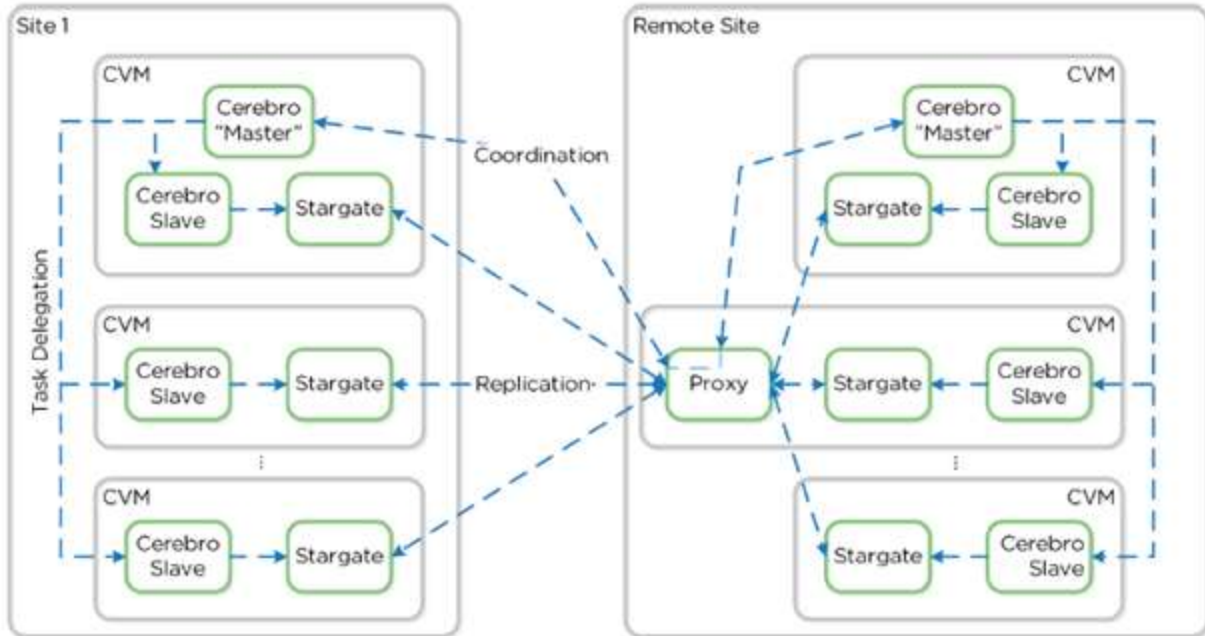


Figure 3.4.20. Replication Architecture - Proxy

Image credit: <https://nutanixbible.com>

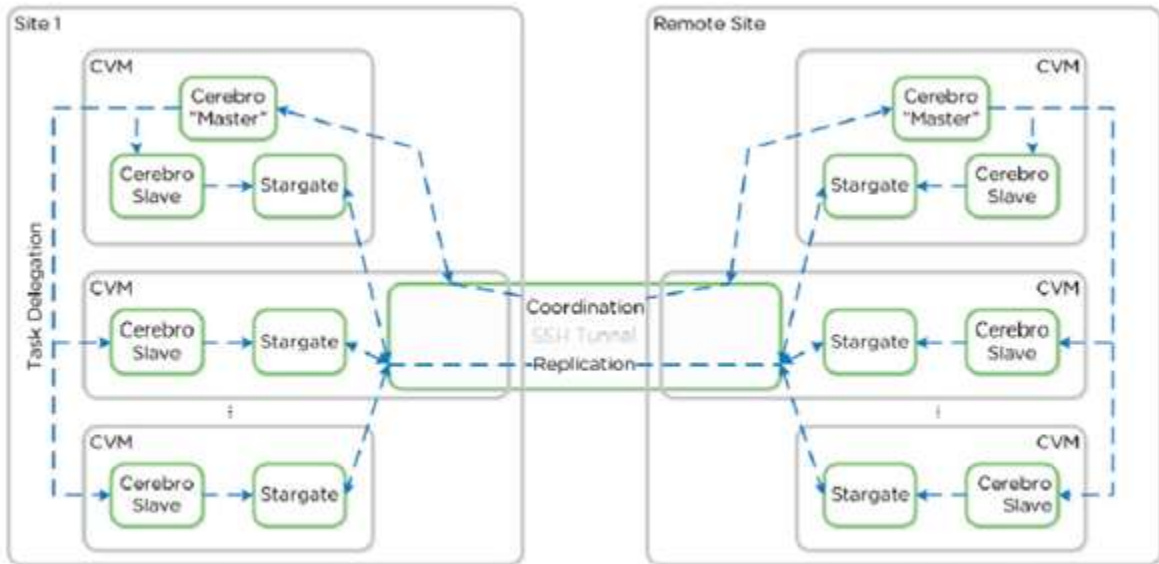


Figure 3.4.21. Replication Architecture - SSH Tunnel

Image credit: <https://nutanixbible.com>

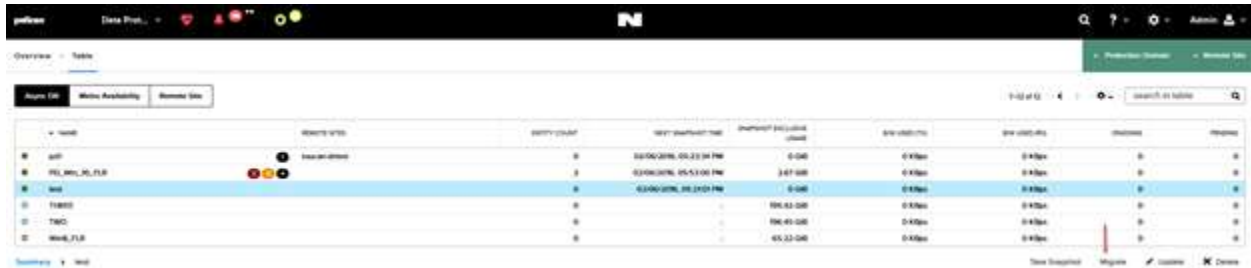
## Explain failover and failback processes

After a protection domain is replicated to at least one remote site, you can carry out a planned migration of the contained entities by failing over the protection domain. You can also trigger failover in the event of a site disaster.

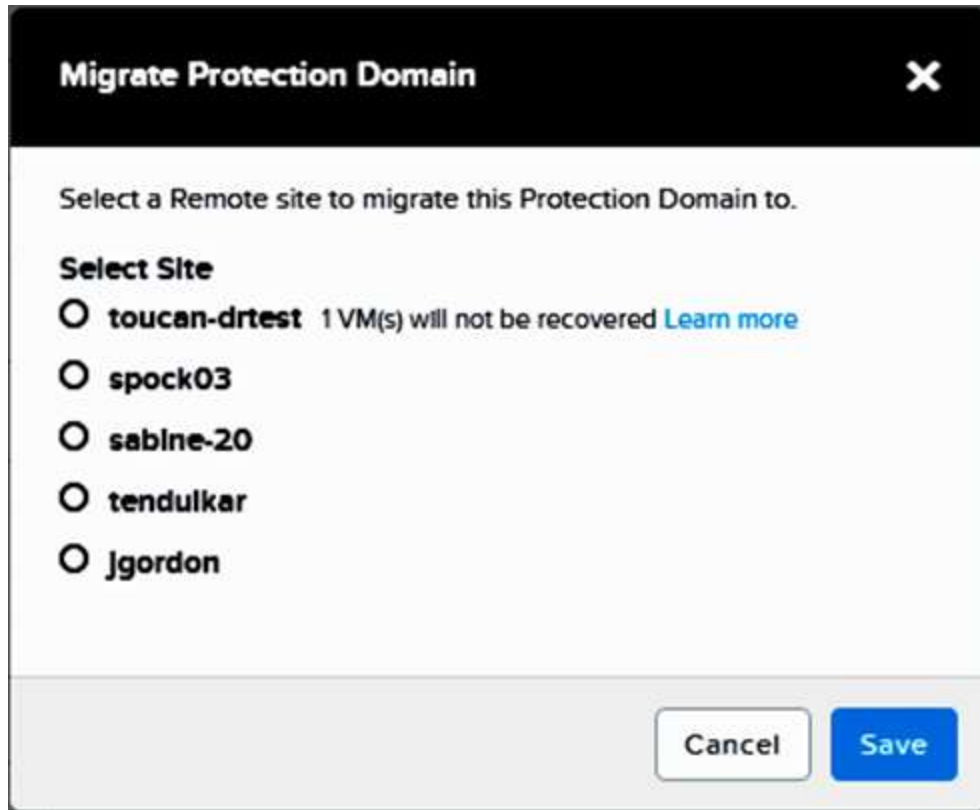
Failover and failback events re-create the VMs and volume groups at the other site, but the volume groups are detached from the iSCSI initiators to which they were attached before the event. After the failover or failback event, you must manually reattach the volume groups to iSCSI initiators and rediscover the iSCSI targets from the VMs.

## Migration (Planned) Failover

System maintenance or expansion might dictate moving a protection domain to another site as a planned event.



NAME	PROTECTION SITE	PROTECTION GROUP	PROTECTION POLICY	PROTECTION POLICY CLASS	PROTECTION POLICY CLASS	PROTECTION POLICY CLASS	PROTECTION POLICY CLASS	PROTECTION POLICY CLASS	PROTECTION POLICY CLASS
all	Insanulman	0	02/00/0196_00.23 (4 PM)	0-000	0-K000	0-K000	0-K000	0-K000	0-K000
RLI_000_00_01.0		0	02/00/0196_00.23 (4 PM)	0-000	0-K000	0-K000	0-K000	0-K000	0-K000
Web		0	02/00/0196_00.23 (4 PM)	0-000	0-K000	0-K000	0-K000	0-K000	0-K000
T1000		0	00.02-000	0-K000	0-K000	0-K000	0-K000	0-K000	0-K000
T000		0	00.02-000	0-K000	0-K000	0-K000	0-K000	0-K000	0-K000
Web_01.0		0	00.22-000	0-K000	0-K000	0-K000	0-K000	0-K000	0-K000



Migrating a protection domain does the following:

- Creates and replicates a snapshot of the protection domain.
- Powers off the VMs on the local site.
- Note: The data protection service waits for 5 minutes for the VM to shutdown. If the VM does not get shutdown within 5 minutes, it is automatically powered off.
- Creates and replicates another snapshot of the protection domain.
- Unregisters all VMs and volume groups and removes their associated files.
- Marks the local site protection domain as inactive.
- Restores all VM and volume group files from the last snapshot and registers them with new UUIDs at the remote site.
- Marks the remote site protection domain as active.

The VMs on the remote site are not powered on automatically. This allows you to resolve any potential network configuration issues, such as IP address conflicts, before powering on the VMs. Additionally, you must **manually reattach the volume groups that were affected by the migration or restore operation, perform in-guest iSCSI attachment**. If you use the nCLI for attaching volume groups, note that the UUIDs of volume groups change

when they are restored at the remote site, and so do their iSCSI target names, which contain the volume group UUID. See Volume Group Configuration.

## Disaster (Unplanned) Failover

When a site disaster occurs, do the following to fail over a protection domain to a remote site:

1. Log into the web console for the target remote site (see Logging Into the Web Console).
2. Go to the Async DR tab of the Data Protection table view (see Data Protection Table View).
3. Select the target protection domain and click the Activate button. A window prompt appears; click the Yes button.

This operation does the following:

- Restores all VM and volume group files from the last fully-replicated snapshot.
- The process first detaches the volume groups that are included in the protection domain or attached to the VMs in the protection domain.
- Registers the VMs and volume groups on the recovery site.
- Marks the failover site protection domain as active.

The VMs are not powered on automatically. This allows you to resolve any potential network configuration issues, such as IP address conflicts, before powering on the VMs. Additionally, you must manually reattach the volume groups that were affected by the migration or restore operation, perform in-guest discovery of the volume groups as iSCSI targets, and log in to the targets. If you use the nCLI for attaching volume groups, note that the UUIDs of volume groups change when they are restored at the remote site, and so do their iSCSI target names, which contain the volume group UUID. See Volume Group Configuration.

The screenshot shows the NetScout VMProtect interface. At the top, there's a navigation bar with 'Data Protection' and 'Remote Site' tabs. Below that, a table lists protection domains. The table has columns for NAME, REMOTE SITE, VM COUNT, NEXT SNAPSHOT TIME, SPACE USED, B/W USED (TB), B/W USED (KBps), ONGOING, and PENDING. The 'Bogota' domain is highlighted. Below the table, there's a 'Summary' section for the 'Bogota' domain, which includes tabs for Replications, VMs, Schedules, Local Snap..., Remote Sna..., Metrics, Alerts, and Events. A red arrow points to the 'Activate' button in the summary section.

NAME	REMOTE SITE	VM COUNT	NEXT SNAPSHOT TIME	SPACE USED	B/W USED (TB)	B/W USED (KBps)	ONGOING	PENDING
Bogota		0	-	0 GiB	0 KBps	0 KBps	0	0
DSVM1_1402418664605	SiteA	0	-	0 GiB	0 KBps	0 KBps	0	0
DSVM2_1402418661653	SiteA	0	-	0 GiB	0 KBps	0 KBps	0	0
DSVM3_1402418658488	SiteA	0	-	0 GiB	0 KBps	0 KBps	0	0
kko_pc		0	-	0 GiB	0 KBps	0 KBps	0	0
MM_1433026417464		0	-	374 KB MB	0 KBps	0 KBps	0	0

## Failing Back a Protection Domain

Perform the following steps to failback a protection domain from remote site to primary site.

1. Login to the Web console of the remote site. The site where the protection domain is currently active.
2. From the Async DR tab under Data Protection, select the protection domain that you want to failback.
3. Click Migrate. The Migrate Protection Domain dialog box appears. Select the site where you want to migrate the protection domain. The VMs that are part of the protection domain, but cannot be recovered on the remote site are also displayed. Click Learn more for additional information.
4. When the field entries are correct, click the Save button.

What to do next:

Manually reattach the volume groups that were affected by the migration or restore operation. Note that the UUIDs of volume groups change when they are restored at the remote site, and so do their iSCSI target names, which contain the volume group UUID.

## Failing Back a Protection Domain (Unplanned)

If an unplanned failure occurs on the primary site, all the entities are failed over to the remote site. After the primary site gets restored, you can failback the entities to the primary site.

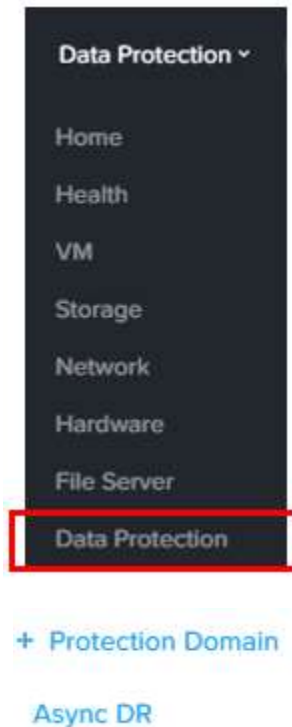
1. Log into the vCenter Server of the primary site.
2. All the hosts are down.
3. Power on all the hosts of the primary site.
  - Controller VMs get automatically restarted and the cluster configuration is established again.
  - All the protection domains that were active before the disaster occurred gets recreated in an active state. However, you cannot replicate the protection domains to the remote site since the protection domains are still active at the remote site.
  - The user VMs get powered on.
4. Log on to one of the Controller VMs at the primary site and deactivate and destroy the VMs by using the following hidden nCLI command.
  - `ncli> pd deactivate-and-destroy-vm name=protection_domain_name`
  - Replace `protection_domain_name` with the name of the protection domain that you want to deactivate and destroy.
  - Caution: Do not use this command for any other workflow. Otherwise, it will delete the VMs and data loss will occur.
  - VMs get unregistered at the primary site and the protection domain is no longer active on the primary site. Remove the orphaned VMs from the inventory of the primary site.
5. (Optional) If you want to schedule frequent replications, log into the remote site and schedule replication to the primary site.
6. Log into the remote site and click Migrate to migrate the protection domain to the primary site. The VMs get unregistered from the secondary site (not removed) and data gets replicated to the primary site, and then the VMs gets registered on the primary site. Protection domain starts replicating from the primary site to the remote site based on the schedule that was originally configured on the primary site.
7. Power on the VMs in the correct sequence and configure the IP address again (if required). Additionally, manually reattach any volume groups that were either included in the protection domain or attached to the VMs in the protection domain. You can start using the VMs from the primary site.
8. Remove the orphaned VMs from the inventory of the secondary site.

## Create and modify a Protection Domain



You configure a protection domain by using Async DR feature (up to 1-hour RPO) or NearSync DR (up to 1-minute RPO ) by defining a group of entities (VMs and volume groups) that are backed up locally on a cluster and optionally replicated to one or more remote sites.

See the DR best practices guide from <https://portal.nutanix.com/#/page/solutions> for guidance to set up DR in your environment.



Name - Virtual Machines - Schedule

A protection domain is a grouping of Virtual Machines for disaster recovery purposes. Enter a name (using alpha numeric characters only) for the protection domain you would like to create. You will then be guided into assigning Virtual Machines to it, and scheduling it.

NAME

Cancel


Create



Protection Domain ✕

Name: Virtual Machines - Schedule

You currently do not have any schedules.

 New Schedule

TYPE	REPEAT ON	START DATE	END DATE	RETENTION POLICY

Previous

Close

Virtual Machines - **Schedule** ? ✕

**Configure your local schedule**

Repeat every  minute(s) ?

Repeat every  hour(s) ?

Repeat every  day(s) ?

Repeat weekly

S  M  T  W  T  F  S

Repeat monthly

Day of month:  ?

Start on  📅 at  🕒

End on  📅 at  🕒

Create application consistent snapshot

**Retention policy**

Local keep the last  snapshots

**Remote Sites**

test\_soak\_1 keep the last  snapshots

upgrade\_45 keep the last  snapshots





aws\_ctr keep the last  snapshots

Cancel Create Schedule

Virtual Machines - Schedule

You currently have 2 schedules . Next snapshot is scheduled on 09/09/15, 02:46:00pm

New Schedule

TYPE	REPEAT ON	START DATE	END DATE	APP CONSISTENT SNAPSHOT	RETENTION POLICY	
Hourly	Every 1 hour	09/08/15, 03:46:00pm	-	No	Local: 2, colossus08: 3	 
Hourly	Every 1 hour	09/09/15, 02:33:00pm	-	Yes	Local: 1, aster: 2	 

Update

Delete

Previous

Close

## Replicate Protection Domain



Select one or more targets to replicate to. This is a one time replication that can start now or at a later time.

LOCAL

### REMOTE SITES

- test\_soak\_1
- upgrade\_45
- aws\_ctr
- AWSAbhiDontDel
- AzureTest
- azure\_monday
- AzureSimpleTest
- AzureTestCheck

### REPLICATION START TIME

Now



### RETENTION TIME

No Expiration



Create application consistent snapshot

Cancel

Save

# Configure a Remote Site

A remote site is the target location to store data replications for protected domains. The remote site can be either another physical cluster or a cluster located in a public cloud. To configure a remote physical cluster that can be used as a replication target, do the following:

+ Remote Site

Physical Cluster

Cloud

Remote Site ×

1. Details    2. Settings

REMOTE SITE NAME

Enable Proxy

CAPABILITIES

Backup     Disaster Recovery

ADDRESSES

    Port: 2020

1. Details 2. Settings

**Bandwidth Throttling**BANDWIDTH THROTTLING 

DEFAULT BANDWIDTH LIMIT

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

BANDWIDTH THROTTLING POLICIES

No policy found

[+ Add Policy](#)**Compression**COMPRESS ON WIRE 

Network and vStore Mappings are used to map source and destination networks, and storage containers for the Source Site to the Remote Site. The network connections, VLANs, and Storage Containers must be configured on both source and destination cluster.

**Mappings****Network Mapping**

Network mapping is only supported between two AHV clusters pair or between an AHV cluster and an ESX cluster pair.

**vStore Name Mapping**

**Remote site is currently not reachable. Please try again later.**



**Mapping**

**NETWORK MAPPING**

Source Cluster                      Destination Cluster

Source Cluster                      Destination Cluster                      +

*Enter unique source and destination name mappings.*

**VSTORE NAME MAPPING**

Source VStore                      Destination VStore

Select a VStore                                           +

*Enter unique source and destination name mappings.*

Cancel                      Save

Cloud Site

? x

Cloud Type    Credentials    Remote Site Settings

Warning! The current license level does not allow for the use of this feature.

CLOUD TYPE

Choose the type of cloud that will act as the remote site.

- AWS
- Azure

Cancel                      Next

# Section 12 – Prism Central

## Identify Prism Central requirements

A Prism Central instance can consist of either a single VM or a set of three VMs. A 3-VM instance increases both the capacity and resiliency of Prism Central at the cost of maintaining the additional VMs. In addition, each VM can be either “large” or “small” in size. Thus, you may choose from four configurations.

VM Size	1-VM Instance	3-VM Instance	Manage
Large	8 vCPUs 32 GB of memory 2500 GiB of storage	24 vCPUs 96 GB of memory 7500 GiB of storage (combined)	2500 VM's (1-VM Instance) 5000 VM's (3-VM Instance)
Small	4 vCPUs 16 GB of memory 500 GiB of storage	12 vCPUs 48 GB of memory 1500 GiB of storage (combined)	12,500 VM's (1-VM Instance) 25,000 VM's (3-VM Instance)

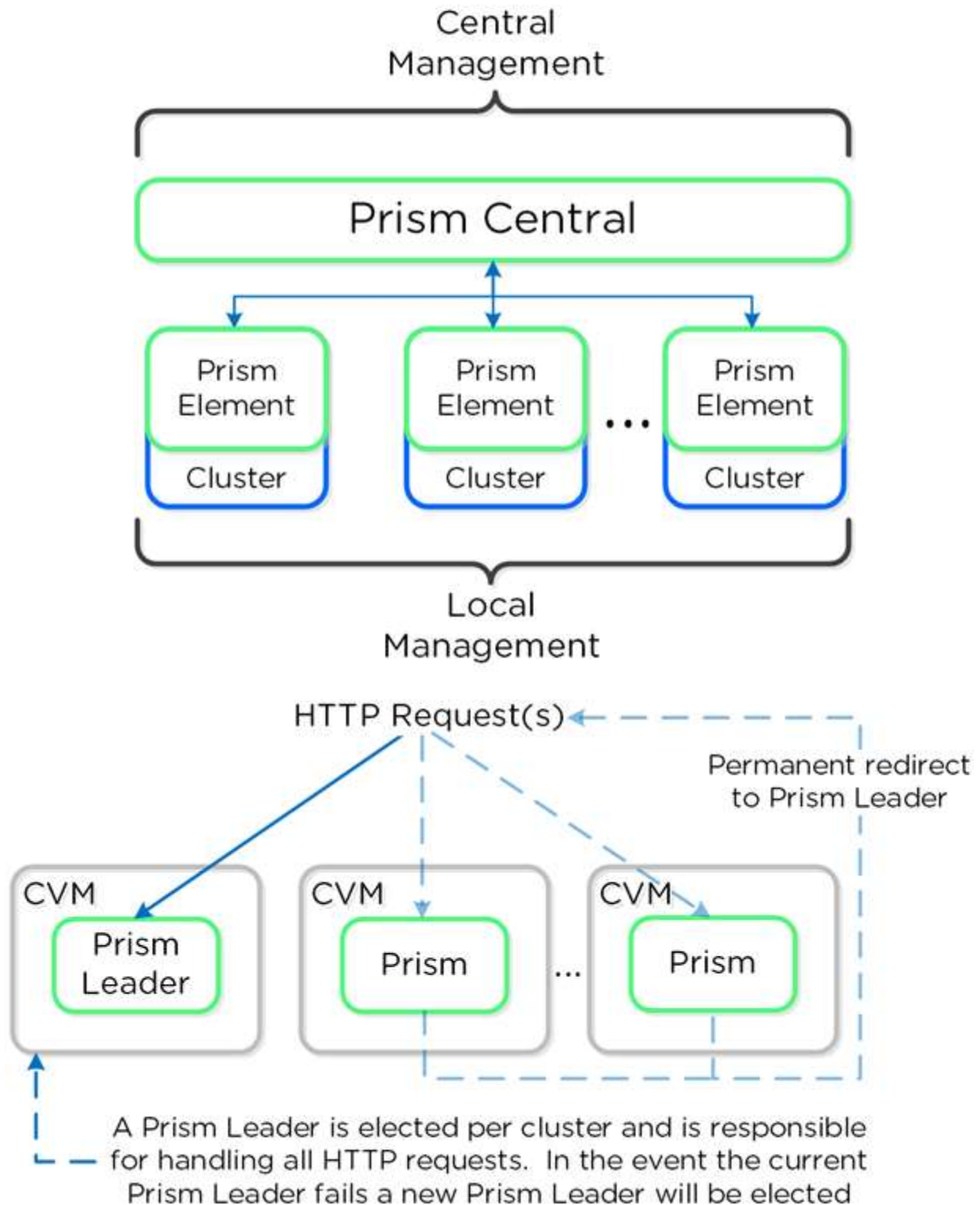
## Describe and differentiate Prism Element and Prism Central

### Prism Central (PC)

- Multi-cluster manager responsible for managing multiple Acropolis Clusters to provide a single, centralized management interface. Prism Central is an optional software appliance (VM) which can be deployed in addition to the Acropolis Cluster (can run on it).
- 1-to-many cluster manager
- Recommended for larger sites (more than one cluster/multiple sites)

### Prism Element (PE)

- Localized cluster manager responsible for local cluster management and operations. Every Acropolis Cluster has Prism Element built-in.
- 1-to-1 cluster manager



- Prism services run on every CVM
- 1 leader responsible for all HTTP requests

- New leader elected if failure
- CVM that is not the leader redirects to leader via 301
- 80 and 9440 (80 redirects to HTTPS on 9440)
- Recommended cluster external IP is used
- If failure, gARP used

## Deploy a Prism Central VM

### Installing Prism Central (1-Click Method)

You can install a Prism Central VM using the “1-click” method. This method employs the Prism web console from a cluster of your choice and creates the Prism Central VM in that cluster.

The “1-click” method is the easiest method to install Prism Central in most cases. However, you cannot use this method when:

- The target cluster runs Hyper-V or XenServer (or mixed hypervisors)
- You do not want to install the Prism Central VM in a Nutanix cluster
- You do not have access to a Nutanix cluster

### Requirements

- The specified gateway must be reachable.
- No duplicate IP addresses can be used.
- The container used for deployment is mounted on the hypervisor hosts.
- When installing on an ESXi cluster:
  - vCenter and the ESXi cluster must be configured properly. See the vSphere Administration Guide for Acropolis (using vSphere Web Client) for details.
  - vCenter must be registered in Prism.
  - DRS must be enabled in vCenter.
  - vCenter is up and reachable during the deployment.

### Installation Process

Select between an existing connection or deploying a new Prism Central

I want to deploy a new Prism Central instance  
I don't have Prism Central or want to deploy a new one

Deploy

I already have a Prism Central instance deployed  
Nutanix recommends connecting this cluster to it

Connect

Close

**Prism Central** ? X

Installation Image

Select an image to install, download the latest version from the Internet or upload one from your computer.

Available versions

5.5	X	<b>Install</b>
5.1.4		Download

Upload installation binary

You can [upload the Prism Central binary](#) instead of downloading from the Internet.

Back Cancel

**Prism Central** ? X

<p><b>Scale-Out Prism Central Cluster</b></p> <p>Capacity: Supports 5k to 25k VMs</p> <p>Added resiliency: RF2</p> <p>Minimum memory required: 48 GB</p> <p><b>Deploy 3-VM PC</b></p>	<p><b>Single-VM Prism Central</b></p> <p>Capacity: Supports 2.5k to 12.5k VMs</p> <p>Default resiliency</p> <p>Minimum memory required: 16 GB</p> <p><b>Deploy 1-VM PC</b></p>
---	--

Back Cancel

Prism Central
? X

General Configuration

VM NAME

SELECT A CONTAINER

SelfServiceContainer

---

VM Sizing

Select the size of your Prism Central VM.

SIZE	VCPUS	MEMORY (GB)	STORAGE (GiB)
<input checked="" type="radio"/> SMALL - (UP TO 2,500 VMs)	4	16	500
<input type="radio"/> LARGE - (UP TO 12,500 VMs)	8	32	2500

---

Network Config

AHV NETWORK + Create Network

Back
Close
Deploy

MetalGear-BLR
admin

Filter

6 Total Tasks

OPERATION REQUEST	ENTITY	PERCENT	STATUS	CREATE TIME	DURATION
Download and deploy prism central	Cluster   Details	<div style="width: 100%; height: 10px; background-color: #92D050;"></div> 100%	Succeeded	02/05/18, 4:36:05 PM	11 minutes
Network create	Virtual network	<div style="width: 100%; height: 10px; background-color: #92D050;"></div> 100%	Succeeded	02/05/18, 4:20:13 PM	few seconds
Cluster upgrade task	Cluster   Details	<div style="width: 100%; height: 10px; background-color: #92D050;"></div> 100%	Succeeded	02/05/18, 5:50:28 AM	25 minutes

MetalGear-BLR
admin

Overview Tasks

NAME	IP	IP ADDRESS	CPUS	MEMORY (GB)	DISK	STATUS	START TIME	END TIME	OPERATION REQUEST	OPERATION STATUS	OPERATION TYPE	OPERATION DURATION	OPERATION START	OPERATION END
PCAgentCore2	192.168.1.10	192.168.1.10	4	16	500	Running	02/05/18, 4:36:05 PM	02/05/18, 4:36:05 PM	Download and deploy prism central	Succeeded	VM	11 minutes	02/05/18, 4:36:05 PM	02/05/18, 4:36:05 PM
PCAgentCore1	192.168.1.11	192.168.1.11	4	16	500	Running	02/05/18, 4:36:05 PM	02/05/18, 4:36:05 PM	Download and deploy prism central	Succeeded	VM	11 minutes	02/05/18, 4:36:05 PM	02/05/18, 4:36:05 PM
PCAgentCore3	192.168.1.12	192.168.1.12	4	16	500	Running	02/05/18, 4:36:05 PM	02/05/18, 4:36:05 PM	Download and deploy prism central	Succeeded	VM	11 minutes	02/05/18, 4:36:05 PM	02/05/18, 4:36:05 PM

# Installing Prism Central (Manually)

When the 1-Click Method is not an option, you can install a Prism Central VM manually.

## Prerequisites

Download the Prism Central file(s) for the desired hypervisor.

The screenshot shows the Nutanix website's Prism Central download page. The navigation menu includes Documentation, Support, Downloads, My Products, and Nutanix Next. A dropdown menu for Prism Central is open, with a red arrow pointing to the 'Prism Central' option. The main content area contains text about Prism Central and a link to download the latest GA release (v4.1.2). To the right, there are three download sections: HYPER-V DOWNLOAD, ESX DOWNLOAD, and AHV DOWNLOAD. Each section includes a date (Feb 22, 2016), a URL copy link, a download button, size, and MD5 hash. There are also labels for 'boot image', 'home image', and 'data image' pointing to their respective download links.

Hypervisor	Download Type	Date	URL Copy	Download Button	Size	MD5
HYPER-V	DOWNLOAD	Feb 22, 2016	Hyper-V Download	Download 4.6 Zip	4.4GB	4487422d43d6f29d4aabcc92c384f3a4
ESX	DOWNLOAD	Feb 22, 2016	ESX Download	Download 4.6 Ova	4.6GB	8da4e05d220cce5353651d46b267475b
AHV	DOWNLOAD	Feb 22, 2016	Boot Image (Recommended)	Download 4.6 Image (AHV)	3.9MB	90505a8e88f798526e2b670437f1e3c
			Home Image (Recommended)	Download 4.6 Image (AHV)	4.6GB	ccf3981d0a14e168cb4e0321f06745b
			Disk Image (Recommended)	Download 4.6 Image (AHV)	19.1MB	5c0175e8e6e56e6852f5aa96257752bf

## VMware

1. Install the OVA file as follows:



- a. Connect to vCenter or an ESXi host using the vSphere client.
- b. Select the OVA file, deploy it, and then start the Prism Central VM.
  - **Note:** Configure the VM to have at least 16 GB of memory, 4 vCPUs, and a NIC. This is the minimum recommended configuration for a Prism Central VM.
2. Log into the Prism Central VM through the vSphere console (user name “nutanix” and password “nutanix/4u”).
3. Assign a static IP address to the Prism Central VM as follows:
  - a. Open the ifcfg-eth0 file for editing.
    - The following command opens the file using the vi editor:

```
$ sudo vi /etc/sysconfig/network-scripts/ifcfg-eth0
```

- Update the NETMASK, IPADDR,BOOTPROTO, and GATEWAY entries as needed.
- **Warning:** Carefully check the file to ensure there are no syntax errors, whitespace at the end of lines, or blank lines in the file.
- Save the changes.
- Remove any existing Nutanix Controller VM entries, that is ones which include “NTNX-<number>-CVM”, from the/etc/hosts file. (Be careful that you do not remove any other entries from the file.)
  - a. To edit the file using vi, enter

```
$ sudo vi /etc/hosts
```

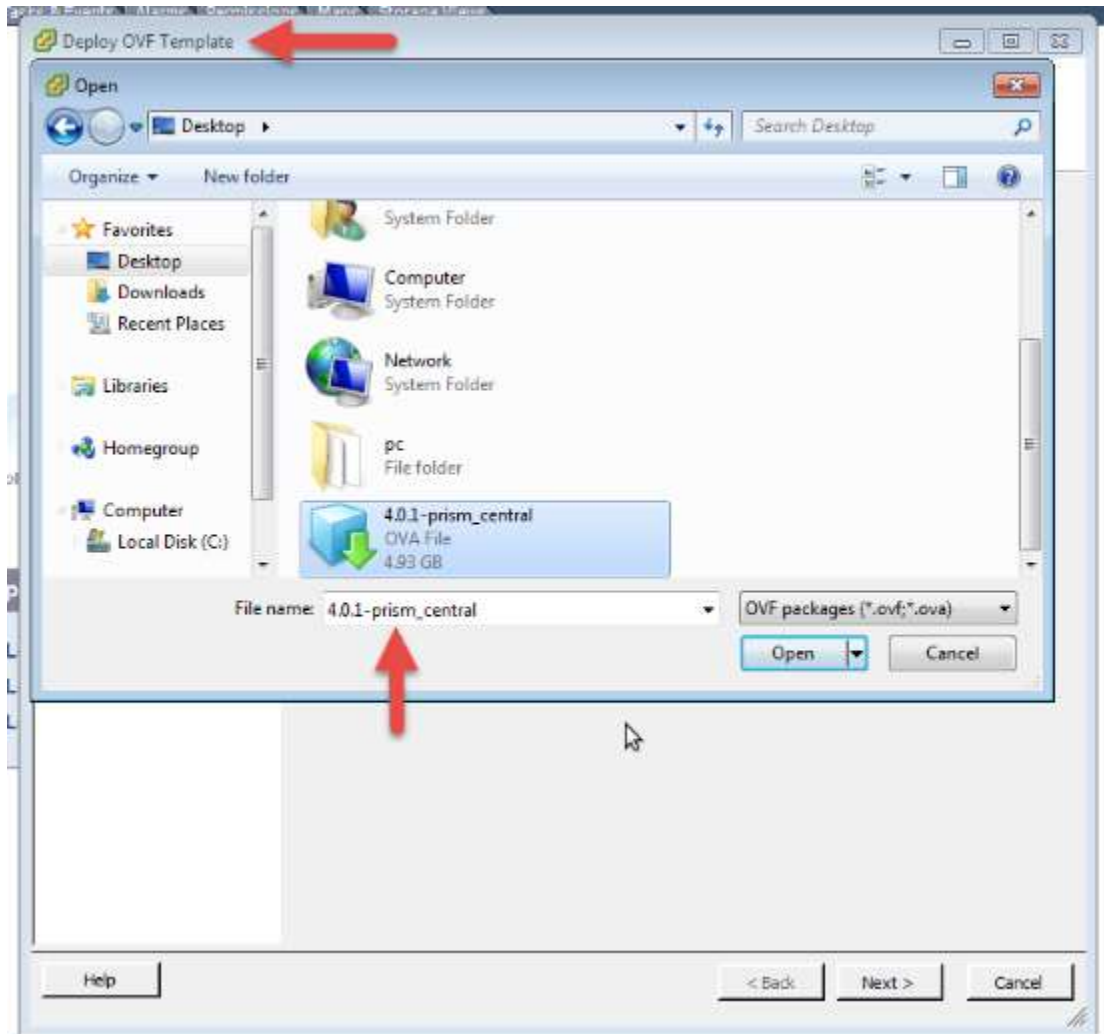
- b. When the Prism Central VM starts the first time and DHCP is enabled in the network, an entry similar to the following is added to the /etc/hosts file. This entry (if present) needs to be removed from the/etc/hosts file before restarting the Prism Central VM in the next step, which generates a new entry if DHCP is enabled.
- c. Restart the Prism Central VM.

```
$ sudo reboot
```

- d. Log into the Prism Central VM again through the vSphere console and then enter the following command to create a cluster:

```
$ cluster -cluster_function_list="multicluster" -s static_ip_address create
```

- e. The *static\_ip\_address* is the Prism Central VM IP address you assigned to the `IPADDR` parameter in step 4.
- b. This completes Prism Central installation. The next step is to register clusters with Prism Central.



## AHV

In this procedure, the Prism Central VM deployment consists of three virtual disks, where rel.# is the release version, deployed in order:

- Boot disk – rel.#-prism\_central-boot.qcow2
  - Home disk – rel.#-prism\_central-home.qcow2
  - Data disk – rel.#-prism\_central-data.qcow2
6. Under the AHV Download heading, click the **Download <release#> Tar (AHV)** button for each image to download the

boot, home, and data image files to your workstation.

The <release#> represents the AOS version number. Repeat this step for the links for the home and data disk images.

7. Log in to the Prism web console of the target cluster and select **Image Configuration** from the task icon pull-down list of the main menu. The Image Configuration window appears. Click the **Upload Image** button
8. Enter a name, for example pc\_boot, and optional description (annotation) for the boot image.
  1. Select **Disk** as the image type.
  2. Select a storage container to use from the pull-down list.
  3. Select **Upload a file** and select the boot image downloaded to your workstation in step 2.
  4. Click the **Save** button.
  5. Repeat this step for the home (pc\_home) and data disk (pc\_data) image files.
9. Create a new VM from the images as follows:
  0. In the VM dashboard, click the **Create VM** button.
  1. In the Create VM window, enter appropriate information in the name, compute details, and memory fields, and (if needed) click the **Add New NIC** button to create a network interface for the VM.
  2. Click the **Add New Disk** button and attach a boot disk.
    0. **Type**: Select **Disk**.
    1. **Operation**: Select **Clone from Image Service**.
    2. **Bus Type**: Select **SCSI**.
    3. **Storage Container**: Select a storage container from the list of available for the Prism
    4. **Image**: Select pc\_boot.img (or whatever you named the boot image, originally downloaded as rel.#-prism\_central-boot.qcow2) from the list of images.
    5. **Size (GiB)**: Enter the disk size (in GiB). This value is populated automatically from the selected image; do not change the value that appears in this field unless directed to do so by Nutanix customer support.
  3. Click the **Add New Disk** button again and attach the home disk.
    - The steps are the same as for the boot disk except the image file name, which is pc\_home.img (or whatever you named the home image).

4. Click the **Add New Disk** button again and attach the data disk.
  - The steps are the same as for the boot disk except the image file name, which is `pc_data.img` (or whatever you named the data image).
10. When all the settings are correct, click the **Save** button to create the VM.
11. Go to the VM dashboard table view, select the new VM (in the table), and then click the **Power on** action link (below the table) to start up the VM.
12. Log into the Prism Central VM through the vSphere console (user name “nutanix” and password “nutanix/4u”).
13. Assign a static IP address to the Prism Central VM as follows:
  0. Open the `ifcfg-eth0` file for editing.
    - The following command opens the file using the vi editor:

```
$ sudo vi /etc/sysconfig/network-scripts/ifcfg-eth0
```
    - Update the `NETMASK`, `IPADDR`, `BOOTPROTO`, and `GATEWAY` entries as needed.
    - **Warning:** Carefully check the file to ensure there are no syntax errors, whitespace at the end of lines, or blank lines in the file.
    - Save the changes.
    - Remove any existing Nutanix Controller VM entries, that is ones which include “`NTNX-<number>-CVM`”, from the `/etc/hosts` file. (Be careful that you do not remove any other entries from the file.)
    - To edit the file using vi, enter

```
$ sudo vi /etc/hosts
```
    - When the Prism Central VM starts the first time and DHCP is enabled in the network, an entry similar to the following is added to the `/etc/hosts` file. This entry (if present) needs to be removed from the `/etc/hosts` file before restarting the Prism Central VM in the next step, which generates a new entry if DHCP is enabled.
14. Restart the Prism Central VM.

```
$ sudo reboot
```

15. Log in to the Prism Central VM again and then enter the following command to create a cluster:

```
cluster -cluster_function_list="multicluster" -s  
static_ip_address create
```

16. The *static\_ip\_address* is the Prism Central VM IP address you assigned to the `IPADDR` parameter in step 4.
17. This completes Prism Central installation. The next step is to register clusters with Prism Central.

### Create Image

NAME

ANNOTATION

IMAGE TYPE

Not Selected

STORAGE CONTAINER

SelfServiceContainer

IMAGE SOURCE

From URL

Upload a file  No file chosen

**Add Disk** ? X

TYPE  
DISK

OPERATION  
CLONE FROM IMAGE SERVICE

BUS TYPE  
SCSI

IMAGE  
pc\_boot.img

SIZE (GiB)  
2.73

Cancel Add

## Register/Unregister a Nutanix Cluster with Prism Central

### Prerequisites

- If you have never logged into Prism Central as the user admin, you need to log in and change the password before attempting to register a cluster with Prism Central.
- Do not enable client authentication in combination with ECDSA certificates on a registered cluster since it causes interference when communicating with Prism Central.
- Port 9440 and 80 need to be open in both directions between the Prism Central VM and any registered clusters.
- A cluster can register with just one Prism Central instance at a time. To register with a different Prism Central instance, first unregister the cluster.

## Register a Nutanix Cluster with Prism Central

# Settings

Setup

Connect to Citrix Cloud

**Prism Central Registration**

Pulse

Rack Configuration

Prism Central

?

Select between an existing connection or deploying a new Prism Central

**I want to deploy a new Prism Central instance**

I don't have Prism Central or want to deploy a new one

Deploy

**I already have a Prism Central instance deployed**

Nutanix recommends connecting this cluster to it

Connect

Several features will be available in Prism Central post registration as part of a larger effort to centralize management. ✕

Some features will be READ-ONLY mode in Prism Element for serviceability and debugging but fully accessible in Prism Central. If an image was created on another Prism Central, the image can only be managed through the original Prism Central.

FEATURE/SERVICE	PRISM ELEMENT	PRISM CENTRAL
Cluster Unregistration	available by scripts only	available by scripts only
Affinity Policies	READ-ONLY	✓
Self-Service Portal	available in PC only	✓

[Back](#)[Cancel](#)[Next](#)



---

**Connect to an existing Prism Central instance**

Please fill in the information needed to establish a connection

Prism Central IP

Port

Optional

Username

Password

---

## Unregister a Nutanix Cluster with Prism Central

**Unregistering** a cluster through the Prism GUI is no longer available. This option was removed to reduce the risk of accidentally unregistering a cluster because several features (including role-based access control, application management, micro-segmentation policies, and self-service capability) require Prism Central to run your clusters. If a cluster is unregistered from Prism Central, not only will these features not be available but the configuration for them may also be erased.

Therefore, only the following procedure is available to unregister a cluster. See [KB 4944](#) for additional details if you have enabled Self Service, Calm, or other special features in Prism Central.

To unregister a cluster from an instance of Prism Central that you have deleted or destroyed, do the following.

1. Log on to any Controller VM of the registered cluster through an SSH session.
2. Run the cluster status command and verify that all services are in a healthy state

After performing these steps you can re-register the cluster with a new or re-created Prism Central instance.

If the clean up process does not complete successfully, try the following:

- Check the logs to indicate if there are any input errors when calling the script. The logs for the unregistration cleanup script can be found under
  - `~/data/logs/unregistration_cleanup.log`.
- If errors occur during script execution, run the cluster status command and check that the cluster services are up and running. Rerun the script and check if it succeeds.


```

acropolis> exit
muban@MTR0-...:~$ ssh -A -C -M 10.5.25.30 -f
muban@MTR0-...:~$ ssh -A -C -M 10.5.25.30 -f cluster status
2017-11-19 21:31:29 INFO zookeeper_session.py:318 cluster is attempting to connect to Zookeeper
2017-11-19 21:31:29 INFO cluster:2428 Creating action status on SWs 10.5.25.30
The state of the cluster: start
Lockdown mode: Disabled

CVM: 10.5.25.30 Ip, ZeusLeader
      Zeus IP [1966, 1995, 1996, 1997, 2043, 2061]
      Scavenger IP [2788, 2756, 2737, 2734]
      SSITerminator IP [2184, 2156, 2337, 2158]
      SecureFileSysc IP [2188, 2162, 2163, 2164]
      Medusa IP [2278, 2312, 2313, 2318, 2421]
      DynamicRingChanger IP [2528, 2660, 2661, 2665]
      Pithos IP [2524, 2586, 2587, 2615]
      Mantle IP [2528, 2590, 2591, 2620]
      Haro IP [2533, 2685, 2689, 2610]
      Stargate IP [2823, 2893, 2894, 3250, 3258]
      InsightsDB IP [2826, 2876, 2877, 3827]
      InsightsDataTransfer IP [2834, 2916, 2917, 3800, 3004, 3005, 3006]
      Fregat IP [2847, 2963, 2964, 2965]
      Cerberus IP [2981, 2984, 2985, 3443]
      Chimera IP [2928, 2999, 3000, 3074]
      Curator IP [2942, 3047, 3048, 3500]
      Prism IP [3024, 3094, 3095, 3001, 3787, 3770]
      CIM IP [3037, 3154, 3155, 3214]
      AlertManager IP [3046, 3253, 3256, 3622]
      AntUnios IP [3083, 3215, 3216, 3488]
      Colalog IP [3186, 3189, 3188, 3190]
      Acropolis IP [3173, 3294, 3295, 3296]
      Iliad IP [3197, 3389, 3388, 3396]
      Srap IP [3198, 3371, 3372, 3395]
      SysStatCollector IP [3268, 3445, 3446, 3447]
      Tunnel IP [3231, 3582, 3583]
      Janus IP [3384, 3485, 3486]
      NutanixGuestTools IP [3517, 3617, 3618, 3682]
      MinervaCVM IP [4029, 4064, 4065, 4066, 4213]

```

```
Cluster Id           : 00055875-0cf5-0f0f-0000-0000000097fc::38908
Cluster Uuid        : 00055875-0cf5-0f0f-0000-0000000097fc
Cluster Name       : johny5-1
Cluster Version    : 5.5
External IP Address :
External Data Services... :
Node Count         : 1
Block Count        : 1
Support Verbosity Level : BASIC_COREDUMP
Lock Down Status   : Disabled
Shadow Clones Status : Enabled
Password Remote Login ... : Enabled
Timezone           : America/Los_Angeles
On-Disk Dedup      : Disabled
Has Self Encrypting Disk : no
Common Criteria Mode : Enabled
Degraded Node Monitoring : Enabled
```



## Section 13 – Cluster Maintenance

### Perform one or more Nutanix Cluster Checks

A set of health checks are run regularly that provide a range of clusters health indicators. You can specify which checks to run and configure the schedulable checks and other parameters for each health check.

The cluster health checks cover a range of entities including AOS, hypervisor, and hardware components. A set of checks are enabled by default, but you can run, disable, or reconfigure any of the checks at any time. To reconfigure one or more health checks, do the following:

## Run Checks



Select checks you want to run:

- All Checks (333)
- Only Failed And Warning Checks (2)
- Specific Checks

Send the cluster check report in the email

Recipients: none

Cancel


Run

View Summary



Summary of Cluster Check Executed on 11/05/18, 12:40:58 PM

[Download Output](#)



CHECK STATUS	NUMBER OF CHECKS
Passed	276
Info	3
Warn	1
Error	0
Failed	0
Total	280

Close

## Install NCC

# Upgrading NCC by Uploading Binary and Metadata Files

The screenshot shows the 'Upgrade Software' interface. At the top, there is a navigation breadcrumb: 'AOS · File Server · Hypervisor · NCC · Foundation', with 'NCC' highlighted. Below this, the 'CURRENT VERSION' is listed as '3.6.2.1'. Under 'AVAILABLE COMPATIBLE VERSIONS', there is a checkbox for 'Hide incompatible versions' which is unchecked. A message box states 'No available versions for upgrade.' Below this, there is a section for 'UPLOAD UPGRADE SOFTWARE BINARY' with the text: 'You can [upload the NCC binary](#) instead of downloading from the Internet.'

## Installing NCC from an Installer File

Installs the Nutanix Cluster Check (NCC) health script to test for potential issues and cluster health.

1. Download the installation file to any controller VM in the cluster and copy the installation file to the /home/nutanix directory.
2. Check the MD5 value of the file. It must match the MD5 value published on the support portal. If the value does not match, delete the file and download it again from the support portal.

```
nutanix@cvm$ md5sum ./ncc_installer_filename.sh
```

3. Perform these steps for NCC versions that include a single installer file (ncc\_installer\_filename.sh)
  1. Make the installation file executable.

```
nutanix@cvm$ chmod u+x ./ncc_installer_filename.sh
```

## 2. Install NCC.

```
nutanix@cvm$ ./ncc_installer_filename.sh
```

3. The installation script installs NCC on each node in the cluster.
  4. NCC installation file logic tests the NCC tar file checksum and prevents installation if it detects file corruption.
  5. If it verifies the file, the installation script installs NCC on each node in the cluster.
  6. If it detects file corruption, it prevents installation and deletes any extracted files. In this case, download the file again from the Nutanix support portal.
4. Perform these steps for NCC versions that include an installer file inside a compressed tar file (ncc\_installer\_filename.tar.gz).
- a. Extract the installation package.

```
nutanix@cvm$ tar xvmf ncc_installer_filename.tar.gz --recursive-unlink
```

- b. Replace ncc\_installer\_filename.tar.gz with the name of the compressed installation tar file.
  - c. The --recursive-unlink option is needed to ensure old installs are completely removed.
  - d. Run the install script. Provide the installation tar file name if it has been moved or renamed.

```
nutanix@cvm$ ./ncc/bin/install.sh [-f install_file.tar]
```

- e. The installation script copies the install\_file.tar tar file to each node and install NCC on each node in the cluster.
  5. Check the output of the installation command for any error messages.
    - If installation is successful, a Finished Installation message is displayed. You can check any NCC-related messages in `/home/nutanix/data/logs/ncc-output-latest.log`.
    - In some cases, output similar to the following is displayed. Depending on the NCC version installed, the installation file might log the output to `/home/nutanix/data/logs/` or `/home/nutanix/data/serviceability/ncc`.

```
Copying file to all nodes [DONE]
+-----+
| State | Count |
+-----+
| Total | 1 |
+-----+
Plugin output written to /home/nutanix/data/logs/ncc-output-latest.log
```

```
[ info ] Installing ncc globally.  
[ info ] Installing ncc on 10.130.45.72, 10.130.45.73  
[ info ] Installation of ncc succeeded on nodes 10.130.45.72, 10.130.45.73.
```

## Configure an HTTP Proxy

If the customer site cannot send traffic to a Nutanix service center directly, an HTTP proxy is required.

**HTTP Proxies** ?

---

Configure one or more HTTP Proxy servers that you would like to use. Proxy servers that have been configured are displayed below.

[+ New Proxy](#)

NAME	ADDRESS	PORT
HTTP Proxies have not been configured.		

---

## Create HTTP Proxy



Name

Address

Port

Username

Password

Protocols

HTTP

HTTPS

Cancel

Save

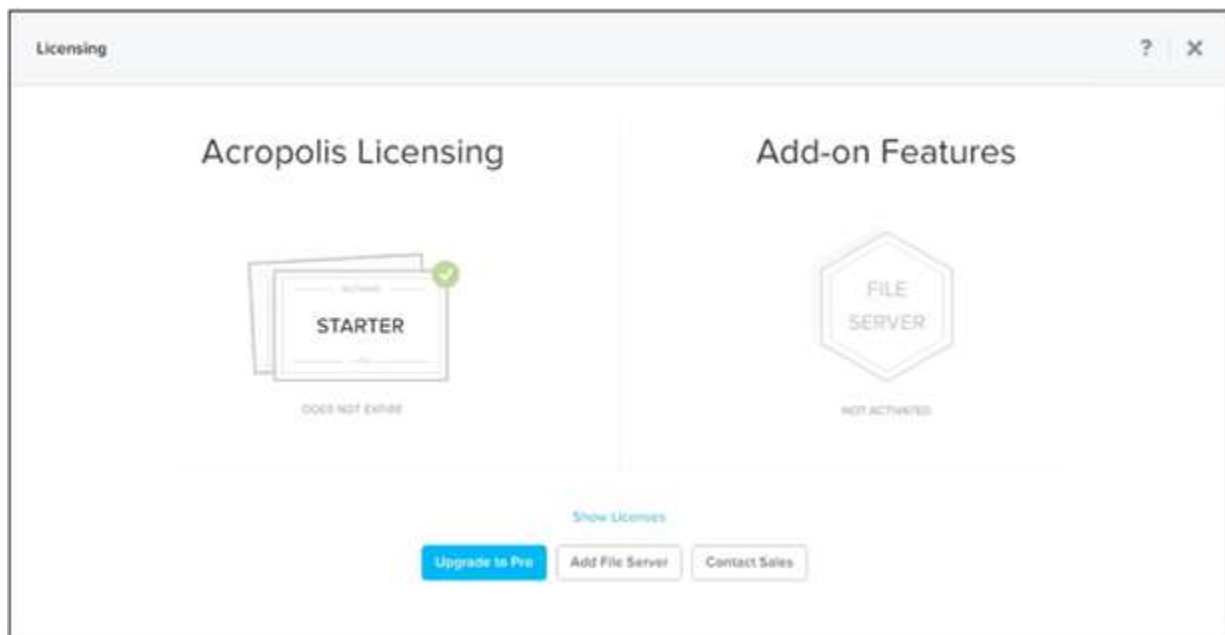
## Section 14 – Lifecycle Operations

### Describe processes and procedures for license management, including AOS and Prism licenses

The Portal Connection feature simplifies license management by integrating the licensing workflow into a single interface in the Prism web console. Once you enable this feature, you can perform most licensing tasks from Prism

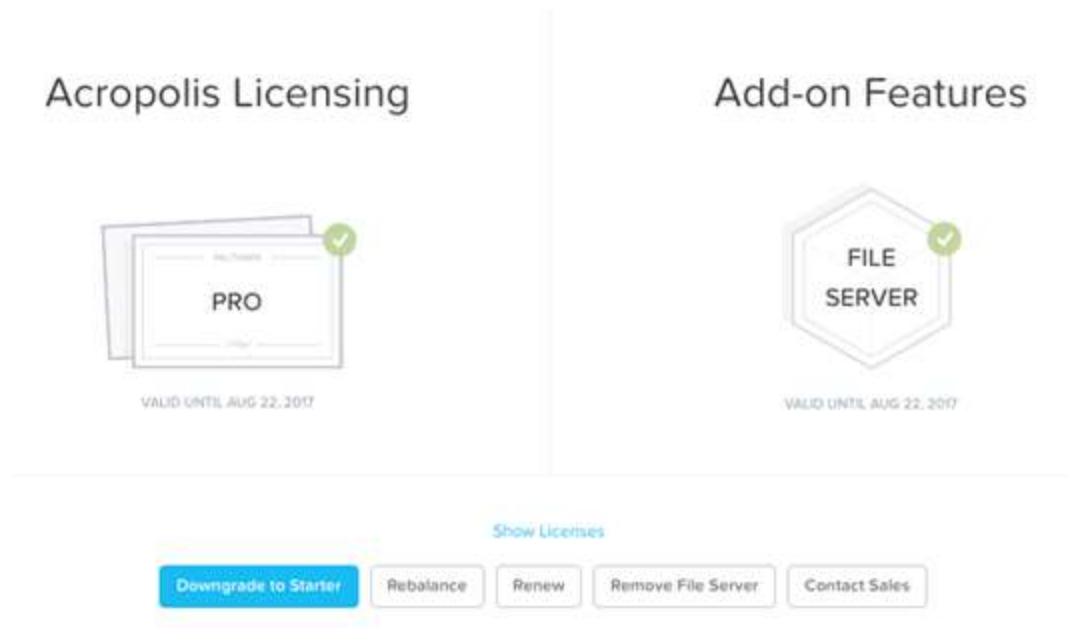


without needing to explicitly log on to the Nutanix Support Portal. It is disabled by default.



## Licensing action in progress...

- **Step 1**     Generate cluster summary file
- **Step 2**    Upload cluster summary file to Support Portal
- **Step 3**    Apply license



## Nutanix Xtreme Computing Platform Licenses

### Starter License

Each Nutanix node and block is delivered with a default Starter license, which does not expire. You are not required to register this license on the Nutanix Customer Portal account assigned to you when you purchased your nodes. These licenses are **automatically applied** whenever you create a cluster, including after you have destroyed a cluster. You do not need to reclaim Starter licenses in this case.

### Pro and Ultimate Licenses

The Pro and Ultimate license types require you to download a license file from the Customer Support Portal and install it on your cluster. When you upgrade to a Pro or Ultimate license or add nodes or clusters to your environment with these licensed features, you must generate the license file, download it, and install it. Define features.

	STARTER	PRO	ULTIMATE
<b>Core Data Services</b>			
Cluster Size	12	Unlimited	Unlimited
vSphere & Hyper-V Support	•	•	•
Heterogeneous Clusters	•	•	•
MapReduce Tiering	•	•	•
Inline Compression	•	•	•
Inline Performance Deduplication	•	•	•
MapReduce Compression		•	•
MapReduce Deduplication		•	•
KVM Support		•	•
<b>Resilience</b>			
Data Path Redundancy	•	•	•
Tunable Redundancy Factor	2	2 or 3	2 or 3
Availability Domains		•	•
<b>Data Protection</b>			
VMCaliber Snapshots	•	•	•
VMCaliber Clones	•	•	•
Single Site DR (1-to-1)	•	•	•
Online Cluster Grow/Shrink	•	•	•
Time Stream		•	•
SRA Integration		•	•
VSS Integration		•	•
Multiple Site DR (1-to-many, many-to-many)			•
Common Access Card			•
Cluster Shield			•
<b>Management &amp; Analytics</b>			
Prism Element	•	•	•
Pulse	•	•	•
Cluster Health	•	•	•
One-click Upgrades	•	•	•
Prism Central		•	•
Rest APIs		•	•

## Prism Central/Prism Pro

AOS 4.6 introduced the Pro license for Prism Central. The Prism Pro license adds additional capabilities to Prism Central, including most of the features

available through the Prism web console of an individual cluster (also known as Prism Element).

## Add-Ons

Individual features known as add-ons can be added to your existing license feature set. When Nutanix makes add-ons available, you can add them to your existing Pro or Ultimate license. For example, Acropolis File Services is an add-on.

## Viewing License Status

The most current information about your licenses is available from the Prism web console. It is also available at the Nutanix Support Portal from the My Products link. You can view information about license types, expiration dates, and any free license inventory (that is, unassigned available licenses). See [Displaying License Features and Details](#).

# Given a scenario, recognize processes to start, stop, and expand a cluster

## Start Cluster

1. Using the vSphere client, take the ESXi hosts out of maintenance mode
2. Power on the CVM's within the cluster
3. SSH into one of the CVM's, and issue the following command:

```
cluster start
```

4. Once completed, it will report back showing that all services are running on all the CVM's within the cluster
5. Validate that the datastores are available and connected to all hosts within the cluster
6. Power on guest VM's

## Stop Cluster

1. Power down all VM's running in the cluster, with the exception of the CVM's

2. SSH into one of the CVM's, and issue the following command:

```
cluster stop
```

3. Once completed, it will report back showing that all services have stopped with the exception of a couple
4. Perform a guest shutdown of the CVM's
5. Place the ESXi hosts into maintenance mode
6. Power down the ESXi hosts

## Check Cluster Status

1. SSH into one of the CVM's, and issue the following command:

```
cluster status
```

2. Once completed, it will report back the status of the cluster

## Expand Cluster

Configuration	Description
Same hypervisor and AOS version	The node is added to the cluster without re-imaging it.
AOS version is different	The node is re-imaged before it is added.
AOS version is same but hypervisor version is different	You are provided with the option to re-image the node before adding it. (Re-imaging is appropriate in many such cases, but in some cases it may not be necessary such as for a minor version difference.)
Data-At-Rest Encryption	You are provided with the option to re-image the node before adding it. (Re-imaging is appropriate in many such cases, but in some cases it may not be necessary such as for a minor version difference.)

# Settings

General

Cluster Details

Configure CVM

Convert Cluster

Expand Cluster

- You can add multiple nodes to an existing cluster at the same time

**Expand Cluster** ? X

1. Host Selection 2. Host Configuration

Newly discovered nodes are displayed below. Select the ones you would like to add and configure their network addresses. Remember to add licenses for all new nodes.

**discovered blocks/nodes**

- NX1020 (Serial Number: 14SM123800)
- null (Serial Number: 10-4-62-93)
- null (Serial Number: 10-4-62-88)

Cancel Next

## Expand Cluster

1. Host Selection 2. Host Configuration

Newly discovered nodes are displayed below. Select the ones you would like to add and configure their network addresses. Remember to use the same IP addresses for all new nodes.

NX1020 (Serial Number: 14SM123800) **block** **node**

D [10.4.60.98] ⓘ

HOST NAME ONLY REQUIRED FOR HYPER-V

Host D

CONTROLLER VM IP SUBNET: 10.4.60.95 / 255.255.252.0

Host D  .  .  .

HYPervisor IP SUBNET: 10.4.60.91 / 255.255.252.0

Host D  .  .  .

IPMI IP SUBNET: 10.4.60.99 / 255.255.252.0



**Expand Cluster** ? X

1. Host Selection · **2. Host Configuration**

Selected nodes are not compatible with existing NOS version: 4.5 X  
or existing hypervisor types or versions: ['esx': ['5.0.0-914586']] on  
cluster nodes

**incompatibility note**

HYPERVISOR(S) NEEDED

1 hypervisor is detected. You might need to upload correct ISO images  
respectively before expanding with selected hosts.

Hypervisor : ESXI REQUIRED BY 1 HOST(S)

Choose File ← **upload ISO image**

Hypervisor ISO Whitelist UPDATED: JUST NOW

**upload whitelist** → Update

Back Cancel Validate **Expand Cluster**

## Install, upgrade and reclaim licenses

# Install

1. Generate a cluster summary file in the Prism web console.
  1. Open **Licensing** from the gear icon in the Prism web console for the connected cluster.
  2. Click **Manage Licenses** in the dialog box, then click **Generate** to create and download a cluster summary file.
  3. The cluster summary file is saved to your browser's download directory or directory you specify.
2. Upload the cluster summary file to the Nutanix support portal.
  1. Click **Support Portal**, log on to the Nutanix support portal, and click **My Products > Licenses**.
  2. Click **License a New Cluster** and the **Upload Cluster Summary File** dialog box is displayed.
  3. Click **Choose File**, then browse to the cluster summary file you just downloaded, select it, and click the Next arrow > Next icon.
  4. The portal automatically assigns the license required according to the information contained in the cluster summary file.
3. Generate and apply the downloaded license file to the cluster.
  1. Click **Generate License File** to download the license file created for the cluster to your browser's download folder or directory you specify.
  2. In the Prism web console, click the upload link in the **Manage Licenses** dialog box.
  3. Browse to the license file you downloaded, select it, and click **Save**.

Cluster UUID - 0004f790-e597-e597-0000-000000000000

Pick a general license type  [Advanced settings](#)

**Model NX-3050**

License Type ▲	# Licenses Assigned ⇅	Expiration Date(s) ⇅
Pro	2	2017-05-01

**Model NX-3061**

License Type ▲	# Licenses Assigned ⇅	Expiration Date(s) ⇅
Pro	1	2017-05-01

[Generate License File](#) [Advanced settings](#) [Close](#)

## Upgrading

- **Step 1** Option 1: [Generate](#) a cluster summary file.  
Option 2: Print or copy down the cluster information below. [Show Info](#)
- **Step 2** Login to the [Support Portal](#) and upload the cluster summary file to generate licenses.
- **Step 3** Download the generated license file and [upload](#) it here.

1. Generate a cluster summary file in the Prism web console.
  1. Open **Licensing** from the gear icon in the Prism web console for the connected cluster.
  2. Click **Update License**.
  3. Click **Generate** to create and save a cluster summary file to your local machine.
  4. The cluster summary file is saved to your browser download directory or directory you specify.
2. Upload the cluster summary file to the Nutanix support portal.
  1. Click **Support Portal**, log on to the Nutanix Support Portal, and click **Products > Node Licenses**.
  2. Click **Licensed Clusters** to display a summary of your licensed clusters.
  3. Find the **Cluster UUID** for the cluster you want to upgrade, then click **Upgrade**.
  4. In the **Manage Licenses** dialog box, click **Choose File**, then browse to the cluster summary file you just downloaded, select it, and click **Next**.
3. Generate and apply the license file to the cluster.
  1. Click **Generate** to download the license file created for the cluster to your browser download folder or directory you specify.
  2. In the Prism web console, click the **upload** link in the **Manage Licenses** dialog box.
  3. Browse to the license file you downloaded, select it, and click **Save**.

## Reclaim

- You can reclaim and optionally re-apply licenses for nodes in your clusters:
  - You must reclaim licenses when you plan to destroy a cluster. First reclaim the licenses, then destroy the cluster.
  - Return licenses to your inventory when you remove one or more nodes from a cluster. Also, if you move nodes from one cluster to another, first reclaim the licenses, move the nodes, then re-apply the licenses.
  - You can reclaim licenses for nodes in your clusters in cases where you want to make modifications or downgrade licenses. For example, applying an Ultimate license to all nodes in a cluster where some nodes are currently licensed as Pro and some nodes are licensed as Ultimate. You might also want to transition nodes from Ultimate to Pro licensing.
1. Generate a cluster summary file in the Prism web console.

1. Open **Licensing** from the gear icon in the Prism web console for the connected cluster.
2. Click **Manage Licenses** in the dialog box, then click **Generate** to create and download a cluster summary file.
3. The cluster summary file is saved to your browser's download directory or directory you specify. To apply the license file to clusters not connected to the Internet, save the file to a USB stick or other media.
2. Upload the cluster summary file to the Nutanix support portal.
  1. Click **Support Portal**, log on to the Nutanix support portal, and click **My Products > Licenses**.
  2. Click **Licensed Clusters** to display a summary of your licensed clusters.
  3. Click **Licensed Clusters**, find your Cluster UUID, and click **Reclaim All**.
  4. In the **Upload Cluster Summary File** dialog box, click **Choose File**, then browse to the cluster summary file you downloaded, select it, and click the Next arrow > Next icon.
3. Click Done.

## Start a node and shut down a node in a Nutanix Cluster

### Start Node VMware

1. Using the vSphere client, take the ESXi hosts out of maintenance mode
2. Power On the CVM
3. SSH into the CVM, and issue the following command:

```
ncli cluster status | grep -A 15 cvm_ip_addr
```

4. Validate that the datastores, are available and connected to all hosts within the cluster

### Stop Node VMware

### GUI

1. Using the vSphere client, place the ESXi host into maintenance mode

2. SSH into CVM and issue the following command:

```
cvm_shutdown -P now
```

3. Once the CVM is powered down, shutdown the host

## CLI

1. SSH into the node being shutdown
2. From the command line, issue the following command:

```
cvm_shutdown -P now
```

3. Login to another CVM, and issue the following commands:

```
~/serviceability/bin/esx-enter-maintenance-mode -s
```

```
~/serviceability/bin/esx-shutdown -s
```

4. Ping the hypervisor IP and confirm that it is powered down

## Start Node AHV

1. Log on to the AHV host with SSH.
2. Find the name of the Controller VM.

```
root@ahv# virsh list --all | grep CVM
```

3. Make a note of the Controller VM name in the second column.
4. Determine if the Controller VM is running.
  - If the Controller VM is off, a line similar to the following should be returned:

```
NTNX-12AM2K470031-D-CVM shut off
```

- Make a note of the Controller VM name in the second column.

- If the Controller VM is on, a line similar to the following should be returned:

```
NTNX-12AM2K470031-D-CVM running
```

- If the Controller VM is shut off, start it.

```
root@ahv# virsh start cvm_name
```

- Replace *cvm\_name* with the name of the Controller VM that you found from the preceding command.
5. If the node is in maintenance mode, log on to the Controller VM and take the node out of maintenance mode.

```
nutanix@cvm$ acli
```

```
<acropolis> host.exit_maintenance_mode AHV-hypervisor-IP-address
```

6. Replace *AHV-hypervisor-IP-address* with the IP address of the AHV hypervisor.

```
<acropolis> exit
```

7. Log on to another Controller VM in the cluster with SSH.
8. Verify that all services are up on all Controller VMs.

```
nutanix@cvm$ cluster status
```

9. If the cluster is running properly, output similar to the following is displayed for each node in the cluster:

```
CVM: 10.1.64.60 Up
      Zeus UP [5362, 5391, 5392, 10848, 10977, 10992]
      Scavenger UP [6174, 6215, 6216, 6217]
      SSLTerminator UP [7705, 7742, 7743, 7744]
      SecureFileSync UP [7710, 7761, 7762, 7763]
      Medusa UP [8029, 8073, 8074, 8176, 8221]
      DynamicRingChanger UP [8324, 8366, 8367, 8426]
      Pithos UP [8328, 8399, 8400, 8418]
```

	Hera	UP	[8347, 8408, 8409, 8410]
	Stargate	UP	[8742, 8771, 8772, 9037, 9045]
	InsightsDB	UP	[8774, 8805, 8806, 8939]
8890]	InsightsDataTransfer	UP	[8785, 8840, 8841, 8886, 8888, 8889,
	Ergon	UP	[8814, 8862, 8863, 8864]
	Cerebro	UP	[8850, 8914, 8915, 9288]
	Chronos	UP	[8870, 8975, 8976, 9031]
	Curator	UP	[8885, 8931, 8932, 9243]
	Prism	UP	[3545, 3572, 3573, 3627, 4004, 4076]
	CIM	UP	[8990, 9042, 9043, 9084]
	AlertManager	UP	[9017, 9081, 9082, 9324]
	Arithmos	UP	[9055, 9217, 9218, 9353]
	Catalog	UP	[9110, 9178, 9179, 9180]
	Acropolis	UP	[9201, 9321, 9322, 9323]
	Atlas	UP	[9221, 9316, 9317, 9318]
	Uhura	UP	[9390, 9447, 9448, 9449]
	Snmp	UP	[9418, 9513, 9514, 9516]
	SysStatCollector	UP	[9451, 9510, 9511, 9518]
	Tunnel	UP	[9480, 9543, 9544]
10301]	ClusterHealth	UP	[9521, 9619, 9620, 9947, 9976, 9977,
	Janus	UP	[9532, 9624, 9625]
	NutanixGuestTools	UP	[9572, 9650, 9651, 9674]
	MinervaCVM	UP	[10174, 10200, 10201, 10202, 10371]
	ClusterConfig	UP	[10205, 10233, 10234, 10236]
	APLOSEngine	UP	[10231, 10261, 10262, 10263]
	APLOS	UP	[10343, 10368, 10369, 10370, 10502,
10503]	Lazan	UP	[10377, 10402, 10403, 10404]
	Orion	UP	[10409, 10449, 10450, 10474]
	Delphi	UP	[10418, 10466, 10467, 10468]

## Stop Node AHV

**Caution:** Verify the data resiliency status of your cluster. If the cluster only has replication factor 2 (RF2), you can only shut down **one node for each cluster**. If an RF2 cluster would have more than one node shut down, shut down the entire cluster.



1. Shut down guest VMs that are running on the node, or move them to other nodes in the cluster.
2. If the Controller VM is running, shut down the Controller VM.
3. Log on to the Controller VM with SSH.
  - List all the hosts in the cluster.

```
acli host.list
```

- Note the value of **Hypervisor address** for the node you want to shut down.
4. Put the node into maintenance mode.

```
nutanix@cvm$ acli host.enter_maintenance_mode Hypervisor address [wait="{ true | false }" ]
```

5. Specify *wait=true* to wait for the host evacuation attempt to finish.
6. Shut down the Controller VM.

```
nutanix@cvm$ cvm_shutdown -P now
```

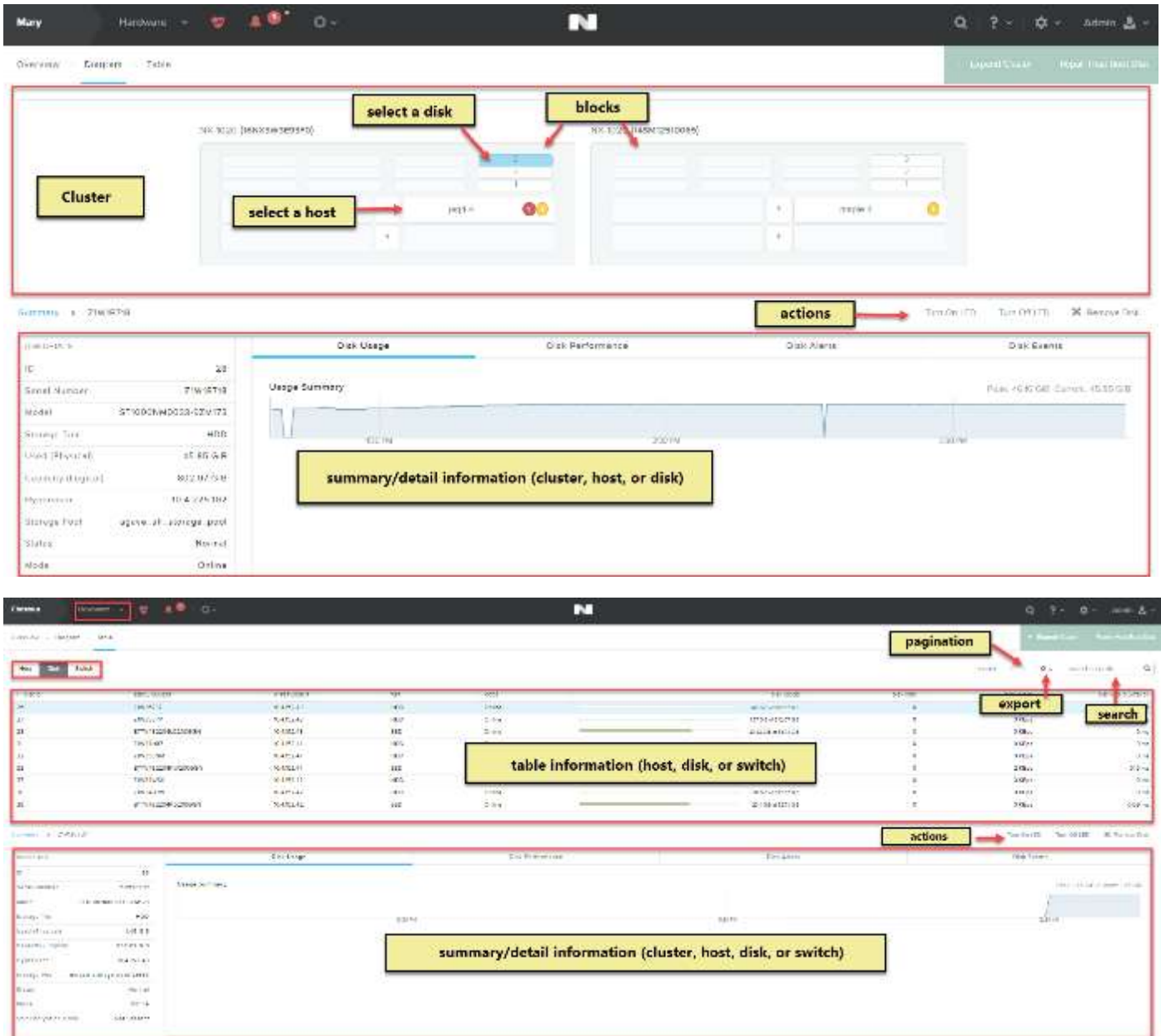
7. Log on to the AHV host with SSH.
8. Shut down the host.

```
root@ahv# shutdown -h now
```

## Eject a node from a Nutanix Cluster

Hardware components (nodes and disks) can be removed from a cluster or reconfigured in other ways when conditions warrant it.

Note: If the Data-at-Rest Encryption is enabled then before removing a drive or node from a cluster, test the certificates again by clicking **Test all nodes** and ensure that testing is successful and the status is **Verified**. In case of an SED drive or node, if the drive or node is not removed as recommended then the drive or node will be locked.



1. To remove a host (node) from the cluster, either select the target host in the diagram (Diagram view) or click the Host tab and select that host in the table (Table view), and click the **Remove Host** link on the right of the Summary line. A dialog box appears to verify the action; click the OK button.
  1. The Prism web console displays a warning message that you need to reclaim the license after you have removed the node.
  2. Removing a host takes some time because data on that host must be migrated to other hosts before it can be removed from the cluster. You can monitor progress through the dashboard

messages. Removing a host automatically removes all the disks in that host. Only one host can be removed at a time. If you want to remove multiple hosts, you must wait until the first host is removed completely before attempting to remove the next host.

3. (Hyper-V only) Initiating a removal of a node running Hyper-V fails if the node is running as a part of a Hyper-V failover cluster and the following message is displayed.
  1. Node node id is a part of a Hyper-V failover cluster. Please drain all the roles, remove the node from the failover cluster and then mark the node for removal.
  2. If this message is displayed in either nCLI or in Web interface, cluster administrators must use the management tools provided by Microsoft such as Failover Cluster Manager to drain all the highly-available roles off the host and then remove the host from the failover cluster followed by removing the host from the AOS cluster.
4. (ESXi only) VMware/cluster administrators must use the management tools provided by VMware such as vCenter to migrate all the VMs off the host and then remove the host from the vCenter cluster followed by removing the host from the AOS cluster.
  - **Caution:** Be sure to migrate VMs in an ESXi cluster prior to removing a host. Verify you have enough available compute capacity in the cluster (ESXi or Hyper-V) before actually migrating the VMs.
  - If you click **Remove Host** without first migrating the VMs, the VM's data may be migrated without any notice and those VMs will lose service.
2. After a node is removed, it goes into an unconfigured state. You can add such a node back into the cluster through the expand cluster workflow.
3. To **add a host into the metadata store**, either select the target host in the diagram (Diagram view) or click the Host tab and select that host in the table (Table view), and click the **Enable Metadata Store link on the right of the Summary** line.
4. Each node includes a disk used for metadata storage, and AOS maintains a metadata store across these disks to ensure uninterrupted resiliency should a metadata disk fail. After such a failure, that node is taken out of the metadata store group and the cluster continues to operate seamlessly without it. Normally, the node is brought back into the metadata store automatically after the failed metadata disk is

replaced. However, under certain (rare) circumstances this might not happen. If the node is ready but was not added back automatically, the alert message Node ready to be added to metadata store is displayed and the Enable Metadata Store link appears indicating you can add the node back into the metadata store manually.